



Etude de l'interactome ARN-ARN chez *Staphylococcus aureus* : caractérisation, fonction et impact sur les réseaux de régulation

Emma Desgranges

► To cite this version:

Emma Desgranges. Etude de l'interactome ARN-ARN chez *Staphylococcus aureus* : caractérisation, fonction et impact sur les réseaux de régulation. Bactériologie. Université de Strasbourg, 2020. Français. NNT : 2020STRAJ076 . tel-03504785

HAL Id: tel-03504785

<https://theses.hal.science/tel-03504785>

Submitted on 29 Dec 2021

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

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

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ (414)

Unité de recherche : Architecture et Réactivité de l'ARN (UPR9002)

Institut de biologie moléculaire et cellulaire (IBMC), Strasbourg

THÈSE présentée par :

Emma DESGRANGES

Soutenue le : **17 septembre 2020**

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

Discipline/S spécialité : Aspects Moléculaires et Cellulaires de la Biologie

**Étude de l'interactome ARN-ARN chez *Staphylococcus aureus* :
caractérisation, fonction et impact sur les réseaux de régulation**

THÈSE dirigée par :

Mme Romby Pascale

&

Mme Caldelari Isabelle

Docteur, université de Strasbourg

Docteur, université de Strasbourg

RAPPORTEURS :

Mr Repoila Francis

Mr Condon Ciarán

Docteur, AgroParisTech et université de Paris-Saclay

Professeur, université de Paris Diderot

EXAMINATEUR INTERNE :

Mme Frugier Magali

Docteur, université de Strasbourg

AUTRE MEMBRE DU JURY :

Mr Mandin Pierre

Docteur, université de Aix-Marseille

Remerciements

Je tiens premièrement à remercier les membres de mon jury, **Ciarán Condon, Francis Repoila, Magalie Frugier et Pierre Mandin**, qui ont gentiment accepté d'évaluer mon travail de thèse.

Je voudrais également remercier les différentes personnes qui ont participé de près ou de loin à mon travail de thèse et/ou créé un environnement agréable au cours de ces quatre dernières années.

Pascale, merci de m'avoir accueilli dans l'équipe au cours de mon Master, et de m'avoir entraîné et encouragé pour pouvoir y réaliser une thèse. Ta passion pour la science, ta rigueur, ta patience, l'implication que tu as pour chacun de nous, sont toutes autant de raisons pour te dire merci.

Isabelle, merci à toi également de m'avoir accepté pour les stages et la thèse. Il a été particulièrement agréable de travailler à tes côtés. Merci également d'avoir toujours su rester optimiste quand souvent je ne l'étais plus. Merci de manière générale pour l'encadrement auquel j'ai eu le droit, de m'avoir fait confiance et laissé être autonome bien que ta porte ait toujours été ouverte quand il y avait besoin de discuter science ou autres d'ailleurs.

Stefano, merci d'avoir pris le temps de m'aider et de m'avoir expliqué à plusieurs reprises les choses concernant les analyses de séquençage. Merci de nous faire partager ta curiosité scientifique.

David, ou papy, lalalou, et même lalaloulila j'entends à l'oreillette, merci d'avoir débarqué au lab pour nous montrer que même en post doc il est possible de galérer sur des choses aussi simples que les northern et les transcriptions in vitro. Merci d'avoir rejoint le projet DPRIM, parce que c'est comme ça qu'il continuera à s'appeler dans nos cœurs.

Laura A, on a commencé cette aventure ensemble depuis le master, merci pour l'entraide, merci pour la chimie, merci pour les histoires sur tes animaux loufoques.

Lucas, merci d'avoir un IMC parfait, merci de m'avoir laissé jouer à tu connais Lucas. Merci pour l'aide dernièrement avec les manips et depuis le début pour toujours avoir été ok de dépanner une solution par-ci par-là.

Noemie Merdier, merci ma petite meuf, ce couple aura été très fusionnel. Merci de te bruler au phenol, de presque tomber dans les pommes en faisant un pH, et pour le presque incendie de la salle radioac. Je ne peux pas croire que je vais louper tes prochaines aventures.

Laura B, merci pour les petits gâteaux que nous imposons que tu nous fasses. Bon courage pour la nouvelle l'aventure qu'est la thèse que tu vas bientôt commencer.

De manière générale merci au 436 box 4 ever, pour les tea time, les pauses clopes, les recherches internet toutes plus wtf les unes que les autres, les montages photo, les bières du mois, les routes du rhum. Et surtout, merci d'avoir écouté mes plaintes et jérémiades sur tout et n'importe quoi.

Roberto & Hirki, thank you for being part of the lab, unfortunately we didn't have a lot of time to know each other. I wish you the best for the following, to both of you.

Anne-Catherine, merci d'être autant impliqué dans l'organisation du labo, ce qui nous simplifie grandement la vie au quotidien.

Delphine, même si tu n'es plus au 436, tu m'as pris sous ton aile dès mes premiers pas au lab, tu m'as tout montré, donné tous les tips, merci pour ça et puis merci pour toutes ces poses café à déblatérer sur nos vies, où nous avons pu nous rendre compte que bien que nous nous sommes rencontrés à Strasbourg nous étions toutes deux né à l'hôpital de bourg de péage dans la Drôme ! J'ai eu grand plaisir à travailler avec toi surtout quand nos sujets de thèse ont commencé à se mélanger.

L'UPR9002, parmi vous un grand nombre a pu m'aider à un moment ou un autre donc merci beaucoup pour vos conseils, vos idées, les coups de main en général. Merci également aux personnes des services communs **Danielle**, **Léandra**, **Anaïs** et **Catherine**. Merci également à **Béatrice** pour les analyses bio-informatique, le temps investi et la pédagogie dont tu as fait preuve pour répondre à toutes mes questions. Merci au buveur de café du matin, merci d'avoir été dans mes pattes chaque jour, le démarrage des journées aurait été beaucoup plus banal sans vous. Merci à la team RU, ceux qui sont encore doctorant et ceux qui ont déjà fini, partager ces repas avec vous presque quotidiennement, toujours dans une bonne ambiance a été un réel plaisir. Merci à tous les doctorants même ceux qui ne mangeaient pas au RU, pour les PhD séminaires suivis du PhD drink, vous entendre expliquer vos sujets respectifs a toujours été intéressant, et la petite bière après un agréable moment.

Il y a également des personnes à remercier en dehors du labo.

Mr Bergasse, je vous remercie de m'avoir permis de découvrir la biologie. Je pense que c'est la façon dont vous enseigniez cette discipline qui m'a donné l'envie de poursuivre mes études dans ce domaine.

Il y a un grand nombre d'amis rencontrés à Strasbourg :

La famille, on en a partagé beaucoup de choses et des noms de conversations, **Arnaud**, **Paula**, **Vincent** et **Robin**, maintenant que la famille a fini sa crise de la quarantaine, la famille passe la thèse (- robin), on a pu aller boire des coups et faire des karaokés ensemble dans plein de pays européens et c'était super cool, on a même des fois pris le temps de visiter entre deux bières.

Arnaud, merci pour ces discussions presque quotidiennes sur les bananes, ce fruit a pris une grande place dans notre alimentation et notre amitié. Merci de nous transmettre ta passion du Racing.

Paula, ma petite Besta, merci pour tous ces bons moments qu'on a eus toutes les deux en vacances, à la piscine, même au Charly's. Je ne boirais plus jamais un bomb-cul-sec sans penser à toi.

Vincent, mon chat merci d'être toujours le dernier à partir, surtout quand tu restes plus longtemps que les autres et qu'il faut te jeter dehors. Merci aussi pour les twerk endiablés et les chansons de Queen B partagé. Merci de m'emmener me balader (certains appellerons ça une randonnée) le dimanche.

Robin, tu es le seul d'entre nous à ne pas avoir fait l'expérience de la thèse, mais malgré ça tu nous as souvent écouté en parler alors merci pour ta patience. Merci également pour toutes les photos mal cadrées. Promis on vient bientôt te faire des gros bisous à Madrid.

Le groupe de la bière, Alors, je ne vais pas vous citer un par un, car je ne voudrais pas risquer d'oublier un seul d'entre vous. Merci pour avoir été globalement très présent à ce rendez-vous incontournable et tant attendu qui est l'A-P-E-R-O du mercredi. Mais aussi les autres jours additionnels quand il y en avait besoin. Merci également d'avoir aussi bien occupé mes vendredis et/ou samedis soir voire les journées quand nous n'avons pas su nous arrêter. Merci pour les week-ends de festoch, pour les week-ends de visite de ceux qui sont déjà partis « loin » à Paris ou Bruxelles. Et bien sûr merci pour les apéros à distance pendant le confinement.

Cooney, mon coronacataire ! Merci d'avoir été là pendant le confinement, de voir ne pas bosser sur ta propre thèse m'a étonnamment motivé à bosser sur la mienne.

Le groupe des filles, Merci d'être d'aussi bonnes copines, pour ces moments rien qu'entre nous quand les garçons ont bien voulu accepter qu'on fasse des soirées sans eux, pour la conversation qui est le meilleur endroit pour partager des questions existentielles et pour les 50 messages non lus quand tu as le malheur de ne pas regarder ton portable 30 minutes pendant l'une d'elles.

La team orthographe, Estelle, Morgane et Cooney merci pour la relecture de dernière minute.

Le BM, bar très charmant, service impeccable, je recommande. **BLAGUE !** Quand j'ai ouvert la porte de ce bar pour la première fois je ne pensais pas y trouver une deuxième maison. Alors, merci le Blue Moon ! Merci **Mika, Mimi**, et **Toto** de nous accueillir aussi bien. Merci pour les madeleines, les Jagger bombe cul-sec, et les pintes de pils que je n'ai même plus besoin de demander en arrivant.

Il y a aussi des copines de longue date.

Aurore et Candice, ... que dire, ça va être dure de résumer en quelques lignes l'importance que vous avez eue jusqu'alors dans ma vie. Alors, je vais simplement vous remercier pour toutes ces années d'amitié à Romans puis à Grenoble et enfin même lorsque la distance physique s'est faite entre nous.

Et bien sûr il y a aussi la famille.

La famille Desgranges, en général, merci à tous mes oncles, tantes, cousins, cousines, pour les bons moments en famille. Merci également pour toutes ces années où vous vous êtes demandé quand est-ce qu'Emma finirait ces études. La route a été longue mais nous y sommes la petite dernière va devenir grande.

Nancy, ma marraine merci pour ta gentillesse et ta générosité, merci d'avoir toujours cru en moi et d'avoir suivi mon parcours avec tant d'intérêt depuis toutes ces années.

Wilfried, merci d'être mon grand frère et merci de t'être toujours intéressé à ce que je faisais. Merci également à **Margaux** qui a rejoint notre famille. Je me réjouis de bientôt pouvoir être plus disponible pour les sorties ski en famille !

Papa et Maman, un merci tout particulier, pour l'excellent travail de parent que vous avez fait, de m'avoir fait grandir en étant constamment encouragé. Je pense que si je fais preuve d'autant de persévérance jusque-là c'est surtout grâce à vous, car malgré les difficultés vous m'avez toujours montré qu'il faut s'accrocher. Merci de m'avoir fait confiance et de m'avoir laissé une aussi grande liberté pour choisir ce que je voulais faire dans la vie. Merci aussi de supporter mon caractère.

Papi et Mami, c'est un merci qui ne sera pas lu, ce qui me fait particulièrement mal au cœur. Merci pour la belle enfance que j'ai eue grâce à vous deux, pour l'implication sans failles de Mami quand il a fallu faire beaucoup de dictée, ces bons repas, sa gentillesse tout simplement, pour la fierté dans les yeux de papi à chaque fois qu'il y avait une bonne note. S'il y a bien deux personnes qui seraient particulièrement fières de moi c'est bien vous et vous rendre fière a toujours été un moteur dans ma vie.

Liste des Abréviations

Nucléotides, acides nucléiques, et coenzyme

A	adénine
ADN	acide désoxyribonucléique
ADNc	ADN complémentaire
ARN	acide ribonucléique
ARNas	ARN anti sens
ARNe	ARN éponge
ARNm	ARN messenger
ARNnc	ARN non codant
ARNr	ARN ribosomique
ARNt	ARN de transfert
ASOs	<i>short antisens oligonucleotides</i> ; petits oligonucléotides anti sens
ATP	adénosine triphosphate
C	cytosine
CTP	cytidine triphosphate
ddNTP	didésoxyribonucléotide
dNTP	désoxyribonucléotide triphosphate
eADN	ADN extracellulaire
FADH₂	flavine adénine dinucléotide réduit
G	guanine
GTP	guanosine triphosphate
miARN	micro ARN
NAD⁺	nicotinamide adénine dinucléotide
NADH	nicotinamide adénine dinucléotide réduit
NADPH	nicotinamide adénine dinucléotide phosphate réduit
NTP	nucléotide triphosphate
Oligo	oligonucléotides
sARN	ARN régulateur
siARN	ARN interférent
T	thymine
TTP	thymidine triphosphate
U	uracile
UTP	uridine triphosphate

Ribonucléase

DNase	désoxyribonucléase
RNase E	endoribonucléase E
RNase III	endoribonucléase III
RNase R	exoribonucléase R

Bactéries

<i>B. subtilis</i>	<i>Bacillus subtilis</i>
<i>E. coli</i>	<i>Escherichia coli</i>
<i>L. monocytogenes</i>	<i>Listeria monocytogenes</i>
SARM	<i>S. aureus</i> résistant à la méthicilline
SARM-AC	souche SARM associés à la communauté
SARM-AH	souche SARM associés à l'hôpital
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
<i>S. argenteus</i>	<i>Staphylococcus argenteus</i>

<i>S. aureus</i>	<i>Staphylococcus aureus</i>
<i>S. capitis</i>	<i>Staphylococcus capitis</i>
<i>S. enterica</i>	<i>Salmonella enterica</i>
<i>S. epidermidis</i>	<i>Staphylococcus epidermidis</i>
<i>S. flexneri</i>	<i>Shigella flexneri</i>
<i>S. simiae</i>	<i>Staphylococcus simiae</i>
SCV	<i>small-colony variant</i> ; variant à petite colonie
<i>V. cholerae</i>	<i>Vibrio cholerae</i>
VISA	<i>vancomycin intermediate S. aureus</i> ; <i>S. aureus</i> intermédiaire à la vancomycine
VRSA	<i>vancomycin resistant S. aureus</i> ; <i>S. aureus</i> résistant à la vancomycine
WT	<i>wild type</i> ; sauvage

Milieu de culture

BHI	<i>brain heart infusion</i> ; bouillon cœur cerveau
LB	Luria-Bertani / <i>lysogeny broth</i> ; bouillon de lysogénie

Sucres

F6P	fructose-6-phosphate
G1P	glucose-1-phosphate
G6P	glucose-6-phosphate
LPS	lipopolysaccharide
M6P	mannose-6-phosphate

Techniques

CLASH	<i>cross-linking, ligation and sequencing of hybrids</i> ; pontage des hybrides, ligation et séquençage
DBRI	<i>Direct determination of bacterial RNA-RNA interactome</i> ; détermination directe de l'interactome ARN-ARN bactérien
LIGR-seq	<i>ligation of interacting RNA and high-throughput sequencing</i> ; ligation de l'ARN en interaction couplé à un séquençage à haut débit
MAPS	<i>MS2-affinity purification coupled with RNA-seq</i> ; purification d'affinité MS2 couplée à un séquençage de l'ARN. MS2-CP capsid du phage MS2. MS2 ARN MS2 du phage MS2
PCR	<i>polymerase chain reaction</i> ; réaction de polymérisation en chaîne
RIL-seq	<i>RNA interaction by ligation and sequencing</i> ; interaction d'ARN par ligation et séquençage
RNA-seq	<i>RNA sequencing</i> ; séquençage de l'ARN
RT-PCR	<i>transcriptase inverse-PCR</i> ; reverse transcription-PCR

Protéines, gènes, génome, voies métaboliques

Agr	<i>accessory gene regulator system</i> ; système régulateur de gène accessoire
AIP	<i>auto-inducing peptide</i> ; peptide auto inducteur
cre	<i>catabolite responsive elements</i> ; éléments sensibles aux catabolites
ETS	<i>external transcribed spacer</i> ; espaceur externe entre transcrits
FT	facteur de transcription
HGT	<i>horizontal gene transfer</i> ; transfert horizontal de gènes
ICR	<i>intercistronic region</i> ; région inter-cistronique
IGR	<i>intergenic region</i> ; région inter-génique

ORF	<i>open reading frame</i> ; cadre ouvert de lecture
PBP	<i>penicillin binding protein</i> ; enzyme de liaison à la pénicilline
PPP	<i>pentose phosphate pathway</i> ; voie des pentoses phosphates
PSM	<i>phenol-soluble modulín</i> ; moduline soluble au phénol
PTS	<i>phosphotransferase system</i> ; système phosphotransférase
PVL	leucocidine de Pantón Valentine
RBS	<i>ribosome binding site</i> ; site de fixation du ribosome
SCCmec	<i>staphylococcal cassette chromosome mec</i> ; cassette chromosomique <i>mec</i> du <i>staphylocoque</i>
SD	séquence Shine et Dalgarno
TCS	<i>two component system</i> ; système à deux composants
TSST-1	<i>superantigenic toxic shock toxin 1</i> ; toxine superantigénique du syndrome de choc toxique 1
UTR	<i>untranslated region</i> ; région non traduite

Espèces réactives oxygénées

NO	oxyde nitrique
ROS	<i>reactive oxygen species</i> ; espèces réactives de l'oxygène

Solutions, réactifs, matériel

AMT	4'-Aminométhyltrioxsalène hydrochloride
BET	bromure d'éthidium
DIG	digoxygénine
DTT	dithiothréitol
EDTA	acide éthylène diamine tétraacétique
P32	phosphore 32
PCI	phénol, chloroforme, alcool isoamylique
PTA	psoralène TEG-azide
RM	rack magnétique
SDS	sodium dodécyl sulfate
SSC	<i>saline-sodium-citrate</i> ; solution saline-citrate de sodium
TBE	tris-borate-EDTA

Autres

CMI	concentration minimale inhibitrice
DO	densité optique
FFL	<i>feedforward loop</i> ; boucle de rétroaction
OMV	<i>outer membrane vesicle</i> ; vésicule de la membrane externe
QS	<i>quorum sensing</i> ; détection du quorum, système détectant la densité cellulaire
Rsa	<i>RNA from S. aureus</i> ; ARN de <i>S. aureus</i>
R-chimère	chimère redondante (chimère retrouvée au moins 5 fois dans l'analyse DBRI)

Généralité

°C	degré Celsius
cpm	coups par minute
G	G-force, force centrifuge relative
g	gramme
l	litre
M	molaire
mol	mole

MT	marqueur de taille
nt	nucléotide
pb	paire de base
pH	potentiel hydrogène
rpm	rotation par minute
UV	ultraviolet
V	volt
vol	volume
W	watt
X	fois

Table des Matières

I. Introduction	1.
1. <i>Staphylococcus aureus</i>	2.
1.1. Informations générales	2.
1.2. Un peu d'histoire	2.
1.3. Les résistances aux antibiotiques chez <i>S. aureus</i>	3.
1.3.1. Généralités	3.
1.3.2. Résistance aux β -lactames	3.
1.3.3. Résistance aux glycopeptides	5.
2. Le métabolisme bactérien	6.
2.1. Généralités	6.
2.2. Le métabolisme central	6.
2.3. Le métabolisme central chez <i>S. aureus</i>	8.
2.4. Le métabolisme de <i>S. aureus</i> en conditions infectieuses	9.
3. Régulation de l'expression des gènes chez <i>S. aureus</i>	10.
3.1. Les facteurs de transcription	10.
3.1.1. Généralités	10.
3.1.2. Les facteurs de transcription reliés au projet de thèse	11.
3.1.2.1. Les facteurs de transcription de la famille des Sar	11.
3.1.2.1.1. SarA	11.
3.1.2.1.2. TcaR	11.
3.1.2.2. Les facteurs de transcription impliqués dans le métabolisme et la virulence	12.
3.1.2.2.1. Rex	13.
3.1.2.2.2. CcpA	13.
3.1.3. L'exemple de la régulation par Rex et CcpA en condition de stress	NO 14.
3.2. Les systèmes à deux composants (TCS)	15.
3.2.1. Généralités	15.
3.2.2. Les TCS reliés au projet de thèse	16.
3.2.2.1. Le système Agr	16.
3.2.2.2. Le système HptRS	16.
3.3. Les ARN régulateurs	17.
4. Les ARN régulateurs	18.
4.1. Un peu d'histoire	18.
4.2. Aspect évolutif	19.
4.2.1. Évolution, conservation des ARN régulateurs	19.

4.2.2. Évolution, conservation des régions 3'UTR	21.
4.3. Les ARN régulateurs chez <i>S. aureus</i>	23.
4.3.1. sARN et carences nutritives	40.
4.3.1.1. RsaC	40.
4.3.1.2. RsaD	40.
4.3.2. RsaE : implication dans la réponse au stress oxydatif, le métabolisme central et la formation de biofilm	41.
4.3.3. sARN et virulence	42.
4.3.4. sARN et antibiotiques	42.
4.4. Les ARN éponges	42.
4.5. Caractérisation et étude des ARN régulateurs	45.
4.5.1. Caractérisation des ARN régulateurs de manière globale et individuelle	45.
4.5.2. Caractérisation du targetome d'un ARN régulateur	55.
4.6. Description des réseaux	75.
4.7. Mécanisme de régulation et de communication en dehors de la cellule par les ARN : Les OMV	76.
4.8. Espoir thérapeutique contre la résistance aux antibiotiques grâce à l'ARN	77.
5. Les infections à <i>S. aureus</i>	78.
5.1. Les infections extracellulaires	78.
5.2. Les infections intracellulaires	79.
5.2.1. Généralités	79.
5.2.2. Adaptation métabolique	80.
5.2.3. Régulation de l'expression des transporteurs : le cas de UhpT	81.
5.2.4. sARN et internalisation	82.
5.2.5. Émergence des SCV	83.
5.2.6. Émergence des bactéries peristantes ou persisteurs	83.
6. Objectifs de la thèse	84.
6.1. Découverte des sARN dans le génome de <i>S. aureus</i>	85.
6.1.1. Caractérisation de RsaI	85.
6.1.2. Caractérisation de RsaG	86.
6.2. Développement d'une technique d'étude de l'interactome ARN-ARN chez <i>S. aureus</i> : DBRI	86.
II. Résultats	88.
1. Chapitre 1 : RsaI, un ARN régulateur aux multiples facettes, module le métabolisme de <i>S. aureus</i>	89.
2. Chapitre 2 : RsaG, un sARN dérivé d'une région 3'UTR régule l'homéostasie redox en réponse au glucose-6-phosphate chez <i>S. aureus</i>	131.

3. Chapitre 3 : Mise au point de la technique DBRI « Direct determination of bacterial RNA-RNA interactome »	189.
3.1. Matériel et méthodes	189.
3.1.1. Culture bactérienne	189.
3.1.2. Mise au point du protocole : variantes 1-4	189.
3.1.2.1. Traitement au psoralène et pontage par irradiation	189.
3.1.2.2. Extraction des ARN	190.
3.1.2.3. Traitement Dnase	190.
3.1.2.4. Ribodéplétion	191.
3.1.2.5. Digestion à la nucléase S1	191.
3.1.2.5.1. Digestion à la nucléase S1 des transcrits in vivo	191.
3.1.2.5.2. Digestion à la nucléase S1 des transcrits in vitro	191.
3.1.2.5.2.1. Transcription in vitro de <i>tcaR</i> et RsaG	191.
3.1.2.5.2.2. Digestion à la nucléase S1 des transcrit in vitro	192.
3.1.2.6. Ligation	192.
3.1.2.6.1. Ligation (1) T4 RNA ligase (ThermoFischer)	192.
3.1.2.6.2. Ligation (2) T4 RNA ligase (NEB)	192.
3.1.2.6.3. Ligation (3) à haute température	192.
3.1.2.6.3.1. Adénylation	192.
3.1.2.6.3.2. Ligation	193.
3.1.2.7. Digestion des ARN par la RNase R	193.
3.1.2.8. Réversion du pontage	193.
3.1.3. Mise au point du protocole : variantes 5 et 6	193.
3.1.3.1. Traitement au psoralène	193.
3.1.3.2. Irradiation aux UV	194.
3.1.3.3. Extraction des ARN et traitement à la DNase	194.
3.1.3.4. Réaction de chimie « click »	194.
3.1.3.5. Purification sur des billes magnétiques de streptavidine	194.
3.1.3.6. Digestion des ARN à la nucléase S1 sur les billes	195.
3.1.3.7. Ligation sur les billes (1) : variante 5	195.
3.1.3.7.1. Adénylation	195.
3.1.3.7.2. Ligation	195.
3.1.3.8. Ligation sur les billes (2) : variante 6	196.
3.1.3.9. Réversion du pontage	196.
3.1.3.9.1. Réversion directe du pontage	196.
3.1.3.9.2. Elution suivie de la réversion du pontage	196.
3.1.4. Séquençage	197.

3.1.4.1.	Séquençage d'une chimère connue résultante du couple d'ARN <i>tcaR</i> -RsaG	197.
3.1.4.1.1.	RT-PCR	197.
3.1.4.1.2.	Clonage des fragments de PCR	197.
3.1.4.2.	Séquençage à haut débit	197.
3.2.	Analyse des résultats	198.
3.2.1.	Explication du choix des conditions de culture	198.
3.2.2.	Variantes 1.1 et 1.2	198.
3.2.3.	Variante 2	200.
3.2.4.	Variante 3	201.
3.2.4.1.	Approche expérimentale utilisée	201.
3.2.4.2.	Analyse du séquençage à haut débit	201.
3.2.4.3.	Conclusion	203.
3.2.5.	Variante 4.1	204.
3.2.6.	Variante 4.2	205.
3.2.6.1.	Approche experimental utilisée	205.
3.2.6.2.	Analyse du séquençage à haut débit	206.
3.2.6.3.	Conclusion	210.
3.2.7.	Variante 5	211.
3.2.8.	Variante 6	212.
3.2.9.	Variante 7	213.
3.3.	Conclusion	214.
III.	Discussion	215.
1.	Coexistence des sARN dans les réseaux de régulation	216.
1.1.	Construction d'un modèle global de régulation impliquant plusieurs sARN dépendant de l'environnement de <i>S. aureus</i>	216.
1.1.1.	Interaction entre RsaI et RsaG	220.
1.1.2.	Redondance possible des cibles entre RsaG et RsaD	224.
1.1.3.	Redondance des cibles entre RsaG et RsaC	227.
1.1.4.	Conclusion	228.
2.	Étude globale de l'interactome ARN/ARN chez <i>S. aureus</i>	229.
2.1.	Intérêt du développement d'une nouvelle technique basée sur le pontage des ARN	229.
2.2.	Les prochaines étapes nécessaires pour améliorer DBRI	230.
2.3.	Intérêt et critique des études à l'échelle globale pour l'étude des cibles ARN	232.
2.4.	Perspectives d'utilisation	235.
3.	Conclusion générale	235.

IV. Matériel & méthodes	237.
1. Liste des oligonucléotides	238.
2. Liste des plasmides	238.
3. Amplification des plasmides Pjet::RsaG- <i>tcaR</i> (DBRI)	238.
3.1. Préparation des cellules compétentes	239.
3.2. Transformation des bacteries chimio-compétentes	239.
3.3. Séquençage	239.
4. Northern blot (DBRI)	239.
4.1. Préparation des ARN totaux	239.
4.2. Séparation des ARN totaux et transfert	239.
4.3. Détection des ARN	240.
4.4. Construction des sondes ARN modifiées par la digoxygénine	240.
5. PCR	241.
6. Transcription in vitro	241.
7. Déphosphorylation de l'extrémité 5' des ARN	241.
8. Marquage d'ARN ou d'oligonucléotides en 5'	242.
9. Retard sur gel	242.
10. Toe-print	243.
V. Références bibliographiques	245.

I. Introduction

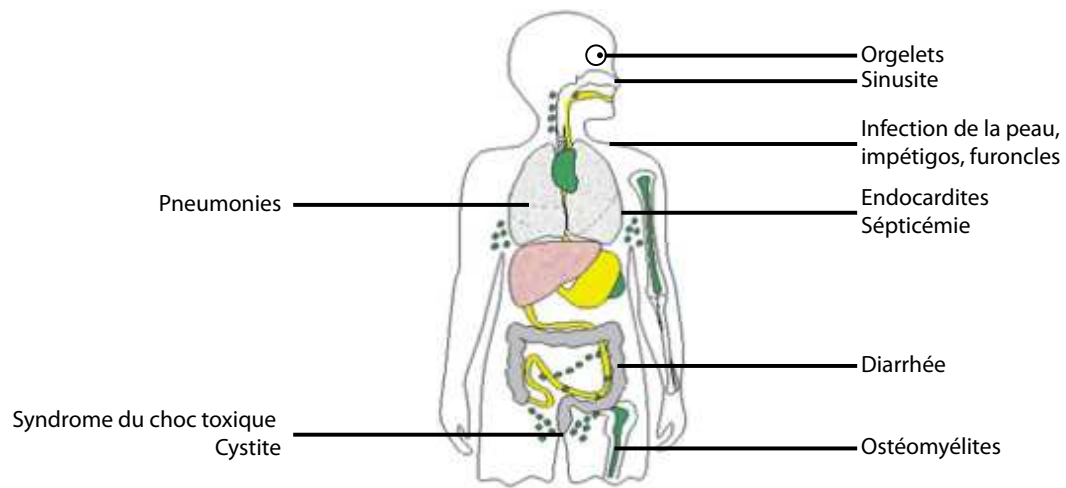


Figure 1 : Large éventail des maladies associées à *S. aureus*

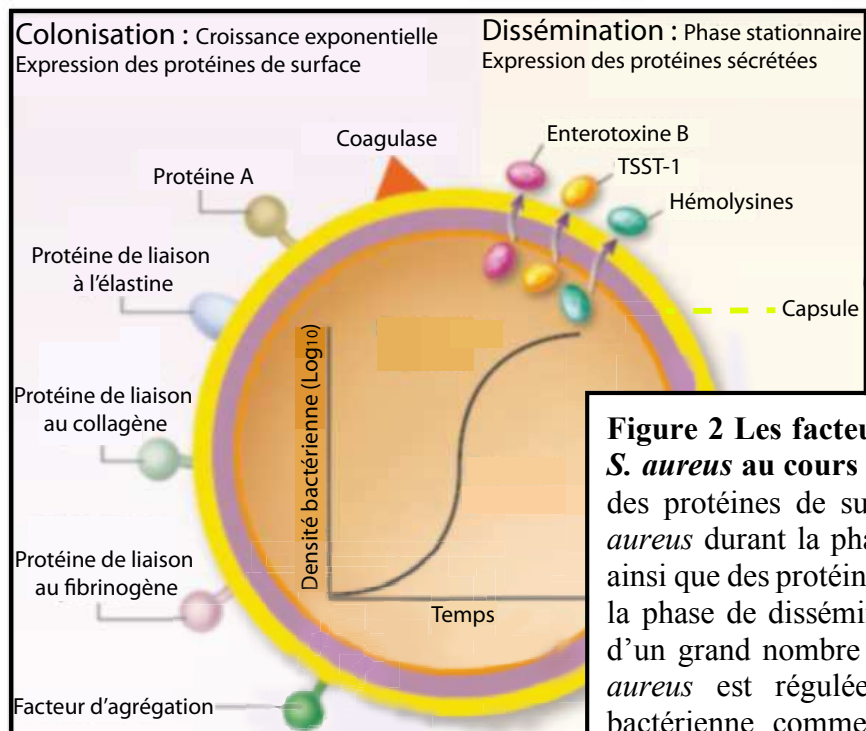


Figure 2 Les facteurs de virulence déployés par *S. aureus* au cours de l'infection : Représentation des protéines de surface/liaison exprimées par *S. aureus* durant la phase de colonisation (à gauche), ainsi que des protéines/toxines sécrétées au cours de la phase de dissémination (à droite). L'expression d'un grand nombre de facteurs de virulence de *S. aureus* est régulée en fonction de la densité bactérienne comme le montre le graphique (au centre) des phases de croissance. D'après Lowy et al., 1998.

1. *Staphylococcus aureus*

1.1. Informations générales

Staphylococcus aureus est une bactérie pathogène opportuniste de l'Homme. Comme commensale, elle est retrouvée à la surface de la peau et des muqueuses d'environ 30% des individus sains (Wertheim et al., 2005). Cependant sa capacité à coloniser un grand nombre de tissus fait d'elle la cause d'un large éventail de maladies infectieuses, allant des infections superficielles de la peau à des maladies plus graves, voire mortelles, comme les pneumonies, endocardites et septicémies (Figure 1).

La capacité de *S. aureus* à exprimer de nombreux facteurs de virulence (200-300) lui est d'une grande aide dans l'établissement des infections. Ceux-ci permettent dans un premier temps de faciliter son attachement, grâce à l'expression de protéines de surface/d'adhésion. Dans un deuxième temps, lorsque la densité cellulaire est plus élevée, les facteurs de virulence exprimés permettent de faciliter l'invasion des tissus, grâce à l'expression de protéines et de toxines excrétées. Le passage de *S. aureus* d'un stade de colonisation à celui de la dissémination est très finement régulé en fonction de la densité bactérienne. Les facteurs de virulence de *S. aureus* lui permettent aussi de compromettre l'efficacité des neutrophiles et des macrophages, qui représentent la première ligne de défense de l'hôte, en partie grâce à l'expression de la protéine A et la formation de la capsule (Figure 2) (Foster, 2005; Guerra et al., 2017; Lowy, 1998).

Par ailleurs, en plus de l'arsenal déployé par *S. aureus* au cours de l'infection, cette bactérie est aussi connue pour développer un grand nombre de résistances aux antibiotiques. Ceci résulte en une complexité dans la mise en place de traitements voire en impasses thérapeutiques, faisant de cette bactérie un problème majeur de santé publique.

1.2. Un peu d'histoire

C'est Alexander Ogston en 1880 qui a été le premier à observer un arrangement de bactéries ressemblant à une grappe de raisin dans une infection purulente, qu'il a donc nommée *Staphylococcus*, du grec *staphyle* « grappe de raisin » et *kokkos* « grain ». Par la suite, en 1884, Anton Julius Friedrich Rosenbach a réussi à établir une culture pure de ces micro-organismes et les a nommés par leurs couleurs... *Staphylococcus aureus* du latin *aurum* pour « doré », et *Staphylococcus albus* du latin pour « blanc », aujourd'hui appelé *Staphylococcus epidermidis*. Les travaux d'Alexander Fleming ont ensuite mis en lumière *S. aureus*. En 1928, Alexander

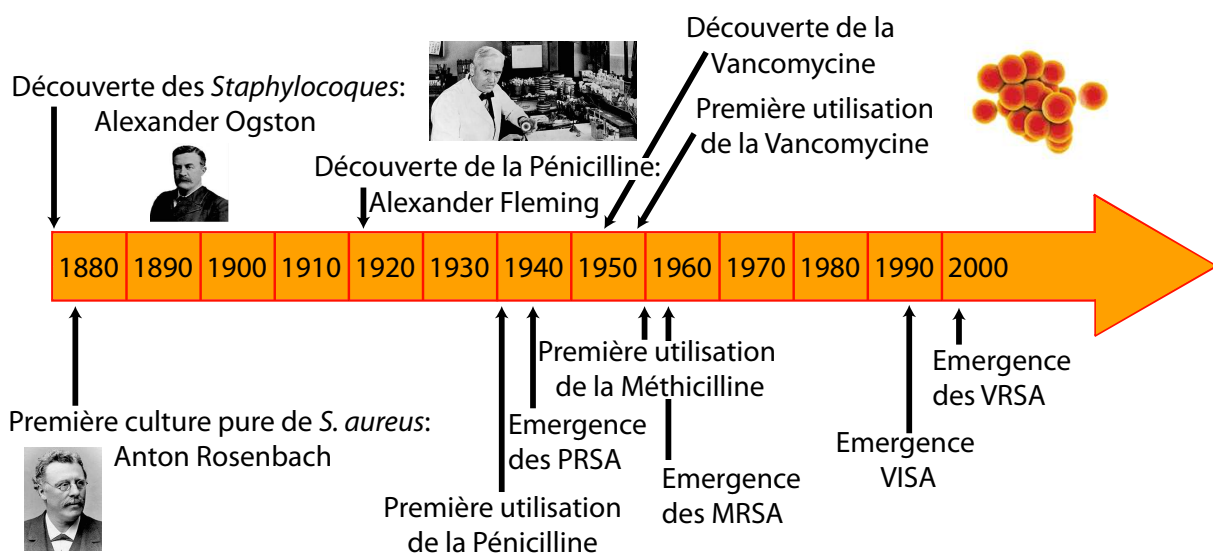


Figure 3 : Frise chronologique de l'histoire de *S. aureus* depuis sa découverte à nos jours.

Fleming découvre le premier antibiotique, la pénicilline, grâce à la contamination d'une de ses cultures de *S. aureus* par le champignon *Penicillium notatum*. Cette découverte prometteuse pour le traitement des infections bactériennes a pavé la voie vers l'âge d'or des antibiotiques. Il reste cependant intéressant de noter que bien que la première utilisation clinique de la pénicilline eu lieu en 1941, un cas de *S. aureus* résistant à la pénicilline avait déjà été rapporté un an auparavant. L'émergence des souches résistantes à la pénicilline augmente au milieu des années 40'. *S. aureus* continue sans cesse à s'adapter très rapidement aux antibiotiques. La méthicilline découverte en 1958, et utilisée pour la première fois en 1959, donne lieu à l'apparition de SARM (*S. aureus* résistant à la méthicilline) dès 1961. De même pour la vancomycine découverte en 1953, et utilisée dès 1958 pour traiter les infections insensibles à la méthicilline, qui entraîne l'apparition des VISA « *vancomycin intermediate S. aureus* » dès 1996, ainsi que des VRSA « *vancomycin resistant S. aureus* » dès 2002. Ces souches résistantes aux antibiotiques sont maintenant dispersées à travers le monde (Figure 3).

1.3. Les résistances aux antibiotiques chez *S. aureus*

1.3.1. Généralités

L'utilisation trop large des antibiotiques en santé humaine, avec parfois des traitements et posologies inadaptés, ainsi que leur utilisation massive dans l'agriculture (dans un but prophylactique) a créé un contexte favorable pour sélectionner des bactéries résistantes aux antibiotiques. L'apparition de résistance aux antibiotiques résulte de quatre mécanismes majeurs : l'expulsion de l'antibiotique vers le milieu extérieur, la production d'une enzyme qui modifie ou détruit l'antibiotique, la modification de la cible de l'antibiotique, ou l'imperméabilisation de la membrane de la bactérie vis-à-vis de cet antibiotique. *S. aureus* présente une grande capacité d'adaptation, due à la plasticité de son génome qui lui a permis de développer des mécanismes variés de résistance aux antibiotiques suite à des mutations ou l'acquisition de gènes par transfert horizontal (HGT). En effet, ce pathogène redoutable a réussi à développer des résistances à tous les antibiotiques introduits au cours du siècle écoulé. Ainsi, nous observons actuellement plus de 95% des souches de *S. aureus* résistants à la pénicilline.

1.3.2. Résistance aux β -lactames

Les β -lactames inhibent la réaction de transpeptidation en se fixant à la protéine de liaison à la pénicilline (PBP), empêchant ainsi la synthèse du peptidoglycane. Il existe deux

	<i>S. aureus</i> résistant à la méthicilline associés à l'hôpital (SARM-AH)				<i>S. aureus</i> résistant à la méthicilline associé à la communauté (SARM-AC)			
Type d'infections	Infections de la peau et des tissus superficiels				Infections profondes ou systemiques			
SCCmec	Type SCCmec	Complexe mec	Compexe de genes <i>ccr</i>	Résistances associées	Type SCCmec	Complexe mec	Compexe de genes <i>ccr</i>	Résistances associées
	I	Classe B	1	Kanamycine Tobramycine	IV	Classe B	2	/
	II	Classe A	2	Kanamycine Tobramycine Erythromycine Spectinomycine	V	Classe C2	5	/
	III	Classe A	3	kanamycine Tobramycine Erythromycine Tracycline Spectinomycine Cadmium Mercure	VI	Classe B	4	/
					VII	Classe C1	5	/
PVL	Généralement non porteurs				Généralement porteurs			
					VIII	Classe A	4	Erythromycine

Tableau 1 : Différences majeures entre SARM-AH et SARM-CA : Le tableau représente les types d'infections produites, la nomenclature des huit types de SCCmec, ainsi que la capacité de production de leucocidine de Panton Valentine (PVL) en fonction du type de SARM.

mécanismes de résistance à la pénicilline, et aux β -lactames de manière générale. Le premier est la synthèse d'une β -lactamase grâce à l'acquisition d'un plasmide. Cette enzyme est responsable de l'hydrolyse des β -lactames, les rendant inefficaces. Ce mécanisme a été découvert rapidement après l'émergence des premiers *S. aureus* résistants à la pénicilline par Kirby, (1944). L'acquisition du gène *mecA*, ainsi que de ses régulateurs *mecI* et *mecRI*, suite au transfert horizontal de l'élément génétique mobile *SCCmec* (« *staphylococcal cassette chromosome mec* ») constitue un second mécanisme de résistance aux β -lactames. En effet, le gène *mecA* est responsable de la synthèse d'une enzyme de liaison à la pénicilline, la PBP2a, qui présente une affinité moindre pour la pénicilline que la PBP. L'expression de cette enzyme entraîne donc la résistance de *S. aureus* à la méthicilline (SARM). La cassette *SCCmec* est composée des gènes *mec* ainsi que des gènes *ccr* qui codent pour des recombinaisons permettant la mobilité de la cassette (Fuda et al., 2005). En fonction des souches, d'autres gènes peuvent être présents sur les éléments mobiles *SCCmec*, tels que des transposons ou des gènes de résistance aux métaux lourds ou à d'autres antibiotiques que les β -lactames. Il existe 8 types de cassettes *SCCmec* qui dépendent de la combinaison du complexe *mec* (4 classes A, B, C1 et C2) et du complexe de gènes *ccr* (5 types, 1 à 5) présent (Dumitrescu et al., 2010 ; tableau 1).

Parmi les infections à *S. aureus* SARM, on peut distinguer les infections associées à l'hôpital (SARM-AH) et celles associées à la communauté (SARM-AC). Les SARM-AH sont plus prédominantes chez des individus avec des facteurs de risque. Les SARM-AC sont apparues plus récemment, et sont la cause d'infections chez des patients n'ayant aucun facteur de risque, et n'ayant pas été en contact avec le milieu hospitalier. Actuellement les SARM-AC sont également retrouvées dans le contexte hospitalier, augmentant ainsi le risque de transmission non seulement pour les patients, mais également pour les individus sains comme le personnel soignant.

De manière intéressante les variants alléliques du *SCCmec* de types I à III sont plus représentés dans les SARM-AH, alors que les types IV à VIII sont identifiés dans des souches SARM-AC. Les cassettes *SCCmec* de types I à III, retrouvées dans les souches SARM-AH, comportent des gènes de résistance à d'autres antibiotiques que les β -lactames, tels que la kanamycine ou l'érythromycine. Il en résulte des souches multi-résistantes aux antibiotiques. A l'inverse, les SARM-AC portent des *SCCmec* de type IV à VIII, de séquence plus courte et ne comportant pas de gènes additionnels entraînant des multi-résistances aux antibiotiques. Une autre différence entre ces deux types de SARM est qu'un grand pourcentage des SARM-AC code pour la leucocidine de Pantone Valentine (PVL) qui est largement absente des génomes des

SARM-AH (Tableau 1). La toxine PVL agit en formant des pores dans la membrane des neutrophiles, monocytes et macrophages, et entraîne également une cytotoxicité des leucocytes par l'induction de la production de médiateurs pro-inflammatoires (Dumitrescu et al., 2010). Il est généralement rapporté que les souches SARM-AC sont plus virulentes que les souches SARM-AH, dû entre autres à la production de la PVL (Montgomery et al., 2018 ; Otto et al., 2013). Cependant d'autres études n'appuient pas cette théorie et montrent que la PVL n'est pas le seul facteur de virulence responsable de l'intensité de la virulence des souches AH et AC. Il a été montré que les souches de *S. aureus* AH et AC sont autant efficaces pour établir une infection mortelle dans un modèle de chenilles *Galleria mellanella* (Boakes et al., 2016). Par ailleurs, ce modèle d'étude est éloigné de l'Homme mais permet de mettre en avant l'importance d'un autre type de facteur de virulence qui sont les PSM (*phenol-soluble modulin*, ce sont des peptides cytolytiques excrétés par *S. aureus*). Tous les PSM sont encodés dans le *core* génome de *S. aureus* à l'exception des PSM*mec*. Le gène *psm-mec* est retrouvé dans les SCC*mec* de types II, III, et VIII, car présent dans le complexe *mec* de classe A. La présence de PSM*mec* (à l'image des complexes *mec* de classe A) est donc partagé entre certains membres des SARM-AH et d'autres membres des SARM-AC (Qin et al., 2016).

1.3.3. Résistance aux glycopeptides

La grande capacité d'adaptation de *S. aureus* a également permis l'apparition de souches VISA. Ce sont des souches qui présentent une sensibilité diminuée à la vancomycine et d'autres glycopeptides. Le mécanisme moléculaire lié à l'apparition des VISA n'est pas clairement compris, mais l'exposition prolongée aux glycopeptides semble être l'élément essentiel à l'origine de diverses mutations qui a permis l'acquisition de ce phénotype. Cela se manifeste entre autres par une augmentation de l'épaisseur de la paroi cellulaire (McGuinness et al., 2017). Les souches VISA se distinguent des souches VRSA, plus rares, qui elles sont pleinement résistantes à la vancomycine, suite à l'acquisition de l'opéron *vanA* localisé sur le transposon Tn1546 originaire des entérocoques. Brièvement l'opéron *vanA* encode toutes les protéines nécessaires pour sentir la présence de vancomycine et en réponse synthétiser un pentapeptide (composant du peptidoglycane) modifié insensible à la vancomycine (Périchon & Courvalin, 2009 ; McGuinness et al., 2017).

Les gènes acquis par *S. aureus*, les mutations apparues suite à l'exposition prolongée aux antibiotiques, ainsi que l'expression de facteurs de virulence particuliers, contribuent ainsi à l'acquisition de ces multi-résistances.

2. Le métabolisme bactérien

2.1. Généralités

Le métabolisme se découpe en quatre grandes étapes :

- 1) Les réactions d'approvisionnement permettent, à partir des nutriments, de fabriquer tous les métabolites précurseurs qui sont au nombre de treize, et de générer du pouvoir réducteur (NADH, NADPH) et de l'énergie (potentiel de proton, ATP) nécessaire à l'ensemble du métabolisme. Le métabolisme central (glycolyse, cycle de Krebs, et voie des pentoses phosphates) fait partie de ces réactions d'approvisionnement.
Chaque bactérie a la capacité de s'adapter à sa niche écologique pour utiliser les nutriments qui s'y trouvent. Cela explique qu'il existe une grande diversité dans ces réactions entre microorganismes bien que le métabolisme central soit majoritairement conservé.
- 2) Les réactions de biosynthèse permettent à partir des treize métabolites précurseurs, du pouvoir réducteur et de l'énergie, de former les briques élémentaires (acides aminés, nucléotides...). Ces réactions sont universelles sauf quelques exceptions comme la synthèse de la lysine ou des composants particuliers (peptidoglycane, LPS).
- 3) Les réactions de polymérisation sont en général des réactions universelles qui conduisent à la formation de macromolécules à partir des briques élémentaires (ARN, ADN, protéines).
- 4) Les réactions d'assemblage permettent la construction des différentes macromolécules pour construire une structure cellulaire (ribosome, paroi, capsule...).

Le métabolisme s'articule comme un nœud papillon. A chaque extrémité du nœud on observe une grande diversité, un grand nombre de nutriments qui permettent d'aboutir sur une grande variété de macromolécules et de structures complexes. Au milieu du nœud se trouve le métabolisme central, très conservé, qui permet d'obtenir les métabolites précurseurs, le pouvoir réducteur et l'énergie qui sont les pivots du métabolisme (Figure 4).

2.2. Le métabolisme central

La glycolyse ou moins communément appelée voie d'Embden-Meyerhof-Parnas est une voie de catabolisme du glucose en pyruvate (Figure 5, à gauche). Elle permet de générer sept

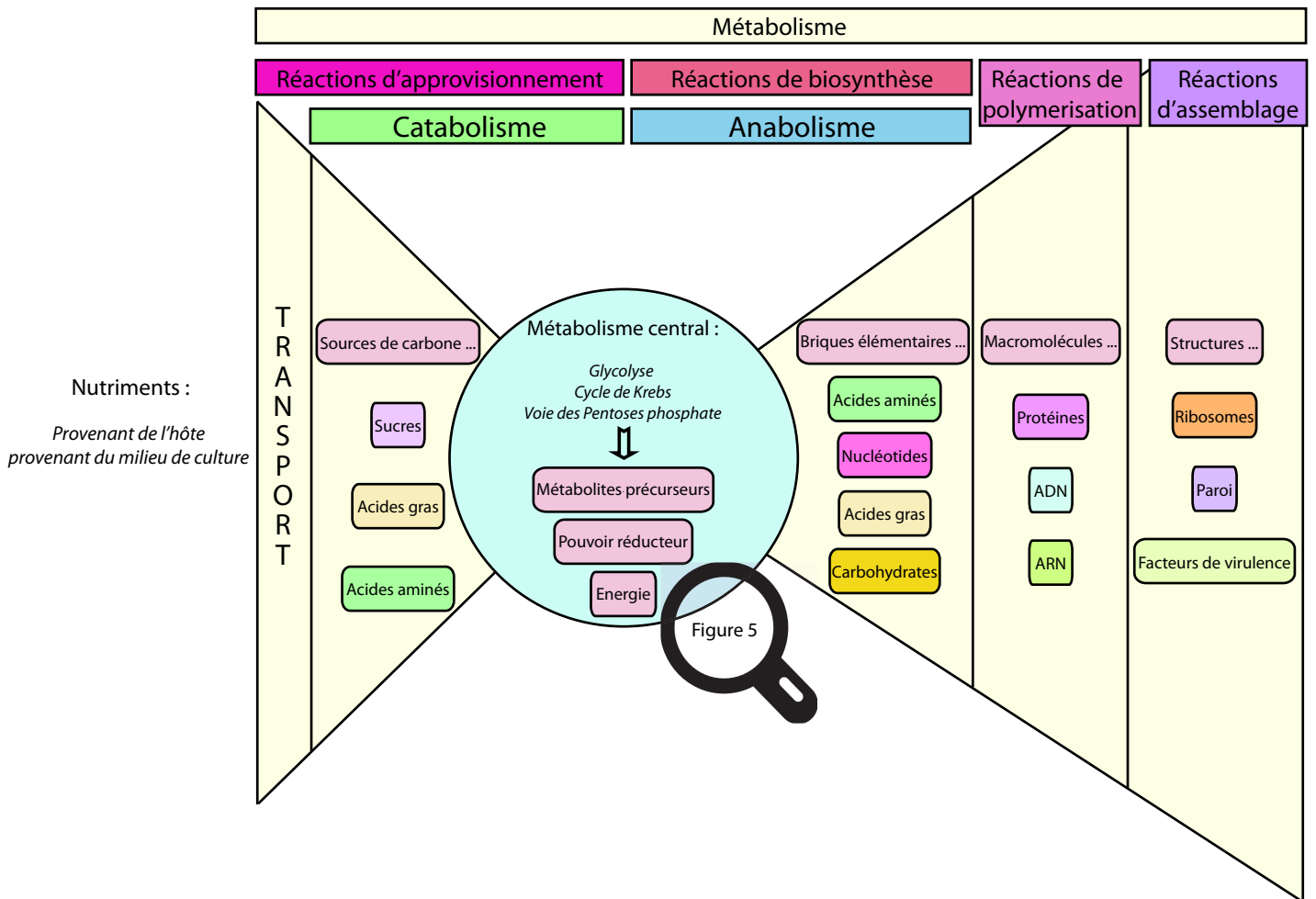


Figure 4 : Vue d'ensemble du métabolisme : Représentation du métabolisme bactérien en architecture de nœud papillon avec une large variété de nutriments à l'entrée pour produire un grand nombre de structures cellulaires. La figure représente les quatre étapes du métabolisme : les réactions d'approvisionnement (regroupées sous le terme de catabolisme), les réactions de biosynthèse (aussi appelé anabolisme), les réactions de polymérisation, et les réactions d'assemblages. La figure 4 appelle la figure 5 pour plus de détails concernant le métabolisme central.

des treize métabolites précurseurs, de l'ATP et du pouvoir réducteur (NADH). La glycolyse est largement alimentée par le système phosphotransférase (PTS) dépendant du phosphoénolpyruvate. Ce mode de transport abouti à la phosphorylation d'un carbohydrate importé dans le cytoplasme. La phosphorylation des sucres à leur entrée à deux avantages. Premièrement, elle permet d'éviter leurs efflux vers le milieu extracellulaire, dans le sens du gradient de concentrations. Deuxièmement, la phosphorylation des sucres permet de les activer, ce qui constitue les premières étapes de la glycolyse. Le glucose peut également être activé intracellulairement par l'action d'hexokinases. Il s'en suit une série de réactions enzymatiques clefs entraînant un flux de carbone. Plusieurs enzymes sont sujettes à une régulation allostérique en fonction de la concentration d'intermédiaires ou de produits de la glycolyse. Le contrôle de l'activité de ces enzymes permet d'agir sur la vitesse du flux de carbone. Le fonctionnement de la glycolyse est indépendant de la présence ou de l'absence d'oxygène (Richardson et al., 2015).

La voie des pentoses phosphate (PPP), ou voie de Warburg-Dickens-Horecker, est également une voie de dégradation du glucose (Figure 5, en haut à droite). Elle produit trois des treize métabolites précurseurs du pouvoir réducteur (NADPH) ainsi que des intermédiaires de la glycolyse. Elle se déroule aussi bien en aérobie qu'en anaérobie, mais est inhibée lorsque la croissance est ralentie. Parmi les métabolites précurseurs produits par la PPP, on retrouve le ribose-5-phosphate et l'érythrose-4-phosphate utilisés dans la synthèse de briques élémentaires d'importance, respectivement les nucléotides (purine) et les acides aminés (aromatiques) (Richardson et al., 2015).

Enfin le cycle de Krebs aussi communément appelé cycle de l'acide citrique (Figure 5, en bas à droite), entraîne la production des trois derniers métabolites précurseurs, de l'énergie (ATP) et du pouvoir réducteur (NADH, $FADH_2$) à partir du pyruvate. Cette voie est présente seulement en présence d'oxygène. Les métabolites précurseurs produits au cours du cycle de Krebs permettent d'initier les voies de biosynthèse d'un grand nombre d'acides aminés ainsi que d'acides gras.

L'altération du flux de carbone dans ces trois voies du métabolisme central a un effet direct sur la concentration des métabolites précurseurs, d'ATP, et sur l'homéostasie rédox. Ces voies sont régulées par des régulateurs. L'activité de ces régulateurs est contrôlée par la concentration intracellulaire en métabolites précurseurs, briques élémentaires, et cofacteurs. Ainsi, la modification du métabolisme permet de convertir les changements environnementaux externes (comme par exemple le changement de la source de carbone) en signaux métaboliques

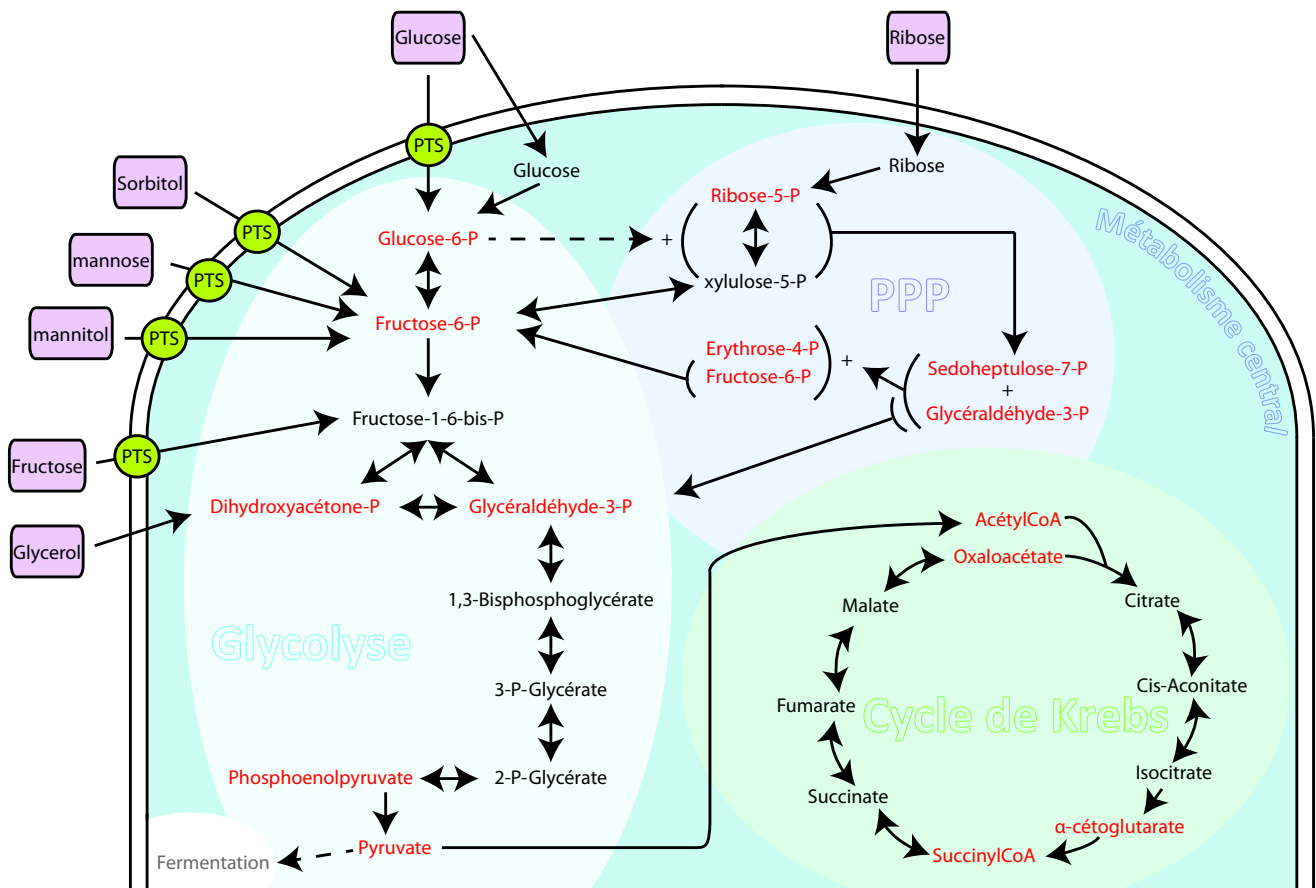


Figure 5 : Vue d'ensemble du métabolisme central : Représentation schématique des trois voies du métabolisme central bactérien que sont la glycolyse, la voie des pentoses phosphates (PPP), et le cycle de Krebs. Le métabolisme central permet de fournir à la cellule les treize métabolites précurseurs, de l'énergie, et du pouvoir réducteur, nécessaires au fonctionnement de l'ensemble du métabolisme.

internes, ce qui va modifier l'activité de ces régulateurs pour faciliter l'adaptation à ce nouvel environnement.

2.3. Le métabolisme central chez *S. aureus*

De manière intéressante alors que la glycolyse et la voie PPP sont globalement très conservées, le cycle de Krebs ne l'est pas nécessairement. Les différences observées dans les capacités du métabolisme central entre bactéries reflètent aussi leur aptitude à persister en dehors de leur organisme hôte. En effet, au cours de l'évolution certaines bactéries ont perdu en partie ou entièrement le cycle de Krebs. Cela a rendu ces bactéries auxotrophes à un grand nombre d'acides aminés et vitamines, puisqu'elles sont incapables de synthétiser les métabolites précurseurs nécessaires à la synthèse de ces briques élémentaires. Un cycle de Krebs incomplet est commun chez les bactéries à Gram positif (par exemple *Streptococcus pyogenes* et *Enterococcus faecium*). *S. aureus* a cependant conservé un métabolisme central complet, ce qui lui apporte un avantage majeur pour s'adapter à un grand nombre d'environnements et permettre sa survie indépendamment d'un hôte (Richardson et al., 2015). De plus, le tropisme tissulaire étendu de *S. aureus* nécessite une grande flexibilité métabolique de sa part. Par ailleurs il a été montré que le cycle de Krebs est indispensable pour la survie de *S. aureus* au cours des infections osseuses (Potter et al., 2020). D'autre part, des mutations affectant le cycle de Krebs influencent la virulence de *S. aureus*, en ayant un effet indirect sur la formation de la capsule (Mei et al., 1997).

De manière préférentielle en présence de glucose ou d'autres sucres rapidement métabolisables (notamment ceux transportés par un PTS), ces sucres sont catabolisés en pyruvate via la glycolyse. Dans la phase exponentielle de croissance, *S. aureus* va shunter le cycle de Krebs en détournant le métabolisme à partir du pyruvate vers la fermentation, indépendamment de la présence d'oxygène. Les produits de fermentations générés (acétate, lactate, éthanol) sont excrétés dans le milieu extracellulaire qui s'acidifie. La fermentation permet également de réoxyder les cofacteurs nécessaires au fonctionnement du métabolisme central (exemple : $\text{NADH} \rightarrow \text{NAD}^+$). Durant la phase stationnaire de croissance, lorsque la source de carbone préférée est épuisée de l'environnement, le cycle de Krebs est déréprimé. Les produits de fermentation précédemment excrétés dans le milieu sont réimportés dans la cellule et utilisés comme source de carbone alternative, permettant ainsi d'alimenter le cycle de Krebs. Un des rôles majeurs du cycle de Krebs est de générer du pouvoir réducteur. L'oxydation

des substrats du cycle de Krebs permet de réduire les cofacteurs (ex : $\text{NAD}^+ \rightarrow \text{NADH}$) (Richardson et al., 2015).

2.4. Le métabolisme de *S. aureus* en conditions infectieuses

S. aureus doit s'adapter rapidement à une large variété de sources de carbone et d'azote au cours du processus infectieux, ainsi qu'à des concentrations variables d'oxygène. En fonction des tissus infectés, la concentration des sources de carbone est très variable. Par exemple, dans le foie et la circulation sanguine le glucose est très abondant, alors que sa concentration est très faible dans les poumons. Dans les tissus où la concentration de glucose est faible, *S. aureus* survit en utilisant des sources de carbone secondaires. *S. aureus* peut ainsi utiliser des sources de carbone telle que des peptides, des acides aminés, des acides gras (seulement pour les incorporer à sa membrane, ils ne sont pas oxydés pour produire de l'énergie), et des acides organiques (Halsey et al., 2017).

S. aureus code pour un grand nombre de transporteurs d'acides aminés et de perméases d'oligopeptides. L'hydrolyse des peptides exogènes ou des acides aminés libres tels que l'alanine, l'arginine, l'aspartate, la cystéine, le glutamate, la glycine, la proline, la sérine, et l'histidine, conduira rapidement à la génération d'énergie. De plus, les acides aminés seront utilisés comme source de carbone. D'autres acides aminés (isoleucine, leucine, lysine, méthionine, phénylalanine, tyrosine, et valine) sont consommés plus lentement car ils sont premièrement utilisés pour la synthèse de protéines (Richardson, 2019). Les voies de catabolisme des acides aminés codées par *S. aureus* lui permettent la synthèse de trois intermédiaires du métabolisme central : le pyruvate, l'oxaloacétate, et l' α -cétoglutarate. Les voies de catabolisme des acides aminés sont soumises à la répression catabolique, les rendant ainsi utilisables seulement dans des conditions pauvres en glucose (Halsey et al., 2017). Il a récemment été montré que la biosynthèse d'aspartate par *S. aureus* est indispensable à sa survie au cours d'une infection osseuse. En effet cela résulte de l'adaptation métabolique de *S. aureus* aux tissus osseux, un environnement très riche en glutamate. L'excès de glutamate inhibe le système de transport de l'aspartate, rendant la biosynthèse de cet acide aminé indispensable (Potter et al., 2020).

Par ailleurs, en plus de la spécificité du métabolisme reliée au type de tissus infectés, l'état physiologique de la personne infectée a aussi une importance. Certaines comorbidités associées, comme le diabète, peuvent augmenter la prévalence des infections à *S. aureus*. En effet, ces patients présentent un taux de sucre élevé dans le sang, ce qui entraîne une

perturbation globale de l'homéostasie du glucose dans tout le corps. *S. aureus* tire avantage de ces conditions pour établir des infections. Dans le cas où la glycémie des patients est contrôlée par un traitement avec de l'insuline, cette prévalence chute, indiquant que ce sont les concentrations élevées en glucose qui sont responsables de l'augmentation de la prévalence des infections à *S. aureus*, plutôt que les déficiences immunitaires qui sont également une conséquence de cette maladie (Kroin et al., 2015; Rich & Lee, 2005; Vitko et al., 2016). Dans ces conditions, *S. aureus* utilise préférentiellement le glucose et réalise un métabolisme glycolytique, ce qui lui donne un avantage certain pour l'établissement de sa virulence. En effet, il a été montré qu'une souche mutante des quatre transporteurs de glucose présents chez *S. aureus* est non virulente dans un modèle murin d'infection de la peau (Vitko et al., 2016). Les infections superficielles de la peau sont les infections les plus retrouvées chez les patients diabétiques.

Pour résumer, en fonction du site d'infection et de la physiologie de l'hôte infecté, on observe une adaptation métabolique de *S. aureus* à son environnement.

3. Régulation de l'expression des gènes chez *S. aureus*

3.1. Les facteurs de transcription

3.1.1. Généralités

Les facteurs de transcription (FT) sont les éléments primaires des réseaux de régulation. Ils peuvent activer ou réprimer l'expression de certains gènes en réponse aux conditions environnementales et physiologiques. Ils permettent entre autres de moduler l'expression de facteurs de virulence ou de gènes du métabolisme en fonction du tissu infecté.

En fonction des souches de *S. aureus*, Ibarra et al. (2013) ont déterminé que le nombre total de FT, incluant les facteurs sigma et les systèmes à deux composantes (TCS), varie de 126 pour la souche RF122 (pathogène de bovin) à 151 pour la souche N315 (SARM). Parmi eux, 112 FT sont présents chez toutes les souches de *S. aureus*. Les FT conservés sont des régulateurs de gènes impliqués autant dans le métabolisme (ArcR, Fur, HutR, GntR, GlnR, CcpA, ArgR, ArcR, FeoA, Fur, PerR, FemC, TreR, GapR, LacR, FapR, CcpA, PurR, HisR) que dans la virulence (AgrA, SarA, SarR, SarS, SarV, SarX, SarY, SarZ, Rot, et MgrA). *S. aureus* arbore des régulateurs essentiels très conservés parmi les bactéries, qui permettent la régulation des processus cellulaires clefs. Des facteurs de transcription spécifiques au genre ou à l'espèce sont également retrouvés chez *S. aureus*, et ont pour rôle de réguler les gènes impliqués dans la virulence ou des voies métaboliques spécifiques.

FT	Cible	Fonction	Publication
SarA	Inhibe sarS, sarT, sarV, active agrA, active icaR	Régulation expression proteines de surface et des toxines sécrétées, régulation formation des biofilm	Manna et al., 1998 ; Cheung et al., 2004 ; Beenken et al., 2003
sarS	Active spa, inhibe hla	Régulation expression proteines de surface et des toxines sécrétées	Tegmark et al., 2000 ; Cheng et al., 2004
SarR	Inhibe sarA et agrA, inhibe sarS	Régulation expression proteines de surface et des toxines sécrétées	Cheung et al., 2004 ; Manna & Cheung 2006
SarT	Active sarS, inhibe sarU	Régulation expression proteines de surface et des toxines sécrétées	Cheung et al., 2004
SarU	Active agrA	Régulation expression proteines de surface et des toxines sécrétées	Cheung et al., 2004
SarV	Inhibe plusieurs gènes impliqués dans l'autolyse (atl, lrgB, et scdA)	Régule l'autolyse	Manna et al., 2004
SarX	Active sa propre expression , inhibe agrA	Régulation expression proteines de surface et des toxines sécrétées	Cheung et al., 2004 ; Manna & Cheung 2006
SarY	?	?	Cheung et al., 2008
SarZ	Active hla, et inhibe sarS	Régulation expression proteines de surface et des toxines sécrétées	Kaito et al., 2006
TcaR	Inhibe l'operon tcaRAB et l'operon icaADBC, active sarS	Implication dans la susceptibilité à la teicoplanine, la régulation de la formation de biofilm, et la régulation de l'expression de proteines de surface	Brandenberger et al., 2000 ; Jefferson et al., 2004 ; McCallun et al., 2004
Rot	Inhibe hla et active spa	Régulation expression proteines de surface et des toxines sécrétées	Bronner et al 2004
MgrA	Active sarX et inhibe sarV	Régulation de la formation de la capsule, régulation de la formation de biofilm, régulation de l'autolyse	Cheung et al., 2008 Romilly et al 2014

Tableau 2 : Les facteurs de transcription de la famille Sar.

3.1.2. Les facteurs de transcription reliés au projet de thèse

3.1.2.1. Les facteurs de transcription de la famille des Sar

La famille de facteurs de transcription Sar est spécifique des *Staphylococaceae*, c'est également la plus grande famille de FT. La famille Sar se divise en 3 sous familles, les protéines à un domaine (SarA, SarR, SarT, SarV, SarX and Rot), les protéines à deux domaines (SarS, SarU et SarY) et les homologues de MarR (MgrA, sarZ, tcaR). Les fonctions majeures des FT de la famille Sar sont listées dans un tableau récapitulatif (Tableau 2, pour revue Cheung et al., 2008).

3.1.2.1.1. SarA

L'exemple type de la famille Sar est SarA, un facteur de transcription largement impliqué dans la virulence de *S. aureus* (Tableau 2). Dans la phase post-exponentielle de croissance, SarA augmente l'expression des toxines α , β , et δ , en régulant directement et indirectement le système Agr (cf I.3.2.2.1). SarA module la résistance aux stress oxydatifs générés par la respiration aérobie ou les défenses immunitaires de l'hôte, en régulant les superoxide dismutase SodA et SodM (respectivement manganèse et fer dépendante). En effet, SarA inhibe SodM en conditions anaérobiques, et SodM et SodA en conditions micro-aérobiques. Ces résultats semblent indiquer que SarA permet de garder une expression faible des protéines Sod lorsque leur expression n'est pas requise (Ballal and Manna, 2009). Enfin, SarA est essentielle pour la formation de biofilms chez *S. aureus* (Valle et al., 2003).

Le locus de *sarA* comporte 3 promoteurs (P1, P2, P3) qui donnent naissance à trois transcrits chevauchants. Les promoteurs P1 et P2 permettent également la transcription d'ARN non codants, respectivement Teg48 et Teg49 (cf revue Desgranges et al., 2019, I.4.3). Comme facteur de transcription, SarA peut réguler l'expression des gènes au niveau transcriptionnel en se liant aux promoteurs spécifiques (Cheung et al., 2004). Cependant de manière plus surprenante, SarA semble également agir au niveau post-transcriptionnel en stabilisant des ARNm pendant la phase exponentielle, par interaction directe avec l'ARN (Morrison et al., 2012; Roberts et al., 2006). Ces derniers résultats n'ont cependant jamais été reproduits

3.1.2.1.2. TcaR

Un autre facteur de transcription de la famille Sar, de la catégorie des homologues de MarR, est TcaR (Tableau 2). TcaR fait partie de l'opéron *tcaRAB* dont il en est le régulateur,

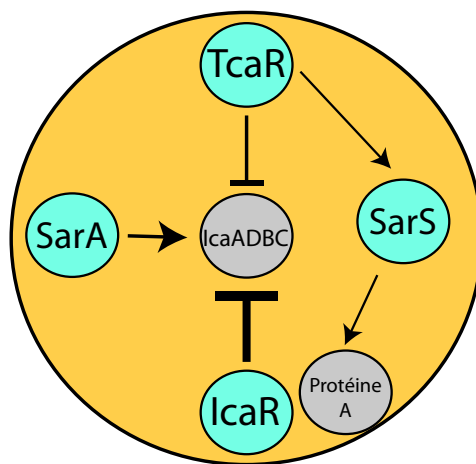


Figure 6 : Réseau de régulation de TcaR : Les protéines représentées en bleu sont des facteurs de transcription. Les flèches symbolisent une activation et les autres connecteurs une inhibition.

cet opéron est impliqué dans la sensibilité à la téicoplanine. La délétion de *tcaA*, codant une protéine transmembranaire, ou de l'opéron entier, entraîne une résistance accrue à la téicoplanine et d'autres glycopeptides comme la vancomycine. Cet opéron pourrait être important dans la synthèse de la paroi cellulaire (Brandenberger et al., 2000; Maki et al., 2004). Cependant, il a clairement été établi que TcaR a un rôle dans la formation de biofilm (Jefferson et al., 2004). Le principal composé des polymères du biofilm est le PNAG (Poly-N-acétylglucosamine). Les enzymes responsables de sa synthèse sont codées par les gènes de l'opéron *icaADBC* qui sont très finement régulés par différents régulateurs. Le gène *icaR* présent dans ce même locus est transcrit de manière divergente à cet opéron et code pour un répresseur de *icaADBC* (Jefferson et al., 2003), alors que SarA (*cf* I.3.1.2.1.1) est un activateur de l'opéron *icaADBC* (Beenken et al., 2003; Valle et al., 2003). TcaR est également capable de se fixer au promoteur de *icaADBC*. La délétion de *tcaR* augmente de cinq fois l'abondance du transcrit *icaAD*, ce qui fait de TcaR un régulateur négatif de la transcription de *icaADBC* (Figure 6). Cependant, la capacité à former des biofilms d'une souche mutante de *tcaR* n'est pas augmentée. L'effet de *tcaR* sur la production de PNAG est visible seulement dans un double mutant *icaR tcaR*. Alors qu'un simple mutant de *icaR* présente une synthèse de PNAG dix fois supérieure, un double mutant voit sa synthèse augmentée de cent fois. Ces résultats montrent que TcaR et IcaR coopèrent dans la répression de la formation des biofilms. Cependant IcaR étant un régulateur plus fort, il masque le phénotype d'un simple mutant *tcaR* (Jefferson et al., 2004). TcaR est également responsable de l'activation indirecte de la synthèse de la protéine A, un facteur de virulence de *S. aureus*, en agissant comme activateur de *sarS* (McCallum et al., 2004 ; Tableau 2 ; Figure 6).

3.1.2.2. Les facteurs de transcription impliqués dans le métabolisme et la virulence

L'expression des facteurs de virulence de *S. aureus* est coordonnée par les FT cités ci-dessus ainsi que AgrA, MgrA et Rot. Cependant, l'établissement de la virulence de *S. aureus* est dépendante des conditions et stimuli environnementaux. C'est pourquoi en parallèle, la mise en place de la virulence dépend également d'autres FT tel que CcpA, CodY, et Rex. Ces FT ont originellement été décrits comme ayant un rôle dans la régulation du métabolisme, démontrant que métabolisme et pathogénicité sont étroitement liés (King et al., 2020; Richardson, 2019).

3.1.2.2.1. Rex

Rex est un régulateur qui réprime un grand nombre de gènes du métabolisme anaérobique (fermentation, respiration anaérobique...) en présence d'oxygène. En se fixant à l'ADN, il empêche la reconnaissance des promoteurs par l'ARN polymérase, inhibant ainsi la transcription de ses gènes cibles (Hecker et al., 2009).

Le premier signe au niveau cellulaire d'un passage de l'aérobie à l'anaérobie est la perturbation du rapport NADH/NAD⁺. Cette modification du potentiel rédox a un effet direct sur Rex car, alors que le NAD⁺ améliore les capacités de répression de Rex, le NADH (qui a une meilleure affinité pour Rex (Wang et al., 2008)) l'inhibe. En présence d'oxygène, la respiration permet d'entretenir un pouvoir réducteur élevé. Rex réprime la fermentation lactique et alcoolique qui sont des voies de régénération du NAD⁺ en conditions d'anaérobie, au profit de l'utilisation de la glycolyse et du cycle de Krebs (Hecker et al., 2009; Pagels et al., 2010). Cependant, la fermentation lactique n'est pas impliquée seulement dans la régénération du pouvoir réducteur en conditions anaérobiques. Chez *S. aureus*, elle lui permet aussi de résister aux radicaux d'oxyde nitrique (NO), un antimicrobien produit par le système immunitaire inné de l'hôte (Pagels et al., 2010; Richardson et al., 2008; cf I.3.1.3).

3.1.2.2.2. CcpA

CcpA est une protéine impliquée dans le contrôle catabolique. Il exerce un rôle essentiel dans l'utilisation de la source carbonée préférée au regard d'autres sources de carbone secondaires, en régulant le métabolisme central. En effet CcpA, un régulateur transcriptionnel de la famille LacI/GalR, active l'expression de gènes codant pour des enzymes de la glycolyse. À l'inverse il réprime l'expression de protéines impliquées dans le cycle de Krebs, ainsi que des gènes impliqués dans l'utilisation des acides aminés comme source de carbone alternative. CcpA, activé en présence de glucose grâce à la formation d'un complexe avec le corépresseur Hpr phosphorylé en position Ser46, est constitutivement exprimé chez *S. aureus*. En présence de glucose extracellulaire, Hpr peut également être phosphorylé au niveau de His15, ce qui lui permet d'agir dans le système phosphotransférase (PTS) pour l'acquisition de sucre. CcpA, lorsqu'il est associé avec Hpr phosphorylé sur Ser46, a une forte affinité pour des séquences *cre* (*catabolite responsive elements*) au niveau des promoteurs de ses gènes cibles. En fonction de la position du site *cre*, CcpA exercera une répression ou une activation de la transcription (Seidl et al., 2009).

CcpA est également impliqué dans la virulence et la résistance aux antibiotiques en régulant l'expression de facteurs de virulence tels que l'ARNIII, *spa* codant pour la protéine A, *hla* codant l'hémolysine- α , et *tst* codant pour TSST-1 (*superantigenic toxic shock toxin 1*) (Seidl et al., 2016 ; Seidl et al., 2008a). D'un point de vue physiopathologique, il a été observé que CcpA augmente le potentiel hémolytique de *S. aureus* dans un modèle de souris diabétiques (Bischoff et al., 2017). CcpA a également un rôle dans la production des biofilms qui chez *S. aureus* est dépendante de fortes concentrations en sucre et en sel. Il active le gène *cidA* qui contribue à la production d'ADN extracellulaire (eADN) pour la matrice du biofilm, et indirectement *icaA* (Seidl et al., 2008b).

Il est maintenant clairement établi que beaucoup de régulateurs des voies métaboliques influencent également la virulence et *vice versa*. Il existe donc un lien étroit entre la régulation du métabolisme et de la virulence (Richardson, 2019).

3.1.3. L'exemple de la régulation par Rex et CcpA en condition de stress NO

Lors d'un stress NO, on observe que *S. aureus* en présence ou en absence d'oxygène produit du lactate par fermentation lactique. Cette caractéristique est propre à *S. aureus* et explique sa résistance au NO, et le distingue d'autres *staphylocoques* commensaux ou moins virulents. En effet, chez *S. aureus*, c'est l'activité de la lactate déshydrogénase Ldh1 qui lui permet de résister au stress NO.

Comme Rex inhibe les gènes de la fermentation (*cf* I.3.1.2.2.1), dont *ldh1*, son inactivation est donc un prérequis en cas de stress NO (Pagels et al., 2010). La transition métabolique vers la fermentation lactique en présence d'oxygène est due à l'inhibition par le NO de l'oxydase terminale de la chaîne de transport d'électrons, arrêtant ainsi la consommation d'oxygène. L'inhibition de la respiration médiée par le NO entraîne une perturbation de l'homéostasie rédox, rétablie par l'activité de Ldh1 (Richardson et al., 2008; Figure 7). Par ailleurs, il a été montré que le glucose est requis pour une activité maximale de Ldh1. La présence de glucose entraîne l'activation de CcpA (*cf* I.3.1.2.2.2). CcpA ne se fixe ni au promoteur de *rex* ni à celui de *ldh1*. Or, il a été montré qu'une délétion de *ccpA* entraîne une diminution de l'activité de Ldh1 et un taux élevé de NADH. Cependant cette délétion accompagnée de celle de *rex* n'affecte pas l'expression de Ldh1. Cela suggère que l'activation de Ldh1 par CcpA est indirecte. Donc l'expression de Ldh1 en présence de NO dépend à la fois de l'inactivation de Rex par le potentiel rédox, et de l'induction par CcpA en présence de glucose (Christmas et al., 2019; Crooke et al., 2013). De manière intéressante, *ldh1* est le seul

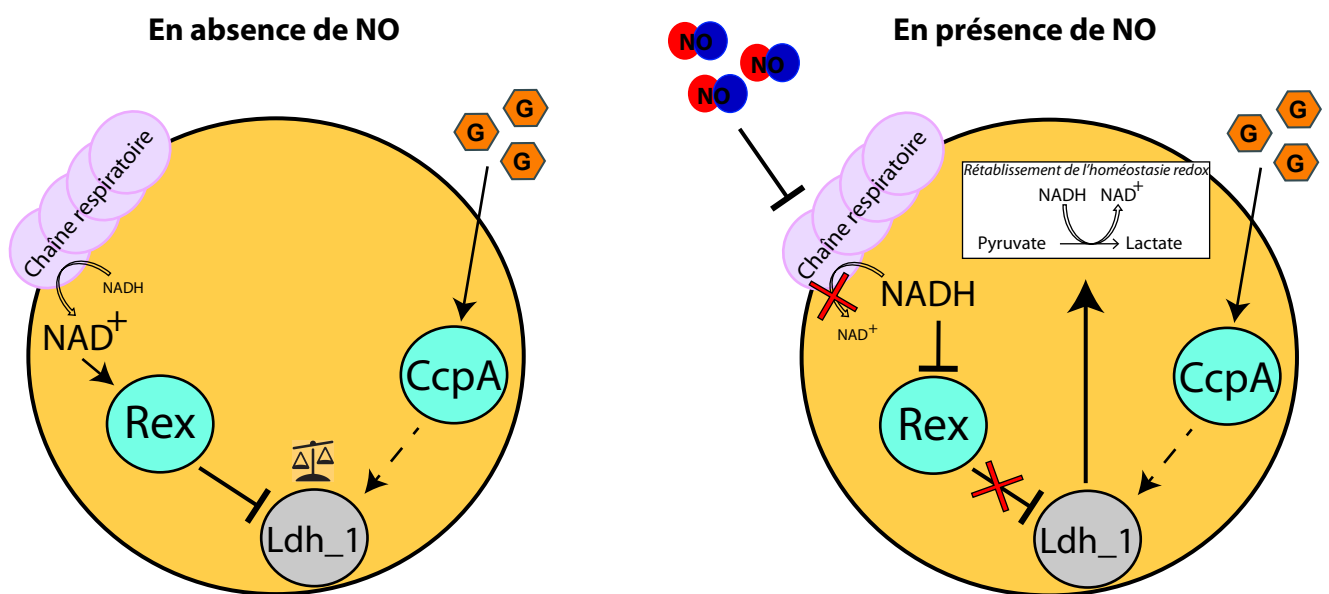


Figure 7 : Réseau de régulation permettant la résistance au stress NO : Les protéines représentées en bleu sont des facteurs de transcription. Les flèches symbolisent une activation, les autres connecteurs une inhibition, et les traits en pointillé représentent une régulation indirecte. Les « G » dans un hexagone représentent le glucose, et les ronds rouges et bleus les radicaux d'oxyde nitrique (NO). La figure présente à gauche la régulation de Ldh_1 en absence de NO et à droite en présence.

gène régulé par Rex qui est également sous le contrôle du glucose et de CcpA (Crooke et al., 2013; Figure 7).

3.2. Les systèmes à deux composants (TCS)

3.2.1. Généralités

Parmi les 135 FT, on retrouve 17 systèmes à deux composantes (TCS). 14 de ces 17 TCS sont strictement conservés dans toutes les souches de *S. aureus* (Ibarra et al., 2013; Tableau 3). Les bactéries utilisent ces TCS pour s'adapter à leurs environnements diversifiés et y donner une réponse à la mesure. Un TCS classique se compose d'une histidine kinase membranaire (senseur) et d'un régulateur de réponse cytoplasmique. Les TCS fonctionnent sur un mécanisme de transduction du signal. Le senseur, en réponse à un stimulus extracellulaire, va subir une autophosphorylation au niveau d'un résidu histidine conservé. L'histidine kinase phosphorylée transfère son groupement phosphate au niveau d'un résidu aspartate conservé du régulateur de réponse. Le régulateur de réponse phosphorylé, sous sa forme active, peut effectuer ses fonctions régulatrices : classiquement la régulation de l'expression des gènes par liaison à l'ADN.

Chaque TCS est une entité de régulation autonome capable de sentir et répondre à une ou plusieurs conditions environnementales particulières. Par exemple, *VraRS* explique à lui seul la résistance aux détergents, *SrrAB* la capacité de croître à 28°C, et *GraRS* la capacité de se développer à faible pH (Villanueva et al., 2018). En revanche, les bactéries rencontrent des environnements variés en fonction des tissus de l'hôte, ainsi l'établissement de la virulence ne peut pas être attribué à un seul TCS. Seul le système *WalRK*, qui régule la division bactérienne, est essentiel pour la viabilité. La délétion d'aucun autre TCS n'a entraîné de défaut de croissance observable (Villanueva et al., 2018), contrairement à ce qu'indiquaient les études de Sun et al., (2005) qui présentaient le TCS *AirSR* (également appelé *YhcSR*) comme essentiel. La fonction de l'ensemble des TCS de *S. aureus* est explicitée dans un tableau récapitulatif (Tableau 3).

La variation du nombre de TCS entre deux espèces peut en partie être expliquée par la nécessité d'adaptation de chacune. Par exemple, *S. saprophyticus*, principalement non pathogène, cause parfois des infections urinaires sans complication. *S. saprophyticus* présente 11 TCS seulement, suffisant à la colonisation de cette niche réduite, à la différence de *S. aureus* qui colonise un grand nombre de tissus. *S. aureus* arbore plus de TCS ce qui lui permet une adaptation spécifique. Des évidences montrent que l'acquisition de TCS additionnels est un

TCS	Nom	Senseur	Régulateur	Fonction régulatrice	Effet sur la virulence	Publication
TCS-1	WalRK	WalK	WalR	Division bacterienne, métabolisme du peptidoglycane	? (essentiel)	Villanueva et al 2018
TCS-2	?	?	?	Spécifique des <i>S. aureus</i> avec SCCmec, homologue de la copie chromosomale de <i>kdpDE</i>	?	Villanueva et al 2018
TCS-3	HptRS	HptS	HptR	Régulation de gènes impliqués dans la consommation d'hexose-6-phosphate (glucose-6 -phosphate)	Survie et répllication intracellulaire	Park et al 2015
TCS-4	LytSR	LytS	LytR	Autolyse	Tolerance à la péniciline	Bronner et al 2004
TCS-5	GraRS	GraS	GraR	Coordonne l'expression des gènes de résistance aux peptides antimicrobiens cationiques	Resistance au systeme immunitaire inné	Farold et al 2011
TCS-6	SaeRS	SaeS	SaeR	Détecte des variations de pH et de concentration en sel(s), contrôle l'expression de facteurs de virulences	Régulation coagulase, et des hemolysines	Bronner et al 2004
TCS-7	?	?	?	?	?	Villanueva et al 2018
TCS-8	ArlRS	ArlS	ArlR	Croissance et division, autolyse	Régulation de facteurs de virulence (lipase, serine protease, coagulase, proteine A...)	Bronner et al 2004
TCS-9	SrrAB	SrrB	SrrA	Régule des proteines impliquées dans le métabolisme energetique, requis lors de la croissance fermentative en conditions anaérobies	Régulation de facteurs de virulence en conditions anaérobie (TSST-1)	Bronner et al 2004
TCS-10	PhoRP	PhoR	PhoP	Adaptation aux faibles concentrations de phosphate inorganique	Un mutant présente une virulence atténuée dans le foie	Kelliher et al 2018
TCS-11	YhcSR (ou AirSR)	YhcS	YhcR	Répond a la présence d'oxygène en régulant un grand nombre de gènes en conditions anaérobies	Modifie la susceptibilité à l'H ₂ O ₂ et plusieurs antibiotiques	Sun et 2013
TCS-12	VraRS	VraS	VraR	Résistance aux dommages de la paroi	Résistance aux détergents, et aux antibiotiques	Villanueva et al 2018
TCS-13	AgrCA	AgrC	AgrA	Sent la densité cellulaire, contrôle l'expression de facteurs de virulence	Régulation des facteurs d'adhesion et de dissémination	Bronner et al 2004
TCS-14	KdpDE	KdpD	KdpE	Sent le K ⁺ extracellulaire, contrôle l'expression de facteurs de virulence	Régulation des facteurs d'adhesion et de dissémination	Xue et al 2011
TCS-15	HssRS	HssS	HssR	Maintien de l'homéostasie cellulaire en hèmes, contrôle l'expression de facteurs de virulence	?	Stauff et al 2007
TCS-16	NreBC	NreB	NreC	Régule le transport et la réduction du nitrate et du nitrite	Aucun	Schlag et al 2008
TCS-17	BraRS	BraS	BraR	Résistance à la nisine et la bacitracine	Résistance à la nisine et la bacitracine	Hiron et al 2011

Tableau 3 : Les TCS de *S. aureus*.

processus dynamique, comme c'est le cas chez certaines souches de *S. aureus* SARM qui ont acquis un nouveau TCS. Le TCS-2 est une duplication chromosomique du TCS KdpDE (Tableau 3), porté sur l'élément mobile SCCmec (cf 1.3.2) (Villanueva et al., 2018).

3.2.2. Les TCS reliés au projet de thèse

3.2.2.1. Le système Agr

Un TCS important chez *S. aureus* pour l'établissement de la virulence est le système à deux composantes AgrCA. Ce TCS sur le locus *agr* (*accessory gene regulator*) code pour un système de densité cellulaire appelé *quorum sensing* (QS). Le QS est un système de communication chimique entre bactéries qui permet de sentir la densité cellulaire. En fonction de la densité cellulaire, une coordination entre les bactéries se met en place, permettant l'expression commune de certains gènes au sein de la population. Ces gènes interviennent aussi bien dans la formation de biofilm, que dans la résistance aux antibiotiques, ou dans la production de facteurs de virulence. Le locus *agr* est composé de deux unités transcriptionnelles divergentes codant pour l'ARNII et l'ARNIII. L'ARNII code pour le TCS AgrCA. AgrC est une histidine kinase membranaire qui sent la concentration de l'auto-inducteur (*tailed thiolactone ring*, AIP), et AgrA est le régulateur de réponse qui permet l'auto-activation du locus *agr*. L'ARNII code également l'auto-inducteur AgrD et la protéase membranaire AgrB qui permet la maturation et l'export de AgrD (Novick & Geisinger, 2008). L'ARNIII est un ARN régulateur qui a largement été décrit car il joue un rôle crucial dans la virulence. Étant régulé par AgrA, il coordonne l'expression de gènes de virulence en fonction de la densité de la population pour permettre le passage de *S. aureus* d'un mode d'adhésion vers un mode de dissémination (Bronesky et al., 2016; Desgranges et al., 2019, I.4. .3 ; Figure 8).

Un grand nombre de facteurs de transcription régulent AgrA. Par exemple, SarA, SarU, SarX et CcpA activent son expression, alors que SrrAB et ArlRS le répriment. Les différents niveaux de contrôle mis en place autour du système *agr* ainsi que le QS reflètent l'importance de ce mécanisme dans la physiologie de *S. aureus* (Novick & Geisinger, 2008).

3.2.2.2. Le système HptRS

Un TCS qui s'est révélé important pour le projet est HptRS (Figure 9) qui est le seul TCS de *S. aureus* répondant à une source de carbone. De manière intéressante, ce TCS répond à la présence de G6P, un sucre clef du métabolisme énergétique et carboné, car il constitue un intermédiaire majeur de la glycolyse. HptS est l'histidine kinase membranaire et HptR est le

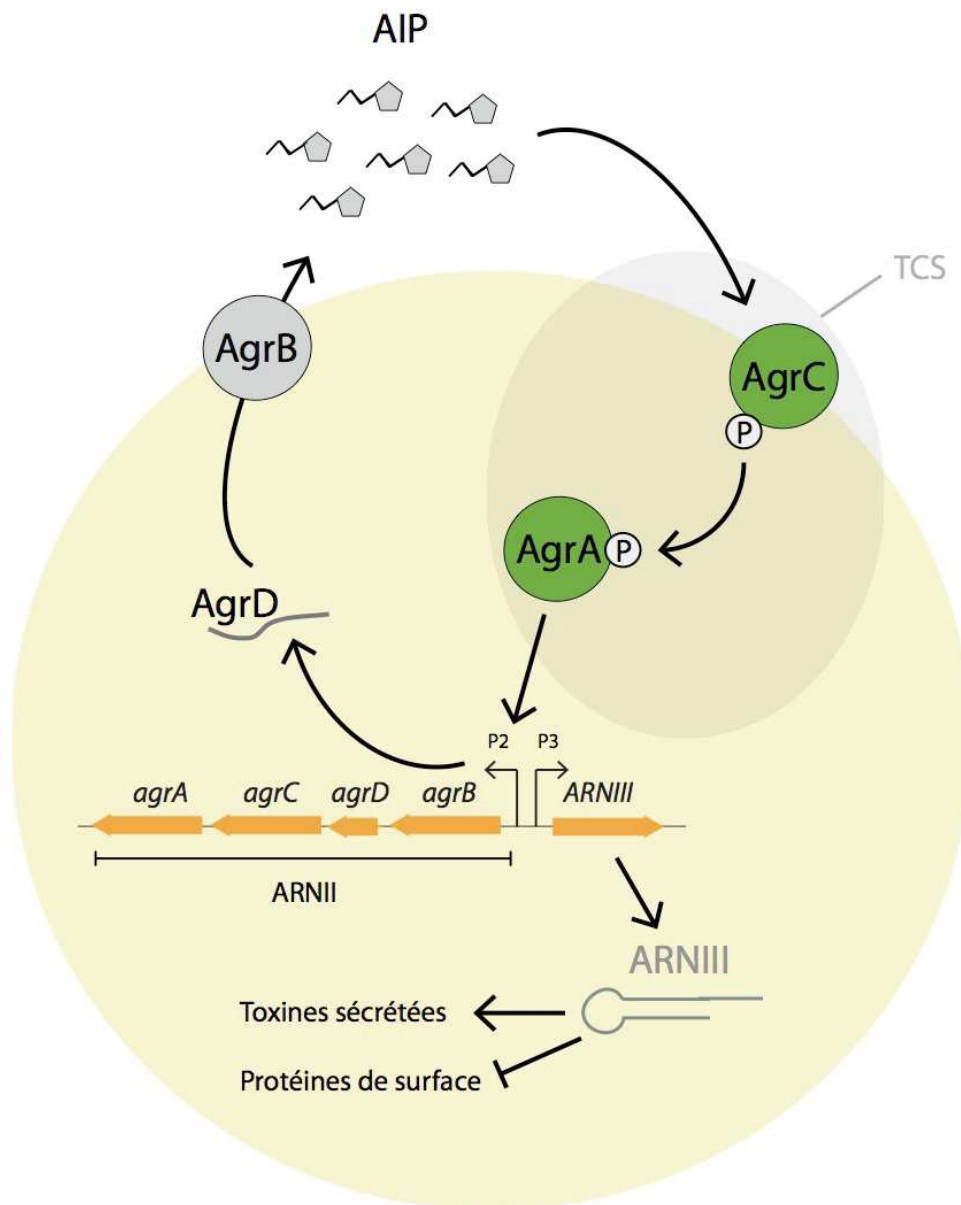


Figure 8 : Le système Agr : Représentation schématique de l'auto-induction du système Agr. Le peptide AgrD est excrété par AgrB sous sa forme mature (AIP). L'AIP est senti par l'histidine kinase AgrC qui après autophosphorylation induit la phosphorylation du régulateur de réponse AgrA. AgrA phosphorylé active les promoteurs P2 et P3, entraînant la transcription de l'ARNII et l'ARNIII. L'ARNIII est un sARN qui régule l'expression des protéines de surface et des toxines sécrétées. En Gris est entourés AgrC et AgrA qui représente le TCS du système Agr. Inspiré de Novick & Geisinger, 2008.

régulateur de réponse cytoplasmique. De plus, une protéine accessoire est également présente dans l'opéron, HptA. C'est une protéine extracellulaire, sécrétée ou accrochée à la membrane, permettant une meilleure perception du G6P et une meilleure transduction du signal par HptRS (Yang et al., 2016). La présence de G6P extracellulaire est donc sentie par HptS permettant l'activation de HptR par sa phosphorylation. HptR activé est capable de reconnaître une séquence particulière appelée HptR BOX. Cette séquence de 30 nucléotides est retrouvée dans le promoteur du gène *uhpT*, localisé entre -67 et -96 de son +1 de transcription. Le gène *uhpT* code pour un transporteur du G6P. La présence de G6P extracellulaire va donc entraîner la transcription *via* HptRS, suivie de la traduction du transporteur UhpT, afin d'importer le G6P dans la bactérie et de l'utiliser comme source de carbone (Yang et al., 2016).

Reed et al., (2018) ont montré que la délétion de ce TCS dans la souche clinique SA564 n'entraîne pas d'effet phénotypique observable sur la croissance, le métabolisme ou le protéome, lorsque cette souche est cultivée en milieu riche (*Tryptic soy broth*, TSB). Cependant, Park et al., (2015) ont clairement montré qu'une souche USA-300 délétée de *hptRS* présente un défaut de croissance dans un milieu pauvre (Muller Hinton II) supplémenté en G6P, en comparaison à une souche sauvage. Ce phénomène est également observable lorsque *S. aureus* est internalisé dans des cellules de l'hôte. En effet, le mutant de *hptRS* présente une survie et une multiplication diminuées en conditions d'internalisation (Park et al., 2015; I.5.2.3).

Enfin, HptRS est également impliqué dans la résistance à la fosfomycine, dans la mesure où *uhpT* est requis pour faire rentrer cet antibiotique dans la cellule. Des mutations dans HptRS s'accompagnent ainsi d'une résistance accrue pour la fosfomycine (Park et al., 2015; Wu et al., 2020).

3.3. Les ARN régulateurs

Les ARN régulateurs (sARN) en complément des FT et des systèmes à deux composants permettent une régulation fine de l'expression génique en répondant à des conditions environnementales changeantes, comme la présence/absence de composés dans l'environnement et les conditions physico-chimiques.

La plupart des sARN sont non-codants, sauf exception de certains sARN qui codent pour des petits peptides et sont dits multifonctionnels, comme c'est le cas pour l'ARNIII chez *S. aureus* qui code pour l'hémolysine δ (Hld). La classe prédominante de sARN agit par appariements de base avec leurs ARN cibles. En général, ils sont codés sur un locus différent

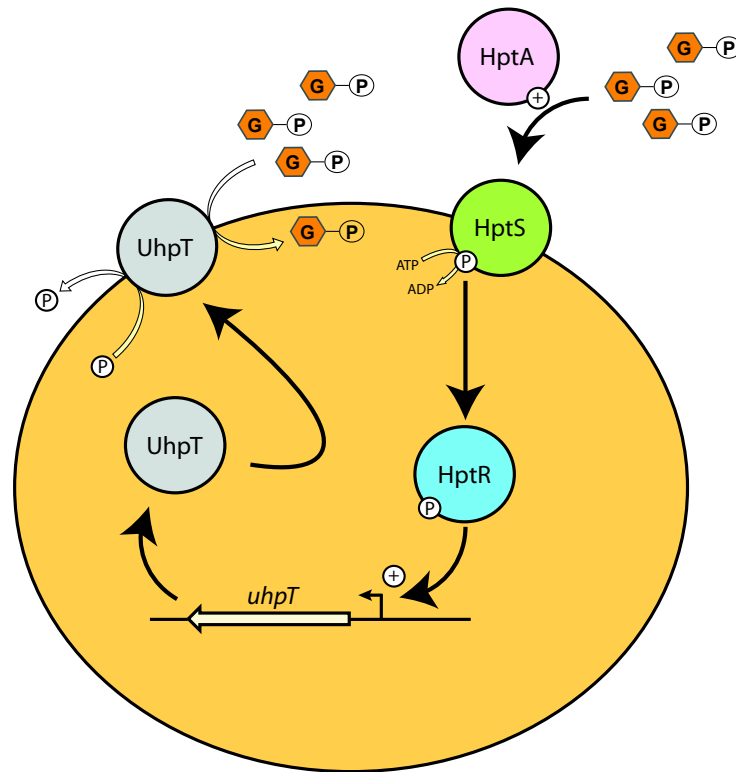


Figure 9 : Transduction du signal par le TCS HptRS : En présence de G6P, le senseur HptS est phosphorylé. La protéine HptA, protéine accessoire encodé dans le locus HptASR (*cf* figure 8), permet d'améliorer la perception du signal. HptS phosphorylé active HptR qui lui-même active la transcription du gène *uhpT* codant pour un transporteur de G6P. La présence de G6P extracellulaire entraîne donc l'augmentation de l'acquisition de G6P.

du gène régulé et forment des appariements de type Watson-Crick non contigus avec leurs cibles. Il en découle qu'un sARN peut réguler plusieurs ARNm, et qu'un ARNm peut être régulé par plusieurs sARN. A l'inverse, les ARN antisens (ARNas) forment un duplex parfait avec leurs cibles puisqu'ils sont produits à partir du brin opposé du gène qu'ils régulent (Caldelari et al., 2013; Desgranges et al., 2019; Richards and Vanderpool, 2011). Les ARN régulateurs sont extrêmement variés. Ils peuvent également agir en cis (riboswitch, T-Box), et sont souvent multi-fonctionnels (antisense et riboswitch, antisense et séquestration de protéine...). Dans ce qui va suivre, je me suis focalisée sur la classe des sARN qui interagissent par appariements de bases avec leurs ARN cibles. Toutefois, la revue Desgranges et al., (2019) (*cf* I.4.3) présente la grande diversité qu'il existe parmi les sARN.

Les sARN représentent la classe majeure des régulateurs post-transcriptionnels. Ils prédominent dans les circuits de régulations établies pour répondre aux changements environnementaux. Cette constatation peut s'expliquer du fait qu'ils apportent de nombreux avantages à la bactérie, et c'est pourquoi ils auraient pu majoritairement être sélectionnés, plutôt que les FT protéiques pour la régulation de certaines voies. Ces avantages sont détaillés par Beisel & Storz, (2010). En voici deux exemples qui me semblent les plus importants. Premièrement, les sARN agissent rapidement en réponse à un stimulus et favorisent un retour à la normale plus rapide après l'arrêt du stimulus. En effet, les sARN, en agissant au niveau post-transcriptionnel, impactent l'expression génique à une étape plus proche de la traduction que les régulateurs transcriptionnels. Par conséquent, moins de temps est requis pour obtenir un impact sur les protéines cibles (Shimoni et al., 2007). L'utilisation d'un sARN réduit le coût énergétique de la régulation. En effet, comme l'indique leurs noms, les ARN régulateurs sont composés d'ARN. Par ailleurs, ils sont souvent de petites tailles majoritairement, allant de 50 à 500 nt (globalement plus courts que les transcrits codant pour des protéines). De plus, ces ARN n'entraînent pas de coût énergétique pour leur traduction à l'inverse des FT. Ainsi, en conditions de stress transitoire, l'utilisation des sARN a sûrement été sélectionnée pour permettre une régulation rapide et à faible coût métabolique (Altuvia & Wagner, 2000; Beisel & Storz, 2010).

4. Les ARN régulateurs

4.1. Un peu d'histoire

Alors qu'il y a 80 ans Jacques Monod et François Jacob ont proposé que l'ARN pourrait jouer le rôle d'intermédiaire pour donner le message de l'information génétique entre l'ADN

et les protéines, le dogme central de la biologie qui repose sur le fait que l'ARN est seulement un porteur de l'information nécessaire à la synthèse des protéines émerge. Ainsi, en 1961, Monod et Jacob, bousculent ce dogme en émettant l'hypothèse qu'un ARN antisens pouvait être à l'origine de la régulation de l'opéron lactose. Même si ce modèle de régulation s'est avéré faux, ils furent néanmoins les premiers à énoncé la notion d'ARN antisens comme régulateur. Ce n'est que dans les années 80, que ce concept a été vérifié expérimentalement. Un des premiers ARN non-codant identifié est l'ARNI de *E. coli*, responsable de la régulation de la réplication du plasmide ColE1 (Tomizawa & Itoh, 1981). Bien que l'ARN 6S ait été découvert 14 ans avant l'ARNI (Hindley, 1967) ce n'est qu'une trentaine d'années après sa découverte que sa fonction a été définie. Il a ainsi été montré que l'ARN 6S agit par séquestration de protéine plutôt que comme ARN antisens et qu'il permet de rediriger l'ARN polymérase vers des facteurs sigma alternatifs en séquestrant le facteur sigma 70 (Wassarman & Storz, 2000).

La notion d'ARN antisens a ainsi été caractérisée chez les bactéries bien avant la découverte des microARN (miARN) et ARN interférents (siARN) chez les eucaryotes. En effet les premiers miARN ont été découverts chez le nématode *Caenorhabditis elegans* entre le début des années 1990 et 2000. Il s'agit de *lin-4* et *let-7*, tous deux impliqués dans le contrôle du développement de *Caenorhabditis elegans* (Lagos-Quintana, 2001; Lee et al., 1993).

Ce n'est que dans les années 2000, que l'on a réellement pris conscience que les parties non codantes des génomes bactériens pouvaient générer des ARN. Avec le développement du séquençage des génomes, trois équipes ont émis l'hypothèse que par analogie aux protéines, les régions intergéniques conservées devaient être fonctionnelles (Argaman et al., 2001; Rivas and Eddy, 2001; Rivas et al., 2001; Wassarman, 2001). Ces études pionnières ont permis de mettre en avant les sARN chez *E. coli*. Depuis, la littérature explose dans ce domaine de recherche, révélant de multiples fonctions et mécanismes d'action des ARN régulateurs dans les trois domaines du vivant.

4.2. Aspect évolutif

4.2.1. Évolution, conservation des ARN régulateurs

Actuellement, l'origine des sARN, leur divergence fonctionnelle au cours de l'évolution, voire même leur disparition des génomes bactériens, est peu comprise. L'étude de l'évolution des sARN est clef dans la compréhension de l'évolution des voies métaboliques, de l'adaptation à l'environnement, et de la virulence, mais elle a aussi été déterminante pour leur découverte et pour l'amélioration de l'annotation des génomes (Dutcher & Raghavan, 2018).

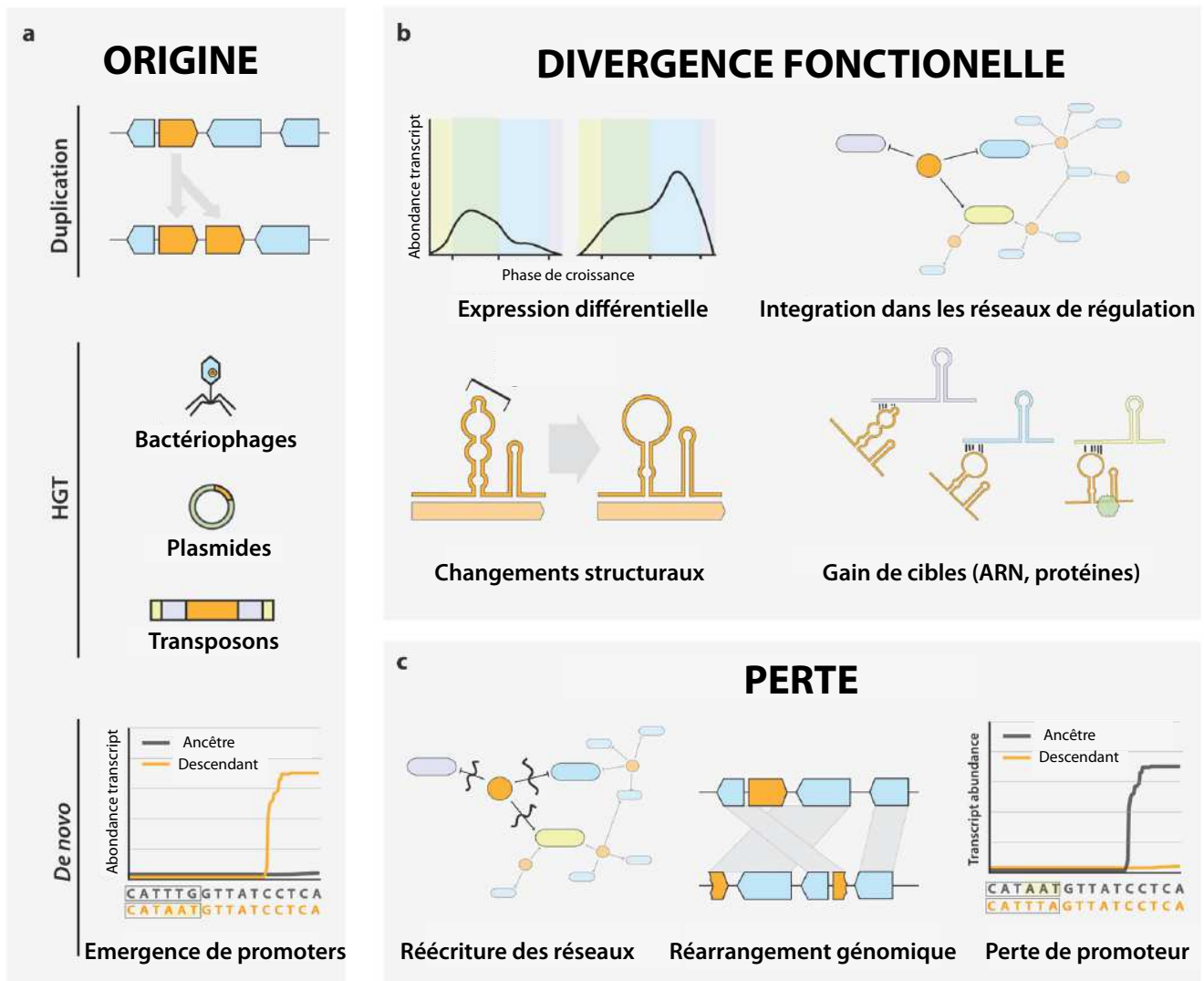


Figure 10 : Évolution des ARN régulateurs au sein des génomes bactériens : a) L'apparition de sARN peut être le résultat d'événements de duplication du génome, de transferts horizontaux (HGT, par des bactériophages, des plasmides, des transposons), ou une émergence *de novo* suite à l'apparition d'un promoteur. b) Les sARN peuvent subir des divergences fonctionnelles telles qu'une modification de leurs conditions d'expression, de leurs cibles, ou de leur structure. c) La perte d'un sARN se produit lorsque des pressions sélectives modifient son réseau d'interaction, que des mutations entraînent la perte de son promoteur, ou que des réarrangements du génome modifient la région intergénique depuis laquelle le sARN est transcrit. D'après Dutcher et al., 2018.

Trois mécanismes peuvent mener à l'apparition de sRNA. Nous parlons ici des ARN non codants agissant en *trans*. Premièrement leur émergence peut être le résultat d'un évènement de duplication, comme c'est souvent le cas pour les séquences codantes. Deuxièmement, le transfert horizontal de gènes (HGT) via des plasmides, des bactériophages ou des éléments transposables mobiles est une source commune d'apparition de sARN (Dutcher & Raghavan, 2018). C'est le cas de SprD chez *S. aureus*, un ARN exprimé depuis un îlot de pathogénicité, qui contient des gènes de virulence et de résistance aux antibiotiques. SprD acquis par transfert horizontal régule l'ARNm *sbi* (*immune evasion proteïne*) encodé par le *core* génome (Chabelskaya et al., 2010). Enfin, une origine *de novo* des sARN est également possible à partir de transcrits, au premier abord, non fonctionnels. Cela est dû, entre autres, à la réorganisation du génome via recombinaison homologue qui peut créer des nouvelles régions inter-géniques (IGR) desquelles de nouveaux sARN peuvent émerger. De plus, des mutations ponctuelles dans les IGR peuvent entraîner la transcription d'un ARN par la création d'un promoteur fonctionnel.

Une fois l'ARN régulateur apparu au cours de l'évolution il peut avoir plusieurs destins. Ainsi de nouvelles mutations (au niveau de la « *seed region* » ou modifiant la structure) vont pouvoir entraîner un gain de fonction (par exemple de nouvelles cibles), une divergence fonctionnelle (par exemple un changement de son contrôle transcriptionnel, permettant l'expression dans de nouvelles conditions) ou une perte du sARN (Dutcher & Raghavan, 2018). Ces différents destins vont s'expliquer par la pression de sélection exercée pour retenir ou perdre ces sRNA. Dans l'exemple d'un sARN qui régulerait une ou plusieurs cibles qui viendraient à disparaître (comme c'est le cas par exemple pour des pathogènes qui au cours de l'évolution changent de style de vie en devenant intracellulaires, cela s'accompagne souvent par la perte de gènes codant pour des protéines), celui-ci va pouvoir évoluer afin d'acquérir de nouvelles cibles, ou alors disparaître à cause de l'absence de pressions sélectives pour son maintien. Par exemple, EcsR1 de *E. coli* est absent de *Salmonella enterica* suite à un réarrangement génétique au niveau de l'IGR qui l'encode. Ce sARN intervient chez *E. coli* dans la régulation de la formation des biofilms, une caractéristique qui n'est pas avantageuse pour le style de vie de *S. enterica* (Raghavan et al., 2015). Au cours de l'évolution, il existe donc une réécriture des réseaux de régulation. Ceux-ci s'agrémentent de nouveaux membres ou se débarrassent de certains, certainement pour permettre une meilleure adaptation à de nouveaux environnements (Dutcher & Raghavan, 2018); Figure 10).

4.2.2. Évolution, conservation des régions 3'UTR

Nous allons porter maintenant une attention particulière à l'évolution des régions 3'UTR. L'analyse du transcriptome de *S. aureus* a révélé qu'au moins un tiers de ces ARNm comportent de longues régions 3'UTR (Menendez-Gil et al., 2020; Ruiz de los Mozos et al., 2013). De plus, ces régions contiennent des éléments régulateurs qui modulent l'expression génique. En effet, on pourra y trouver des *riboswitch*, des promoteurs internes pour la transcription de sARN, ainsi que des sites de clivage de RNases pour libérer des ARN régulateurs (Miyakoshi et al., 2015).

Un exemple de région 3'UTR ayant un rôle régulateur, est la région 3'UTR de l'ARNm *icaR* de *S. aureus* qui code pour un répresseur des biofilms. La région 3'UTR de *icaR* est longue de 390 nucléotides, et agit en *cis* sur son propre ARNm. La région 3'UTR de *icaR* en interagissant par appariement de bases avec le site de reconnaissance du ribosome en 5'UTR, empêche ainsi sa propre traduction, et induit la dégradation de son ARNm par l'endoribonucléase III (RNase III), qui cible spécifiquement les duplex ARN-ARN. Ce mécanisme original permet alors de moduler la synthèse de ce facteur de transcription en réponse à un signal encore méconnu (Ruiz de los Mozos et al., 2013). Un exemple similaire d'interaction entre régions 3'UTR et 5'UTR a également été décrit pour l'ARNm *hbs* de *Bacillus subtilis*. Cette interaction permet de coupler la stabilité de l'ARNm *hbs* à l'initiation de sa traduction (Braun et al., 2017). La présence d'une longue région 3'UTR peut également entraîner une régulation en *trans*, grâce à l'émergence d'un sARN depuis cette région (Chao and Vogel, 2016; Chao et al., 2012; De Mets et al., 2019; Kim et al., 2014; Lalaouna et al., 2019; López-Garrido et al., 2014; Wang et al., 2020). Cet aspect des ARN régulateurs est présenté dans la suite de l'introduction (revue Desgranges et al., 2019 ; I.4.3), et la caractérisation et l'évolution de l'un d'eux sera abordée avec un exemple concret dans la partie résultat (cf II.2).

Bien que le rôle régulateur des régions 3'UTR est communément admis, la compréhension de leur évolution et leur implication dans la divergence des espèces bactériennes est très limitée. Un effort de compréhension de l'évolution de ces régions chez les *staphylocoques* a été réalisé récemment dans l'étude de Menendez-Gil et al., (2020). Cette étude a montré qu'il existe une grande divergence de séquence et de taille des régions 3'UTR de gènes orthologues (gènes communs à différentes espèces, provenant d'une évolution commune, et ayant conservé leurs fonctions). Par exemple, un manque de conservation des régions 3'UTR entre *S. aureus* et *S. argenteus* est perceptible, bien que *S. argenteus* soit l'espèce la plus proche de *S. aureus*. La perte de la conservation est considérablement augmentée chez des espèces

phylogénétiquement plus éloignées telles que *S. simiae*, *S. capitis* et *S. epidermidis*. Ces espèces ont conservé seulement quelques 3'UTR en commun avec *S. aureus*. Ce constat a également été élargi à d'autres genres, montrant que de manière générale, la conservation des ARNm étudiés est perdue en aval de leur région codante. Les régions 3'UTR sont donc plus aptes aux variations que les séquences codantes et que les régions 5'UTR qui portent la séquence Shine et Dalgarno. Par exemple, l'ARNm *icaR* présente de grandes variations en termes de séquence et de taille en fonction des espèces de *staphylococques* étudiées. Chez *S. aureus*, l'inhibition de la traduction de *icaR* est rendue possible par l'interaction entre le motif UCCCC de la région 3'UTR avec le site de reconnaissance du ribosome (RBS) en 5'UTR (*cf* ci-dessus). Cependant ce motif n'est pas retrouvé chez tous les *staphylocoques*. Il est présent chez les espèces *S. simiae* et *S. argenteus* et absent de *S. epidermidis* et *S. capitis*, ce qui montre que d'une espèce à une autre les variations dans les régions 3'UTR peuvent créer des différences d'expression génique via des modifications dans les mécanismes de régulation. De plus, la construction d'ARNm chimère de *S. aureus*, où le 3'UTR a été remplacé par celui d'un gène orthologue a montré que la variation de cette région ne permet pas de retrouver le niveau d'expression de la protéine encodée par cet ARNm chez *S. aureus*. Cela suggère, qu'en fonction des espèces les 3'UTR pourraient avoir des rôles distincts (Menendez-Gil et al., 2020).

Les modifications dans les régions 3'UTR sont dues à des mutations ponctuelles, des réarrangements du génome et des phénomènes de HGT (Menendez-Gil et al., 2020). L'étude de l'évolution de ces régions, au vu de leur grande divergence entre espèces et des éléments régulateurs qui s'y cachent, est importante. Elles pourraient nous donner des clefs pour mieux comprendre pourquoi certains *staphylocoques* sont commensaux alors que d'autres sont pathogènes avec des niveaux de virulence variables d'une espèce à l'autre.

4.3. Les ARN régulateurs chez *S. aureus*

Les ARN régulateurs, largement distribués dans les génomes bactériens, sont également présents chez les bactéries pathogènes pour acheminer des fonctions similaires à celles des sARN chez les bactéries non pathogènes ainsi que des fonctions additionnelles pour le contrôle de la virulence. La présente revue « Non-coding RNA » a pour but de présenter de manière non exhaustive les différentes classes d'ARN non-codants retrouvés chez *S. aureus*. Cette revue résume ce qui était connu sur les sARN de *S. aureus* jusqu'en 2018. À la suite de cette revue, je ferai une mise à jour des connaissances récentes sur les sARN de *S. aureus* qui seront utiles pour la suite du manuscrit.

Noncoding RNA

E. DESGRANGES,¹ S. MARZI,¹ K. MOREAU,²
P. ROMBY,¹ and I. CALDELARI¹

¹Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, F-67000 Strasbourg, France

²CIRI, International Center for Infectiology Research, Inserm, Université Claude Bernard Lyon 1, CNRS, UMR5308, École Normale Supérieure de Lyon, Hospices Civils de Lyon, University of Lyon, F-69008, Lyon, France

ABSTRACT Regulatory RNAs, present in many bacterial genomes and particularly in pathogenic bacteria such as *Staphylococcus aureus*, control the expression of genes encoding virulence factors or metabolic proteins. They are extremely diverse and include noncoding RNAs (sRNA), antisense RNAs, and some 5' or 3' untranslated regions of messenger RNAs that act as sensors for metabolites, tRNAs, or environmental conditions (e.g., temperature, pH). In this review we focus on specific examples of sRNAs of *S. aureus* that illustrate how numerous sRNAs and associated proteins are embedded in complex networks of regulation. In addition, we discuss the CRISPR-Cas systems defined as an RNA-interference-like mechanism, which also exist in staphylococcal strains.

GENERAL INTRODUCTION

Regulatory RNAs have been identified in many bacteria and in pathogenic bacteria such as *Staphylococcus aureus*, where they play major roles in the regulation of virulence or the synthesis of metabolic proteins, besides transcriptional factors and two-component systems (1–4). Most of them are noncoding RNAs (sRNAs), but some of them express small peptides. Certain sRNAs, acting in *cis*, are situated at the 5' untranslated regions (UTRs) of mRNAs and act as sensors of metabolites, tRNA, or environmental stimuli (e.g., temperature, pH) or are situated at the 3' UTR. In contrast, the genes encoding sRNAs, which act *in trans*, sit on the opposite strand of the regulated mRNA or at genomic locations distant from the mRNAs they regulate. *Cis*-encoded sRNAs, also called antisense RNAs (asRNAs), are fully complementary to their targets. In contrast, *trans*-encoded sRNAs share only partial complementarity, and as a consequence, they can regulate many mRNAs. Most of them are encoded mainly in the core genome, while a few of them are localized within mobile

elements, pathogenic islands, or plasmids. In this review, we will focus on the most recent mechanisms of RNA regulation discovered in *S. aureus* and how regulatory RNAs are part of sophisticated networks that allow the bacteria to adapt quickly to their environment or survive in their host.

mRNA 5' UTRs: RIBOSWITCHES, T-BOXES, AND THERMOSENSORS

5' UTRs of mRNAs contain riboswitches, T-boxes, or thermosensors with potential impacts for novel anti-biotherapy. Riboswitches and T-boxes are found in the 5' UTR of some mRNAs and contain highly structured domains, which recognize metabolites such as cofactors, vitamins, amino-acids, nucleotides, second messenger cyclic di-GMP, Mg²⁺, or nonaminoacylated tRNAs (5). Binding of these metabolites induces structural changes that modify the expression of the downstream mRNA,

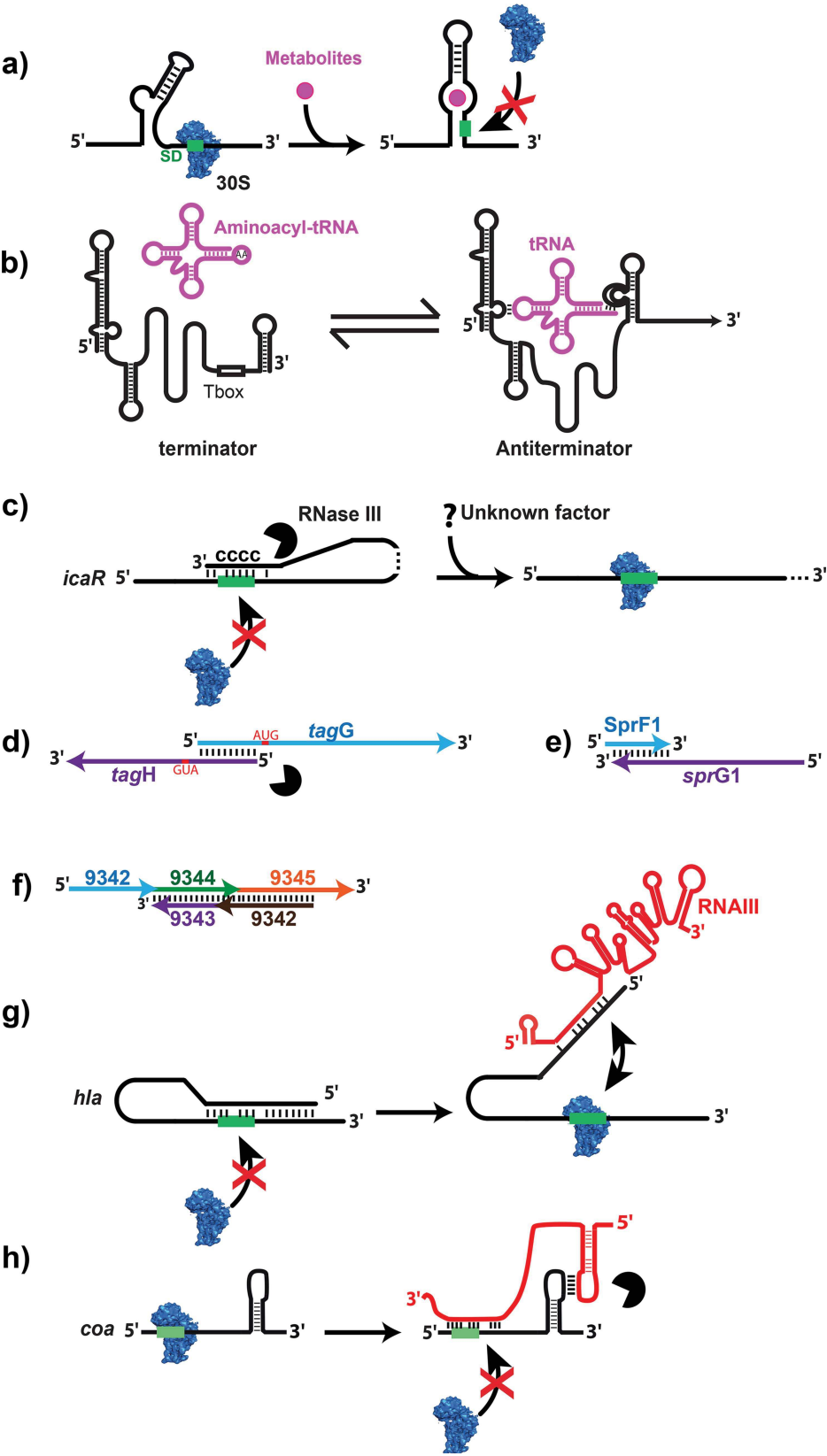
Received: 7 August 2018, **Accepted:** 25 October 2018,

Published: 19 April 2019

Editors: Vincent A. Fischetti, The Rockefeller University, New York, NY; Richard P. Novick, Skirball Institute for Molecular Medicine, NYU Medical Center, New York, NY; Joseph J. Ferretti, Department of Microbiology & Immunology, University of Oklahoma Health Science Center, Oklahoma City, OK; Daniel A. Portnoy, Department of Molecular and Cellular Microbiology, University of California, Berkeley, Berkeley, CA; Miriam Braunstein, Department of Microbiology and Immunology, University of North Carolina-Chapel Hill, Chapel Hill, NC, and Julian I. Rood, Infection and Immunity Program, Monash Biomedicine Discovery Institute, Monash University, Melbourne, Australia

Citation: Desgranges E, Marzi S, Moreau K, Romby P, and Caldelari I. 2018. Non-coding RNA. *Microbiol Spectrum* 7(2):GPP3-0038-2018. doi:10.1128/microbiolspec.GPP3-0038-2018.

Correspondence: Isabelle Caldelari, i.caldelari@ibmc-cnrs.unistra.fr
© 2018 American Society for Microbiology. All rights reserved.



for example, by inducing premature transcription arrest, repression/activation of translation, or cleavage (Fig. 1a). A T-box senses the aminoacylation status of tRNAs and mainly controls transcription of downstream genes that encode proteins involved in biosynthesis, transport of amino acids, or aminoacylation of tRNAs (Fig. 1b). Based on sequence and structure conservation, most of the T-boxes and riboswitches were predicted in *S. aureus* genomes (4). A large proportion of riboswitches control the expression of genes involved in metabolic pathways. Because these genes are often essential for growth, they represent interesting targets for the development of alternative antimicrobial drugs in the battle against multidrug-resistant *S. aureus*. This strategy was used with the guanine and glucosamine-6-phosphate (GlcN6P) riboswitches. Mulhbach et al. (6) identified a pyrimidine derivative that binds to the guanine riboswitch and represses the expression of *guaA*. This compound significantly attenuated *S. aureus* infections in a mouse model. In Gram-positive bacteria, the *glmS* mRNA is both a ribozyme, which catalyzes its own cleavage, and a riboswitch responding to GlcN6P. Its product encodes an essential enzyme, which converts fructose-6-phosphate into GlcN6P, a building block of bacterial peptidoglycan. Tight regulation of *glmS* mRNA is crucial to maintain a homeostatic level of GlcN6P in the cell. At high concentrations of GlcN6P, its binding to the 5' UTR of *glmS* leads to site-specific self-cleavage, which generates a 5' hydroxylated end molecule rapidly degraded by the RNase J1 (7, 8). A recent study led to the design, synthesis, and characterization of a GlcN6P analogue, carba-GlcN6P, which constitutively activates the

glmS ribozyme of vancomycin-resistant *S. aureus* and destabilizes its mRNA (7). This compound was experimentally shown to induce the efficient self-cleavage of the *glmS* mRNA in a similar fashion as the natural metabolite and thus represented an important step in the development of antibiotics with a new mode of action. Very recently, a new approach called "Term-seq" revealed that several antibiotic resistance genes are under the control of riboswitches responding to antibiotics commonly used against Gram-positive pathogenic bacteria such as *Listeria* spp. and *Enterococcus faecalis* (9). These results suggest that the same phenomenon could exist in *S. aureus* and that RNA-mediated regulation could play a broader role in antibiotic resistance mechanisms than has been envisioned.

Not surprisingly, the presence of antibiotics can also modulate the regulatory activity of T-boxes. In *S. aureus*, an unusual *glyS* T-box regulates transcription antitermination (Fig. 1b) of the unique glycyl-tRNA synthetase (*glyRS*) gene responsible for catalyzing the aminoacylation of the five tRNA^{Gly} isoacceptors, independently of their anticodon (GCC or UCC) and with different binding affinities (10). Thereafter, the T-box senses the availability of glycine not only for its incorporation into nascent polypeptide chains during translation but also for the formation of pentaglycine bridges into the peptidoglycan molecule, linking two essential pathways. Antibiotics targeting the small ribosomal subunit stabilized the T-box/tRNA complex and induced a read-through of transcription, while chloramphenicol and linezolid attenuated *glyS* transcription (11). The outcome depended on the binding sites of the protein syn-

FIGURE 1 Several mechanisms of RNA regulation in *S. aureus*. **(a)** Schematic drawing of the flavin mononucleotide riboswitch. The 5' UTR adopts a particular structure recognized by the flavin mononucleotide, which in turn leads to the stabilization of a stem-loop structure sequestering the SD sequence to inhibit translation. 30S is for the small ribosomal subunit. **(b)** An example of a T-box motif as found in the 5' UTR of many mRNAs encoding aminoacyl-tRNA synthetases. Nonaminoacylated tRNA binds to the leader region at two sites and stabilizes an antiterminator structure, allowing transcription of the downstream gene. The drawing is adapted from reference 4. **(c)** The 3' UTR of the biofilm repressor *IcaR* possesses a cytosine-rich motif, which binds to the SD sequence and hinders ribosomes from its binding site on the mRNA (see text for details). **(d)** Overlapping 5' UTRs of *tagG* and *tagH* mRNAs are processed by the endoribonuclease III (Rnase III). Shorter 5' ends might facilitate ribosome recruitment. **(e)** The antitoxin RNA *SprF1* interacts at the 3' end of the toxin encoded by *sprG1* and triggers its degradation. **(f)** A cluster of five sRNAs was sequenced in the *S. aureus* Newman strain that encodes a putative toxin-antitoxin system (see text for details). **(g,h)** sRNAs act by an antisense mechanism. Binding of the 5' UTR of *RNAIII* to the 5' UTR of *hla* mRNA liberates its SD and activated translation (g), whereas the 3' domain of *RNAIII* acts as a repressor domain, which contains C-rich motifs for base-pairing with the SD sequence of mRNA as *coa* mRNA depicted in the figure (h). Green bar, SD sequence; black circle, RNase III (for references and more details, see text).

thesis inhibitors (11). Although T-boxes can be direct targets for antibiotics against *S. aureus*, it was also reported that a high concentration of the antibiotic tetracycline might induce possible off-target inhibition of the RNA polymerase (11).

RNA thermosensors are regulatory elements often localized at the 5' UTR of mRNAs encoding heat or cold shock proteins and virulence factors (for review see 12). Briefly, at low temperatures, the mRNA cannot be translated, since the Shine and Dalgarno (SD) sequence is trapped in a hairpin structure, which melts gradually when the temperature increases. The best-studied example in Gram-positive bacteria is the thermosensor regulating the expression of the transcriptional factor *prfA*, which activates the expression of most of the virulence genes in *Listeria monocytogenes* at high temperatures (13, 14). No such example has yet been demonstrated for other Gram-positive bacteria, including *S. aureus*.

3' UTRs OF mRNAs ACT IN CIS OR ARE RESERVOIRS OF sRNAs

Transcriptome analysis of the human pathogen *S. aureus* revealed that at least one-third of mRNAs carry long 3' UTRs and thus might display multiple regulatory functions (15). Some of them have direct action on the expression of their own mRNA (15, 16). Long 3' UTRs (>100 nucleotides) can end at an intrinsic Rho-independent terminator of transcription (TT) and can be generated from a specific RNase cleavage or from a termination read through of the RNA polymerase. Remarkably, several 3' UTRs contain riboswitches functioning in metabolite-sensing regulation (16). Indeed, the TT of the riboswitch in an OFF conformation also serves as the TT of the gene encoded upstream of the riboswitch.

A singular example of a newly identified posttranscriptional regulatory mechanism was shown for the *icaR* mRNA (Fig. 1c), which possesses an unusually long 3' UTR (390 bp). In this case, the expression of the mRNA is modulated through a long-distance interaction between 3' UTR and 5' UTR (16). Because *icaR* codes for a transcriptional repressor of the *icaADBC* operon, encoding enzymes involved in the synthesis of PIA-PNAG, the main polysaccharides of the biofilm matrix, this regulation has a direct impact on biofilm formation. Base-pairing interactions between the long 3' UTR of *icaR* mRNA and the SD sequence of the same mRNA hindered efficient translation initiation (16). This long-range RNA duplex generated a specific site for the

double-stranded endoribonuclease III (RNase III) for cleavage. As a consequence of this cleavage, PIA-PNAG synthesis increased. However, the mechanism allowing IcaR translation is not known to prevent the action of the 3' UTR as a *cis*-acting antisense RNA (Fig. 1c).

It is also postulated that the 3' UTR length provides other types of transcript-specific regulation. Indeed, in *Salmonella* spp., the 3' UTRs are also reservoirs for sRNAs, which originate either by transcription from an internal promoter or by processing. In both cases, the sRNA generated from the 3' UTR regulates trans-encoded mRNA targets. For instance, the sRNA CpxQ is generated by RNase E cleavage of the 3' UTR of *cpxP* and represses the translation of mRNAs encoding a family of envelope proteins, whereas the sRNA DapZ is transcribed from an internal promoter within the *dapB* gene and inhibits translation of the major ABC transporters, DppA and OppA (17, 18; for review see 19). Recent work has shown that such 3' UTR-derived sRNAs also exist in *S. aureus*, although their functions remain to be addressed (E. Desgranges, S. Marzi, P. Romby and I. Caldelari, unpublished data).

asRNAs IN PERVASIVE TRANSCRIPTION AND ACTING AS ANTITOXINS

asRNAs are transcribed from the opposite strand of the mRNAs they regulate, so that they display perfect complementarities with their targets. Short asRNAs are often encoded on mobile elements such as plasmids, transposons, and phage-like elements. Such elements can potentially be transferred horizontally to other bacterial species or be duplicated (20). In *S. aureus*, they were first described to control plasmid conjugation and replication (21). The size of asRNAs can vary from 10 to thousands of nucleotides, because these RNAs can overlap with part of a gene (3' or 5' ends), the entire gene, or a group of genes (Fig. 1d to f). This phenomenon is called pervasive transcription. Initially, pervasive transcription was considered a nonfunctional transcriptional noise. However, considering the large number of asRNAs expressed from the entire genome and in several bacterial species, these RNAs might play an important role in the regulation of gene expression. Genome-wide analysis of *S. aureus* highlighted that the expression of a significant proportion (75%) of antisense transcripts to annotated open reading frames are synthesized from the complementary strand and that these sense/antisense duplexes are digested by RNase III, generating short fragments all along the genome (22). Another study using the RIP-Seq approach confirmed the involvement of RNase III in the

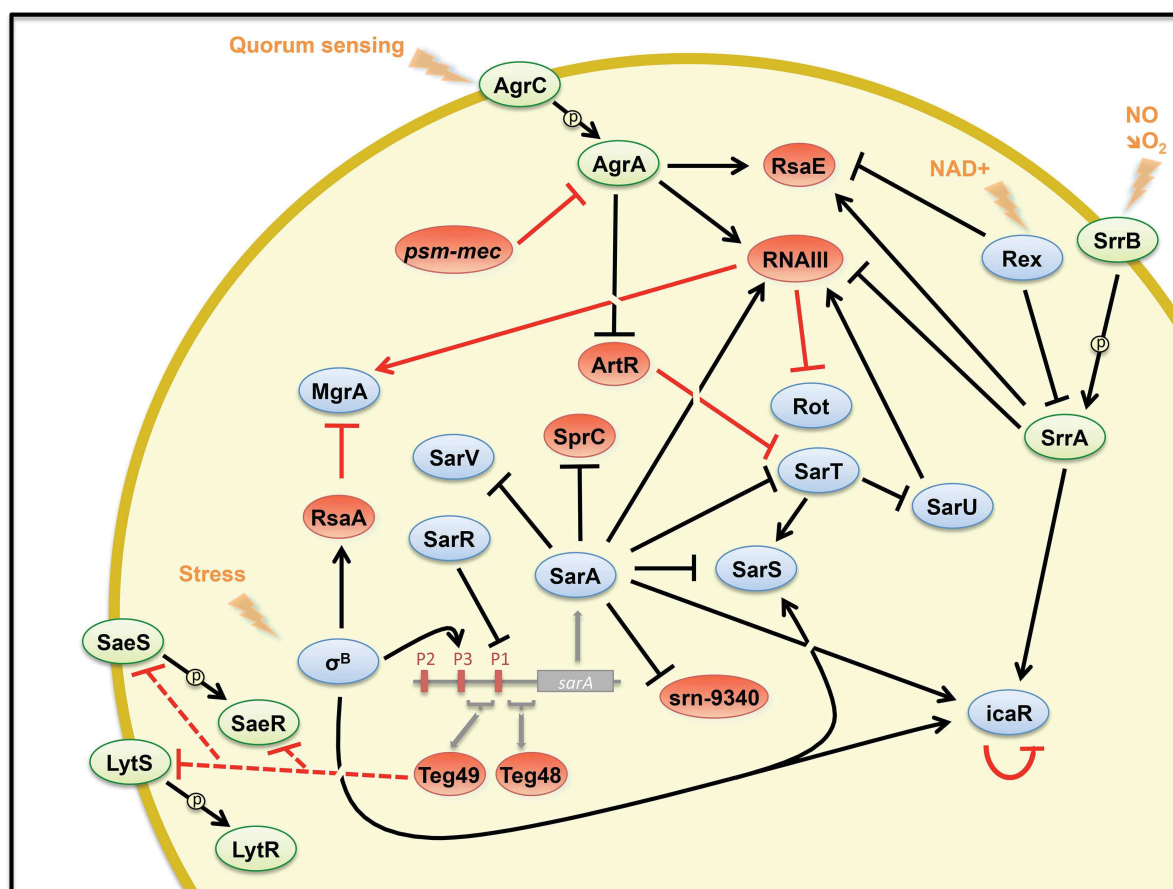


FIGURE 2 Examples of the complex network between sRNAs and transcriptional factors in *S. aureus* in response to stress. Arrows show activation and bars show repression. Blue, transcriptional regulators; green, two-component systems; red, regulatory sRNAs. Red lines corresponded to posttranscriptional regulation, and black lines, to transcriptional regulation. Dotted lines are for the target mRNAs that were not experimentally validated. Only sRNA-dependent mRNA targets encoding transcriptional factors are depicted in the figure.

regulation of sense/antisense transcripts and overlapping UTRs (23). The situation might be even more complex, because recent results suggested that the termination factor Rho plays a major role in preventing pervasive transcription in *Bacillus subtilis*, but also in *S. aureus* (24, 25). Although the biological outcome of pervasive transcription is not clearly understood, some of the asRNAs produced are functional and control several biological processes (15).

Type I toxin-antitoxin (TA) systems are particular cases of short asRNAs, in which the antitoxin is an asRNA regulating the translation or the fate of the toxin-encoding mRNA, whereas in the type III system, the antitoxin sequesters the toxin (reviewed in [26](#), [27](#)). In *S. aureus*, several type I TA module systems have been described (reviewed in [28](#)). One of them, called SprF1/

SprG1, expresses SprF1 as the asRNA and *sprG1* mRNA (Fig. 1e), which encodes two short secreted peptides with hemolytic and antibacterial activity (29). The antitoxin SprF1 binds to the 3' end of *sprG1* mRNA, which leads to mRNA degradation and inhibition of peptide synthesis to protect cells against lethality (29). The SprA1/asSprA1 pair is another intriguing and unconventional system. The asSprA1 is transcribed from the opposite strand of the *sprA1* mRNA, producing a cytolytic peptide. Their 3' ends overlapped by 35 nucleotides, but experimental data indicated that the functional domain of asSprA1 is outside the complementary sequence with *sprA1* mRNA. Both RNAs were expressed concomitantly, and asSprA1 5' base-paired with the ribosome-binding site (RBS) of *sprA1*, impairing its translation (30). Thus, asSprA1 acts as a *trans* regulator with its

complementary target, suggesting that it can potentially interact with other RNA targets. Finally, a cluster of five genes encoding sRNAs specific to the *S. aureus* Newman strain contained a putative TA system (31) (Fig. 1f). Three genes were transcribed from the positive strand and two from the negative strand. Moreover, one small open reading frame was detected within one of the genes from the minus strand and coded for a secreted peptide with similarity to the RelE toxin (32). Whether these two overlapping genes corresponded to a novel TA system remains to be addressed. Interestingly, this locus was expressed in a growth-phase-dependent manner, in nutrient starvation, and in oxidative stress. Type I TA systems have been involved in many functions (e.g., membrane depolarization, plasmid maintenance), mainly in *Escherichia coli*, including the persistence phenomenon (for review see 33), but not in *S. aureus* until now (28).

TRANSCRIPTIONAL FACTORS AND sRNAs BUILD COMPLEX REGULATORY NETWORKS

The sRNAs belong to intricate networks of regulation, and their synthesis is often dependent on transcription factors or on two-component systems. In addition, sRNAs can also control transcription factors at the posttranscriptional level (Fig. 2). Typical examples will be described below. In addition, *trans*-acting sRNAs regulate mRNAs by imperfect base-pairings, which signifies that one sRNA can modulate several targets and one target can be controlled by several sRNAs. In *S. aureus*, the annealing region between sRNA and mRNAs are often longer than in *E. coli* and mostly targets the RBS of mRNAs affecting translation. In several cases, a second distinct site of interaction occurs in the coding region. Unlike Gram-negative bacteria, in which Hfq and ProQ proteins participate in sRNA regulation by stabilizing and facilitating their pairings with mRNA targets (see reviews in 34, 35), no RNA chaperones have yet been identified in *S. aureus*. Indeed, there is no identifiable *proQ* homolog, and the role of Hfq is still unclear. Recent work has shown that the rim domain of Hfq has an amino acid composition (low in arginine) incompatible with RNA annealing activity compared to *E. coli* Hfq (36).

AgrA, the response regulator of the *agr* quorum sensing system, activates the transcription of RNAIII. The bifunctional RNAIII codes for δ -hemolysin and regulates the expression of virulence genes at the posttranscriptional level (see below). It interacts with various mRNAs either to activate or to repress translation (Fig. 1g,h). The

RBS of *hla* mRNA encoding α -hemolysin is embedded in a hairpin, which prevents ribosome binding and blocks the start of translation. The 5' UTR of RNAIII possesses complementary sequences to the leader region of *hla*. Interaction between these RNAs enables the recruitment of ribosomes to initiate Hla translation. In its 3' UTR, RNAIII carries conserved UCCC motifs that are used as the seed sequences to bind with the RBS of the target mRNAs coding for protein A, coagulase, Sbi protein, the transcription factor Rot (a repressor of exotoxins), and the endopeptidase LytM (for review see 37). Moreover, AgrA represses the sRNA ArtR, which inhibits translation of the SarA homolog SarT (38), and activates RsaE (see 39) (Fig. 2).

The staphylococcal accessory regulator SarA is synthesized from three distinct promoters (P1, P2, P3) and binds DNA or RNA (1). As a transcription factor, it regulates many genes involved in virulence, autolysis, biofilm formation, stress response, antibiotic resistance, and metabolism, but also two sRNAs, SprC and Srn_9340, located on the same pathogenicity island (40). SprC prevents ribosome binding to the SD of the *atl* mRNA coding for an autolysin. Deletion of the *sprC* gene causes enhanced phagocytosis of *S. aureus* by monocytes and macrophages, and this effect was found to be partly due to the deregulation of *atl* expression (41). SarA represses both *sprC* and *srn_9340* transcription and requires an ATTTTAT sequence in its binding site (40). However, while the SarA level remains relatively constant during bacterial growth, the expression of SprC fluctuates, which suggests that additional factors might control its synthesis and that a mechanism of derepression should coexist under specific conditions. These are the first examples of two sRNAs regulated by the same transcription factor.

In the following, we will describe an sRNA whose transcription is controlled by three independent transcription factors. The sRNA RsaE possesses two consensus sequence motifs UCCC as found in RNAIII, which interact with the RBS of several mRNAs involved in central metabolism to repress their translation (39, 42). RsaE is highly conserved between the *Staphylococcaceae* and *Bacillaceae* families. Not only is the sequence of RsaE conserved between *B. subtilis* and *S. aureus* species, but so are its regulation and functions. Recent studies have shown that its transcription is activated by AgrA (see above), the two-component system SrrAB (staphylococcal respiratory response) in response to NO, and a binding site for the redox sensing repressor Rex has been predicted (39, 43, 44). In *B. subtilis*, the Rex-repression of RsaE (also called RoxS) has been

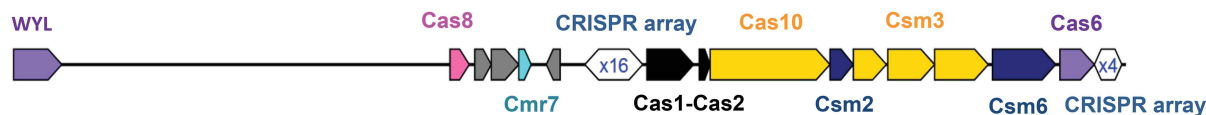
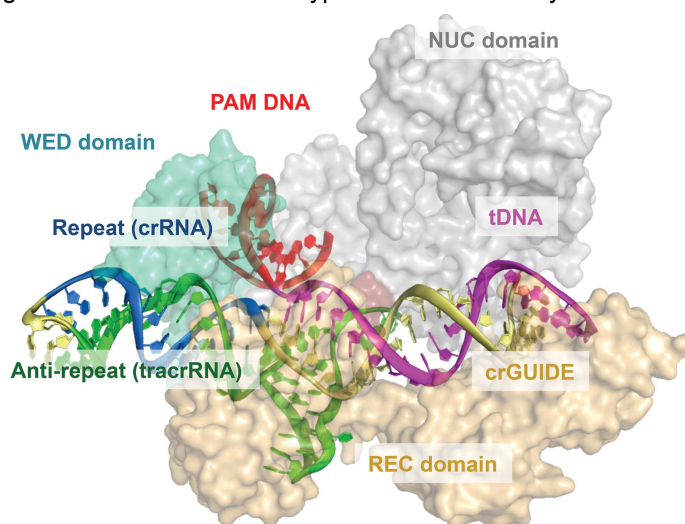
a) *S. aureus* type III-A CRISPR-Cas system (08BA02176 strain)b) *S. aureus* type II-C CRISPR-Cas system (M06/0171 strain)c) *S. aureus* Cas9-sgRNA-DNA structure from type II CRISPR-Cas system

FIGURE 3 (a) Genomic organization of the loci for the type III-A CRISPR system of *S. aureus* strain 08BA02176. Type III is the typical *S. aureus* CRISPR organization. The scheme was obtained using CRISPRone (72), and the genome sequence was deposited in GenBank (accession number 08BA02176; RefSeq accession number GCF_000296595.1). (b) Genomic organization of the loci for the type II-C CRISPR system of *S. aureus* strain M06/0171. The CRISPR-Cas genes were found on an SCCmec inserted into the 3' end of the chromosomally located *orfX* gene. The scheme was obtained using CRISPRone (72), and the SCCmec sequence was deposited in GenBank (GenBank accession number HE980450.1). (c) Cartoon (RNA and DNA) and surface (Cas9) representations of the SaCas9-sgRNA-target DNA complex (pdb file 5AXW) (80). The SaCas9 sgRNA consists of the crRNA guide region (crGUIDE represented in pale yellow) forming a heteroduplex with the target DNA strand (tDNA in magenta) and the repeat/antirepeat helix (blue, the repeat crRNA-derived strand, green, the antirepeat tracrRNA-derived strand). The protospacer adjacent region-containing DNA duplex is red. Cas9 domains are colored as follows: cyan, WED domain; pale orange, REC domain; grey, NUC domain. Molecular graphics images were prepared using PyMol.

proposed to readjust the cellular balance of NAD⁺/NADH on various signals (44).

The alternative sigma B factor (σ^B) together with RNA polymerase guides transcription of genes mainly in the stationary phase of growth and under stress conditions. Its regulon comprises more than 200 genes,

including several virulence factors, transcription factors, and sRNAs (Fig. 2). Among the sRNAs induced by σ^B are SbrA, B, and C activated by KOH (45), and RsaA (39). The stability of RsaA depends on RNase III and the endoribonuclease RNase Y (23, 46). RsaA acts as an acute virulence attenuator in *S. aureus* (see below) by

inhibiting translation of the MgrA transcription factor (Fig. 2) (47), which in turn causes the activation of the synthesis of several surface proteins (48). RsaA possesses two UCCC motifs, which in the case of *mgrA* mRNA, bind to two distant regions, involving an imperfect duplex masking the SD sequence of the mRNA and a loop-loop interaction occurring downstream in the coding region. These two distant binding sites are required for efficient repression and RNase III-dependent degradation of the repressed mRNA (47). Finally, the sRNA Teg49 is transcribed from the σ^B -dependent P3 promoter of *sarA* and is probably processed by RNase III and RNase Y with the help of the helicase CshA (49–51). Transcriptomic analyses revealed that besides genes involved in virulence and autolysis, Teg49 might posttranscriptionally affect the SaeRS and LytRS two-component systems, yet the exact mechanism is not known (51). In addition, another sRNA, Teg48, whose role is not known, is transcribed from the P1 promoter of *sarA* (Fig. 2). Even if the maturation process and the function of Teg48 and Teg49 are not clearly established, the long 5' UTR of *sarA* mRNA encoding a master regulator of virulence in *S. aureus* represents a putative reservoir for novel sRNAs.

crRNA, tracrRNA, AND THE CRISPR-Cas ADAPTIVE IMMUNITY SYSTEMS IN *S. AUREUS*

Phages are the most abundant forms of life on earth and the natural killers of bacteria because in most cases their lytic life cycle ends with the death of the bacterial cell. Outnumbering their microbial hosts, phages impose selective pressure for the diversification of microbial defense systems. These include various innate phage-resistance mechanisms such as restriction/modification enzymes, receptor masking, blocking DNA injection, abortive infection (52, 53), and the adaptive defense mechanism based on clustered, regularly interspaced short palindromic repeat (CRISPR) and CRISPR-associated (*cas*) genes (54, 55). The latter RNA-interference-like mechanism relies on small noncoding RNAs, CRISPR RNA (crRNA), and in some cases trans-activating crRNA (tracrRNA), through which prokaryotic hosts (bacteria and archaea) can acquire heritable resistance to genetic parasites such as phages, but also plasmids and transposons (for reviews see 53, 56). To date, CRISPR-Cas systems have been found in about 50% of bacterial genomes and 95% of archaeal genomes (57, 58). The CRISPR and the *cas* loci are often located next to each other in the genomes, sometimes organized into operons, but a significant num-

ber of genomes have also isolated *cas* loci and/or CRISPRs (59).

Despite the large diversity of the CRISPR-cas systems, they share common features. Briefly, the CRISPR loci are characterized by an array of short and palindromic repetitive sequences interspaced by sequences called “spacers” that are derived from plasmid and viral DNAs (in some cases, also RNAs). During the initial infection with a virus or plasmid, these spacers are first integrated into the CRISPR array in the host genome to provide the host with immunity (acquisition step [60, 61]). During a second infection event they are transcribed and used as guides to inactivate the viral or plasmid genome. This two-step pathway involves a variety of Cas proteins, leading to several major types of CRISPR-cas systems (reviewed in 58, 62). However, the acquisition step involves two highly conserved Cas1 and Cas2 proteins (58, 63). The Cas1-Cas2 integrase is a heterohexameric complex of four Cas1s and two Cas2s which preferentially incorporates foreign DNA at the first CRISPR repeat and participates in the discrimination against self-DNA and in the minimization of off-targeting insertions (64–66). Both proteins have been found in several *S. aureus* strains (57). During the second step, activation of transcription from a promoter located in an AT-rich leader sequence preceding the first CRISPR repeat (67, 68) leads to expression of the whole array into precursor CRISPR transcripts (pre-crRNA). The pre-crRNAs are then processed into mature crRNAs consisting of partial repeat(s) and a single spacer sequence, each complementary to a unique invader sequence (69). Different endonucleases participate in the maturation step, which might vary in different bacteria. Type I and III systems perform the function by a multisubunit Cas protein complex and are characterized by Cas6 processing (69). Type II uses another sRNA, tracrRNA, to direct RNase III-dependent maturation of the pre-crRNA in the presence of Cas9, the hallmark protein of the type II system (70). Cas9 endonuclease remains associated with the dual-tracrRNA:crRNA structure, which during the interference phase, guides the cleavage of site-specific cognate target DNA (71). Type I and II CRISPR-Cas systems target DNA (72, 73), whereas type III systems provide immunity against DNA and RNA (74).

With the completion of the genome sequences of several *S. aureus* strains, it appears that the CRISPR-Cas systems are not highly prevalent, and only a few of them have been experimentally demonstrated. A CRISPR Finder analysis (57) of the 115 sequenced *S. aureus* genomes present in the CRISPRdb (<http://crispr.i2bc>

.paris-saclay.fr) showed that the majority of CRISPR-like loci contain only a few spacers (1 or 2), a few have between 3 and 10 spacers, and only 1 genome (from methicillin-resistant *S. aureus* [MRSA] 08BA 02176, isolated from a patient) has 15 spacers. Using CRISPRone (<http://omics.informatics.indiana.edu/CRISPRone>), several genes belonging to the type III-A CRISPR system (including Csm2, Cas1, Cas2, and Cas6) (Fig. 3a) have been recently predicted (72) in different isolates (MSHR1132, JS395, CIG290, and 21252). Multilocus sequence typing performed on one of these isolates (MSHR1132) has shown that it belongs to a divergent clonal complex, which appears to be closely related to *Staphylococcus epidermidis* and *Staphylococcus lugdunensis* (75), and which has been renamed *Staphylococcus argenteus* (76). It has thus been hypothesized that CRISPR/cas was present in a common ancestor of *S. epidermidis*, *S. lugdunensis*, and *S. aureus* and was later lost in most conventional *S. aureus* strains. It is noteworthy that numerous repeat-spacer-like structures resemble CRISPR elements but lack spacer diversity and have been classified as false-CRISPR (72). *S. aureus* repeat-like elements (GC-rich direct repeats) belong to this class of RNAs, for which the functions remain to be addressed (77, 78).

More recently, a type II-C CRISPR-Cas system was found in an MRSA strain isolated from an Irish patient (Fig. 3b). This CRISPR element is located on a pseudo staphylococcal cassette chromosome *mec* (SCC*mec*) composite island, which was probably horizontally acquired from an *S. epidermidis* strain (79). The peculiarity of this system (SaCas9) is that it contains a shorter version of Cas9 protein. The crystal structure of a complex containing SaCas9, the sgRNA (single-guide RNA; an artificial fusion product of a crRNA and a tracrRNA [71]) and its target DNA provided a model to understand how crRNA and tracrRNA guide Cas9 on the target DNA and prepare it for double-stranded DNA cleavages (80). In this structure, the DNA duplex of the protospacer-adjacent region contains the signals recognized by the Cas9 protospacer-adjacent region-interacting (PI) domain, which discriminates the invader DNA against self-DNA and facilitates unwinding of the target DNA, leading to the formation of a heteroduplex. The beginning of the RNA-DNA heteroduplex ("seed" region) adopts a distorted structure critical for Cas9-catalyzed DNA cleavage. This conformation results from the RNA-DNA helix structure, the interactions with protein residues, and the RNA helix formed by the repeat/antirepeat regions mimicking the tracrRNA-crRNA interactions (Fig. 3c). The discovery of the

smaller Cas9 protein led to recent improvements in genome editing (81). The seminal works of Jennifer Doudna and Emmanuelle Charpentier on the use of CRISPR-Cas systems for genome editing have inspired many studies showing the incredible potency of the system (82–84). Thousands of publications have reported the use of *Streptococcus pyogenes* Cas9 (85) directed by the sequence of a sgRNA for site-specific genome modifications, gene knockouts or replacements, gene expression control, and functional genome screenings in over 40 species. Interestingly, SaCas9 has been successfully used for eukaryotic genome editing since its smaller size makes it easier to be delivered via adeno-associated virus vectors to somatic tissues (81).

The fact that CRISPR-Cas has been mapped on mobile elements in MRSA confirms the importance of horizontal gene transfer from other cocolonizing bacteria in the acquisition of novel functions and in the evolution of *S. aureus* strains. Interestingly, transcription of the CRISPR-Cas genes can be highly regulated and induced upon infection (86–88) by membrane stress (89) and, in Gram-negative bacteria, by quorum sensing signaling (90, 91). Recent work has developed genetic engineering tools to apply the CRISPR/Cas9 system as an antimicrobial strategy against *S. aureus* (92). Clearly, the CRISPR/Cas9 system offers an alternative therapy to conventional antibiotics.

ROLE OF sRNAs IN PHYSIOPATHOLOGY

S. aureus pathogenesis can take different forms depending on the infected tissue and the invading bacterial strain. This is often accompanied by the expression of various virulence factors involved in the colonization and the alteration of the tissue but also by the capacity to escape from the host immune response. Among key regulators of virulence, several sRNAs have been shown to modulate the synthesis of virulence factors in a dynamic manner, and some of them contribute to specific aspects of bacterial virulence in animal models of infection (1).

The most studied RNA in *S. aureus*, *agr*-RNAIII, is the main intracellular effector of the quorum sensing system *agr*. As described above, RNAIII regulates expression of virulence factors known to be associated with infectious diseases. For instance, it represses the synthesis of protein A, which triggers inflammatory signaling pathways and contributes to evasion of the immune response. Conversely, RNAIII induces the synthesis of a battery of toxins, which contribute to the degradation of tissues and subversion of host defenses, such as the pore-

forming toxins, and peptides with proinflammatory and lytic activities. Recent modeling of the quorum sensing system and of its regulators has illustrated the importance of the *agr* system in promoting dissemination of the bacteria from biofilms or dense populations (93, 94). Nevertheless, despite the fact that many clinical isolates from acute infections express RNAIII, its steady state level varies considerably among them (95, 96). In particular, a higher level of RNAIII has been observed in the community-acquired MRSA strains with increased virulence compared to other *S. aureus* lineages (97, 98). Interestingly, a recent study showed that the level of RNAIII is lower in strains isolated from patients with sepsis than in those from commensal carrier patients (99). Perhaps more surprisingly, heterogeneity has been reported in patients in which *agr*-positive and *agr*-negative strains coexisted. This has been proposed as a factor that might modulate the outcome of the infections (100, 101).

S. aureus is also frequently exposed to other microbes during colonization and infection, providing opportunities to acquire mobile genetic elements that contribute to the evolution of the genome. Some of these genomic islands play key roles in pathogenesis through their possession of new virulence factors (pathogenicity islands) or through the synthesis of novel regulators modulating the expression of genes of the core genome. As an example, SprD is an important small regulatory sRNA (142 nucleotides) expressed from a pathogenicity island, which significantly promoted *S. aureus* diseases in a mouse sepsis model of infection (102, 103). SprD interacts through base-pairings with the *sbi* mRNA, which encodes an immune evasion molecule protecting the bacteria from the host immune responses (104, 105). However, the phenotype of the Δ *sprD* mutant strain was not linked to the SprD-dependent regulation of *sbi* since the Δ *sbi* mutant strain behaved like the wild-type strain in the mouse sepsis model (102). Therefore, these data strongly suggest that SprD might regulate the expression of other proteins important for infection.

SSR42 (for small stable RNAs) is an 891-nucleotide-long sRNA whose stability is greatly enhanced in the stationary phase of growth (106, 107). It regulated the expression of approximately 80 mRNAs in 2 genetically different *S. aureus* strain backgrounds. While it increased the expression of capsule Cap5a, SSR42 down-regulated the expression of protein A, α and γ hemolysin, and Panton-Valentin leukocidin (107). Because no direct binding was evidenced between SSR42 and mRNAs encoding virulence determinants, the effect was predicted to be indirect through the modulation of the

expression of a transcriptional regulator. Phenotypically, the deletion of the SSR42 gene affected erythrocyte lysis, resistance to opsonization killing, and pathogenesis in a murine model of skin and soft tissue infections (107). More recently, SSR42 was identified as an important effector of intracellular virulence by screening of a transposon mutant library pool. After internalization in epithelial cells, the Δ SSR42 mutant strain was significantly enriched in the intracellular fraction, most likely due to an attenuated cytotoxicity (108).

In contrast to RNAIII, SprD, and SSR42, which contribute to enhanced virulence of *S. aureus*, other regulatory RNAs behave as attenuators of virulence. This is, for instance, the case of another encoded pathogenicity island sRNA, the so-called SprC (103). Indeed, the virulence of the isogenic strain lacking SprC was significantly and reproducibly enhanced in a mouse systemic model. Furthermore, SprC reduces *S. aureus* susceptibility to phagocytosis by human monocytes and macrophages (41). Another example is the *psm-mec* RNA, which is a bifunctional RNA located in the SCCmec mobile genetic element. It encodes a phenol-soluble modulins (PSMmec) cytolytic toxin and acts as a translational repressor through direct binding with *agrA* mRNA (109). The transcription of all *psm* genes is positively regulated by AgrA. This activation is linked to a specific binding of the phosphorylated form of AgrA to the promoter sequences upstream of the *psm* genes, except that this binding was not yet demonstrated for the *psm-mec* gene (reviewed in 110). The deletion of the *psm-mec* RNA increased the expression of AgrA, which resulted in an increase of toxin and PSM α production and enhanced virulence in mice (109). Interestingly, community-acquired MRSA that does not carry the *psm-mec* gene has been shown to be more virulent than the hospital-associated MRSA that harbors the gene (109, 111). Finally, a mutant strain that did not express the sRNA, RsaA, was attenuated in the severity of acute systemic infection in a mouse model (47). This deletion is linked to the deregulation of MgrA, a master regulator of capsule synthesis and clumping (112). This phenotype in pathogenesis is probably linked to the high sensitivity of the mutant strain to opsonophagocytosis by host polymorphonuclear leukocytes. Because the expression of these three sRNAs is detrimental for bacterial spreading within colonized host organisms, one may suggest that during evolution they have favored commensalism with the host.

To evaluate the impact of sRNA expression in the context of host infection, two studies explored the possible relationships between infection severity and RNA

expression levels. In the first study, the expression levels of five sRNAs (RNAIII, RsaA, RsaE, RsaG, and RsaH) were analyzed in samples from acute cutaneous infection, cystic fibrosis sputum, or nasal colonization. The expression profiles did not correlate with the type of infection, but the authors noticed that the expression of these five RNAs was more homogeneous in the nasal colonization isolates than in those responsible for infection (96). More recently, the expression levels of RNAIII and SprD were measured in 40 strains cultivated from patients with sepsis or septic shock and compared to 21 strains isolated from asymptomatic colonized carriers. It appeared that strains from septic shock had significantly lower levels of RNAIII and to a lesser extent for SprD (99). It is important to note that this analysis was performed on clinical isolates cultured *in vitro* and does not necessarily reflect the expression of these RNAs during infection within the host. In fact, it is very difficult to obtain reliable data from *in vivo* studies to assess the role and importance of RNAs in the establishment or evolution of infection. The great variability of *S. aureus* strains, the difficulty in obtaining highly controlled cohorts of patients, and the reliability of sampling protocols, sample processing, and RNA expression analysis are all obstacles to be overcome. Moreover, biological variables may influence the analysis since the relationships between host immune systems and microbes seem to be particularly individualized and can influence the disease outcome (113, 114). Furthermore, interspecies interactions between bacterial pathogens and the commensal microbiota, as well as limited nutrients, play major roles in promoting or preventing *S. aureus* colonization (113). Interestingly, it was shown that the *agr* system is repressed by high concentrations of hemoglobin in the nasal fluids, leading to the expression of several cell surface proteins and favoring nasal colonization (115). Similar data were obtained when *S. aureus* was cocultivated with the nasal strain of *Corynebacterium striatum* (116). A recent study also demonstrated that the commensal *S. epidermidis* can influence the expression of one ncRNA of *S. aureus* (117).

Clearly, we are just beginning to better appreciate the roles of regulatory sRNAs during colonization and in the pathophysiology of *S. aureus* infections.

ACKNOWLEDGMENTS

This work was supported by the Centre National de la Recherche Scientifique (CNRS) (to P.R.), by the Agence Nationale de la Recherche (ANR, ANR-16-CE11-0007-0-RIBOSTAPH) (to P.R.), and by LABEX: ANR-10-LABX-0036 NETRNA (to P.R.), funding from the state managed by the French National Research Agency as part of the investments for the future program.

REFERENCES

1. Bischoff M, Romby P. 2016. Genetic regulation, p 301–334. In *Staphylococcus: Genetics and Physiology*. Caister Academic Press, Poole, United Kingdom.
2. Tomasini A, François P, Howden BP, Fechter P, Romby P, Caldeleri I. 2014. The importance of regulatory RNAs in *Staphylococcus aureus*. *Infect Genet Evol* 21:616–626 <http://dx.doi.org/10.1016/j.meegid.2013.11.016>.
3. Guillet J, Hallier M, Felden B. 2013. Emerging functions for the *Staphylococcus aureus* RNome. *PLoS Pathog* 9:e1003767 <http://dx.doi.org/10.1371/journal.ppat.1003767>.
4. Caldeleri I, Fechter P, Lioliou E, Romilly C, Chevalier C, Gaspin C, Romby P. 2011. A current overview of regulatory RNAs in *Staphylococcus aureus*, p 51–75. In Marchfelder A, Hess W (ed), *Regulatory RNAs in Prokaryotes*. Springer-Verlag, Vienna, Austria.
5. Quereda JJ, Cossart P. 2017. Regulating bacterial virulence with RNA. *Annu Rev Microbiol* 71:263–280 <http://dx.doi.org/10.1146/annurev-micro-030117-020335>.
6. Mulhbach J, Brouillette E, Allard M, Fortier L-C, Malouin F, Lafontaine DA. 2010. Novel riboswitch ligand analogs as selective inhibitors of guanine-related metabolic pathways. *PLoS Pathog* 6:e1000865 <http://dx.doi.org/10.1371/journal.ppat.1000865>.
7. Lünse CE, Schmidt MS, Wittmann V, Mayer G. 2011. Carba-sugars activate the glmS-riboswitch of *Staphylococcus aureus*. *ACS Chem Biol* 6:675–678 <http://dx.doi.org/10.1021/cb200016d>.
8. Collins JA, Irnov I, Baker S, Winkler WC. 2007. Mechanism of mRNA destabilization by the glmS ribozyme. *Genes Dev* 21:3356–3368 <http://dx.doi.org/10.1101/gad.1605307>.
9. Dar D, Shamir M, Mellin JR, Koutero M, Stern-Ginossar N, Cossart P, Sorek R. 2016. Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. *Science* 352:aad9822 <http://dx.doi.org/10.1126/science.aad9822>.
10. Apostolidi M, Saad NY, Drainas D, Pournaras S, Becker HD, Stathopoulos C. 2015. A *glyS* T-box riboswitch with species-specific structural features responding to both proteinogenic and nonproteinogenic tRNA^{Gly} isoacceptors. *RNA* 21:1790–1806 <http://dx.doi.org/10.1261/rna.052712.115>.
11. Stamatoopoulou V, Apostolidi M, Li S, Lamprinou K, Papakyriakou A, Zhang J, Stathopoulos C. 2017. Direct modulation of T-box riboswitch-controlled transcription by protein synthesis inhibitors. *Nucleic Acids Res* 45:10242–10258 <http://dx.doi.org/10.1093/nar/gkx663>.
12. Kortmann J, Narberhaus F. 2012. Bacterial RNA thermometers: molecular zippers and switches. *Nat Rev Microbiol* 10:255–265 <http://dx.doi.org/10.1038/nrmicro2730>.
13. Johansson J, Mandin P, Renzoni A, Chiaruttini C, Springer M, Cossart P. 2002. An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. *Cell* 110:551–561 [http://dx.doi.org/10.1016/S0092-8674\(02\)00905-4](http://dx.doi.org/10.1016/S0092-8674(02)00905-4).
14. Loh E, Memarpour F, Vaitkevicius K, Kallipolitis BH, Johansson J, Söndén B. 2012. An unstructured 5'-coding region of the *prfA* mRNA is required for efficient translation. *Nucleic Acids Res* 40:1818–1827 <http://dx.doi.org/10.1093/nar/gkr850>.
15. Lasa I, Toledo-Arana A, Gingeras TR. 2012. An effort to make sense of antisense transcription in bacteria. *RNA Biol* 9:1039–1044 <http://dx.doi.org/10.4161/rna.21167>.
16. Ruiz de los Mozos I, Vergara-Irigaray M, Segura V, Villanueva M, Bitarte N, Saramago M, Domingues S, Arraiano CM, Fechter P, Romby P, Valle J, Solano C, Lasa I, Toledo-Arana A. 2013. Base pairing interaction between 5'- and 3'-UTRs controls *icaR* mRNA translation in *Staphylococcus aureus*. *PLoS Genet* 9:e1004001 <http://dx.doi.org/10.1371/journal.pgen.1004001>.
17. Chao Y, Papenfort K, Reinhardt R, Sharma CM, Vogel J. 2012. An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir

- of regulatory small RNAs. *EMBO J* 31:4005–4019 <http://dx.doi.org/10.1038/emboj.2012.229>.
18. Chao Y, Vogel J. 2016. A 3' UTR-derived small RNA provides the regulatory noncoding arm of the inner membrane stress response. *Mol Cell* 61:352–363 <http://dx.doi.org/10.1016/j.molcel.2015.12.023>.
 19. Miyakoshi M, Chao Y, Vogel J. 2015. Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *EMBO J* 34:1478–1492 <http://dx.doi.org/10.15252/emboj.201490546>.
 20. Fozo EM, Hemm MR, Storz G. 2008. Small toxic proteins and the antisense RNAs that repress them. *Microbiol Mol Biol Rev* 72:579–589 <http://dx.doi.org/10.1128/MMBR.00025-08>.
 21. Novick RP, Iordanescu S, Projan SJ, Kornblum J, Edelman I. 1989. pT181 plasmid replication is regulated by a countertranscript-driven transcriptional attenuator. *Cell* 59:395–404 [http://dx.doi.org/10.1016/0092-8674\(89\)90300-0](http://dx.doi.org/10.1016/0092-8674(89)90300-0).
 22. Lasa I, Toledo-Arana A, Dobin A, Villanueva M, de los Mozos IR, Vergara-Irigaray M, Segura V, Fagegaltier D, Penadés JR, Valle J, Solano C, Gingeras TR. 2011. Genome-wide antisense transcription drives mRNA processing in bacteria. *Proc Natl Acad Sci U S A* 108:20172–20177 <http://dx.doi.org/10.1073/pnas.1113521108>.
 23. Lioliou E, Sharma CM, 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 <http://dx.doi.org/10.1371/journal.pgen.1002782>.
 24. Bidnenko V, Nicolas P, Grylak-Mielnicka A, Delumeau O, Auger S, Aucouturier A, Guerin C, Repoilé F, Bardowski J, Aymerich S, Bidnenko E. 2017. Termination factor Rho: from the control of pervasive transcription to cell fate determination in *Bacillus subtilis*. *PLoS Genet* 13:e1006909 <http://dx.doi.org/10.1371/journal.pgen.1006909>.
 25. Mäder U, Nicolas P, Depke M, Pané-Farré J, Debarbouille M, van der Kooi-Pol MM, Guérin C, Dérozier S, Hiron A, Jarmer H, Leduc A, Michalik S, Reilman E, Schaffer M, Schmidt F, Bessières P, Noirot P, Hecker M, Msadek T, Völker U, van Dijk JM. 2016. *Staphylococcus aureus* transcriptome architecture: from laboratory to infection-mimicking conditions. *PLoS Genet* 12:e1005962 <http://dx.doi.org/10.1371/journal.pgen.1005962>.
 26. Goeders N, Chai R, Chen B, Day A, Salmond GP. 2016. Structure, evolution, and functions of bacterial type III toxin-antitoxin systems. *Toxins (Basel)* 8:282 <http://dx.doi.org/10.3390/toxins8100282>.
 27. Coray DS, Wheeler NE, Heinemann JA, Gardner PP. 2017. Why so narrow: distribution of anti-sense regulated, type I toxin-antitoxin systems compared with type II and type III systems. *RNA Biol* 14:275–280 <http://dx.doi.org/10.1080/15476286.2016.1272747>.
 28. Brielle R, Pinel-Marie M-L, Felden B. 2016. Linking bacterial type I toxins with their actions. *Curr Opin Microbiol* 30:114–121 <http://dx.doi.org/10.1016/j.mib.2016.01.009>.
 29. Pinel-Marie M-L, Brielle R, Felden B. 2014. Dual toxic-peptide-coding *Staphylococcus aureus* RNA under antisense regulation targets host cells and bacterial rivals unequally. *Cell Reports* 7:424–435 <http://dx.doi.org/10.1016/j.celrep.2014.03.012>.
 30. Sayed N, Jousselin A, Felden B. 2011. A cis-antisense RNA acts in trans in *Staphylococcus aureus* to control translation of a human cytolytic peptide. *Nat Struct Mol Biol* 19:105–112 <http://dx.doi.org/10.1038/nsmb.2193>.
 31. Bronsard J, Pascreau G, Sassi M, Mauro T, Augagneur Y, Felden B. 2017. sRNA and cis-antisense sRNA identification in *Staphylococcus aureus* highlights an unusual sRNA gene cluster with one encoding a secreted peptide. *Sci Rep* 7:4565 <http://dx.doi.org/10.1038/s41598-017-04786-3>.
 32. Neubauer C, Gao Y-G, Andersen KR, Dunham CM, Kelley AC, Hentschel J, Gerdes K, Ramakrishnan V, Brodersen DE. 2009. The structural basis for mRNA recognition and cleavage by the ribosome-dependent endonuclease RelE. *Cell* 139:1084–1095 <http://dx.doi.org/10.1016/j.cell.2009.11.015>.
 33. Berghoff BA, Wagner EGH. 2017. RNA-based regulation in type I toxin-antitoxin systems and its implication for bacterial persistence. *Curr Genet* 63:1011–1016 <http://dx.doi.org/10.1007/s00294-017-0710-y>.
 34. Updegrave TB, Zhang A, Storz G. 2016. Hfq: the flexible RNA matchmaker. *Curr Opin Microbiol* 30:133–138 <http://dx.doi.org/10.1016/j.mib.2016.02.003>.
 35. Attaiech L, Glover JNM, Charpentier X. 2017. RNA chaperones step out of Hfq's shadow. *Trends Microbiol* 25:247–249 <http://dx.doi.org/10.1016/j.tim.2017.01.006>.
 36. Zheng A, Panja S, Woodson SA. 2016. Arginine patch predicts the RNA annealing activity of Hfq from Gram-negative and Gram-positive bacteria. *J Mol Biol* 428:2259–2264 <http://dx.doi.org/10.1016/j.jmb.2016.03.027>.
 37. Bronesky D, Wu Z, Marzi S, Walter P, Geissmann T, Moreau K, Vandenesch F, Caldelari I, Romby P. 2016. *Staphylococcus aureus* RNAIII and its regulon link quorum sensing, stress responses, metabolic adaptation, and regulation of virulence gene expression. *Annu Rev Microbiol* 70:299–316 <http://dx.doi.org/10.1146/annurev-micro-102215-095708>.
 38. Xue T, Zhang X, Sun H, Sun B. 2014. ArtR, a novel sRNA of *Staphylococcus aureus*, regulates α -toxin expression by targeting the 5' UTR of sarT mRNA. *Med Microbiol Immunol (Berl)* 203:1–12 <http://dx.doi.org/10.1007/s00430-013-0307-0>.
 39. 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 <http://dx.doi.org/10.1093/nar/gkp668>.
 40. Mauro T, Rouillon A, Felden B. 2016. Insights into the regulation of small RNA expression: SarA represses the expression of two sRNAs in *Staphylococcus aureus*. *Nucleic Acids Res* 44:10186–10200.
 41. Le Pabic H, Germain-Amiot N, Bordeau V, Felden B. 2015. A bacterial regulatory RNA attenuates virulence, spread and human host cell phagocytosis. *Nucleic Acids Res* 43:9232–9248 <http://dx.doi.org/10.1093/nar/gkv783>.
 42. Bohn C, Rigoulay C, Chabelskaya S, Sharma CM, 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 <http://dx.doi.org/10.1093/nar/gkq462>.
 43. Durand S, Tomasini A, Braun F, Condon C, Romby P. 2015. sRNA and mRNA turnover in Gram-positive bacteria. *FEMS Microbiol Rev* 39:316–330 <http://dx.doi.org/10.1093/femsrev/fuv007>.
 44. Durand S, Braun F, Helfer A-C, Romby P, Condon C. 2017. sRNA-mediated activation of gene expression by inhibition of 5'-3' exonucleolytic mRNA degradation. *eLife* 6:e23602 <http://dx.doi.org/10.7554/eLife.23602>.
 45. Nielsen JS, Christiansen MHG, Bonde M, Gottschalk S, Frees D, Thomsen LE, Kallipolitis BH. 2011. Searching for small σ B-regulated genes in *Staphylococcus aureus*. *Arch Microbiol* 193:23–34 <http://dx.doi.org/10.1007/s00203-010-0641-1>.
 46. Marincola G, Schäfer T, Behler J, Bernhardt J, Ohlsen K, Goerke C, Wolz C. 2012. RNase Y of *Staphylococcus aureus* and its role in the activation of virulence genes. *Mol Microbiol* 85:817–832 <http://dx.doi.org/10.1111/j.1365-2958.2012.08144.x>.
 47. Romilly C, Lays C, Tomasini A, Caldelari I, Benito Y, Hammann P, Geissmann T, Boisset S, Romby P, Vandenesch F. 2014. A non-coding RNA promotes bacterial persistence and decreases virulence by regulating a regulator in *Staphylococcus aureus*. *PLoS Pathog* 10:e1003979 <http://dx.doi.org/10.1371/journal.ppat.1003979>.
 48. Tomasini A, Moreau K, Chicher J, Geissmann T, Vandenesch F, Romby P, Marzi S, Caldelari I. 2017. The RNA targetome of *Staphylococcus aureus* non-coding RNA RsaA: impact on cell surface properties

- and defense mechanisms. *Nucleic Acids Res* 45:6746–6760 <http://dx.doi.org/10.1093/nar/gkx219>.
49. 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 <http://dx.doi.org/10.1371/journal.pone.0010725>.
50. Kim S, Reyes D, Beaume M, Francois P, Cheung A. 2014. Contribution of teg49 small RNA in the 5' upstream transcriptional region of *sarA* to virulence in *Staphylococcus aureus*. *Infect Immun* 82:4369–4379 <http://dx.doi.org/10.1128/IAI.02002-14>.
51. Manna AC, Kim S, Cengher L, Corvaglia A, Leo S, Francois P, Cheung AL. 2018. Small RNA teg49 is derived from a *sarA* transcript and regulates virulence genes independent of *SarA* in *Staphylococcus aureus*. *Infect Immun* 86:e00635-17.
52. Labrie SJ, Samson JE, Moineau S. 2010. Bacteriophage resistance mechanisms. *Nat Rev Microbiol* 8:317–327 <http://dx.doi.org/10.1038/nrmicro2315>.
53. Westra ER, Swarts DC, Staals RHJ, Jore MM, Brouns SJJ, van der Oost J. 2012. The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity. *Annu Rev Genet* 46:311–339 <http://dx.doi.org/10.1146/annurev-genet-110711-155447>.
54. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315:1709–1712 <http://dx.doi.org/10.1126/science.1138140>.
55. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. 2006. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct* 1:7 <http://dx.doi.org/10.1186/1745-6150-1-7>.
56. Wiedenheft B, Sternberg SH, Doudna JA. 2012. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482:331–338 <http://dx.doi.org/10.1038/nature10886>.
57. Grissa I, Vergnaud G, Pourcel C. 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* 8:172 <http://dx.doi.org/10.1186/1471-2105-8-172>.
58. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, Barrangou R, Brouns SJJ, Charpentier E, Haft DH, Horvath P, Moineau S, Mojica FJM, Terns RM, Terns MP, White MF, Yakunin AF, Garrett RA, van der Oost J, Backofen R, Koonin EV. 2015. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* 13:722–736 <http://dx.doi.org/10.1038/nrmicro3569>.
59. Haft DH, Selengut J, Mongodin EF, Nelson KE. 2005. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLOS Comput Biol* 1:e60 <http://dx.doi.org/10.1371/journal.pcbi.0010060>.
60. Horvath P, Barrangou R. 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327:167–170 <http://dx.doi.org/10.1126/science.1179555>.
61. Nuñez JK, Lee ASY, Engelman A, Doudna JA. 2015. Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* 519:193–198 <http://dx.doi.org/10.1038/nature14237>.
62. Marraffini LA. 2015. CRISPR-Cas immunity in prokaryotes. *Nature* 526:55–61 <http://dx.doi.org/10.1038/nature15386>.
63. Yosef I, Goren MG, Qimron U. 2012. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res* 40:5569–5576 <http://dx.doi.org/10.1093/nar/gks216>.
64. Levy A, Goren MG, Yosef I, Auster O, Manor M, Amitai G, Edgar R, Qimron U, Sorek R. 2015. CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature* 520:505–510 <http://dx.doi.org/10.1038/nature14302>.
65. Modell JW, Jiang W, Marraffini LA. 2017. CRISPR-Cas systems exploit viral DNA injection to establish and maintain adaptive immunity. *Nature* 544:101–104 <http://dx.doi.org/10.1038/nature21719>.
66. Wright AV, Liu J-J, Knott GJ, Doxzen KW, Nogales E, Doudna JA. 2017. Structures of the CRISPR genome integration complex. *Science* 357:1113–1118 <http://dx.doi.org/10.1126/science.aao0679>.
67. Jansen R, Embden JD, Gaastra W, Schouls LM. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* 43:1565–1575 <http://dx.doi.org/10.1046/j.1365-2958.2002.02839.x>.
68. Pougach K, Semenova E, Bogdanova E, Datsenko KA, Djordjevic M, Wanner BL, Severinov K. 2010. Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Mol Microbiol* 77:1367–1379 <http://dx.doi.org/10.1111/j.1365-2958.2010.07265.x>.
69. Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Doudna JA. 2010. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* 329:1355–1358 <http://dx.doi.org/10.1126/science.1192272>.
70. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E. 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471:602–607 <http://dx.doi.org/10.1038/nature09886>.
71. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816–821 <http://dx.doi.org/10.1126/science.1225829>.
72. Zhang Q, Ye Y. 2017. Not all predicted CRISPR-Cas systems are equal: isolated cas genes and classes of CRISPR like elements. *BMC Bioinformatics* 18:92 <http://dx.doi.org/10.1186/s12859-017-1512-4>.
73. Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuys RJH, Snijders APL, Dickman MJ, Makarova KS, Koonin EV, van der Oost J. 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321:960–964 <http://dx.doi.org/10.1126/science.1159689>.
74. Strutt SC, Torrez RM, Kaya E, Negrete OA, Doudna JA. 2018. RNA-dependent RNA targeting by CRISPR-Cas9. *eLife* 7:e32724 <http://dx.doi.org/10.7554/eLife.32724>.
75. Holt DC, Holden MTG, Tong SYC, Castillo-Ramirez S, Clarke L, Quail MA, Currie BJ, Parkhill J, Bentley SD, Feil EJ, Giffard PM. 2011. A very early-branching *Staphylococcus aureus* lineage lacking the carotenoid pigment staphyloxanthin. *Genome Biol Evol* 3:881–895 <http://dx.doi.org/10.1093/gbe/evr078>.
76. Tong SYC, Schaumburg F, Ellington MJ, Corander J, Pichon B, Leendertz F, Bentley SD, Parkhill J, Holt DC, Peters G, Giffard PM. 2015. Novel staphylococcal species that form part of a *Staphylococcus aureus*-related complex: the non-pigmented *Staphylococcus argenteus* sp. nov. and the non-human primate-associated *Staphylococcus schweitzeri* sp. nov. *Int J Syst Evol Microbiol* 65:15–22 <http://dx.doi.org/10.1099/ijis.0.062752-0>.
77. Cramton SE, Schnell NF, Götz F, Brückner R. 2000. Identification of a new repetitive element in *Staphylococcus aureus*. *Infect Immun* 68:2344–2348 <http://dx.doi.org/10.1128/IAI.68.4.2344-2348.2000>.
78. Purves J, Blades M, Arafat Y, Malik SA, Bayliss CD, Morrissey JA. 2012. Variation in the genomic locations and sequence conservation of STAR elements among staphylococcal species provides insight into DNA repeat evolution. *BMC Genomics* 13:515 <http://dx.doi.org/10.1186/1471-2164-13-515>.
79. Kinnevey PM, Shore AC, Brennan GI, Sullivan DJ, Ehrlich R, Monecke S, Slickers P, Coleman DC. 2013. Emergence of sequence type 779 methicillin-resistant *Staphylococcus aureus* harboring a novel pseudo staphylococcal cassette chromosome *mec* (SCCmec)-SCC-SCC_{CRISPR} composite element in Irish hospitals. *Antimicrob Agents Chemother* 57:524–531 <http://dx.doi.org/10.1128/AAC.01689-12>.
80. Nishimasu H, Cong L, Yan WX, Ran FA, Zetsche B, Li Y, Kurabayashi A, Ishitani R, Zhang F, Nureki O. 2015. Crystal structure of

- Staphylococcus aureus* Cas9. *Cell* 162:1113–1126 <http://dx.doi.org/10.1016/j.cell.2015.08.007>.
81. Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS, Koonin EV, Sharp PA, Zhang F. 2015. *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature* 520:186–191 <http://dx.doi.org/10.1038/nature14299>.
82. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823 <http://dx.doi.org/10.1126/science.1231143>.
83. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. 2013. RNA-programmed genome editing in human cells. *eLife* 2:e00471 <http://dx.doi.org/10.7554/eLife.00471>.
84. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013. RNA-guided human genome engineering via Cas9. *Science* 339:823–826 <http://dx.doi.org/10.1126/science.1232033>.
85. Sakuma T, Nishikawa A, Kume S, Chayama K, Yamamoto T. 2015. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci Rep* 4:5400.
86. Agari Y, Sakamoto K, Tamakoshi M, Oshima T, Kuramitsu S, Shinkai A. 2010. Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. *J Mol Biol* 395:270–281 <http://dx.doi.org/10.1016/j.jmb.2009.10.057>.
87. Quax TEF, Voet M, Sismeiro O, Dillies M-A, Jagla B, Coppée J-Y, Sezonov G, Forterre P, van der Oost J, Lavigne R, Prangishvili D. 2013. Massive activation of archaeal defense genes during viral infection. *J Virol* 87:8419–8428 <http://dx.doi.org/10.1128/JVI.01020-13>.
88. Young JC, Dill BD, Pan C, Hettich RL, Banfield JF, Shah M, Fremaux C, Horvath P, Barrangou R, Verberkmoes NC. 2012. Phage-induced expression of CRISPR-associated proteins is revealed by shotgun proteomics in *Streptococcus thermophilus*. *PLoS One* 7:e38077 <http://dx.doi.org/10.1371/journal.pone.0038077>.
89. Perez-Rodriguez R, Haitjema C, Huang Q, Nam KH, Bernardis S, Ke A, DeLisa MP. 2011. Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in *Escherichia coli*. *Mol Microbiol* 79:584–599 <http://dx.doi.org/10.1111/j.1365-2958.2010.07482.x>.
90. Patterson AG, Jackson SA, Taylor C, Evans GB, Salmond GPC, Przybiski R, Staats RHJ, Fineran PC. 2016. Quorum sensing controls adaptive immunity through the regulation of multiple CRISPR-Cas systems. *Mol Cell* 64:1102–1108 <http://dx.doi.org/10.1016/j.molcel.2016.11.012>.
91. Høyland-Kroghsbo NM, Paczkowski J, Mukherjee S, Broniewski J, Westra E, Bondy-Denomy J, Bassler BL. 2017. Quorum sensing controls the *Pseudomonas aeruginosa* CRISPR-Cas adaptive immune system. *Proc Natl Acad Sci U S A* 114:131–135 <http://dx.doi.org/10.1073/pnas.1617415113>.
92. Kang YK, Kwon K, Ryu JS, Lee HN, Park C, Chung HJ. 2017. Nonviral genome editing based on a polymer-derivatized CRISPR nano-complex for targeting bacterial pathogens and antibiotic resistance. *Bioconjug Chem* 28:957–967 <http://dx.doi.org/10.1021/acs.bioconjug.6b00676>.
93. Nitzan M, Fechter P, Peer A, Altuvia Y, Bronesky D, Vandenesch F, Romby P, Biham O, Margalit H. 2015. A defense-offense multi-layered regulatory switch in a pathogenic bacterium. *Nucleic Acids Res* 43:1357–1369 <http://dx.doi.org/10.1093/nar/gkv001>.
94. Audretsch C, Lopez D, Srivastava M, Wolz C, Dandekar T. 2013. A semi-quantitative model of quorum-sensing in *Staphylococcus aureus*, approved by microarray meta-analyses and tested by mutation studies. *Mol Biosyst* 9:2665–2680 <http://dx.doi.org/10.1039/c3mb70117d>.
95. 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 <http://dx.doi.org/10.1016/j.ijmm.2009.07.003>.
96. Song J, Lays C, Vandenesch F, Benito Y, Bes M, Chu Y, Lina G, Romby P, Geissmann T, Boisset S. 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 <http://dx.doi.org/10.1371/journal.pone.0037294>.
97. Montgomery CP, Boyle-Vavra S, Adem PV, Lee JC, Husain AN, Clasen J, Daum RS. 2008. Comparison of virulence in community-associated methicillin-resistant *Staphylococcus aureus* pulsotypes USA300 and USA400 in a rat model of pneumonia. *J Infect Dis* 198:561–570 <http://dx.doi.org/10.1086/590157>.
98. Montgomery CP, Boyle-Vavra S, Daum RS. 2010. Importance of the global regulators Agr and SaeRS in the pathogenesis of CA-MRSA USA300 infection. *PLoS One* 5:e15177 <http://dx.doi.org/10.1371/journal.pone.0015177>.
99. Bordeau V, Cady A, Revest M, Rostan O, Sassi M, Tattevin P, Donnio P-Y, Felden B. 2016. *Staphylococcus aureus* regulatory RNAs as potential biomarkers for bloodstream infections. *Emerg Infect Dis* 22:1570–1578 <http://dx.doi.org/10.3201/eid2209.151801>.
100. Painter KL, Krishna A, Wigneshweraraj S, Edwards AM. 2014. What role does the quorum-sensing accessory gene regulator system play during *Staphylococcus aureus* bacteremia? *Trends Microbiol* 22:676–685 <http://dx.doi.org/10.1016/j.tim.2014.09.002>.
101. Pollitt EJG, West SA, Crusz SA, Burton-Chellew MN, Diggle SP. 2014. Cooperation, quorum sensing, and evolution of virulence in *Staphylococcus aureus*. *Infect Immun* 82:1045–1051 <http://dx.doi.org/10.1128/IAI.01216-13>.
102. Chabelskaya S, Gaillot 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 <http://dx.doi.org/10.1371/journal.ppat.1000927>.
103. 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 U S A* 102:14249–14254 <http://dx.doi.org/10.1073/pnas.0503838102>.
104. Haupt K, Reuter M, van den Elsen J, Burman J, Hälbig S, Richter J, Skerka C, Zipfel PF. 2008. The *Staphylococcus aureus* protein Sbi acts as a complement inhibitor and forms a tripartite complex with host complement factor H and C3b. *PLoS Pathog* 4:e1000250 <http://dx.doi.org/10.1371/journal.ppat.1000250>.
105. Zhang L, Jacobsson K, Vasi J, Lindberg M, Frykberg L. 1998. A second IgG-binding protein in *Staphylococcus aureus*. *Microbiology* 144:985–991 <http://dx.doi.org/10.1099/00221287-144-4-985>.
106. Anderson KL, Roberts C, Disz T, Vonstein V, Hwang K, Overbeek R, Olson PD, Projan SJ, Dunman PM. 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 <http://dx.doi.org/10.1128/JB.00609-06>.
107. Morrison JM, Miller EW, Benson MA, Alonzo F III, Yoong P, Torres VJ, Hinrichs SH, Dunman PM. 2012. 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 <http://dx.doi.org/10.1128/JB.06708-11>.
108. Das S, Lindemann C, Young BC, Muller J, Österreich B, Ternette N, Winkler A-C, Paprotka K, Reinhardt R, Förstner KU, Allen E, Flaxman A, Yamaguchi Y, Rollier CS, van Diemen P, Blättner S, Remmele CW, Selle M, Dittrich M, Müller T, Vogel J, Ohlsen K, Crook DW, Massey R, Wilson DJ, Rudel T, Wyllie DH, Fraunholz MJ. 2016. Natural mutations in a *Staphylococcus aureus* virulence regulator attenuate cytotoxicity but permit bacteremia and abscess formation. *Proc Natl Acad Sci U S A* 113:E3101–E3110 <http://dx.doi.org/10.1073/pnas.1520255113>.
109. 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 <http://dx.doi.org/10.1371/journal.ppat.1003269>.
110. Qin L, McCausland JW, Cheung GYC, Otto M. 2016. PSM-Mec-A virulence determinant that connects transcriptional regulation, virulence, and antibiotic resistance in staphylococci. *Front Microbiol* 7:1293 <http://dx.doi.org/10.3389/fmicb.2016.01293>.
111. 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 <http://dx.doi.org/10.1371/journal.ppat.1001267>.
112. Crosby HA, Schlievert PM, Merriman JA, King JM, Salgado-Pabón W, Horswill AR. 2016. The *Staphylococcus aureus* global regulator MgrA modulates clumping and virulence by controlling surface protein expression. *PLoS Pathog* 12:e1005604 <http://dx.doi.org/10.1371/journal.ppat.1005604>.
113. Krismer B, Weidenmaier C, Zipperer A, Peschel A. 2017. The commensal lifestyle of *Staphylococcus aureus* and its interactions with the nasal microbiota. *Nat Rev Microbiol* 15:675–687 <http://dx.doi.org/10.1038/nrmicro.2017.104>.
114. Brown AF, Leech JM, Rogers TR, McLoughlin RM. 2014. *Staphylococcus aureus* colonization: modulation of host immune response and impact on human vaccine design. *Front Immunol* 4:507 <http://dx.doi.org/10.3389/fimmu.2013.00507>.
115. Pynnonen M, Stephenson RE, Schwartz K, Hernandez M, Boles BR. 2011. Hemoglobin promotes *Staphylococcus aureus* nasal colonization. *PLoS Pathog* 7:e1002104 <http://dx.doi.org/10.1371/journal.ppat.1002104>.
116. Ramsey MM, Freire MO, Gabriliska RA, Rumbaugh KP, Lemon KP. 2016. *Staphylococcus aureus* shifts toward commensalism in response to *Corynebacterium* species. *Front Microbiol* 7:1230 <http://dx.doi.org/10.3389/fmicb.2016.01230>.
117. Hermansen GMM, Sazinas P, Kofod D, Millard A, Andersen PS, Jelsbak L. 2018. Transcriptomic profiling of interacting nasal staphylococci species reveals global changes in gene and non-coding RNA expression. *FEMS Microbiol Lett* 365: <http://dx.doi.org/10.1093/femsle/fny004>.

4.3.1. sARN et carences nutritives

4.3.1.1. RsaC

Au laboratoire, les fonctions régulatrices de l'ARN non codant RsaC ont été en partie résolues (Lalaouna et al., 2019). RsaC est un ARN long de 850 à 1116 nucléotides en fonction des souches de *S. aureus*, à l'image de l'ARNIII montrant que la taille des sARN varie énormément. De manière intéressante, RsaC est co-transcrit avec l'opéron *mntABC*, le transporteur ABC majeur du manganèse (un type de transporteur qui utilise l'ATP comme énergie pour le transport de manganèse), en réponse à une carence en manganèse (Mn). La RNaseIII serait responsable de la maturation de cet sARN issue de la région 3'UTR de *mntC*. Les ARNm cibles régulés par RsaC sont impliqués dans la réponse au stress oxydatif ainsi que l'homéostasie des métaux. En absence de manganèse, RsaC inhibe la synthèse de l'enzyme SodA, une superoxide dismutase manganèse-dépendante, nécessaire dans la réponse aux stress oxydatifs générés par la respiration aérobie ou par les défenses immunitaires de l'hôte. L'inhibition de SodA par RsaC permet de rediriger le peu de manganèse présent vers d'autres protéines essentielles l'ayant pour cofacteur. Indirectement, RsaC régule positivement SodM (en réprimant *sarA*, cf III.1.1.3), une autre superoxide dismutase utilisant le fer comme cofacteur, permettant ainsi de maintenir une voie de détoxification des ROS (« *reactive oxygen species* ») en absence de manganèse (Lalaouna et al., 2019) (Figure 11).

4.3.1.2. RsaD

RsaD est un des petits sARN qui répond à divers stress. Sa transcription est régulée à quatre niveaux. Premièrement, celle-ci est sous la dépendance du facteur sigma B impliqué dans la réponse au stress chez *S. aureus*, et est plus fortement exprimée suite à un choc froid de température (Geissmann et al., 2009). Deuxièmement, la synthèse de RsaD est négativement régulée par CodY, un facteur de transcription responsable de la répression d'un grand nombre de gènes impliqués dans l'import et la biosynthèse d'acides aminés en fonction de la concentration intracellulaire en acides aminés ramifiés, tels que isoleucine, leucine, et valine et en GTP (Ratnayake-Lecamwasam et al., 2001; Shivers & Sonenshein, 2004). CodY régule également un grand nombre de gènes impliqués dans la virulence de *S. aureus* (Majerczyk et al., 2010). Troisièmement, RsaD est régulé par le système à deux composants SrrAB dans des conditions d'expositions au NO (Bronesky et al., 2019). Enfin, l'expression de RsaD est aussi augmentée dans un mutant du système Agr, qui répond à la densité cellulaire (Geissmann et al., 2009). Récemment, il a été montré que RsaD déstabilise l'ARNm *alsS* codant pour l'acétolactate synthase, et inhibe sa traduction. L'opéron *alsSD* permet la production d'acétoine

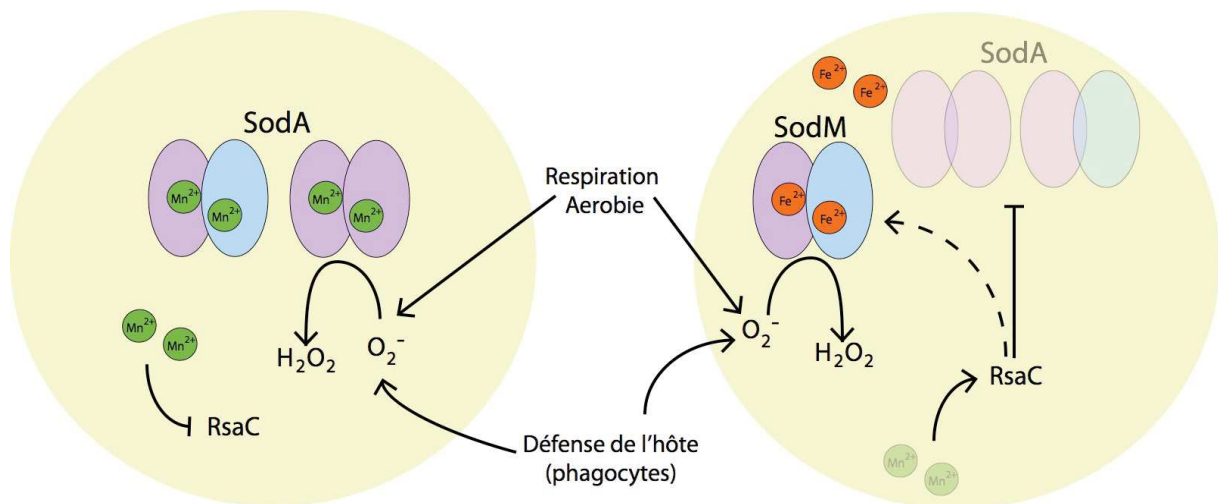


Figure 11 : L'ARN régulateur RsaC module la réponse au stress oxydatif en conditions de carence en manganèse (Mn) : À gauche la superoxide dismutase manganèse dépendante, SodA, permet la protection contre les stress oxydatifs résultant de la respiration aérobie ou du système immunitaire de son hôte. À droite le sARN RsaC est activé en absence de Mn, et inhibe la synthèse de la protéine SodA inactive dans ces conditions. En parallèle afin de maintenir une réponse aux stress oxydatif RsaC active indirectement SodM, une superoxide dismutase fer dépendante. Inspiré de Lalaouna et al., 2019.

à partir du pyruvate et est impliqué dans la résistance au stress acide. Ainsi l'inhibition de *alsS* par RsaD a pour effet de rediriger le flux de carbone vers la production d'acétate (Augagneur et al., 2020).

4.3.2. RsaE : implication dans la réponse au stress oxydatif, le métabolisme central et la formation de biofilm

De nouvelles cibles du sARN RsaE ont été récemment décrites. Ce sARN est l'un des seuls à être très conservé chez les firmicutes. Cette conservation inhabituelle entre différents genres suggère une pression de sélection positive pour préserver la structure et la séquence de RsaE, soulignant l'importance de ce régulateur dans la physiologie des bactéries à Gram positif (Marincola et al., 2019). RsaE a été largement étudié chez *Bacillus subtilis* où il porte le nom de RoxS. Il est impliqué dans le maintien de l'homéostasie rédox en régulant des gènes impliqués dans le métabolisme central comme le cycle de Krebs, le métabolisme du folate, et le métabolisme des acides aminés (Figure 12). RoxS est également impliqué dans la réponse au stress oxydatif, ce qui a entraîné son changement d'appellation de RsaE à RoxS (pour « *related to oxydative stress* ») (Durand et al., 2015a, 2015b). Chez *S. aureus* RsaE régule comme chez *B. subtilis* des gènes impliqués dans le métabolisme du folate et dans le cycle de Krebs (Bohn et al., 2010; Geissmann et al., 2009, figure 12). L'étude de Rochat et al., (2018), a montré que RsaE est également un régulateur du catabolisme des acides aminées (Figure 12). En inhibant différentes enzymes utilisant le NAD⁺, RsaE réduit le catabolisme des acides aminés et l'approvisionnement du cycle de Krebs. Cette fonction donne à RsaE la capacité de réguler le rapport NAD⁺/NADH, comme RoxS chez *B. subtilis* (Rochat et al., 2018). Aussi bien chez *S. aureus* que chez *B. subtilis*, l'expression de RsaE/RoxS est activée par le TCS SrrAB chez *S. aureus* et son homologue ResDE chez *B. subtilis*. SrrAB/ResDE répondent à l'hypoxie et au stress NO (Durand et al., 2015a, 2015b). De plus, le facteur de transcription Rex (*cf* 3.1.2.2.1) réprime l'expression de RoxS en conditions aérobies, et la présence de malate ou de glucose lève la répression de RoxS par Rex (Durand et al., 2017). Chez *Staphylococcus epidermidis*, RsaE est également impliqué dans la régulation du cycle de Krebs. Récemment il a été démontré que chez *S. epidermidis*, RsaE est exprimé de manière hétérogène au sein d'une population de bactéries formant un biofilm. De par son expression hétérogène, RsaE pourrait faciliter une diversité au sein du biofilm puisque les cellules qui expriment RsaE influenceraient la composition de la matrice du biofilm. En effet, RsaE permet de faciliter le relargage d'ADN dans la matrice, en inhibant le répresseur *lrgAB*. En conséquence il activerait indirectement CidABC pour induire l'autolyse. RsaE inhibe également la synthèse du répresseur de biofilm

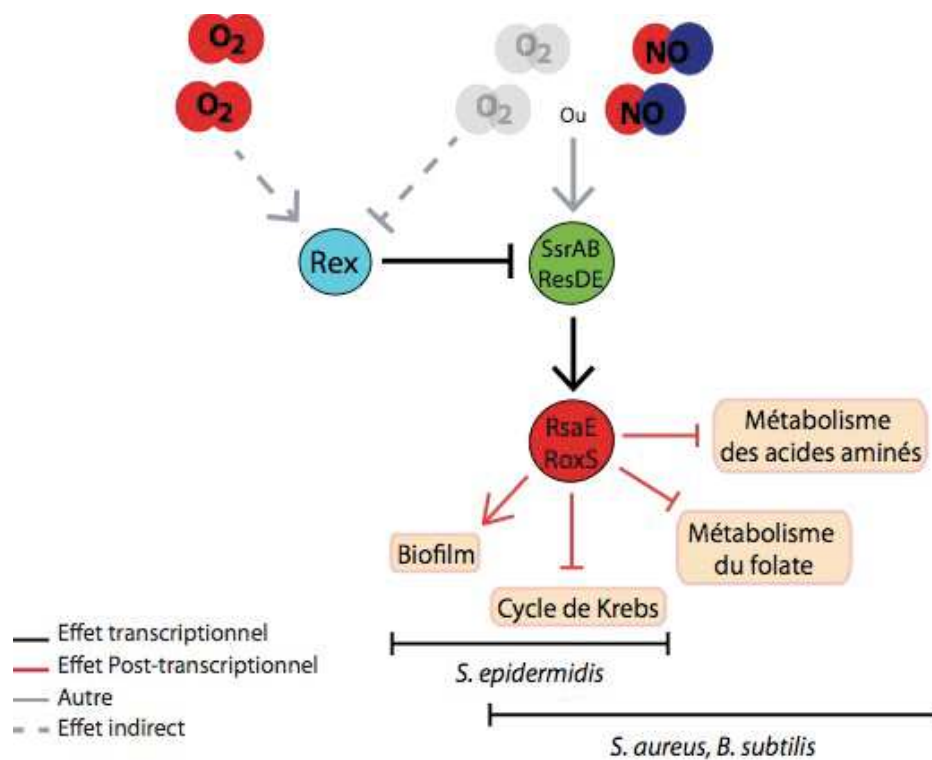


Figure 12 : Voies régulées par RsaE/RoxS au sein des firmicutes : Représentation schématique de la régulation de RsaE (*S. aureus*, *S. epidermidis*) ou son homologue RoxS (*B. subtilis*) en fonction de la présence de NO et d'oxygène par le TCS SsrAB (*S. aureus*, *S. epidermidis*) ou ResDE (*B. subtilis*). Les voies dans lesquelles RsaE/RoxS exerce une régulation sont représentées dans un carré orange.

IcaR ce qui a pour effet de stimuler indirectement la synthèse de PIA-PNAG par l'opéron *icaADBC* (Marincola et al., 2019; Schoenfelder et al., 2019, figure 12).

4.3.3. sARN et virulence

L'implication du sARN Teg41 dans la virulence de *S. aureus* a été étudiée. Le sARN Teg41 est transcrit de manière divergente du locus qui code pour les exotoxines, alpha *phenol-soluble modulin* (α PSM). L'interaction entre la région 3' de Teg41 et l'ARNm α PSM, induirait une stabilisation de ce transcrit par un mécanisme non élucidé, ce qui influencerait positivement la production des α PSM. Une souche délétée de la région 3' de Teg41 présente une diminution de son activité hémolytique et une virulence atténuée dans un modèle d'infection murin (Zapf et al., 2019).

4.3.4. sARN et antibiotiques

Enfin, un set de 10 sARN a été identifié comme répondant à l'exposition aux antibiotiques. Cependant la délétion individuelle de ces ARN n'a pas montré de phénotype, que ce soit au niveau de la CMI (concentration minimale inhibitrice), ou de la croissance en présence de linézolide (Gao et al., 2020; Howden et al., 2013). Ces résultats suggèrent qu'il existe une redondance dans la fonctionnalité des sARN, ou que puisque les sARN agissent dans des réseaux complexes, la perte de l'un d'eux serait compensée par d'autres acteurs de ces réseaux. Parmi ces sARN, RsaI est fortement induit lorsque les cellules sont soumises à une exposition à la vancomycine (Howden et al., 2013).

4.4. Les ARN éponges

Comme explicité précédemment, les sARN régulent principalement leurs cibles ARNm au niveau post-transcriptionnel. Cependant, depuis quelques années, il est également apparu que des ARNnc interagissaient entre eux, pour réguler l'activité de l'un des deux ARN de même classe. Ces ARNnc sont appelés ARN éponges (ARNe). En les séquestrant, les ARNe régulent le taux du sARN libre dans la cellule. La séquestration aura pour effet d'empêcher le sARN de remplir ses fonctions régulatrices, soit en bloquant l'interaction avec leurs cibles ARNm, soit en induisant la dégradation du sARN cible (Denham, 2020).

Le premier exemple d'ARNe décrit est l'ARN anti-ChiX (*chbBC-ICR*) de *Salmonella*. Cet ARN est issu de la région inter-cistronique (ICR) entre les gènes *chiB* et *chiC* de l'opéron

chb codant pour des protéines impliquées dans le métabolisme des chitooligosaccharides. *chbBC-ICR* interagit avec le sARN ChiX qui cible l'ARNm *chiP*. ChiX entraîne l'inhibition de la traduction et la dégradation par la RNase E de *chiP*. ChiP est une porine permettant l'import d'oligosaccharides dérivés de la chitine. En l'absence d'inducteur, l'expression de *chiP* est contenue par l'expression constitutive du sARN ChiX. De plus ChiX, n'est pas co-dégradé avec son ARNm cible *chiP*, ce qui permet le maintien de sa concentration intracellulaire. En présence de chitooligosaccharides, deux phénomènes indépendants permettent l'expression de ChiP. Premièrement le facteur de transcription ChbR, de l'opéron *chb*, subit une activation allostérique en présence de chitobiose, lui permettant d'activer la transcription de son propre opéron et de *chiP*. Deuxièmement l'activation de l'expression de l'opéron *chb* entraîne la transcription de l'ARNe *chbBC-ICR*, qui interagit avec ChiX. Cette interaction titre le sARN ChiX libre, et favorise sa dégradation. Ces deux phénomènes conjoints permettent de lever la répression de *chiP* par ChiX, permettant la synthèse de porine à chitobiose, lorsque ce composé est présent dans l'environnement (Figueroa-Bossi et al., 2009).

Par ailleurs, l'origine des ARNe est très diverse. En effet il a été montré que les ARNe peuvent émerger de régions inter-cistroniques ou inter-géniques. Ils peuvent également être issus de transcrits indépendants, d'une maturation depuis les régions UTR des ARNm, d'une terminaison prématurée de la transcription, d'un promoteur interne, ou bien d'ARN antisens (Denham, 2020). De manière inattendue, certains espaceurs entre transcrits d'ARNt (ETS *external transcribed spacer*) peuvent agir comme ARNe (Lalaouna et al., 2015). De prime abord, ces produits de maturation, souvent rapidement dégradés, ont été longtemps considérés comme des déchets cellulaires. Cependant, certains d'entre eux sont stables questionnant sur leur possible fonctionnalité. Chez *E. coli*, le 3'ETS^{leuZ} fait partie de cette famille d'ARNe, puisqu'il interagit avec les sARN RyhB et RybB, dont l'expression est activée respectivement lors d'une perturbation de l'homéostasie du fer et de la disruption de l'intégrité de la membrane. Ces deux sARN régulent leurs cibles ARNm pour répondre à ces stress. En l'absence de stress, les synthèses de RyhB et de RybB sont alors réprimées. Cependant, malgré cette répression, il demeure un bruit de fond transcriptionnel. Le 3'ETS^{leuZ} en interagissant avec RyhB et RybB élimine ce bruit de fond, les rendant complètement inactifs lorsqu'ils sont à faible concentration (Lalaouna et al., 2015). En revanche, lorsque l'expression de RyhB et RybB est activée, un seuil de concentration suffisant est alors atteint qui n'est plus titrable par les 3'ETS^{leuZ}.

Enfin, il a également été montré qu'un ARNe peut être bifonctionnel, en assurant la séquestration d'un sARN, mais également en régulant des ARNm cibles. En effet, les deux sARN, ArcZ et CyaR, possèdent chacun des cibles ARNm, mais interagissent également

ensemble. La région 5' de CyaR interagit avec la séquence « *seed* » de ArcZ, en résulte une régulation unidirectionnelle de CyaR par ArcZ. La fonction biologique de cette interaction n'est pas encore établie mais celle-ci a été montrée comme entraînant une déstabilisation de CyaR et une diminution de la régulation de ces cibles ARNm (Iosub et al., 2020). Outre le fait que les ARNe peuvent avoir des origines diverses, il est également intéressant de noter qu'ils peuvent présenter un éventail de fonctions différentes.

4.5. Caractérisation et étude des ARN régulateurs

4.5.1. Caractérisation des ARN régulateurs de manière globale et individuelle

La présente revue s'adresse aux experts comme aux débutants qui voudraient intégrer le domaine d'étude des sARN. Elle explique les subtilités qui existent pour pouvoir effectuer un choix parmi les technologies actuelles permettant d'étudier les sARN. Que l'on travaille sur un ARN régulateur en particulier ou que l'on commence par une étude globale pour ensuite étudier un ou plusieurs d'entre eux, certaines étapes sont nécessaires à la caractérisation des ARN, tant du point de vue de la détermination de leurs bornes, de leur structure ou encore de leurs fonctions.



Contents lists available at ScienceDirect

BBA - Gene Regulatory Mechanisms

journal homepage: www.elsevier.com/locate/bbagrm

Navigation through the twists and turns of RNA sequencing technologies: Application to bacterial regulatory RNAs[☆]

Emma Desgranges, Isabelle Caldelari, Stefano Marzi, David Lalaouna^{*}

Université de Strasbourg, CNRS, ARN UPR 9002, F-67000 Strasbourg, France

ARTICLE INFO

Keywords:

Small regulatory RNA
RNA sequencing
Post-transcriptional regulation
Targetome
Regulatory network
Bacteria

ABSTRACT

Discovered in the 1980s, small regulatory RNAs (sRNAs) are now considered key actors in virtually all aspects of bacterial physiology and virulence. Together with transcriptional and translational regulatory proteins, they integrate and often are hubs of complex regulatory networks, responsible for bacterial response/adaptation to various perceived stimuli. The recent development of powerful RNA sequencing technologies has facilitated the identification and characterization of sRNAs (length, structure and expression conditions) and their RNA targets in several bacteria. Nevertheless, it could be very difficult for non-experts to understand the advantages and drawbacks related to each offered option and, consequently, to make an informed choice. Therefore, the main goal of this review is to provide a guide to navigate through the twists and turns of high-throughput RNA sequencing technologies, with a specific focus on those applied to the study of sRNAs.

This article is part of a Special Issue entitled: RNA and gene control in bacteria edited by Dr. M. Guillier and F. Repoila.

1. Introduction

The continuous advances in the characterization of bacterial small regulatory RNAs (sRNAs) have allowed the appreciation of their diversity regarding their origins, sizes and functions [1,2]. Usually, the transcription of sRNAs relies on two-component systems and/or transcriptional factors responding directly to environmental stimuli or metabolites [2]. This heterogeneous group of RNAs includes not only *bone fide* sRNAs, acting in *trans* and having their own promoter and transcriptional terminator, but also RNA fragments released from precursor RNAs. Typically, these fragments are the result of an RNase-dependent cleavage (e.g. RNase E, RNase III) of 3' untranslated region (UTR) of mRNAs [3], 3' external transcribed spacer of tRNA transcripts [4] or sRNAs, which are processed to give mature forms. For example, ArcZ sRNA is transcribed as a 120 nt-long fragment in *Escherichia coli* and then processed into two shorter forms, a low abundant (88 nts) and a stable (55 nts) fragments, sharing the same 3' end [5]. The genesis of an sRNA (proper transcription or maturation) can be easily discriminated by analyzing their distinctive 5' ends. Primary RNAs, which are produced from their own promoter, have a 5' triphosphorylated (5'PPP) end, while secondary RNAs which are processed from a precursor transcript, have a 5' monophosphorylated (5'P) end. Hence, 5' ends are characterized as transcription start sites (TSS) and processing

start sites (PSS), respectively.

sRNAs do not only differ in their origins but also in their lengths. They are commonly described as ranging from 50 to 500 nts. However, recent findings revealed that they are even more heterogeneous in size, from dozens to over a thousand nucleotides. Indeed, very small RNA fragments (such as tRNA-derived fragments [6]) and large transcripts (e.g. 1116 nt-long RsaC sRNA in *Staphylococcus aureus* [7]) have been described as trans-acting regulators.

sRNAs act mainly by forming non-contiguous pairings with their target mRNAs, thus regulating their translation and/or stability. Chaperone proteins like Hfq or ProQ are often involved in sRNAs stability and/or functions [8,9]. Remarkably, an sRNA can regulate many mRNA targets, constituting its targetome, and one target can be regulated by numerous sRNAs. As a result, sRNAs post-transcriptionally control the expression of their multiple mRNA targets involved in all kinds of cellular processes and can even regulate their own transcription factors *via* feedback loops. They are often non-coding, meaning that they do not bear ribosome binding signals and in frame start and stop codons, the key features of coding sequences. Yet few are named bifunctional or dual-function sRNAs, as they encode peptides like SgrS in *E. coli* and RNAIII in *S. aureus* [10], coding for the SgrT peptide and the Hld toxin, respectively. In addition, a few sRNAs buffer other sRNA transcripts [4,11] or cellular proteins [12,13]. In a nutshell, sRNAs

[☆] This article is part of a Special Issue entitled: RNA and gene control in bacteria edited by Dr. M. Guillier and F. Repoila.

^{*} Corresponding author.

E-mail address: d.lalaouna@ibmc-cnrs.unistra.fr (D. Lalaouna).

<https://doi.org/10.1016/j.bbagrm.2020.194506>

Received 14 November 2019; Received in revised form 11 February 2020; Accepted 13 February 2020

Available online 14 February 2020

1874-9399/ © 2020 Elsevier B.V. All rights reserved.

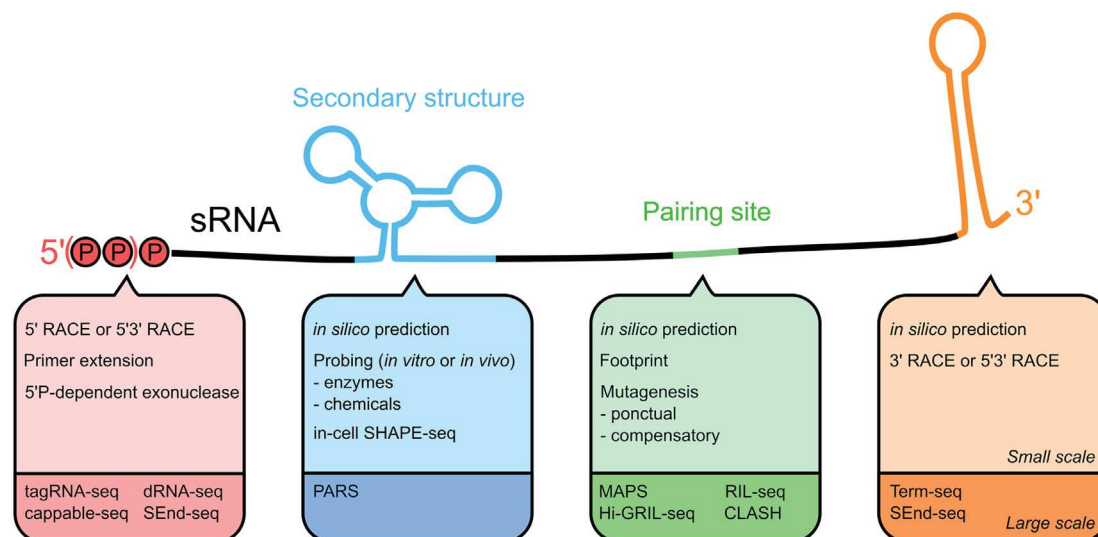


Fig. 1. Schematic representation of the main methods (either on a single RNA, upper panels, or globally, lower panels) used to characterize an sRNA.

integrate various internal and/or external signals and are at the center of complex regulatory networks, impacting every aspect of bacterial physiology/virulence.

This review will provide an overview of cutting-edge technologies/approaches for the complete characterization of bacterial sRNAs, their whole targetome and their cellular functions. Due to the profusion of available algorithms/methods, we provide here a non-comprehensive list, and we sincerely apologize to authors of uncited works.

2. Identification of sRNA pool in a bacterial genome

Historically, sRNAs have been mainly identified by computational searches (potential promoters and rho-independent terminators in intergenic region), comparative genomics (sequence conservation) or co-purification with RNA-binding proteins such as Hfq and CsrA/RsmA [14,15]. However, the introduction of tiling microarrays and RNA sequencing (RNA-seq)-based approaches has revolutionized our understanding of bacterial transcriptomes, including sRNAs. With the development of next-generation sequencing (NGS) technologies, the cost associated to transcriptomic analysis has decreased at a dramatic rate, making RNA-seq more affordable. As a consequence, we have recently observed an explosion of transcriptomics information. Indeed, the study of sRNAs is no longer restricted to model organisms such as *E. coli* or *Salmonella Typhimurium*. The transcriptome of many pathogenic strains such as *S. aureus* [16], *Helicobacter pylori* [17], *Mycobacterium tuberculosis* [18], *Streptococcus pneumonia* [19], *Clostridium difficile* [20], *Pseudomonas aeruginosa* [21], but also of too often neglected environmental strains [22–24], has led to the discovery of many non-coding transcripts with potential regulatory functions. Remarkably, the improvement in sensitivity of sequencing technologies has revealed that bacterial sRNAs are not only restricted to the cytoplasm but are also packaged in secreted membrane vesicles to suppress host immune defenses [6,25]. Indeed, *P. aeruginosa* produces outer membrane vesicles containing sRNAs, including a tRNA-derived fragment, which is released in the host cell and reduces IL-8 secretion [6]. Similarly, the intracellular pathogen *Listeria monocytogenes* secretes both “naked” RNAs and intravesicular RNAs [25].

RNA-seq data mining and analysis is still arduous and time consuming. Algorithms and tools such as Rockhopper [26], ANNOgesic [27], sRNA-Detect [28], and APERO [29] were developed to facilitate the genome-wide identification of sRNAs which are inferred from RNA sequencing data. Leonard et al. [29], compared commonly used algorithms using *E. coli* and *S. Typhimurium* RNA-seq data, and concluded that the number of identified sRNAs is extremely variable, from

hundreds to thousands. This is typically due to bias of respective detection methods. Indeed, most of them are using read coverage distribution to define sRNAs, searching either for a uniform coverage or setting a coverage threshold. Even if these approaches are quite efficient for eukaryotic sRNAs (e.g. miRNAs and siRNAs), bacterial sRNAs are longer and often exhibit a non-homogeneous coverage due to their strong secondary structure, their processing by endoribonucleases (e.g. RNase E and RNase III) and exoribonucleases (e.g. PNPase) or the presence of repetitive elements. This ultimately leads to an over-estimation of their number and an incorrect deduction of their extremities. The type of library preparation (size-selection, fragmentation, adaptor ligation or random priming, single- or paired-end sequencing) can also impact the outcome. Among the above-mentioned analysis algorithms, a main advantage of APERO (Analysis of Paired-End RNA-Seq Output) is to directly use individually sequenced non-fragmented transcripts (paired-end). Hence, APERO circumvents the lack of coverage homogeneity, and infers more accurately the 5' and 3' extremities of sRNAs. The size-selection of non-fragmented transcripts in the sRNA library preparation (50 to 200 nts) also helps to identify RNA fragments deriving from processed precursors, which are otherwise often lost in the mass of fragmented reads corresponding to the respective primary transcripts in a global transcriptomic approach. However, a main drawback of this method is the size-selection, with the possible exclusion of very small and large regulatory RNAs. Tested experimental conditions can also limit sRNAs identification, as specific stimuli trigger their production. To circumvent the limitation, RNA extracted in different growth conditions could be pooled before the analysis.

Despite the improvements made in sRNA prediction tools based on RNA-seq data, the proper identification of sRNAs and their boundaries still require dedicated experimental protocols and *in vivo* validations.

3. Characterization of sRNAs at a single molecule or a global scale

Full sRNAs characterization is a prerequisite for the experimental design to unveil their targetome and functions (Section 4). For example, a careful determination of the 5' and 3' end of sRNAs (and their mRNA targets) is required for cloning and mutagenesis purposes and, consequently, for the characterization of sRNA-dependent regulatory mechanisms. This is also important for the study of sRNAs genesis and expression.

In the following paragraphs, we will describe suitable methods allowing the accurate determination of sRNA extremities and their secondary structure, at a single molecule or a global level (Fig. 1). To note, most mentioned methods are not specific to sRNAs and could be applied

to any RNA transcript.

3.1. Characterization of their 5' extremity

One of the biggest challenges in the sRNAs characterization is to determine their extremities. For this purpose, specific protocols have been developed to identify 5' ends and discriminate between transcription and processing start sites, while other methods allow the direct sequencing of 3' ends (Section 3.2). Interestingly, both extremities can also be simultaneously determined (Section 3.3).

Several approaches are available for the determination of the 5' end of individual transcripts, such as 5' RACE (Rapid amplification of cDNA ends) and primer extension, which are here concisely described. There are many variants of the 5' RACE method. Commercial kits are also available, such as 5' RACE System for Rapid Amplification of cDNA Ends (Invitrogen). Here, we describe an efficient and user-friendly protocol routinely used in many laboratories [30]. This method is inexpensive and relatively easy to implement. A complementary DNA (cDNA) is obtained by reverse transcription (RT) using a sequence-specific oligonucleotide from total RNA. The design of this primer is a decisive step. It must be highly selective for your RNA of interest and located in the vicinity of the +1 (50–150 nucleotides downstream), which requires prior knowledge concerning the putative 5' end. If not, different oligos can be used. An adaptor of known sequence, acting as an anchor sequence, is then linked at the 3' end of the single stranded cDNA by the T4 RNA ligase. A complementary oligonucleotide is used to initiate the synthesis of the second strand. The double-stranded cDNA is amplified by nested PCR and finally inserted into a cloning vector. The nested PCR corresponds to two successive rounds of PCR using two distinct and specific primers along with the oligonucleotide complementary to the 3' adaptor. The first amplicon is used as template for the second PCR primed by an internal oligonucleotide. This approach highly reduces non-specific amplifications. After sequencing, the +1 of your favorite RNA is identified as the first nucleotide downstream the adaptor sequence. As reverse transcriptases are subject to premature termination, the 5' end could be misattributed. Hence, the ligation of an RNA adaptor directly to total RNA 5' ends using T4 RNA ligase prior to RT could be another strategy for reducing non-specific amplifications [31]. However, in return, the reaction will mainly depend on ligation efficiency and RNA abundance.

Primer extension assays are also commonly used to determine the +1 of specific transcripts by RT run-off signal analysis. A detailed protocol is provided by Carey et al. [32]. Unlike 5' RACE approach, primer extension assays are generally performed by using a P^{32} radio-labeled DNA oligonucleotide to prime the RT reaction. The 5' terminus of a sequence-specific DNA primer is thus radiolabeled with $\gamma\text{-}^{32}\text{P}$ ATP by the T4 polynucleotide kinase (PNK). In this way, the chosen RNA is reverse transcribed from total RNA, leading to the synthesis of a radioactive single-stranded cDNA. The sample is then loaded on a high-resolution polyacrylamide gel next to sequencing ladders (G, A, T and C) obtained by Sanger method. Hence, the +1 of the starting transcript is identified without the need for cloning or sequencing. The method can be tricky if the RNA of interest is weakly expressed (due to the absence of PCR amplification), or if it is highly structured (detection of several bands due to premature terminations of RT). A safer and time-saving alternative is to use fluorescently labeled primers and DNA sequencing apparatus [33].

In the last decade, genome-wide RNA sequencing approaches such as dRNA-seq [34], Cappable-seq [35] and TagRNA-seq [36] have been developed to identify the 5' extremity of all transcripts. They also allow investigators to distinguish between primary and processed RNAs and, by extension, between transcription and post-transcriptional maturation at a global scale. In parallel, an *in silico* prediction of putative bacterial promoter elements (–10 and –35 boxes) upstream the identified TSS can be performed.

The first developed method, called differential RNA-seq (dRNA-seq)

[34], identifies TSS at single-base resolution, and discriminates the different types of 5' end by comparing a control library, representing all transcripts, with a library obtained after specific degradation of processed transcripts. This depletion is achieved using a Terminator 5'-phosphate-dependent exonuclease (TEX; Lucigen Corporation) treatment. In both cases, all 5' triphosphates are converted to 5' monophosphates by the tobacco acid pyrophosphatase (TAP; Epicentre) before cDNA synthesis to allow the ligation of an RNA adapter. In addition, RNAs are poly(A) tailed by the poly(A) polymerase, to synthesize the first-strand cDNA using a complementary oligo(T) primer. dRNA-seq method has been widely used to characterize sRNAs in bacteria such as *Helicobacter pylori* [37] and *Neisseria meningitidis* [38], and especially those generated by RNase-dependent processing like 3'UTR-derived sRNAs. The major limitation of this approach is that modified RNAs (5' capped RNAs, see below) or RNAs with a strong secondary structure such as tRNAs, are insensitive to the exonuclease attacks and indistinguishable from 5' triphosphorylated transcripts.

Cappable-seq [35] is using an opposite strategy. Instead of degrading 5' monophosphorylated RNAs, Cappable-seq captures biotinylated 5'triphosphorylated RNAs. Total RNA is first treated by the vaccinia capping enzyme, resulting in the capping of 5'PPP ends with a biotinylated GTP. After enrichment using streptavidin-coated magnetic beads, capped-5' extremities are converted to 5' monophosphorylated ends by RppH treatment. As a control, the streptavidin-dependent enrichment step is skipped. For cDNA libraries preparation, an adaptor is ligated to both 5' and 3' extremities. Remarkably, Cappable-seq was applied to the mouse gut microbiota [35] and the *Wolbachia* bacterial endosymbiont of nematodes [39], enabling TSS determination in complex environments. A main advantage is that this fishing step implies no need for rRNAs depletion. Still, a contamination with processed RNAs is observed (at least 3–5% of enriched RNAs).

TagRNA-seq method [36] is based on the ligation of different tags to PSS and TSS, enabling to discriminate both types of 5' extremities in a single experiment. This represents a main advantage when compared to other technologies which eliminated 5' monophosphorylated extremities and then compare two distinct libraries, a source of bias. 5' monophosphorylated RNAs are first labeled with a specific tag. Subsequently, 5' triphosphates are converted to 5' monophosphates (TAP treatment) and newly obtained 5'P end are labeled with a different tag. In theory, TagRNA-seq clearly distinguishes TSS and PSS, but in practice it is not so straightforward due to several bias (e.g. aspecific ligation, spontaneous cleavage). Consequently, a mixture of TSS and PSS is generally mapped at the same position. Additional bioinformatic analyses are then required.

Remarkably, a cap-like structure has been detected at the 5' end of some sRNAs, which is notably involved in their stability. According to Bird and colleagues [40], the non-canonical initiating nucleotide (NCIN) is incorporated into RNA during transcription initiation by RNA polymerase. The percentage of capped sRNAs seems to be rather limited. Indeed, a few *E. coli* sRNAs are known to be modified in their 5' with nicotinamide adenine dinucleotide (NAD⁺) [41]. In addition, this phenomenon is restricted to a small proportion of the sRNA population (e.g. ~13% of RNAI sRNA). This discovery was made possible by the development of chemo-enzymatic capture and next-generation sequencing NAD captureSeq method [41]. Essentially, the NAD⁺ group is transglycosylated and then biotinylated to enable the streptavidin-dependent purification of NAD-modified 5' end. Another approach, called CapZyme-Seq [42], proposes to extend the study to other nucleoside-containing metabolites, such as NADH, 5'-desphospho coenzyme A (dpCoA) and flavin adenine dinucleotide (FAD), which could potentially function as NCIN capping [40]. To globally detect and quantify NCIN-capped RNAs, CapZyme-Seq method compares two datasets generated by selective enzymatic processing: NCIN-capped 5' ends are treated by the RNA-decapping enzyme NudC and uncapped 5' ends by the RNA 5' polyphosphatase (Rpp) to obtain 5' monophosphorylated RNAs. Hence, Vvedenskaya et al. notably revealed three major capped-

sRNAs, SibE, SibD and OxyS (13–22%) in *E. coli*. A main advantage of CapZyme-Seq is to concurrently identify the sRNA 5' end at the single-nucleotide resolution, but 5'PPP and 5'P are not discriminated. While CapZyme-Seq allows the detection of RNAs carrying other NCIN-caps than NAD⁺, the nature of the non-canonical initiating nucleotide remains unknown. Indeed, NudC recognizes at least four distinct NCIN-caps: NAD⁺, NADH, dpCoA, and FAD. For this purpose, the use of additional and specific processing enzymes should be considered. Another limitation is that processing efficiency could be altered by the presence of strong secondary structures.

3.2. Characterization of their 3' extremity

The vast majority of sRNA possesses a Rho-independent terminator (or intrinsic terminator) [43], composed of a GC-rich palindrome sequence followed by a stretch of U residues (≈ 6 –8 residues). This structural element has the obvious role of marking the 3' end of an sRNA during transcription, but also of stabilizing it, notably being a recognition site of Hfq [44]. A non-negligible proportion of termination is mediated by the ATP-dependent DNA translocase Rho, which likely interacts with the RNA polymerase to induce its release [45]. To our knowledge, no sRNA has been yet described to possess a Rho-dependent terminator. The sequence recognized by Rho (Rho utilization site, rut) is still poorly defined (≈ 60 –80 nts, C/G > 1, C residues every 11–13 nts), making it difficult to accurately identify Rho-dependent termination. Remarkably, the 3' end of an sRNA can differ notably due to transcriptional read-through (e.g. RsaA sRNA possesses two distinct terminators in *S. aureus* [46]).

Intrinsic terminator can be spotted by various bioinformatic tools such as TransTermHP [47], RNIE [48] or iTerm-PseKNC [49]. Very recently, RhoTermPredict was developed to predict Rho-dependent termination, notably by searching for putative rut sites followed by a pause site for the RNA polymerase [50]. Because of a significant rate of false positives, *in silico* predicted terminators must be experimentally validated.

A high-throughput technology called Term-seq has been developed to identify the 3' extremity of every RNA fragment at a single-base resolution [51]. Briefly, an adaptor is first added to each 3' end. RNAs are sheared and then reverse transcribed using a complementary oligonucleotide. Here, another adapter is ligated to the 3' end of the newly synthesized single-strand cDNA to produce a double-stranded cDNA. Thus, the obtained cDNA library is sequenced. Term-seq has notably allowed the discovery of several transcriptional riboswitches. These regulatory elements are found in the 5'UTRs of several mRNAs and adopt a distinct secondary structure in presence of a specific ligand, affecting the transcription or the translation of controlled gene(s).

3.3. Simultaneous determination of 5' and 3' ends

As it could be laborious to separately determine 5' and 3' ends, several techniques have been developed to concurrently characterize both boundaries of RNAs.

The main advantage of the 5'3' RACE (also named circular RACE) is to determine both 5' and 3' ends which does not rely on the ligation of adapters. First, the 5' triphosphorylated end is converted to 5' monophosphate by TAP or RppH treatment. The RNA of interest is then circularized by ligation of its 5'P and 3'OH extremities, which is catalyzed by RNA Ligase [52]. To note, polyethylene glycol (PEG) can be used to improve the ligation efficiency. The reverse transcription is performed using a sequence-specific primer. Then, the site of ligation is amplified by nested PCR, and is finally sequenced.

Library preparation by previously described methods generally require the fragmentation of transcripts, which prevent the determination of both 5' and 3' ends. Remarkably, the high-throughput approach called simultaneous 5' and 3' end sequencing (SEnd-seq) allows the mapping of both extremities at a global scale and without size-selection

[53]. First, a 3' adaptor is added to total RNA using T4 RNA Ligase 1. The first strand cDNA is obtained by RT using a biotinylated primer complementary to the 3' adaptor. After 5' phosphorylation by T4 PNK, cDNAs are circularized by TS2126 RNA Ligase I, favoring intramolecular ligation. Circularized cDNAs are fragmented and the second strand is obtained using the Second Strand cDNA Synthesis Kit (NEB). Biotin-labeled DNA fragments, containing the 5'3' junction, are purified by streptavidin beads and sequenced. Interestingly, SEnd-seq also allows to distinguish between 5'PPP and 5'P extremities by a specific enrichment of primary or processed transcripts. Performed in *E. coli*, SEnd-seq method identified sRNAs of various lengths such as CyaR (87 nts), GcvB (204 nts), CsrB (365 nts). Oddly, SEnd-seq obtained results that differ from previously characterized lengths of GcvB (205 nts) and CsrB (369 nts). As putative technical artefacts cannot be ruled out, these results must be individually validated by some of the previously mentioned methods before drawing any conclusions.

3.4. Determination of their secondary structure

Intramolecular interactions (formation of hydrogen bonds between complementary nucleotides, usually A-U, G-C and G-U) generate stem-loops and more complex structures (e.g. pseudoknots), which are crucial for sRNAs stability, protein recognition and constrain pairing with their targets [54].

Software using individual (e.g. UNAFold [55] and RNAFold [56]) or comparative structure predictions (e.g. RNAalifold [57]) predict RNA secondary structure based on consensus structures, co-variation and minimum free energy. Interestingly, some *in silico* tools integrate information provided by structure probing assays such as SHAPE (selective 2'-hydroxyl acylation analyzed by primer extension; e.g. RNAs-structure [58]). A more comprehensive list of available algorithms is provided by Wright et al. [59].

In vitro RNA probing techniques using chemicals (CMCT, DMS, Pb (II), BzCN) or RNases (A, T1, T2, V1) are performed with any kinds of RNA molecules from any organisms (for review [60]). Direct radioactive labeling of small RNAs (< 300 nucleotides) followed by loading of cleaved RNAs (from RNases action or from chemical treatment of modified bases) on a denaturing polyacrylamide gel is effective, while probing of longer RNAs requires primer extension by reverse transcriptase.

Several high-throughput methods have appeared in the last years to study RNA structure at a large scale. PARS (Parallel Analysis of RNA Structures) has been implemented in bacteria to determine the secondary structure of *in vitro* refolded total RNA, using two distinct cocktails of enzymes cleaving either single stranded or double-stranded RNAs [61]. Digested RNA fragments are phosphorylated with T4 PNK prior to library preparation and deep sequencing. In *Yersinia pseudotuberculosis*, Righetti et al. identified multiple RNA riboswitches, called thermometers, controlling gene expression in a temperature-dependent manner [61]. Remarkably, PARS-derived structures of abundant sRNAs such as CsrB and CsrC were also determined. However, a major drawback is that RNA structures are formed *in vitro* after denaturation of total RNA, which could generate artefactual interactions (both intra- and intermolecular). The cellular abundance of sRNAs is also a limiting factor. The ultimate goal is obviously to explore the secondary structure of RNAs in a cellular context. For this purpose, several *in vivo* approaches have been recently developed and are reviewed by Ignatova and Narberhaus [62]. Among them, in-cell SHAPE-seq uses the ability of 1-methyl-7-nitroisatoic anhydride (1M7) to cross membranes and penetrate living cells [63]. 1M7 is a SHAPE reagent which modifies the 2'-hydroxyl group of flexible (unpaired) RNA nucleotides. Modified 2'OH groups, and by extent, reactive nucleotides, are then detected by primer extension using their ability to arrest reverse transcription. Here, Watters and colleagues used specific RT primers to target RNAs of interest such as the *btuB* mRNA riboswitch domain in *E. coli*, which facilitate data analysis. However, to our knowledge, no suitable

approach exists to determine the *in vivo* secondary structure of sRNAs at a global scale. Indeed, the presence of multiple interaction sites (intermolecular interactions with RNA binding proteins and mRNAs) could lead to a misinterpretation of RNA flexibility and, thus drastically complexify data analysis. Therefore, the *in vivo* characterization of sRNAs structure still represents a big challenge in the field.

4. Characterization of their targetome

Several features make the characterization of sRNAs targetome difficult: (1) *trans*-acting sRNAs often interact with multiple target mRNAs (e.g. > 30 targets for GcvB sRNA in *E. coli* [64]) using limited and non-contiguous sequence complementarities; (2) a single sRNA can bear multiple pairing sites, also named “seed” sequences; (3) RNA chaperone proteins such as Hfq [9] and ProQ [8] are often described as promoting the pairing between an sRNA and its partners especially in *Enterobacteriaceae*. During the last decade, several algorithms and experimental approaches have been developed to tackle this major problem.

The pairing site between an sRNA and its targets can be predicted by *in silico* tools such as TargetRNA2 [65], sTarPicker [66], IntaRNA [67] and CopraRNA [68]. These tools are widely compared in the following reviews [59,69,70]. According to Pain et al. [69], CopraRNA outperforms other algorithms, due to the integration of both the accessibility of the interaction sites and the evolutionary conservation of sRNA:mRNA interaction. IntaRNA is an attractive alternative, notably if sequence comparison by different organisms is not possible [69,70]. Remarkably, a recently developed software pipeline, called SPOT for “sRNA target prediction organizing tool” [71], regroups all previously cited algorithms and also integrates experimental data (from MAPS or RIL-seq, see below) to improve targets prediction. However, the false-positive rate remains high. Moreover, a significant part of the RNA target is overlooked, as most algorithms focus on a specific window around the start codon (−200 to +100; default parameters). They also focus only on coding RNAs and completely ignore sRNAs, tRNAs or even rRNAs as putative targets. However, such interactions have been recently described [4,72].

During the last decade, many experimental approaches have been developed to deeply characterize the targetome of sRNAs. The major techniques are based either on the purification of a tagged sRNA with its partners or on the pull-down of all RNAs interacting with an RNA binding protein (RBP). They have been recently reviewed by Saliba et al. [73].

4.1. Deep analysis of a specific sRNA targetome

In numerous organisms, no sRNA-associated protein (e.g. Hfq or ProQ homologs) has been identified. Hence, an elegant strategy is to use an sRNA as bait to co-purify all interacting RNAs. MAPS (MS2-affinity purification coupled with RNA sequencing) technology was performed in several organisms like *E. coli* [4,64,74,75], *Salmonella* Typhimurium [76] and *S. aureus* [77,78]. Detailed protocols are available [79,80]. Briefly, the sRNA of interest is fused to an MS2 aptamer and synthesized *in vivo* (Fig. 2). Interacting RNAs are then co-purified using the strong and specific affinity of MS2-MBP (maltose binding protein) fusion protein (immobilized on an amylose column) for MS2 aptamer. Enriched RNAs are finally eluted by addition of maltose, a binding competitor. MAPS technology deeply characterizes and zooms in on a specific targetome. The co-purification step enables the detection of even low abundant targets. Major limitations are (1) the addition of the MS2 tag could affect function, stability and/or folding of sRNA. Consequently, all MS2-sRNA constructs must be scrupulously validated; (2) some targets might be missed by the use of improper growth conditions; (3) MS2-sRNA are often overexpressed when compared to endogenous expression; (4) indirect co-purification could occur *via* interactions with RNA-binding proteins or other RNAs, then an experimental validation is

still necessary.

An alternative approach called Hybrid-trap-seq [81] is based on the co-purification of interacting RNAs with a 3'-biotinylated RNA *in vitro* (Fig. 2). The sRNA of interest is *in vitro* transcribed and 3' end biotinylated. Total RNA is extracted in various experimental conditions and mixed with the synthetic sRNA bait. *In vitro* formed complexes are then isolated using streptavidin magnetic beads. Notably, Hybrid-trap-seq was performed in *S. aureus* to identify several targets of RsaE sRNA. Major limitations of this approach are: (1) the use of a synthetic RNA. Improper RNA folding and the absence of post-transcriptional modifications could have a detrimental impact on the results; (2) RNA:RNA complexes are formed *in vitro* after denaturation. RNA folding and pairing conditions (RNAs renatured altogether at room temperature) could generate some interactions with regions which are paired *in vivo*. (3) RNA chaperone proteins, which could be involved in RNA folding and RNA:RNA interaction, are not present during *in vitro* annealing.

4.2. sRNA:RNA interactions associated with an RBP

RIL-seq (RNA interaction by ligation and sequencing) [82–84] and CLASH (UV-crosslinking, ligation and sequencing of hybrids) [85] aim to define the global RNA-RNA interactome associated with a specific RBP in bacteria. Both methods share major similarities. They rely on UV crosslinking between an RBP (e.g. ProQ, Hfq or RNase E) and its RNA partners, followed by a ligation between proximate RNAs (Fig. 2). The RNA:RNA chimeras are sequenced and identified by mapping reads on the reference genome. By comparing the first two studies (Hfq/RIL-seq [82] and RNaseE/CLASH [85]), Hör and Vogel (2017) [86] concluded that even if two different bait proteins were used, the number of sRNA:mRNA pairs is very similar, suggesting that Hfq and RNase E act together on these specific hybrids. This is consistent with previous studies showing that Hfq recruits RNase E to induce sRNA-dependent mRNA decay [74,87]. Major limitations of both methods are (1) they require tagging of the RBP of interest, which can prevent their activity; (2) the identification of sRNA:RNA interactions is restricted to sRNAs that function with the RBP of interest, in specific growth conditions; (3) the number of interactions might be under-estimated as crosslinking and ligation steps are not fully efficient; (4) the binding of RBP to RNA duplexes might be transient.

4.3. Unravelling all RNA:RNA interactions

Various approaches are now developed to eliminate the co-purification step (either with an interacting protein or RNA) and provide atlases of every RNA:RNA interaction.

Hi-GRIL-seq (High-throughput global small non-coding RNA target identification by ligation and sequencing) is based on the expression under an inducible promoter of the T4 RNA ligase, which covalently ligates the 3'-OH of an RNA to the 5'-monophosphate of another one (Fig. 2) [88,89]. After removal of rRNAs, the ligated mixture is sequenced. In contrast to RIL-seq, CLASH or MAPS, the method does not require affinity purification steps, avoiding the disassembly of duplexes. However, two major drawbacks of the experiment are the overproduction of the T4 RNA ligase *in vivo*, which could lead to unspecific RNA pairs and the absence of trimming of 5' triphosphate RNAs before ligation. Indeed, without a trimming step, the proximity ligation rate might be variable, mostly depending on the distance between the pairing site and RNA extremities. Here, only 0.26–0.29% of reads are chimeric fragments. Among them, the regulon of the sRNA PrrF1 of *Pseudomonas aeruginosa* was extended and new RNA:RNA interactions revealed [88].

Liu and co-workers [90] developed in *E. coli* a modified version of the CLASH method, which does not require a co-purification step using a specific tagged-protein. Therefore, this approach has the potential to detect every RNA-RNA interaction and, consequently, to unravel the whole interactome. Basically, cells are treated with 4'-aminomethyl

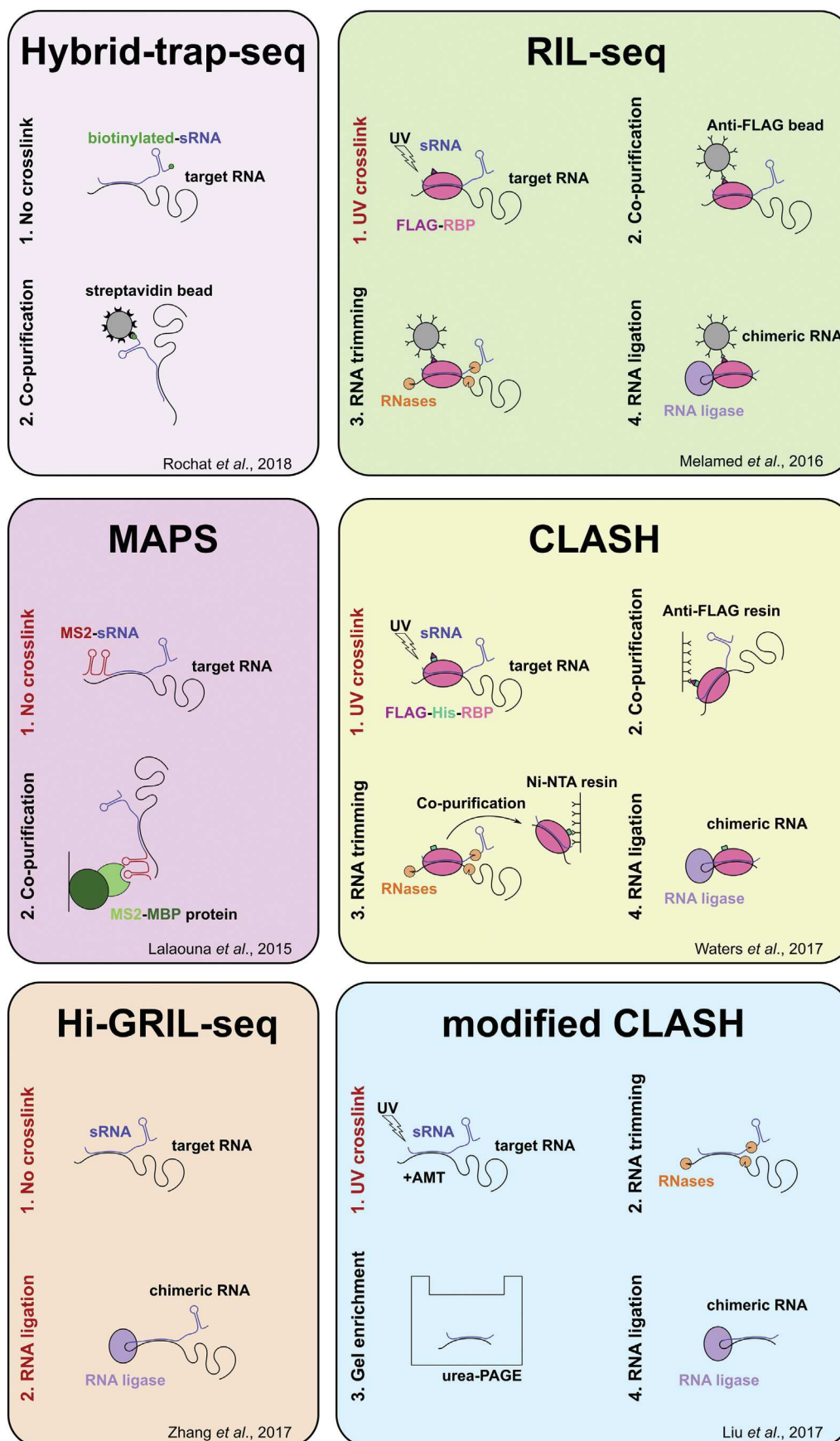


Fig. 2. Comparison of Hybrid-trap-Seq, MAPS, RIL-seq, CLASH, Hi-GRIL-seq and modified CLASH approaches. Steps indicated in red are performed *in vivo*, while those in black are performed *in vitro*. More details are provided in [Section 4](#). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

trioxsalen (AMT) and then exposed to UV (365 nm), which leads to the formation of adducts between opposing uridine residues (Fig. 2). AMT specifically crosslinks RNA:RNA interactions (both intra- and inter-molecular), and not RNA:protein or protein:protein interactions such as formaldehyde. After extraction, total RNA is trimmed with RNase T1, that specifically degrades single-stranded RNA at G residues. Residual single-stranded RNAs are eliminated by hybridization of 20-mer oligonucleotides and RNase H treatment. RNA duplexes are purified on denaturing gel according to their size (40–100 nts). Free RNA extremities are proximity ligated by T4 RNA Ligase 1, before crosslink reversal (UV irradiation at 254 nm). Chimeric fragments are finally submitted to RNA sequencing. To assess bias of the modified CLASH method, several controls are used (e.g. without AMT and/or T4 RNA ligase). After bioinformatics analysis, only 29 RNA:RNA interactions were detected in *E. coli* by Liu et al., mainly involving rRNAs and tRNAs, reflecting the absence of rRNA removal before library construction. Consequently, the deepness of sequencing was probably too low to detect any sRNA-mRNA interactions.

In view of previously cited approaches, we are far from visualizing the whole cellular interactome in bacteria. However, the comparison of recently developed technologies, notably applied to the study of eukaryotic or viral RNA:RNA interactions will certainly help to achieve this goal. Indeed LIGR-Seq, PARIS, SPLASH and COMRADES rely on the same strategy of ligation between chemically cross-linked RNAs with psoralen derivatives (such as AMT). LIGR-Seq method is based on the removal of non-ligated RNAs by enzymatic digestions (such as modified CLASH), while other methods include an additional purification step either by gel enrichment or pull-down of cross-linked RNA hybrids (for review see Schönberger et al. [91]). The main advantage of this purification step is to enrich true cross-linked RNA:RNA interactions and to decrease the level of false positives, resulting in an increase of specificity. After proximity ligation, crosslinks are reversed at 254 nm, and RNA samples are subjected to RNA sequencing. However, due to the weak yield of crosslink and poor ligation efficiency, not all RNA:RNA interactions are revealed. Moreover, AMT has a preference to intercalate into opposite uracil bases. Thus, binding sites with a high proportion of GC residues can be underrepresented. Interestingly, in LIGR-seq method, the circRNA Ligase is favored over the T4 RNA Ligase to allow reaction at high temperature (60 °C instead of 16 °C), which reduces unspecific RNA:RNA interaction and ligation *in vitro*. Yet it will promote circularization of RNAs, which might be excluded in the downstream steps of the library preparation. A good compromise would be to use another RNA Ligase operating at high temperature. The method used to prepare cDNA libraries is also a critical parameter. The adapter ligation method is more suitable than random priming to facilitate the mapping of both chimera extremities and identify interacting partners. In the same vein, samples fragmentation should be avoided. The development of a technology inspired by previously mentioned methods and taking into account suggested improvements will definitively help defining the whole RNA:RNA interactome.

4.4. Validation of candidates

Global analysis of RNA interactome raise many potential RNA:RNA pairs that need to be validated both *in vitro* and *in vivo* in order to decipher their functions. Jagodnik et al. [92] reviewed extensively all the techniques used in the field. Especially, the binding sites of recognition of sRNAs with their mRNA targets can be analyzed by gel retardation assays (also named electrophoretic mobility shift assay, EMSA). Chemicals or enzymatic probing can be used to examine the structure of regulatory complexes. Finally, the mechanisms of regulation can be characterized at the molecular level by analyzing the impact of the RNA:RNA complexes on RNA stability (Northern blot, rifampicin assays) and translation (toe-print, Western blot, gene reporter fusions). Here, we only focused on two recent technological advances that enable to measure at the genomic scale either RNA half-life or translation

efficiency.

The analysis of variations in RNA decay rate in absence/presence of the studied sRNA not only helps understanding regulatory mechanisms, but also scouting new targets. To evaluate the dynamics of RNA transcription, processing and decay at the transcriptome-wide level, metabolic labeling experiments coupled with RNA sequencing have proven to be the *de facto* standard approaches [93]. In most protocols, 4-thiouridine (4sU) label, which is incorporated into newly synthesized transcripts, but not into pre-existing molecules, helps distinguishing both RNA populations. Remarkably, pulse-chase experiments allow the determination of both the speed of transcription and RNAs half-life. For this purpose, two main approaches are used: (1) an enrichment of the metabolically labeled RNAs by biochemical separation is achieved before library preparation. After total RNA isolation, thiol groups of 4sU-labeled RNAs are biotinylated and, thus efficiently purified. This approach, named 4sU-seq, has become quite popular in RNA biology, as exemplified by the increasing number of studies (see [94] for review). (2) 4sU nucleotides are converted into cytosines, which can be observed as point mutations in RNA-seq data (T to C transitions) allowing the partition during RNA-seq analysis of the newly synthesized transcripts from the non-mutated ones [95]. Current protocols are similar to the ones developed for bisulfite sequencing, which also uses chemically induced mutations to provide insight into RNA or DNA methylation. Here, the absence of any biochemical separation makes this method more accessible due to lower input amounts and a less laborious protocol. Although not yet used to study the impact of sRNAs on global mRNA turnover, these promising approaches, analyzing the kinetics of transcription and RNA decay, could lead to a more comprehensive understanding of the sRNA-dependent regulation.

Ribosome profiling experiments (Ribo-seq; [96]) precisely and quantitatively measure the translation rate of mRNAs in the cell, providing the opportunity to address the effect of sRNAs on the translation of mRNA targets. Commonly, sRNAs impact translation at the initiation steps either by repressing ribosome loading (occlusion of the Shine-Dalgarno sequence and/or start codon) or by stimulating it. In the latter case, the sRNA binding to an upstream region of the targeted transcript promotes secondary structure rearrangements around the ribosome binding site, which becomes more accessible. Ribo-seq method includes the purification of ribosomes bound to actively translated mRNAs from cell lysates. After RNase treatment, the ~30 nts nuclease-protected mRNA fragments are sequenced, enabling the determination of ribosome positions on transcripts. Wang and colleagues [97] have adopted Ribo-seq method to experimentally identify regulatory targets of RyhB sRNA in *E. coli*. In addition to validate previously characterized targets, Ribo-seq increased the regulatory targetome of RyhB by unveiling novel mRNA partners. Remarkably, Ribo-seq coupled to transcriptomic analysis gives a rather complete overview of all RNAs regulated by an sRNA at the post-transcriptional and/or translational levels. Still, observed results are not systematically a consequence of direct interactions. Another valuable advantage of this approach is to identify bifunctional sRNAs encoding small open reading frame (ORF) as published recently [98].

5. Conclusion

The recent and considerable evolution in high-throughput sequencing technologies has not only uncovered knowledge on sRNAs characterization and expression, but also on their cellular functions. This enables to build vast and complex regulatory networks integrating perceived stimuli and the impact on bacterial physiology. To go further, cell-to-cell diversity should be considered. Indeed, heterogeneous behavior between bacterial individuals is usually observed, notably in specific conditions such as biofilm or dormancy, providing a benefit to adapt more quickly to environmental changes. Until now, a unique approach addresses it by combining super-resolution imaging with fluorescence *in situ* hybridization (smFISH). This notably revealed the

subcellular localization of a sRNA:mRNA complex and the pairing kinetic in fixed cells [99]. Moreover, very recently Kannaiah and colleagues (2019) [100] were able to perform a transcriptomic analysis of different cellular compartments (cytosol, membrane and poles). They noticed that several Hfq-dependent sRNAs were specifically enriched at the poles together with their mRNA targets. The study revealed that not only the direct interaction between the sRNA and its targets and their relative stoichiometry should be considered but also their cellular co-localization to define regulatory functions.

In the future, continuous progresses in the sensitivity of RNAseq technologies will definitely allow analysis of the expression profiling of sRNAs at single-cell level and unravel their targetome. This will certainly explain cell-to-cell variation within a cell population and help revealing a target hierarchy.

CRediT authorship contribution statement

Emma Desgranges: Writing - review & editing. **Isabelle Caldelari:** Writing - review & editing. **Stefano Marzi:** Writing - review & editing. **David Lalaouna:** Writing - review & editing, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Pascale Romby (PR) for helpful advice and discussions. We also thank the Integrative Molecular and Cellular Biology (IMCBio) graduate school.

Funding

This work was supported by the “Agence Nationale de la Recherche” (ANR, Grant ANR-16-CE11-0007-01, RIBOSTAPH, and ANR-18-CE12-0025-04, CoNoCo, to PR). It has also been published under the framework of the labEx NetRNA ANR-10-LABX-0036 and of ANR-17-EURE-0023 (to PR), as funding from the state managed by ANR as part of the investments for the future program. DL was supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 753137-SaRNAReg. ED was supported by the “Fondation pour la Recherche Médicale” (FDT201904007957).

References

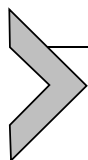
- M.C. Carrier, D. Lalaouna, E. Masse, Broadening the definition of bacterial small RNAs: characteristics and mechanisms of action, *Annu. Rev. Microbiol.* 72 (2018) 141–161.
- A. Brosse, M. Guillier, Bacterial small RNAs in mixed regulatory networks, *Microbiology Spectrum* 6 (2018).
- Y. Chao, K. Papenfort, R. Reinhardt, C.M. Sharma, J. Vogel, An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs, *EMBO J.* 31 (2012) 4005–4019.
- D. Lalaouna, M.C. Carrier, S. Semsey, J.S. Brouard, J. Wang, J.T. Wade, E. Masse, A 3' external transcribed spacer in a tRNA transcript acts as a sponge for small RNAs to prevent transcriptional noise, *Mol. Cell* 58 (2015) 393–405.
- P. Mandin, S. Gottesman, Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA, *EMBO J.* 29 (2010) 3094–3107.
- K. Koeppen, T.H. Hampton, M. Jarek, M. Scharfe, S.A. Gerber, D.W. Mielcarz, E.G. Demers, E.L. Dolben, J.H. Hammond, D.A. Hogan, B.A. Stanton, A novel mechanism of host-pathogen interaction through sRNA in bacterial outer membrane vesicles, *PLoS Pathog.* 12 (2016) e1005672.
- D. Lalaouna, J. Baude, Z. Wu, A. Tomasini, J. Chicher, S. Marzi, F. Vandenesch, P. Romby, I. Caldelari, K. Moreau, RsaC sRNA modulates the oxidative stress response of *Staphylococcus aureus* during manganese starvation, *Nucleic Acids Res.* 47 (18) (2019) 9871–9887.
- A. Smirnov, C. Schneider, J. Hor, J. Vogel, Discovery of new RNA classes and global RNA-binding proteins, *Curr. Opin. Microbiol.* 39 (2017) 152–160.
- R.F. Dos Santos, C.M. Arraiano, J.M. Andrade, New molecular interactions broaden the functions of the RNA chaperone Hfq, *Curr. Genet.* 65 (6) (2019) 1313–1319.
- M. Raina, A. King, C. Bianco, C.K. Vanderpool, Dual-function RNAs, *Microbiology Spectrum* 6 (2018), <https://doi.org/10.1128/microbiolspec.RWR-0032-2018>.
- M. Miyakoshi, Y. Chao, J. Vogel, Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA, *EMBO J.* 34 (2015) 1478–1492.
- T. Romeo, C.A. Vakulskas, P. Babitzke, Post-transcriptional regulation on a global scale: form and function of Csr/Rsm systems, *Environ. Microbiol.* 15 (2013) 313–324.
- K. Bharwad, S. Rajkumar, Rewiring the functional complexity between Crc, Hfq and sRNAs to regulate carbon catabolite repression in *Pseudomonas*, *World journal of microbiology & biotechnology* 35 (2019) 140.
- J. Vogel, C.M. Sharma, How to find small non-coding RNAs in bacteria, *Biol. Chem.* 386 (2005) 1219–1238.
- J. Sridhar, P. Gunasekaran, Computational small RNA prediction in bacteria, *Bioinformatics and biology insights* 7 (2013) 83–95.
- R.K. Carroll, A. Weiss, W.H. Broach, R.E. Wiemels, A.B. Mogen, K.C. Rice, L.N. Shaw, Genome-wide annotation, identification, and global transcriptomic analysis of regulatory or small RNA gene expression in *Staphylococcus aureus*, *MBio* 7 (2016) (e01990-01915).
- S.R. Pernitzsch, C.M. Sharma, Transcriptome complexity and riboregulation in the human pathogen *helicobacter pylori*, *Front. Cell. Infect. Microbiol.* 2 (2012) 14.
- M. Wang, J. Fleming, Z. Li, C. Li, H. Zhang, Y. Xue, M. Chen, Z. Zhang, X.E. Zhang, L. Bi, An automated approach for global identification of sRNA-encoding regions in RNA-Seq data from *mycobacterium tuberculosis*, *Acta Biochim. Biophys. Sin.* Shanghai 48 (2016) 544–553.
- D. Sinha, K. Zimmer, T.A. Cameron, D.B. Rusch, M.E. Winkler, N.R. De Lay, Redefining the small regulatory RNA transcriptome in *Streptococcus pneumoniae* Serotype 2 Strain D39, *J. Bacteriol.* 201 (2019).
- O.A. Soutourina, M. Monot, P. Boudry, L. Saujet, C. Pichon, O. Sismeiro, E. Semenova, K. Severinov, C. Le Bouguenec, J.Y. Coppee, B. Dupuy, I. Martin-Verstraete, Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*, *PLoS Genet.* 9 (2013) e1003493.
- M. Gomez-Lozano, R.L. Marvig, S. Molin, K.S. Long, Genome-wide identification of novel small RNAs in *Pseudomonas aeruginosa*, *Environ. Microbiol.* 14 (2012) 2006–2016.
- M.P. Grull, L. Pena-Castillo, M.E. Mulligan, A.S. Lang, Genome-wide identification and characterization of small RNAs in *Rhodobacter capsulatus* and identification of small RNAs affected by loss of the response regulator CtrA, *RNA Biol.* 14 (2017) 914–925.
- G. Pei, T. Sun, S. Chen, L. Chen, W. Zhang, Systematic and functional identification of small non-coding RNAs associated with exogenous biofuel stress in *Cyanobacterium Synechocystis* sp. PCC 6803, *Biotechnology for biofuels* 10 (2017) 57.
- C. del Val, E. Rivas, O. Torres-Quesada, N. Toro, J.I. Jimenez-Zurdo, Identification of differentially expressed small non-coding RNAs in the legume endosymbiont *Sinorhizobium meliloti* by comparative genomics, *Mol. Microbiol.* 66 (2007) 1080–1091.
- R. Frantz, L. Teubner, T. Schultze, L. La Pietra, C. Muller, K. Gwozdziński, H. Pillich, T. Hain, M. Weber-Gerlach, G.D. Panagiotidis, A. Mostafa, F. Weber, M. Rohde, S. Pleschka, T. Chakraborty, M. Abu Mraheil, The secRNome of *Listeria monocytogenes* harbors small noncoding RNAs that are potent inducers of Beta interferon, *MBio* 10 (2019).
- B. Tjaden, De novo assembly of bacterial transcriptomes from RNA-seq data, *Genome Biol.* 16 (1) (2015).
- S.H. Yu, J. Vogel, K.U. Forstner, ANNOgesic: a Swiss army knife for the RNA-seq based annotation of bacterial/archaeal genomes, *GigaScience* (2018) 7.
- L. Pena-Castillo, M. Gruell, M.E. Mulligan, A.S. Lang, Detection of bacterial small transcripts from RNA-Seq data: A comparative assessment, *Pacific Symposium on Biocomputing*, Pacific Symposium on Biocomputing, vol. 21, 2016, pp. 456–467.
- S. Leonard, S. Meyer, S. Lacour, W. Nasser, F. Hommais, S. Reverchon, APERO: a genome-wide approach for identifying bacterial small RNAs from RNA-Seq data, *Nucleic Acids Res.* 47 (15) (2019) e88.
- D. Tillett, B.P. Burns, B.A. Neilan, Optimized rapid amplification of cDNA ends (RACE) for mapping bacterial mRNA transcripts, *BioTechniques* 28 (2000) 448 (450, 452–443, 456).
- M. Fromont-Racine, E. Bertrand, R. Pictet, T. Grange, A highly sensitive method for mapping the 5' termini of mRNAs, *Nucleic Acids Res.* 21 (1993) 1683–1684.
- M.F. Carey, C.L. Peterson, S.T. Smale, The primer extension assay, *Cold Spring Harb Protoc* 2013 (2013) (pdb.prot071902).
- R.A. Fekete, M.J. Miller, D.K. Chatteraj, Fluorescently labeled oligonucleotide extension: a rapid and quantitative protocol for primer extension, *BioTechniques* 35 (2003) 90–94, 97–98.
- C.M. Sharma, J. Vogel, Differential RNA-seq: the approach behind and the biological insight gained, *Curr. Opin. Microbiol.* 19 (2014) 97–105.
- L. Ettwiller, J. Buswell, E. Yigit, I. Schildkraut, A novel enrichment strategy reveals unprecedented number of novel transcription start sites at single base resolution in a model prokaryote and the gut microbiome, *BMC Genomics* 17 (2016) 199.
- N. Innocenti, M. Golumbeanu, A. Fouquier d'Herouel, C. Lacoux, R.A. Bonnin, S.P. Kennedy, F. Wessner, P. Serror, P. Boulou, F. Repoila, E. Aurell, Whole-genome mapping of 5' RNA ends in bacteria by tagged sequencing: a comprehensive view in *Enterococcus faecalis*, *Rna* 21 (2015) 1018–1030.
- C.M. Sharma, S. Hoffmann, F. Darfeuille, J. Reignier, S. Findeiss, A. Sittka, S. Chabas, K. Reiche, J. Hackermüller, R. Reinhardt, P.F. Stadler, J. Vogel, The primary transcriptome of the major human pathogen *Helicobacter pylori*, *Nature* 464 (2010) 250–255.
- N. Heidrich, S. Bauriedl, L. Barquist, L. Li, C. Schoen, J. Vogel, The primary transcriptome of *Neisseria meningitidis* and its interaction with the RNA chaperone Hfq, *Nucleic Acids Res.* 45 (2017) 6147–6167.
- A.N. Luck, B.E. Slatko, J.M. Foster, Removing the needle from the haystack: enrichment of *Wolbachia* endosymbiont transcripts from host nematode RNA by

- Cappable-seq, *PLoS One* 12 (2017) e0173186.
- [40] J.G. Bird, Y. Zhang, Y. Tian, N. Panova, I. Barvik, L. Greene, M. Liu, B. Buckley, L. Krasny, J.K. Lee, C.D. Kaplan, R.H. Ebright, B.E. Nickels, The mechanism of RNA 5' capping with NAD⁺, NADH and desphospho-CoA, *Nature* 535 (2016) 444–447.
 - [41] M.L. Winz, H. Cahova, G. Nubel, J. Frindert, K. Hofer, A. Jaschke, Capture and sequencing of NAD-capped RNA sequences with NAD captureSeq, *Nat. Protoc.* 12 (2017) 122–149.
 - [42] I.O. Vvedenskaya, J.G. Bird, Y. Zhang, Y. Zhang, X. Jiao, I. Barvik, L. Krasny, M. Kiledjian, D.M. Taylor, R.H. Ebright, B.E. Nickels, CapZyme-Seq comprehensively defines promoter-sequence determinants for RNA 5' capping with NAD, *Mol. Cell* 70 (2018) 553–564 (e559).
 - [43] J. Chen, T. Morita, S. Gottesman, Regulation of transcription termination of small RNAs and by small RNAs: molecular mechanisms and biological functions, *Front. Cell. Infect. Microbiol.* 9 (2019) 201.
 - [44] H. Otake, H. Ishikawa, T. Morita, H. Aiba, PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action, *Proc. Natl. Acad. Sci. U. S. A.* 108 (2011) 13059–13064.
 - [45] P. Mitra, G. Ghosh, M. Hafeezunnisa, R. Sen, Rho protein: roles and mechanisms, *Annu. Rev. Microbiol.* 71 (2017) 687–709.
 - [46] E. Lioliou, C.M. Sharma, I. Caldeleri, A.C. Helfer, P. Fechter, F. Vandenesch, J. Vogel, P. Romby, Global regulatory functions of the *Staphylococcus aureus* endoribonuclease III in gene expression, *PLoS Genet.* 8 (2012) e1002782.
 - [47] C.L. Kingsford, K. Ayanbule, S.L. Salzberg, Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake, *Genome Biol.* 8 (2007) R22.
 - [48] P.P. Gardner, L. Barquist, A. Bateman, E.P. Nawrocki, Z. Weinberg, RNIE: genome-wide prediction of bacterial intrinsic terminators, *Nucleic Acids Res.* 39 (2011) 5845–5852.
 - [49] C.Q. Feng, Z.Y. Zhang, X.J. Zhu, Y. Lin, W. Chen, H. Tang, H. Lin, iTerm-PseKNC: a sequence-based tool for predicting bacterial transcriptional terminators, *Bioinformatics* 35 (2019) 1469–1477.
 - [50] M. Di Salvo, S. Puccio, C. Peano, S. Lacour, P. Alifano, RhoTermPredict: an algorithm for predicting Rho-dependent transcription terminators based on *Escherichia coli*, *Bacillus subtilis* and *Salmonella enterica* databases, *BMC Bioinformatics* 20 (2019) 117.
 - [51] D. Dar, M. Shamir, J.R. Mellin, M. Koutero, N. Stern-Ginossar, P. Cossart, R. Sorek, Term-seq reveals abundant rpo-regulation of antibiotics resistance in bacteria, *Science* 352 (2016) aad9822.
 - [52] R.A. Britton, T. Wen, L. Schaefer, O. Pellegrini, W.C. Uicker, N. Mathy, C. Tobin, R. Daou, J. Szyk, C. Condon, Maturation of the 5' end of *Bacillus subtilis* 16S rRNA by the essential ribonuclease YkqC/RNase J1, *Mol. Microbiol.* 63 (2007) 127–138.
 - [53] X. Ju, D. Li, S. Liu, Full-length RNA profiling reveals pervasive bidirectional transcription terminators in bacteria, *Nat. Microbiol.* 4 (2019) 1907–1918.
 - [54] N. Peterman, A. Lavi-Itzkovitz, E. Levine, Large-scale mapping of sequence-function relations in small regulatory RNAs reveals plasticity and modularity, *Nucleic Acids Res.* 42 (2014) 12177–12188.
 - [55] N.R. Markham, M. Zuker, UNAFold: software for nucleic acid folding and hybridization, *Methods Mol. Biol.* 453 (2008) 3–31.
 - [56] R. Lorenz, S.H. Bernhart, C. Honer Zu Siederdissen, H. Tafer, C. Flamm, P.F. Stadler, I.L. Hofacker, ViennaRNA package 2.0, *Algorithms for molecular biology: AMB* 6 (2011) 26.
 - [57] I.L. Hofacker, M. Fekete, P.F. Stadler, Secondary structure prediction for aligned RNA sequences, *J. Mol. Biol.* 319 (2002) 1059–1066.
 - [58] C.E. Hajdin, S. Bellaousov, W. Huggins, C.W. Leonard, D.H. Mathews, K.M. Weeks, Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots, *Proc. Natl. Acad. Sci. U. S. A.* 110 (2013) 5498–5503.
 - [59] P.R. Wright, M. Mann, R. Backofen, Structure and interaction prediction in prokaryotic RNA biology, *Microbiology Spectrum* 6 (2018).
 - [60] E. Mailler, J.C. Paillart, R. Marquet, R.P. Smyth, V. Vivet-Boudou, The evolution of RNA structural probing methods: from gels to next-generation sequencing, *Wiley Interdisciplinary Reviews. RNA* 10 (2019) e1518.
 - [61] F. Righetti, A.M. Nuss, C. Twittenhoff, S. Beele, K. Urban, S. Will, S.H. Bernhart, P.F. Stadler, P. Dersch, F. Narberhaus, Temperature-responsive in vitro RNA structure of *Yersinia pseudotuberculosis*, *Proc. Natl. Acad. Sci. U. S. A.* 113 (2016) 7237–7242.
 - [62] Z. Ignatova, F. Narberhaus, Systematic probing of the bacterial RNA structure to reveal new functions, *Curr. Opin. Microbiol.* 36 (2017) 14–19.
 - [63] K.E. Watters, T.R. Abbott, J.B. Lucks, Simultaneous characterization of cellular RNA structure and function with in-cell SHAPE-Seq, *Nucleic Acids Res.* 44 (2016) e12.
 - [64] D. Lalaouna, A. Eyraud, A. Devinck, K. Prevost, E. Masse, GcvB small RNA uses two distinct seed regions to regulate an extensive targetome, *Mol. Microbiol.* 111 (2019) 473–486.
 - [65] M.B. Kery, M. Feldman, J. Livny, B. Tjaden, TargetRNA2: identifying targets of small regulatory RNAs in bacteria, *Nucleic Acids Res.* 42 (2014) W124–W129.
 - [66] X. Ying, Y. Cao, J. Wu, Q. Liu, L. Cha, W. Li, sTarPicker: a method for efficient prediction of bacterial sRNA targets based on a two-step model for hybridization, *PLoS One* 6 (2011) e22705.
 - [67] M. Mann, P.R. Wright, R. Backofen, IntaRNA 2.0: enhanced and customizable prediction of RNA-RNA interactions, *Nucleic Acids Res.* 45 (2017) W435–W439.
 - [68] P.R. Wright, J. Georg, M. Mann, D.A. Sorescu, A.S. Richter, S. Lott, R. Kleinkauf, W.R. Hess, R. Backofen, CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains, *Nucleic Acids Res.* 42 (2014) W119–W123.
 - [69] A. Pain, A. Ott, H. Amine, T. Rochat, P. Boulloc, D. Gautheret, An assessment of bacterial small RNA target prediction programs, *RNA Biol.* 12 (2015) 509–513.
 - [70] D. Lai, I.M. Meyer, A comprehensive comparison of general RNA-RNA interaction prediction methods, *Nucleic Acids Res.* 44 (2016) e61.
 - [71] A.M. King, C.K. Vanderpool, P.H. Degnan, sRNA Target Prediction Organizing Tool (SPOT) integrates computational and experimental data to facilitate functional characterization of bacterial small RNAs, *mSphere*, 2019, p. 4.
 - [72] D. Bronesky, E. Desgranges, A. Corvaglia, P. Francois, C.J. Caballero, L. Prado, A. Toledo-Arana, I. Lasa, K. Moreau, F. Vandenesch, S. Marzi, P. Romby, I. Caldeleri, A multifaceted small RNA modulates gene expression upon glucose limitation in *Staphylococcus aureus*, *EMBO J.* 38 (2019).
 - [73] A.E. Saliba, S. C Santos, J. Vogel, New RNA-seq approaches for the study of bacterial pathogens, *Curr. Opin. Microbiol.* 35 (2017) 78–87.
 - [74] D. Lalaouna, K. Prevost, G. Laliberte, V. Houe, E. Masse, Contrasting silencing mechanisms of the same target mRNA by two regulatory RNAs in *Escherichia coli*, *Nucleic Acids Res.* 46 (2018) 2600–2612.
 - [75] D. Lalaouna, A. Morissette, M.C. Carrier, E. Masse, DsrA regulatory RNA represses both hns and rbsD mRNAs through distinct mechanisms in *Escherichia coli*, *Mol. Microbiol.* 98 (2015) 357–369.
 - [76] I.J. Silva, S. Barahona, A. Eyraud, D. Lalaouna, N. Figueroa-Bossi, E. Masse, C.M. Arraiano, SraL sRNA interaction regulates the terminator by preventing premature transcription termination of rho mRNA, *Proc. Natl. Acad. Sci. U. S. A.* 116 (2019) 3042–3051.
 - [77] D. Bronesky, E. Desgranges, A. Corvaglia, P. François, C.J. Caballero, L. Prado, A. Toledo-Arana, I. Lasa, K. Moreau, F. Vandenesch, S. Marzi, P. Romby, I. Caldeleri, A Dual sRNA in *Staphylococcus aureus* Induces a Metabolic Switch Responding to Glucose Consumption, *bioRxiv*, 2018.
 - [78] A. Tomasini, K. Moreau, J. Chicher, T. Geissmann, F. Vandenesch, P. Romby, S. Marzi, I. Caldeleri, The RNA targetome of *Staphylococcus aureus* non-coding RNA RsaA: impact on cell surface properties and defense mechanisms, *Nucleic Acids Res.* 45 (2017) 6746–6760.
 - [79] D. Lalaouna, E. Desgranges, I. Caldeleri, S. Marzi, Chapter sixteen - MS2-affinity purification coupled with RNA sequencing approach in the human pathogen *Staphylococcus aureus*, *Methods Enzymol.* 612 (2018) 393–411.
 - [80] D. Lalaouna, K. Prevost, A. Eyraud, E. Masse, Identification of unknown RNA partners using MAPS, *Methods* 117 (2017) 28–34.
 - [81] T. Rochat, C. Bohn, C. Morvan, T.N. Le Lam, F. Razvi, A. Pain, C. Toffano-Nioche, P. Ponien, A. Jacq, E. Jacquet, P.D. Fey, D. Gautheret, P. Boulloc, The conserved regulatory RNA RsaE down-regulates the arginine degradation pathway in *Staphylococcus aureus*, *Nucleic Acids Res.* 46 (2018) 8803–8816.
 - [82] S. Melamed, A. Peer, R. Faigenbaum-Romm, Y.E. Gatt, N. Reiss, A. Bar, Y. Altuvia, L. Argaman, H. Margalit, Global mapping of small RNA-target interactions in bacteria, *Mol. Cell* 63 (2016) 884–897.
 - [83] S. Melamed, R. Faigenbaum-Romm, A. Peer, N. Reiss, O. Shechter, A. Bar, Y. Altuvia, L. Argaman, H. Margalit, Mapping the small RNA interactome in bacteria using RIL-seq, *Nat. Protoc.* 13 (2018) 1–33.
 - [84] S. Melamed, P.P. Adams, A. Zhang, H. Zhang, G. Storz, RNA-RNA interactomes of ProQ and Hfq reveal overlapping and competing roles, *Mol. Cell* 77 (2) (2019) 411–425.e7.
 - [85] S.A. Waters, S.P. McAteer, G. Kudla, I. Pang, N.P. Deshpande, T.G. Amos, K.W. Leong, M.R. Wilkins, R. Strugnelli, D.L. Gally, D. Tollervy, J.J. Tree, Small RNA interactome of pathogenic *E. coli* revealed through crosslinking of RNase E, *EMBO J.* 36 (2017) 374–387.
 - [86] J. Hor, J. Vogel, Global snapshots of bacterial RNA networks, *EMBO J.* 36 (2017) 245–247.
 - [87] T. Morita, H. Aiba, RNase E action at a distance: degradation of target mRNAs mediated by an Hfq-binding small RNA in bacteria, *Genes Dev.* 25 (2011) 294–298.
 - [88] Y.F. Zhang, K. Han, C.E. Chandler, B. Tjaden, R.K. Ernst, S. Lory, Probing the sRNA regulatory landscape of *P. aeruginosa*: post-transcriptional control of determinants of pathogenicity and antibiotic susceptibility, *Mol. Microbiol.* 106 (2017) 919–937.
 - [89] K. Han, B. Tjaden, S. Lory, GRIL-seq provides a method for identifying direct targets of bacterial small regulatory RNA by in vivo proximity ligation, *Nat. Microbiol.* 2 (2016) 16239.
 - [90] T. Liu, K. Zhang, S. Xu, Z. Wang, H. Fu, B. Tian, X. Zheng, W. Li, Detecting RNA-RNA interactions in *E. coli* using a modified CLASH method, *BMC Genomics* 18 (2017) 343.
 - [91] B. Schonberger, C. Schaal, R. Schafer, B. Voss, RNA interactomics: recent advances and remaining challenges, *F1000Research* 7 (2018).
 - [92] J. Jagodnik, A. Brosse, T.N. Le Lam, C. Chiaruttini, M. Guillier, Mechanistic study of base-pairing small regulatory RNAs in bacteria, *Methods* 117 (2017) 67–76.
 - [93] A. Uvarovskii, I.S. Naarmann-de Vries, C. Dieterich, On the optimal design of metabolic RNA labeling experiments, *PLoS Comput. Biol.* 15 (2019) e1007252.
 - [94] L. Wachutka, J. Gagneur, Measures of RNA metabolism rates: toward a definition at the level of single bonds, *Transcription* 8 (2017) 75–80.
 - [95] T. Yamada, N. Akimitsu, Contributions of regulated transcription and mRNA decay to the dynamics of gene expression, *Wiley interdisciplinary reviews. RNA* 10 (2019) e1508.
 - [96] N.T. Ingolia, S. Ghaemmaghami, J.R. Newman, J.S. Weissman, Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling, *Science* 324 (2009) 218–223.
 - [97] J. Wang, W. Rennie, C. Liu, C.S. Carmack, K. Prevost, M.P. Caron, E. Masse, Y. Ding, J.T. Wade, Identification of bacterial sRNA regulatory targets using ribosome profiling, *Nucleic Acids Res.* 43 (2015) 10308–10320.
 - [98] J. Weaver, F. Mohammad, A.R. Buskirk, G. Storz, Identifying small proteins by ribosome profiling with stalled initiation complexes, *mBio* 10 (2019).
 - [99] J. Fei, D. Singh, Q. Zhang, S. Park, D. Balasubramanian, I. Golding, C.K. Vanderpool, T. Ha, RNA biochemistry. Determination of in vivo target search kinetics of regulatory noncoding RNA, *Science* 347 (2015) 1371–1374.
 - [100] S. Kannaiah, J. Livny, O. Amster-Choder, Spatiotemporal Organization of the *E. coli* transcriptome: translation Independence and engagement in regulation, *Mol. Cell* 76 (2019) 574–589 (e577).

4.5.2. Caractérisation du targetome d'un ARN régulateur

La technique de MAPS pour « *MS2-affinity purification coupled with RNA-seq* » repose sur le principe de l'isolation des complexes formés *in vivo* entre le sARN d'intérêt et ses ARN cibles à l'aide d'une étiquette MS2 exprimée en 5' du sARN. Cette étiquette est spécifiquement reconnue par la protéine de capsid du phage MS2 (MS2-CP), qui est elle-même fixée sur une résine d'amylose pour retenir spécifiquement les complexes contenant le sARN-MS2 modifié. En couplant cette chromatographie d'affinité au séquençage à haut débit (Mi-seq, Illumina), les ARN liés à l'ARN-MS2 sont caractérisés.

La première utilisation de cette approche remonte à 2009 (Said et al., 2009) où elle avait été utilisée principalement pour la caractérisation des protéines reconnaissant les sARN de *Salmonella enterica* par une analyse de spectrométrie de masse. La technique a ensuite été adaptée à *Escherichia coli* (Lalaouna & Massé, 2015; Lalaouna et al., 2015) et à *S. aureus* (Tomasini et al., 2017) pour déterminer les ARN cibles par séquençage haut débit des ARN.



MS2-Affinity Purification Coupled With RNA Sequencing Approach in the Human Pathogen *Staphylococcus aureus*

David Lalaouna, Emma Desgranges, Isabelle Caldelari¹,
Stefano Marzi¹

Architecture et Réactivité de l'ARN, Université de Strasbourg, IBMC-CNRS, Strasbourg, France

¹Corresponding authors: e-mail address: i.caldelari@ibmc-cnrs.unistra.fr; s.marzi@ibmc-cnrs.unistra.fr

Contents

1. Introduction	394
1.1 <i>Staphylococcus aureus</i> : A Major Human Opportunistic Pathogen	394
1.2 sRNA-Mediated Regulation in Bacteria	395
2. MAPS Protocol Applied to <i>S. aureus</i>	396
2.1 Plasmids, Strains, and Growth Conditions	396
2.2 Preparation of Crude Bacterial Extracts	397
2.3 Purification of MS2-MBP Protein	397
2.4 Affinity Purification of Regulatory Complexes Containing MS2-Tagged RNA	397
2.5 Transcriptomics and RNA Sequencing	399
2.6 RNA Sequencing Analysis	399
2.7 Validation of sRNA Target and Mechanism of Regulation	401
2.8 Frequently Asked Questions	402
Acknowledgments	407
Funding	407
References	407

Abstract

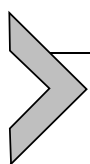
Staphylococcus aureus is a Gram-positive major human pathogen involved in a wide range of human infectious diseases (from minor skin infections to septicemia, endocarditis or toxic shock syndrome). The treatment of *S. aureus* infections is very challenging due to the emergence of multiple antibiotic-resistant isolates. The high diversity of clinical symptoms caused by *S. aureus* depends on the precise expression of numerous virulence factors and stress response pathways, which are tightly regulated at every level (transcriptional, posttranscriptional, translational, and posttranslational).

During the last two decades, it has become evident that small regulatory RNAs (sRNAs) play a major role in fast adaptive responses, mainly by targeting mRNA

translation. sRNAs act as antisense RNAs by forming noncontiguous pairings with their target mRNAs and their mechanisms of action vary according to the interaction site. To obtain a global and detailed view of the regulatory networks involved in the adaptive processes of *S. aureus*, we have adapted the MAPS approach to get individual sRNA targetomes. We also set up different strategies to validate MAPS results and establish sRNA regulatory activities. As this method has been first developed in Gram-negative bacteria, we provide here a protocol for its application in *S. aureus* and highlight underlying differences. Finally, we discuss several points that have been and could be further improved and provide a workflow file for the automatic analysis of the sequencing in Galaxy.

Highlights

- Detailed protocol for MAPS application in *S. aureus* to define sRNA targetome.
- Detailed procedures for both qualitative and quantitative data analysis.
- Analysis of advantages and drawbacks of MAPS.
- Comparison with MAPS applied to Gram-negative bacteria.
- Discussions of critical steps.
- Downloadable Galaxy Workflow for complete MAPS analysis.



1. INTRODUCTION

1.1 *Staphylococcus aureus*: A Major Human Opportunistic Pathogen

Staphylococcus aureus is an opportunistic pathogen, which is the major cause of nosocomial and community infections. This versatile bacterium can invade many tissues and then causes a wide spectrum of infections (cutaneous abscesses, endocarditis, septic shock, etc.). Moreover, the emergence of multidrug resistant strains makes urgent the exploration of new antimicrobial therapies. Its pathogenesis depends on the production of numerous wall-associated or secreted virulence factors (Dale, Doherty-Kirby, Lajoie, & Heinrichs, 2004; Goldmann & Medina, 2017; Novick & Geisinger, 2008; Oliveira, Borges, & Simoes, 2018; Penades, Chen, Quiles-Puchalt, Carpena, & Novick, 2015). They contribute to defense mechanisms against the host immune system, participate in adhesion, invasion, and dissemination in host tissues and also in the acquisition of nutrients, for example, by producing siderophores. The expression of these factors is finely regulated at several levels and requires many actors such as two-component systems, transcription factors, and regulatory RNAs.

All these macromolecules form complex networks that allow *S. aureus* to sense and respond to a wide variety of signals by coordinating the expression of virulence genes dynamically (Tomasini et al., 2014).

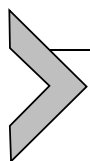
1.2 sRNA-Mediated Regulation in Bacteria

Small regulatory RNAs (sRNAs) have been identified and characterized in numerous bacteria, from Proteobacteria to Firmicutes (Wagner & Romby, 2015). Bacterial genomes, including the human pathogen *S. aureus*, encode a few hundred sRNAs, which are mainly encoded in the core genome, while few are localized within mobile elements, pathogenic islands or plasmids (Abu-Qatouseh et al., 2010; Beaume et al., 2010; Bohn et al., 2010; Geissmann et al., 2009; Pichon & Felden, 2005). These pleiotropic and powerful regulators are involved in every aspect of cell physiology (e.g., primary or secondary metabolism, motility, iron starvation). In *S. aureus*, sRNAs also have a crucial role in virulence (Le Pabic, Germain-Amiot, Bordeau, & Felden, 2015; Manna et al., 2018; Nitzan et al., 2015; Quereda & Cossart, 2017; Romilly et al., 2014), and potentially in antibiotic resistance (Eyraud, Tattevin, Chabelskaya, & Felden, 2014; Howden et al., 2013; Qin, McCausland, Cheung, & Otto, 2016).

Bacterial sRNAs are commonly short (ranging from 50 to 300 nt), non-coding RNA molecules. Some exceptions exist, notably in *S. aureus*, where large coding sRNAs are found (e.g., RNAIII, a 516 nt long sRNA coding for the δ -hemolysin) (Novick et al., 1993; Vanderpool, Balasubramanian, & Lloyd, 2011). Most sRNAs act in *trans* via short, and imperfect base pairings in the vicinity of the ribosome binding site (RBS) and thus interfere with the translational initiation of multiple target mRNAs. Many *S. aureus* sRNAs carry a consensus UCCC motif hybridizing with the Shine-Dalgarno sequence of target mRNAs (Bohn et al., 2010; Geissmann et al., 2009; Romilly et al., 2014). Remarkably, *S. aureus* sRNAs often bind to a second site located either in the 5' untranslated region or in the coding region of the target mRNA. As a consequence, these binding characteristics complicate the identification of their target mRNAs using in silico predictions.

Recent developments in high-throughput RNA sequencing made it possible to draw a comprehensive interaction map of sRNAs (Saliba, C Santos, & Vogel, 2017). Commonly, an RNA binding protein such as the RNA chaperone Hfq is used as bait to copurify sRNA:mRNA duplexes. More recently, Melamed et al. (2016) improved this approach by ligation of sRNA and mRNAs pairing sites (RIL-seq). Although Hfq is involved in several aspects of sRNA-dependent regulation in most studied organisms

(Kavita, de Mets, & Gottesman, 2018; Santiago-Frangos & Woodson, 2018), its role is rather limited in *S. aureus* (Bohn, Rigoulay, & Boulloc, 2007; Liu et al., 2010). As an alternative, MS2-affinity purification coupled with RNA-Seq (MAPS) approach uses a MS2-tagged sRNA to copurify its target mRNAs (Lalaouna, Carrier, et al., 2015; Lalaouna, Prevost, Eyraud, & Masse, 2017). As MAPS has already revealed the targetome of several sRNAs in *Escherichia coli*, we have decided to adapt it to the study of *S. aureus* sRNAs (Tomasini et al., 2017).



2. MAPS PROTOCOL APPLIED TO *S. aureus*

The principle of MAPS is to isolate complexes formed in vivo between sRNA and its target mRNAs using an MS2 tag fused to the 5' end of sRNA. This tag is specifically recognized by the phage MS2 capsid protein (MS2-CP) fused to MBP (maltose binding protein) which is itself bound to an amylose resin to specifically retain complexes containing the modified MS2-sRNA. By coupling this affinity chromatography to high-throughput sequencing, RNAs-bound to a specific sRNA can be characterized in depth.

2.1 Plasmids, Strains, and Growth Conditions

Experiments are performed in the model strain HG001, a derivative of the RN1 strain (NTC8325) with a restored *rsbU* gene, an activator of SigB, and in sRNA deletion mutants obtained by homologous recombination. The genome of HG001 strain has been recently sequenced (Caldelari et al., 2017) (GenBank accession number CP018205.1). Staphylococcal strains are grown routinely in BHI (brain heart infusion) broth at 37°C under agitation. MS2-sRNA plasmids are prepared using pCN51 as template shuttle vector (Charpentier et al., 2004). The MS2 tag alone or MS2-tag fused to the 5' end of sRNAs is cloned under the control of the *agr*-dependent P3 promoter (replacing the cadmium inducible promoter). This allows an accumulation of sRNAs at the late exponential phase of growth, which correspond to most sRNAs natural expression (Geissmann et al., 2009). All plasmids are constructed in *E. coli* strain DC10B (Monk, Shah, Xu, Tan, & Foster, 2012). Plasmids extracted from *E. coli* are directly transformed into *S. aureus* by electroporation.

2.2 Preparation of Crude Bacterial Extracts

For preparation of *S. aureus* crude extracts expressing MS2-RNAs or MS2 alone, cells are grown in 50 mL BHI medium (Sigma) containing 10 µg/mL of erythromycin to an OD₆₀₀ of 5 (~6 h of culture at 37°C), immediately chilled on ice, and then pelleted by centrifugation (3,273 × *g*, 15 min, 4°C). Supernatants are carefully removed, and the pellets are resuspended in 4 mL ice-cold Buffer A (20 mM Tris-HCl pH 8, 150 mM KCl, 1 mM MgCl₂, 1 mM DTT). Cells are then transferred into Lysis Matrix B Tubes (MP Biomedicals) and lysed with FastPrep apparatus (MP Biomedicals). Samples are centrifuged 15 min (15,700 × *g*, 4°C). The supernatants (about 4 mL) are carefully removed and placed in new ice-cold tubes that are maintained at 4°C until subsequent use for the affinity chromatography or for the total RNA extraction by phenol-chloroform (Fig. 1).

2.3 Purification of MS2-MBP Protein

The MS2 coat protein fused to maltose binding protein (MS2-MBP) is expressed in *E. coli* BL21DE (+ pLysS) containing the pMBP-MS2 plasmid (Addgene #65104) (Macias, Bragulat, Tardiff, & Vilardell, 2008) by isopropyl-1-thio-β-D-galactopyranoside (IPTG) induction and then purified as followed. After 3 h induction with 1 mM IPTG, bacterial cells are lysed by sonication in Buffer AB1 (20 mM Tris-HCl pH 8, 200 mM KCl, 1 mM EDTA) supplemented with protease inhibitors and DNase I. Cleared lysate is loaded on an AB1-equilibrated MBPTrap HP column (GE Healthcare). Washes are performed first with Buffer AB1 and then Buffer AB2 (20 mM Tris-HCl pH 8, 20 mM KCl, 1 mM EDTA). For elution, 10 mM of maltose are added to Buffer AB2. Fractions containing MS2-MBP are pooled and loaded on a monoQ anion exchange column (GE Healthcare, monoQ 5/50 GL) equilibrated in Buffer AB2 to remove unspecifically bound nucleic acids. MS2-MBP is then eluted with a gradient of Buffer B (20 mM Tris-HCl pH 8, 1 M KCl, 1 mM EDTA) and dialyzed in Buffer AB2.

2.4 Affinity Purification of Regulatory Complexes Containing MS2-Tagged RNA

Affinity chromatography purification steps are performed at 4°C. For each cell extract (control or MS2-sRNA), 300 µL of amylose resin (New England BioLabs) are packed into a Poly-Prep empty chromatography column

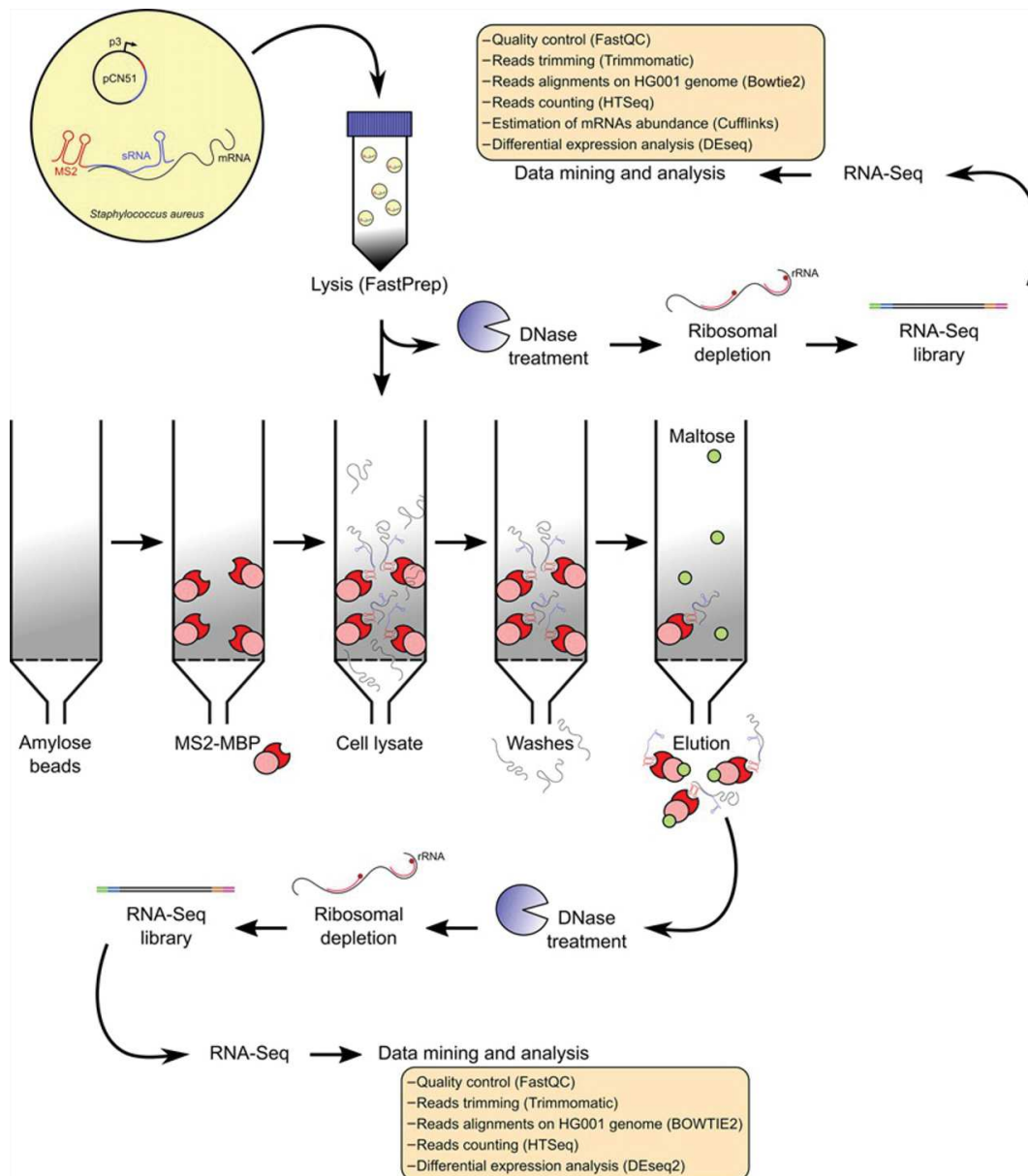


Fig. 1 Schematic representation of MAPS approach in *Staphylococcus aureus*. The MS2-sRNA construct (or MS2 alone) is expressed in vivo (under the control of a quorum sensing-dependent P3 promoter). Bacterial cells are then lysed with FastPrep apparatus. Lysates are loaded onto a column where the MS2-MBP (maltose binding protein) fusion protein is immobilized (amylose resin). The MS2 protein binds with high specificity and affinity to the MS2 aptamer and, therefore, enables the copurification of RNAs interacting with the MS2-sRNA. Several washes allow removing unspecifically bound RNAs (column, beads, MS2-MBP). RNA:RNA duplexes are eluted by the addition of maltose. After DNase I treatment and ribosomal depletion, cDNA libraries are built (ScriptSeq Complete Kit (Bacteria), Illumina) and sequenced (Mi-Seq, Illumina). Data are analyzed using Galaxy Project (Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Cech, M., et al. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 Update. Nucleic Acids Research, 44(W1), W3–W10. doi:10.1093/nar/gkw343).

(Biorad). The columns containing the amylose resin are washed three times with 10 mL Buffer A, and 3600 pmol of MS2-MBP (diluted in 6 mL of Buffer A) are then immobilized. The crude bacterial extracts (4 mL) are directly loaded, followed by three washes with 10 mL Buffer A. Finally, RNAs are eluted with 1 mL of Buffer E (20 mM Tris-HCl pH 8, 150 mM KCl, 1 mM MgCl₂, 1 mM DTT, 0.1% Triton X-100, 12 mM maltose). Eluted RNAs are extracted with phenol:chloroform:isoamylalcohol (25:24:1 (v/v), Roth) and precipitated with three volumes of cold absolute ethanol. MS2-affinity purifications are performed in duplicates.

2.5 Transcriptomics and RNA Sequencing

In order to take into account the expression levels of mRNA candidates, total RNA is purified from bacterial cell extract (Fig. 1) by phenol:chloroform:isoamylalcohol (25:24:1 (v/v), Roth) and precipitated with three volumes of cold absolute. Then both total RNA and MS2-eluted RNAs are treated with DNase I (0.1 U/μL) for 1 h at 37°C to remove genomic DNA. RNAs are then purified by phenol:chloroform:isoamylalcohol and precipitated with ethanol in the presence of sodium acetate 0.3 M at pH 5. RNA pellets are resuspended in milliQ water. RNAs are quantified by Qubit (Life Technologies) and their integrity is assessed with Bioanalyzer (Agilent Technologies). RNAs were treated to deplete abundant rRNAs, and the cDNA libraries are prepared using the ScriptSeq complete kit (bacteria) from Illumina.

2.6 RNA Sequencing Analysis

The libraries are sequenced with Illumina Mi-seq using a V4 chemistry sequencing kit (Illumina), which preserves the information about the orientation of transcripts and produces reads of 150 nt. The reads are then analyzed using Galaxy (Afgan et al., 2016). The workflow we have developed can be downloaded from the Supplementary Material in the online version at <https://doi.org/10.1016/bs.mie.2018.08.022> session and used directly in Galaxy. The utilities which are needed, available in most of the Galaxy instances, are the following: FastQC (Galaxy Version 0.62), Trimmomatic (Galaxy Version 0.32.3), Bowtie2 (Galaxy Version 2.2.6), Flagstat (Galaxy Version 1.0.1), HTSeq-count (Galaxy Version 0.6.1), and SARTools DEseq2 (Galaxy Version 0.99.2). Briefly, after having assessed the quality of the reads with FastQC, they are processed to remove adapter sequences

and bad quality reads using Trimmomatic (Bolger, Lohse, & Usadel, 2014). For this step, the parameters are given by (i) ILLUMINACLIP, which allows a maximum of two mismatches between the read and the TruSeq3 adapter sequences; (ii) SLIDINGWINDOW, which scans from the 5' end of the read and trims once the average Illumina quality score of 4 consecutive bases falls below the 20 threshold; (iii) MINLEN, which drops the reads shorter than 20 nucleotides. The reads are then aligned on the HG001 genome (GenBank accession number CP018205.1) using Bowtie2 (Langmead & Salzberg, 2012). The alignment is done using the option `very-sensitive-local`. In this mode, Bowtie2 uses a seed length of 20 with 0 mismatch allowed, and might “trim” or “clip” some read characters from one or both ends of the alignment if doing so maximizes the alignment score. In this way, reads belonging to MS2-sRNA can also be recovered. To obtain statistics on the alignment including the percentage of mapped reads we use Flagstat (Li et al., 2009). Finally, the number of reads mapping to each annotated feature is counted with HTSeq (Anders, Pyl, & Huber, 2015) using the interception nonempty protocol. To estimate the enrichment values for the MAPS experiment or the differential expression analysis for the transcriptomic experiment, we use DEseq2 (Varet, Brillet-Gueguen, Coppee, & Dillies, 2016). DEseq2 uses one of the most common approaches in comparative analysis of transcriptomics data. It assumes that the count data follow a negative binomial distribution testing the null hypothesis that the logarithmic fold change (LFC) between MS2-sRNA and MS2 for the recovering of a given transcript is exactly zero, i.e., that the target mRNA is not at all enriched by the purification via the sRNA. The statistical analysis process includes data normalization, graphical exploration of raw and normalized data, estimation of the dispersion of the data and test for differential enrichment for each feature between the MS2 and MS2-sRNA samples, raw *P*-value adjustment, and export of lists of features having a significant differential enrichment (threshold *P*-value = 0.05; fold change threshold = 2) between the conditions.

Since there is a great variation in the expression of the different transcripts purified with the MS2-sRNA bait, the rank obtained with the LFCs does not necessarily reflect the affinities between the sRNA and its targets. Thus, we also estimate the cellular abundance of each of the transcripts of the targetome. In combination with the enrichment values of the MAPS analysis, the *in vivo* affinities (designated as affinity rank) between sRNA and its targets can be then evaluated. Affinity rank is obtained by the ratios between

the counted fragments (normalized for the gene length and the deepness of the sequencing—“FPKM”) of the sRNA and those of each of mRNAs issued from the MAPS (Tomasini et al., 2017).

2.7 Validation of sRNA Target and Mechanism of Regulation

Putative target mRNAs identified by MAPS are subsequently validated using methods described in Fig. 2 (also reviewed by Jagodnik, Brosse, Le Lam, Chiaruttini, & Guillier, 2017).

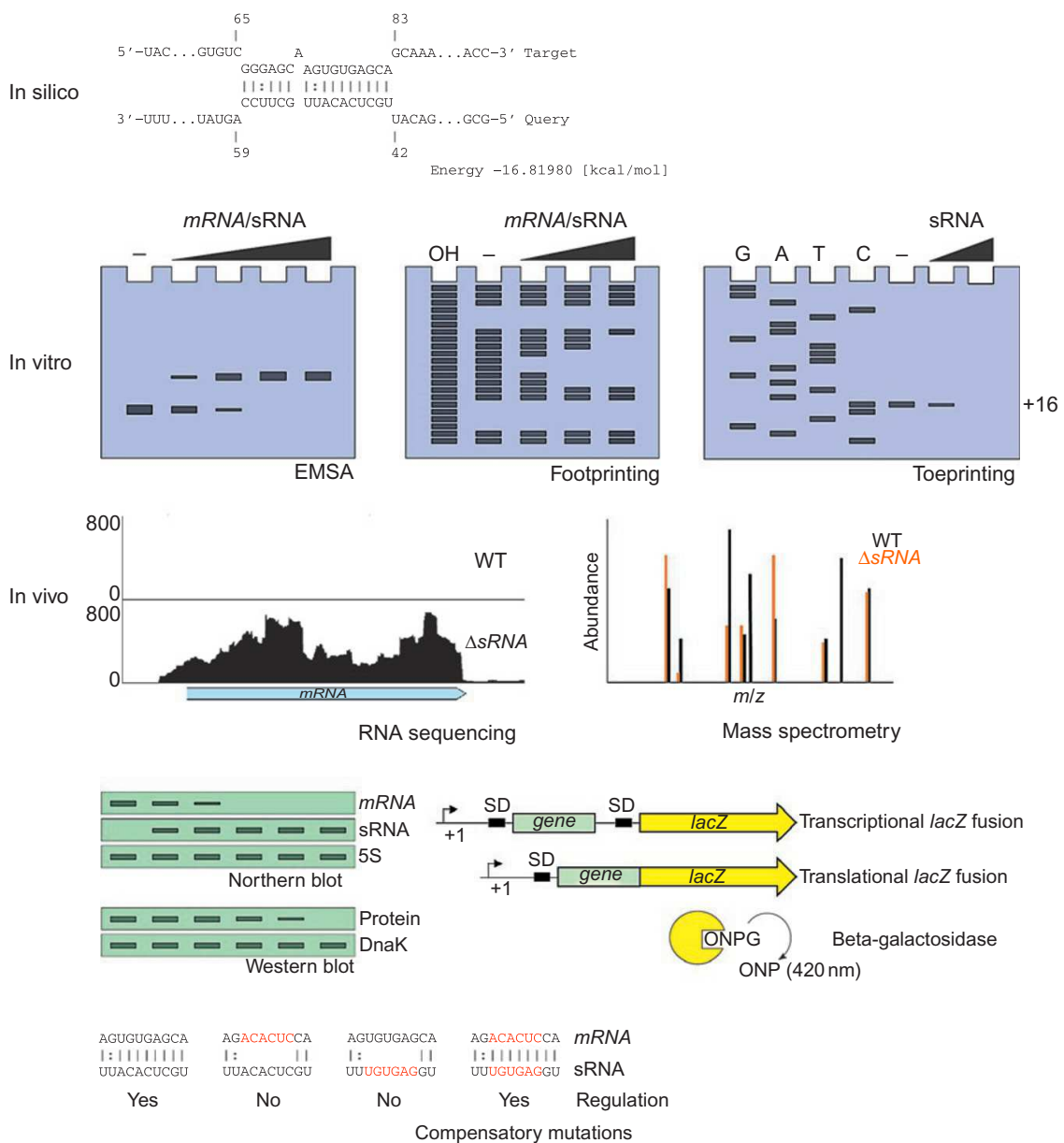


Fig. 2 Overview of methods commonly used to validate mRNA candidates identified by MAPS.

Gel retardation assays are first used to visualize direct interactions between RNAs. Briefly 5' end radiolabelled sRNA and cold mRNAs are renatured separately at room temperature. Then, increasing concentrations of cold mRNA are mixed with labeled sRNA in a buffer containing 20 mM Tris-HCl pH 7.5, 10 mM MgCl₂, and 60 mM KCl. Complex formation is carried out at 37°C for 15 min. After incubation, 10 µL of glycerol blue is added and the samples are loaded on a 6% polyacrylamide gel under nondenaturing conditions (5 h, 300 V, 4°C). To characterize sRNAs regulatory mechanism and notably their impact on target mRNA translation, we usually perform toeprinting assays to monitor the formation of a simplified initiation ribosomal complex (Fechter et al., 2009), but also in vitro translation and in vivo β-galactosidase assays (using *lacZ* translational fusion in cells expressing or not the sRNA of interest) (see Tomasini et al., 2017). Moreover, more precise interactions can be mapped by footprinting using enzymatic hydrolysis and/or chemical modification. To prove direct interaction beyond doubt, the seed region of both sRNA and target mRNA is punctually mutated. Both mutations must individually impair sRNA-dependent regulation, but they must restore sRNA activity together (compensatory mutations). This can be performed either in vitro or in vivo.

2.8 Frequently Asked Questions

2.8.1 How to Fuse the MS2 Aptamer to Your Favorite sRNA?

This is an essential question since to modify an sRNA is tantamount to walk in a minefield. Said et al. (2009) tried to add the MS2 tag at various positions, notably just before the native transcriptional terminator of the sRNA, which is often Rho-independent (a stem-loop followed by a poly-U tail). They even replaced this latter by an exogenous terminator, originating from *vrrA* sRNA gene in *Vibrio cholerae* (Song et al., 2008). Unfortunately, they noticed the transcription of RNA fragments of variable sizes due to aberrant processing and transcriptional read-through. We, and Massé's group, chose to add the MS2 sequence at the 5'-end of sRNA genes (Bronesky et al., 2018; Lalaouna, Carrier, et al., 2015; Lalaouna, Morissette, Carrier, & Masse, 2015; Lalaouna, Prevost, Laliberte, Houe, & Masse, 2018; Tomasini et al., 2017). In this case, a major risk mentioned by Said et al. (2009) is that the addition of MS2 tag could disrupt sRNA function when the seed region is located in its 5'-end (e.g., *E. coli* RybB sRNA). However, such a limit can be circumvented by the addition of a 6-nt long spacer between the MS2 aptamer and the +1 of sRNA (Lalaouna et al., 2017). This has been validated notably using RybB sRNA (Lalaouna, Carrier, et al., 2015).

It is also possible to add the MS2 tag in other regions of the sRNA, but it is important to mention that the regulatory RNAs are composed of distinct functional regions that must be preserved which are sometimes difficult to localize:

- (1) A seed region, which generally initiates the recognition with the target mRNAs by sequence complementarity. Note that several *S. aureus* sRNAs carry multiple seed sequences such as RsaA (Bronesky et al., 2018) or RNAIII (Bronesky et al., 2016). A behavior already observed for some *E. coli* sRNAs like DsrA (Lalaouna & Masse, 2016) and GcvB (Coornaert, Chiaruttini, Springer, & Guillier, 2013).
- (2) A possible binding site of an RNA chaperone. Even if no RNA chaperone proteins have been so far identified for sRNA regulation in *S. aureus*, we cannot exclude that specific regions could constitute a binding site for an unidentified protein. For Gram-negative (*E. coli* and *Salmonella*) the importance of RNA chaperone proteins has been amply demonstrated. This is the case for *E. coli* Hfq, which uses two anchorage points on the sRNAs (Schu, Zhang, Gottesman, & Storz, 2015). The involvement of Hfq in sRNA-dependent regulation is still controversial in *S. aureus* (Bohn et al., 2007; Liu et al., 2010) and no homologs of ProQ, another described RNA chaperone protein (Olejniczak & Storz, 2017; Smirnov, Schneider, Hor, & Vogel, 2017), have been identified.
- (3) Remaining sequences which may be involved in the secondary structure of the sRNA, which is a key determinant in its stability and functions.
- (4) The above-mentioned terminator, which could also be involved in RNA:RNA interactions (Susan Gottesman, 5th Meeting of Regulating with RNA in Bacteria & Archaea).

2.8.2 How to Validate MS2 Constructs?

Addition of a MS2 aptamer (43-nt long) to regulatory RNAs (ranging from 50 to 500 nucleotides) can generate certain risks notably concerning the impact of the nucleotides of the tag on sRNA folding. Thus, MS2-sRNA constructs must be carefully validated before subsequent applications. A crucial step is to determine whether the MS2 aptamer affects cellular level and the functions of chimeric sRNA (MS2-sRNA) compared to untagged sRNA (control). Classically, MS2-sRNA level is compared to sRNA control in vivo using Northern blot analysis. For its activity, even if informative, a simple in silico prediction of the sRNA folding in the presence and

the absence of the MS2 aptamer is not enough. If possible, at least one previously described target mRNA should be used to verify that MS2-sRNA is able to fulfill its regulatory role in a similar manner than untagged sRNA using Northern blot analysis. MS2-sRNA construct could be also validated using in vitro translation assays on specific target mRNAs (Tomasini et al., 2017).

2.8.3 Endogenous Expression or Overexpression of MS2 Constructs?

MS2-sRNA and cognate controls are generally overexpressed from plasmids to ensure their sufficient expression and increase the yield of MS2-affinity purification (Bronesky et al., 2018; Lalaouna et al., 2018; Lalaouna, Carrier, et al., 2015; Lalaouna, Morissette, et al., 2015; Said et al., 2009; Tomasini et al., 2017). For *S. aureus* the plasmid is a modified pCN51 (quorum sensing-dependent P3 promoter) with a copy number of 20–25 copies/cell (Bronesky et al., 2018; Charpentier et al., 2004; Tomasini et al., 2017). This is similar to the plasmid used in *S. Typhimurium*, pZ plasmid (constitutive PLlacO-1 promoter) with a copy number of 20–30 copies/cell (Lutz & Bujard, 1997; Said et al., 2009), or *E. coli*, pBAD24 plasmid (arabinose pBAD promoter) with a copy number of 30 copies/cell (Guzman, Belin, Carson, & Beckwith, 1995; Lalaouna, Carrier, et al., 2015). Although the induction from plasmids only slightly overexpressed RsaA sRNA (1.7-fold) or RyhB (2.7-fold) compared to their endogenous expression (Lalaouna, Carrier, et al., 2015; Tomasini et al., 2017), other *S. aureus* sRNAs are drastically overproduced (unpublished data). Therefore, identified mRNA candidates could be pure artifacts, for example, due to indirect copurification through associated proteins such as RNA binding proteins (e.g., Hfq or other RNA chaperones). Intriguingly, we and other noticed that some highly enriched mRNA candidates only respond to overexpressed sRNAs (unpublished data). Indeed, no significant effect was observed when a Δ sRNA strain is compared to WT. A possible explanation is that the overexpression of sRNAs may sequestered some RNA chaperone proteins and impair the stability and function of other RNA chaperone protein-dependent sRNAs (Moon & Gottesman, 2011).

To settle this problem, we could envisage to use chromosomally encoded MS2-sRNA constructs (under the control of native promoter). Said et al. (2009) successfully copurified Hfq protein using endogenously transcribed MS2-RybB. It remains to be determined whether this approach enables the copurification of target mRNAs in detectable amounts.

2.8.4 Constitutive or Pulsed Expression?

The advantage of pulse expression compared to constitutive expression is to restrict sRNA activity in a narrow window of time.

- (1) In Gram-negative bacteria, many sRNAs negatively regulate their target mRNAs at the posttranscriptional level by recruiting ribonucleases such as RNase E and thus actively inducing mRNA decay (Lalaouna, Simoneau-Roy, Lafontaine, & Masse, 2013; Mohanty & Kushner, 2016). If sRNA genes are constitutively expressed, the level of many, if not most, target mRNAs will drastically drop. In *S. aureus*, there is no homolog of RNase E. Other RNases such as RNase III, but also RNase Y, J1, and J2, might be involved in sRNA-mediated mRNA decay (Durand, Tomasini, Braun, Condon, & Romby, 2015). To slow down sRNA-dependent mRNA decay in *E. coli*, an RNA degradosome assembly mutant (*rne131*) of RNase E is used (Lalaouna et al., 2017). The use of an *rne131* strain significantly increases the enrichment ratio of negatively regulated targets, as confirmed with MS2-DsrA (WT *rne* vs *rne131*; D. Lalaouna, unpublished data). In *S. aureus*, we work in a wild-type background, which may principally shed light on stable interactions and translation regulation instead of regulation by targeting the mRNA stability. The mutation of the nonessential RNase III could be envisioned (Lioliou et al., 2012).
- (2) sRNA synthesis all along the cell growth is detrimental, enabling sRNAs to regulate target mRNAs in nonphysiological conditions, but also to meet new transcripts (e.g., no more separation in time, exponential vs stationary phase of growth). Most sRNAs are solely produced under specific conditions (e.g., iron starvation, host invasion, osmotic stress). Even transcriptional noise could be damaging for bacterial cells which use sRNAs sponge to avoid it (Lalaouna, Carrier, et al., 2015).

2.8.5 Which Genetic Background Should Be Used?

To optimize MS2-affinity purification, the endogenous sRNA gene should be deleted to avoid competition for targets.

2.8.6 What Are Suitable Controls for MS2-Affinity Purification?

Two main controls can be used to validate specific enrichment of mRNA candidates: a sRNA without MS2 tag (untagged sRNA) or a MS2 tag alone.

- (1) Untagged sRNA control enables to deduct false positives resulting from unspecific interactions with the column or agarose beads. Theoretically no RNA should be purified in the presence of untagged sRNA.

Unfortunately, numerous RNAs are eluted such as ribosomal RNAs (which could represent more than 90% of copurified RNAs). Accordingly, an additional step of ribosomal depletion (Ribo-Zero rRNA Removal Kit (Bacteria), Illumina) has been added to the original protocol established by [Lalaouna, Carrier, et al. \(2015\)](#) (see [Bronesky et al., 2018](#); [Tomasini et al., 2017](#)).

- (2) MS2 tag control expressed in the parental strain enables to remove unspecific RNAs interacting with the column, agarose beads, and most importantly with the MS2 aptamer. To note, the choice of the transcriptional terminator is primordial. [Said et al. \(2009\)](#) used a transcriptional terminator originating from *V. cholerae* to lower the probability to be recognized in *S. Typhimurium*. We, and Massé's group, used sRNA gene terminators. For example, both MS2-*RsaI* and MS2 tag constructs share the same transcriptional terminator ([Tomasini et al., 2017](#)).

Massé's group mainly used an enrichment ratio corresponding to MS2-sRNA/untagged sRNA. In addition to eliminate unspecifically bound RNAs, this enables a comparable sRNA-induced regulation in both samples. In [Lalaouna, Morissette, et al. \(2015\)](#), they also published a list of contaminants that bound to the MS2 aptamer in tested conditions (ratio pBAD-MS2/empty vector).

For *S. aureus*, we use an enrichment ratio corresponding to MS2-sRNA/MS2 tag. Since the MS2 tag alone is expressed in the wild-type background, the untagged sRNA is also present. Nevertheless, as we discussed above, the level of endogenous sRNA and plasmid expressed MS2-sRNA might differ. We also extracted and sequenced total RNAs before the MS2-affinity purification. The comparison of both datasets enables to determine relative level of sRNA and mRNA target candidates before and after purification.

2.8.7 How Reliable are MAPS Results?

Massé's group commonly performs MAPS only once. For our part, we perform MAPS in duplicate, adding statistical analysis, and thus robustness. However, we all agree that MAPS results are more like snapshot views. Any variation in medium, harvesting time, or even experimentalist will influence results.

2.8.8 Are Highly Enriched mRNAs Always Direct Targets?

As previously described, some of mRNA candidates could be artifacts due to unspecific interactions with the column or the MS2 aptamer. Moreover, mRNAs could be copurified by indirect interactions, notably through

RNA binding proteins or RNAs having the same targets. Hence, every mRNA candidate must be validated using in vivo or in vitro approaches (Fig. 2). Prediction of basepairing interactions might also be useful to distinguish primary from unspecific target RNAs. We often used CopraRNA and/or IntaRNA (Wright et al., 2014).

2.8.9 Is It Possible to Infer the Prioritization of Target RNAs From MAPS Data?

Performing MAPS can potentially generate a long list of RNA candidates sorted according to fold change (or enrichment ratio). Hence, we could imagine that sRNAs have a greater affinity for highly enriched target RNAs. However, it is crucial to obtain information on their relative or absolute levels before MS2-affinity purification. For example, a poorly transcribed RNA which is highly enriched will be recognized with a higher affinity than a highly transcribed one. For this purpose, we performed total RNA sequencing simultaneously with MAPS (Bronesky et al., 2018; Tomasini et al., 2017).

ACKNOWLEDGMENTS

We thank Pascale Romby for helpful advice and discussions. We are grateful to Joseph Vilardell for the gift of the plasmid expressing the MS2-MBP. RNA-Seq analyses have been done using the Roscoff (France) instance of Galaxy (<http://galaxy.sb-roscoff.fr/>).

FUNDING

This work was supported by the Agence Nationale de la Recherche (ANR, Grant ANR-16-CE11-0007-01, RIBOSTAPH, to Pascale Romby [P.R.]). It has also 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. D.L. was supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 753137—SaRNAReg.

REFERENCES

- Abu-Qatouseh, L. F., Chinni, S. V., Seggewiss, J., Proctor, R. A., Brosius, J., Rozhdestvensky, T. S., 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. *Journal of Molecular Medicine*, 88(6), 565–575. <https://doi.org/10.1007/s00109-010-0597-2>.
- Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Cech, M., et al. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 Update. *Nucleic Acids Research*, 44(W1), W3–W10. <https://doi.org/10.1093/nar/gkw343>.

- Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31(2), 166–169. <https://doi.org/10.1093/bioinformatics/btu638>.
- Beaume, M., Hernandez, D., Farinelli, L., Deluen, C., Linder, P., Gaspin, C., et al. (2010). Cartography of methicillin-resistant *S. aureus* transcripts: Detection, orientation and temporal expression during growth phase and stress conditions. *PLoS One*, 5(5), e10725. <https://doi.org/10.1371/journal.pone.0010725>.
- Bohn, C., Rigoulay, C., & Boulloc, P. (2007). No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. *BMC Microbiology*, 7, 10. <https://doi.org/10.1186/1471-2180-7-10>.
- Bohn, C., Rigoulay, C., Chabelskaya, S., Sharma, C. M., Marchais, A., Skorski, P., et al. (2010). Experimental discovery of small RNAs in *Staphylococcus aureus* reveals a riboregulator of central metabolism. *Nucleic Acids Research*, 38(19), 6620–6636. <https://doi.org/10.1093/nar/gkq462>.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Bronesky, D., Desgranges, E., Corvaglia, A., François, P., Caballero, C. J., Prado, L., et al. (2018). A dual sRNA in *Staphylococcus aureus* induces a metabolic switch responding to glucose consumption. *bioRxiv*. <https://doi.org/10.1101/278127>.
- Bronesky, D., Wu, Z., Marzi, S., Walter, P., Geissmann, T., Moreau, K., et al. (2016). *Staphylococcus aureus* RNAIII and its regulon link quorum sensing, stress responses, metabolic adaptation, and regulation of virulence gene expression. *Annual Review of Microbiology*, 70, 299–316. <https://doi.org/10.1146/annurev-micro-102215-095708>.
- Caldelari, I., Chane-Woon-Ming, B., Noirot, C., Moreau, K., Romby, P., Gaspin, C., et al. (2017). Complete genome sequence and annotation of the *Staphylococcus aureus* strain HG001. *Genome Announcements*, 5(32). <https://doi.org/10.1128/genomeA.00783-17>.
- Charpentier, E., Anton, A. I., Barry, P., Alfonso, B., Fang, Y., & Novick, R. P. (2004). Novel cassette-based shuttle vector system for Gram-positive bacteria. *Applied and Environmental Microbiology*, 70(10), 6076–6085. <https://doi.org/10.1128/aem.70.10.6076-6085.2004>.
- Coornaert, A., Chiaruttini, C., Springer, M., & Guillier, M. (2013). Post-transcriptional control of the *Escherichia coli* PhoQ-PhoP two-component system by multiple sRNAs involves a novel pairing region of GcvB. *PLoS Genetics*, 9(1), e1003156. <https://doi.org/10.1371/journal.pgen.1003156>.
- Dale, S. E., Doherty-Kirby, A., Lajoie, G., & Heinrichs, D. E. (2004). Role of siderophore biosynthesis in virulence of *Staphylococcus aureus*: Identification and characterization of genes involved in production of a siderophore. *Infection and Immunity*, 72(1), 29–37.
- Durand, S., Tomasini, A., Braun, F., Condon, C., & Romby, P. (2015). sRNA and mRNA turnover in Gram-positive bacteria. *FEMS Microbiology Reviews*, 39(3), 316–330. <https://doi.org/10.1093/femsre/fuv007>.
- Eyraud, A., Tattevin, P., Chabelskaya, S., & Felden, B. (2014). A small RNA controls a protein regulator involved in antibiotic resistance in *Staphylococcus aureus*. *Nucleic Acids Research*, 42(8), 4892–4905. <https://doi.org/10.1093/nar/gku149>.
- Fechter, P., Chevalier, C., Yusupova, G., Yusupov, M., Romby, P., & Marzi, S. (2009). Ribosomal initiation complexes probed by toeprinting and effect of trans-acting translational regulators in bacteria. *Methods in Molecular Biology*, 540, 247–263. https://doi.org/10.1007/978-1-59745-558-9_18.
- Geissmann, T., Chevalier, C., Cros, M. J., Boisset, S., Fechter, P., Noirot, C., et al. (2009). A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved

- sequence motif for regulation. *Nucleic Acids Research*, 37(21), 7239–7257. <https://doi.org/10.1093/nar/gkp668>.
- Goldmann, O., & Medina, E. (2017). *Staphylococcus aureus* strategies to evade the host acquired immune response. *International Journal of Medical Microbiology*. <https://doi.org/10.1016/j.ijmm.2017.09.013>.
- Guzman, L. M., Belin, D., Carson, M. J., & Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of Bacteriology*, 177(14), 4121–4130.
- Howden, B. P., Beaume, M., Harrison, P. F., Hernandez, D., Schrenzel, J., Seemann, T., et al. (2013). Analysis of the small RNA transcriptional response in multidrug-resistant *Staphylococcus aureus* after antimicrobial exposure. *Antimicrobial Agents and Chemotherapy*, 57(8), 3864–3874. <https://doi.org/10.1128/aac.00263-13>.
- Jagodnik, J., Brosse, A., Le Lam, T. N., Chiaruttini, C., & Guillier, M. (2017). Mechanistic study of base-pairing small regulatory RNAs in bacteria. *Methods*, 117, 67–76. <https://doi.org/10.1016/j.ymeth.2016.09.012>.
- Kavita, K., de Mets, F., & Gottesman, S. (2018). New aspects of RNA-based regulation by Hfq and its partner sRNAs. *Current Opinion in Microbiology*, 42, 53–61. <https://doi.org/10.1016/j.mib.2017.10.014>.
- Lalaouna, D., Carrier, M. C., Semsey, S., Brouard, J. S., Wang, J., Wade, J. T., et al. (2015). A 3' external transcribed spacer in a tRNA transcript acts as a sponge for small RNAs to prevent transcriptional noise. *Molecular Cell*, 58(3), 393–405. <https://doi.org/10.1016/j.molcel.2015.03.013>.
- Lalaouna, D., & Masse, E. (2016). The spectrum of activity of the small RNA DsrA: Not so narrow after all. *Current Genetics*, 62(2), 261–264. <https://doi.org/10.1007/s00294-015-0533-7>.
- Lalaouna, D., Morissette, A., Carrier, M. C., & Masse, E. (2015). DsrA regulatory RNA represses both hns and rbsD mRNAs through distinct mechanisms in *Escherichia coli*. *Molecular Microbiology*, 98(2), 357–369. <https://doi.org/10.1111/mmi.13129>.
- Lalaouna, D., Prevost, K., Eyraud, A., & Masse, E. (2017). Identification of unknown RNA partners using MAPS. *Methods*, 117, 28–34. <https://doi.org/10.1016/j.ymeth.2016.11.011>.
- Lalaouna, D., Prevost, K., Laliberte, G., Houe, V., & Masse, E. (2018). Contrasting silencing mechanisms of the same target mRNA by two regulatory RNAs in *Escherichia coli*. *Nucleic Acids Research*, 46(5), 2600–2612. <https://doi.org/10.1093/nar/gkx1287>.
- Lalaouna, D., Simoneau-Roy, M., Lafontaine, D., & Masse, E. (2013). Regulatory RNAs and target mRNA decay in prokaryotes. *Biochimica et Biophysica Acta*, 1829(6–7), 742–747. <https://doi.org/10.1016/j.bbagrmm.2013.02.013>.
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>.
- Le Pabic, H., Germain-Amiot, N., Bordeau, V., & Felden, B. (2015). A bacterial regulatory RNA attenuates virulence, spread and human host cell phagocytosis. *Nucleic Acids Research*, 43(19), 9232–9248. <https://doi.org/10.1093/nar/gkv783>.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- Lioliou, E., Sharma, C. M., Caldelari, I., Helfer, A. C., Fechter, P., Vandenesch, F., et al. (2012). Global regulatory functions of the *Staphylococcus aureus* endoribonuclease III in gene expression. *PLoS Genetics*, 8(6), e1002782. <https://doi.org/10.1371/journal.pgen.1002782>.
- Liu, Y., Wu, N., Dong, J., Gao, Y., Zhang, X., Mu, C., et al. (2010). Hfq is a global regulator that controls the pathogenicity of *Staphylococcus aureus*. *PLoS One*, 5(9). <https://doi.org/10.1371/journal.pone.0013069>.

- Lutz, R., & Bujard, H. (1997). Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Research*, 25(6), 1203–1210.
- Macias, S., Bragulat, M., Tardiff, D. F., & Vilardell, J. (2008). L30 binds the nascent RPL30 transcript to repress U2 snRNP recruitment. *Molecular Cell*, 30(6), 732–742. <https://doi.org/10.1016/j.molcel.2008.05.002>.
- Manna, A. C., Kim, S., Cengher, L., Corvaglia, A., Leo, S., Francois, P., et al. (2018). Small RNA teg49 is derived from a sarA transcript and regulates virulence genes independent of SarA in *Staphylococcus aureus*. *Infection and Immunity*, 86(2). <https://doi.org/10.1128/iai.00635-17>.
- Melamed, S., Peer, A., Faigenbaum-Romm, R., Gatt, Y. E., Reiss, N., Bar, A., et al. (2016). Global mapping of small RNA-target interactions in Bacteria. *Molecular Cell*, 63(5), 884–897. <https://doi.org/10.1016/j.molcel.2016.07.026>.
- Mohanty, B. K., & Kushner, S. R. (2016). Regulation of mRNA decay in bacteria. *Annual Review of Microbiology*, 70, 25–44. <https://doi.org/10.1146/annurev-micro-091014-104515>.
- Monk, I. R., Shah, I. M., Xu, M., Tan, M. W., & Foster, T. J. (2012). Transforming the untransformable: Application of direct transformation to manipulate genetically *Staphylococcus aureus* and *Staphylococcus epidermidis*. *MBio*, 3(2), 1–11. <https://doi.org/10.1128/mBio.00277-11>.
- Moon, K., & Gottesman, S. (2011). Competition among Hfq-binding small RNAs in *Escherichia coli*. *Molecular Microbiology*, 82(6), 1545–1562. <https://doi.org/10.1111/j.1365-2958.2011.07907.x>.
- Nitzan, M., Fechter, P., Peer, A., Altuvia, Y., Bronesky, D., Vandenesch, F., et al. (2015). A defense-offense multi-layered regulatory switch in a pathogenic bacterium. *Nucleic Acids Research*, 43(3), 1357–1369. <https://doi.org/10.1093/nar/gkv001>.
- Novick, R. P., & Geisinger, E. (2008). Quorum sensing in staphylococci. *Annual Review of Genetics*, 42, 541–564. <https://doi.org/10.1146/annurev.genet.42.110807.091640>.
- 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. *The EMBO Journal*, 12(10), 3967–3975.
- Olejniczak, M., & Storz, G. (2017). ProQ/FinO-domain proteins: Another ubiquitous family of RNA matchmakers? *Molecular Microbiology*, 104(6), 905–915. <https://doi.org/10.1111/mmi.13679>.
- Oliveira, D., Borges, A., & Simoes, M. (2018). *Staphylococcus aureus* toxins and their molecular activity in infectious diseases. *Toxins (Basel)*, 10(6). <https://doi.org/10.3390/toxins10060252>.
- Penades, J. R., Chen, J., Quiles-Puchalt, N., Carpena, N., & Novick, R. P. (2015). Bacteriophage-mediated spread of bacterial virulence genes. *Current Opinion in Microbiology*, 23, 171–178. <https://doi.org/10.1016/j.mib.2014.11.019>.
- Pichon, C., & 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(40), 14249–14254. <https://doi.org/10.1073/pnas.0503838102>.
- Qin, L., McCausland, J. W., Cheung, G. Y., & Otto, M. (2016). PSM-Mec-A virulence determinant that connects transcriptional regulation, virulence, and antibiotic resistance in staphylococci. *Frontiers in Microbiology*, 7, 1293. <https://doi.org/10.3389/fmicb.2016.01293>.
- Quereda, J. J., & Cossart, P. (2017). Regulating bacterial virulence with RNA. *Annual Review of Microbiology*, 71, 263–280. <https://doi.org/10.1146/annurev-micro-030117-020335>.

- Romilly, C., Lays, C., Tomasini, A., Caldelari, I., Benito, Y., Hammann, P., et al. (2014). A non-coding RNA promotes bacterial persistence and decreases virulence by regulating a regulator in *Staphylococcus aureus*. *PLoS Pathogens*, 10(3), e1003979. <https://doi.org/10.1371/journal.ppat.1003979>.
- Said, N., Rieder, R., Hurwitz, R., Deckert, J., Urlaub, H., & Vogel, J. (2009). In vivo expression and purification of aptamer-tagged small RNA regulators. *Nucleic Acids Research*, 37(20), e133. <https://doi.org/10.1093/nar/gkp719>.
- Saliba, A. E., C Santos, S., & Vogel, J. (2017). New RNA-seq approaches for the study of bacterial pathogens. *Current Opinion in Microbiology*, 35, 78–87. <https://doi.org/10.1016/j.mib.2017.01.001>.
- Santiago-Frangos, A., & Woodson, S. A. (2018). Hfq chaperone brings speed dating to bacterial sRNA. *Wiley Interdisciplinary Reviews. Ribonucleic Acid*, 9(4), e1475. <https://doi.org/10.1002/wrna.1475>.
- Schu, D. J., Zhang, A., Gottesman, S., & Storz, G. (2015). Alternative Hfq-sRNA interaction modes dictate alternative mRNA recognition. *The EMBO Journal*, 34(20), 2557–2573. <https://doi.org/10.15252/embj.201591569>.
- Smirnov, A., Schneider, C., Hor, J., & Vogel, J. (2017). Discovery of new RNA classes and global RNA-binding proteins. *Current Opinion in Microbiology*, 39, 152–160. <https://doi.org/10.1016/j.mib.2017.11.016>.
- Song, T., Mika, F., Lindmark, B., Liu, Z., Schild, S., Bishop, A., et al. (2008). A new *Vibrio cholerae* sRNA modulates colonization and affects release of outer membrane vesicles. *Molecular Microbiology*, 70(1), 100–111. <https://doi.org/10.1111/j.1365-2958.2008.06392.x>.
- Tomasini, A., Francois, P., Howden, B. P., Fechter, P., Romby, P., & Caldelari, I. (2014). The importance of regulatory RNAs in *Staphylococcus aureus*. *Infection, Genetics and Evolution*, 21, 616–626. <https://doi.org/10.1016/j.meegid.2013.11.016>.
- Tomasini, A., Moreau, K., Chicher, J., Geissmann, T., Vandenesch, F., Romby, P., et al. (2017). The RNA targetome of *Staphylococcus aureus* non-coding RNA RsaA: Impact on cell surface properties and defense mechanisms. *Nucleic Acids Research*, 45(11), 6746–6760. <https://doi.org/10.1093/nar/gkx219>.
- Vanderpool, C. K., Balasubramanian, D., & Lloyd, C. R. (2011). Dual-function RNA regulators in bacteria. *Biochimie*, 93(11), 1943–1949. <https://doi.org/10.1016/j.biochi.2011.07.016>.
- Varet, H., Brillet-Gueguen, L., Coppee, J. Y., & Dillies, M. A. (2016). SARTools: A DESeq2- and EdgeR-based R pipeline for comprehensive differential analysis of RNA-Seq data. *PLoS One*, 11(6), e0157022. <https://doi.org/10.1371/journal.pone.0157022>.
- Wagner, E. G., & Romby, P. (2015). Small RNAs in bacteria and archaea: Who they are, what they do, and how they do it. *Advances in Genetics*, 90, 133–208. <https://doi.org/10.1016/bs.adgen.2015.05.001>.
- Wright, P. R., Georg, J., Mann, M., Sorescu, D. A., Richter, A. S., Lott, S., et al. (2014). CopraRNA and IntaRNA: Predicting small RNA targets, networks and interaction domains. *Nucleic Acids Research*, 42(Web Server issue), W119–W123. <https://doi.org/10.1093/nar/gku359>.

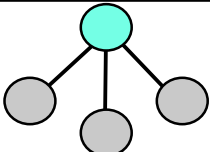
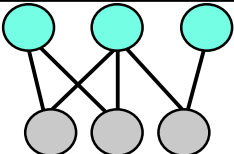
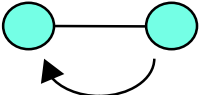
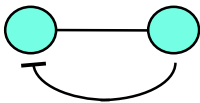
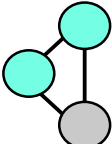
Circuits	Diagrammes	Avantages
<i>Single-input module</i>		• Coordonne la régulation de plusieurs cibles
<i>Overlapping regulon</i>		• Intègre des signaux multiples
<i>Positive feedback loop</i>		• Amplifie le signal
<i>negative feedback loop</i>		• Entraîne une réponse rapide
<i>Feedforward loop</i>		• Modifie la dynamique de régulation des cibles • Intègre des signaux multiples

Figure 13 : Exemples de circuits de régulation retrouvés dans les réseaux de régulation bactériens : Les régulateurs sont représentés en bleu et les cibles en gris. Inspiré de Beisel et al., 2010.

4.6. Description des réseaux

Les différentes techniques présentées par ces deux revues successives apportent les clés pour établir les cibles des ARN régulateurs. La caractérisation des cibles des sARN est nécessaire pour définir leurs fonctions. Ainsi, ces études permettent de placer les sARN dans des réseaux de régulation impliquant d'autres régulateurs, les TF et les TCS. Cette vue globale de la régulation permet de mieux appréhender la physiologie des bactéries. De plus, la connaissance des conditions d'expression de ces différents régulateurs permet de mieux appréhender la dynamique des réseaux et les stratégies pour adapter leur physiologie en fonction des conditions. En effet, en réponse à un stimulus, la transduction du signal à travers les réseaux de régulation va conduire à la coordination de l'expression des gènes pour permettre l'adaptation à ce stimulus (Beisel & Storz, 2010). Les réseaux de régulation sont caractérisés par plusieurs circuits de régulation. Un circuit de régulation est composé de nœuds, qui correspondent à des gènes, et de lignes qui connectent ces nœuds et qui correspondent à la régulation exercée entre les gènes. Ces circuits peuvent être de divers types et leur utilisation présente des avantages différents (Figure 13). Par exemple, certains circuits permettent d'induire une réponse suite à l'intégration de plusieurs stimulus, d'autres coordonnent l'expression de plusieurs gènes en réponse à un même stimuli ou induisent un effet sur la dynamique de la réponse, en modifiant sa vitesse ou sa force (Beisel & Storz, 2010).

Dans le cas où un réseau fait intervenir différents types de régulateurs, ces circuits de régulations sont dits mixtes. Dans le cas d'un circuit appelé « *feedforward loop* » mixte soit le sARN, soit le facteur de transcription peut être au sommet de la régulation. La régulation des gènes *spa* (codant le facteur d'adhésion, protéine A) et *hla* (codant l'hémolysine alpha) par l'ARNIII effecteur intracellulaire de la densité cellulaire, est un exemple type de « *feedforward loop* » mixte où le sARN est au sommet de la cascade de régulation. Ce réseau est approprié pour changer le comportement d'une population de cellules, et dans ce cas, pour permettre le passage de *S. aureus* d'un mode d'adhésion à un mode de dissémination (Nitzan et al., 2015). Le réseau de l'ARNIII présenté en figure 14 fait intervenir la combinaison de deux « *feedforward loop* » mixte FFL1 et FFL2. L'augmentation de la densité cellulaire se traduit par une augmentation progressive de l'ARNIII (*cf* I.3.2.2.1 et I.4.3). Dans la FFL1, l'inhibition de *spa* par l'ARNIII s'effectue au niveau post-transcriptionnel de manière directe et au niveau transcriptionnel de manière indirecte via l'inhibition du FT Rot. Dans la FFL2, cette configuration permet une activation directe de *hla* par l'ARNIII au niveau traductionnel, et indirecte au niveau transcriptionnel par l'inhibition de Rot. L'association des deux FFL

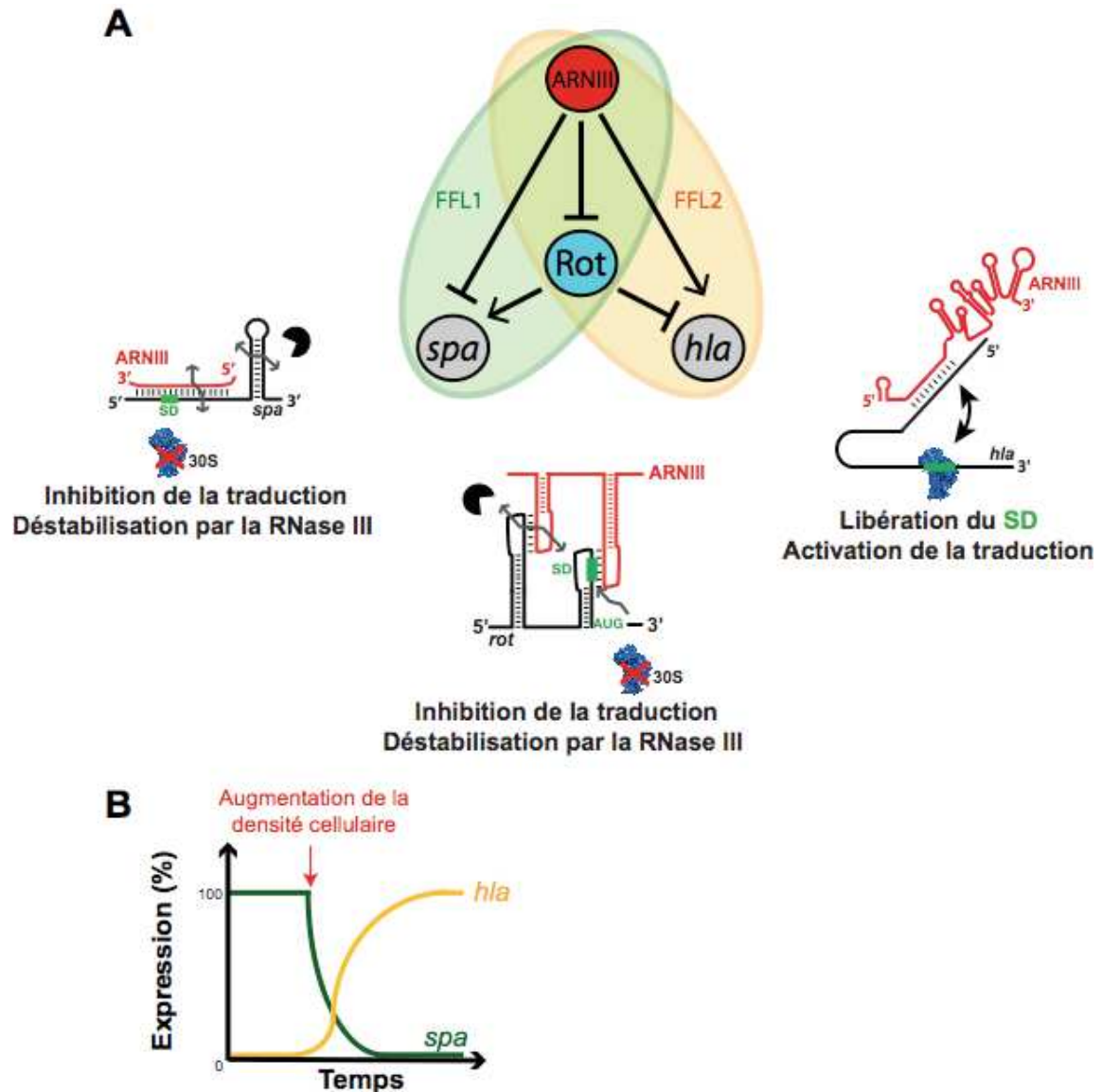


Figure 14 : Exemple du double circuit « *feedforward loop* » mixte faisant intervenir l'ARNIII : A) L'ARNIII réprime la traduction du FT Rot en bloquant la séquence Shine et Dalgarno (SD) et le codon d'initiation (AUG) de la traduction de *rot*, cette interaction entraîne la déstabilisation du messager suite au recrutement de la RNaseIII. Rot régule au niveau transcriptionnel *spa* et *hla*. L'ARNIII est également capable de réguler ces messagers au niveau post-transcriptionnel. L'ARNIII inhibe la traduction de *spa* en se fixant à la séquence SD et entraîne la déstabilisation du messager par la RNaseIII. À l'inverse l'ARNIII active la traduction de *hla* en libérant la séquence SD. L'ARNIII et Rot interviennent donc dans deux « *feedforward loop* » mixtes FFL1 et FFL2. B) Le double circuit « *feedforward loop* » contrôle l'expression inverse de *spa* et *hla* en réponse à une augmentation de la densité cellulaire. Un changement d'expression transitoire de ces deux ARNm est observé. Inspiré de Bronesky et al., 2016 ; Nitzan et al., 2017.

coordonne la régulation asymétrique de ces deux cibles. De plus, cela garantit une régulation fine et transitoire car la régulation indirecte de *spa* et *hla* par l'ARNIII via *rot* entraînera un délai dans la réponse, dû au temps nécessaire pour éliminer le pool de Rot. Ainsi, la transition entre l'arrêt de l'expression de *spa* et le début de l'expression de *hla* suit une dynamique progressive, permettant un passage coordonné de la population bactérienne d'un phénotype d'adhésion à un phénotype de dissémination (Nitzan et al., 2015, 2017; figure 14). Ce type de circuit a également été identifié dans d'autres systèmes biologiques. Par exemple, chez *E. coli*, un double circuit « *feedforward loop* » mixte permet, en fonction des nutriments disponibles, d'orienter les bactéries vers un mode de vie sessile en biofilm ou vers une mobilité des cellules, via la production de flagelles. Cela montre que ce type de circuit est couramment utilisé pour opérer un changement de style de vie dans les populations bactériennes (Nitzan et al., 2015b). Au vu des propriétés apportées par chaque type de circuit de régulation, connaître la place des sARN dans ces circuits permet de mieux comprendre leur impact sur la réponse cellulaire en raison des changements dans l'environnement.

4.7. Mécanisme de régulation et de communication en dehors de la cellule par les ARN : Les OMV

Comme les cellules eucaryotes et les archées, les bactéries sécrètent des vésicules membranaires dans leur environnement extracellulaire. Ce mécanisme est conservé au sein de tous les domaines du vivant. Les vésicules extracellulaires bactériennes sont des bicouches de protéolipides sphériques contenant des protéines, des lipides, des acides nucléiques et des facteurs de virulence. Chez les bactéries à Gram négatif, ces vésicules sont appelées OMV pour « *outer membrane vesicles* » et ont été observées pour la première fois dans les années 60. Ces vésicules sont originaires de la membrane externe de l'enveloppe bactérienne. A l'inverse chez les bactéries à Gram positif, ne portant pas de membrane externe, ce type de vésicules n'a été découvert que récemment et leur mode de biogenèse n'a pas encore été établi (Lee et al., 2009).

Ce système de sécrétion fait partie des stratégies de communication entre bactéries, mais aussi de communication entre l'hôte et le pathogène. Un exemple est la sécrétion de vésicules extracellulaires par *S. aureus* qui est corrélée avec l'apparition d'eczéma. En effet, la présence d' α -hémolysine dans les vésicules sécrétées par *S. aureus*, entraîne, après fusion des vésicules avec des kératinocytes, leur apoptose ainsi qu'une inflammation, créant des dommages de la barrière cutanée (Hong et al., 2014).

De manière intéressante ces vésicules peuvent aussi contenir des ARN. Le séquençage des ARN contenus dans les vésicules de *Vibrio cholerae* a révélé que les ARN les plus

abondants correspondent à des régions non codantes (Sjöström et al., 2015), bien que leurs fonctions demeurent peu comprises.

Dans la mesure où les vésicules libérées par une bactérie peuvent être reprises par d'autres bactéries ou par des cellules eucaryotes, une hypothèse logique est de penser que les ARN associés aux OMV peuvent avoir des effets régulateurs dans leur nouvelle cellule hôte. C'est effectivement ce qui a été montré chez *Pseudomonas aeruginosa*. L'infection par *P. aeruginosa* stimule la sécrétion d'IL-8. En revanche, le sRNA52320 présent dans les OMV produites par ce pathogène atténue cette sécrétion, ce qui diminue la réponse immunitaire de l'hôte vis-à-vis de *P. aeruginosa*, en limitant le recrutement des neutrophiles (Koeppen et al., 2016, pour revue Guerrero-Mandujano et al., 2017). Enfin, il a été montré que le pathogène intracellulaire *Listeria monocytogenes* sécrète des ARN dans le cytoplasme de ses cellules hôtes, par le biais ou non de vésicules, appelés secRNome (Frantz et al., 2019). L'un de ces sARN, rli32, contribue à induire des hauts taux d'interférons de type I via la régulation de RIG-I, ce qui offre un environnement propice à la réplication de *L. monocytogenes* et à son passage de cellule en cellule pour une infection efficace (Frantz et al., 2019). Cette nouvelle fonction des sARN, qui est d'agir dans des cellules eucaryotes via leur sécrétion par les OMV, fait naturellement penser aux virus qui expriment des miARN afin de manipuler les cellules qu'ils infectent (revue Girardi et al., 2018).

4.8. Espoir thérapeutique contre la résistance aux antibiotiques grâce à l'ARN

Comme traité précédemment, l'utilisation d'antibiotiques entraîne l'apparition de souches qui y sont résistantes (cf I.1.3). De plus, la plupart des antibiotiques couramment utilisés sont à spectre large et vont donc entraîner dans un même temps l'élimination des bactéries ciblées, mais également du microbiote. L'altération du microbiote a des effets néfastes sur la santé humaine puisqu'il contribue au système immunitaire de son hôte, aide à la digestion de la nourriture, permet la production de certaines vitamines, empêche la colonisation des pathogènes par compétition, et confère bien d'autres avantages (Ubeda et al., 2017; Wang et al., 2017).

C'est pour cela que des stratégies plus spécifiques ont été mises en œuvre. L'une de ces approches utilise des « *short antisens oligonucleotides* » (ASOs) afin de contrer l'apparition des résistances aux antibiotiques, tout en limitant la perturbation du microbiote. Les ASOs consistent en un oligonucléotide de 10-12 bases antisens d'un ARNm cible, qui correspond à un gène essentiel de la bactérie ciblée. Afin d'être spécifique d'une espèce, l'ARNm cible doit

être spécifique de cette espèce ou bien encore l'oligonucléotide ASO devra reconnaître une région non conservée d'un ARNm conservé entre différentes espèces. Les ASOs sont couplés à un peptide pour permettre leurs transports dans les bactéries. Ce type d'antibiotique, basé sur l'utilisation d'ARN, est prometteur et a déjà fait ses preuves dans des essais sur des modèles murins (Vogel, 2020).

L'étude de Yanagihara et al. (2006) s'inscrit dans le même genre d'effort de l'utilisation d'ARN pour lutter contre les bactéries pathogènes. Ils ont montré, *in vitro*, en utilisant une souche de *S. aureus* SARM, que sa pathogénicité était atténuée grâce à la transfection de petits ARN interférants (siARN). Les siARN utilisés ciblent l'ARNm *coa* codant pour la staphylocoagulase, un important facteur de virulence de *S. aureus*.

Les régions 5'UTR des ARNm peuvent contenir des riboswitches, ils reconnaissent des métabolites (cofacteurs, vitamines, acides aminés...) et des seconds messagers, ce qui induit un changement structural conduisant à une modification de l'expression de l'ARNm en aval. Une grande partie des riboswitches contrôlent l'expression de gènes impliqués dans des voies métaboliques. Les riboswitches régulant des gènes essentiels à la croissance, pourraient ainsi être utilisés pour le développement d'antibiothérapie alternative (Deigan & Ferré-D'Amaré, 2011). L'exemple de deux molécules identifiées comme ciblant des riboswitches de *S. aureus* est donné dans la revue Desgranges et al. (2019) (*cf* I.4.3).

Les ASO, siRNA et riboswitches pourraient donc être une nouvelle option thérapeutique pour traiter des infections aux souches bactériennes résistantes aux antibiotiques.

5. Les infections à *S. aureus*

Les types d'infections produites par *S. aureus* sont vastes. Dans le but de mieux les traiter, outre de nouvelles stratégies thérapeutiques, il est également important d'étudier comment elles s'établissent et de comprendre les mécanismes qui permettent à *S. aureus* de favoriser un mode de vie commensal ou pathogène. Ceci implique de caractériser les régulateurs impliqués afin d'imaginer de nouveaux traitements plus efficaces que les antibiotiques et en étant plus spécifiques du type d'infection.

5.1. Les infections extracellulaires

Il existe plusieurs modes de colonisation d'un individu par *S. aureus*. Dans le cas des porteurs sains de *S. aureus*, la colonisation est passive et n'entraîne pas immédiatement de

maladies c'est le cas par exemple des portages nasaux. Il est cependant observé que la proportion de ces individus à être infectés est augmentée selon la souche qui les colonise. Par ailleurs, la colonisation de *S. aureus* au niveau de la peau et des muqueuses peut être à l'origine d'une infection lorsqu'une brèche survient. La disruption des barrières, que sont la peau et les muqueuses, permet le passage des bactéries dans les tissus adjacents, donnant lieu à une réaction inflammatoire entraînant par exemple la formation d'abcès. Les bactéries peuvent également se propager dans la circulation sanguine, ce qui entraîne des infections sévères au niveau de sites secondaires, par exemple des endocardites. Pour les colonisations entraînant une infection, le fait que celle-ci soit contenue ou se disperse, va dépendre des facteurs de virulence exprimés par *S. aureus* et des réponses immunitaires mises en place par l'hôte pour les contrer (Ferens & Bohach, 2000; Lowy, 1998). Cependant, *S. aureus* est également capable de jouer avec notre système immunitaire. La production de superantigène par *S. aureus* bouleverse le fonctionnement du système immunitaire de l'hôte entraînant une forte létalité suite au développement du syndrome de choc toxique responsable du dysfonctionnement de plusieurs organes (Kulhankova et al., 2014). Les entérotoxines produites par *S. aureus* présentent une activité superantigénique. Lorsque de la nourriture contaminée avec une quantité suffisante d'entérotoxines est ingérée, les entérotoxines peuvent être responsables d'un empoisonnement alimentaire, menant à des nausées, des vomissements et des diarrhées (Argudín et al., 2010).

Enfin un autre moyen pour *S. aureus* d'établir une infection, est rendu possible grâce à l'utilisation de dispositifs médicaux, tels que les cathéters, indispensables pour le traitement de certaines maladies. Les cathéters, lors de leur mise en place sont rapidement recouverts de protéines de matrice de l'hôte, rendant cette surface favorable pour l'attachement des bactéries qui ont passé la barrière de la peau au moment de la pose du cathéter. *S. aureus* est ensuite capable de former un biofilm. Une fois le biofilm mature, ou suite à un choc mécanique ou chimique, les cellules qui se sont nouvellement divisées vont se détacher entraînant leur dissémination dans la circulation sanguine. Ceci résulte en une infection du système sanguin (Esposito et al., 2013).

5.2. Les infections intracellulaires

5.2.1. Généralités

Pendant les premiers temps de l'infection, *S. aureus* adhère aux cellules épithéliales grâce à l'expression d'adhésines, ce qui donne lieu à l'apparition des infections extracellulaires précédemment décrites, qui sont majoritaires. Cependant, après l'adhésion *S. aureus* peut être

phagocyté par les cellules endothéliales. *S. aureus* a traditionnellement été considéré comme un pathogène extracellulaire. Cependant, au cours du temps, il a été rapporté comme pouvant envahir et survivre à l'intérieur de divers types cellulaires : des cellules non immunitaires (cellules épithéliales, myoblastes, ostéoblastes...) et des cellules immunitaires (neutrophiles, macrophages, lymphocytes T...). À la vue de ces observations, *S. aureus* apparaît donc comme un pathogène intracellulaire facultatif (Hamza & Li, 2014). Les bactéries internalisées pourront se répliquer et induire la lyse des cellules hôtes afin de se disséminer à d'autres cellules ou persister dans leur cellule hôte. Les infections avec internalisation sont difficiles à traiter, car une fois internalisé, *S. aureus* est protégé des défenses immunitaires de l'hôte et des effets des antibiotiques. Cependant, l'incapacité des antibiotiques à éradiquer les *S. aureus* intracellulaires ne peut pas pleinement être attribuée à la faible pénétration des antibiotiques (Rowe et al., 2020 ; cf I. 5.2.6). L'adaptation, menant à une prolifération intracellulaire active ou à la persistance, est encore faiblement comprise. L'internalisation dépend du type de cellule envahies, et de l'état physiologique de chacun des acteurs. Par ailleurs, il est clair qu'il existe une hétérogénéité dans la population bactérienne, et qu'une fraction est dormante pendant que l'autre est répliquative (Palma Medina et al., 2019).

5.2.2. Adaptation métabolique

Une fois la bactérie internalisée, l'hôte et le pathogène doivent s'adapter à cette nouvelle situation, et rentrent en compétition pour les ressources disponibles. Cela entraîne une adaptation de la part des deux parties. Cette adaptation évolue au cours du temps car dans un premier temps les bactéries sont enfermées dans des vésicules et se divisent activement les 12 premières heures (Michalik et al., 2017). Puis, elles commencent à exprimer et sécréter un ensemble de toxines entraînant la formation de pores (HlgB, HlgC, Hla, LukH et LukG) dans la membrane des vésicules, permettant aux bactéries de s'échapper et de se retrouver libres dans le cytosol de la cellule hôte. Cette relocalisation entraîne une compétition directe entre les bactéries pour les nutriments, entraînant un impact significatif sur le métabolisme de la cellule hôte, de même que sur les bactéries qui voient leur croissance devenir très lente après 12 h d'internalisation (Michalik et al., 2017). De plus, les bactéries présentent un changement métabolique en faveur de la fermentation après 24 h d'internalisation dans cet environnement micro aérobie qu'est la cellule. Cependant, une culture en conditions de micro aérobie ne permet pas de résumer l'environnement intracellulaire car celui-ci comporte des sucres qui lui sont propres. En condition aérobie, Rex est le FT impliqué dans l'inhibition de la

fermentation (*cf* I.3.1.2.2.1) via l'inhibition de plusieurs lactates déshydrogénases dont *ldh1* (Pagels et al., 2010). Cependant, Rex est inactif dans des conditions de micro aérobie. En condition de micro aérobie, dans un milieu contenant du glucose, la protéine Ldh1 est activée par CcpA en raison de la présence de glucose, mais aussi suite à l'inactivation de Rex par le potentiel rédox (Christmas et al., 2019; Crooke et al., 2013, *cf* I.3.1.3). Par ailleurs, en conditions d'internalisation, bien que les bactéries optent pour un métabolisme fermentaire, *ldh1* pourtant déréprimée par Rex n'est pas activée due à l'absence de glucose dans le milieu intracellulaire (Michalik et al., 2017 ; Figure 15). Cette absence de glucose redirige le métabolisme de la bactérie vers les nutriments disponibles dans le cytosol de la cellule hôte, comme par exemple les acides aminés. Dès le début de l'internalisation, les acides aminés sont utilisés comme source alternative de carbone et d'azote. En effet, la dégradation des acides aminés permet la production de pyruvate qui génère une source d'énergie, mais également d'ammoniaque qui permet d'augmenter le pH de l'environnement intracellulaire. L'augmentation du pH empêche la fusion des endosomes avec les lysosomes, processus qui serait létal pour les bactéries (Palma Medina et al., 2019).

5.2.3. Régulation de l'expression des transporteurs : le cas de UhpT

A l'image du protéome, le transcriptome lui aussi change en réponse à l'internalisation puisque 40% des gènes voient leur expression modifiée. Pendant les premières heures de l'internalisation (2 à 6 h), il a été montré que l'expression d'un grand nombre de transporteurs est augmentée. C'est le cas de l'ARNm *uhpT* qui code pour un transporteur du G6P. Cette activation est explicable car le glucose est phosphorylé après son transport dans les cellules eucaryotes, formant ainsi un pool de G6P senti par HptRS (Garzoni et al., 2007; Park et al., 2015). *uhpT* est régulé par le système à deux composants HptRS (*cf* I.3.2.2.2) qui a été montré comme jouant un rôle important dans l'adaptation de *S. aureus* au milieu intracellulaire. En effet, un mutant de délétion de ce système présente une absence de réplication et un pourcentage de survie largement diminué dans divers types cellulaires (macrophages, monocytes, cellules épithéliales de poumon) (Park et al., 2015). UhpT a par ailleurs été montré dans d'autres bactéries intracellulaires telles que *Shigella flexneri* et *Listeria monocytogenes* comme étant fortement exprimé en conditions d'internalisation. Chez *S. flexneri*, c'est le gène le plus induit en condition intracellulaire en comparaison à une culture liquide. De manière intéressante, une souche de *S. flexneri* mutante de *uhpT* se multiplie et se propage de la même façon qu'une souche sauvage, signifiant que d'autres sources de carbone que le G6P sont disponibles dans le

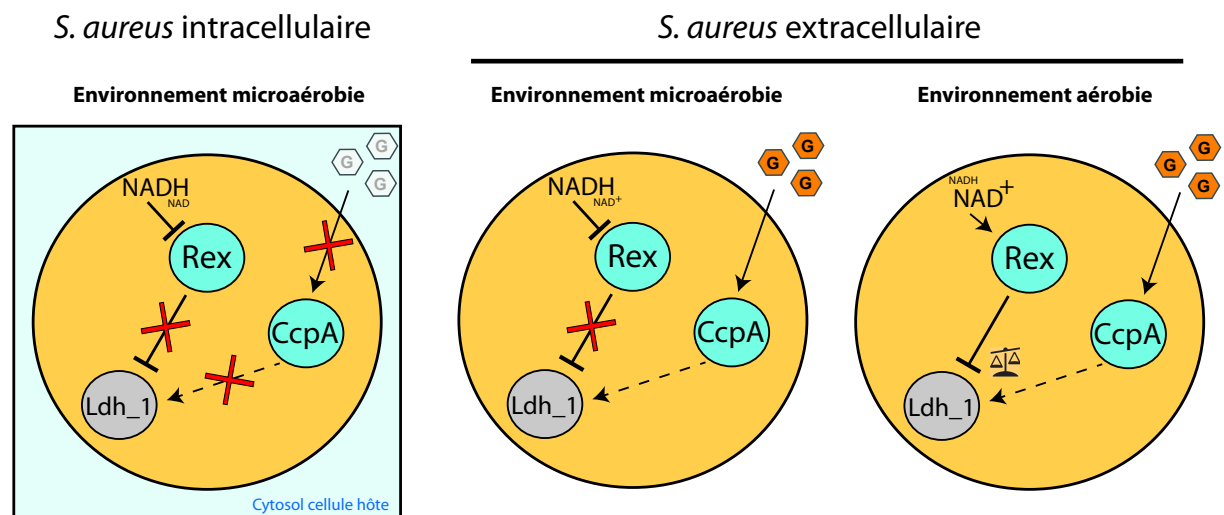


Figure 15 : Régulation de la lactate déshydrogénase Ldh_1 en fonction de l'environnement de *S. aureus* : Les FT Rex et CcpA sont représentés en bleu et leur expression respective est régulée par le ratio NADH/NAD⁺ et la présence de glucose, ce qui influence l'expression de Ldh_1 en gris. À gauche le réseau de régulation représente *S. aureus* internalisé dans le cytosol d'une cellule hôte, où l'anaérobiose entraîne une augmentation du NADH, tandis que le glucose est absent. À droite en condition extracellulaire avec ou sans oxygène en présence de glucose.

cytosol de l'hôte (Runyen-Janecky & Payne, 2002). Ce phénomène n'est pas si clair chez *S. aureus* dans la mesure où, comme expliqué précédemment, un mutant du TCS HptRS activateur de *uhpT* voit sa survie intracellulaire diminuée (Park et al., 2015). Nous pouvons émettre deux hypothèses. Premièrement, le métabolisme de *S. aureus* ne permet pas l'utilisation d'autres sources de carbone que le G6P. Deuxièmement, HptRS ne régulerait pas seulement *uhpT*, et c'est l'absence de régulation des autres gènes qui entraîne un défaut de survie. Nous pouvons aussi émettre l'hypothèse que le G6P n'est qu'un signal (et non une source de carbone) indiquant aux bactéries qu'elles sont internalisées, et qu'une adaptation est nécessaire pour leur survie. Ceci est corrélé avec l'observation de Palma Medina et al. (2019) qui montre que les protéines impliquées dans la glycolyse ne présentent pas un niveau d'expression supérieur en internalisation, comme si l'apport de G6P (qui normalement rentre dans la glycolyse dès la deuxième étape) par le cytosol de la cellule hôte n'était pas utilisé dans cette voie métabolique. De manière intéressante, chez *L. monocytogenes*, *uhpT* (HPT) est régulé par le régulateur PrfA qui induit l'expression d'un grand nombre de facteurs de virulence, dès l'entrée de la bactérie dans une cellule hôte. L'absence de *uhpT* rend *L. monocytogenes* incapable de proliférer intracellulairement et le rend moins virulent dans un modèle murin d'infection. En contrepartie, la délétion de *prfA* rend impossible l'utilisation d'hexoses-phosphates (G1P, G6P, F6P, M6P) qui sont des sucres principalement retrouvés dans l'environnement cytosolique. Chez *L. monocytogenes* il existe donc un lien très clair entre l'établissement de la virulence et la présence d'hexose-phosphate. Cette observation va dans le sens de l'hypothèse selon laquelle la présence de sucres phosphorylés est un signal pour indiquer aux bactéries qu'elles sont internalisées (Chico-Calero et al., 2002).

A la vue de ce qui se passe chez d'autres bactéries, bien que le rôle du G6P ne soit pas encore bien décrit dans le contexte de l'internalisation chez *S. aureus*, il constituerait probablement un stimulus clef dans l'adaptation à ce milieu.

5.2.4. sARN et internalisation

A la vue de la fonction clef des sARN dans les mécanismes d'adaptation au stress, il était donc évident que ceux-ci puissent exercer une action dans le processus d'internalisation. L'ARN régulateur SprC permet le maintien du commensalisme car son expression induit un désavantage pour l'entrée des bactéries dans l'environnement intracellulaire. Lorsque *S. aureus* n'exprime pas cet ARN, son taux d'internalisation est en effet augmenté (Le Pabic et al., 2015). RsaA est un autre sARN qui agit comme un atténuateur de la virulence, en rendant les bactéries

plus sensibles à l'opsonophagocytose par les polynucléaires neutrophiles (Romilly et al., 2014). Cet ARN contribue au commensalisme et aux interactions saprophytes avec l'hôte. Il reste à déterminer si d'autres sARN participent à maintenir l'internalisation de *S. aureus* dans les cellules hôtes.

5.2.5. Émergence des SCV

Une autre caractéristique des infections de *S. aureus* avec internalisation est l'apparition de « *small-colony variants* » (SCV). Ces variants ne sont pas particulièrement virulents mais peuvent persister durablement dans les cellules de l'hôte. Pendant des dizaines d'années, le patient reste asymptomatique (Proctor et al., 1998). Ce réservoir caché de *S. aureus* est bien évidemment un problème car il laisse à penser que l'infection a été traitée, alors que malheureusement elle pourra se déclarer à nouveau des années plus tard. Les SCV s'adaptent à l'environnement intracellulaire en arborant une activité métabolique réduite. Ce phénotype est rendu possible par l'apparition de mutations dans des gènes du métabolisme, entraînant l'apparition d'auxotrophies. Ce sont les conditions environnementales, telles que les traitements antibiotiques prolongés, qui sélectionnent l'apparition de SCV. Par ailleurs, l'environnement intracellulaire rend les SCV plus résistants aux antibiotiques. La réduction de leur métabolisme leur permet également de diminuer l'entrée des antibiotiques, à cause de la diminution de la production d'ATP et du gradient de proton au niveau de leur membrane (Melter & Radojevič, 2010).

5.2.6. Émergence des bactéries persistantes ou persisteurs

L'utilisation d'antibiotiques entraîne l'apparition de bactéries persistantes. Les « persisteurs » sont une sous-catégorie de cellules au sein de la population bactérienne, qui ne se répliquent pas et résistent à des concentrations mortelles d'antibiotiques. Les persisteurs adoptent ce phénotype de manière transitoire, lorsque les antibiotiques sont présents dans l'environnement. Il existe également des persisteurs intracellulaires comme l'a récemment montré l'étude de Peyrusson et al. (2020). L'étude du transcriptome des persisteurs intracellulaires a révélé qu'ils modifient profondément leur métabolisme, entraînant une diminution de l'abondance de transcrits codant pour des protéines impliquées dans la glycolyse et la fermentation. Les gènes codant pour des enzymes du cycle de Krebs ne semblent pas quant à eux être affectés dans ces conditions. Ce « *switch* » dans le métabolisme entraîne une croissance lente avec absence de réplication, mais un maintien de la production d'énergie

nécessaire pour la mise en place des mécanismes de défense et de réponses aux stress (Peyrusson et al., 2020). L'étude de Wang et al. (2018) allait également dans ce sens en montrant que la formation de persisteurs ne corrélait pas avec une modification du niveau cellulaire d'ATP. En revanche, ils auraient établi que l'inhibition de certaines enzymes du cycle de Krebs (par exemple les sous-unités SucA et SucB de l' α -cétoglutarate) augmente considérablement l'émergence des persisteurs, en affectant directement le potentiel de la membrane et la force protomotrice (Wang et al., 2018). Enfin, il a aussi été montré que les ROS peuvent induire la tolérance à de multiples antibiotiques. La persistance induite par les ROS survient *in vivo* lorsque *S. aureus* est internalisé par des macrophages, mais que ces derniers ne sont pas efficaces pour l'éradiquer (Rowe et al., 2020). Les macrophages, en essayant de tuer *S. aureus* internalisé sans y parvenir, vont induire une persistance aux antibiotiques de cette population bactérienne par la production de ROS. Les ROS inactivent le cycle de Krebs, ce qui impacterait la respiration et la concentration intracellulaire d'ATP, permettant ainsi la tolérance aux antibiotiques (Rowe et al., 2020).

Les études des paramètres nécessaires à l'établissement des persisteurs, et des conséquences physiologiques sur cette population bactérienne, demandent donc encore à être mieux définies.

6. Objectifs de la thèse

Au cours de ma thèse je me suis focalisée à élucider la fonction de deux sARN, RsaI et RsaG. L'étude de RsaI avait été débutée au laboratoire par Delphine Bronesky. Ses expériences ont montré que cet sARN est exprimé seulement dans un milieu carencé en glucose, suite à la levée de sa répression par CcpA. Le MAPS a permis d'identifier des cibles ARN de RsaI impliquées dans le métabolisme du glucose. Dans la première partie de ma thèse, j'ai participé à valider les cibles *in vivo* afin de mieux comprendre les fonctions biologiques régulées par RsaI. En revanche, pour RsaG, j'ai déterminé ses conditions d'expression, étudié sa biogénèse, analysé sa conservation au sein des *staphylocoques*, et déterminé sa fonction biologique par la découverte de ces cibles. L'étude de ces deux sARN a permis de les intégrer dans des réseaux de régulation complexes faisant intervenir des FT et des TCS qui ont permis de mieux mettre en lumière les liens existant entre le métabolisme et la virulence chez *S. aureus*.

Mon deuxième projet, que j'ai débuté tard, consistait à adapter une technique dérivée du LIGR-seq (pour « *ligation of interacting RNA and high-throughput sequencing* », mise en place pour les cellules eucaryotes) (Sharma et al., 2016) chez les bactéries. Cette technique permet



Figure 16 Structure secondaire de RsaI : Les séquences en rouge représentent les régions riches en G, les séquences encadrées représentent les régions conservées chez les staphylocoques. La structure a été prédite en utilisant RNAfold. À partir de Geissmann et al., 2009.

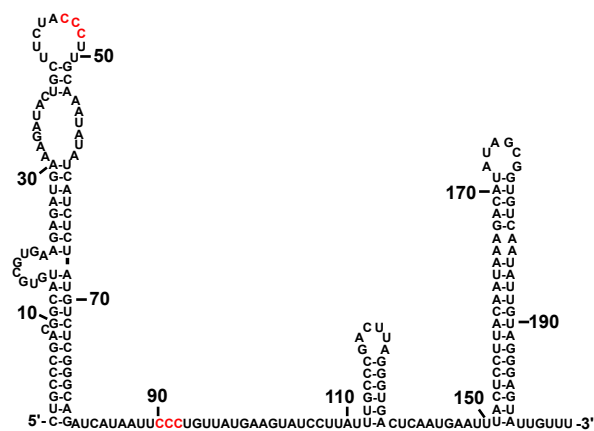


Figure 17 : Structure secondaire de RsaG : Les séquences en rouge représentent les régions riches en C accessibles. La structure a été identifiée par cartographie en solution de l'ARN. À partir de Geissmann et al., 2009.

d'établir de manière globale l'ensemble des interactions ARN-ARN *in vivo*. Dans un premier temps, l'étude individuelle des sARN RsaG et RsaI a servi de preuve de concept lors de l'élaboration du protocole. A plus long terme, une fois la technique mise au point, elle sera appliquée dans diverses conditions de stress (stress oxydatif, anaérobie, antibiotiques...) ou divers types d'environnements rencontrés par *S. aureus* au cours du processus infectieux (modèles cellulaires, modèles animaux...). Cette méthode puissante permettra d'établir de manière globale la place des sARN de *S. aureus* dans les réseaux de régulation, dans le but de mieux comprendre la physiologie de ce pathogène opportuniste.

6.1. Découverte des sARN dans le génome de *S. aureus*

Depuis la découverte de l'ARNIII par Richard Novick en 1993, *S. aureus* a émergé comme organisme modèle pour l'étude des sARN. Mon équipe a participé à la découverte d'un grand nombre de sARN de *S. aureus*. Parmi eux, nous retrouvons une classe de 11 sARN nommés RsaA à RsaK pour « *RNA from S. aureus* ». Ces ARN ont été sélectionnés car ils étaient présents dans des IGR conservés dans plusieurs *staphylocoques* ou spécifiques de *S. aureus*, et qu'ils comportaient un terminateur rho-indépendant, ainsi que des structures ou séquences particulières. Ils contiennent tous, sauf RsaI (cf I.6.1.1), au moins un motif riche en cytosines qui est défini comme une signature des sARN interagissant avec la séquence SD des ARNm cibles. Ces 11 ARN présentaient en général une expression augmentée en phase stationnaire de croissance (Geissmann et al., 2009).

6.1.1. Caractérisation de RsaI

RsaI est un ARN de 144 nucléotides. Il possède deux régions riches en G situées dans des boucles (en rouge, Figure 16) ainsi qu'une région riche en U/C située dans une longue région en simple brin. Ces séquences sont conservées au sein de tous les *staphylocoques* (encadré, Figure 16). Les résultats présentés dans Geissmann et al. (2009) indiquaient que RsaI s'accumulait en fin de phase exponentielle dans un milieu riche, cependant dans un milieu carencé, il était exprimé de manière constitutive (Bronsky et al., 2019). De plus, son expression était augmentée sous diverses conditions de stress (stress oxydatif, variation de température, pH acide) (Geissmann et al., 2009) ainsi qu'en présence de vancomycine (Howden et al., 2013). Tant de conditions d'expression intéressante qui ont initié la caractérisation de cet ARN régulateur.

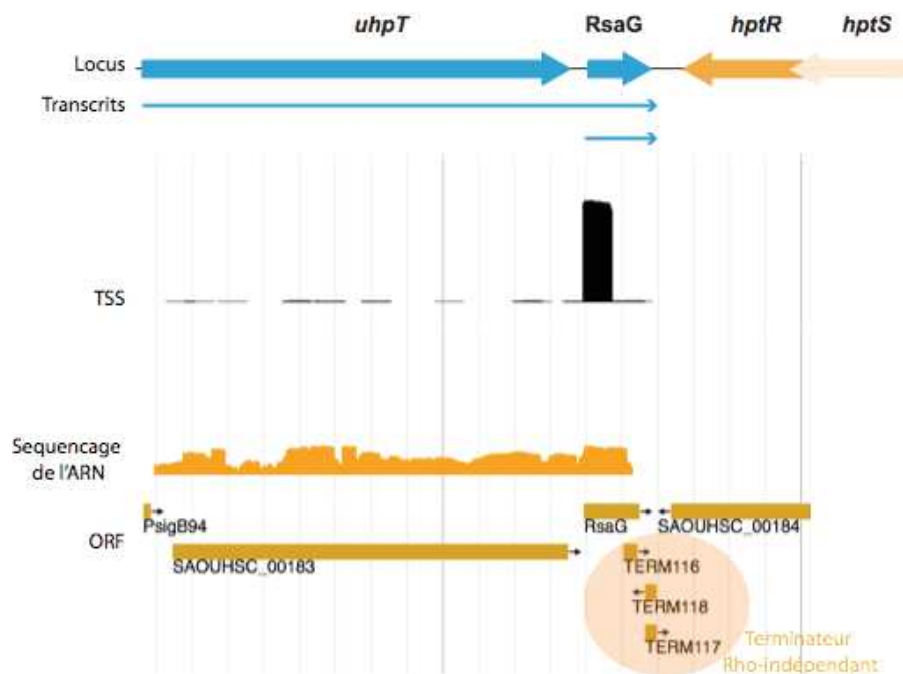


Figure 18 : Organisation du locus *uhpT-rsaG* : Est représenté le locus *uhpT-rsaG* transcrit de manière divergente au TCS *hptRS*. En dessous sont représentés de haut en bas : la localisation du site de démarrage de la transcription (TSS) de *RsaG* ; les lectures correspondantes au séquençage de l'ARN ; les différents cadres ouverts de lecture (ORF) présents dans le locus, ainsi que le promoteur (*PsigB94*) et les terminateurs (TERM) Rho-indépendants prédits. Inspiré de <http://staph.unavarra.es/>.

6.1.2. Caractérisation de RsaG

RsaG est un ARN long de 194 nucléotides, il possède deux régions riches en C accessibles, l'une dans une boucle apicale et l'autre dans une région interhélicoïdale en simple brin (en rouge, Figure 17). RsaG est exprimé dès l'entrée en phase stationnaire de croissance dans un milieu de culture riche. Cependant, dans un milieu pauvre (NZM), son expression n'a pas été détectée par Northern blot ou dans des conditions diverses de stress (Geissmann et al., 2009).

RsaG a été découvert dans une IGR, région qui correspond à toutes les séquences présentes en dehors des séquences codantes des gènes ainsi que leurs régions flanquantes non traduites 5'UTR et 3'UTR (Geissmann et al., 2009). Comme le montre la Figure 18, RsaG est localisé en 3' du gène *uhpT* codant pour un transporteur de G6P, et de manière divergente au système à deux composants HptRS qui répond à la concentration extracellulaire de G6P (Yang et al., 2015). Des données de séquençage de l'ARN (RNA-seq) ont suggéré que RsaG pourrait être compris dans la région 3'UTR de *uhpT* (Figure 18 et Lasa et al., 2011; Ruiz de los Mozos et al., 2013). En effet, les régions séquencées couvrent le transcrit *uhpT* ainsi que la région génomique de RsaG. Aucun terminateur de transcription n'est décrit en aval de la séquence codante de *uhpT*, alors que deux terminateurs Rho-indépendants sont prédits en aval de RsaG (le TERM116 étant celui de RsaG et constitue sa troisième et dernière tige boucle, cf Figure 17). De plus, cette analyse montre qu'il existe un site de démarrage de la transcription (TSS) dans la région 3'UTR de *uhpT*. Cette étude ne permet pas de distinguer les transcrits primaires des transcrits secondaires, ainsi nous ne pouvons pas conclure si RsaG est généré par la présence d'un promoteur interne dans la région 3'UTR de *uhpT*, ou suite à une maturation d'un ARN contenant *uhpT* et RsaG (Figure 18).

La localisation particulière de cet ARN, ainsi que sa conservation restreinte à *S. aureus* constituaient les raisons qui ont motivé son étude approfondie.

6.2. Développement d'une technique d'étude de l'interactome ARN-ARN chez *S. aureus* : DBRI

Récemment, le développement de méthodes globales a permis de définir les cibles directes de plusieurs sARN de *S. aureus*. Ces méthodes sont malheureusement spécifiques d'un sARN à la fois, et requièrent souvent l'addition d'une étiquette à l'ARN d'intérêt. De plus, les outils qui se basent sur la pêche de couples ARN-ARN liés à une protéine telle que le RIL-seq ou CLASH (Melamed et al., 2016; Waters et al., 2017) n'est pas applicable à *S. aureus*, puisqu'à

ce jour aucune protéine facilitant les interactions ARN-ARN n'a été montrée comme indispensable chez *S. aureus*. La seule protéine à ce jour impliquée dans la régulation par les ARN est l'endoribonucléase III. De nombreux ARN, dont des sARN, ont été identifiés comme partenaires de la RNase III par la technique de RIP-Seq (Lioliou et al., 2012).

Dans le but de construire les réseaux de régulation de *S. aureus* et leur dynamique, DBRI pour « *Direct determination of bacterial RNA-RNA intercatome* » est une optimisation de techniques déjà existantes et utilisées chez les eucaryotes et/ou les bactéries, telles que le LigR (Sharma et al., 2016, Durand et al., 2019 (bioRxiv)), le SPLASH (Aw et al., 2016), le PARIS (Lu et al., 2016) ou le « *modified CLASH* » (Liu et al., 2017). Toutes ces méthodes se basent sur le pontage aux UV des liaisons ARN-ARN grâce au psoralène ou l'un de ses dérivés, suivie d'une ligation de proximité permettant la formation de chimères entre les ARN en interaction, et enfin le séquençage à haut débit de ces ARN hybrides. Bien que l'analyse des données soit complexe, elle devrait nous conduire à mieux définir les réseaux faisant intervenir des sARN, et à déterminer quelle mesure ces réseaux sont modifiés en fonction des conditions de culture.

II. Résultats

1. Chapitre 1 : RsaI, un ARN régulateur aux multiples facettes, module le métabolisme de *S. aureus*

RsaI, un ARN régulateur aux multiples facettes, module le métabolisme du pathogène opportuniste *Staphylococcus aureus*

Emma Desgranges¹, Delphine Bronesky^{1*}, Anna Corvaglia², Patrice François², Carlos Caballero³, Laura Prado³, Alejandro Toledo-Arana³, Inigo Lasa⁴, Karen Moreau⁵, François Vandenesch⁵, Stefano Marzi¹, Pascale Romby¹, Isabelle Caldelari¹

¹Université de Strasbourg, CNRS, Architecture et réactivité de l'ARN, UPR9002, F-67000 Strasbourg, France.

²Laboratoire de recherche en génomique, département de spécialités médicales, Genève hôpitaux universitaires, université de Genève, Genève, Suisse.

³Instituto de agrobiotecnología. CSIC-UPNA-GN, 31192-Mutilva, Navarra, Espagne.

⁴Navarrabiomed, universidad pública de Navarra, departamento de salud, IDISNA, Pamplona, Espagne.

⁵CIRI, centre international de recherche en infectiologie, Inserm U1111, université Lyon1, école normale supérieure de Lyon, CNRS UMR5308, F-69008 Lyon, France.

*Adresse actuelle : ARNA, U1212, UMR 5320, université de Bordeaux, 146 rue Leo Saignat, Carreire zone nord Bat 3A 1^{er}, BP12, 33076 Bordeaux, France.

i.caldelari@unistra.fr



► *Staphylococcus aureus* est une bactérie commensale retrouvée chez environ 30 % des individus sains dont elle colonise la peau et la muqueuse nasale. Cependant, c'est également une bactérie pathogène opportuniste responsable d'infections diverses telles que orgelet, ostéomyélite, endocardite, ou encore septicémie en envahissant un grand nombre de tissus et d'organes. Cette bactérie est capable de s'adapter à des conditions hostiles et variées, telles que carence nutritive et stress osmotique, oxydant, ou thermique, ainsi qu'à la réponse immunitaire de l'hôte, car elle produit une grande diversité de facteurs de virulence. La synthèse de ces facteurs est finement régulée par des protéines et des ARN régulateurs majoritairement non codants, souvent désignés par l'abréviation sARN (dérivée de l'anglais, *small RNA*). Les facteurs de transcription et les systèmes à deux composants contrôlent l'expression des gènes impliqués non seulement dans le métabolisme, mais aussi dans la réponse au stress et la virulence [1]. Par exemple, la protéine du contrôle catabolique (*carbon catabolite control protein A*, CcpA) a un rôle essentiel dans le choix de la source carbonée en régulant le métabolisme central de la bactérie ainsi que la virulence [2, 3]. CcpA se fixe à une séquence promotrice spécifique appelée *cre* (*catabolite-responsive element*),

qui est très conservée chez les bactéries à Gram positif [2]. Quant aux sARN, ils interagissent principalement avec leurs ARN messagers (ARNm) cibles. L'hybridation peut conduire à la stabilisation/déstabilisation de l'ARNm ou à l'activation/répression de sa traduction [4]. Nous avons montré que la transcription du sARN RsaI (*RNA Staphylococcus aureus I*) est réprimée par CcpA en présence de glucose [5]. L'induction de la synthèse de RsaI signale que la concentration en glucose diminue dans le milieu extracellulaire et que la croissance des bactéries est ralentie. En interagissant avec ses ARNm cibles ou d'autres sARN, il permet à la population bactérienne de modifier son métabolisme lorsque la source carbonée primaire est consommée.

RsaI est inhibé par CcpA en présence de glucose

RsaI est un sARN de 144 nucléotides, très conservé chez les *staphylococcaceae*, et qui comporte deux motifs riches en guanosine ainsi qu'une longue région simple brin riche en pyrimidines. Un motif *cre* de liaison à CcpA a été détecté dans la région promotrice du gène codant RsaI. Ce sARN est exprimé dès l'entrée en phase stationnaire de croissance des bactéries dans un milieu riche contenant 2 % de glucose [6]. Néanmoins, dans un

milieu sans glucose ou dans une souche délétée du gène *ccpA*, RsaI est exprimé de manière constitutive, suggérant que CcpA est le principal répresseur de RsaI en présence de glucose et que cette répression est levée après consommation du glucose (en début de phase stationnaire).

L'identification des cibles de RsaI révèle ses fonctions

La technique MAPS (*MS2-affinity purification coupled with RNA sequencing*) a été utilisée afin d'identifier l'ensemble des cibles ARN directes de RsaI (Figure 1) [7]. Cette technique permet d'isoler les complexes formés entre RsaI et ses ARN associés *in vivo*. RsaI a été au préalable fusionné à une étiquette MS2 placée en 5' du sARN. Cette étiquette est spécifiquement reconnue par la protéine de capside du phage MS2 (MS2-CP), qui est elle-même fixée sur une résine d'amylose pour retenir spécifiquement les complexes contenant RsaI-MS2. En couplant à cette chromatographie d'affinité une technique de séquençage à haut débit (Miseq, Illumina), les ARN liés à RsaI-MS2 sont caractérisés. L'interaction directe entre RsaI et ses cibles potentielles, ainsi que les conséquences de son action sur la régulation du métabolisme de la bactérie ont été ensuite validées par une combinaison d'approches *in*

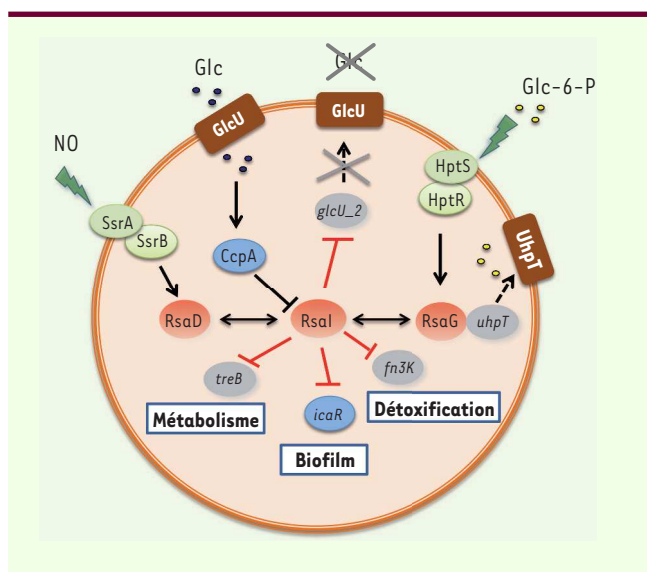


Figure 1. Représentation schématisque des réseaux de régulation centrés sur l'ARN régulateur RsaI. Les sARN sont colorés en rouge, les facteurs de transcription en bleu, les systèmes à deux composants en vert, les ARNm en gris, et les transporteurs en marron. Les connecteurs noirs ou rouges indiquent respectivement un contrôle transcriptionnel ou post-transcriptionnel. Les flèches symbolisent une activation, tandis que les barres transversales indiquent une inhibition. sARN : small RNA ; RsaI : RNA *Staphylococcus aureus* I.

vivo et *in vitro*. Ainsi, nous avons révélé l'existence de deux classes de cibles pour RsaI : des ARNm impliqués dans l'importation du glucose (*glcU_2*), le métabolisme des sucres (*treB*), la détoxification des protéines (*fn3K*), et la formation des biofilms (*icaR*), et des sARN possédant des séquences riches en cytosine, complémentaires de la séquence du site de fixation du ribosome sur leurs ARNm cibles (séquence de Shine-Dalgarno). Par ailleurs, le transcriptome différentiel indique que RsaI inhiberait la voie des pentoses phosphate et, à l'inverse, activerait des enzymes des voies de fermentation ainsi que d'autres protéines requises pour résister au monoxyde d'azote (NO), un métabolite libéré par les cellules endothéliales, les polynucléaires neutrophiles, et les macrophages de l'hôte.

RsaI contient deux domaines fonctionnels distincts

La mutagenèse dirigée a permis de distinguer deux motifs fonctionnels de RsaI. Sa région en simple brin riche en adénine et uridine interagit avec les séquences de Shine-Dalgarno de ses ARNm cibles, inhibant ainsi leur traduction. C'est le cas pour les transcrits codant le transporteur du glucose *GlcU_2*, le transporteur du tréhalose *TreB*, et la fructosamine kinase *Fn3K* impliquée dans

interagissent quant à eux avec les motifs riches en cytosine des sRNA *RsaG* et *RsaD*. *RsaG* est induit par la présence de glucose-6-phosphate (G6P) dans le milieu extracellulaire sous le contrôle du système à deux composants *HptRS* (hexose phosphate transport RS, R désignant le régulateur et S le senseur), alors que *RsaD* répond au NO sous la dépendance du système *SsrAB* (staphylococcal respiratory response AB) (Figure 1). Étonnamment, *RsaG* serait co-transcrit avec l'ARNm de *uhpT* (uptake of hexose phosphate transporter) codant le transporteur du G6P (Figure 1). Nous avons montré que la fixation de *RsaG* n'affecte pas les fonctions régulatrices de RsaI, telles que la répression traductionnelle de *glcU_2* (Figures 1 et 2B). Il reste à déterminer si RsaI peut agir comme une « éponge à sARN » afin d'empêcher leur action.

RsaI module la synthèse de PIA-PNAG, un composant de la matrice des biofilms

S. aureus est la cause principale des infections nosocomiales chez les patients opérés pour la pose d'une prothèse, où se développent facilement des biofilms bactériens résistants au traitement antibiotique. L'opéron *icaADBC* code les enzymes responsables de la synthèse de l'exopolysaccharide PIA-PNAG (polysac-

charide intercellular adhesin-poly-N-acetylglucosamine), composant majeur de la matrice des biofilms. Ces gènes sont réprimés par le facteur de transcription *IcaR*. L'ARNm *icaR* a la particularité d'inhiber sa propre traduction en induisant sa circularisation. En effet, une séquence riche en cytosine (anti-SD) située dans la région 3' non-traduite (3'UTR pour 3' untranslated region) s'hybride à la séquence SD pour empêcher la fixation du ribosome et induire la dégradation de l'ARNm *icaR* par l'endoribonucléase III (Figure 2C) [8]. L'ARNm *icaR* a été identifié comme une cible de RsaI par la technique MAPS. Nous avons montré que RsaI interagit avec la 3'UTR de l'ARNm *icaR* en aval de la séquence anti-SD (Figure 2C). Des expériences permettant d'évaluer la production de PIA-PNAG indiquent que RsaI et la région 3'UTR de l'ARNm *icaR* sont nécessaires pour inhiber le répresseur *IcaR* et activer la synthèse de PIA-PNAG (Figure 2C).

Conclusion

Ces résultats illustrent le lien étroit entre la régulation du métabolisme, l'adaptation au stress, et la formation de biofilms bactériens chez *S. aureus*, impliquant le facteur de transcription *CcpA* et le sARN *RsaI*, et l'utilisation des sources carbonées. Cet ARN non-codant aux multiples facettes pourrait être principalement produit au niveau des sites d'inflammation où la concentration en glucose est faible (e.g., poumons, reins), et intervenir dans la persistance des bactéries lorsque les conditions sont défavorables à leur multiplication. Il reste encore beaucoup à apprendre sur le rôle des sARN et sur leurs réseaux

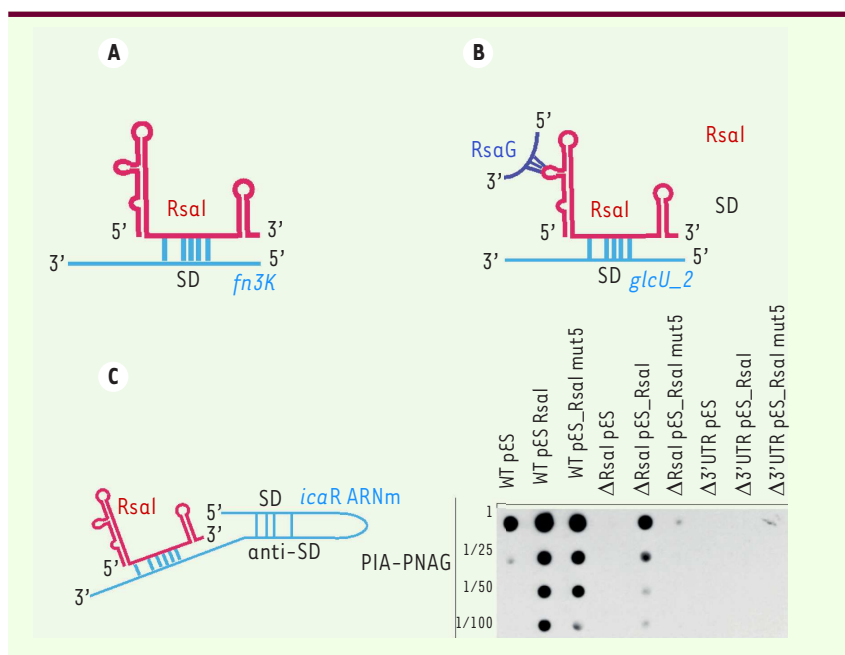


Figure 2. Divers mécanismes de régulation impliquant RsaI.

A. Schéma du complexe formé entre RsaI et l'ARNm de *fn3K*. RsaI inhibe la traduction de cet ARNm en se fixant à sa séquence de Shine-Dalgarno (SD). **B.** Schéma du complexe formé entre RsaI, l'ARNm de *glcU_2*, et RsaG. La liaison de RsaG n'empêche pas RsaI de se fixer à son ARNm cible et d'y exercer son contrôle. **C.** Effet de RsaI sur la synthèse des exopolysaccharides PIA-PNAG *in vivo*. À gauche, schéma représentant RsaI se liant à la région 3' de l'ARNm *icaR*. L'interaction entre les séquences complémentaires SD et anti-SD est également indiquée. À droite, la biosynthèse de PIA-PNAG a été quantifiée par la technique de dot blot. Des échantillons ayant subi une série de dilutions (jusqu'à 1:100) ont été déposés sur une membrane de nitrocellulose, et la production de PIA-PNAG a été détectée avec des anticorps

spécifiques de cet exopolysaccharide dans la souche 132 de *S. aureus* (WT, pour *wild-type*), la souche délétée pour le gène *rsal* ($\Delta rsal$), et la souche portant une délétion de la région *icaR* 3'UTR ($\Delta 3'UTR$). Ces trois souches ont été transformées soit avec un plasmide pES (pES), soit avec le plasmide recombinant exprimant *rsal* (pES_ *rsal*) ou sa forme mutée *rsal* mut5 (pES_ *rsal* mut5), dont la séquence d'interaction avec l'ARNm *icaR* a été inactivée par substitution (d'après [5]).

d'interactions dans le contrôle des voies métaboliques essentielles à l'invasion, la survie, et la dormance de *S. aureus*. Elucider ces mécanismes complexes de régulation permettra de mieux comprendre la pathogénicité de *S. aureus*, et de développer de nouvelles stratégies thérapeutiques contre les infections redoutables dont elle est parfois responsable. ♦

RsaI, a multifaceted regulatory RNA, modulates the metabolism of the opportunistic pathogen *Staphylococcus aureus*

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

- Richardson AR, Somerville GA, Sonenshein AL. Regulating the intersection of metabolism and pathogenesis in Gram-positive bacteria. *Microbiol Spectr* 2015 ; 3 : 1-27.
- Seidl K, Stucki M, Ruegg M, et al. *Staphylococcus aureus* CcpA affects virulence determinant production and antibiotic resistance. *Antimicrob Agents Chemother* 2006 ; 50 : 1183-94.
- Li C, Sun F, Cho H, et al. CcpA mediates proline auxotrophy and is required for *Staphylococcus aureus* pathogenesis. *J Bacteriol* 2010 ; 192 : 3883-92.
- Wagner EGH, Romby P. Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Adv Genet* 2015 ; 90 : 133-208.

- Bronesky D, Desgranges E, Corvaglia A, et al. A multifaceted small RNA modulates gene expression upon glucose limitation in *Staphylococcus aureus*. *EMBO J* 2019 ; 38 : e99363.
- Geissmann T, Chevalier C, Cros M-J, et al. A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation. *Nucleic Acids Res* 2009 ; 37 : 7239-57.
- Lalaouna D, Desgranges E, Caldelari I, et al. MS2-affinity purification coupled with RNA sequencing approach in the human pathogen *Staphylococcus aureus*. *Meth Enzymol* 2018 ; 612 : 393-411.
- 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.



Tarifs d'abonnement m/s - 2019

Abonnez-vous

à médecine/sciences

> Grâce à m/s, vivez en direct les progrès des sciences biologiques et médicales

Bulletin d'abonnement

page 1230 dans ce numéro de m/s



A multifaceted small RNA modulates gene expression upon glucose limitation in *Staphylococcus aureus*

Delphine Bronesky^{1,†}, Emma Desgranges^{1,†}, Anna Corvaglia², Patrice François², Carlos J Caballero³, Laura Prado³ , Alejandro Toledo-Arana³ , Inigo Lasa⁴, Karen Moreau⁵, François Vandenesch⁵, Stefano Marzi¹ , Pascale Romby^{1,*}  & Isabelle Caldelari^{1,**} 

Abstract

Pathogenic bacteria must rapidly adapt to ever-changing environmental signals resulting in metabolism remodeling. The carbon catabolite repression, mediated by the catabolite control protein A (CcpA), is used to express genes involved in utilization and metabolism of the preferred carbon source. Here, we have identified RsaI as a CcpA-repressed small non-coding RNA that is inhibited by high glucose concentrations. When glucose is consumed, RsaI represses translation initiation of mRNAs encoding a permease of glucose uptake and the FN3K enzyme that protects proteins against damage caused by high glucose concentrations. RsaI also binds to the 3' untranslated region of *icaR* mRNA encoding the transcriptional repressor of exopolysaccharide production and to sRNAs induced by the uptake of glucose-6 phosphate or nitric oxide. Furthermore, RsaI expression is accompanied by a decreased transcription of genes involved in carbon catabolism pathway and an activation of genes involved in energy production, fermentation, and nitric oxide detoxification. This multifaceted RNA can be considered as a metabolic signature when glucose becomes scarce and growth is arrested.

Keywords carbon metabolism; catabolite control protein A; pathogenic bacteria; regulatory RNAs; sRNA; translational regulation

Subject Categories Metabolism; Microbiology, Virology & Host Pathogen Interaction; RNA Biology

DOI 10.15252/embj.201899363 | Received 6 March 2018 | Revised 17 December 2018 | Accepted 21 January 2019 | Published online 13 February 2019

The EMBO Journal (2019) 38: e99363

Introduction

All bacteria require a carbon source, providing energy for their growth and for the synthesis of all macromolecules. Besides, pathogenic bacteria during the host infectious process must cope with hostile conditions such as nutrient deficiency, temperature, and oxidative and osmotic shocks, and must overcome innate immune responses. For instance, *Staphylococcus aureus* uses carbohydrates to grow under high nitric oxide (NO) and anaerobiosis (Vitko *et al*, 2016). To survive in these complex environments and to counteract the host defense, *S. aureus* produces a plethora of virulence factors. The synthesis of these factors is fine-tuned by intricate interactions between multiple regulators involving both proteins and RNAs (Novick, 2003). Additionally, biosynthetic intermediates, generated from the central metabolism of *S. aureus*, have strong impacts on the synthesis of virulence factors. Besides, several metabolite-sensing regulatory proteins (CcpA, CodY, Rex, and RpiR) act as key regulatory factors to coordinate the synthesis of genes involved in metabolic pathways, in stress responses, and in pathogenesis (Somerville & Proctor, 2009; Richardson *et al*, 2015). Through the adaptation of the metabolism of the bacteria to specific host microenvironment, these proteins contribute to *S. aureus* pathogenesis (Richardson *et al*, 2015).

Among these proteins, the catabolite control protein A (CcpA) acts as the main global regulator of carbon catabolite repression, allowing the bacteria to use the preferred carbon source (i.e., glucose) in a hierarchical manner (Seidl *et al*, 2008a, 2009). Catabolite control protein A belongs to the LacI repressor family and binds to a specific DNA sequence, called the *cre* (catabolite-responsive element) sequence, which is conserved in many Gram-positive bacteria. Transcription of CcpA is constitutive, and the protein is activated through the binding of its co-regulator histidine-containing

¹ Architecture et Réactivité de l'ARN, CNRS, Université de Strasbourg, Strasbourg, France

² Genomic Research Laboratory, Department of Medical Specialties, Geneva University Hospitals, University of Geneva, Geneva, Switzerland

³ Instituto de Agrobiotecnología (IdAB), CSIC-UPNA-GN, Navarra, Spain

⁴ Navarrabiomed-Universidad Pública de Navarra-Departamento de Salud, IDISNA, Pamplona, Spain

⁵ CIRI, Centre international de Recherche en Infectiologie, Inserm, U1111, Université Claude Bernard Lyon 1, CNRS, UMR5308, École Normale Supérieure de Lyon, Hospices Civils de Lyon, Univ Lyon, Lyon, France

*Corresponding author. Tel: +33 388417068; Fax: +33 388602218; Email: p.romby@ibmc-cnrs.unistra.fr

**Corresponding author. Tel: +33 388417068; Fax: +33 388602218; Email: i.caldelari@ibmc-cnrs.unistra.fr

[†]These authors contributed equally to this work

phosphocarrier protein HPr in the presence of glucose. Inactivation of *ccpA* gene decreases transiently the growth rate and yield as long as glucose is present in the medium (Seidl *et al*, 2006) and affects the expression of a large number of metabolic encoding genes in a glucose-dependent and glucose-independent manner (Seidl *et al*, 2008a). Additionally, CcpA has a strong impact on the expression of *S. aureus* virulon. It enhances the yield of the quorum-sensing-induced RNAIII, which represses protein A and various adhesion factors at the post-transcriptional level, and conversely activates the synthesis of many exotoxins. However, CcpA also modulates the transcription of mRNAs encoding protein A, α -hemolysin (*hla*), and toxic shock syndrome toxin (TSST; Seidl *et al*, 2006, 2008b), represses capsule formation, and activates biofilm formation in a glucose-rich environment (Seidl *et al*, 2008a). Indeed, a *S. aureus* *ccpA* deletion mutant strain was less pathogenic than the wild-type strain in a murine abscess model (Li *et al*, 2010) and *ccpA* inactivation increased the susceptibility of hyperglycemic animals to acute pneumonia infections (Bischoff *et al*, 2017). Nevertheless, the mechanism by which CcpA affected *S. aureus* pathogenesis cannot be simply resumed as a modulation of the RNAIII-dependent regulatory networks. Therefore, it has been suggested that CcpA can also act indirectly on gene expression through the action of other regulatory proteins or small non-coding RNAs (sRNAs; Somerville & Proctor, 2009; Richardson *et al*, 2015).

In Enterobacteriaceae, several sRNAs have been shown as key actors of the uptake and the metabolism of carbohydrates (reviewed in Bobrovskyy & Vanderpool, 2016). For instance, they participate in the regulation of the galactose operon and carbon catabolite repression, metabolism of amino acids, and contribute to bacterial survival during phospho-sugar stress. The importance of sRNAs in regulatory networks is now well recognized to rapidly adjust cell growth to various stresses and changes in the environment. Thus, they are obvious candidates creating the links between virulence and metabolism. One example is RsaE, a sRNA conserved in *S. aureus* and *Bacillus subtilis*, that controls enzymes involved in the TCA cycle (Geissmann *et al*, 2009; Bohn *et al*, 2010) and in arginine degradation pathway (Rochat *et al*, 2018). In *B. subtilis*, the transcriptional repressor ResD represses RoxS (homologous to *S. aureus* RsaE) to readjust the pool of NAD^+/NADH in responses to various stress and stimuli (Durand *et al*, 2017). Its promoter is also highly conserved among *Staphylococceae* and is recognized by the orthologous response regulator SrrA in *S. aureus*. Responding to reactive oxygen species through SrrAB, *S. aureus* RsaE may also intervene in the survival of cells against host immune reactions (Durand *et al*, 2015, 2017).

Here, we have identified a signaling pathway responding to glucose uptake, which involves a sRNA, called RsaI (or RsaOG). This 144 nucleotides-long sRNA is highly conserved among *Staphylococceae* and carries two conserved regions including two G-track sequences and a long unpaired interhelical region rich in pyrimidines (Geissmann *et al*, 2009; Marchais *et al*, 2010). The expression of RsaI is observed exclusively at the stationary phase of growth in rich medium (Geissmann *et al*, 2009) and is enhanced after vancomycin exposure (Howden *et al*, 2013). In this study, we revealed that CcpA is the main repressor of RsaI expression in the presence of glucose and that this inhibition is alleviated after the utilization of glucose. The identification of the targetome of RsaI using the MS2-affinity purification approach coupled with RNA sequencing

(MAPS) unexpectedly showed two classes of RNA targets, including mRNAs involved in glucose uptake, sugar metabolism, and biofilm formation, as well as various sRNAs. Using site-directed mutagenesis, we identified two distinct and functional regions of RsaI. All in all, our data showed the existence of two sRNAs involved in essential pathways responding to either glucose or glucose-6 phosphate uptake, with RsaI acting as a signature for a metabolic switch when the preferred carbon source is metabolized.

We will discuss the importance of sRNA-mediated regulation in *S. aureus* to fine tune the expression of genes according to essential nutrient availability and their possible consequences on metabolism adaptation and virulence.

Results

The expression of RsaI is inhibited by glucose and by the catabolite control protein A

We have previously shown that the synthesis of RsaI is high at the stationary phase of growth in BHI medium while its expression was constitutive in MHB medium (Geissmann *et al*, 2009). These data suggested that RsaI expression is regulated in a manner dependent on nutrient or biosynthetic intermediate availability. A major difference between BHI and MHB composition is their carbon source, glucose, and starch, respectively. We wondered whether the expression of RsaI might be dependent on the available carbon source. For this purpose, Northern blot experiments were performed on total RNAs prepared from HG001 (wild-type) strain grown in MHB medium at various time points (Fig 1A). In MHB, where the glucose is not immediately available as the carbon source, RsaI was constitutively and highly expressed (Fig 1A). Conversely, when glucose was added, either at the beginning of the culture or after 3 h of growth, the steady-state level of RsaI was immediately dropped, indicating the repressing effect of this sugar (Fig 1A).

As CcpA sensed the intracellular concentration of glucose (Seidl *et al*, 2006), we analyzed whether the expression of RsaI was CcpA-dependent. Northern blot analysis was performed on total RNA extracts prepared from HG001 strain and an isogenic mutant deleted of *ccpA* gene (ΔccpA). In parallel, we have also tested CodY, another global regulator of metabolism and virulence, which was shown as a direct regulator of amino acid biosynthesis and transport of macromolecules (Pohl *et al*, 2009; Majerczyk *et al*, 2010). The data showed that in the absence of glucose, the yield of RsaI was similar in all strains. However, in MHB medium supplemented with glucose, the yield of RsaI declined dramatically in the WT and ΔcodY strains, whereas it was still high in ΔccpA strain (Fig 1B). In BHI medium, RsaI was expressed after 4 h of growth in the WT strain, while its expression became constitutive in the mutant ΔccpA strain (Figs 1C and EV1A).

We also analyzed the expression of RsaI in MHB medium supplemented with various sugars such as fructose and xylose (Fig 1D). The data showed that expression of RsaI was very low at the beginning of growth in MHB supplemented with either glucose or fructose but not with xylose. These data show that CcpA senses both glucose and fructose (Fig 1).

Overall, we showed that CcpA represses RsaI expression in the presence of glucose. In accordance with this observation, a

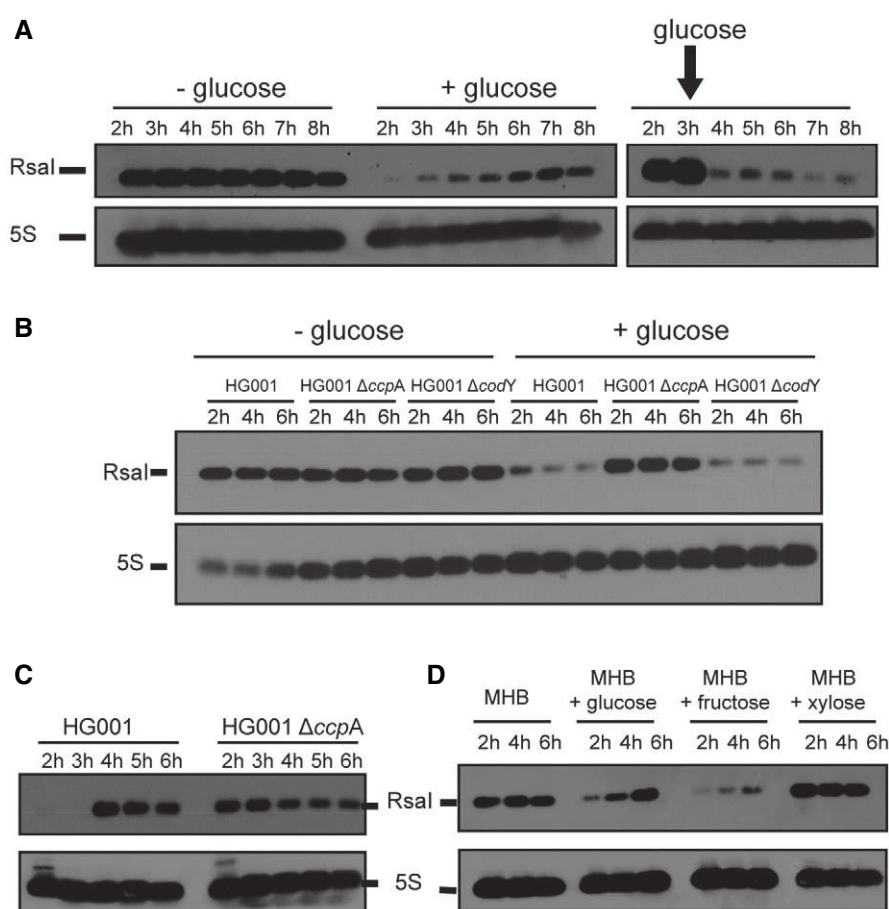


Figure 1. RsaI responds to glucose through the transcriptional factor CcpA.

- A Northern blot experiments show the expression of RsaI during growth phase in the HG001 strain in MHB medium with or without the addition of 1.5 g/l of *D*-glucose. Glucose was added either at the beginning of growth (+ glucose, left panel) or after 3 h of growth (right panel).
- B Northern blot experiment shows the expression of RsaI during growth phase in the HG001, Δ ccpA mutant strain, and Δ codY mutant strain, in MHB medium with (+) or without (–) the addition of 1.5 g/l of *D*-glucose.
- C Northern blot experiment shows the expression of RsaI in the HG001 and Δ ccpA mutant strains. Total RNA was prepared after 2, 3, 4, 5, and 6 h of culture in BHI medium at 37°C.
- D Northern blot analysis of RsaI in the HG001 strain grown in MHB medium with or without the addition of 1 g/l of glucose, fructose, or xylose. For all the experiments, loading controls were done using the expression of 5S rRNA (5S) as revealed after hybridization of the membranes with a specific probe. However for these controls, we used aliquots of the same RNA preparations, but the migration of the samples was performed in parallel to the experiments on a separate agarose gel because RsaI and 5S rRNA have very similar sizes.

Source data are available online for this figure.

conserved *cre* (GGAAAcGcTTACAT) sequence was found at position –30 upstream the transcriptional start site of RsaI (Fig EV1B). This region was sufficient to confer repression by glucose in the complemented strain containing pCN51::PrsaI (Fig EV1C).

The targetome of RsaI as revealed by the MAPS approach

The MAPS approach (“MS2 affinity purification coupled with RNA sequencing”) was used to purify *in vivo* regulatory complexes involving RsaI. MAPS has been successful to identify the RNA targets of any sRNAs in *Escherichia coli* (Lalaouna et al, 2015) and more recently of RsaA sRNA in *S. aureus* (Tomasini et al, 2017). Briefly, the MS2-tagged version of RsaI was expressed from a plasmid under the control of the *agr*-dependent P3 promoter, allowing an accumulation of RsaI at the stationary phase of growth in the

Δ rsaI mutant strain. RsaI was detected by Northern blot using total RNAs extracted at 2, 4, and 6 h of growth in BHI medium. Using a DIG-labeled RsaI probe, we showed that the steady-state levels of MS2-RsaI were very similar to the wild-type (WT) RsaI and that MS2-RsaI was specifically and strongly retained by the MS2-MBP fusion protein after the affinity chromatography (Fig EV1D). The RNAs were then extracted from the eluted fraction to be sequenced. The data were analyzed using Galaxy (Afgan et al, 2016), and the sequencing reads were mapped, counted per feature, and normalized using the HG001 genome as reference (Herbert et al, 2010; Caldelari et al, 2017). The enrichment of putative targets was measured by comparing the number of reads obtained from the MS2-RsaI purification and the MS2 alone as control. Since the MS2 tag alone was produced in the WT HG001 background, the untagged RsaI was also expressed under the conditions of growth but was

poorly retained on the affinity chromatography (Fig EV1D). The data were sorted by a decreasing fold change. In the following study, we have considered as RsaI targets, the RNAs that were enriched at least fourfold and were reproducibly detected in two independent experiments (Tables EV1 and EV2).

Two classes of RNAs were co-purified with RsaI, including mRNAs and sRNAs. On the one hand, the most enriched mRNA encodes IcaR, the repressor of the *icaADBC* operon, which encodes the enzymes required for the synthesis of the poly- β -1,6-N-acetylglucosamine polymer (PIA-PNAG), the main staphylococcal exopolysaccharide constituting biofilms. In addition, several mRNAs encoding proteins linked to sugar utilization and transport, such as *glcU_2* encoding a major transporter of glucose, and *fn3K* encoding fructosamine 3-kinase were significantly enriched. It is noteworthy that other less enriched mRNAs (< 4-fold) encode proteins related to sugar metabolism such as the trehalose-specific PTS transporter (TreB), a sugar phosphatase (YidA), and a maltose transport system permease (Table EV2). Finally, in addition to *icaR*, several mRNAs express transcriptional regulators (Xre type, the maltose regulatory protein GlvR, SlyA, SigS). On the other hand, we found the RsaD, RsaE, and RsaG sRNAs as putative RsaI targets (Table EV1), which all contained at least one conserved C-rich motif (Geissmann *et al*, 2009). RsaH could be also included in this group though it was less enriched (< 4-fold; Table EV2).

The MAPS data suggested that RsaI is involved in networks of RNA pairings.

RsaI contains two distinct regulatory domains

Following the MAPS experiments, we searched for intermolecular base-pairing interactions between RsaI and its potential RNA targets using IntaRNA. Stable interactions were predicted for most of the enriched RNAs (Table EV1). The CU-rich unpaired region of RsaI was predicted to form base-pairings with most of the mRNAs. They were located close or at the ribosome-binding site of most of the mRNAs except for *icaR* and *isaA*, which involves nucleotides in their 3' untranslated regions (Table EV1). A second domain of interaction corresponds to the G-track sequences located in the first hairpin domain of RsaI and the C-rich sequences of the sRNAs RsaD, RsaE, RsaG, and RsaH (Fig 2A).

These results suggested that RsaI contains at least two functional domains, the unpaired CU-rich region dedicated to mRNA pairing and the G-track used to interact with other sRNAs. To validate whether RsaI effectively binds to the RNA candidates identified by MAPS through these motifs, we performed gel retardation assays

using WT and mutated RsaI molecules (Figs 2B and C, and EV2). For that, *in vitro* 5' end-labeled RsaI was first incubated with increasing concentrations of various mRNAs encoding proteins involved in biofilm formation (IcaR), sugar uptake and metabolism (GlcU_2 and Fn3K), and several sRNAs (RsaG, RsaE, and RsaD). For these experiments, we used the full-length mRNAs and sRNAs (Table EV3). Complex formation was performed with RNAs, which were renatured separately in a buffer containing magnesium and salt. As expected, the data showed that RsaI formed complexes with high affinity (between 20 and 100 nM, Table EV1) with many RNAs such as *icaR*, *glcU_2*, and *fn3K* mRNAs, and RsaG sRNA (Fig 2B). The stability of other complexes (e.g., *treB* mRNA and RsaD sRNA) was significantly lowered (> 250 nM) (Fig EV2).

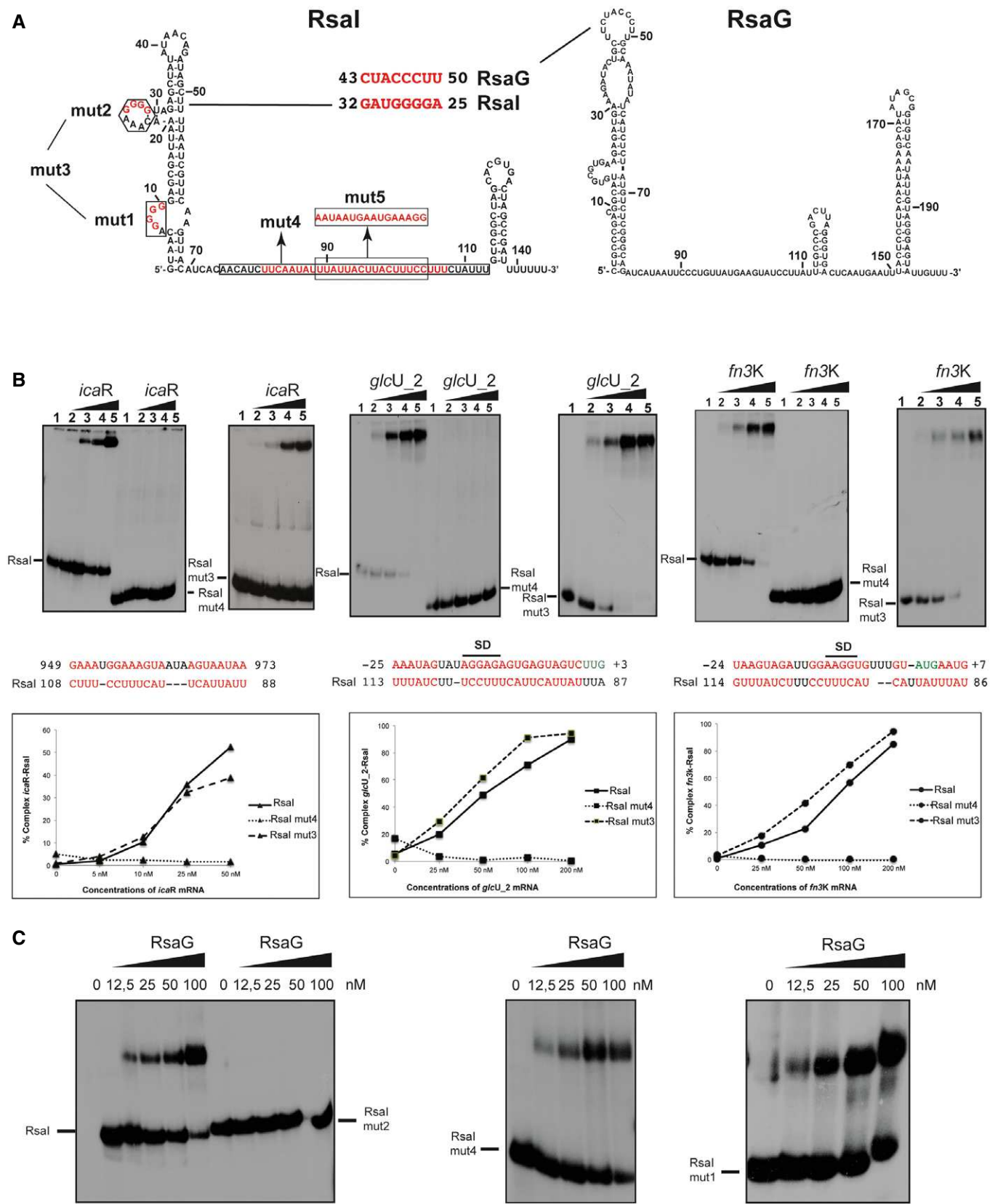
Based on the base-pairing predictions, mutations have been introduced into *rsaI* to map its functional regulatory regions. The two conserved G-track sequences were mutated separately (RsaI mut1: Δ G7-G10, RsaI mut2: Δ G26-G29) or together (RsaI mut3: Δ G7-G10/ Δ G26-G29), while several nucleotides (RsaI mut4: Δ U81-U107) were deleted in the conserved interhelical unpaired sequence (Fig 2A). We analyzed the ability of mutated RsaI derivatives to form complexes with *glcU_2*, *fn3K*, and *icaR* mRNAs, and RsaG sRNA (Fig 2B). Quantification of the data showed that RsaI mut3 binds to all three mRNAs similarly to the WT RsaI, while complex formation was completely abolished with RsaI mut4. Only the mutation in the second G-track sequences (RsaI mut2) strongly altered the binding to RsaG, while the two other mutated RsaI (mut1 and mut4) recognized RsaG with an equivalent binding affinity as the WT RsaI (Fig 2C).

To demonstrate that the RNA-RNA interactions were also occurring *in vivo*, MAPS approach was used to monitor the effect of the RsaI mutations on its target RNAs. The RsaI mut2 and mut4, which independently affected *in vitro* binding of sRNAs and mRNAs, respectively, were tagged with MS2 and expressed in the HG001 Δ rsaI mutant strain. As described above, the enrichment of putative targets was calculated using a similar procedure (Table EV4). For this analysis, we have used the data coming from three independent experiments performed with the WT MS2-RsaI, the two MS2 controls and the MS2-RsaI mutants (mut2 and mut4) taking into account that the MAPS was done in two distinct sets of experiments (see Supplementary Material and Methods). This allowed us to directly compare the fold changes between the WT and the mutant RsaI versions (Table EV5). In the fraction containing MS2-RsaI mut4, most of the mRNA targets including *icaR*, *fn3K*, and *glcU_2* were strongly reduced while RsaG was still significantly enriched. Conversely, we observed that the three mRNAs were still

Figure 2. RsaI binds to *icaR*, *glcU_2*, *fn3K* mRNA, and the sRNA RsaG.

- A Secondary structure model of RsaI. In red, are the nucleotides deleted in the RsaI mutants (mut1 to mut4), and the nucleotides, which were substituted (mut5). The potential base-pairings between RsaI and RsaG are shown. Squared nucleotides are conserved sequences in RsaI.
- B Gel retardation assays show the formation of the complex between RsaI and *icaR*, *glcU_2* and *fn3K* mRNAs. The 5' end-labeled wild-type RsaI (RsaI), RsaI mutant 3 (RsaI mut3, deletion of the two G-track motifs), and RsaI mutant 4 (RsaI mut4, deletion of the C/U-rich unpaired interhelical region) were incubated with increasing concentrations of mRNAs: lane 1, 0 nM; lane 2, 25 nM; lane 3, 50 nM; lane 4, 100 nM; and lane 5, 200 nM. Below the gels, the predicted interactions between RsaI and its targets are shown. Translation start codons are in green, and SD is for Shine and Dalgarno sequence. Graphs represented the % of complex formed between either RsaI or its two mutant forms (RsaI mut3 and RsaI mut4) and the target mRNA (*icaR*, *glcU_2*, *fn3K*) as the function of mRNA concentrations.
- C Gel retardation assays show the formation of the complex between RsaI and RsaG. The 5' end-labeled wild-type RsaI (RsaI), RsaI mutant 2 (RsaI mut2), RsaI mutant 4 (RsaI mut4), and RsaI mutant 1 (RsaI mut1) were incubated with increasing concentrations of RsaG given in nM on the top of the autoradiographies.

Source data are available online for this figure.



co-purified together with the MS2-RsaI mut2 at a level close to the WT RsaI while RsaG was strongly reduced in the fraction containing MS2-RsaI mut2 (Table EV5). In addition, with the two MS2-RsaI mutants, RsaE was almost lost while the enrichment of the transcriptional regulator XRE was strongly reduced. Finally, *dck* mRNA remained enriched at the same level as for the WT MS2-RsaI (Table EV5). These data strongly suggested that mutations at the G-track motif of RsaI had a major effect on RsaG/RsaE sRNAs binding, whereas the deletion of the C/U-rich sequence in RsaI mut4 had a major impact on the recognition of many mRNAs. The loss of RsaE in the MAPS performed with MS2-RsaI mut4 suggested that the sRNA might also be associated with one of the RsaI-dependent mRNA target. In support to the hypothesis, base-pairings have been recently predicted between RsaE and the 5'UTR of *icaR* mRNA (Rochat et al, 2018).

Taken together, our data reveal that RsaI has at least two distinct regulatory domains that directly interact either with mRNAs or with sRNAs.

RsaI inhibits ribosome binding by masking the Shine and Dalgarno sequence of mRNAs

Because the C/U unpaired region of RsaI was predicted to form base-pairings with the Shine and Dalgarno (SD) sequence of several mRNAs (Figs 2 and 3A), we analyzed whether RsaI would have the ability to compete with the ribosome for binding to *glcU_2* and *fn3K* mRNAs. Using toe-printing assays, we analyzed the effect of RsaI on the formation of the ternary initiation complex constituting of mRNA, the initiator tRNA, and the 30S subunit. Binding of the 30S on the two mRNAs is illustrated by the presence of a toe-print signal at position +16 (+1 being the initiation codon). For *fn3K* mRNA, two toe-print signals were observed at +16 and at +20, most probably corresponding to the presence of two AUG codons distant of 5 nucleotides. However, only the first AUG is used *in vivo* (Gemayel et al, 2007). The addition of increasing concentrations of RsaI together with the 30S strongly decreased the toe-print signals for both mRNAs showing that RsaI is able to form a stable complex with the mRNA sufficient to prevent the binding of the 30S subunit (Fig 3A).

The *in vivo* relevance of RsaI-dependent repression of several mRNAs (*glcU_2*, *fn3K*, *treB*, *HG001_01242*, *HG001_02520*) was then analyzed using gene reporter assays. The whole leader regions of the mRNAs (54 nts in *glcU_2*, 34 nts in *fn3K*, 24 nts in *treB*, 35 nts

in *HG001_01242*, and 72 nts in *HG001_02520*) including 100 nucleotides of their coding sequences were cloned in frame with *lacZ* under the control of the strong promoter *PrpB*. The synthesis of β -galactosidase was analyzed in the Δ *rsaI* mutant strain and in the same strain transformed with the plasmid carrying the *lacZ* reporter and RsaI under its own promoter. For the *fn3K-lacZ* fusion, we also analyzed its expression in the mutant strain expressing RsaI mut4. The β -galactosidase activity was reduced almost eight times in cells expressing wild-type RsaI for all reporter constructs (Fig 3B). In addition, the expression of RsaI mut4 did not cause any repression of the β -galactosidase expressed from *fn3K-lacZ* (Fig 3B). Therefore, disrupting the interaction between RsaI mut4 and *fn3K* alleviated the repression of the reporter gene suggesting that the regulation of *fn3K* is occurring at the translational level through specific RNA-RNA interactions.

These data showed that RsaI hinders ribosome binding on *glcU_2* and *fn3K* mRNAs in agreement with the fact that the C/U-rich region of RsaI would bind to the SD sequence of the mRNAs. We propose a similar mechanism for *treB*, *HG001_01242*, and *HG001_02520* mRNAs considering that the predicted base-pairings include the RBS of these mRNAs (Tables EV1 and EV2).

RsaI interacts with the 3'UTR of *icaR* mRNA and affects PIA-PNAG synthesis

The MAPS approach revealed that *icaR* mRNA, which encodes the repressor of the main exopolysaccharidic compound of *S. aureus* biofilm matrix, is the most enriched mRNA co-purified with MS2-RsaI. This mRNA is of particular interest because it contains a large 3'UTR that is able to bind to its own SD sequence through the anti-SD UCCCCUG motif (Ruiz de los Mozos et al, 2013). Consequently, the long-range interaction provokes an inhibitory effect on translation and generates a cleavage site for RNase III (Ruiz de los Mozos et al, 2013). RsaI is predicted to form base-pairings with the 3'UTR of *icaR* downstream of the anti-SD sequence (Table EV1, Fig 4A). We first monitored whether the long-range interaction might be critical for RsaI binding. Previous work showed that the substitution of the anti-SD UCCCCUG sequence by AGGGGAC significantly destabilized the long-range interaction to enhance *icaR* translation (Ruiz de los Mozos et al, 2013). However, gel retardation assays showed that the WT and the mutant *icaR* mRNAs bind to RsaI with an equivalent binding affinity (Fig 4B), suggesting that the anti-SD sequence is not

Figure 3. RsaI inhibits *glcU_2* and *fn3K* mRNA translation.

- A Toe-print assays showing the effect of RsaI on the formation of the ribosomal initiation complex of *glcU_2* and *fn3K* mRNAs, respectively. Lane 1: incubation control of mRNA alone; lane 2: incubation control of mRNA with 30S subunits; lane 3: incubation control of mRNA with RsaI; lane 4: formation of the ribosomal initiation complex containing mRNA, 30S, and the initiator tRNA^{Met} (tRNA_i); lanes 5–9: formation of the initiation complex in the presence of increasing concentrations of RsaI, respectively: 50 nM (lane 5), 100 nM (lane 6), 150 nM (lane 7), 300 nM (lane 8), and 400 nM (lane 9). Lanes T, A, C, G: sequencing ladders. The Shine and Dalgarno (SD) sequence, the start site of translation (START), and the toe-printing signals (+16) are indicated. At the bottom of the gels are shown the predicted interactions between RsaI and its targets. Translation start codons are in green, and the Shine and Dalgarno (SD) sequence is underlined, and the arrowheads depict the toe-printing signals.
- B The β -galactosidase activity (Miller Units) has been measured from various fusions expressed from a plasmid which also carries *rsaI* gene under its own promoter: *PrpB::glcU::lacZ::rsaI*, *PrpB::fn3K::lacZ::rsaI*, *PrpB::treB::lacZ::rsaI*, *PrpB::HG001_1242::lacZ::rsaI*, and *PrpB::HG001_2520::lacZ::rsaI* expressed in the mutant strain *HG001-ΔrsaI*. The same constructs were made in the absence of *rsaI* gene. For the *fn3K-lacZ* fusion, we also used an additional construct *PrpB::fn3K::lacZ::rsalmut4* expressing both the fusion FN3K-LacZ protein and RsaI mut4. The β -galactosidase activity was normalized for bacterial density, and the results represented the mean of four independent experiments. The error bars are standard deviations, and the statistical significance was determined using the Student's t-test. **P* < 0.05, ***P* < 0.005, ****P* < 0.0005, *****P* < 0.0001, ns is for not significant.

Source data are available online for this figure.

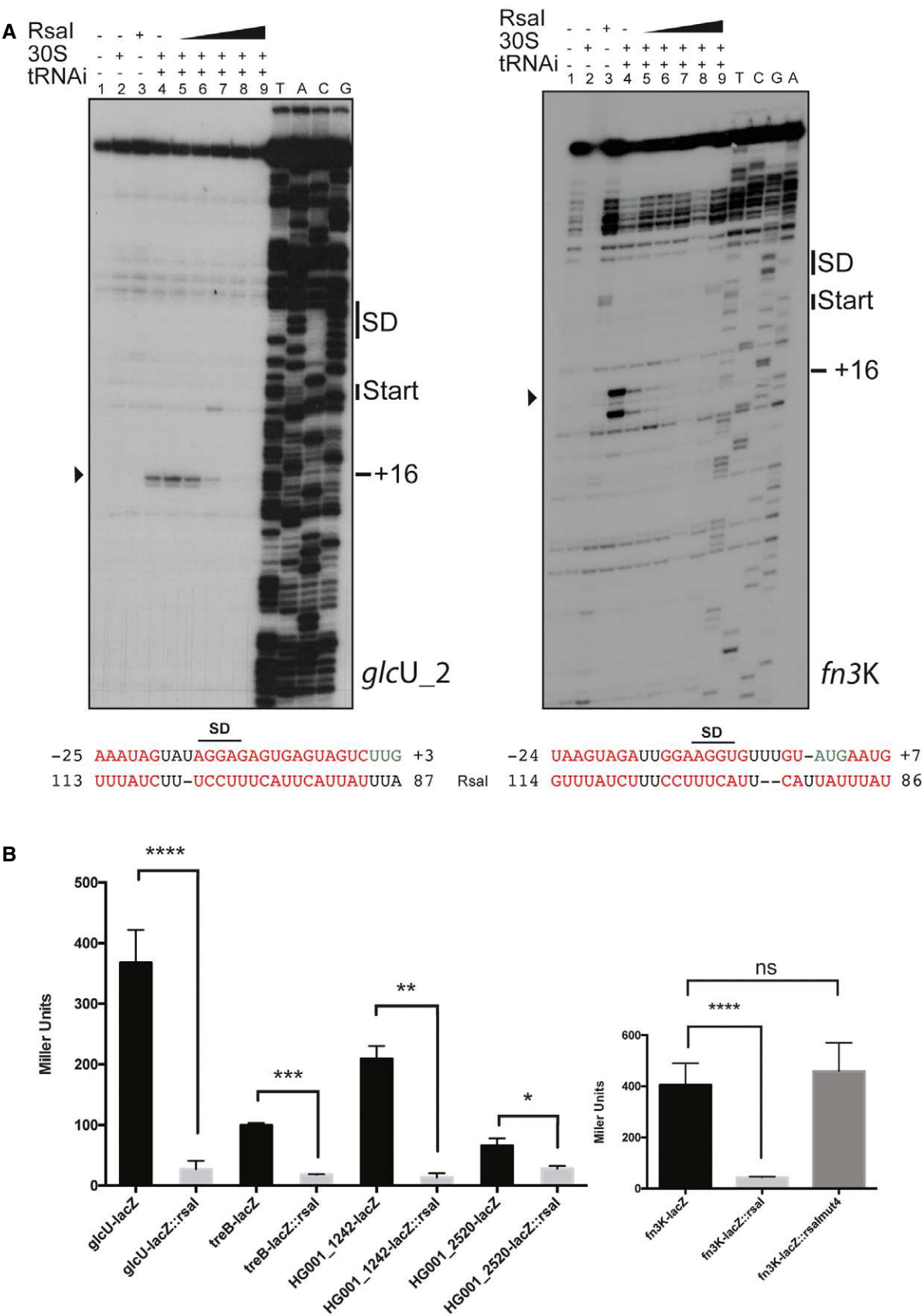


Figure 3.

required for RsaI binding. Instead, the 3'UTR of *icaR* retained the full capacity to efficiently bind RsaI (Fig 4B, Table EV1). We then analyzed whether RsaI binding to the 3'UTR would alter the circularization of *icaR* mRNA. Gel retardation assay was performed with radiolabelled 3'UTR in the presence of the 5'UTR at a concentration of 500 nM leading to almost 50% of binding (Fig EV3A, lane 2). The addition of increasing concentrations of RsaI led to the formation of a 3'UTR-RsaI binary complex but also of a ternary complex involving the three RNA species (Fig EV3A, lanes 4–8). Hence, RsaI does not prevent the formation of the interaction between the two UTRs of *icaR* mRNA *in vitro*.

Then, to analyze the *in vivo* effect of the RsaI interaction with the *icaR* mRNA, we monitored the production of PIA-PNAG as a natural reporter, assuming that variations on IcaR expression would be directly reflected on the production of this exopolysaccharide. For that, dot-blot assays were performed with anti-PIA-PNAG-specific antibodies to monitor the levels of PIA-PNAG in the strain 132, which expresses *rsaI* and produces high levels of this exopolysaccharide compared to HG001 in our experimental conditions. We measure the PIA-PNAG levels in WT 132 strain and the isogenic mutant strains Δ *rsaI* and Δ 3'UTR *icaR*, which carried a deletion of the 3'UTR of *icaR*. Bacteria were grown for 6 h in TSB containing NaCl, which is required to enhance PIA-PNAG synthesis in 132 strain (Vergara-Irigaray *et al*, 2009) (Fig 4C). The results showed that both RsaI and the 3'UTR of *icaR* are required for efficient production of PIA-PNAG because only the WT strain produces significant levels of exopolysaccharides. In addition, the three strains were transformed with a plasmid overexpressing RsaI or RsaI mut5 carrying a substitution of nucleotides 88–103 (UUAUUACUUA CUUUC to AAUAAUGAAUGAAAGG) under the control of a strong and constitutive promoter. This mutation decreased the stability of RsaI-*icaR* duplex (Fig EV3B). Northern blots confirmed RsaI overexpression in these strains (Fig 4C). Dot-blot results revealed that production of PIA-PNAG could be restored in Δ *rsaI* strain when RsaI is constitutively expressed. However, complementation is observed neither in the Δ *rsaI* expressing the RsaI mut5 nor in the Δ 3'UTR mutant strains overexpressing RsaI or RsaI mut5. These results suggested that the interaction of RsaI with 3'UTR is required to promote PIA-PNAG synthesis (Fig 4C). Interestingly, overexpression of RsaI in the WT strain considerably increased PIA-PNAG production. Although this effect is less pronounced in the WT expressing RsaI mut5, it suggested the presence of additional regulatory pathways.

Taken together, these data suggest that RsaI would contribute to PIA-PNAG synthesis by at least reducing the IcaR repressor protein levels through a specific interaction with the 3'UTR of *icaR* mRNA.

Two sRNAs responded to sugar uptake

The MAPS experiments revealed that RsaG was the most highly co-purified sRNA together with RsaI (Table EV1), and the complex formed between RsaI and RsaG is very stable (Fig 2C). Interestingly, *rsaG* is localized just downstream *uhpT* encoding the hexose phosphate transporter (Fig EV4A), whose transcription is activated by the two-component system HptRS in response to extracellular glucose-6 phosphate (G-6P), another major carbon source produced by host cells (Park *et al*, 2015). We therefore analyzed whether RsaG expression would also respond to the cellular concentration of G-6P. Northern blot experiments were performed on total RNA extracts produced from the HG001 strain grown in BHI medium supplemented with G-6P. Under these conditions, the synthesis of RsaG is strongly enhanced (Fig EV4B, left panel), in contrast to RsaI, which is completely inhibited under the same conditions (Fig EV4B, right panel). The deletion of *hptRS* considerably reduced the levels of RsaG (Fig EV4C). Therefore, these data indicated that RsaG is activated by HptRS upon G-6P signaling together with *uhpT* and that RsaG sRNA might be derived from the *uhpT* 3'UTR.

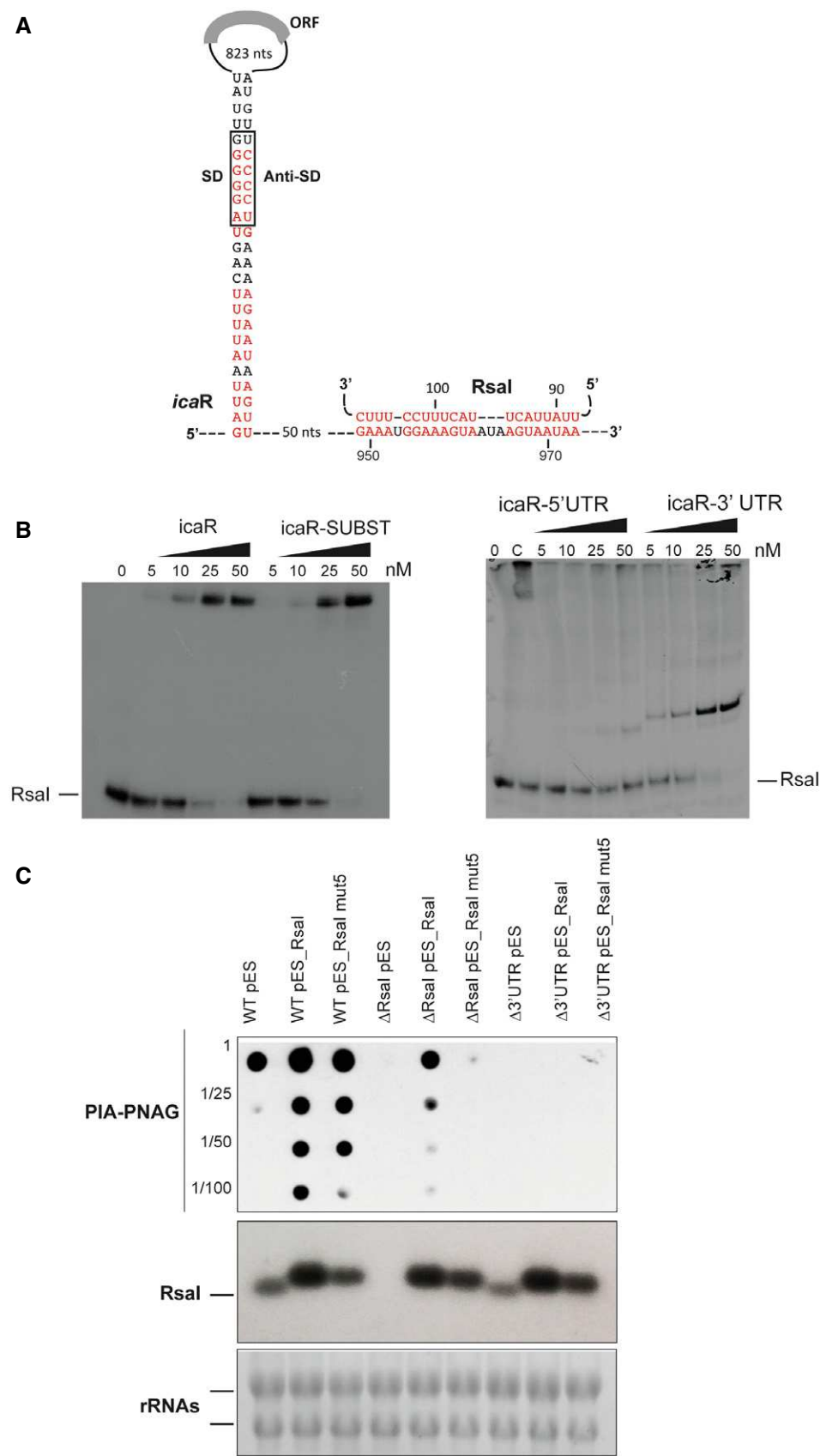
We then addressed the consequences of RsaI-RsaG pairings on target recognition. Using gel retardation assays, we analyzed whether RsaG is able to form a ternary complex with RsaI and one of its target mRNA or whether RsaG competes with the mRNA for RsaI binding. These experiments were conducted for three mRNA targets (*glcU_2*, *HG001_1242*, and *HG001_0942*), which all formed stable complexes with the 5' end-labeled RsaI (Figs 2, 5A and EV2, Table EV1). On the opposite, RsaG is not able to interact with *glcU_2* and *HG001_1242* as shown the gel retardation assays performed with 5' end-labeled RsaG (Fig 5B). We then performed competition experiments between RsaI and RsaG for mRNA binding. The experiments were performed in the presence of a single concentration of unlabeled RsaG (50 nM) sufficient to bind most of 5' end-labeled RsaI in the presence of increasing concentrations of mRNA (Fig 5A). For the three mRNAs, the addition of RsaG causes the formation of a high molecular weight complex most likely formed by RsaG, RsaI, and the mRNA (Fig 5A). Hence, RsaG does not interfere with the mRNA binding to RsaI.

Furthermore, *in vivo*, the synthesis of the β -galactosidase expressed from the *fn3K-lacZ* fusion was measured in the WT and

Figure 4. Interaction of RsaI to *icaR* mRNA induces biofilm production.

- Secondary structure model of *icaR* mRNA. The nucleotides in red are involved in the long-range interaction between the 3' and 5'UTRs or with RsaI. SD and anti-SD are, respectively, for the Shine and Dalgarno sequence and for the sequence complementary to the SD sequence.
- Gel retardation assays show the formation of the complex between RsaI and *icaR* full-length, *icaR* SUBST, *icaR*-5'UTR, and *icaR*-3'UTR. The 5' end-labeled of RsaI was incubated with increasing concentrations of the various mRNAs. UTR is for untranslated region, and SUBST stands for the substitution of UCCCCUG sequence by AGGGGAC. This mutation, which is located in the 3'UTR of *icaR*, significantly destabilizes the long-range interaction and enhances *icaR* translation. Lane C represents binding between radiolabelled RsaI and full-length *icaR* mRNA (50 nM). Due to the rather short migration of the gel, the full-length mRNA is still observed in the pocket.
- In vivo* effect of RsaI on PIA-PNAG synthesis in the *S. aureus* wild-type (WT) 132 strain, the Δ *rsaI* mutant, and the strain carrying a deletion of *icaR* 3'UTR (Δ 3'UTR). This last mutant strain has been transformed with the pE5 plasmids expressing *rsaI* or *rsaI* mut5. The PIA-PNAG exopolysaccharide biosynthesis was quantified using dot-blot assays. Serial dilutions (1/5) of the samples were spotted onto nitrocellulose membranes, and PIA-PNAG production was detected with specific anti-PIA-PNAG antibodies. RsaI was detected in the same samples by Northern blot using a probe directed against RsaI. Ethidium bromide staining of rRNA was used as loading controls of the same gel.

Source data are available online for this figure.



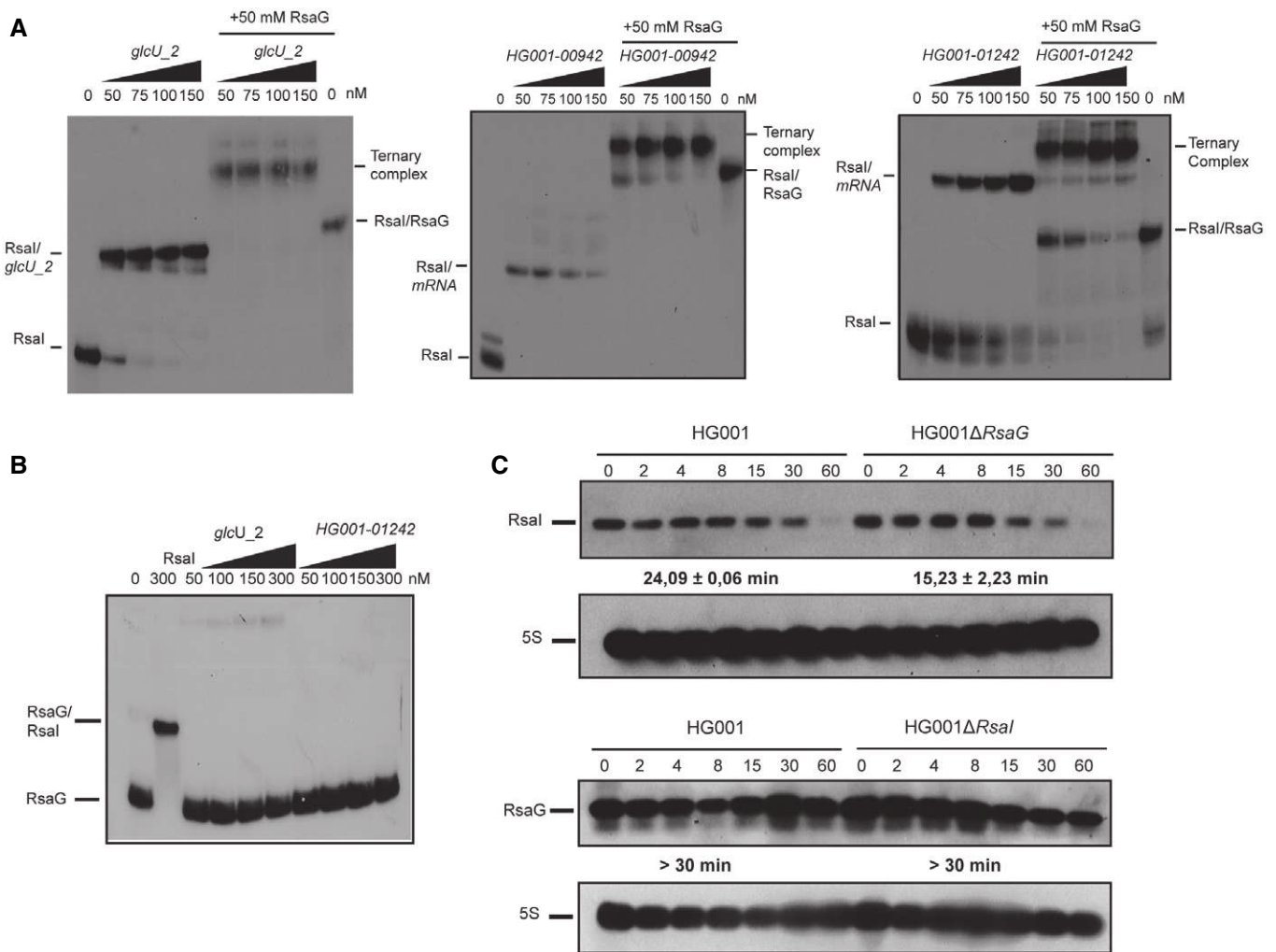


Figure 5. Regulatory function of RsaI is not impaired by binding of RsaG.

A Ternary complex formation between RsaI, various mRNAs (*glcU_2*, *HG001-00942* and *HG001-01242*), and RsaG. The 5' end-labeled RsaI was incubated with increasing concentrations of the target mRNA alone or in the presence of 50 nM of RsaG. The various complexes are notified on the sides of the autoradiographies.

B RsaG does not form stable complexes with *glcU_2* and *HG001-01242*. Binding assays were done in the presence of 5' end-labeled RsaG and either cold RsaI (300 nM) or increasing concentrations of *glcU_2* and *HG001-01242*.

C Measurements of the half-lives of RsaI and RsaG in *HG001-ΔrsaG* or *HG001-ΔrsaI* mutant strains using Northern blot experiments. The cells were treated with rifampicin at 4 h of growth, and total RNAs were extracted after 2, 4, 8, 15, 30, and 60 min at 37°C in BHI. 5S rRNA was probed to quantify the yield of RNAs in each lane using the same samples, which were however run on two different gels. Calculated half-lives are shown beneath the autoradiographies and are the average of 2 experiments. The data were normalized to 5S rRNA.

Source data are available online for this figure.

ΔrsaG mutant strains. The expression of the fusion was strictly identical in both strains. The overexpression of RsaI WT from a plasmid causes a weak but specific inhibition of β-galactosidase synthesis because the RsaI-dependent repression was alleviated in the strain transformed with a plasmid expressing RsaI mut4 (Fig EV4D). This experiment showed that the expression of RsaG does not significantly impact the synthesis of one of the RsaI target.

Using rifampicin treatment, the half-lives of RsaI and RsaG were measured in the WT and the mutant strains (Fig 5C). Quantification of the autoradiographies showed that the two RNAs are highly stable as it was previously shown (Geissmann et al, 2009). However, the stability of RsaI was reproducibly higher in the WT strain (24 min) than in the *ΔrsaG* mutant strain (15 min; Fig 5C).

Taken together, the data show that the synthesis of RsaG and RsaI is regulated in response to different but related carbon sources, G-6P and glucose, respectively. In addition, their binding has no major consequences on the RsaI target recognition and regulation, although RsaG has a slight effect on RsaI stability.

RsaD, the other sRNA target candidate of RsaI, responds to nitric oxide

The signaling pathway of RsaD, another potential sRNA partner of RsaI, has been studied. RsaD binds to RsaI *in vitro* albeit with a lower affinity than RsaG. Although RsaD was significantly enriched in the MAPS experiments using the wild-type MS2-RsaI (Table EV2),

the data were unfortunately not reliable with the mutant MS2-RsaI mut2 and mut4. This is most probably due to the fact that RsaD is weakly expressed at the stationary phase of growth in BHI medium as revealed by the transcriptomic analysis (Table EV6).

Upstream the *rsaD* gene, we identified a conserved motif AGTGACAA that has been described as a binding site for the response regulator SrrA (Pragman *et al*, 2004; Kinkel *et al*, 2013). SrrAB is a major two-component system affecting the temporal expression of the virulence factors (Pragman *et al*, 2004) while it promotes resistance to nitrosative stress and hypoxia (Richardson *et al*, 2006; Kinkel *et al*, 2013). We demonstrated that the expression of RsaD drops considerably in a Δ srrAB mutant strain (Fig EV4E) while its expression is not under the control of the two-component system SaeRS, another global regulator of virulence (Fig EV4E; Liu *et al*, 2016). Because SrrAB is able to sense and respond to nitric oxide (NO) and hypoxia (Kinkel *et al*, 2013), we tested the effect of NO on RsaD synthesis by adding diethylamine-NONOate, as it was previously described for RsaE (Durand *et al*, 2015). We observed a significant and reproducible enhanced expression of RsaD in the WT HG001 strain about 10 min after the addition of diethylamine-NONOate to BHI medium (Fig EV4E). RsaD expression decreased after 20 min due to the short half-life of diethylamine-NONOate. Noteworthy, RsaE, which was significantly enriched together with MS2-RsaI, was also shown to be under the control of SrrAB (Durand *et al*, 2015).

These data suggested that through the binding of sRNAs, a network of sRNAs would coordinate sugar metabolism pathways, energy production, and NO stress responses.

The effect of RsaI on global gene expression in *Staphylococcus aureus*

Before to perform the comparative transcriptomic analysis, we first analyzed whether the deletion or the overexpression of RsaI might cause growth phenotype. Unexpectedly, we did not succeed to overexpress RsaI from a strong promoter in the WT HG001 strain. This result was not due to technical problems as we managed to express RsaI under the control of its own promoter on the same plasmid. Hence, the growth phenotypic experiment was performed in the 132 strain. The WT 132 and the Δ rsaI mutant strains harbor a similar growth curve in BHI medium (Fig EV5A). However, when RsaI was expressed under the control of a strong promoter, a slight but reproducible delay in growth was observed only in the WT 132 strain (Fig EV5A). Northern blot analysis was performed to measure the steady-state levels of RsaI during growth in the various strains (Fig EV5B). Quantification of the autoradiographies and normalization with 5S rRNA showed that the highest yields of RsaI were observed in the WT 132 strain expressing RsaI from the strong promoter, and only in this context, a significant amount of RsaI was detected at the beginning of growth (Fig EV5B).

In order to get a global overview of RsaI impact on gene regulation, a comparative transcriptomic analysis was performed on total RNAs extracted from the WT HG001 strain, the isogenic HG001 Δ rsaI mutant strain, and the same mutant strain complemented with a plasmid expressing RsaI under the control of its own promoter in order to avoid side effects of a too strong expression of RsaI on growth (Tables EV6 and EV7, and Fig EV5C). The cultures were done in triplicates with high reproducibility in BHI medium at 37°C until 6 h, under the conditions allowing the

expression of RsaI (Fig 1). We have considered a gene to be regulated by RsaI if the ratio between two strains was at least twofold. Changes in gene expression were more pronounced between the mutant Δ rsaI versus the same strain expressing RsaI from a plasmid rather than between the mutant Δ rsaI versus the WT strain (Table EV7). This effect is most likely due to the higher expression of RsaI in the complemented strain in comparison with the WT strain (Fig EV5C). The levels of 26 and 50 mRNAs were significantly decreased and enhanced, respectively, when the complemented strain was compared to the mutant Δ rsaI (Table EV7). The complete data set is presented in Table EV6. Beside genes encoding phage-related proteins, most of the RsaI-dependent activation was observed for genes involved in fermentation processes, in NO resistance, and in energy-generating processing (in red in Table EV7). In addition, weaker effects were also observed for mRNAs encoding proteins involved in iron-sulfur cluster repair (ScdA), in NO detoxification (qoxABCD, hmp), and in various metabolic pathways. The level of RsaG was also enhanced in the complemented strain. Conversely, the overexpression of RsaI caused a reduced expression of genes that are functionally related. Several of them are involved in glycolysis (*fba1*) and pentose phosphate pathway (*gnd*, *tkt*), in thiamine co-factor synthesis (*thiW*, *tenA*), and in arginine catabolism (*arcABCD*, *arcR*) (in blue in Table EV7). Additionally, significant repression was also observed for *miaB* encoding tRNA-specific modification enzyme, *tyrS* encoding tryptophanyl-tRNA synthetase, *ebpS* encoding the cell surface elastin binding protein, and to a weaker extent for *rex* encoding redox-sensing transcriptional regulator. Interestingly, the MAPS approach revealed that several of these mRNAs (i.e., *qoxABCD* operon, *tyrS*, and *plfA-plfB*) were also enriched together with RsaI (Table EV2). Surprisingly, the most enriched RsaI targets identified by MAPS (i.e., *icaR*, *glcU_2*, and *fn3K*) did not show significant mRNA level variations when RsaI was deleted or overexpressed suggesting that RsaI would primarily regulate their translation (Table EV1).

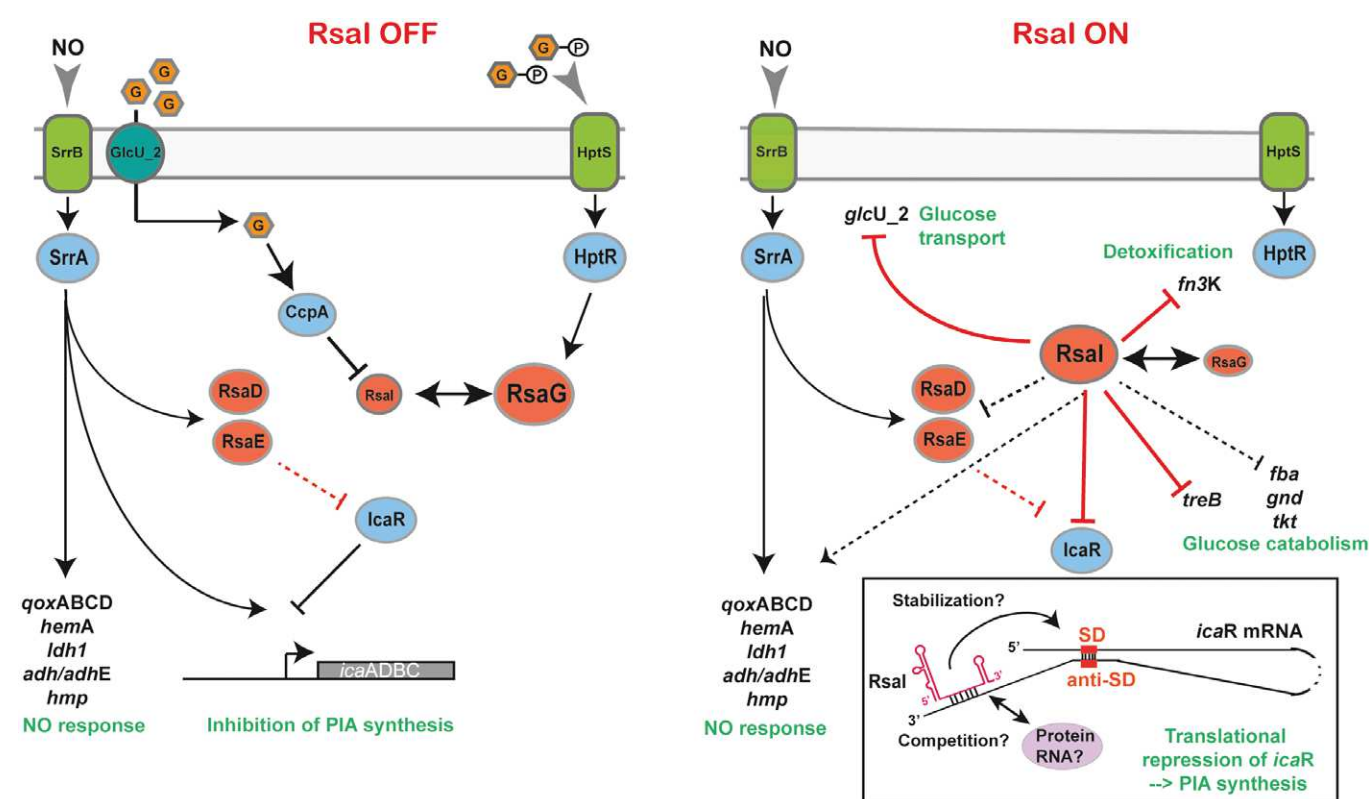
These data showed that high concentrations of RsaI affected the mRNA levels of several enzymes involved in sugar metabolism, in arginine catabolism, in the pentose phosphate pathway, and in various processes linked to NO detoxification, energy production, and fermentation. However, the small overlap between the two approaches suggested that many of the RsaI-dependent effects as revealed in the transcriptomic data resulted from indirect effects.

Discussion

In this work, we have investigated the cellular functions of one abundant sRNA, called RsaI (or RsaOG), which is highly conserved among *Staphylococcaeae* (Geissmann *et al*, 2009; Marchais *et al*, 2010). In contrast to many sRNAs that contained C-track motifs, this RNA has the particularity to carry two conserved G-rich sequences and a large unpaired CU-rich interhelical region (Fig 2). This RNA was proposed to fold into a pseudoknot structure involving the two conserved regions limiting the access of regulatory regions (Marchais *et al*, 2010). Combining several global approaches (MAPS, transcriptomic analysis), our data show that RsaI is linked to a large regulon involved in sugar uptake and metabolism,

DNA-binding affinity to the *cre*-binding site to repress or activate the target genes. RsaI repression requires the presence of CcpA, most probably through the binding at a *cre* motif located at -45/-30 (GGAAAACGCTTACAT) from the *rsaI* transcriptional start site (Figs 1 and EV1B). Interestingly, deletion of *ccpA* gene affected vancomycin resistance (Seidl *et al*, 2009) and recent observations showed that sub-inhibitory treatment of cells with vancomycin led to an enhanced expression of RsaI (Howden *et al*, 2013). We showed that the repression of RsaI is alleviated as soon as the concentration of glucose is strongly reduced. Therefore, RsaI might be a signature of a metabolic switch of the bacterial population. Using the MAPS approach and gel retardation assays, we could identify several mRNAs that strongly bind to RsaI with its long unpaired and conserved CU-rich region. Binding of RsaI to *glcU*_2 encoding a glucose permease and to *fn3K* encoding the fructosamine 3 kinase hindered the formation of the initiation ribosomal complex. Because their steady-state yields are not affected, we suggest that both mRNAs are regulated at the translational level. We propose that this mechanism is most likely common to all mRNAs that could form base-pairings between their ribosome-binding sites and the CU-rich region of RsaI. This is the case for mRNAs encoding a

The cellular level of RsaI is tightly controlled during growth in rich BHI medium. In this study, we showed that RsaI expression is strongly and rapidly repressed at the transcriptional level through the activity of CcpA in response to glucose availability. Carbon catabolite repression is a universal regulatory phenomenon that allows the cells to use the preferred carbon source to produce energy and to provide the building blocks for macromolecules. Concomitantly, it represses genes that are involved in the metabolism of less-preferred carbon sources. To do so, CcpA has to be activated through a cascade of events involving its co-regulator histidine-containing phosphocarrier protein (HPr), which has been phosphorylated by its cognate kinase/phosphorylase HprK/P activated in the presence of glycolytic intermediates. It is thought that binding of the phosphorylated HPr to CcpA enhances its



In the presence of glucose or glucose 6-phosphate (left panel), the expression of RsaI is inhibited, in contrast to RsaG. In the absence of glucose or when glucose is metabolized (right panel), repression of RsaI is alleviated to regulate its RNA targets. In blue, are the transcriptional protein regulators, and in red, the regulatory sRNAs. Gray arrows represent functional links, black arrows are for activation, and bars for repression. Red lines corresponded to post-transcriptional regulation and black lines to transcriptional regulation. Dotted lines represented the regulatory events for which the regulation is not yet demonstrated. A schematic view of the post-transcriptional *icaR* mRNA regulation is represented in the insert. The 3'UTR of *icaR* contains an anti-Shine and Dalgarno (anti-SD) sequence able to bind the SD sequence in the 5'UTR (Ruiz de los Mozos *et al.*, 2013). The CU-rich unpaired sequence of RsaI (in red) binds to the 3'UTR of *icaR* mRNA downstream of the anti-SD and represses the translation indirectly either by stabilizing the interaction between the two UTRs or by preventing the action of trans-acting activators (protein or RNA).

transcriptional regulator of the XRE family, the PTS system trehalose-specific EIIBC component TreB, a peptidase, a cell wall-binding lipoprotein, and DegV containing protein (Table EV2). Noteworthy, *in vivo*, most of these mRNAs were no more found enriched using MS2-RsaI mut4 in which the CU-rich region has been deleted (Table EV5). Among these proteins, two of them are directly involved in sugar uptake and metabolism (GlcU_2, TreB). In *S. aureus*, the PTS (phosphotransferase system)-dependent and PTS-independent transport systems ensure efficient glucose transport (reviewed in Götz *et al*, 2006). If a rapidly metabolizable sugar (such as glucose) is used during growth at a rather low concentration, the transport will occur via the PTS system, and concomitantly, the carbon catabolite repression system will be activated through CcpA protein. At high concentration of glucose, it is assumed that the sugar will be transported by the PTS system and in addition by the permease GlcU_2. The glucose transported by the permease will be phosphorylated within the cell by the glucose kinase GlkA. Hence, it is reasonable to propose that RsaI would repress the synthesis of GlcU_2 when the glucose concentration dropped. In addition, RsaI represses the expression of *treB* mRNA most probably through base-pairings with its SD sequence. TreB is the enzyme specific for trehalose, a diholoside which is transported exclusively by PTS (Bassias & Brückner, 1998). Transcriptomic analysis also revealed that RsaI strongly repressed various key enzymes involved in glucose catabolism pathway such as fructose-biphosphate aldolase (*fba1*), 6-phosphogluconate dehydrogenase 2C decarboxylase (*gnd*), and the thiamine-dependent enzyme transketolase (*tkt*). Additionally, we showed that RsaI represses the synthesis of fructosamine 3-kinase (Fn3K), which deglycated products of glycation formed from ribose 5-phosphate or erythrose 4-phosphate produced by the pentose phosphate pathway (Gemayel *et al*, 2007). This enzyme is part of a repair machinery to protect the cells from damages caused by glycation as the results of high glucose concentrations (Deppe *et al*, 2011). The pentose phosphate pathway is also an alternative route for glucose metabolism and provides the source of pentose phosphates necessary for nucleotide synthesis. Although it is not known whether RsaI regulated their expression in a direct or indirect manner, base-pairings were predicted between RsaI and the ribosome-binding site of *fba1* and *tkt* mRNAs. The transketolase is a key enzyme of the pentose phosphate pathway, which requires thiamine diphosphate as a co-factor. Interestingly, the thiamine pathway is also repressed in strain expressing RsaI (Table EV6). Although, we have observed very similar pathways that are deregulated by RsaI using the MAPS and RNA-seq approaches, not so many overlaps were identified. It is possible that the conditions of the MAPS approach performed on the WT strain expressing all the ribonucleases have preferentially enriched the mRNAs that are regulated at the translational level. Therefore, it is tempting to deduct that RsaI would inhibit the synthesis of the major permease of glucose uptake, of enzymes involved in the glycolysis, of unnecessary enzymes involved in detoxification of high glucose concentration, and of the pentose phosphate pathway when glucose concentration decreases (Fig 6).

RsaI interacts with sRNA containing C-rich motifs

The MAPS approach revealed that several sRNAs (RsaG, RsaD, RsaE) were enriched significantly with RsaI. We demonstrated here

that the second G-track sequence located in the first hairpin domain of RsaI is able to form a highly stable complex with RsaG and to a lesser extent with RsaD. Additionally, mRNA targets interacting with the CU-rich domain of RsaI did not disturb the binding of RsaG suggesting that the two functional domains of RsaI are independent. Conversely, preliminary data suggested that the apical loop of the first hairpin of RsaG contains the C-rich motif that is recognized by RsaI, but which is also used to regulate the expression of several mRNAs (Desgranges E, Marzi S, Romby P, Caldelari I, unpublished data). RsaG is part of the 3'UTR of *uhpT* mRNA, and its expression was strongly enhanced by the extracellular concentration of G-6P. Under these conditions, the levels of RsaI are much lower than for RsaG (Fig EV4B). These two sRNAs are thus involved in pathways related to the use of the preferred carbon sources. Indeed, RsaI negatively controls glucose uptake when glucose is consumed or absent from the medium while RsaG responds to the extracellular concentration of G-6P. Although the functions of RsaG remain to be addressed, we hypothesized that the sRNA might regulate either the expression of unnecessary genes, or of genes involved in sugar metabolism, or of genes required to protect cells against damages linked to sugar-phosphate uptake and metabolism (Bobrovskyy & Vanderpool, 2016). Noteworthy, RsaI sequence is conserved in the *Staphylococcus* genus while RsaG is only conserved in the *S. aureus* species. Our data suggested that RsaG has no major impact on the regulation of RsaI targets. Its slight stabilizing effect on RsaI strongly supports the MAPS data showing that the two sRNAs interact *in vivo* and potentially would form ternary complexes in the presence of the RsaI targets. Whether RsaI might regulate the functions of RsaG remained to be addressed.

In addition, the overproduction of RsaI induced several changes into the transcriptome of *S. aureus* that resembled the regulon of the two-component system SrrAB, which was demonstrated as the essential system responding to both NO and hypoxia (Kinkel *et al*, 2013). The SrrAB regulon has also been shown to confer to the cells the ability to maintain energy production, to promote repair damages, and NO detoxification (Richardson *et al*, 2006; Kinkel *et al*, 2013). We observed here that RsaI enhances the expression of genes encoding cytochrome biosynthesis (*qoxABCD*), as well as genes involved in anaerobic metabolism (*pflAB*, *ldh1*, *focA*, *adh*), in iron-sulfur cluster repair (*scdA*), and most importantly in NO detoxification (*qoxABCD*, *hmp*) and NO resistance (*ldh1*). These effects are most likely indirect and might be due to the interaction between RsaI and the SrrAB-dependent sRNAs RsaD and RsaE, which both contained a typical *srrA* site upstream their genes. These two sRNAs present a C-rich sequence that can potentially form base-pairings with the G-track sequences of RsaI (Table EV1), but the formation of complexes with RsaI is less efficient than with RsaG. It cannot be excluded that an RNA-binding protein might be required in these cases to stabilize the base-pairings. We also do not exclude that the two sRNAs were pooled down because they might share similar mRNA targets with RsaI. RsaD and RsaE are both upregulated by the presence of NO in the cellular medium (Durand *et al*, 2015; Fig EV4D). Although the functions of RsaD require additional studies, *S. aureus* RsaE was previously shown to coordinate the down-regulation of numerous metabolic enzymes involved in the TCA cycle and the folate-dependent one-carbon metabolism (Geissmann *et al*, 2009; Bohn *et al*, 2010) and of enzymes involved in arginine degradation pathway (Rochat *et al*, 2018). Additionally, in

B. subtilis, the homologue of RsaE called RoxS is under the control of the NADH-sensitive transcription factor Rex, and the Rex binding site is also conserved in *S. aureus* *rsaE* gene (Durand *et al.*, 2017). Hence, it was proposed that RsaE downregulated several enzymes of the central metabolism under non-favorable conditions and in addition would contribute to readjust the cellular NAD⁺/NADH balance under stress conditions (Bohn *et al.*, 2010; Durand *et al.*, 2015, 2017).

Many of the RsaI-dependent effects, which have been monitored by the transcriptomic analysis, are indirect, and we propose that some of these effects might result from the interaction of RsaI with sRNAs when the preferred carbon source became scarce.

Physiological consequences of RsaI regulation

Our data suggested that several sRNAs in *S. aureus* are part of intricate regulatory networks to interconnect in a dynamic manner various metabolic pathways following sugar metabolism and uptake. The concept that sRNAs are key actors to coordinate the regulation of metabolic enzymes has been largely demonstrated for *Enterobacteriaceae* (reviewed in Papenfort & Vogel, 2014). These sRNA-dependent regulations often induced significant growth phenotypes in response to the availability of carbon sources and nutrient. For instance, *E. coli* and *Salmonella* SgrS contribute to stress resistance and growth during glucose–phosphate stress by inhibiting synthesis of sugar transporters or activating dephosphorylation and efflux of sugars (e.g., Vanderpool & Gottesman, 2004, 2007; Kawamoto *et al.*, 2005; Rice *et al.*, 2012; Papenfort *et al.*, 2013). The repertoire of SgrS targets has been recently expanded to mRNAs involved in key metabolic pathways, which allow restoring metabolic homeostasis during sugar–phosphate stress and growth recovery (Richards *et al.*, 2013; Bobrovskyy & Vanderpool, 2016). *Salmonella* GcvB regulon is required for growth if peptides represent the unique carbon source (Miyakoshi *et al.*, 2015), while *E. coli* Spot42 is important for optimal utilization of carbon sources and its overproduction inhibited growth on several non-preferred carbon sources (Møller *et al.*, 2002; Beisel & Storz, 2011). In *S. aureus*, the overproduction of RsaE induces a growth defect, which was partially alleviated by acetate (Bohn *et al.*, 2010). These examples illustrate how the yields of these sRNAs should be tightly controlled during growth in order to optimize the use of the preferred carbon sources, to restore metabolic homeostasis during stress, and to avoid cell damages or metabolites depletion caused by the intracellular production of glucose–phosphate. Interestingly, in *Bacillus subtilis*, the sRNA SR1 was found repressed under glycolytic conditions mainly by CcpN and to a lesser extent by CcpA (Licht *et al.*, 2005; Heidrich *et al.*, 2007; Gimpel *et al.*, 2012). SR1 expression is induced by L-arginine and acts as a main regulator or arginine catabolism through the translational repression of *ahrC* mRNA (Heidrich *et al.*, 2007). In this case, the deletion or the overexpression of SR1 did not affect growth (Licht *et al.*, 2005). The deletion of *rsaI* gene does not cause a growth phenotype, and only high levels of RsaI constitutively expressed from a plasmid in the WT strain cause a weak growth delay (Fig EV5A). Additional works will be required to identify the growth conditions under which the function of RsaI can be studied in a more relevant way.

One possible track resides in the fact that RsaI interconnects various metabolic pathways. By acting as a post-transcriptional

regulator, it could play an important role during the infection process. *S. aureus* has the ability to generate infections through the colonization of many different metabolic host niches. Several studies have shown that both glycolysis and gluconeogenesis systems are mandatory for the infection process, and moreover, *S. aureus* appears to be resistant to NO radicals that are heavily produced by the macrophages. Interestingly, glycolysis is an important process that contributes to persist within the macrophages and to protect the intracellular bacteria against NO (Vitko *et al.*, 2015). However, if the bacteria escape the macrophages or lyse the host cells, *S. aureus* is thought to form aggregates at the center of highly inflamed and hypoxic abscesses. Under these conditions, the host cells consumed a large amount of glucose to fight the inflammation. Hence, glucose will become scarce for *S. aureus* suggesting that lactate and amino acids derived from the host might be used as the major sources of carbon to enter gluconeogenesis (Richardson *et al.*, 2015). These conditions favored the activation of the two-component system SrrAB, which in turn activates genes required for anaerobic metabolism, cytochrome, and heme biosynthesis, and NO radical detoxification should play an essential role in the survival of the bacteria (Kinkel *et al.*, 2013). Because the CcpA-dependent repression of RsaI is alleviated under conditions where the glucose is strongly reduced, and because many SrrAB-dependent genes are also induced when RsaI is expressed at high levels, it is thus expected that RsaI might also contribute to metabolic adaptations of the cells to the dynamic nature of the host immune environment. Besides, it is also tempting to propose that RsaI might be involved in the dormant state of bacterial cells while environmental conditions are unfavorable (Lennon & Jones, 2011). Interestingly enough, we also observed a RsaI-dependent activation of the expression of the mRNA encoding EsxA, a type VII-secreted virulence factor required for the release of the intracellular *S. aureus* during infection (Korea *et al.*, 2014). We also showed here that RsaI activates the synthesis of the PIA-PNAG exopolysaccharide most probably by decreasing the translation of IcaR, which is the transcriptional repressor of *icaADBC* operon (Figs 4 and 6). We demonstrated that RsaI base-pairs *icaR* mRNA without interfering the interaction between *icaR* 3'UTR and 5'UTR (Fig EV3). Although the precise molecular mechanism is not yet defined, we propose that RsaI would inhibit the synthesis of IcaR by stabilizing the circularization of the mRNA and/or by counteracting with the binding of an activation factor of *icaR* mRNA translation (Fig 6). Interestingly, RsaE, which is another RsaI target, also binds *icaR* mRNA at two distinct positions close to the SD sequence and in the coding sequence through C-rich motifs (Rochat *et al.*, 2018). These motifs are almost identical to the anti-SD sequence (UCCCC) comprised in the *icaR* 3'UTR (Ruiz de los Mozos *et al.*, 2013). This opens the possibility of a competition for hindering the *icaR* SD by RsaE or the 3'UTR. Thus, RsaI and RsaE might interact individually or simultaneously with *icaR* mRNA to repress its expression in response to changes in cellular medium of glucose and NO. Though further investigations are needed to decipher the physiological consequences of such complex interactions, it indicates that the regulation of the synthesis of the PIA-PNAG is rather complexed and multifactorial and is tightly controlled according to the metabolic state of the bacterial population (Cue *et al.*, 2012). The signals modifying the yield of these RNAs would be crucial to determine the levels of IcaR repressor and PIA-PNAG synthesis. We also do not exclude that RsaI might regulate another regulator of the synthesis

of PIA-PNAG. For instance, inactivation of the TCA cycle resulted in a massive derepression of the PIA-PNAG biosynthetic enzymes to produce the exopolysaccharide (Sadykov *et al*, 2008), and glucose enhances PIA-PNAG-dependent biofilm formation (You *et al*, 2014), while SrrAB appears as an inducer of PIA-PNAG-dependent biofilm (Ulrich *et al*, 2007). It remains to be elucidated if the biofilm induction mediated by SrrAB occurs through activation of RsaE that might promote interaction with *icaR* mRNA and/or RsaI. Hence, *icaR* can be considered as a hub gene, integrating both transcriptional regulatory networks and sRNA-mediated post-transcriptional regulatory signals to control biofilm formation in *S. aureus*.

Although our study shed light on the regulatory activities of this multifaceted sRNA, there is still much to be learned on how RsaI and other sRNAs can be integrated into the networks regulating the metabolic pathways that are essential for *S. aureus* biofilm formation, survival, persistence, and invasion within the host.

Materials and Methods

Plasmids and strains constructions

All strains and plasmids constructed in this study are described in Table EV8. The oligonucleotides designed for cloning and for mutagenesis are given in Table EV3. *Escherichia coli* strain DC10B was used as a host strain for plasmid amplification before electroporation in *S. aureus*. Plasmids were prepared from transformed *E. coli* pellets following the NucleoSpin Plasmid kit protocol (Macherey-Nagel). Transformation of both *E. coli* and *S. aureus* strains was performed by electroporation (Bio-Rad Gene Pulser).

The *rsaI* deletion mutant was constructed by homologous recombination using plasmid pMAD in *S. aureus* HG001 and 132 (Arnaud *et al*, 2004). The deletion comprises nucleotides 2,376,101–2,376,306 according to HG001 genome (Caldelari *et al*, 2017). Experimental details are given in supplementary materials.

The vector pCN51::PrsaI was constructed by ligating a PCR-amplified fragment (Table EV3) containing *rsaI* with 166 pb of its promoter region and digested by *SphI* and *PstI* into pCN51 vector digested with the same enzymes. The vector pUC::T7rsaI was constructed by ligating a PCR-amplified fragment (Table EV3) containing *rsaI* with the T7 promoter sequence and digested by *EcoRI* and *PstI* into pUC18 vector digested with the same enzymes. Mutagenesis of pUC::T7rsaI was performed with QuikChange XL Site-directed mutagenesis (Stratagene) leading to pUC::T7rsaI mut1, 2, 3, and 4 (Table EV3). To obtain the plasmid pCN51::P3::MS2-rsaI, a PCR product containing the MS2 tag fused to the 5' end of *rsaI* was cloned into pCN51::P3 by digestion of both PCR fragments and of the plasmid by *PstI/BamHI* (Table EV3). Plasmids from positive clones were sequenced (GATC Biotech) before being transformed in DC10B, extracted, and electroporated into *S. aureus* strains. The plasmids pCN51::P3::MS2-rsaI mut2 and mut4 were generated by QuikChange mutagenesis as above. Construction of plasmids pES::rsaI and pES::rsaI mut5 is described in supplementary materials.

Growth conditions

Escherichia coli strains were cultivated in Luria-Bertani (LB) medium (1% peptone, 0.5% yeast extract, 1% NaCl) supplemented with

ampicillin (100 µg/ml) when necessary. *Staphylococcus aureus* strains were grown in brain-heart infusion (BHI), tryptic soy broth (TSB), or Muller Hinton broth (MHB) media (Sigma-Aldrich) supplemented with erythromycin (10 µg/ml) for plasmid selection. When needed, MHB was complemented with either 5 g/l of glucose 6-phosphate or 1 g/l of glucose, fructose, or xylose (Sigma-Aldrich). NO production was induced by addition of 100 µM Na-diethylamine-NONOate (Sigma-Aldrich) as previously described (Durand *et al*, 2015).

Preparation of total RNA extracts

Total RNAs were prepared from *S. aureus* cultures taken at different times of growth. After centrifugation, bacterial pellets were resuspended in 1 ml of RNA Pro Solution (MP Biomedicals). Lysis was performed with FastPrep and 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 (500 ng) was performed on 1.5% agarose gel containing 25 mM guanidium thiocyanate. After migration, RNAs were vacuum-transferred on nitrocellulose membrane. Hybridization with specific digoxigenin (DIG)-labeled probes complementary to RsaI, RsaG, RsaD, 5S, *ccpA* sequences, followed by luminescent detection, was carried out as previously described (Tomasini *et al*, 2017).

MAPS experiments, and transcriptomic and RNA-seq analysis

Crude bacterial extract was prepared and purified by affinity chromatography as previously described (Tomasini *et al*, 2017). The eluted 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. The experiments were carried out with the tagged wild-type RsaI and two mutant forms (mut2 and mut4). RNAs were treated to deplete abundant rRNAs, and the cDNA libraries were performed using the ScriptSeq complete kit (bacteria) from Illumina. The libraries were sequenced using Illumina MiSeq with a V3 Reagent kit (Illumina), which preserves the information about the orientation of the transcripts and produces reads of 150 nts, which map on the complementary strand. Each RNA-seq was performed at least in duplicates. The reads were then processed to remove adapter sequences and poor-quality reads by Trimmomatic (Bolger *et al*, 2014), converted to the FASTQ format with FASTQ Groomer (Blankenberg *et al*, 2010), and aligned on the HG001 genome (Caldelari *et al*, 2017) 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. All processing steps were performed using the Galaxy platform (Afgan *et al*, 2016).

For total RNA extracts and MS2-eluted RNAs, DNase I (0.1 U/µl) treatment was performed 1 h at 37°C. The reactions mixtures were then purified by phenol:chloroform:isoamylalcohol 25:24:1 (v/v) and subsequent ethanol precipitation. RNA pellets were resuspended

in sterile milliQ water. RNA was quantified by Qubit (Life Technologies), and the integrity was assessed with the Bioanalyzer (Agilent Technologies). For transcriptomic, 1 µg of total RNA was ribo-depleted with the bacterial RiboZero kit from Illumina. The TruSeq total RNA-stranded kit from Illumina was used for the library preparation. Library quantity was measured by Qubit, and its quality was assessed with a TapeStation on a DNA High sensitivity chip (Agilent Technologies). Libraries were pooled at equimolarity and loaded at 7 pM for clustering. The 50 bases oriented single-read sequencing was performed using TruSeq SBS HS v3 chemistry on an Illumina HiSeq 2500 sequencer.

Preparation of RNAs for *in vitro* experiments

Transcription of RsaI, RsaI mutants, and RsaG was performed using linearized pUC18 vectors. PCR fragments containing the 5'UTR of the coding region of selected mRNA targets were directly used as templates for *in vitro* transcription using T7 RNA polymerase. The RNAs were then purified using a 6% or 8% 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 labeling of the 5' end of dephosphorylated RNAs (RsaI/RsaI mutants) and DNA oligonucleotides was performed with T4 polynucleotide kinase (Fermentas) and [γ - 32 P] ATP. Before use, cold or labeled 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 MgCl₂, 60 mM KCl, 1 mM DTT).

Gel retardation assays

Radiolabeled purified RsaI or RsaI mutants (50,000 cps/sample, concentration < 1 pM) and cold mRNAs were renatured separately as described above. For each experiment, increasing concentrations of cold mRNAs were added to the 5' end-labeled wild-type or RsaI mutants in a total volume of 10 µl containing the ToeP⁺ buffer. Complex formation was performed at 37°C for 15 min. After incubation, 10 µl of glycerol blue was added and the samples were loaded on a 6% or 8% polyacrylamide gel under non-denaturing conditions (300 V, 4°C). Under these conditions where the concentration of the labeled RNA is negligible, the K_D dissociation constant can be estimated as the concentration of the cold RNA that showed 50% of binding.

Toe-printing 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 previously described (Fechter *et al.*, 2009). Increasing concentrations of RsaI were used to monitor their effects on the formation of the initiation complex with *glcU*₂ and *fn3K* mRNAs.

In vivo β-galactosidase assays

Translation fusions were constructed with plasmid pLUG220, a derivative of pTCV-*lac*, a low-copy-number promoter-less *lacZ* vector, containing the constitutive *rpoB* promoter (Table EV8). The

whole leader regions of *glcU*₂ (−54/+99), *fn3K* (−33/+99), *treB* (−23/+99), HG001_01242 (−34/+99), and HG001_02520 (−71/+99) (Table EV3) were cloned downstream the *PrpoB* in frame with *lacZ*. The whole gene encoding RsaI or RsaI mut4 (obtained synthetically from IDT) with 166 bp of its promoter region was digested with *PstI* and ligated at the unique *PstI* of pLUG220::PrpoB vector. The final constructs were transformed into the strain HG001Δ*rsaI*. β-galactosidase activity was measured four times as previously described (Tomasini *et al.*, 2017).

PIA-PNAG quantification

Cell surface PIA-PNAG exopolysaccharide levels were monitored according to Cramton *et al.* (1999). Overnight cultures were diluted 1:50 in TSB-3% NaCl, and bacteria were grown at 37°C. Samples were extracted at 6 h after inoculation. The same number of cells of each strain was resuspended in 50 µl of 0.5 M EDTA (pH 8.0). Then, cells were incubated for 5 min at 100°C and centrifuged 17,000 g for 5 min. Supernatants (40 µl) were incubated with 10 µl of proteinase K (20 mg/ml) (Sigma) for 30 min at 37°C. After the addition of 10 µl of the buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.01% bromophenol blue), serial dilutions 1:25 were performed in the same buffer. Then, 10 µl was spotted on a nitrocellulose membrane using a Bio-Dot microfiltration apparatus (Bio-Rad). The membrane was blocked overnight with 5% skimmed milk in phosphate-buffered saline (PBS) with 0.1% Tween 20 and incubated for 2 h with specific anti-PNAG antibodies diluted 1:20,000 (Maira-Litrán *et al.*, 2005). Bound antibodies were detected with peroxidase-conjugated goat anti-rabbit immunoglobulin G antibodies (Jackson ImmunoResearch Laboratories, Inc., Westgrove, PA) diluted 1:10,000, using the SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific).

Data availability

The transcriptomics and MAPS data have been deposited in NCBI's Gene Expression Omnibus (Barrett *et al.*, 2013) and are accessible through the GEO Series accession number GSE122092 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE122092>).

Expanded View for this article is available online.

Acknowledgements

We thank David Lalaouna for critical reading of the manuscript and helpful discussions, Thomas Geissmann for helpful advices, Marie Beaume for the construction of the mutant strain HG001-Δ*rsaI*, and Eve-Julie Bonetti and Anne-Catherine Helfer for excellent technical assistance. We are grateful to Joseph Vilardell for the gift of the plasmid expressing the MS2-MBP, Aurélie Hiron and Tarek Masdek for providing us the HG001-Δ*hptRS* and -Δ*srrAB* mutant strains, and Christiane Wolz for the HG001-Δ*ccpA* and -Δ*codY* mutant strain. RNA-Seq analyses have been partially done using the Roscoff (France) instance of Galaxy (<http://galaxy.sb-roscoff.fr/>). This work was supported by the Centre National de la Recherche Scientifique (CNRS) to [P.R.] and by the Agence Nationale de la Recherche (ANR, grant ANR-16-CE11-0007-01, RIBOS-TAPH, to [P.R.]). It has also 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. The work is financed by a "Projet international de coopération

scientifique" (PICS) No. PICS07507 between France and Spain to [I.C.]. D. Bronesky was supported by Fondation pour la Recherche Médicale (FDT20160435025). A. T-A is financed by the Spanish Ministry of Economy and Competitiveness (BFU2014-56698-P) and the European Research Council Consolidator Grant (646869-ReguloBac-3UTR).

Author contributions

DB, ED, and IC performed the MAPS experiments and the validation of the data. CJC, LP, AT-A, and IL constructed several mutant strains and performed the biofilm assays. AC and PF have done the transcriptomic analysis. SM analyzed all the RNA-Seq analysis and quantification of the data. KM and FV contributed to phenotypic analysis of some of the mutant strains. PR and IC were responsible of the project and wrote the manuscript. All authors contributed to the analysis, interpretation of the experiments, and writing of the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Eberhard C, Grüning B, Guerler A, Hillman-Jackson J, Von Kuster G, Rasche E, Soranzo N, Turaga N, Taylor J, Nekrutenko A, Goecks J (2016) The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res* 44: W3–W10
- Anders S, Pyl PT, Huber W (2015) HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* 31: 166–169
- Arnaud M, Chastanet A, Débarbouillé M (2004) New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl Environ Microbiol* 70: 6887–6891
- Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A (2013) NCBI GEO: archive for functional genomics data sets-update. *Nucleic Acids Res* 41: D991–D995
- Bassias J, Brückner R (1998) Regulation of lactose utilization genes in *Staphylococcus xylosus*. *J Bacteriol* 180: 2273–2279
- Beisel CL, Storz G (2011) The base-pairing NA spot42 participates in a multioutput feedforward loop to help enact catabolite repression in *Escherichia coli*. *Mol Cell* 41: 286–297
- Bischoff M, Wonnemberg B, Nippe N, Nyffenegger-Jann NJ, Voss M, Beisswenger C, Sunderkötter C, Molle V, Dinh QT, Lammert F, Bals R, Herrmann M, Somerville GA, Tschernig T, Gaupp R (2017) CcpA affects infectivity of *Staphylococcus aureus* in a hyperglycemic environment. *Front Cell Infect Microbiol* 7: 1–10
- Blankenberg D, Gordon A, Von Kuster G, Coraor N, Taylor J, Nekrutenko A, Team G (2010) Manipulation of FASTQ data with galaxy. *Bioinformatics* 26: 1783–1785
- Bobrovskyy M, Vanderpool CK (2016) Diverse mechanisms of post-transcriptional repression by the small RNA regulator of glucose-phosphate stress. *Mol Microbiol* 99: 254–273
- Bohn C, Rigoulay C, Chabelskaya S, Sharma CM, 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
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120
- Caldelari I, Chane-Woon-Ming B, Noirot C, Moreau K, Romby P, Gaspin C, Marzi S (2017) Complete genome sequence and annotation of the *Staphylococcus aureus* strain HG001. *Genome Announc* 5: e00783-17
- Cramton SE, Gerke C, Schnell NF, Nichols WW, Gotz F (1999) The intercellular adhesion (ica) locus is present in *Staphylococcus aureus* and is required for biofilm formation. *Infect Immun* 67: 5427–5433
- Cue D, Lei MG, Lee CY (2012) Genetic regulation of the intercellular adhesion locus in Staphylococci. *Front Cell Infect Microbiol* 2: 38
- Deppe VM, Bongaerts J, O'Connell T, Maurer KH, Meinhardt F (2011) Enzymatic deglycation of Amadori products in bacteria: mechanisms, occurrence and physiological functions. *Appl Microbiol Biotechnol* 90: 399–406
- Durand S, Braun F, Helfer AC, Romby P, Condon C (2017) sRNA-mediated activation of gene expression by inhibition of 5'-3' exonucleolytic mRNA degradation. *Elife* 6: e23602
- Durand S, Braun F, Lioliou E, Romilly C, Helfer AC, Kuhn L, Quittot N, Nicolas P, Romby P, Condon C (2015) A nitric oxide regulated small RNA controls expression of genes involved in redox homeostasis in *Bacillus subtilis*. *PLoS Genet* 11: e1004957
- Fechter P, Chevalier C, Yusupova G, Yusupov M, Romby P, Marzi S (2009) Ribosomal initiation complexes probed by toeprinting and effect of trans-acting translational regulators in bacteria. *Methods Mol Biol* 540: 247–263
- Geissmann T, Chevalier C, Cros MJ, 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
- Gemayel R, Fortpied J, Rzem R, Vertommen D, Veiga-da-Cunha M, Van Schaftingen E (2007) Many fructosamine 3-kinase homologues in bacteria are ribulosamine/erythrulosamine 3-kinases potentially involved in protein deglycation. *FEBS J* 274: 4360–4374
- Gimpel M, Pries H, Barth E, Gramzow L, Brantl S (2012) SR1-a small RNA with two remarkable conserved functions. *Nucleic Acids Res* 40: 11659–11672
- Götz F, Bannerman T, Schleifer K (2006) The Genera *Staphylococcus* and *Micrococcus*. In *The Prokaryotes*, Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E (eds), pp 5–75. New York, NY: Springer
- Heidrich N, Moll I, Brantl S (2007) *In vitro* analysis of the interaction between the small RNA SR1 and its primary target *ahrC* mRNA. *Nucleic Acids Res* 35: 4331–4346
- Herbert S, Ziebandt AK, Ohlsen K, Schäfer T, Hecker M, Albrecht D, Novick R, Götz F (2010) Repair of global regulators in *Staphylococcus aureus* 8325 and comparative analysis with other clinical isolates. *Infect Immun* 78: 2877–2889
- Howden BP, Beaume M, Harrison PF, Hernandez D, Schrenzel J, Seemann T, Francois P, Stinear TP (2013) Analysis of the small RNA transcriptional response in multidrug-resistant *Staphylococcus aureus* after antimicrobial exposure. *Antimicrob Agents Chemother* 57: 3864–3874
- Kawamoto H, Morita T, Shimizu A, Inada T, Aiba H (2005) Implication of membrane localization of target mRNA in the action of a small RNA: mechanism of post-transcriptional regulation of glucose transporter in *Escherichia coli*. *Genes Dev* 19: 328–338
- Kinkel TL, Roux CM, Dunman PM, Fang FC (2013) The *Staphylococcus aureus* SrrAB two-component system promotes resistance to nitrosative stress and hypoxia. *MBio* 4: e00696-13
- Korea CG, Balsamo G, Pezzicoli A, Merakou C, Tavarini S, Bagnoli F, Serruto D, Unnikrishnan M (2014) Staphylococcal Esx proteins modulate apoptosis and release of intracellular *Staphylococcus aureus* during infection in epithelial cells. *Infect Immun* 82: 4144–4153

- Lalaouna D, Carrier MC, Semsey S, Brouard JS, Wang J, Wade JT, Masse E (2015) A 3' external transcribed spacer in a tRNA transcript acts as a sponge for small RNAs to prevent transcriptional noise. *Mol Cell* 58: 393–405
- Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Bowtie: an ultrafast memory-efficient short read aligner. *Genome Biol* 10: R25
- Lennon JT, Jones SE (2011) Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat Rev Microbiol* 9: 119–130
- Li C, Sun F, Cho H, Yelavarthi V, Sohn C, He C, Schneewind O, Bae T (2010) CcpA mediates proline auxotrophy and is required for *Staphylococcus aureus* pathogenesis. *J Bacteriol* 192: 3883–3892
- Licht A, Preis S, Brantl S (2005) Implication of CcpN in the regulation of a novel untranslated RNA (SR1) in *Bacillus subtilis*. *Mol Microbiol* 58: 189–206
- Liu Q, Yeo WS, Bae T (2016) The SaeRS two-component system of *Staphylococcus aureus*. *Genes* 7: E81
- Maira-Litrán T, Kropec A, Goldmann DA, Pier GB (2005) Comparative opsonic and protective activities of *Staphylococcus aureus* conjugate vaccines containing native or deacetylated Staphylococcal Poly-N-acetyl-beta-(1-6)-glucosamine. *Infect Immun* 73: 6752–6762
- Majerczyk CD, Dunman PM, Luong TT, Lee CY, Sadykov MR, Somerville GA, Bodi K, Sonenshein AL (2010) Direct targets of CodY in *Staphylococcus aureus*. *J Bacteriol* 192: 2861–2877
- 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
- Miyakoshi M, Chao Y, Vogel J (2015) Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *EMBO J* 34: 1478–1492
- Møller T, Franch T, Udesen C, Gerdes K, Valentin-Hansen P (2002) Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. *Genes Dev* 16: 1696–1706
- Novick RP (2003) Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Mol Microbiol* 48: 1429–1449
- Papenfors K, Sun Y, Miyakoshi M, Vanderpool CK, Vogel J (2013) Small RNA-mediated activation of sugar phosphatase mRNA regulates glucose homeostasis. *Cell* 153: 426–437
- Papenfors K, Vogel J (2014) Small RNA functions in carbon metabolism and virulence of enteric pathogens. *Front Cell Infect Microbiol* 4: 91
- Park JY, Kim JW, Moon BY, Lee J, Fortin YJ, Austin FW, Yang SJ, Seo KS (2015) Characterization of a novel two-component regulatory system, HptRS, the regulator for the hexose phosphate transport system in *Staphylococcus aureus*. *Infect Immun* 83: 1620–1628
- Pohl K, Francois P, Stenz L, Schlink F, Geiger T, Herbert S, Goerke C, Schrenzel J, Wolz C (2009) CodY in *Staphylococcus aureus*: a regulatory link between metabolism and virulence gene expression. *J Bacteriol* 191: 2953–2963
- Pragman AA, Yarwood JM, Tripp TJ, Schlievert PM (2004) Characterization of virulence factor regulation by SrrAB, a two-component system in *Staphylococcus aureus*. *J Bacteriol* 186: 2430–2438
- Rice JB, Balasubramanian D, Vanderpool CK (2012) Small RNA binding-site multiplicity involved in translational regulation of a polycistronic mRNA. *Proc Natl Acad Sci USA* 104: 20454–20459
- Richards GR, Patel MV, Llyod CR, Vanderpool CK (2013) Depletion of glycolytic intermediates plays a key role in glucose-phosphate stress *Escherichia coli*. *J Bacteriol* 195: 4816–4825
- Richardson AR, Somerville GA, Sonenshein AL (2015) Regulating the intersection of metabolism and pathogenesis in gram-positive bacteria. *Microbiol Spectr* 3: 1–27
- Richardson AR, Dunman PM, Fang FC (2006) The nitrosative stress response of *Staphylococcus aureus* is required for resistance to innate immunity. *Mol Microbiol* 4: 927–939
- Rochat T, Bohn C, Morvan C, Le Lam TN, Razvi F, Pain A, Toffano-Nioche C, Ponien P, Jacq A, Jacquet E, Fey PD, Gautheret D, Bouloc P (2018) The conserved regulatory RNA RsaE down-regulates the arginine degradation pathway in *Staphylococcus aureus*. *Nucleic Acids Res* 46: 8803–8816
- Ruiz de los Mozos I, Vergara-Irigaray M, Segura V, Villanueva M, Bitarte N, Saramago M, Domingues S, Arraiano CM, Fechter P, Romby P, Valle J, Solano C, Lasa I, Toledo-Arana A (2013) Base pairing interaction between 5'- and 3'-UTRs controls icaR mRNA translation in *Staphylococcus aureus*. *PLoS Genet* 9: e1004001.
- Sadykov MR, Olson ME, Halouska S, Zhu Y, Fey PD, Powers R, Somerville GA (2008) Tricarboxylic acid cycle-dependent regulation of *Staphylococcus epidermidis* polysaccharide intercellular adhesin synthesis. *J Bacteriol* 190: 7621–7632
- Seidl K, Bischoff M, Berger-Bächi B (2008a) CcpA mediates the catabolite repression of *tst* in *Staphylococcus aureus*. *Infect Immun* 76: 5093–5099
- Seidl K, Goerke C, Wolz C, Mack D, Berger-Bächi B, Bischoff M (2008b) *Staphylococcus aureus* CcpA affects biofilm formation. *Infect Immun* 76: 2044–2050
- Seidl K, Müller S, François P, Kriebitzsch C, Schrenzel J, Engelmann S, Bischoff M, Berger-Bächi B (2009) Effect of a glucose impulse on the CcpA regulon in *Staphylococcus aureus*. *BMC Microbiol* 9: 95
- Seidl K, Stucki M, Ruegg M, Goerke C, Wolz C, Harris L, Berger-Bächi B, Bischoff M (2006) *Staphylococcus aureus* CcpA affects virulence determinant production and antibiotic resistance. *Antimicrob Agents Chemother* 50: 1183–1194
- Somerville GA, Proctor RA (2009) At the crossroads of bacterial metabolism and virulence factor synthesis in staphylococci. *Microbiol Mol Biol Rev* 73: 233–248
- Tomasini A, Moreau K, Chicher J, Geissmann T, Vandenesch F, Romby P, Marzi S, Caldelari I (2017) The RNA targetome of *Staphylococcus aureus* non-coding RNA RsaA: impact on cell surface properties and defense mechanisms. *Nucleic Acids Res* 45: 6746–6760
- Ulrich M, Bastian M, Cramton SE, Ziegler K, Pragman AA, Bragonzi A, Memmi G, Wolz C, Schlievert PM, Cheung A, Döring G (2007) The staphylococcal respiratory response regulator SrrAB induces *ica* gene transcription and polysaccharide intercellular adhesin expression, protecting *Staphylococcus aureus* from neutrophil killing under anaerobic growth conditions. *Mol Microbiol* 65: 1276–1287
- Vanderpool CK, Gottesman S (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol Microbiol* 54: 1076–1089
- Vanderpool CK, Gottesman S (2007) The novel transcription factor SgrR coordinates the response to glucose-phosphate stress. *J Bacteriol* 189: 2238–2248
- Varet H, Brillet-Guéguen L, Coppée JY, Dillies MA (2016) SARTools: a DESeq2- and edgeR-based R pipeline for comprehensive differential analysis of RNA-Seq data. *PLoS ONE* 11(6): e0157022
- Vergara-Irigaray M, Valle J, Merino N, Latasa C, Garcia B, Ruiz de los Mozos I, Solano C, Toledo-Arana A, Penades JR, Lasa I (2009) Relevant role of fibronectin-binding proteins in *Staphylococcus aureus* biofilm-associated foreign-body infections. *Infect Immun* 77: 3978–3991
- Vitko NP, Grosser MR, Khatri D, Lance TR, Richardson AR (2016) Expanded glucose import capability affords *Staphylococcus aureus* optimized glycolytic flux during infection. *MBio* 7: e00296-16
- Vitko NP, Spahich NA, Richardson AR (2015) Glycolytic dependency of high-level nitric oxide resistance and virulence in *Staphylococcus aureus*. *MBio* 6: e00045-15
- You Y, Xue T, Cao L, Zhao L, Sun H, Sun B (2014) *Staphylococcus aureus* glucose-induced biofilm accessory proteins, GbaAB, influence biofilm formation in a PI-A-dependent manner. *Int J Med Microbiol* 304: 603–612

Appendix file

A multifaceted small RNA modulates gene expression upon glucose limitation in *Staphylococcus aureus*.

Delphine Bronesky, Emma Desgranges, Anna Corvaglia, Patrice François, Carlos J. Caballero, Laura Prado, Alejandro Toledo-Arana, Inigo Lasa, Karen Moreau, François Vandenesch, Stefano Marzi, Pascale Romby and Isabelle Caldelari

Table of contents

Appendix Material and Methods.....	2
Appendix References.....	3

APPENDIX MATERIALS AND METHODS

Plasmids and Strains Constructions

To delete *rsal* in HG001, we amplified 967 nucleotides and 1000 nucleotides respectively upstream and downstream to the deleted region by using the following specific primers: Teg24-5'F-MluI and Teg24-5'R-NcoI (for the 5'-region PCR fragment), Teg24-3'F-NcoI and Teg24-3'R-BglII (for the 3'-region PCR fragment) (Table S2). The 5' - and 3'-region fragments were digested with *NcoI*, ligated and re-amplified using primers Teg24-5'F-MluI and Teg24-3'R-BglII. The generated 1.96 kb fragment was digested using *BglII* and *MluI* restriction enzymes and cloned into the plasmid pMAD CM (Memmi et al., 2008). The resulting plasmid was transformed into *E. coli* strain DH5 α for subsequent electroporation into *S. aureus* strain RN4220. Clones were selected on Mueller-Hinton broth (MHB) supplemented with 10 μ g/mL chloramphenicol and the plasmid was transduced in *S. aureus* strain HG001. Plasmid integration into the chromosome was selected after growth at a non-permissive temperature (42°C) while maintaining antibiotic selective pressure. A second recombination event was screened at the permissive temperature (30°C). Recombinant genomes were checked by PCR and sequenced using appropriate primers (Table S2). The mutant allele for deletion of *rsal* gene in *S. aureus* 132 strain was generated by two-step overlapping PCR using the appropriate oligonucleotides (Table S2). The resulting PCR product was purified from agarose gel using NucleoSpin® Gel and PCR Clean-up Macherey-Nagel kit and cloned into a LIC-modified pMAD vector using the Ligation-independent cloning kit (Novagen) (Arnaud et al., 2004). The pMAD- Δ *rsal* plasmid was introduced into *S. aureus* 132 strain by electroporation (Lee, 1995). Then, gene deletion was performed by a two-step procedure replacing the wild type chromosomal region by the corresponding mutant allele contained into the pMAD plasmid (Valle et al. 2003). The resulting modified strain was verified by PCR and fully sequenced using Sanger sequencing (Table S2).

Plasmids pES::*rsal* and pES::*rsal* mut5 were constructed as followed. The promoter of *blaZ* gene (*PblaZ*) was excised using *SphI* and *BamHI* from the pCN57 and cloned into pCN47 generating the pES plasmid (Charpentier et al., 2004). The wild type *rsal* gene was amplified by PCR using the *rsal_BamHI* and *rsal_EcoRI* oligonucleotides. The *RsaI* sRNA carrying the mutation at the predicted IcaR binding site was generated by two-step overlapping PCR using appropriate oligonucleotides (Table S2). The resulting *RsaI* and *RsaI_mut5* PCR products were purified from agarose gel using NucleoSpin® Gel and PCR Clean-up Macherey-Nagel kit, ligated into the pJET 1.2 vector (Thermo Scientific) and cloned in *E. coli* XL1-Blue (Stratagene). The pJET-*rsal* and pJET *rsal_mut5* plasmids were purified from overnight cultures with the NucleoSpin® Plasmid Macherey-Nagel kit. The DNA fragments were excised with *BamHI* and *EcoRI* FastDigest restriction enzymes (Thermo

Scientific), purified from agarose gel, and ligated using the Rapid Ligation Kit (Thermo Scientific) into a pES plasmid digested with the same enzymes. The generated pES-*rsal* and pES-*rsal_mut5* plasmids were sequenced for verification and then introduced into *S. aureus* 132 derivative strains by electroporation (Lee, 1995).

MAPS experiments and RNA-seq analysis

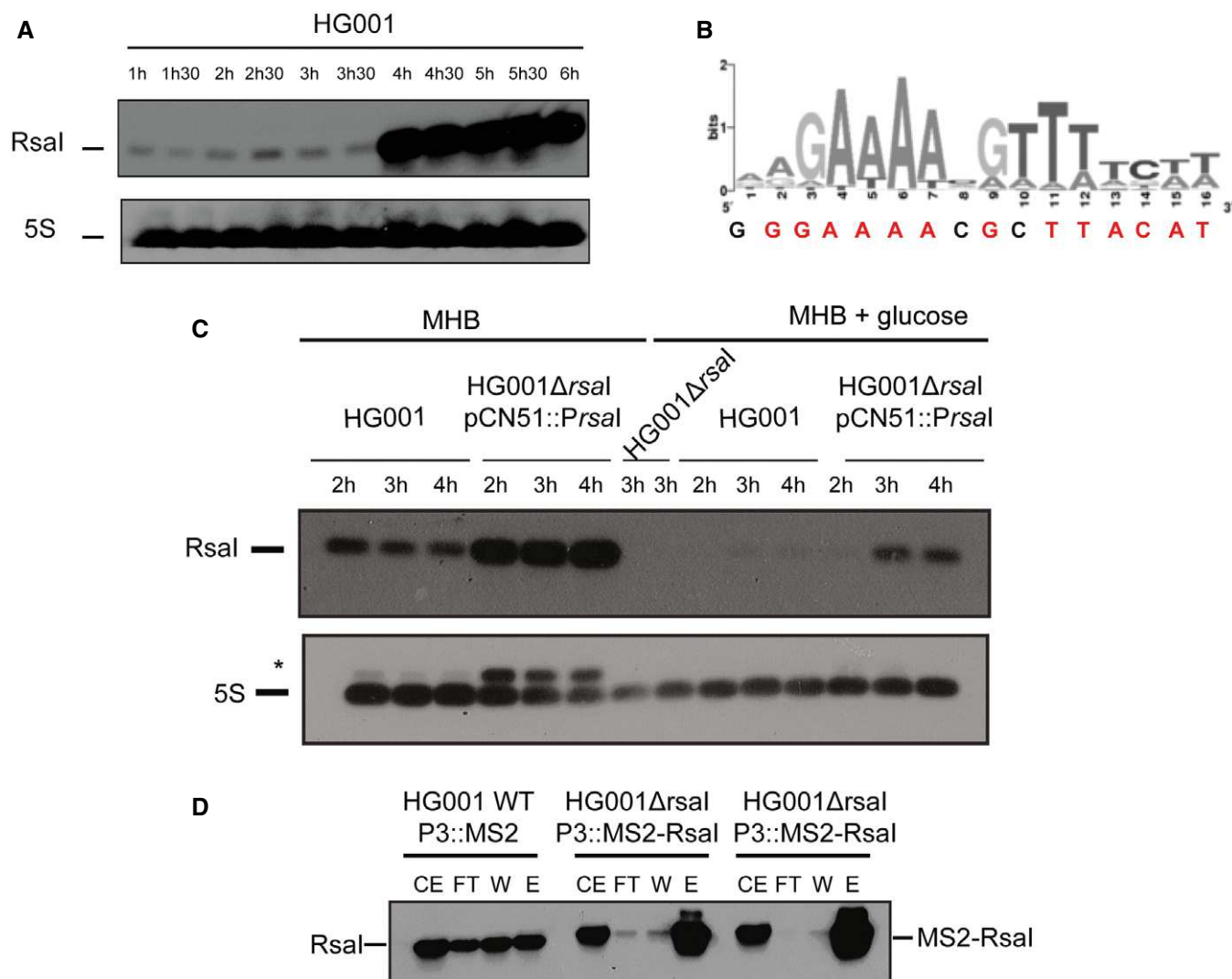
To analyse the effect of specific mutations in *RsaI* on the target recognition, MAPS approach was performed on a set of experiments including 1 MS2 control, 1 wild-type MS2-*RsaI*, and one experiment for each *RsaI* mutant (MS2-*RsaI* mut2 carrying deletion of C22 to G29, and MS2-*RsaI* mut4 carrying deletion of nucleotides 75 to 113). To get a more relevant analysis of the data, we have also included the previous experiments performed on the MS2-*RsaI*. Therefore, in total seven MAPS experiments (2 MS2 controls, 3 MS2-*RsaI*, 1 MS2-*RsaI* mut2 and 1 MS2-*RsaI* mut4) have been analysed together using Deseq2. For this analysis, we have considered four categorical variables or factors (MS2 contr, MS2-*RsaI*-, MS2-*RsaI* mut2 and MS2-*RsaI* mut4). Because the MS2 purifications were done in two separate days (1 MS2 control and two MS2-*RsaI* in one experiment; 1 MS2 control, 1 MS2-*RsaI*, 1 MS2-*RsaI* mut2 and 1 MS2-*RsaI* mut4 in the other sets of experiments), some aleatory variables were expected to occur. The data pairing has been performed by assigning a blocking factor (batch effect) in order to be able to directly compare the obtained fold changes. Hence, the shrinkage estimation for dispersions and fold changes has been calculated considering the seven samples together. Under these conditions, the amount of shrinkage was significantly enhanced when compared to the analysis presented in Tables 1 and S3. It is inversely proportional to the residual degrees of freedom = number of samples - number of parameters to estimate. In this case, we added two more parameters, the groups MS2-*RsaI* mut2 and MS2-*RsaI* mut4, and four more samples. Because the shrinkage is increased, we applied a lower cut-off for the consideration of the significantly enriched RNAs, which is 3 for this analysis.

APPENDIX REFERENCES

- Arnaud M, Chastanet A, Débarbouillé M (2004) New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl. Environ. Microbiol.* 70, 6887–6891
- Charpentier E, Anton AI, Barry P, Alfonso B, Fang Y, Novick R.P (2004) Novel cassette-based shuttle vector system for gram-positive bacteria. *Appl. Environ. Microbiol.* 70, 6076–6085
- Herbert S, Ziebandt A.K, Ohlsen K, Schäfer T, Hecker M, Albrecht D, Novick R, Götz F (2010) Repair of global regulators in *Staphylococcus aureus* 8325 and comparative analysis with other clinical isolates. *Infect. Immun.* 78: 2877–2889
- Kreiswirth B.N, Lofdahl S, Betley M.J, O'Reilly M, Schlievert P.M, Bergdoll M.S, Novick R.P (1983) The toxic shock syndrome exotoxin structural gene is not detectably transmitted

- by a prophage. *Nature* 305: 709-712
- Lee, J.C (1995) Electrotransformation of Staphylococci. *Methods Mol Biol* 47, 209–216
- Memmi G, Filipe S.R, Pinho M.G, Fu Z, Cheung A (2008) *Staphylococcus aureus* PBP4 is essential for β -lactam resistance in community-acquired methicillin-resistant strains. *Antimicrob. Agents Chemother* 52 : 3955–3966
- Monk I.R, Shah I.M, Xu M, Tan M.W, Foster T.J (2012) Transforming the untransformable: Application of direct transformation to manipulate genetically *Staphylococcus aureus* and *Staphylococcus epidermidis*. *MBio* 20: e00277-11
- Pohl K, Francois P, Stenz L, Schlink F, Geiger T, Herbert S, Goerke C, Schrenzel J, Wolz C (2009) CodY in *Staphylococcus aureus*: A regulatory link between metabolism and virulence gene expression. *J. Bacteriol* 191: 2953–2963
- Romilly C, Lays C, Tomasini A, Caldelari I, Benito Y, Hammann P, Geissmann T, Boisset S, Romby P, Vandenesch F (2014) A Non-Coding RNA Promotes Bacterial Persistence and Decreases Virulence by Regulating a Regulator in *Staphylococcus aureus*. *PLoS Pathog.* Mar 20:e1003979
- Taylor R. G, Walker D. C, McInnes R. R (1993) *E. coli* host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Res.* 11:1677-8
- Valle J, Toledo-Arana A, Berasain C, Ghigo J.M, Amorena B, Penadés J.R, Lasa I (2003) SarA and not sigma B is essential for biofilm development by *Staphylococcus aureus*. *Mol Microbiol* 48: 1075–1087
- van der Voort M, Kuipers O.P, Buist G, de Vos W.M, Abee T (2008) Assessment of CcpA-mediated catabolite control of gene expression in *Bacillus cereus* ATCC 14579. *BMC Microbiol* 8: 62
- Villanueva M, García B, Valle J, Rapún B, Ruiz de Los Mozos I, Solano C, Martí M, Penadés J.R, Toledo-Arana A, Lasa I (2018) Sensory deprivation in *Staphylococcus aureus*. *Nat Commun* 6: 523

Expanded View Figures

**Figure EV1. Analysis of the RsaI expression.**

- A** Northern blot experiments showing the expression of RsaI in HG001 strain during growth phase. Total RNA was prepared from samples taken after various time points of growth in BHI medium at 37°C. Hybridization against 5S rRNA was used as loading control on the same samples which have been migrated on another agarose gel because RsaI and 5S rRNA have very similar sizes.
- B** The *cre*-site consensus was defined in *Bacillus cereus* (van der Voort *et al*, 2008). The sequence found in the 5' region of RsaI is represented below the graph, the conserved residues are in red.
- C** Northern blot experiments showing the expression of RsaI in HG001 strain, the isogenic HG001ΔrsaI mutant strain, and the same mutant strain complemented with a plasmid expressing RsaI under the control of its own promoter (HG001ΔrsaI pCN51::PrsaI). Total RNA was prepared from samples taken after various time points of growth in MHB medium in the absence or presence of 1% glucose. For the mutant HG001ΔrsaI strain (used as a negative control), total RNA was prepared after 4 h of growth in MHB medium in the absence or in the presence of glucose. Hybridization against 5S rRNA was used as loading control on the same membrane. *Traces of RsaI signal after re-hybridization of the membrane with the 5S probe were still observed.
- D** MS2-RsaI is specifically retained by affinity chromatography containing the MS2-MBP protein. The Northern blot was performed using a DIG-labeled RsaI probe (Table EV3) to visualize RsaI and MS2-RsaI following the MS2 chromatography affinity. CE is for crude extract, FT for flow-through, W for washing, and E for elution. For CE/FT/W samples, 5 μg of total RNA was loaded on a 1.5% agarose gel while for E sample, only 0.5 μg of total RNA was used. The two replicates are shown for the RNA purified from the strain HG001ΔrsaI::MS2-RsaI.

Source data are available online for this figure.

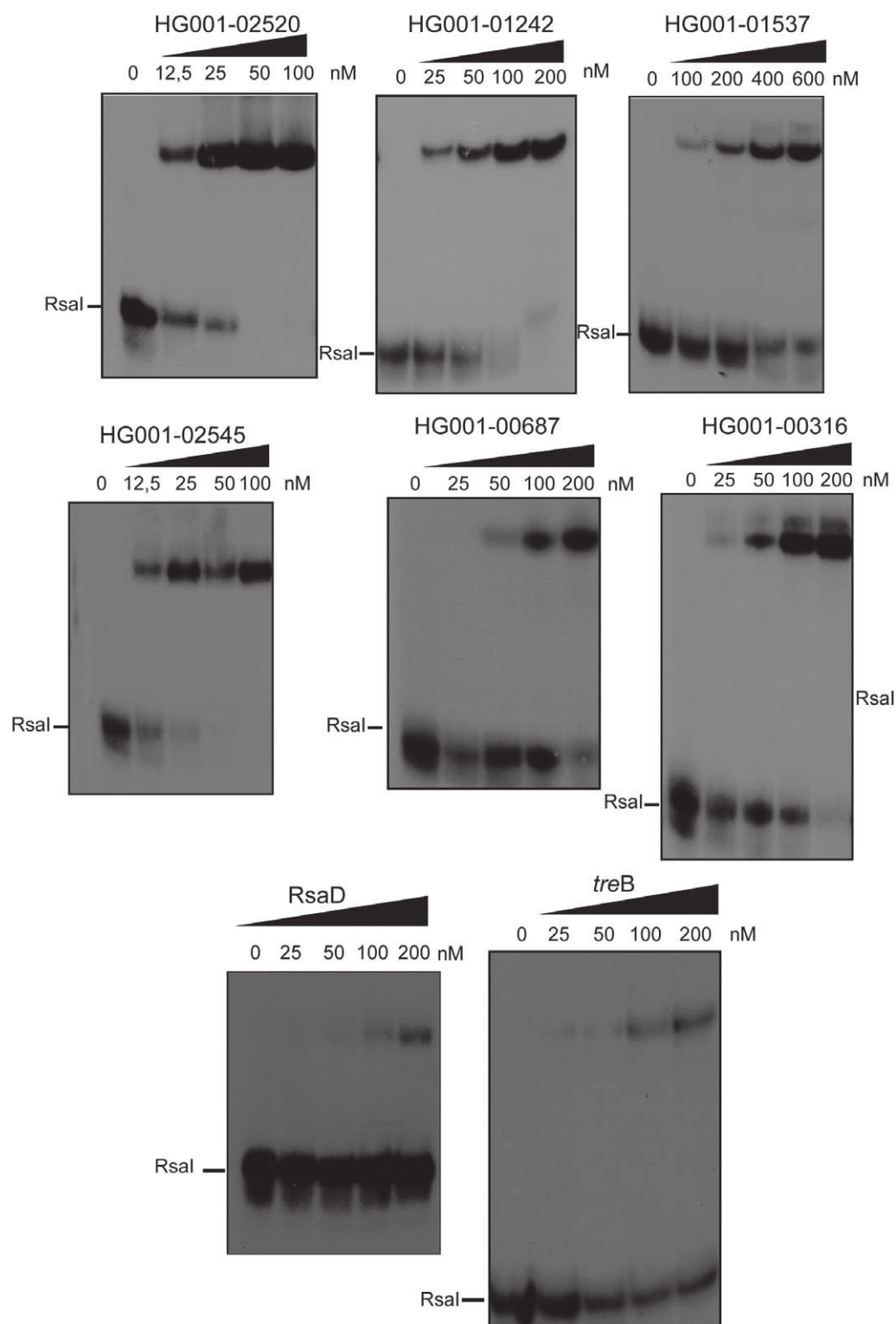


Figure EV2. RsaI interacts with several RNA species.

Gel retardation assays showing the formation of complexes between RsaI and several mRNAs, and the sRNA RsaD. The 5' end-labeled RsaI was incubated with increasing concentrations of target mRNAs or RsaD as shown on the top of the autoradiographies. Quantification of the autoradiographies showed that RsaI forms stable complexes with many of the mRNAs ($K_d < 50$ nM) except for *TreB* and RsaD ($K_d > 200$ nM).

Source data are available online for this figure.

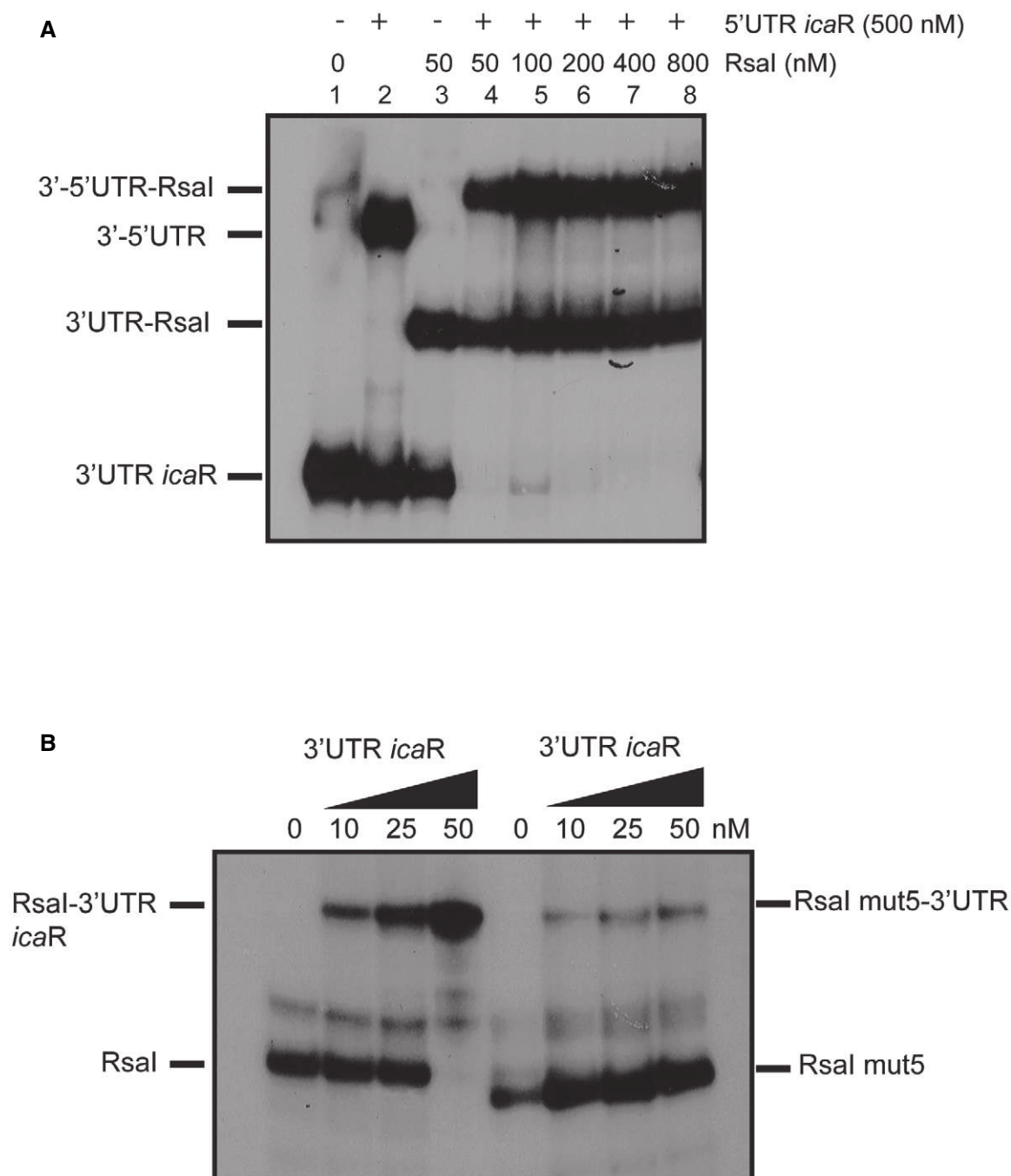


Figure EV3. RsaI interacts with the 3'UTR of *icaR* mRNA.

A Analysis of the effect of RsaI on the interaction between the 5'UTR and the 3'UTR of *icaR*. The 5' end-labeled of the 3'UTR of *icaR* was incubated with 500 nM of the 5'UTR of *icaR* alone or in the presence of increasing concentrations of RsaI, as shown on the top of the autoradiography.

B Gel retardation assays showing the formation of complexes between RsaI or RsaI mut5 and the 3'UTR of *icaR*. RsaI or RsaI mut5 labeled at their 5' ends was incubated with increasing concentrations of the 3'UTR of *icaR* (10, 25, and 50 nM). In both gels, the various RNA species present in the different bands are given on one side of the autoradiographies.

Source data are available online for this figure.

Figure EV4. Analysis of the expression of RsaG and RsaD in response to stress.

- A Genomic organization of the *uhpT* locus in HG001 strain. This locus is conserved in all *S. aureus* strains (Geissmann *et al*, 2009). The transcription of *uhpT* containing RsaG is induced by the two-component system HptRS, which senses the extracellular concentration of glucose-6 phosphate (G-6P).
- B Northern blot experiment showing RsaG (left panel) or RsaI (right panel) expression in the HG001 strain. Total RNA was extracted at 2, 4, and 6 h of growth in BHI or MHB with or without addition of glucose-6 phosphate (+G-6P). 5S rRNA was used as loading control using the same samples, which have been migrated on another agarose gel.
- C Northern blot experiment showing RsaG expression in the HG001 Δ *hptRS* deleted mutant. Total RNA was extracted at 2, 4, and 6 h of growth in BHI with or without addition of glucose-6 phosphate (+G-6P). 5S rRNA was used as loading control using the same samples, which have been migrated on another agarose gel.
- D The β -galactosidase activity (Miller Units) have been measured from *PrpoB::fn3K::lacZ*, *PrpoB::fn3K::lacZ::rsal*, and *PrpoB::HG001_fn3K::lacZ::rsal* mut4 expressed in HG001 and HG001 Δ *rsaG* strains. The β -galactosidase activity was normalized for bacterial density and the results represented the mean of four independent experiments. **** $P < 0.0001$, ns: not significant.
- E Northern blot experiment showing the steady-state level of RsaD. On the left, RsaD expression was analyzed at 2, 4, and 6 h of growth, in HG001 wild-type strain and in Δ *saeRS* and Δ *srrAB* mutant strains. On the right, the expression of RsaD was monitored in response to nitric oxide (NO) stress. The culture performed in BHI medium until OD₆₀₀ = 0.2 at 37°C was then treated with the addition of 100 μ M Na-diethylamine (+DEA NONOate). 5S rRNA was used as loading control using the same samples, which have been migrated on another agarose gel.

Source data are available online for this figure.

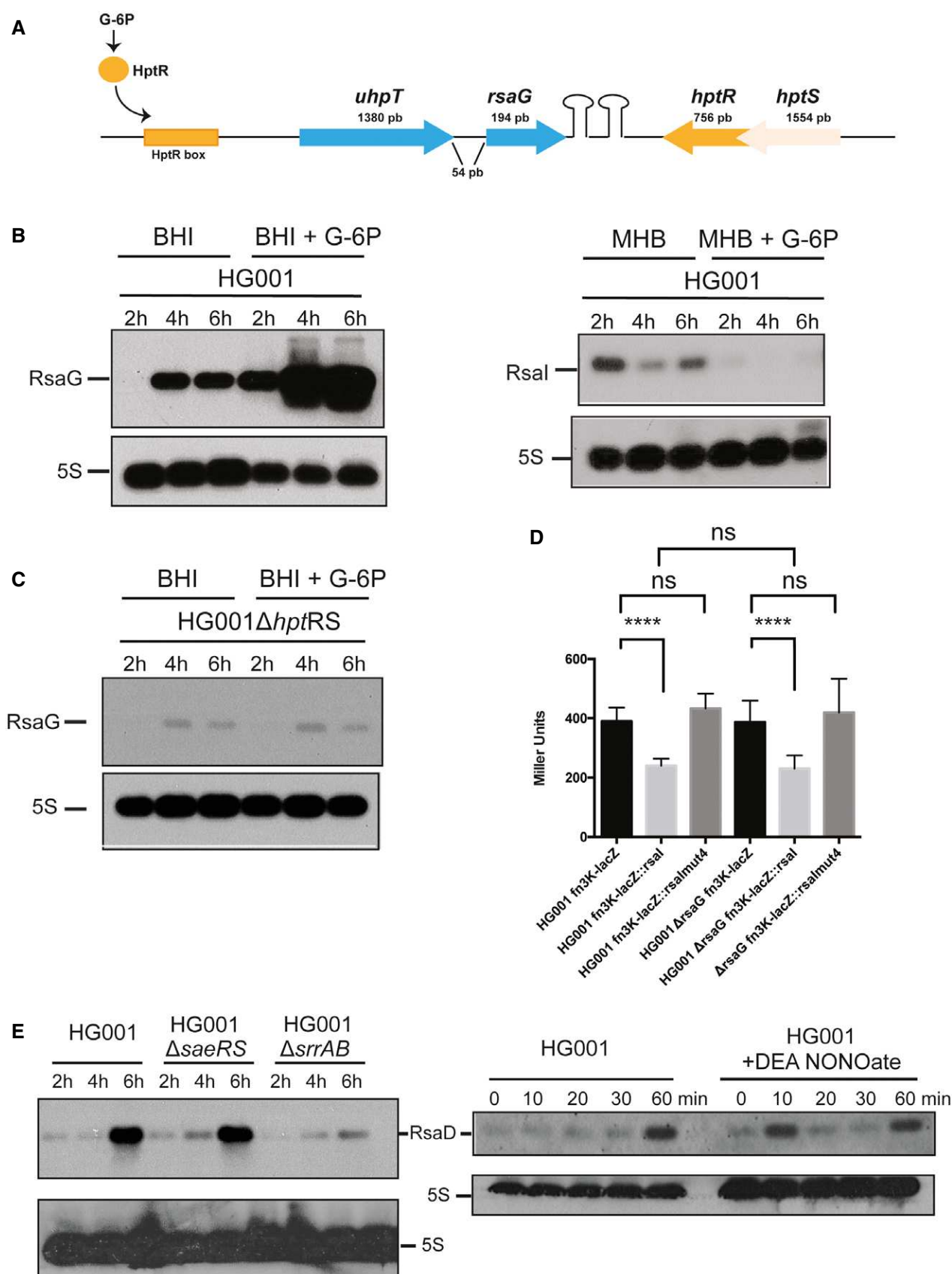


Figure EV4.

Figure EV5. RsaI represents a signature of a metabolic change as the result of glucose consumption.

- A Growth curves of the wild-type 132 (132) and mutant Δ rsaI (132D) strains containing the vector pES (empty) or pES::rsaI (expressing RsaI from a constitutive promoter) in BHI.
- B Northern blot experiment showing RsaI expression in the wild-type 132 (132) or mutant Δ rsaI (132 Δ rsaI) strains containing the vector pES or pES::rsaI. Total RNA was extracted at 2, 4, and 6 h of growth in BHI. 5S rRNA was used as the loading control using the same samples, which have been migrated on another agarose gel. Quantification of RsaI normalized to 5S rRNA was done with ImageQuant TL software (GE Healthcare Life Sciences).
- C Northern blot experiments showing the expression of RsaI in the wild-type HG001 strain, the mutant HG001 Δ rsaI strain, and the same mutant strain complemented with a plasmid expressing RsaI from its own promoter (HG001 Δ rsaI pCNS1::PrsaI). All the other strains have been transformed with the pCNS1 plasmid. Total RNA was prepared from samples taken after 6 h of growth in BHI medium. Hybridization against 5S rRNA was used as loading control using the same membrane. Three replicates were carried out (lanes 1–3), and the same samples were used for the differential transcriptomic analysis.

Source data are available online for this figure.

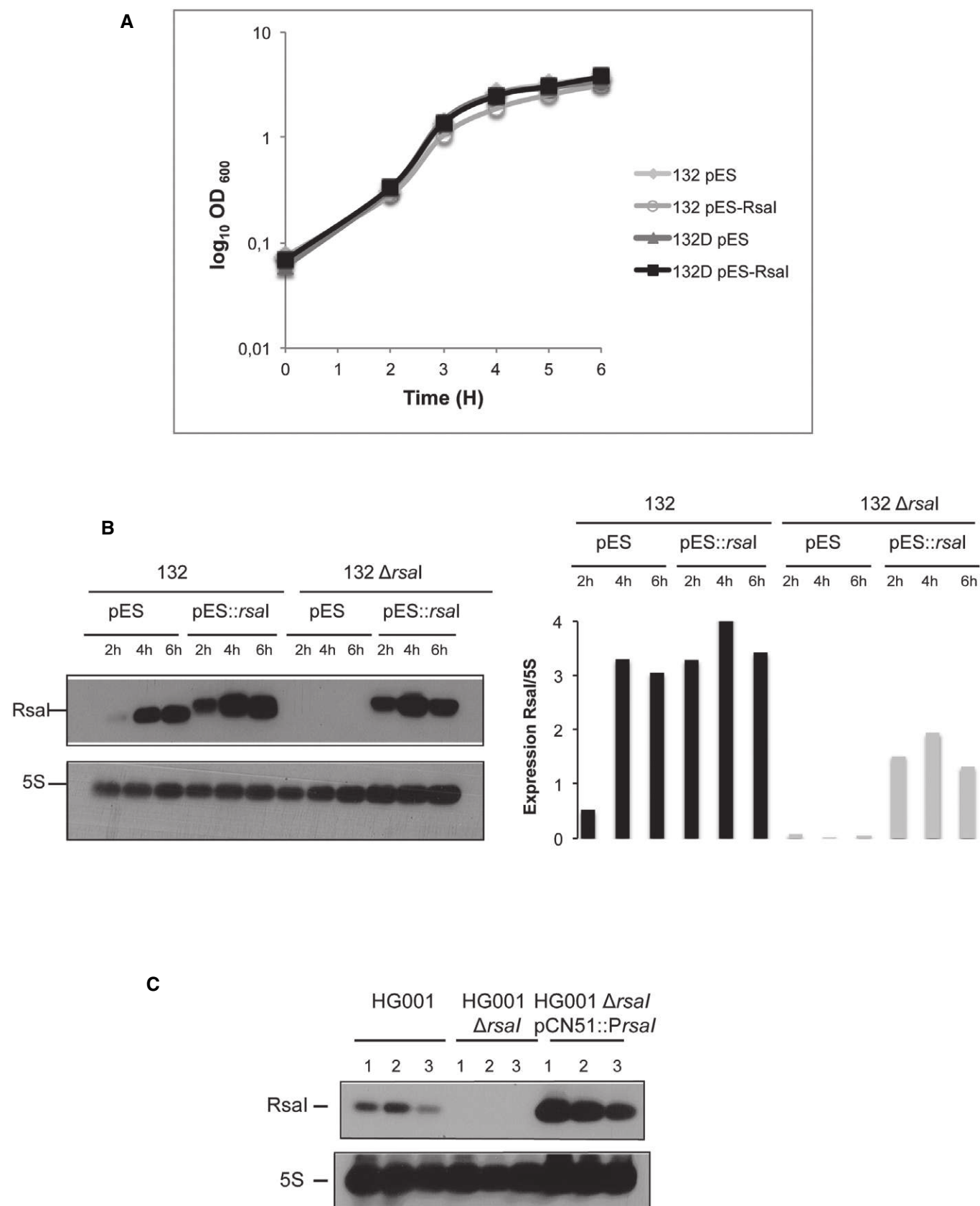


Figure EV5.

Table EV1: MAPS data revealed the best RNA candidates found enriched with MS2-Rsal.

Target	Name of the gene and function	Base-pairing interactions prediction	ΔG (kcal/mol)	Fold change* MS2-Rsal/MS2	Binding	Fold change Δ rsal/complemented
HG001_02704	IcaR Biofilm operon <i>icaADBC</i> HTH-type negative transcriptional regulator	949 GAAAUGGAAAGUAAUAAGUAAUA 973 108 CUUU-CCUUUCAU---UCAUUUAU 88	-14,92	33,976	i	0,839
HG001_02520	XRE family transcriptional regulator	-21 AUAACAUAUGGAG-UGA--UGAUUAUG 3 92 UAUUUUAUACUUCUACAACACUACAUAU 66	-12,10	22,978	i	0,479
HG001_01242	Hypothetical protein	-29 AAGGAAGUAAUGAAUAAA--GGAGAUG 3 106 UUCCUUUCAU-CAUUUUUAUAAUUCUACA 76	-16,27	17,369	i	0,818
HG001_02756	RsaG	43 CUACCCUU 50 32 GAUGGGGA 25	-8,82	13,621	i	0,484
HG001_02360	GlvR	602 UACCCUU+608 31 AUGGGGA 25	-6,78	9,961	ni	1,0342
HG001_01537	RsmG Putative O-methyltransferase	-22 GGGGAUCCAGCAGUAGUAAUGUAUGUGAAU 8 106 UUCCUUUCAU-UCAUUUAU-UUAU-ACUUC 80	-9,74	7,137	i	1,536
HG001_02609	<i>isaA</i>	741 UAUAUAGCUCU-UCCC 755 41 AUAUAUCGAGAUUGGG 26	-11,36	6,76	ni	1,054
HG001_00847	RsaE	59 CCCCUUUCUU 68 29 GGGGAAACAA 20	-14,21	5,476	ni	0,951
HG001_02545	Hypothetical protein	-2 AAUUGGAAAGAAUUUGAAGUAAU 21 113 UUUUAUCUUUCUUUCAUUCAUUA 91	-11,16	5,262	i	1,103
HG001_00316	Peptidase propeptide and YPEB domain protein	-23 AAUAGGAAGAGAGAGUGAUUUUUG 3 113 UUUUAUCUUUC-CUUUCAUUAUUUAU 89	-15,55	5,234	i	1,342
HG001_02210	Hypothetical protein	-26 AAUUAUUUAAAGGAUGU -10 90 UUUUAUAC--UUCUACA 76	-8,33	4,978	ni	1,212
HG001_02293	GlcU_2 putative glucose uptake protein	-25 AAUAGUAUAGGAGUGAGUAGUCUUG 3 113 UUUUAUCUU-UCCUUUCAUUAUUUAU 87	-18,25	4,769	i	1,061
HG001_00974	Hypothetical protein	401 CUACCCU 407 32 GAUGGGG 26	-7,18	4,532	ni	1,219
HG001_02628	Fn3K Fructosamine Kinase	-24 UAAGUAGAUUGGAGGUGUUUGU-AUGAAUG 7 114 GUUUUAUCUUCCUUUCAU--CAUUAUUUAU 86	-16,44	4,37	i	0,935
HG001_00942	Putative cell wall binding lipoprotein	-24 AAUAGCAAGUGGAGGUACAAGUAUGAAAUUGGA 12 113 UUUUAUCUUUC-CUUUCAU--UCAUUUUUAUAAUUCU 81	-16,06	4,091	i	0,676

HG001_00584	RsaD	42 CUCCUUUGU 50 29 GGGGAAACA 21	-7,97	4,078	i	1,015
-------------	------	------------------------------------	-------	-------	---	-------

The red bases are involved in base-pairing interactions. In green are shown the Shine and Dalgarno sequence and the translation start codon AUG (UUG for *g/cU_2* mRNA). ΔG for RNA hybrids were predicted with IntaRNA. The enrichment value as deduced from the MAPS data corresponds to MS2-RsaI versus the MS2 control. *For all assays, the p-value was inferior to 0,00005 (Table S3). Binding of the RNA candidates to RsaI was monitored by gel retardation assays: n.i, means no interaction, and i for validated interaction. The last column corresponds to the transcriptomic data (Table S6) and is represented by the ratio between the mutant Δ rsal strain versus the same strain complemented with a plasmid expressing RsaI.

Table EV2: The list of RNAs sequenced using MS2-RsaI as the bait. The enrichment values corresponded to the fold changes (ratio MS2-RsaI/MS2 control), and the p-values have been calculated using DEseq2. We have considered in this study all RNAs, which were enriched at least 4-fold. Duplicates have been carried out for each assay.

<https://www.embopress.org/doi/full/10.15252/emboj.201899363>

Table EV3: List of oligonucleotides used in the study

Name	Sequences 5'-3'	Use
RsaI For	TAACAGGGGGAGCGATTAAA	DIG probe
RsaI Rev	TAATACGACTCACTATAGGGGCAC GTGCTAGCCGACAAATA	
RsaD For	ACTTCTTCAATAGAGAAGTT	DIG probe
RsaD Rev	TAATACGACTCACTATAGGGACAT GGTAATACACTTGGCT	
RsaG For	TGCCCCGACGGCATGTGCGA	DIG probe
RsaG Rev	TAATACGACTCACTATAGGGGACA CCGCTATATGTC	
<i>g/cU</i> For	TAATACGCATCACTATAGGGCAAG TGTCAATTAAGGGGAGCACT	<i>In vitro</i> transcription
<i>g/cU</i> Rev	CTATTTCAAATTACCTAGAATTATA GCAG	
<i>fn3K</i> For	TAATACGACTCACTATAGGGAAGA TTAAATCATCCTAATTAGGCAA	<i>In vitro</i> transcription
<i>fn3K</i> Rev	CTAACTTGTTGTATCTTGTAATC TT	
RsaD For	TAATACGACTCACTATAGTCTGCT ATAATG	<i>In vitro</i> transcription
RsaD Rev	AAAAACCAAGTGACATGGTAA	
HG001_02520 For	TAATACGACTCACTATAGGGTAAG AAGTTAATAAATAGA	<i>In vitro</i> transcription
HG001_02520 Rev	CAATCACACTTCAATTGC	

HG001_00687 For	TAATACGACTCACTATAGGGT GAG TATATTTTAGGAGGA	<i>In vitro</i> transcription
HG001_00687 Rev	AAAATTTTATAATTATGTTAATCTTA TTT	
HG001_00316 For	TAATACGACTCACTATAGGG ATAT AAGCTTTTGTTATG	<i>In vitro</i> transcription
HG001_00316 Rev	CACTCTTTAATACTTTACCGTT	
HG001_01242 For	TAATACGACTCACTATAGGG TATA ATAAAATGTATAC	<i>In vitro</i> transcription
HG001_01242 Rev	CTTTCGAAATAATTGTTAC	
HG001_02545 For	TAATACGACTCACTATAGGG AAAG GTTACCCCGAA	<i>In vitro</i> transcription
HG001_02545 Rev	TGCTTAATTGATAGAGATGTTATAC G	
icaR 5' for	TAATACGACTCACTATAGGG AAAT ATTTGTAATTGCCAACTTAATTT	<i>In vitro</i> transcription
icaR 5' rev	ATCATCAAGTGTTGTACCGTCATAC CCCTT	
icaR 3' for	TAATACGACTCACTATAGGG CTTT ATATTTGTGAATGGTTAAGTTTGTC TTTGAAC	<i>In vitro</i> transcription
icaR 3' rev	AAAAAGCGCCTATGTCATGATTTAC CATCA	
icaR 5'long rev	TCAAAACAACATTTAACACTTTG	<i>In vitro</i> transcription
<i>treB</i> For	TAATACGACTCACTATAGGG TAAG ACATTAAGGGAGTGAA	<i>In vitro</i> transcription
<i>treB</i> Rev	TGGTATAAAAATATCCCCC	
<i>glcU</i> _2 Rev	AGCGTCGTACCACGA	Toeprint experiment
<i>fn3K</i> Rev	AAATGTATCCGTATC	Toeprint experiment
<i>SphI</i> - <i>RsaI</i> For	<u>CGCGCATGCG</u> ACAGCCACCTCCTA AAAG	Clonage <i>RsaI</i> and its promoteur in pCN51
<i>PstI</i> - <i>RsaI</i> Rev	<u>CGCCTGCAG</u> GAATAACAATACAAC TGAAGAC	
<i>RsaI</i> _A_LIC	GACGACGACAAGAGTCGTACGATA ACATTGGTTTACC	Deletion of <i>RsaI</i>
<i>RsaI</i> _B	CCTTTCTATTTGTCGGCTAG	
<i>RsaI</i> _C	CTAGCCGACAAATAGAAAGGCCAC AAAATTTTACGATTTTTTT	
<i>RsaI</i> _D_LIC	GAGGAGAAGCCCGGTCGTGTTTGT GGAAGACTTAA	
<i>RsaI</i> _E	GCACCTTCTCCTCGTATTGCT	Verification of deletion in strain <i>S. aureus</i> 132
<i>RsaI</i> _F	GCATTGGGAGTGTCATATGTTG	

BamHI-Rsal	<u>GGATCC</u> GATATTACTTGTAACAGG GG	Cloning of Rsal in pES
EcoRI-Rsal	<u>GAATTC</u> GAATAACAATACTGAA GAC	
ovlp_Rsal_mut5_Rev	CCTTTCATTTCATTATTATATTGAAG ATGTTGTGATGTAAGTT	Mutagenesi s of <i>rsal</i> in pES
ovlp_Rsal_mut5_For	CATCACAACATCTTCAATATAATAA TGAATGAAAGGTTTCTATTTGTCGG CTAGCACG	
Teg24-5'F-MluI	CCGACGCGTGTAATAATTACATGTT AATACGTAGTATTAATGGC	Deletion of <i>rsal</i> in <i>S.</i> <i>aureus</i> HG001
Teg24-5'R-NcoI	GCGCCATGGTTTTTACGATTTTTTT TAAAAGATATTGAAAATG	
Teg24-3'F-NcoI	ACTCCATGGGAAGAAGTCTTCAGT TGTATTGTTGATTCT	
Teg24-3'R-BglII	AATAGATCTTTTACTATTCCTTACAT ACTTTGTCTGTTTTCT	
EcoRI-T7-Rsal For	CGCGAATTCT TAATACGACTCACTA TAGTAACAGGGGGAGCGATTAA	Cloning of <i>rsal</i> under the T7 promoter in pUC18
PstI-Rsal-Rev	CGCCTGCAGAAAAAATCGGCTAG TCACG	Cloning of <i>rsal</i> under the T7 promoter in pLUG220
PstI-Rsal For	CGCCTGCAGGACAGCCACCTCCTA AAAG	
PstI-MS2-Rsal For	CGCCTGCAGCGTACACCATCAGG GTACGTTTTTCAGACACCATCAGG GTCTGTTTGTAACAGGGGGAGCGA	Constructio n MS2- Rsal in pCN51- P3
BamHI-MS2-Rsal Rev	CGCGGATCCAACAATACTGAA GAC	
Rsal mut1 For	TTCTAATACGACTCACTATAGTAAC AGAGCGATTAAA	Quickchang e Mutagenesi s of <i>rsal</i>
Rsal mut1 Rev	TTTAATCGCTCTGTTACTATAGTGA GTCGTATTAGA	
Rsal mut2 For	GAGCGATTAAACAAATAGAGCTAT ATAACAGATAGC	Quickchang e mutagenesi s of <i>rsal</i>
Rsal mut2 Rev	GCTATCTGTTATATAGCTCTATTTG TTTAATCGCTC	
Rsal mut4 For	TTACATCACAACATCCTATTTGTCG GCTAGCAC	Quickchang e mutagenesi s of <i>rsal</i>
Rsal mut4 Rev	GTGCTAGCCGACAAATAGGATGTT GTGATGTAA	
BamHI-02520-lacZ For	CGCGGATCCGTAAGAAGTTAATAA ATA	Translation al fusion
BamHI-02520-lacZ Rev	CGCGGATCCGATGTCACCAATTAA	

	TCCT	
BamHI-01242-lacZ For	CGC <u>GGATCC</u> ACAAGGAAAGTAATG AAATA	Translation al fusion
BamHI-01242-lacZ Rev	CGC <u>GGATCCC</u> CATTTATTTAAACT AAGC	
BamHI- <i>treB</i> -lacZ For	CGC <u>GGATCC</u> AAGACATTAAGGGAG TGAAA	Translation al fusion
BamHI- <i>treB</i> -lacZ Rev	CGC <u>GGATCC</u> TAAACGTAATCGTGT TACAC	
BamHI- <i>g/cU</i> _2-lacZ For	GCGGGAT <u>CCCA</u> AGTGTCAATTAAG GGGAG	Translation al fusion
BamHI- <i>g/cU</i> _2-lacZ Rev	CGC <u>GGATCC</u> TTGGTTGTAAGGTCC ACCGC	
BamHI- <i>fn3K</i> -lacZ For	CGC <u>GGATCC</u> TATAAAGTAAGTAGA TTGGA	Translation al fusion
BamHI- <i>fn3K</i> -lacZ Rev	GCGGGAT <u>CCTG</u> TTTCGACTCGATA TGCTT	

Restriction sites sequences are underlined, the T7 promoter sequence is in bold. The MS2 sequence is highlighted in light grey

Table EV4: The list of RNAs sequenced using either MS2-Rsal or two mutant RNAs (MS2-Rsal mut2, and MS2-Rsal mut4) as the baits. As described in the supplementary methods, we have considered all the MAPS experiments comprising three MS2-Rsal, two MS2 controls, and one experiment for each mutant (MS2-Rsal mut2 and mut4). The fold change and p-values were calculated using the same procedure with DEseq2. The whole data sets obtained for the MS2-Rsal, MS2-Rsal mut2, and MS2-Rsal mut4 were given in the panels named MS2RsalvsMS2contr.complete-anno, MS2Rsalmut2vsMS2contr.complete-anno, and MS2Rsalmut4vsMS2contr.complete-anno, respectively. The listed RNAs that were significantly enriched with a good p-value (< 0,001) together with MS2-Rsal, MS2-Rsal mut2, and MS2-Rsal mut4 were given in the panels named MS2RsalvsMS2contr.up-anno, MS2Rsalmut2vsMS2contr.up-anno, and MS2Rsalmut4vsMS2contr.up-anno, respectively.

<https://www.embopress.org/doi/full/10.15252/emj.201899363>

Table EV5: RNAs sequenced using MS2-Rsal, MS2-Rsal mut2 or mut 4 as baits. The MAPS experiments were carried out with the wild-type MS2-Rsal, and two RsaI mutants (nucleotides C22 to G29 were deleted in mut2, while nucleotides 75 to 113 were deleted in mut4). The enrichment value corresponding to the fold changes (MS2-Rsal /MS2 control) were calculated using DEseq2 (Varet et al., 2016). In these experiments, a significant RNA enrichment was considered at a threshold of 3. The fold changes below 3 are represented by red-bold-underlined characters while the fold changes that have been significantly affected in the experiments performed with RsaI mutants are in bold characters. The whole data sets are shown in Table EV5

Target	Product	Fold Change MS2-Rsal/MS2	Fold Change MS2-Rsal mut2/MS2	Fold Change MS2-Rsal mut4/MS2
HG001_02520	XRE family transcriptional regulator	23,583	9,208	5,431
HG001_02704	IcaR	17,598	15,356	<u>2,068</u>
HG001_02347	RsaI	15,029	5,793	12,719
HG001_02756	RsaG	10,446	<u>1,941</u>	3,885
HG001_01242	Hypothetical protein	8,817	12,158	<u>1,621</u>
HG001_00847	RsaE	5,276	<u>0,49</u>	<u>0,652</u>
HG001_02628	Fn3K	4,609	3,119	<u>1,618</u>

HG001_00316	Peptidase propeptide and YPEB domain protein	4,002	2,733	<u>1,969</u>
HG001_00687	DegV domain-containing protein	3,759	5,047	<u>1,429</u>
HG001_00487	dck Deoxyadenosine(dk)deoxycytidine kinase	3,393	3,151	3,047
HG001_02293	GlcU_2	3,183	3,087	<u>2,006</u>
HG001_02545	Hypothetical protein	3,012	3,328	<u>1,236</u>
HG001_00942	Putative cell wall binding lipoprotein	2,992	3,354	<u>0,628</u>

Table EV6: Comparative transcriptomic analysis. The data were obtained from RNA samples prepared from the wild-type strain HG001, and the mutant HG001 Δ rsal strain and from the same mutant strain transformed with a plasmid expressing RsaI under the control of its own promoter (HG001 Δ rsal pCN51::PrsaI). These experiments were carried out in triplicates. The comparison of the data were done with DSeq2: total RNAs sequenced from HG001 and HG001 Δ rsal are listed in tab “Total DrsaI/WT”, total RNAs sequenced between HG001 Δ rsal and HG001 Δ rsal pCN51::PrsaI are listed in tab “Total DrsaI/comp”, and total RNAs sequenced between HG001 and HG001 Δ rsal pCN51::PrsaI are listed in tab “Total comp /WT”.

<https://www.embopress.org/doi/full/10.15252/emboj.201899363>

Table EV7: RsaI-dependent deregulation. Only the RNAs, which have a significant expression change as the result of the deletion of rsal, are represented. The data are coming from the transcriptomic analysis obtained from HG001 Δ rsal and HG001 Δ rsal pCN51::PrsaI (Table S6). The RsaI-dependent repression is in blue while the RsaI-dependent activation is in red. The known biological functions are also given. Only the RNAs, which have a two-fold change ratio between the two conditions, have been considered.

<https://www.embopress.org/doi/full/10.15252/emboj.201899363>

Table EV8: Strains and plasmids used in the study

Strains	Relevant characteristics	Reference
<i>Escherichia coli</i>		
DC10B	<i>dcm</i> mutant ; used for cloning of plasmids which can be directly transformed in <i>S. aureus</i> cells	Monk et al. (2012)
DH5 α	Mutant <i>endA1</i> , <i>hsdR17</i> , <i>recA</i>	Taylor (1993)
XL1-blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proABlacIqZ</i> Δ M15 Tn10 (Tet)]	Stratagene
<i>Staphylococcus aureus</i>		
HG001	Derivative of 8325-4, <i>rsbU</i> restored RN1, <i>agr</i> positive	Herbert et al. (2010)
4220	Restriction mutant of 8325-4 that accepts foreign DNA	Kreiswirth et al. (1983)
132	MRSA clinical strain with the capacity to form proteinaceous biofilm in TSB-glucose, and PIA-PNAG-dependent biofilm in TSB-NaCl	Vergara-Irigaray et al. (2009)
HG001 Δ <i>rsal</i>	In frame deletion of <i>rsal</i> in HG001	This study
132 Δ <i>rsal</i>	In frame deletion of <i>rsal</i> in 132	This study
HG001 Δ <i>rsaG</i>	In frame deletion of <i>rsaG</i> in HG001	This study
HG001 Δ <i>ccpA</i>	In frame deletion of <i>ccpA</i> in HG001	Seidl et al. (2006)
HG001 Δ <i>codY</i>	In frame deletion of <i>codY</i> in HG001	Pohl et al. (2009)
HG001 Δ <i>ssrAB</i>	In frame deletion of <i>ssrAB</i> in HG001	Villanueva et al. (2018)
HG001 Δ <i>saeRS</i>	In frame deletion of <i>sarRS</i> in HG001	Villanueva et al. (2018)
HG001 Δ <i>hptRS</i>	In frame deletion of <i>hptRS</i> in HG001	Villanueva et al. (2018)
Plasmids		
pCN51	Shuttle vector	Charpentier et al. (2004)
pLUG220:: <i>rpoB</i>	pLUG220:: <i>PrpoB</i>	Romilly et al. (2014)
pMAD	Thermosensitive origine of replication, constitutively expressed <i>bgaB</i> gene	Arnaud et al. (2004)
pMAD-CM	pMAD containing the <i>cat</i> gene from pSK236 using NaeI	Memmi et al. (2008)
pES	pCN47:: <i>PblaZ</i>	This study
pUC18:: <i>T7-rsaG</i>	<i>rsaG</i> under T7 promoter	Geissmann et al. 2009
pUC18:: <i>T7-rsal</i>	<i>rsal</i> under T7 promoter	This study
pUC18:: <i>T7-rsal</i> mut1,2,3,4	<i>rsal</i> mut1, 2, 3, 4 under T7 promoter	This study
pCN51:: <i>Prsal</i>	pCN51 expressing <i>Rsal</i> under its own promoter	This study
pES:: <i>Prsal</i>	pES expressing <i>Rsal</i> under the <i>PblaZ</i> promoter	This study
pES:: <i>rsal</i> mut5	pES expressing <i>Rsal</i> mut5 the <i>PblaZ</i> promoter	This study
pCN51:: <i>P3</i>	pCN51 with P3 promoter	Tomasini et al. (2017)
pCN51:: <i>P3-MS2</i>	pCN51 expressing MS2 tag under P3 promoter	Tomasini et al. (2017)
pCN51:: <i>P3-MS2-rsal</i>	pCN51 expressing MS2- <i>Rsal</i> under P3 promoter	This study

pCN51::P3::MS2- <i>rsal</i> mut4	pCN51::P3::MS2- <i>rsal</i> mut4 under P3 promoter	This study
pUT7:: <i>icaR</i>	T7 promoter- <i>icaR</i> allowing the in vitro T7 transcription of the full-length <i>icaR</i> mRNA	Ruiz de los Mozos <i>et al.</i> (2013)
pUT7:: <i>icaR</i> SUBST	T7 promoter- <i>icaR</i> allowing the in vitro T7 transcription of the full-length <i>icaR</i> mRNA carrying the substitution of the UCCCCUG motif by AGGGGAC	This study
pLUG220::PrpoB:: <i>glcU_2::lacZ</i>	pLUG220 with translational fusion <i>glcU::lacZ</i>	This study
pLUG220::PrpoB:: <i>glcU_2::lacZ::rsal</i>	pLUG220 with translational fusion <i>glcU::lacZ</i> and expressing <i>Rsal</i> under its own promoter	This study
pLUG220::PrpoB:: <i>fn3K::lacZ</i>	pLUG220 with translational fusion <i>fn3K::lacZ</i>	This study
pLUG220::PrpoB:: <i>fn3K::lacZ::rsal</i>	pLUG220 with translational fusion <i>fn3K::lacZ</i> and expressing <i>Rsal</i> under its own promoter	This study
pLUG220::PrpoB:: <i>fn3K::lacZ::rsal</i> mut4	pLUG220 with translational fusion <i>HG001_fn3K::lacZ</i> and expressing <i>Rsal</i> mut4 under its own promoter	This study
pLUG220::PrpoB:: <i>treB::lacZ</i>	pLUG220 with translational fusion <i>treB::lacZ</i>	This study
pLUG220::PrpoB:: <i>treB::lacZ::rsal</i>	pLUG220 with translational fusion <i>treB::lacZ</i> and expressing <i>Rsal</i> under its own promoter	This study
pLUG220::PrpoB:: <i>HG001_01242::lacZ</i>	pLUG220 with translational fusion <i>HG001_01242::lacZ</i>	This study
pLUG220::PrpoB:: <i>HG001_01242::lacZ::rsal</i>	pLUG220 with translational fusion <i>HG001_01242::lacZ</i> and expressing <i>Rsal</i> under its own promoter	This study
pLUG220::PrpoB:: <i>HG001_02520</i>	pLUG220 with translational fusion <i>HG001_02520::lacZ</i>	This study
pLUG220::PrpoB:: <i>HG001_02520::lacZ::rsal</i>	pLUG220 with translational fusion <i>HG001_02520::lacZ</i> and expressing <i>Rsal</i> under its own promoter	This study

2. Chapitre 2 : RsaG, un sARN dérivé d'une région 3'UTR régule l'homéostasie rédox en réponse au glucose-6-phosphate chez *S. aureus*

Lors du processus infectieux *S. aureus* doit s'adapter à l'environnement hostile de son hôte en se protégeant contre les réactions immunitaires, et en modulant son métabolisme pour répondre à la diversité nutritive ou aux carences dépendantes du tissu infecté. Les régulateurs protéiques du métabolisme chez des bactéries non pathogènes voient leurs fonctions étendues chez les bactéries pathogènes pour réguler à la fois le métabolisme et la virulence (Richardson, 2019). Par ailleurs, les sARN qui agissent au niveau post-transcriptionnel, sont aussi impliqués dans la régulation de la physiologie et de la réponse adaptative ce qui fait d'eux de bons candidats pour lier la régulation du métabolisme à celle de la virulence.

Dans cette étude, nous montrons que le sARN RsaG dérive d'une région 3'UTR, qui est essentiellement conservée chez *S. aureus* et les espèces très proches. La transcription de RsaG est induite par le G6P lorsque *S. aureus* est internalisé dans deux types de cellules humaines, et lorsque *S. aureus* est mis en présence de mucus intestinal ou pulmonaire humain contenant des sucres phosphorylés (Sit et al., 2015). Nous avons défini le targetome de RsaG en utilisant la technique MAPS qui consiste à une purification d'affinité de RsaG avec ses cibles, suivie d'un séquençage de l'ARN à haut débit (Tomasini et al., 2017). La validation des interactions ARN-ARN a montré que RsaG utilise différents mécanismes pour réguler l'expression de ses cibles qui sont des ARNm ainsi qu'un autre sARN, RsaI. RsaG contrôle des gènes impliqués dans le métabolisme comme *ccpA*, *rex*, *fn3K*, les opérons *arcABDCR*, *thiDME* ou *ldhI*, divers transporteurs d'ions (*mscL*, *kdpBAF/kdpED*) ainsi que d'autres gènes impliqués dans la virulence, *sarA* et *tcaR*. RsaG semble ainsi coordonner le métabolisme carboné et la virulence de *S. aureus* en fonction de l'homéostasie rédox et en réponse à l'entrée du G6P. Il contribue non seulement à étendre le régulon de HptRS mais aussi à ajouter un autre niveau de régulation au TCS.

The 3'UTR-derived sRNA RsaG in *Staphylococcus aureus* coordinates redox homeostasis and metabolism adaptation in response to G6P entry

Emma Desgranges¹, Lucas Herrgott¹, Stefano Marzi¹, Alejandro Toledo-Arana², Karen Moreau³, François Vandenesch³, Pascale Romby¹ and Isabelle Caldelari¹

¹Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, UPR9002, Strasbourg, France

²Université Claude Bernard Lyon 1, Inserm, Centre international de recherche en infectiologie, U1111, Lyon, France

³Public University of Navarra, Instituto de Agrobiotecnología, Pamplelona, Spain

*To whom correspondence should be addressed. E-mail: i.caldelari@ibmc-cnrs.unistra.fr; Tel: +33388417051

SUMMARY

RsaG is a 3' untranslated region (3'UTR) derived sRNA from the universal *uhpT* gene in the opportunistic pathogen *Staphylococcus aureus*. UhpT encodes a glucose-6-phosphate (G6P) transporter expressed in response to extracellular G6P. The transcript *uhpT-RsaG* underwent a rapid degradation from 5' to 3' end by the action of two exoribonucleases, J1 and J2, which were blocked by a stable hairpin structure at the 5' end of RsaG leading to its accumulation. RsaG together with *uhpT* was induced when bacteria are internalized into host cells or in presence of mucus-secreting cells. We identified the transcriptional factors Rex, CcpA, SarA and the sRNA RsaI, as the main RsaG-dependent regulated RNA targets. These regulatory circuits allow the bacteria to control redox homeostasis and to adjust the metabolism to changing environmental conditions. RsaG inactivates the lactate fermentation and conversely activates amino acid catabolism, acetoin synthesis, and alcohol fermentation. RsaG used different molecular mechanisms either to stabilize, to degrade, or to repress translation of its mRNA targets. Finally, while RsaG is conserved only in closely related species as *S. argenteus* and *S. schweitzeri*, we found that the *uhpT* 3'UTR of the ape pathogen *S. simiae* encoded a sRNA, whose sequence is highly different, and which did not sense G6P. Our results support the hypothesis that 3'UTRs of bacterial mRNAs evolved to enable adaptation to host niches.

INTRODUCTION

Staphylococcus aureus is usually described as an extracellular opportunistic pathogen, infecting a wide range of organs and tissues. However, it also invades and replicates in various phagocytotic or non phagocytotic host cells (Hamza and Li, 2014). To be successful as a pathogen, this bacterium needs to adapt with hostile environments of the host and must acquire imposed nutriment for its survival. As a consequence, pathogens have to re-configure their metabolism in order to produce their virulence factors at the appropriate time and space. Similarly, to any bacteria, staphylococcal genome encodes several transporters for metabolites (sugars, metals, amino-acids, etc...), among them, glucose transporters are primordial in virulence as a mutant deprived of the four of them is avirulent in a murine skin infection model (Vitko et al., 2016). These transporters are strictly regulated depending on the available nutrient sources. For instance, in glucose-limiting conditions, when *S. aureus* penetrates into host cells, it could utilise cytosolic glucose-6-phosphate (G6P), whose uptake relies on the expression of the hexose phosphate antiporter UhpT stimulated by the two-component system (TCS) HptRS (see below, and Garzoni et al., 2007; Park et al., 2015; Yang et al., 2016).

Metabolic or virulence genes are tightly controlled by transcriptional factors (TF), or TCS to adapt to various environmental situations (Richardson, 2019; Villanueva et al., 2018). Indeed, *S. aureus* possesses a plethora of TF such as CcpA, CodY and Rex, responding to diverse stimuli (for review see Richardson, 2019). In response to glucose, the catabolite control protein, CcpA, represses expression of genes, whose products act in TCA cycle, gluconeogenesis, or amino acid catabolism and activates glycolytic genes (Seidl et al., 2009). It mediates the catabolic repression allowing bacteria to use the preferred carbon source. Besides, CodY, sensing both branched amino acids and GTP, represses genes involved in amino acid synthesis and secreted virulence factors (Richardson, 2019). Rex is considered as the central regulator of anaerobic metabolism and inhibits genes encoding fermentative enzymes such as lactate deshydrogenase Ldh_1 or alanine deshydrogenase Ald1 under aerobic conditions (Pagels et al., 2010). At high NADH/NAD level, when Rex is inhibited, survival of *S. aureus* depends on the switch to a fermentative metabolism in response to NO or to face the aerobic low redox potential of host cytosol (Christmas et al., 2019; Crooke et al., 2013). While immune radicals damage the terminal oxidase of the electron transport chain leading to the disruption of redox homeostasis, *S. aureus* produces predominantly L or D-lactate as defences. The three lactate deshydrogenases Ldh_1, Ldh_2 or Ddh generate respectively L or D-lactate from glucose fermentation and support NAD⁺ recycling. Therefore, CcpA activates transcription of the key enzyme *ldh_1* only if *rex* is repressed (Crooke et al., 2013). In contrast, when bacteria are internalized and glucose unavailable, Ldh_1 is not active,

and the amino acids are preferred as nutrients reflecting the derepression of the CodY and CcpA regulons (Michalik et al., 2017). Besides the staphylococcal core genome encodes 17 two- or three-component systems (TCS); among them, only *walRK* regulating bacterial division is essential for growth (Martin et al., 1999; Villanueva et al., 2018). These TCS allow bacteria to sense and respond to diverse stimuli and to various metabolites or small compounds in order to adapt to specific niches (Villanueva et al., 2018). In *S. aureus*, when extracellular G6P is the alternative carbon source, HptRS senses it and activates the *uhpT* transporter by binding of the response regulator HptR onto a consensus sequence called HptR box, localized between -67 and -96 before the beginning of the transcriptional start site (Yang et al., 2016). The disruption of *hptRS* gene impairs survival into various kinds of host cells, suggesting that sensing and uptake of G6P is critical for *S. aureus* adaptation into host cells (Park et al., 2015).

Post-transcriptionally, regulatory RNAs (sRNA) cooperate with TF and TCS to quickly adjust the bacterial physiology to the surrounding conditions and especially to the availability of carbon sources (Bobrovskyy and Vanderpool, 2013; Desgranges et al., 2019; Wagner and Romby, 2015). They hybridize to their RNA targets by imperfect complementarities to affect positively or negatively their stability or translation. They differ from their origin, sizes and functions. The vast majority of sRNAs derives from intergenic regions (IGR) and possesses their own promoter, otherwise they are processed from 5' or 3' untranslated regions (UTR) of mRNAs (Desgranges et al., 2019; Miyakoshi et al., 2015). Their transcription is often induced by sensing of external stimuli via TCS or transcriptional factors (Brosse and Guillier, 2018). Recently the targetomes of different staphylococcal sRNAs were identified and reflected the diversity of pathways in which they intervene (e.g., Augagneur et al., 2020; Bronesky et al., 2019; Lalaouna et al., 2019; Rochat et al., 2018; Tomasini et al., 2017). Several of them are part of large regulatory networks that connect major regulatory proteins of virulence gene expression (CodY, CcpA, SigB, AgrA, SrrAB) in response to wide arrays of metabolic and environmental signals (Desgranges et al., 2019). For instance, the sRNA RsaI is repressed by CcpA in presence of glucose. When glucose decreases, RsaI downregulates glucose uptake and activates fermentation, energy production, and NO detoxification. Interestingly, RsaI binds to other sRNAs, such as RsaD and RsaE (Bronesky et al., 2019). SrrA, which senses nitric oxide, activates RsaE to repress many TCA enzymes, which reduces NADH production. In turn, low activity of the TCA cycle has a positive effect on the *agr* system to adjust virulence factor production under stress conditions (for a review, Marincola et al., 2019).

Here we show that the 3'UTR-derived sRNA, RsaG, conserved in *S. aureus* and in very close relative species, is induced specifically by G6P either produced by the host cells when *S. aureus* is internalized or in presence of mucus composed of phosphorylated sugars (Sit et al., 2015). This sRNA uses different mechanisms to regulate its mRNA targets, such as *ccpA*, *rex*

or *ldh_1*, connecting glucose utilization, redox homeostasis, and metabolism adaptation. The consequences of these regulatory networks on *S. aureus* pathogenesis will be discussed.

RESULTS

RsaG expression is induced by G6P upon the *uhpT* promoter

RsaG was first predicted as an intergenic region conserved only in *S. aureus* (Geissmann et al., 2009). RsaG is located downstream the *uhpT* gene and was shown to accumulate in the late exponential phase of growth in rich medium (Geissmann et al., 2009) or in presence of G6P (Bronesky et al., 2019). In a first experiment, RT-PCR experiments were performed with oligonucleotides amplifying either the *uhpT* mRNA alone or a transcript containing both *uhpT* and RsaG. This approach did not detect a full transcript and only *uhpT* coding sequence was identified by RT-PCR (Figure S1). Because no obvious predicted promoter sequence was detected for *rsaG* and knowing that *uhpT* is induced in response to extracellular G6P, we made the hypothesis that RsaG might be the result of a specific processing within the 3'UTR of *uhpT*.

We then tested RsaG expression by Northern blot analysis in MHB without or with G6P in the wild-type (HG001), the deleted mutant of the HptRS system (HG001 $\Delta hptRS$) or of the *uhpT* promoter containing the hptR box recognized by the transcriptional regulator HptR (HG001 $\Delta PuhpT$) (Figure 1A). In MHB, in which starch is added as carbon source, RsaG is constitutively and weakly expressed. In contrast when the medium is supplemented with G6P, we observed a huge stimulation of RsaG transcription, which is almost completely abolished in the HG001 $\Delta hptRS$ or HG001 $\Delta PuhpT$ mutant strains (Figure 1A). Furthermore, we also detected bands with higher molecular weights suggesting that a longer transcript containing RsaG exists. We then constructed various transcriptional fusions carrying *gfp* in place of RsaG and determined the effect of successive deletions within *uhpT* and its promoter on the GFP synthesis (Figure 1B). All the constructs were expressed from a plasmid that was transformed into HG001. The data were analysed using either fluorescent microscopy (Figure 1D) or western blot using an anti-GFP antibody (Figure 1C). Both approaches provided similar results as GFP or fluorescent bacteria were detected only with *uhpT* full length containing its promoter and the *hptR* box.

Taken together, these data evoked that RsaG is transcribed together with *uhpT* under the control of the same promoter responding to G6P via HptRS, and that RsaG accumulated after a rapid degradation of *uhpT*, while we were unsuccessful to detect the co-transcript by RT-PCR (Figure S1).

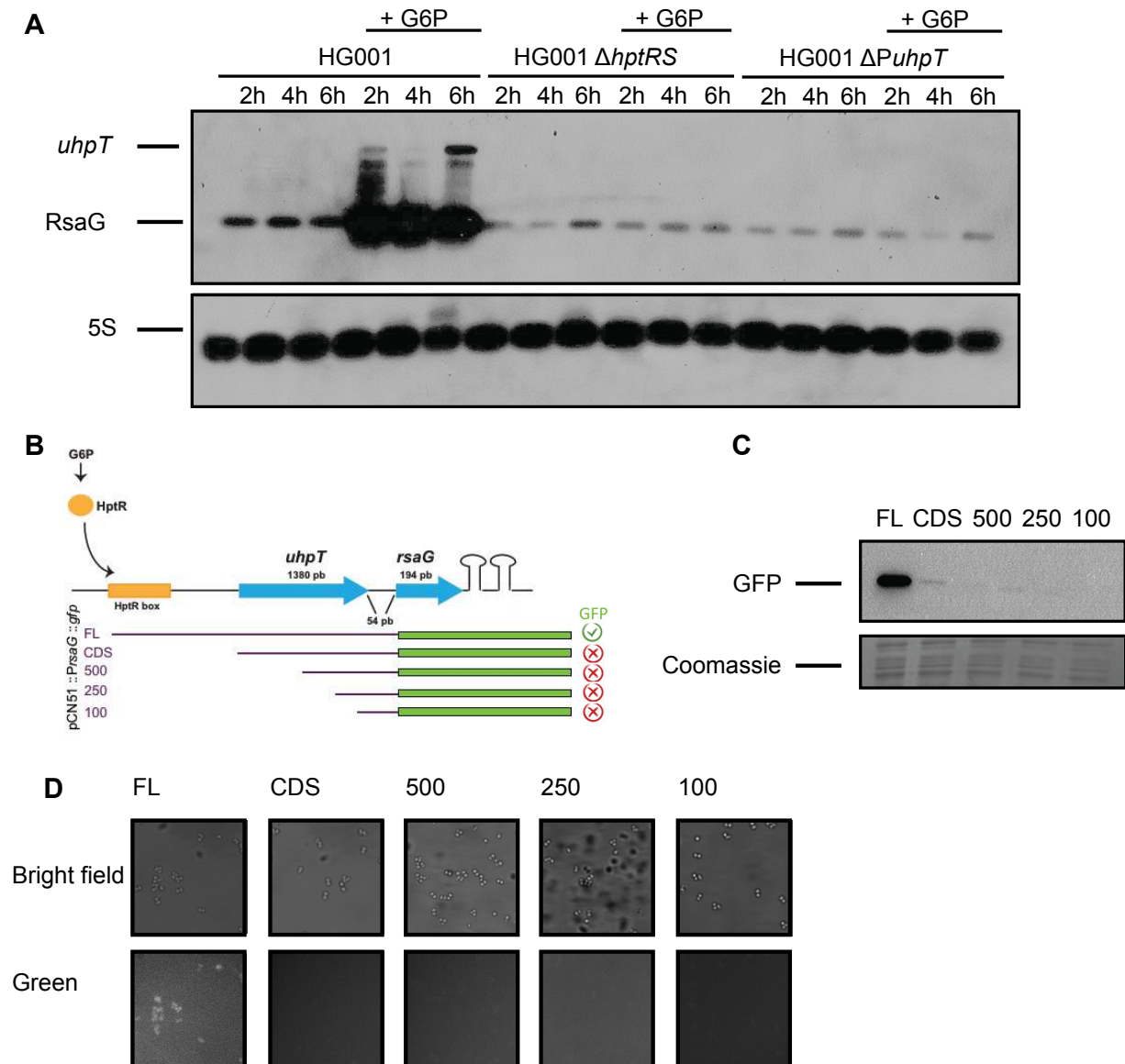


Figure 1

Figure 1: RsaG is expressed under *uhpT* promoter activation by the HptRS two component system upon sensing of external glucose-6-phosphate (G6P).

A. Northern blot analysis of RsaG in HG001 wild type strain, HG001 $\Delta hptRS$ and HG001 $\Delta PuhpT$. Total RNA was extracted at 2, 4 and 6 h of growth in MHB medium with or without the addition of 0.5% G6P. 5S rRNA (5S) was used as loading control. However, for this control, we used aliquots of the same RNA preparations, but the migration of the samples was performed in parallel to the experiments on a separate agarose gel because RsaG and 5S rRNA have very similar sizes. B. Genomic context of *uhpT*-RsaG locus and mapping of the different DNA fragments transcriptionally fused to *gfp* in pCN51 plasmid. C. Western blot experiment detecting GFP synthesis in the different transcriptional fusion constructs (FL, CDS,

500, 250, 100) expressed in HG001 wild-type strain. Total proteins were separated by SDS-PAGE (10%) and were revealed using a GFP-specific antibody. We used Coomassie staining as loading control. However, for this control, we used aliquots of the same protein preparations, but the migration of the samples was performed in parallel to the experiments on a separate SDS-PAGE gel. D. Epifluorescence microscopy and bright field images of HG001 containing different promoters constructions with *gfp*. FL: full length, CDS: coding sequence.

RsaG is derived from the maturation of the 3'UTR of *uhpT*

While only primary transcripts possess a 5'-triphosphate end, we used Terminator™ to discriminate the phosphorylation status of the 5' end of RsaG as the enzyme processes just 5' monophosphate RNA. Total RNA extracted from HG001 grown in BHI for 4 h (when RsaG is present) was treated with Terminator or not and run on an agarose gel. As control experiments, ethidium bromide staining revealed that the mature 16S and 23S rRNAs were fully degraded (Figure 2A, left panel). As a positive control, Northern blot analysis showed that the *bona fide* non-coding RNA RsaI is resistant to the Terminator treatment (Figure 2A, middle panel). The experiment performed with RsaG showed that its level was only slightly reduced in presence of Terminator (see shorter exposure time of the autoradiography). This effect might be due to the fact that the activity of terminator is strongly altered by the presence of a helical structure present at the 5' end, as found for 5S rRNA, which is also fully resistant to the enzymatic treatment (Figure 2A). In addition, the *uhpT*-RsaG transcript and shorter fragments were completely digested by Terminator (Figure 2A right panel, longer exposure). This data suggested that the degradation of the long *uhpT* transcript might be processed by a 5' exoribonuclease, which might be partially blocked by the first stem-loop structure of RsaG (Figure 2B).

In the following experiments, we have tested the involvement of the major ribonucleases from *S. aureus*, namely RNase P, RNase III, RNases J1/J2 and RNase Y (Figures 2C, S2). An *in vitro* transcribed *uhpT*-*rsaG* RNA was used in the assay with the RNA component of RNase P, which is sufficient to induce cleavage (Guerrier-Takada and Altman, 1984), The data showed that the co-transcript was not cut in contrast to the pre-mature tRNA, which was used as a positive control (Figure S1A). For the other enzymes, Northern blot experiments were performed on total RNAs prepared from various mutant strains carrying deletion at specific genes encoding RNase III (HG001 Δ *rnC*), RNase Y (HG001 Δ *my*), and RNases J1 or J2 (RN4220 Δ *j1*, Sa624 Δ *j1* or Sa624 Δ *j2*), and their parent wild-type strains (HG001, RN4220 or Sa624). The strains were grown in MHB containing G6P, and RsaG was revealed using a specific probe against RsaG. While RNase Y and RNase III did not alter the maturation of *uhpT*-RsaG transcript (Figure S2C), a significant decrease of RsaG level was observed in the single deletion mutants of RNases J1 or J2 (Δ *j1*, Δ *j2*) and concomitantly an

accumulation of *uhpT*-RsaG was visualized (Figures 2C, S2B). We did not manage to obtain relevant results with the double mutant strain ($\Delta j1\Delta j2$), due to the fact that the growth was strongly impaired (Linder et al., 2014).

Taken together the data strongly suggested that RsaG is a product of the rapid degradation of *uhpT* mediated by the 5'-3' exoribonuclease activities of RNases J1/J2, and that its 5' hairpin structure most probably blocks the progression of the exoribonuclease.

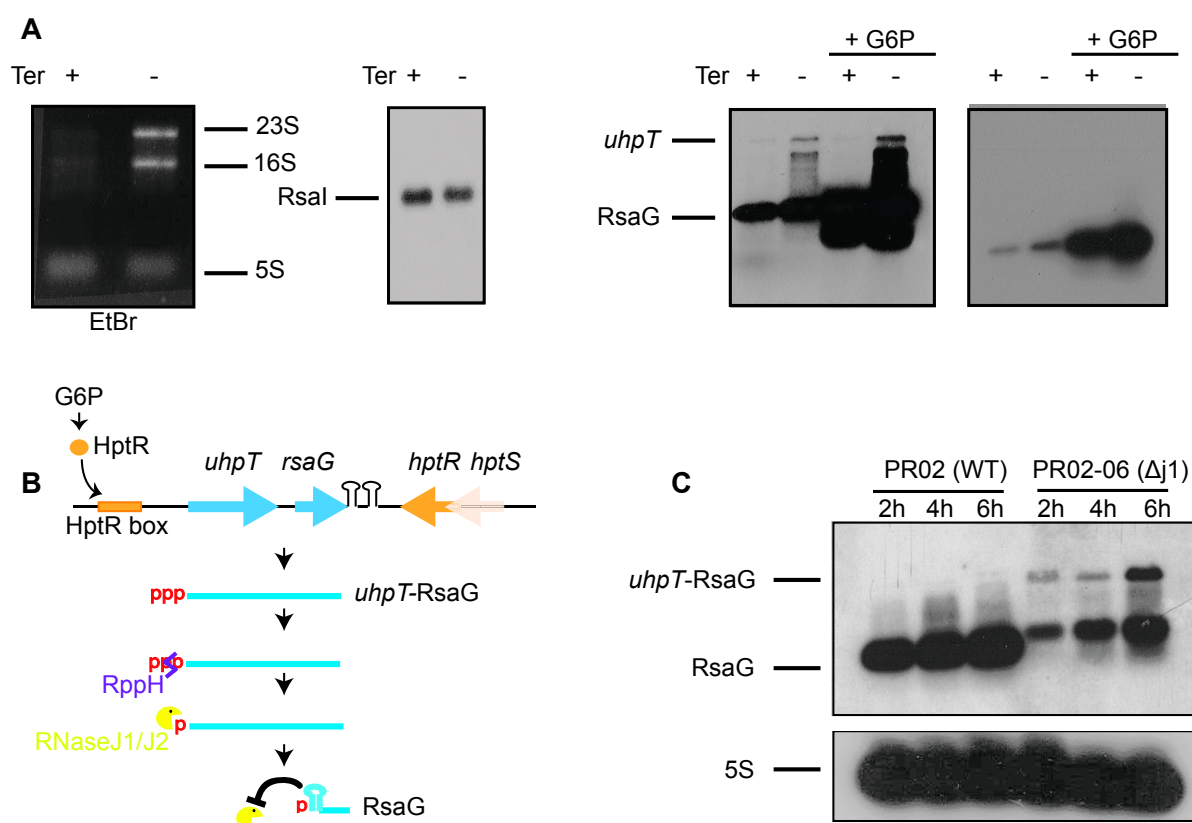


Figure 2

Figure 2: RsaG is matured from *uhpT* by 5'-3'exo-ribonuclease J1-J2.

A. Determination of the 5' end status of RsaG sRNA. 10 μ g of total RNA extracted after 4h of growth in MHB medium (+/- glucose-6-phosphate, G6P) was treated with the Terminator™ 5'-Phosphate-Dependent Exonuclease (+/- Ter). RsaI and RsaG were revealed by Northern blot analysis using specific probes. 16S and 23 rRNAs were used as positive controls and 5S rRNA as a negative control. EtBr: Ethidium bromide staining. Two different exposures are presented for RsaG to detect the full *uhpT*-RsaG co-transcript (left panel). B. Genomic context of *uhpT*-RsaG locus and the sequential steps involved in the maturation. C. Northern blot analysis of RsaG in PR02 wild type strain and PR02-06 ($\Delta j1$). Total RNA was extracted at 2, 4 and 6 h of growth in MHB medium supplemented with 10 mg/L uracil and 0.5% glucose-6-phosphate. 5S rRNA (5S) was used as loading control. However, for this control, we used aliquots of the same

RNA preparations, but the migration of the samples was performed in parallel to the experiments on a separate agarose gel because RsaG and 5S rRNA have very similar sizes.

The *uhpT* mRNA and RsaG are enhanced under conditions mimicking the infection

G6P is the activated form of glucose in numerous cellular metabolic pathways, and hexose-phosphate sugars are predominant in the cytosolic environment (Chico-Calero et al., 2002). Upon internalization into host cells, *S. aureus* senses cytosolic G6P and may consume it as a carbon source. We first compared the growth rate of HG001, HG001 Δ rsaG, and HG001 Δ hptRS strains in LB (without glucose) or in presence of 0,5 % G6P. The growth speeds of both the wild type HG001 strain and the Δ rsaG deletion mutant were identical in presence of G6P but significantly faster than the growth rate of HG001 Δ hptRS under identical conditions of growth (Figure 3A). These data indicated that RsaG is not required for G6P catabolism *in vitro*. Then, we performed internalization assays with HG001 and HG001 Δ hptRS strains in (non-phagocytotic) myoblasts and (phagocytotic) macrophages cell lines cultured in K-MEM or D-MEM medium, respectively. Quantification of RsaG levels in intracellular bacteria was monitored by qRT-PCR. The data clearly showed that RsaG is induced 100-fold under these conditions and that the activation is dependent of HptRS, which allows the expression of *uhpT*-RsaG co-transcript (Figures 3B and S3A).

Because *S. aureus* is primarily an extracellular pathogen, we postulated that bacteria might face G6P in contact to host cells producing mucus either in intestine or lungs. Colon cell lines HT-29 and HT-29 MTX cells were cultivated for 14 days to allow differentiation and production of mucus as verified by dot blot (right panel of Figure 3C). Subsequently, we incubated non-differentiated (day 4) or differentiated (day 14) HT-29 and HT-29 MTX cells (Behrens et al., 2001; von Kleist et al., 1975) with either *S. aureus* HG001 or HG001 Δ hptRS strains. Bacteria were gently recovered and total RNA was extracted. Northern blot experiments were performed with a probe specific for RsaG. The data showed that expression of RsaG and of a higher band corresponding most probably to the co-transcript *uhpT*-RsaG was highly induced when HG001 was mixed with HT-29 and HT29-MTX differentiated cells (day 14), in contrast to HG001 incubated with undifferentiated cells (day 4) (Figure 3C). This effect is strictly dependent on the two component system HptRS, the sensor of the extracellular concentration of G6P, since no significant signal was observed for RsaG in the mutant strain HG001 Δ hptRS. Furthermore, we observed the same strong induction of RsaG by comparing the incubation of HG001 strain with HG001 Δ hptRS strain in the presence of cultivated A549 airway epithelial cells for 4 days, which is associated with a strong mucin production measured with two different antibodies MUC-5AC and MUC-2 (Figure 3D). Conversely, only weak signal

for RsaG was observed when *S. aureus* strains were incubated with liver HU-H7 cells that do not synthesize mucin (Figure 3D).

In summary, we defined two different environmental conditions in which *uhpT* mRNA and consequently RsaG are induced. Because RsaG is not essential for growth in the presence of G6P, we proposed that RsaG is not required for G6P uptake and catabolism but might contribute to adapt the cells in new environment during the infection, i.e. presence of mucus and/or internalization into host cells.

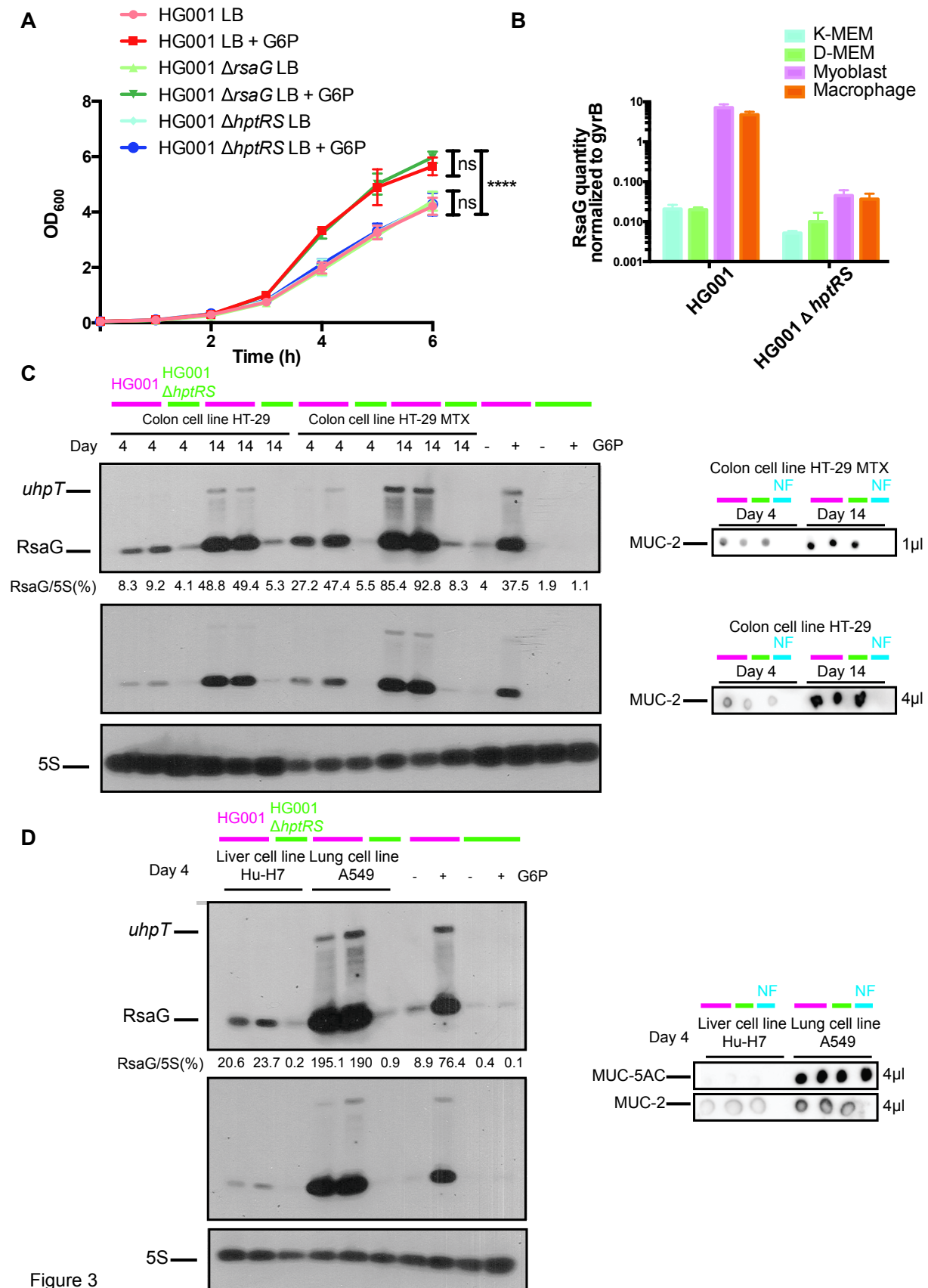


Figure 3

Figure 3: The *uhpT* mRNA and RsaG are enhanced under conditions mimicking the infection.

A. Growth curve of HG001 wild type strain, HG001 Δ *rsaG* and HG001 Δ *hptRS* strains in LB medium with or without the addition of 0.5% glucose-6-phosphate (G6P). ns: non significant, Statistical analysis with Anova, **** $p < 0,0005$. B. Levels of RsaG in HG001 and HG001 Δ *hptRS* strains, determined using qRT-PCR. Samples were taken after 1 h of *S. aureus* internalization into RAW 264.7 macrophages or CTi400 myoblasts. As control, RsaG relative amount was also quantified for HG001 wild type or HG001 Δ *hptRS* cultured for 1 h in K-MEM or D-MEM medium. Data were normalized to *gyrB* and represent the mean of three independent experiments $\pm$ SD. C. Left panel: Northern blot analysis of RsaG in HG001 (pink bars) and HG001 Δ *hptRS* (green bars). Total RNA was extracted after 1 h incubation with human colon epithelial cell lines HT-29 and HT-29MTX at days 4 or 14 of differentiation. As controls, RNA was extracted from HG001 and HG001 Δ *hptRS* incubated 1 h in MHB medium supplemented or not with 0.5% glucose-6-phosphate (-/+ G6P). 5S rRNA (5S) was used as loading control and to perform the quantification of RsaG, below the gel. However, for this control, we used aliquots of the same RNA preparations, but the migration of the samples was performed in parallel to the experiments on a separate agarose gel because RsaG and 5S rRNA have very similar sizes. Right panel: Mucin 2 protein (MUC-2) was quantified by dot blot assay. 1 μ l or 4 μ l of supernatant from the *S. aureus*-HT-29 or HT-29 MTX cells incubated with HG001 (pink bars) and HG001 Δ *hptRS* (green bars) were spotted on a nitrocellulose membrane and MUC-2 was specifically detected with anti-Muc-2 antibody. NF: non-infected HT-29 or HT-29 MTX cells. D. Left panel: Northern blot analysis of RsaG in HG001 (pink bars) and HG001 Δ *hptRS* (green bars). Total RNA was extracted after 1 h incubation with liver cell line Hu-H7 and lung cell line A549 at days 4 of culture. As control, RNA was extracted from HG001 and HG001 Δ *hptRS* incubated 1 h in MHB medium supplemented or not with 0.5% glucose-6-phosphate (-/+ G6P). 5S rRNA (5S) was used as loading control and to perform the quantification of RsaG, below the gel. However, for this control, we used aliquots of the same RNA preparations, but the migration of the samples was performed in parallel to the experiments on a separate agarose gel because RsaG and 5S rRNA have very similar sizes. Right panel: Mucin 2 (MUC-2) and Mucin 5AC (MUC-5AC) proteins were quantified by dot blot assay. 1 μ l of supernatant from the Hu-H7 or A549 cells incubated with HG001 (pink bars) and HG001 Δ *hptRS* (green bars) were spotted on a nitrocellulose membrane. MUC-2 and MUC-5AC were specifically detected with anti-MUC-2 or anti-MUC-5AC antibodies. NF: non-infected Hu-H7 or A549 cells.

Defining the RsaG targetome and its regulated pathways

In order to identify RsaG functions, the MAPS approach (“MS2 affinity purification coupled to RNA sequencing”) was applied to identify the RsaG targetome (Lalaouna et al., 2018). Briefly, the MS2 tagged version of RsaG was expressed from a plasmid under the control of the P3 promoter in the Δ rsaG mutant strain. To mimic the inducible conditions of RsaG, the growth was performed for 5 h in BHI medium and then G6P 0,5 % was added for 1 h before harvesting. As controls, we showed that the MS2-RsaG is specifically retained by the column (Figure S4A) and is as stable as the chromosomal copy (Figure S4B). After elution and extraction, the RNAs were sequenced, and the data were analyzed using Galaxy (Afgan et al., 2016). As previously described, the sequencing reads were mapped, counted per feature, and normalized using the HG001 genome (Tomasini et al., 2017). The enrichment of possible RsaG targets was derived by comparison of the number of reads obtained from the MS2-RsaG purification and the MS2 alone as control. In the latter case, the MS2 tag alone was expressed in the wild-type HG001 strain. In the following study, we have considered as RsaG targets, the RNAs that were enriched at least two-fold and were reproducibly and significantly detected in two independent experiments (Table S1). Among the most enriched RNAs, we identified several mRNAs encoding transcription regulatory proteins (SarA, SlyA, SarX, SarV, Rex, TcaR, CcpA, RpiR). Several metabolic operons were enriched with RsaG like the PTS (phosphoenolpyruvate (PEP)-dependent phosphotransferase system) encoded by the genes *ptsH* and *ptsI*, which allows carbohydrates transport and phosphorylation before entry into glycolysis, the *arcABDCR* which is involved in arginine catabolism when *S. aureus* grows in anaerobic conditions (Makhlin et al., 2007), and the Thi operon in thiamine biosynthesis, an indispensable cofactor of enzymes involved in amino acid and carbohydrate metabolism. Besides, we also have pulled out several polycistronic transcripts encoding the mechanosensitive ion channel MscL, the ABC transporters of K⁺ KdpBAF regulated by the TCS KdpED (Xue et al., 2011), and the ABC transporters of Mn²⁺ MntABC repressed by MntR (Ando et al., 2003). Various dehydrogenases (Ald2, Ndh, Hom) appeared in the list of enriched RNAs. Related to virulence, the Pmt operon encodes the transporter, which exports the Psm (phenol soluble modulins) toxins causing host cell lysis (Chatterjee et al., 2013), and the *srdM-sepA* transcript encodes multi-drug efflux pumps. Finally, although RsaI must be weakly expressed in presence of G6P, it is among the best candidates enriched with RsaG (Bronesky et al., 2019). Interestingly, we co-purified three of its principal targets, *glcU_2*, *icaR* and *HG001_02520*, which probably do not interact directly with RsaG (Bronesky et al., 2019).

We then analyzed whether RsaG is able to form stable complexes with several of these RNAs. In the top 20 candidates, stable pairings between RsaG and the RBS of several mRNAs were predicted, i.e., *mscL*, *sarA*, *slyA*, *arcR*, *tcaR*, *ndh*, suggesting that RsaG would inhibit their translation. Likewise, RsaG by prediction, interacts in the 5'UTR of several mRNAs, i.e., *rpiR*,

kdpA, *sarX*, which could also have a consequence on their translation. In gel retardation assays experiments, RsaG was labelled at its 5' end and complexes were formed in the presence of increasing amounts of 29 target RNAs using gel retardation assays. The data showed that RsaG forms stable pairings with 25 among the 29 tested RNAs with binding affinities ranging from 50 nM to 600 nM (Table S1). Four were not able to bind RsaG *in vitro*, and among them, *glcU_2*, *icaR* and the hypothetical protein HG001_02520 were shown to be targets of RsaI (Bronesky et al., 2019). Other poorly enriched RNAs were also tested. Among them, *ldh_1* encoding the lactate dehydrogenase (and its homolog *ldhD1*), that is repressed by Rex and activated by CcpA, hybridized with RsaG too (Table S1).

Furthermore, comparative transcriptomic experiments were conducted in duplicates with the strains HG001, HG001 Δ *rsaG* and HG001 Δ *rsaG* complemented with a plasmid expressing *rsaG* under the P3 promoter. Bacteria were cultivated in BHI for 5 h and then G6P 0,5 % was added for 1 h before harvesting. Total RNA was purified, DNase-treated and processed for library construction before sequencing. As controls, we demonstrated by Northern blot that RsaG was well expressed in the wild-type strain containing the empty plasmid, was absent in the deletion mutant, and was overexpressed in the deletion mutant strain complemented with the plasmid P3::*rsaG* (Figure S4C). Only the comparison between the complemented and the mutant Δ *rsaG* strains revealed significant changes in the mRNA yields. Particularly, the expression of four genes belonging to the *norB_4-steT-tdcB-ald1* operon is strongly enhanced by RsaG (Table S2). *NorB_4* is a quinolone resistance protein, *SteT*, a serine-threonine exchanger, *TdcB*, a L-threonine dehydratase and *Ald1*, an alanine dehydrogenase. Although this operon was not significantly enriched with RsaG using MAPS, RsaG is able to form stable complexes with *tdcB*, *steT*, *norB_4*, and *ald1* (not shown). The yield of *sodM* mRNA encoding the iron and manganese dependent superoxide dismutase SodM was also significantly increased by RsaG, but was not specifically enriched in MAPS, suggesting that the RsaG-dependent effect is indirect. Interestingly, this effect might be due to SarA, which was shown as a negative regulator of SodM synthesis (Ballal and Manna, 2009). Conversely, only two mRNAs encoding the adhesins *Ssp* and *SdrC*, known as virulence factors, were repressed by RsaG, but not enriched in MAPS, suggesting again that the regulation is most probably indirect (Tables S1, S2). None of the previously identified direct targets of RsaG by MAPS presented differential RNA levels in transcriptomic analysis, including RsaI. However, the latter is only weakly expressed in the presence of G6P (Table S2) (Bronesky et al., 2019).

The fact that little overlap existed between MAPS and transcriptomic data suggested that RsaG might predominantly affect the translation of the target mRNAs. These data further support the hypothesis that RsaG is not required for G6P catabolism but adapts the metabolism of the bacteria in response to G6P uptake.

Figure 4: RsaG binds to *tcaR*, *sarA* and *ccpA* mRNAs and inhibits their translation.

A. Secondary structure model of RsaG. The red C are the nucleotides substituted by G (in blue) in the RsaG mutants (mut1 or mut2). The potential base-pairings between RsaG and its targets is depicted. B. Electrophoretic mobility shift assays (EMSA) show the formation of the complex between RsaG and *tcaR* (left panel), *sarA* (middle panel) and *ccpA* (right panel) mRNAs. 5' end-labeled RsaG was incubated with increasing concentrations of cold mRNA (written on the top of the gel for each transcript, in nM). Below the gels, the predicted interaction sites between RsaG and mRNA targets are depicted. The translation start codon sequence is underlined, and Shine and Dalgarno (SD) sequences are in bold characters. C. Toe-print assays showing RsaG effect on the formation of the ribosomal initiation complex of *tcaR* (left panel), *sarA* (middle panel) and *ccpA* (right panel) mRNAs. Lane 1 : incubation control of mRNA alone; lane 2 : incubation control of mRNA with 30S subunits; lane 3: incubation control of mRNA with RsaG; lane 4 : formation of the ribosomal initiation complex containing mRNA, 30S and the initiator tRNA^{fMet} (tRNA_i); lanes 5 to 8 : formation of the initiation complex in the presence of increasing concentrations of RsaG, respectively : 50 nM (lane 5), 100 nM (lane 6), 200 nM (lane 7), 400 nM (lane 8) for *tcaR* and *ccpA*, lanes 5 to 8 : formation of the initiation complex in the presence of increasing concentrations of RsaG, respectively : 50 nM (lane 5), 150 nM (lane 6), 300 nM (lane 7), 400 nM (lane 8) for *sarA*. Lanes T, A, C, G: sequencing ladders. The Shine and Dalgarno (SD) sequence, the start site of translation (ATG) and the toe-printing signals (+16) are indicated.

RsaG fine-tuned its target mRNAs through multiple mechanisms

Using CopraRNA and IntaRNA, basepairing interactions were predicted between the RNA targets (Figure 4AB), which form stable complexes with RsaG. In several cases, the C-rich sequence in the interhelical region (C84-U89) of RsaG is found complementary to the SD sequences of the mRNAs encoding the biofilm repressor TcaR (Brandenberger et al., 2000; Jefferson et al., 2004) and the accessory regulator SarA (Cheung et al., 2008), while the G39-U44 in the hairpin loop of RsaG is complementary to the SD sequence of *ccpA* mRNA encoding the catabolite control protein A (Seidl et al., 2009) (Figure 4B). Because the stability of these mRNAs was not affected in the strain overexpressing RsaG, we postulated that the regulation might occur at the translation level. The consequence of the SD sequestration by the sRNA on translation initiation was addressed by toe-printing assays with *tcaR*, *sarA* and *ccpA* mRNAs using primer extension with reverse transcriptase (Figure 4C). For the three mRNAs, the addition of the initiator tRNA^{fMet} and of the *S. aureus* 30S subunit causes a pause of the enzyme at position +16, the so-called toe-print. The addition of increasing concentrations of RsaG added to the ternary ribosomal initiation complex strongly decreased the toe-print signal showing that RsaG is able to hinder the formation of the initiation ribosomal complexes.

The redox transcriptional repressor Rex was also highly enriched with RsaG in MAPS. *In vitro* binding assays showed that the two RNAs formed a stable complex (around 150 nM, Figure 5A). Interestingly, base-pairings were predicted between the 5' end of *rex* and the unpaired large interhelical region of RsaG (nts 98 to 105) while a second region located in the coding sequence of *rex* might potentially interact with the 5' hairpin loop (nts 42 to 51) of RsaG (Figure 4A, 5, S5A). Binding of *rex* mRNA also causes protection of these two regions of RsaG against lead(II)-induced cleavages (Figure S5A). Using Toe-printing assays, the addition of increasing concentrations of RsaG did not alter the formation of the initiation ribosomal complex with *rex* mRNA (Figure S5B). Therefore, we analysed whether RsaG can have an impact on *rex* mRNA stability. Using rifampicin assays, the half-life of *rex* mRNA was measured in the wild-type strain HG001 grown in absence or presence of G6P to stimulate RsaG or in the deleted strain of *rsaG* complemented with a plasmid over-expressing RsaG under the P3 promoter (HG001 Δ *rsaG* pCN51::P3*rsaG*). In absence of G6P, the half-life of *rex* mRNA was less than 1 min (33,4 sec), but when RsaG was overexpressed, the mRNA was significantly stabilized. Indeed, the half-life of *rex* reaches 77.9 sec in presence of G6P and 92.6 sec when RsaG is over-expressed from a plasmid (Figure 5B, upper panel). Interestingly, the deletion of the endoribonuclease III causes a significant effect on *rex* mRNA stability in MHB under conditions where the yield of RsaG is low (Figure 5B, bottom panel). The 5' end of *rex* mRNA (103 nucleotides) including 207 nucleotides of its coding sequence was then cloned in frame with *lacZ* under the control of the strong promoter *PrpoB* while *rsaG* was cloned under the constitutive *blaZ* promoter. The synthesis of β -galactosidase was analyzed in the Δ *rsaG* mutant strain transformed with the plasmid carrying the *lacZ* reporter alone or in the same strain transformed with the plasmid containing the *lacZ* and *rsaG*. The β -galactosidase activity was reproducibly increased by two-fold in cells expressing RsaG (Figure S5C) meaning that even if RsaG does not improve the binding of the ribosome on the SD, the stabilisation of *rex* mRNA increases the translation rate.

Interestingly the main target of Rex, the lactate deshydrogenase *Ldh_1* was enriched in MAPS albeit with a threshold below of 2 fold. Prediction of pairing between *ldh_1* and RsaG indicated that the large interhelical and single strand region (nucleotides 85 to 110) of RsaG might form basepairings either with the SD of *ldh_1* mRNA or in the coding region (Figure 5C). We confirmed the binding site at the SD sequence with a competitive gel shift experiment using a single strand DNA oligonucleotide comprising 34 nucleotides of *ldh_1* mRNA (from -14 to +20) (Figure S5D). Because RsaG formed a stable complex with *ldh_1* mRNA which partially covered the ribosome binding site, we analyzed whether RsaG would be able to compete with the formation of the initiation ribosomal complex using toe-printing experiment. The quantification of the data showed that the Toe-print did not decrease at all concentrations of RsaG in the assays (Figure 5CD and S5E). In addition, the stability of *ldh_1* mRNA was

measured using rifampicin treatment after 4 h of growth in absence or presence of G6P from HG001 strain and from the Δ *rsaG* deleted strain complemented with a vector expressing RsaG under the P3 promoter. RNA was extracted and *ldh_1* mRNA detected by Northern blot. The half-life of *ldh_1* was significantly decreased from around 10 min to 2-3 min when RsaG expression was strongly induced (Figure 5E). In response to G6P, RsaG would act at two levels either by activating *rex* or repressing *ldh_1* in order to avoid glycolytic fermentation in favour of amino acid catabolism, when glucose is absent.

Taken together, our data suggested that RsaG acts as a translational repressor of many target mRNAs and also as a post-transcriptional regulator, which modulates positively or negatively the stability of target mRNAs.

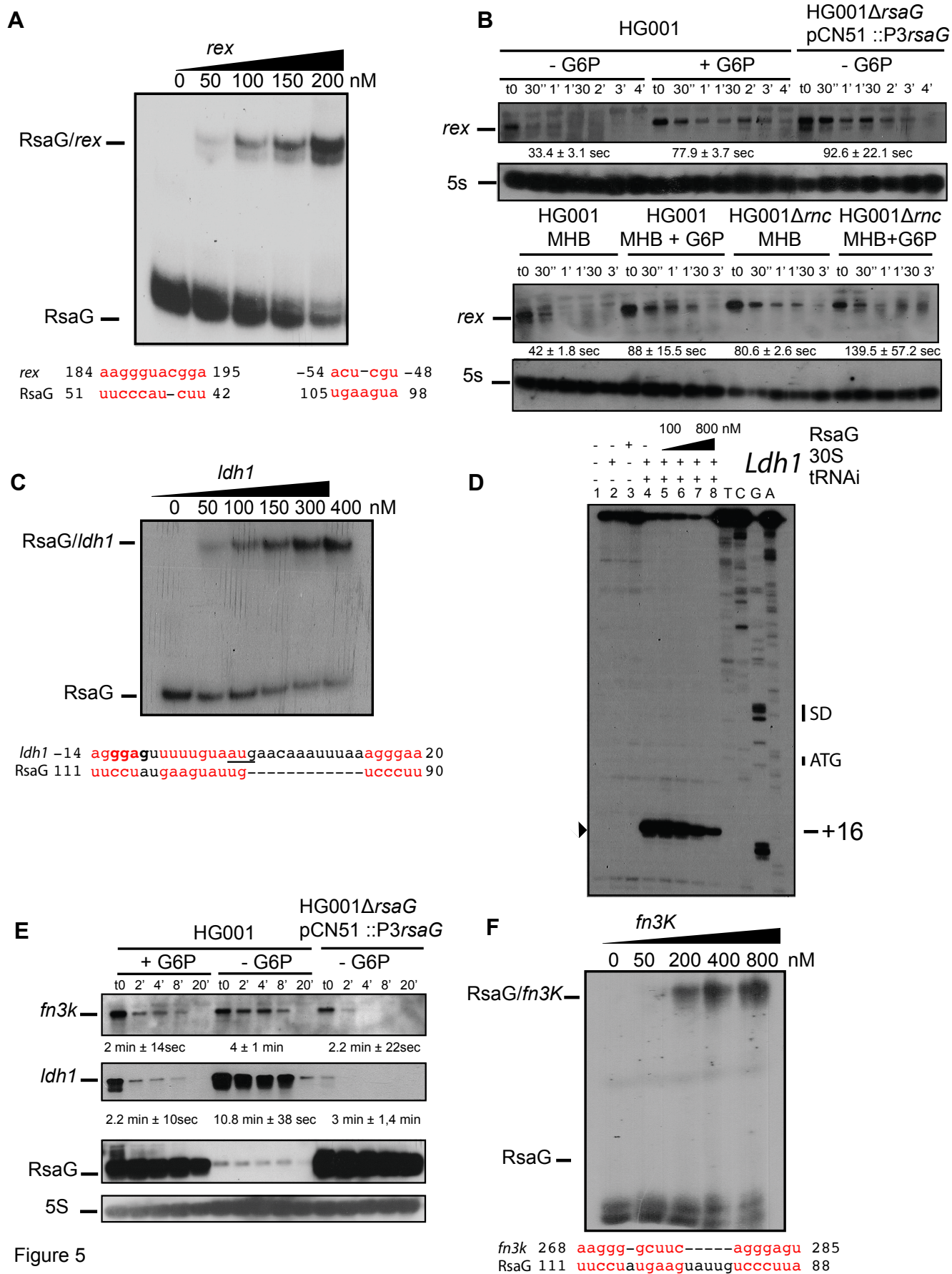


Figure 5

Figure 5: RsaG modulates stability of *rex*, *fn3K* and *ldh_1* mRNAs.

A. Electrophoretic mobility shift assays (EMSA) showing the formation of the complex between RsaG and *rex* mRNA. 5' end-labeled RsaG was incubated with increasing concentration of cold mRNA (written on the top, in nM). Below the gel, the predicted interaction sites are depicted. B. Measurements of the half-lives of *rex* in HG001 and HG001 Δ *rsaG*::pCN51::P3*rsaG* (top) or HG001 and HG001 Δ *rnc* (bottom) strains using Northern blot experiments. Bacterial cultures were grown in MHB medium containing or not glucose-6-phosphate (+/- G6P) treated with rifampicin at 4h of growth. Total RNA was extracted after 2, 4, 8, 15, 30 and 60 min at 37°C. 5S rRNA was probed to quantify the yield of RNAs in each lane. Calculated half-lives are shown beneath the autoradiography and are the average of 2 experiments. The data were normalized to 5S rRNA. C. EMSA showing the formation of the complex between RsaG and *ldh_1* mRNA. 5' end-labeled RsaG was incubated with increasing concentration of cold mRNA (written on the top, in nM). Below the gels, the predicted interaction site is depicted. The translation start codon sequence is underlined, and Shine and Dalgarno (SD) sequences are in bold characters. D. Toe-print assays showing the RsaG effect on the formation of the ribosomal initiation complex of *ldh_1*. Lane 1 : incubation control of mRNA alone; lane 2 : incubation control of mRNA with 30S subunits; lane 3: incubation control of mRNA with RsaG; lane 4 : formation of the ribosomal initiation complex containing mRNA, 30S and the initiator tRNA^{fMet} (tRNA_i); lanes 5 to 8 : formation of the initiation complex in the presence of increasing concentrations of RsaG, respectively : 100 nM (lane 5), 300 nM (lane 6), 600 nM (lane 7), 800 nM (lane 8). Lanes T, A, C, G: sequencing ladders. The Shine and Dalgarno (SD) sequence, the start site of translation (ATG) and the toe-printing signals (+16) are indicated. E. Measurements of the half-lives of *fn3K* and *ldh_1* in HG001 and HG001 Δ *rsaG*::pCN51::P3*rsaG* strains using Northern blot experiments. Bacterial cultures were grown in MHB medium containing or not glucose-6-phosphate (+/- G6P) treated with rifampicin at 4 h of growth. Total RNA was extracted after 2, 4, 8, 15, 30 and 60 min at 37°C. 5S rRNA was probed to quantify the yield of RNAs in each lane. Calculated half-lives are shown beneath the autoradiography and are the average of two experiments. F. EMSA showing the formation of the complex between RsaG and *fn3K* mRNA. 5' end-labeled RsaG was incubated with increasing concentration of cold mRNA (written on the top, in nM). Below the gels, the predicted interaction site is depicted.

RsaI in RsaG-mediated regulation

We previously identified that the second G rich motif of RsaI binds to RsaG. This interaction did not prevent RsaI to bind to its main target mRNA *glcU_2* (Bronesky et al., 2019). By substituting the two C-rich motifs of RsaG with guanines, we could decipher that only the first C-rich motif of RsaG forms stable basepairings with RsaI *in vitro* (Figure 6AB). The

consequence of RsaG-RsaI pairing on RsaG target recognition was then monitored by gel retardation assays. Radiolabelled RsaG was mixed together with a concentration of *sarA* mRNA sufficient to bind most of RsaG. The addition of increasing amount of RsaI induces the appearance of a supershift indicating the formation of a ternary complex RsaI-RsaG-*sarA* (Figure 6C) and supporting our data that *sarA* and RsaI bind to distinct regions of RsaG. Conversely, when the same experiment was performed between RsaG, RsaI, *rex* mRNA, or *steT* mRNA, a competition was observed between RsaI and the two mRNAs (Figure 6DE). This is in agreement with the observation that RsaI and the mRNAs share similar binding sites on RsaG.

In the list of targets enriched by MAPS (below the threshold of 2 fold, table 1), we noticed *fn3K*, whose translation was reduced by RsaI, because the pairings included the ribosome binding site (Bronesky et al., 2019). Pairing sequences between RsaG and *fn3K* predicted binding within the coding region (Figure 5F). Gel retardation assays showed that the radiolabelled RsaG did not hybridize to an *in vitro* transcribed RNA corresponding to the 5' UTR of *fn3K* (not shown) but was able to form a stable complex with an *in vitro* transcribed RNA including the open reading frame of *fn3K* (Figure 5F). Moreover, an oligonucleotide of 17 nucleotides from 268 to 285 was able to compete with *fn3K* bound to RsaG *in vitro*, suggesting that RsaG binds specifically to this region (Figure S5F). We postulated that RsaG would act at the post-transcriptional level to regulate *fn3K*. The stability of the mRNA was measured using rifampicin treatment after 4 h of growth. Total RNAs were extracted from HG001 strain in absence or presence of G6P, and from the Δ *rsaG* deleted strain complemented with a vector expressing RsaG under the P3 promoter. The data shows that the expression of RsaG significantly destabilizes *fn3K* as the half-life decreases from 4 min to 2 min (Figure 5E). Another common target between RsaG and RsaI is *tdcB* mRNA, whose predicted binding sites are different. Gel retardation assay with radiolabelled RsaG mixed together with a concentration of *tdcB* mRNA sufficient to bind most of RsaG and with an increasing amount of RsaI showed a supershift indicating the formation of a ternary complex RsaI-RsaG-*tdcB* (Figure 6F).

In summary, RsaI does not act as a classical sponge sRNA as it affects the binding of only part of the mRNA targets.

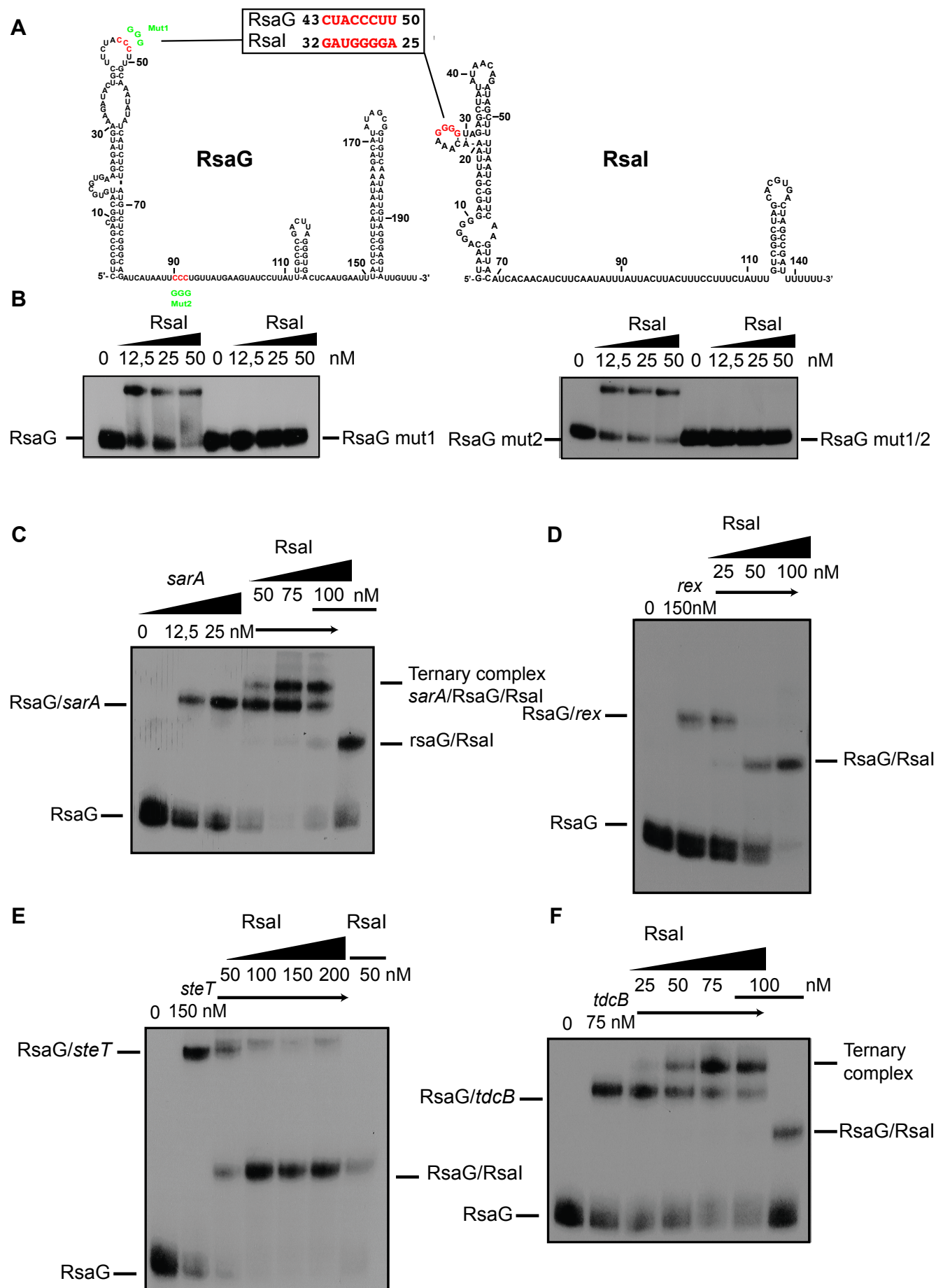


Figure 6

Figure 6: RsaI does not act as a classical sponge sRNA as it affects only part of the regulatory functions of RsaG.

A. Secondary structure models of RsaG and RsaI. The red C are the nucleotides substituted by G (in green) in the RsaG mutants (mut1 or mut2). The potential base-pairings between RsaI and RsaG is depicted. B. Gel retardation assays show the formation of the complex between RsaG and RsaI. The 5' end-labeled wild-type RsaG (RsaG), RsaG mutant 1 (RsaG mut1), RsaG mutant 2 (RsaI mut2), and the RsaG double mutant 1 and 2 (RsaG mut1/2) were incubated with increasing concentrations of RsaI given in nM on the top of the autoradiographies. C. Ternary complex formation between RsaG, *sarA* and RsaI. The 5'-end labeled RsaG was incubated with increasing concentrations of the target mRNA alone or with increasing concentrations of RsaI in the presence of 25 nM of *sarA*. The various complexes are notified on the side of the autoradiography. D. Complexes formation between RsaG, *rex* and RsaI. The 5'-end labeled RsaG was incubated with increasing concentrations of the target mRNA alone or with increasing concentrations of RsaI in the presence of 150 nM of *rex*. The various complexes are notified on the side of the autoradiography. E. Complexes formation between RsaG, *steT* and RsaI. The 5'-end labeled RsaG was incubated with increasing concentrations of the target mRNA alone or with increasing concentrations of RsaI in the presence of 150 nM of *steT*. The various complexes are notified on the side of the autoradiography. D. Ternary complex formation between RsaG, *tdcB* and RsaI. The 5'-end labeled RsaG was incubated with increasing concentrations of the target mRNA alone or with increasing concentrations of RsaI in the presence of 75 nM of *tdcB*. The various complexes are notified on the side of the autoradiography.

Conservation of RsaG in *Staphylococcaceae*

The *uhpT* gene sequence is highly conserved in the *Staphylococcaceae* family (between 83 to 91% similarity) as the locus organization including the presence of the two-component system on the complementary strand (Figure 7A). We then performed sequence alignment between the 3'UTRs using Blast. Surprisingly, the 3'UTR of *uhpT* appears to be conserved in *S. argenteus* and *S. schweitzeri* that are closely related to *S. aureus*, while the sequence and the size diverged a lot in *S. epidermidis* and *S. simiae* (Figure 7A). Using specific probe against *S. aureus* RsaG sequence, we identified a transcript in *S. argenteus* and *S. schweitzeri* from total RNA extracted when G6P was added to the bacterial cultures (Figure 7B). In contrast, we did not detect a processed 3'UTR of *uhpT* mRNA of *S. epidermidis* (using a specific probe), which has a shorter size, even if *uhpT* mRNA is expressed (not shown). However, a signal corresponding to a larger processed 3'UTR was visualized in *S. simiae* using a specific probe against *S. simiae* 3'UTR (Figures 7B and S6). The +1 of transcription was then determined using reverse transcription assays with total RNA extracted from *S. simiae* cultures (Figure

7D). The start site was localized 195 bp downstream of the stop codon of *uhpT* resulting to an apparent 3'UTR derived sRNA of 726 nt in *S. simiae* (294 nt in *S. aureus*). As the HptR box is rather conserved, the induction of *S. simiae* RsaG in absence or presence of G6P was analyzed (Figure 7C). Unexpectedly, the detection of the 3'UTR was independent of the presence of G6P. Although, the 3'UTR sequence of *S. simiae uhpT* has diverged, the maturation process may have been conserved.

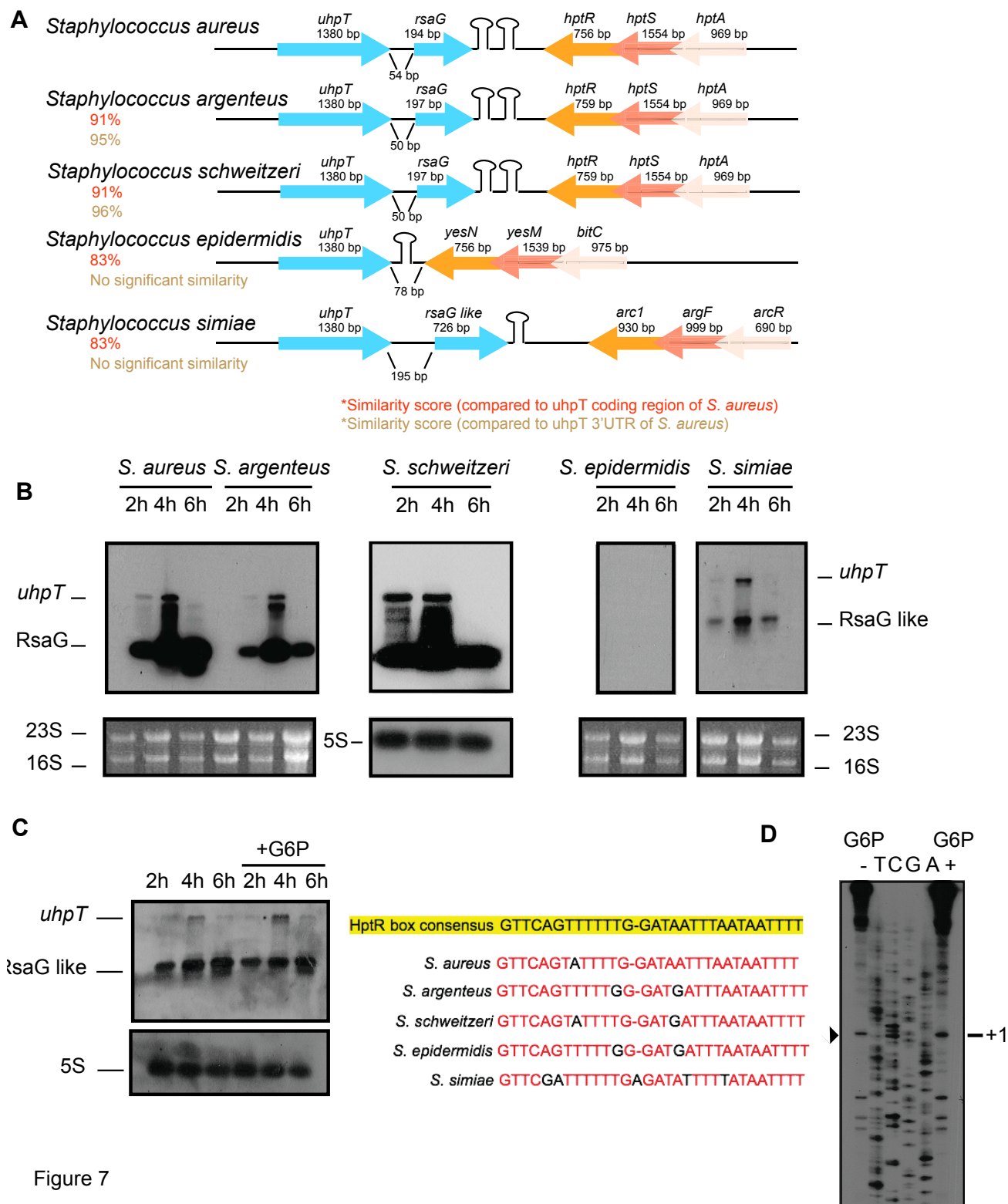


Figure 7

Figure 7: Maturation of a 3'UTR derived sRNA from *uhpT* mRNA is not conserved in all *staphylococcaceae*.

A. Genomic context of *uhpT* locus in *Staphylococcus aureus* HG001, *S. argenteus* MSHR11, *S. schweitzeri* FSA084, *S. epidermidis* ATCC35983 and *S. simiae* CCM7213T. Percentages in orange represent the similarity score of *uhpT* coding region and in brown the similarity score of *uhpT* 3'UTR compared to *S. aureus*. The similarity score was obtained by blast (NCBI). B. Northern blot analysis of RsaG in *S. aureus*, *S. argenteus*, *S. schweitzeri*, *S. epidermidis* and *S. simiae*. Total RNA was extracted at 2, 4 and 6h of growth in BHI medium containing 0.5% glucose-6-phosphate. RsaG from *S. aureus*, *S. argenteus* and *S. schweitzeri* was detected with a specific probe against *S. aureus* RsaG sequence, while for *S. epidermidis* and *S. simiae*, the probe was designed against their respective sequence. 16S, 23S or 5S rRNA were used as loading controls. C. Northern blot analysis of RsaG in *S. simiae*. Total RNA was extracted at 2, 4 and 6 h of growth in BHI medium without or with 0.5% glucose-6-phosphate (+G6P) (left). Multiple sequence alignment of the HptR box in *S. aureus*, *S. argenteus*, *S. schweitzeri*, *S. epidermidis* and *S. simiae* compared to the consensus sequence underlined in yellow (Yang et al., 2016) using blast (NCBI) with default settings (right). D. Primer extension performed on total RNA using a 5'-radiolabelled oligonucleotide complementary to *S. simiae uhpT* 3'UTR sequence. Total RNA was extracted from *S. simiae* at 4 h of growth in BHI with or without 0.5% glucose-6-phosphate (-/+ G6P). The primer is located at nucleotides +108 to +125 (from the +1 of the putative RsaG). T, C, G, A correspond to sequencing ladders.

DISCUSSION

In this study, we have identified RsaG as a new 3' UTR-derived sRNA processed from *uhpT* mRNA encoding the transporter of G6P. The synthesis of UhpT is activated by the TCS HptRS, which senses the external concentration of G6P. Although G6P is critical for *S. aureus* when it is internalized into host cells as a mutant deprived of *hptRS* is not able to survive (Park et al., 2015), RsaG is nevertheless not essential for growth when G6P is used as the solely carbon source. However, *uhpT* mRNA and RsaG are highly induced by G6P when *S. aureus* is internalised into host cytosol and in presence of pulmonary or intestinal mucus (Figure 3; Garzoni et al., 2007). Here we showed that RsaG expands the action of HptRS/UhpT by fine-tuning the cell metabolism in response to G6P entry. Indeed, using different mechanisms of regulation, and C-rich motifs targeting the SD sequences of messengers, RsaG inactivates the lactate fermentation by activating *rex* and repressing *ccpA* expression (Figure 8).

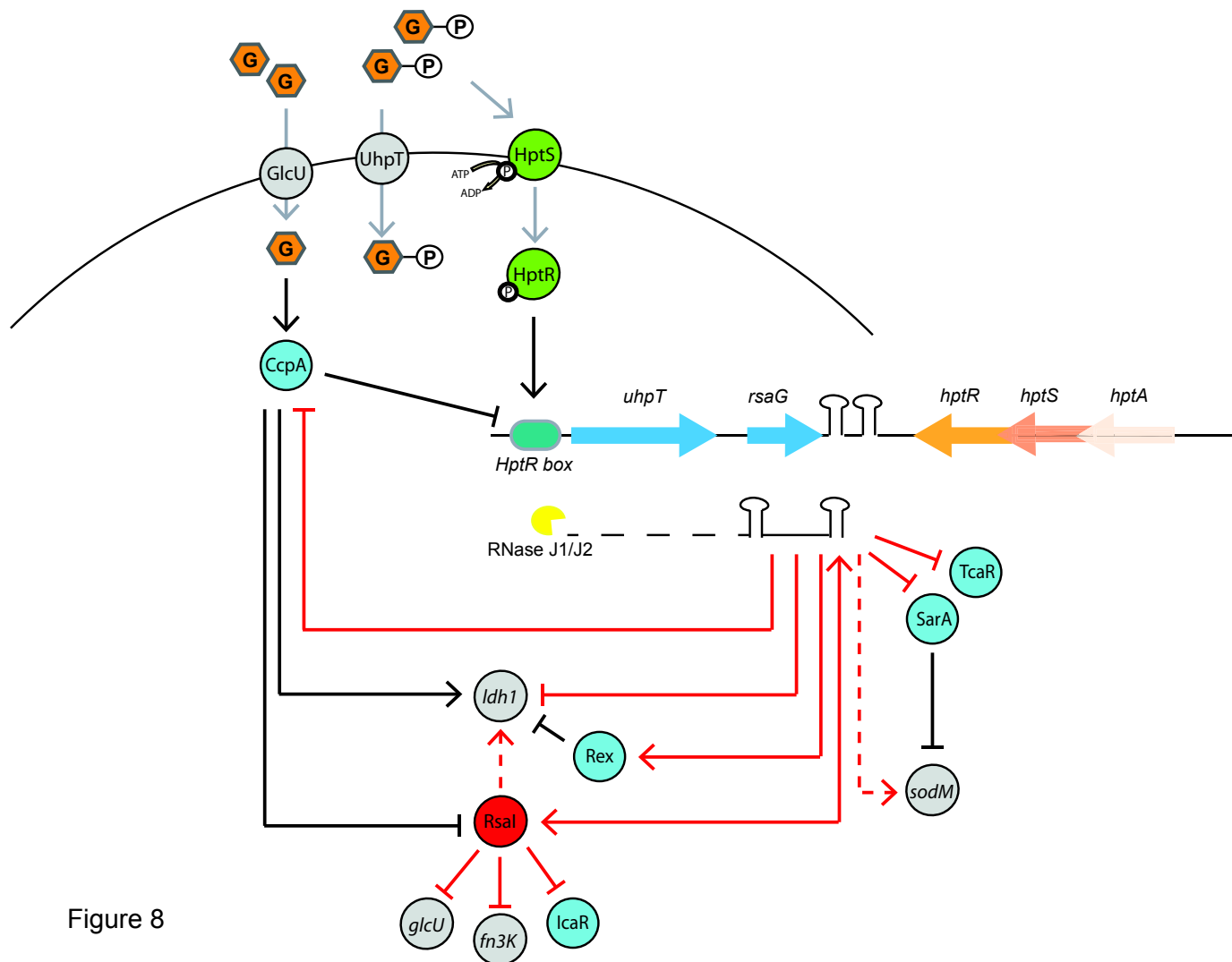


Figure 8

Figure 8: Model of RsaG regulation networks in *Staphylococcus aureus*.

The two-component systems are shown in green, the transcriptional regulatory proteins are in blue. Red bars and arrows are for repression and activation at the post-transcriptional level, respectively. Black bars and arrows are for repression and activation at the transcriptional level.

The dual ribonuclease J1/J2 generates the 3'UTR-derived RsaG sRNA

Recent work has highlighted the diversity in length and sequence of 3'UTRs in *S. aureus* and proposed that these regions contribute to regulate gene expression, affect mRNA turnover or modulate the expression of downstream mRNAs (Menendez-Gil et al., 2020). This hypothesis is also in agreement with the role of 3'UTRs in eukaryotes, which contributed to the divergence of species by evolving regulatory elements. Noteworthy, large 3'UTRs containing multiple miRNA regulatory elements have been particularly described in genes that are spatially or temporally expressed (Cheng et al., 2009). In Gram-negative bacteria, 3'UTRs are also reservoirs of sRNA acting in trans after RNase E cleavage (Eisenhardt et al., 2018; Kim et al.,

2014; Wang et al., 2020). In *S. aureus*, RsaC accumulates in response to Mn starvation and is processed by RNase III from the polycistronic mRNA encoding the MntABC transporter. Here, we describe the first example of an sRNA for which RNase J1/J2 is the primary enzyme required for its processing from the *uhpT* mRNA, while the mRNA encoding *uhpT* is rapidly degraded. Hence RNase J1/J2 is not only essential for stress responses by controlling mRNA degradation but also to generate sRNA.

Intriguingly, RsaG is the third example of a 3'UTR-derived sRNA issued from a mRNA encoding metabolically active transporters, as RsaC processed from the staphylococcal MntABC transporter of manganese and NarS from the NarK transporter of nitrate in *E. coli* (Lalaouna et al., 2019; Wang et al., 2020). In the case of RsaC, it is induced in Mn starvation to repress SodA activity which needs Mn as cofactor. In anaerobiosis, NarK transports nitrate into the cytoplasm, which is reduced in nitrite by NarG. To limit nitrite toxicity, NarS is transcribed together with *narK* and reduced the expression of *nirC* encoding the nitrite transporter. Overall, sRNAs generated from mRNAs contribute to the same metabolic pathway or help bacteria to choose the most appropriate way to adapt to different metabolic niches in response to the effector entry. It is tempting to propose that such dual partners (transporter and sRNA) are more widespread than expected. While RsaG is cotranscribed with the UhpT G6P transporter, we were unable to show its direct role in G6P catabolism (see below). In other pathogenic bacteria such as *Shigella flexneri* and *Listeria monocytogenes*, their corresponding *uhpT* mRNAs do not possess large 3'UTR and were highly overexpressed during internalization (Chico-Calero et al., 2002; Runyen-Janecky and Payne, 2002). In *Listeria* as a virulence factor, it is regulated by the major regulator PrfA, and the deletion mutant of *uhpT* did not proliferate intracellularly or survive in a murine model of infection (Chico-Calero et al., 2002). Although the role of UhpT in staphylococcal virulence still requires investigation, we hypothesized that the accumulation of RsaG would regulate additional metabolic functions when *S. aureus* is internalized into the microaerophilic cytoplasm of host cells or in presence of mucus-secreting cells.

RsaG expands the regulon of HptRS

In order to scrutinize the functional impact of RsaG, we performed differential transcriptomic analysis and MAPS in presence of G6P to mimic inducing conditions. Among the target RNAs, none of them are directly linked to the G6P catabolism *per se* but interestingly, some of them are involved in carbohydrate-dependent metabolic pathway (*ccpA*, *ldh_1*), in amino acid metabolism (*stet-tdcB-ald1* operon) while others are linked to redox homeostasis (*rex*). Interestingly, in the presence of glucose, CcpA repressed *uhpT* and the effect was more pronounced in the late exponential phase of growth (Reed et al., 2018). In contrast HptRS activated the same gene in presence of G6P (Park et al., 2015). Comprehensively CcpA and

HptRS share the same regulon functioning in an opposite manner to respond to changes in the availability of glucose or G6P (Reed et al., 2018). It means that *uhpT* activation by HprRS need the alleviation of CcpA inhibition in presence of G6P. Hence in these conditions, RsaG adds another level of regulation and may repress *ccpA* mRNA translation. Our MAPS analysis revealed direct or indirect interactions with other CcpA-regulated genes, such as *ldh_1*, *ald1*, *adh*, *treB*, *buA1B* (*alsS/D*), *hemC-hemX* suggesting that expression of these genes would be under RsaG control (Table S1 and Seidl et al., 2009). While CcpA inhibited transcription of *ald1*, *adh*, *budA1B* but activated *Ldh_1* activity, the global functions of RsaG would lead to the inhibition of lactate fermentation while it activates amino acid catabolism, acetoin synthesis and alcohol fermentation.

RsaG targeted Rex, which senses the bacterial redox through changes in the NADH-NAD⁺ ratio of the bacteria which can be changed dramatically under different metabolic status with or without a change in oxygen viability. For instance, NADH increases during glycolysis, during the activation of the TCA cycle, and under anaerobia conditions. It is known that Rex-NADH dissociates from its promoter allowing transcription of numerous genes involved in electron transport, in nitrogen and anaerobic metabolism (*nirC*, *narG*, *arcA*, *pflB*, *adhE*, *adh*, *ldh_1*, *ald1*) (Pagels et al., 2010). Interestingly among the potential direct targets, RsaG can also interact with *nirC*, *arcA*, *adh* (Table S1) and directly degrade *ldh_1* mRNA (Figure 4A). By activation of *rex* and direct binding to *ldh_1*, *nirC* or *adh*, RsaG functions at a different regulatory level to monitor the redox status.

Moreover, RsaG interfered with several members of the Sar family. Indeed, we showed clearly that the initiation complex of translation with *sarA* is impaired in presence of RsaG. Phenotypically, it is indirectly linked to the upregulation of *sodM*, examined when RsaG is active (Table S2). *SodM* is repressed by *SarA* (or *CcpA*) and participates to the oxidative stress resistance in presence of methyl viologen or diamide (Ballal and Manna, 2009; Seidl et al., 2009). Additionally, when staphylococcal cultures were starved in Mn, the atypical sRNA *RsaC* blocked translation of *sodA*, a Mn-dependent superoxide dismutase and *SodM* compensated (Lalaouna et al., 2019). In the same study, the authors suggested that *RsaC*, by inhibiting *sarA* translation, alleviated its repression on *sodM*. It is then tempting to conclude that high G6P concentrations sensed by the TCS HptRS induced RsaG, which links oxidative stress to virulence. Transcriptomic analyses revealed an increase in *sdrC* and *ssp* mRNAs levels in the *rsaG* deletion mutant, meaning that RsaG indirectly had a negative post-transcriptional impact (Table S2). The *ssp* gene, encoding the extracellular matrix protein-binding protein, also known as Emp, is an adhesin which binds to fibronectin, fibrinogen and vitronectin and is activated by *SarA* (Harraghy et al., 2005). Then RsaG would act as an inhibitor of bacterial adhesion. Besides, RsaG interacted with two other members of the Sar family, *SarX*, involved in biofilm formation by downregulated *IcaR* and *SarV*, an inducer of autolysis. Of note, *in silico*

predictions of bindings site between RsaG and *sarX* is localised upstream of the SD, or onto the SD of *SarV*, suggesting that RsaG would inhibit their translation (Table S1). Hence, RsaG impact on *SarX* and *SarV* function still need further investigations like its role in virulence.

RsaG and its two functional domains can be modulated by other sRNA

Most of the targets that we validated *in vitro* hybridised specifically at two different functional domains of RsaG, the first hairpin loop or the interhelical single strand region, both containing a C-rich motif (Geissmann et al., 2009). Using substituted mutant in the first C-rich motif of RsaG, we observed that *tdcB* (probably together with *norB_4* and *steT* belonging to the same transcript), *rex* and the sRNA *RsaI* hybridized at the first motif (Figure 4A). The *rex* mRNA is stabilised by RsaG probably by protecting the RNA from RNase III cleavage (Figure 5). A noteworthy observation is that *RsaI* competed with *rex* mRNA by binding to the first C-rich motif of RsaG. Recently, we have shown that *RsaI* indirectly favoured glucose fermentation in lactate, probably indirectly by modulating *ldh_1* mRNA (Bronesky et al., 2019), while RsaG inhibited the same pathway by degradation of *ldh_1* mRNA and stimulation of *rex*. Because, both sRNAs have an opposite effect on *ldh_1*, in certain conditions, *RsaI* would impair *rex* activation by RsaG. The mechanism of regulation for the *norB_4-steT-tdcB-ald1* operon is not known. Pairing was predicted in the coding region for *norB_4*, *steT* and *tdcB* questioning on the mechanism, which led to increase levels of RNAs observed by differential transcriptomic analysis (Table S2). Moreover, *tdcB* was defined as a target of *RsaI* by MAPS (Bronesky et al., 2019) and is predicted to anneal in the coding region but at a different location than RsaG (Figure 6F). *TdcB* initiates the anaerobic degradation of L-threonine to pyruvate. Hence, we proposed that under the conditions of internalization, RsaG would be induced due to the presence of G6P to activate the amino acid catabolism. Under this context where oxygen is limited, the expression of *RsaI* was recently shown to be strongly decreased (Sorensen et al., 2020). Noteworthy, *RsaI* also binds to other sRNAs such as *RsaD* and *RsaE*, and all of these sRNAs regulate several metabolic pathways (Bronesky et al., 2019; Augagneur et al., 2020). Interestingly, *RsaG* and *RsaD* share similar target mRNAs such as *alsS* but are expressed under various stress conditions. Interestingly many genes encoding enzymes and metabolic proteins are often regulated through RNAs at the post-transcriptional level providing a fast response to many different stimuli and environmental changes.

We were able to validate *in vitro* the binding site of *fn3K*, *ldh_1*, *treB* and *nirC* on the second C-rich motif of RsaG leading to mRNA degradation for *fn3K* and *ldh_1*. In contrast, while the predictions of pairings with *sarA* or *tcaR* were localised on the same motif of RsaG and that the Toe-print signal was clearly impaired in presence of RsaG (Figure 4), the substituted mutant of RsaG still fixed the mRNAs in gel retardation assays indicating that a secondary binding site might exist. The fructosamine 3-kinase (*fn3k*), which in bacteria plays

a role in protein deglycation by phosphorylating preferentially proteins bound-ribulosamines and erythrulosamines but not fructosamines, is a common target between RsaG and RsaI. It has been suggested that bacterial fructosamine kinases inactivate exogenous toxic compound like the antibiotics aminoglycosides or macrolides, but the internal substrate remains unknown (Gemayel et al., 2007). The interaction of RsaG or RsaI to *fn3K* mRNA leads respectively to degradation when G6P is present or inhibition of translation in the absence of glucose.

Not only RsaG regulated common targets with RsaI, but also targets, which are involved in common pathways. For instance, RsaI by binding to the 3'UTR of *icaR* mRNA induced the production of biofilm polysaccharides from the *icaADBC* operon, while RsaG would diminish TcaR synthesis, another transcriptional repressor of the same operon. Interestingly a deletion of the *tcaR* locus did not impair biofilm formation in contrast to the double deletion $\Delta icaR/\Delta tcaR$ (Jefferson et al, 2004). We suggest that when RsaI and RsaG are both present, they could cooperate to activate biofilm formation. All in all, RsaI partially prevented RsaG regulation.

Finally, RsaG is conserved only in closely related species as *S. argenteus* and *S. schweitzeri*, while *uhpT* is universal. Interestingly, in the ape pathogen, *S. simiae*, we detected a long (726 nucleotides) and stable RNA probably processed from the 3'UTR of *uhpT* mRNA, but whose sequence is not homolog to RsaG and whose expression is not induced in presence of G6P. The study of RsaG from *S. simiae* (RsaG-like) would need deeper investigations, especially the determination of its secondary structure and its targetome. However, we detected several putative C-rich motifs, but none of them were predicted to interact with the RsaG major targets, *rex* and *ccpA*. Nevertheless, sites were expected to bind RsaG-like upstream of the SD on *ccpA*. This suggests that the emergence of RsaG-mediated mRNA crosstalk evolved with pathogenesis to facilitate a better adaptation to the host niches.

In conclusion, acquisition of glucose transporters such as GlcC transporters was described as an adaptative advantage for *S. aureus* to survive within host tissue when O₂ is rare (Vitko et al., 2016). By analogy, we could postulate that *uhpT* together with RsaG have evolved because it is necessary to uptake G6P, and concomitantly to adapt metabolism and redox homeostasis when *S. aureus* is internalized or in presence of mucus-producing cells.

MATERIALS AND METHODS

Strains and plasmids

All strains and plasmids, which were constructed and used in this study, are described in Table S3. The oligonucleotides sequences are given in Table S4. *Escherichia coli* strain DH5 α was used for cloning purposes. *E. coli* strain DC10B, a DNA cytosine methyltransferase negative

mutant and *Staphylococcus aureus* RN4220 were used for plasmid amplification before *S. aureus* HG001 transformation. Transformation of *E. coli* was performed by heat shock and *S. aureus* by electroporation (Bio-Rad Gene Pulser). The plasmids were extracted from *E. coli* or *S. aureus* with the Nucleospin Plasmid kit (Macherey-Nagel) adding a mechanical breakage of staphylococcal cells with FastPrep in P1 buffer.

Deletions of *rsaG* and *uhpT* promoter comprised respectively removal of nucleotides 201739 to 201936 and nucleotides 200154 to 200282 (Caldelari et al., 2017) and were constructed by homologous recombination using the thermosensitive vector pMAD (Araud et al., 2004). Chromosomal regions upstream or downstream of *rsaG* or *uhpT* promoter were amplified by PCR (see table S4 for primers) and cloned into pMAD. The resulting plasmid was electroporated first into RN4220 recipient strain and then into HG001. Transformants were grown at non-permissive temperature (44°C), followed by several subcultures at 28°C and 37°C to favour double crossing over as previously described (Araud et al., 2004). To generate the plasmid expressing *rsaG* for complementation, the entire gene (194 bp) was PCR amplified with the following oligonucleotides: RsaG *BamHI* rev and RsaG *PstI* for (Table S4), digested by *PstI* and *BamHI* and ligated into pCN51::P3 digested with the same enzymes (Tomasini et al., 2017) to obtain pCN51::P3::rsaG. The MS2 tag was fused to the 5' end of *rsaG* using oligonucleotides *PstI*-MS2-RsaG for and RsaG-*BamHI* rev cloned into *PstI*/*BamHI*-digested pCN51::P3. The pUC-T7::rsaG mut1, mut2 and mut1/2 vectors were obtained by QuickChange XL Site-directed mutagenesis (Stratagene) from pUC-T7::rsaG.

For transcriptional fusions, different length fragments of the *uhpT* upstream region (i.e. 1585, 1456, 500, 250, 100 bp before RsaG transcription start site) (see Table S4 for primers sequences et description) were cloned into *SphI*-*BamHI*-digested pEW-GFP plasmid containing the 5'UTR_{hly} from *Listeria monocytogenes* (Menendez-Gil et al., 2020). Translational fusion to *lacZ* was constructed in pLUG220::rpoB::lacZ (Romilly et al., 2014). First, a 309 bp fragment containing the 5'UTR region of *rex* mRNA and 69 codons (-103/+207 bp) was amplified with primers *rex* for *BamHI* and *rex* rev *BamHI* and cloned into *BamHI*-digested pLUG220::rpoB::lacZ to generate pLUG220::rpoB::rex::lacZ. Then an amplicon containing *rsaG* under the *blaZ* promoter was produced by two distinct PCR. The pCN40 plasmid was used as template to amplify *blaZ* promoter with primers *PstI*-pBlaZ-pES for/RsaG'-pBlaZ-pES3' rev and HG001 genomic DNA to amplify *rsaG* with primers RsaG for/*PstI*-RsaG rev. The two PCR products served as template for a PCR with oligonucleotides *PstI*-pBlaZ-pES for/ *PstI*-RsaG rev and ligated into the *PstI*-digested pLUG220::rpoB::rex::lacZ.

Growth conditions

E. coli strains were grown in Lysogeny-Broth (LB, Roth) medium supplemented with ampicillin (100 µg/mL) or kanamycin (100 µg/mL) when necessary. *S. aureus* strains were cultivated in Brain-Heart infusion (BHI, Sigma) or Muller Hinton Broth (MHB, Sigma) media containing 10 µg/mL of erythromycin when needed. To induce RsaG, MHB or BHI were complemented with 0.5% of glucose-6-phosphate (G6P) (Sigma).

Northern blot

Purification of total RNAs extracts was performed strictly following the procedure described for the FastRNA pro blue kit (MP Biomedicals) with the Fastprep apparatus (MP Biomedicals). Electrophoresis of 10 µg of total RNA was run in a 1% TBE-agarose gel, containing 25 mM guanidium thiocyanate (Sigma). After migration, RNAs were transferred on Hybond N+ nitrocellulose membrane (GE Healthcare Life Sciences) by vacuum with the vacuum-blot system (Whayman biometra) or by capillarity and fixed by UV-crosslinking (Stratalinker 1800 Stratagene). RNA detection was performed by hybridization with specific digoxigenin (DIG)-labelled probes complementary to each targeted RNA as described previously (Tomasini et al., 2017).

Western blot and microscopy

Total proteins were extracted from cultures grown for 4 hours in BHI before harvesting by centrifugation. Bacterial pellets were resuspended in 1X PBS and lysed using the Fastprep (MP Biomedicals). Quantification of proteins was performed with Bradford reagent following the manufacture protocol (BioRad) and BSA as the standard. Equal amounts (0.5 µg) of total proteins were migrated on 12% polyacrylamide-SDS- gels and transferred onto PVDF membranes with a Trans-blot Turbo Transfer system (BioRad). Blots were incubated with anti-GFP at the final dilution 1:5000 followed by an incubation with a goat anti-mouse peroxidase (HRP) (Biorad) at the final dilution 1:5000. Gels were stained by Coomassie blue as loading controls of samples.

For fluorescent microscopy, 10 µl of the same cultures was spread between slide and coverslip and observed with Epifluorescence microscope TiE, Nikon (excitation source: LED Spectra X, Lumencor, camera: Orca-Flash IV, objectif: Plan Apo x100 oil, numerical aperture: 1.45) under excitation at 475nm and emission at 514 +/- 24nm.

MAPS and transcriptomic analyses

Crude bacterial extracts were prepared in duplicates from cultures of Δ rsaG strain expressing MS2-RsaG or from WT strain expressing MS2 grown for 5 h in BHI and then 0.5% G6P was added for another hour of incubation before harvesting and lysis as previously described

(Lalaouna et al., 2018). After MS2-affinity chromatography, RNA was purified from the elution fraction and used either for Northern Blot or DNase I treated prior to RNA-seq analysis as described (Lalaouna et al., 2018). The enrichment values were calculated by DEseq2 (P-adj < 0.05; fold change > 2). For transcriptomic, total RNA was extracted from the WT and Δ rsaG strains containing the empty plasmid pCN51::P3 or Δ rsaG with pCN51::P3::rsaG. Growth cultures were performed as in MAPS and in duplicates. RNA was DNase I treated and depleted with the bacterial RiboZero kit from Illumina. The cDNA libraries were constructed with the ScriptSeq complete kit (bacteria) from Illumina and sequenced using HiSeq 1000 with a V3 Reagent kit (Illumina). We used DEseq2 for differential expression analysis (P-adj < 0.05; fold change > 1,5), as published (Bronesky et al., 2019).

Preparation of RNAs for *in vitro* experiments

Transcription of RsaG or RsaI was achieved with linearized pUC18 vectors (Bronesky et al., 2019; Geissmann et al., 2009). PCR fragments containing RsaG mut1, mut2 and mut1/2, or the 5'UTR of selected mRNAs downstream the sequence of the T7 promoter were used as templates for *in vitro* transcription using T7 RNA polymerase (see Table S4 for oligonucleotides and sizes of DNA fragments). RNAs were purified from 6% or 8% polyacrylamide-8 M urea gel, eluted with 0.5 M ammonium acetate pH 6.5, 1 mM EDTA and 0.1 % SDS and finally precipitated in cold absolute ethanol. Dephosphorylated RNAs were labelled with T4 polynucleotide kinase (Fermentas) and [γ ³²P] ATP.

Gel retardation assay

Radiolabelled RsaG, the mutants RsaG mut1, mut2, mut1/2 (10'000 cps/sample, < 1 pM) and cold mRNAs were renatured separately by incubation at 90°C for 1 min in 100 mM Tris-HCl pH 7,5, 300 mM KCl, 200 mM NH₄Cl, cooled down 1 min on ice, and incubated at room temperature 10 min in presence of 10 mM MgCl₂. Complexes were formed at 37°C for 15 min, then 10 μ l of glycerol blue was added and the samples loaded on a native 6% polyacrylamide gel containing 10 mM MgCl₂ (4-6 h, 300 V, 4°C).

Toe-printing assays

The preparation of 30S subunits of *S. aureus* was described elsewhere (Khusainov et al., 2017), likewise the extension inhibition conditions (Fechter et al., 2009). Increasing concentrations of RsaG were used to monitor its effect on the simplified translational initiation complex formed with *S. aureus* 30S ribosomal subunits, initiator tRNA^{fmet} and mRNAs (*ccpA*, *tcaR*, *sarA*, *ldh_1*, *rex*).

Primer extension assays

In order to determine the transcriptional start site of RsaG from *S. simiae*, 15 µg of total RNA from a bacterial culture grown in presence or absence of G6P was reversed transcribed with the AMV reverse transcriptase (NEB) and the 5'-radiolabelled oligonucleotide RsaG-like rev. The reaction was performed and analysed as described in Lalaouna et al. (2019). The sequencing ladder was obtained with a PCR product obtained with oligonucleotides RsaG-like for/RsaG-like rev.

Monitoring of the 5' extremity of RsaG

Terminator™ 5'-Phosphate-dependent exonuclease Terminator (Epicentre) enzyme degrades secondary transcripts carrying a mono-phosphate 5' extremity but has no effect on primary transcript displaying a tri-phosphate 5' extremity and also on highly structured transcript as 5S rRNA. Total RNA (10 µg) was extracted from HG001 strain grown for 4 h culture in BHI with or without G6P 0.5 % and incubated for 1 h at 30°C with Terminator Exonuclease (1U) and Terminator 1X Reaction buffer A. Furthermore, RNAs were purified with phenol-chloroforme alcohol isoamyl extraction and ethanol precipitation. Northern blot experiment was performed on 1% agarose gel containing 25 mM guanidium thiocyanate.

Co-culture experiments

Human colon epithelial HT29 and HT29MTX, human tumorigenic lung epithelial A549 and human tumorigenic liver HU-H7 cell lines were maintained in Dulbecco's Modified Eagle's Medium (DMEM) (Thermofisher) medium with 10% FCS and 1% penicillin-streptomycin (Gibco) at 37°C under 5% CO₂ atmosphere. At 4 days or 14 days (HT29 and HT29MTX cells), cells are plated in 100 X 20 mm Petri dishes at 0.75x10⁶ cells for HU-H7, 10⁶ cells for A549 and 1.5x10⁶ cells for HT-29 and HT-29MTX. Then 10 ml of HG001 or HG001 $\Delta hptRS$ grown for 4 h in MHB, when RsaG is poorly expressed (ca. OD_{600nm} = 1) are added to human cells for 1 h at 37°C. As controls, fresh 10 ml MHB was added to cells and 10 ml of HG001 bacterial culture was incubated with or without 0.5% G6P. Then supernatants of human cells were carefully collected and centrifuged. Bacterial pellets were used to extract total RNA, which is analysed by Northern blot and supernatants were examined for the presence of mucus by dot blot experiments.

Dot Blot assay

1 µl (or 4 µl supernatant of infected HT-29 cells) of co-cultures supernatants or non-infected supernatants was drop on a PVDF membrane activated with ethanol 95% and equilibrated in TBS 1X tween 0.1%. The membranes were incubated with a rabbit anti-Muc2 or a mouse anti-

Muc5AC at the final dilution of 1:1000 followed by respectively a goat anti-rabbit and a goat anti-mouse at the final dilution of 1:10000.

Quantification of intracellular expression

The human myoblasts CTi400 were cultured in KMEM medium (1v M199, 4v DMEM), 20% fetal bovine serum (v/v), 25 µg/ml Fetuin, 0.5 ng/ml bFGF, 5 ng/ml EGF, 5 µg/ml Insulin, 0.2 mg/ml dexamethasone at 37°C under 5% CO₂ atmosphere. The murine macrophages RAW 264.7 were cultured in DMEM (Thermofisher) supplemented with 10% FBS at 37°C under 5% CO₂ atmospheres.

The intracellular infection of cells was performed using gentamycin protection assay as described in Trouillet et al. (2011) with modifications. Cells were seeded at 80,000 cells per well in 24-well tissue culture plates. After 24 h, cells were washed twice with 1 mL of PBS and infected at a multiplicity of infection (MOI) of 10:1 with bacterial culture (9 h of growth) diluted in cell medium antibiotic free. The MOI was confirmed by CFU counting upon agar plate inoculation. After 2 h of infection at 37°C, cells were washed twice with 1 mL of PBS and incubated for 1 h in medium containing 200 µg/mL gentamicin and 10 µg/mL lysostaphin to kill extracellular bacteria. After the antibiotic treatment, 3 wells were pooled, cells and bacteria were harvested by trypsin detachment and centrifugation. As controls, bacteria were incubated for 1 h in KMEM medium or DMEM under the same conditions as above.

To quantify RsaG, pellets were treated with 20 µg lysostaphin and RNA isolation was performed using the RNeasy Plus mini kit (QIAGEN) according to the manufacturer's instructions. The RNA was quantified using a NanoDrop spectrophotometer, and 150 ng of total RNA was reverse transcribed into cDNA using Reverse Transcriptase System (Promega). 1 µL of 1/5 diluted cDNA was used as template for the real-time PCR amplification using PowerUp SYBR® Green Master Mix and a StepOne Plus system (Applied Biosystem) with specific primers shown in Table S4. Relative RsaG amount analysis was performed by using Δ Ct methods using *gyrB* gene as an internal standard and confirmed by *hu* and 16S gene (see Table S4 for oligonucleotides).

ACKNOWLEDGMENTS

The cell lines HT-29 cells or HT-29 MTX, as the antibodies anti-MUC-2 and anti-MUC-5CA, were kindly supplied by Dr. Benoît Marteyn, (IBMC, Strasbourg), A549 cell line by Dr. Alain Lescure, (IBMC, Strasbourg), and Hu-H7 cell line by Dr. Catherine Schuster (Institut de virologie, Strasbourg). We are thankful to Anne-Catherine Helfer who has performed some of the footprinting experiments.

FUNDING

This work was supported by the “Agence Nationale de la Recherche” (ANR, Grant ANR-16-CE11-0007-01, RIBOSTAPH, to PR). It has also been published under the framework of the labEx NetRNA ANR-10-LABX-0036 and of ANR-17-EURE- 0023 (to PR), as funding from the state managed by ANR as part of the investments for the future program. ED was supported by the “Fondation pour la Recherche Médicale” (FDT201904007957).

REFERENCES

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*, gkw343.

Ando, M., Manabe, Y.C., Converse, P.J., Miyazaki, E., Harrison, R., Murphy, J.R., and Bishai, W.R. (2003). Characterization of the Role of the Divalent Metal Ion-Dependent Transcriptional Repressor MntR in the Virulence of *Staphylococcus aureus*. *Infection and Immunity* *71*, 2584–2590.

Arnaud, M., Chastanet, A., and Débarbouillé, M. (2004). New Vector for Efficient Allelic Replacement in Naturally Nontransformable, Low-GC-Content, Gram-Positive Bacteria. *Appl. Environ. Microbiol.* *70*, 6887–6891.

Augagneur, Y., King, A.N., Germain-Amiot, N., Sassi, M., Fitzgerald, J.W., Sahukhal, G.S., Elasri, M.O., Felden, B., and Brinsmade, S.R. (2020). Analysis of the CodY RNome reveals RsaD as a stress-responsive riboregulator of overflow metabolism in *Staphylococcus aureus*. *Mol Microbiol* *113*, 309–325.

Ballal, A., and Manna, A.C. (2009). Regulation of Superoxide Dismutase (sod) Genes by SarA in *Staphylococcus aureus*. *JB* *191*, 3301–3310.

Behrens, I., Stenberg, P., Artursson, P., and Kissel, T. (2001). Transport of lipophilic drug molecules in a new mucus-secreting cell culture model based on HT29-MTX cells. *Pharmaceutical Research* *18*, 1138–1145.

Bobrovskyy, M., and Vanderpool, C.K. (2013). Regulation of Bacterial Metabolism by Small RNAs Using Diverse Mechanisms. *Annu. Rev. Genet.* *47*, 209–232.

Brandenberger, M., Tschierske, M., Giachino, P., Wada, A., and Berger-Bächi, B. (2000). Inactivation of a novel three-cistronic operon *tcaR-tcaA-tcaB* increases teicoplanin resistance in *Staphylococcus aureus*. *Biochimica et Biophysica Acta (BBA) - General Subjects* *1523*, 135–139.

Bronesky, D., Desgranges, E., Corvaglia, A., François, P., Caballero, C.J., Prado, L., Toledo-Arana, A., Lasa, I., Moreau, K., Vandenesch, F., et al. (2019). A multifaceted small RNA modulates gene expression upon glucose limitation in *Staphylococcus aureus*. *The EMBO Journal* e99363.

Brosse, A., and Guillier, M. (2018). Bacterial Small RNAs in Mixed Regulatory Networks. *Microbiol Spectr* *6*.

Caldelari, I., Chane-Woon-Ming, B., Noirot, C., Moreau, K., Romby, P., Gaspin, C., and Marzi, S. (2017). Complete Genome Sequence and Annotation of the *Staphylococcus aureus* Strain

HG001. Genome Announc. 5, e00783-17.

Chatterjee, S.S., Joo, H.-S., Duong, A.C., Dieringer, T.D., Tan, V.Y., Song, Y., Fischer, E.R., Cheung, G.Y.C., Li, M., and Otto, M. (2013). Essential *Staphylococcus aureus* toxin export system. *Nat Med* 19, 364–367.

Cheng, C., Bhardwaj, N., and Gerstein, M. (2009). The relationship between the evolution of microRNA targets and the length of their UTRs. *BMC Genomics* 10, 431.

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

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 Mol. Biol.* 540, 215–232.

Chico-Calero, I., Suarez, M., Gonzalez-Zorn, B., Scotti, M., Slaghuis, J., Goebel, W., The European *Listeria* Genome Consortium, and Vazquez-Boland, J.A. (2002). Hpt, a bacterial homolog of the microsomal glucose- 6-phosphate translocase, mediates rapid intracellular proliferation in *Listeria*. *Proceedings of the National Academy of Sciences* 99, 431–436.

Christmas, B.A.F., Rolfe, M.D., Rose, M., and Green, J. (2019). *Staphylococcus aureus* adaptation to aerobic low-redox-potential environments: implications for an intracellular lifestyle. *Microbiology* 165, 779–791.

Crooke, A.K., Fuller, J.R., Obrist, M.W., Tomkovich, S.E., Vitko, N.P., and Richardson, A.R. (2013). CcpA-Independent Glucose Regulation of Lactate Dehydrogenase 1 in *Staphylococcus aureus*. *PLoS ONE* 8, e54293.

Desgranges, E., Marzi, S., Moreau, K., Romby, P., and Caldelari, I. (2019). Noncoding RNA. *Microbiology Spectrum* 7(2).

Eisenhardt, K.M.H., Reuscher, C.M., and Klug, G. (2018). PcrX, an sRNA derived from the 3'-UTR of the *Rhodobacter sphaeroides* *puf* operon modulates expression of *puf* genes encoding proteins of the bacterial photosynthetic apparatus. *Molecular Microbiology* 110, 325–334.

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 Mol. Biol.* 540, 247–263.

Garzoni, C., Francois, P., Huyghe, A., Couzinet, S., Tapparel, C., Charbonnier, Y., Renzoni, A., Lucchini, S., Lew, D.P., Vaudaux, P., et al. (2007). A global view of *Staphylococcus aureus* whole genome expression upon internalization in human epithelial cells. *BMC Genomics* 8, 171.

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. *Nucl. Acids Res.* 37, 7239–7257.

Gemayel, R., Fortpied, J., Rzem, R., Vertommen, D., Veiga-da-Cunha, M., and Van Schaftingen, E. (2007). Many fructosamine 3-kinase homologues in bacteria are ribulosamine/erythrusamine 3-kinases potentially involved in protein deglycation: Bacterial fructosamine 3-kinase homologues. *FEBS Journal* 274, 4360–4374.

Guerrier-Takada, C., and Altman, S. (1984). Catalytic activity of an RNA molecule prepared by transcription in vitro. *Science* 223, 285–286.

Hamza, T., and Li, B. (2014). Differential responses of osteoblasts and macrophages upon *Staphylococcus aureus* infection. *BMC Microbiol* 14, 207.

Harraghy, N., Kormanec, J., Wolz, C., Homerova, D., Goerke, C., Ohlsen, K., Qazi, S., Hill, P., and Herrmann, M. (2005). *sae* is essential for expression of the staphylococcal adhesins Eap and Emp. *Microbiology* 151, 1789–1800.

Jefferson, K.K., Pier, D.B., Goldmann, D.A., and Pier, G.B. (2004). The Teicoplanin-Associated Locus Regulator (TcaR) and the Intercellular Adhesin Locus Regulator (IcaR) Are Transcriptional Inhibitors of the *ica* Locus in *Staphylococcus aureus*. *Journal of Bacteriology* 186, 2449–2456.

Khusainov, I., Vicens, Q., Ayupov, R., Usachev, K., Myasnikov, A., Simonetti, A., Validov, S., Kieffer, B., Yusupova, G., Yusupov, M., et al. (2017). Structures and dynamics of hibernating ribosomes from *Staphylococcus aureus* mediated by intermolecular interactions of HPF. *EMBO J* 36, 2073–2087.

Kim, H.M., Shin, J.-H., Cho, Y.-B., and Roe, J.-H. (2014). Inverse regulation of Fe- and Ni-containing SOD genes by a Fur family regulator Nur through small RNA processed from 3'UTR of the *sodF* mRNA. *Nucleic Acids Research* 42, 2003–2014.

von Kleist, S., Chany, E., Burtin, P., King, M., and Fogh, J. (1975). Immunohistology of the Antigenic Pattern of a Continuous Cell Line From a Human Colon Tumor. *JNCI: Journal of the National Cancer Institute* 55, 555–560.

Lalaouna, D., Desgranges, E., Caldelari, I., and Marzi, S. (2018). MS2-Affinity Purification Coupled With RNA Sequencing Approach in the Human Pathogen *Staphylococcus aureus*. In *Methods in Enzymology*, (Elsevier), pp. 393–411.

Lalaouna, D., Baude, J., Wu, Z., Tomasini, A., Chicher, J., Marzi, S., Vandenesch, F., Romby, P., Caldelari, I., and Moreau, K. (2019). RsaC sRNA modulates the oxidative stress response of *Staphylococcus aureus* during manganese starvation. *Nucleic Acids Research* 47, 9871–9887.

Linder, P., Lemeille, S., and Redder, P. (2014). 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* 10, e1004207.

Makhlin, J., Kofman, T., Borovok, I., Kohler, C., Engelmann, S., Cohen, G., and Aharonowitz, Y. (2007). *Staphylococcus aureus* ArcR Controls Expression of the Arginine Deiminase Operon. *JB* 189, 5976–5986.

Marincola, G., Wencker, F.D.R., and Ziebuhr, W. (2019). The Many Facets of the Small Non-coding RNA RsaE (RoxS) in Metabolic Niche Adaptation of Gram-Positive Bacteria. *Journal of Molecular Biology* 431, 4684–4698.

Martin, P.K., Li, T., Sun, D., Biek, D.P., and Schmid, M.B. (1999). Role in Cell Permeability of an Essential Two-Component System in *Staphylococcus aureus*. *J. Bacteriol.* 181, 3666–3673.

Menendez-Gil, P., Caballero, C.J., Catalan-Moreno, A., Irurzun, N., Barrio-Hernandez, I.,

Caldelari, I., and Toledo-Arana, A. (2020). Differential evolution in 3'UTRs leads to specific gene expression in *Staphylococcus*. *Nucleic Acids Research* 48, 2544–2563.

Michalik, S., Depke, M., Murr, A., Gesell Salazar, M., Kusebauch, U., Sun, Z., Meyer, T.C., Surmann, K., Pförtner, H., Hildebrandt, P., et al. (2017). A global *Staphylococcus aureus* proteome resource applied to the in vivo characterization of host-pathogen interactions. *Sci Rep* 7, 9718.

Miyakoshi, M., Chao, Y., and Vogel, J. (2015). Regulatory small RNAs from the 3' regions of bacterial mRNAs. *Current Opinion in Microbiology* 24, 132–139.

Pagels, M., Fuchs, S., Pané-Farré, J., Kohler, C., Menschner, L., Hecker, M., McNamarra, P.J., Bauer, M.C., Von Wachenfeldt, C., Liebeke, M., et al. (2010). Redox sensing by a Rex-family repressor is involved in the regulation of anaerobic gene expression in *Staphylococcus aureus*: Redox sensitive gene regulation in *S. aureus*. *Molecular Microbiology* 76, 1142–1161.

Park, J.Y., Kim, J.W., Moon, B.Y., Lee, J., Fortin, Y.J., Austin, F.W., Yang, S.-J., and Seo, K.S. (2015). Characterization of a Novel Two-Component Regulatory System, HptRS, the Regulator for the Hexose Phosphate Transport System in *Staphylococcus aureus*. *Infection and Immunity* 83, 1620–1628.

Reed, J.M., Olson, S., Brees, D.F., Griffin, C.E., Grove, R.A., Davis, P.J., Kachman, S.D., Adamec, J., and Somerville, G.A. (2018). Coordinated regulation of transcription by CcpA and the *Staphylococcus aureus* two-component system HptRS. *PLoS ONE* 13, e0207161.

Richardson, A.R. (2019). Virulence and Metabolism. *Microbiology Spectrum* 7(2).

Rochat, T., Bohn, C., Morvan, C., Le Lam, T.N., Razvi, F., Pain, A., Toffano-Nioche, C., Ponien, P., Jacq, A., Jacquet, E., et al. (2018). The conserved regulatory RNA RsaE down-regulates the arginine degradation pathway in *Staphylococcus aureus*. *Nucleic Acids Research* 46, 8803–8816.

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 Pathog* 10.

Runyen-Janecky, L.J., and Payne, S.M. (2002). Identification of Chromosomal *Shigella flexneri* Genes Induced by the Eukaryotic Intracellular Environment. *Infection and Immunity* 70, 4379–4388.

Seidl, K., Müller, S., François, P., Kriebitzsch, C., Schrenzel, J., Engelmann, S., Bischoff, M., and Berger-Bächi, B. (2009). Effect of a glucose impulse on the CcpA regulon in *Staphylococcus aureus*. *BMC Microbiol* 9, 95.

Sit, B., Crowley, S.M., Bhullar, K., Lai, C.C.-L., Tang, C., Hooda, Y., Calmettes, C., Khambati, H., Ma, C., Brumell, J.H., et al. (2015). Active Transport of Phosphorylated Carbohydrates Promotes Intestinal Colonization and Transmission of a Bacterial Pathogen. *PLoS Pathog* 11, e1005107.

Sorensen, H.M., Keogh, R.A., Wittekind, M.A., Caillet, A.R., Wiemels, R.E., Laner, E.A., and Carroll, R.K. (2020). Reading between the Lines: Utilizing RNA-Seq Data for Global Analysis of sRNAs in *Staphylococcus aureus*. *MSphere* 5, e00439-20

Tomasini, A., Moreau, K., Chicher, J., Geissmann, T., Vandenesch, F., Romby, P., Marzi, S.,

and Caldelari, I. (2017). The RNA targetome of *Staphylococcus aureus* non-coding RNA RsaA: impact on cell surface properties and defense mechanisms. *Nucleic Acids Research* 45, 6746–6760.

Trouillet, S., Rasigade, J.-P., Lhoste, Y., Ferry, T., Vandenesch, F., Etienne, J., and Laurent, F. (2011). A novel flow cytometry-based assay for the quantification of *Staphylococcus aureus* adhesion to and invasion of eukaryotic cells. *Journal of Microbiological Methods* 86, 145–149.

Villanueva, M., García, B., Valle, J., Rapún, B., Ruiz de los Mozos, I., Solano, C., Martí, M., Penadés, J.R., Toledo-Arana, A., and Lasa, I. (2018). Sensory deprivation in *Staphylococcus aureus*. *Nat Commun* 9, 523.

Vitko, N.P., Grosser, M.R., Khatri, D., Lance, T.R., and Richardson, A.R. (2016). Expanded Glucose Import Capability Affords *Staphylococcus aureus* Optimized Glycolytic Flux during Infection. *MBio* 7, e00296-16.

Wagner, E.G.H., and Romby, P. (2015). Chapter Three - Small RNAs in Bacteria and Archaea: Who They Are, What They Do, and How They Do It. In *Advances in Genetics*, J.C.D. and S.F.G. *Advances in Genetics*, pp. 133–208.

Wang, C., Chao, Y., Matera, G., Gao, Q., and Vogel, J. (2020). The conserved 3' UTR-derived small RNA NarS mediates mRNA crossregulation during nitrate respiration. *Nucleic Acids Research* 48, 2126–2143.

Xue, T., You, Y., Hong, D., Sun, H., and Sun, B. (2011). The *Staphylococcus aureus* KdpDE Two-Component System Couples Extracellular K⁺ Sensing and Agr Signaling to Infection Programming. *Infect. Immun.* 79, 2154–2167.

Yang, Y., Sun, H., Liu, X., Wang, M., Xue, T., and Sun, B. (2016). Regulatory mechanism of the three-component system HptRSA in glucose-6-phosphate uptake in *Staphylococcus aureus*. *Med Microbiol Immunol* 205, 241–253.

SUPPLEMENTARY MATERIALS AND METHODS

RT-PCR

RT-PCR was accomplished using OneStep RT-PCR kit (Qiagen) according to the protocol provided by the manufacture. Total RNAs extracted after 6 h of bacterial growth in BHI containing or not G6P as templates and primers are listed in Table S4. The same experiment amplified the 16S rRNA or was run without RT as controls.

RNase P cleavage assay

The *pretRNA^{tyr}*, *tRNA^{tyr}*, *rnpB* and *uhpT-RsaG* DNA fragments were amplified by PCR from genomic DNA with primers listed in Table S4 and used as template in *in vitro* transcription assay (see above). RNAs were denatured 3 minutes at 95°C, cooled until reaching room temperature and renatured 15 minutes at 37°C in Tris-HCl 50 mM pH 8, MgCl₂ 100 mM, NH₄Cl 800 mM. Then 10 pmol of *pretRNA^{tyr}*, *tRNA^{tyr}*, or 5 pmol of *uhpT-RsaG* were annealed with 5 to 20 pmol of the catalytic RNA *rnpB* for 15 min at 37°C. The reaction was stopped by adding one volume of RNA loading dye (0.5 % formamide, 0.025 % SDS, 0.02 5% bromophenol blue, 0.025 % xylene blue and 0.5 mM EDTA). The *pretRNA^{tyr}* and *tRNA^{tyr}* with *rnpB* reactions were migrated on a 10% acrylamide-8 M urea gel (300 V, room temperature, 3 h 30) and the *uhpT-RsaG* with *rnpB* reaction on a 6% acrylamide-8 M urea gel (300V, 20°C, 4 h 30). After BET staining, gel was exposed on UV (gel doc EZ system, BioRad).

In vivo β-galactosidase assays

The pLUG220::*rpoB::rex::lacZ* and pLUG220::*rpoB::rex::lacZ::blaZ::rsaG* plasmids were transformed into the HG001Δ*rsaG* strain. Furthermore, β-galactosidase activity was measured four times as described in Tomasini et al. (2017).

Probing of RsaG-*rex* interaction using Lead(II)

Probing was accomplished as described in Chevalier et al. (2009). Two Lead(II) concentrations (4 or 8 mM) and increasing concentrations of *rex* were used to monitor its effect on RsaG structure (400 to 1600 nM).

SUPPLEMENTARY FILES

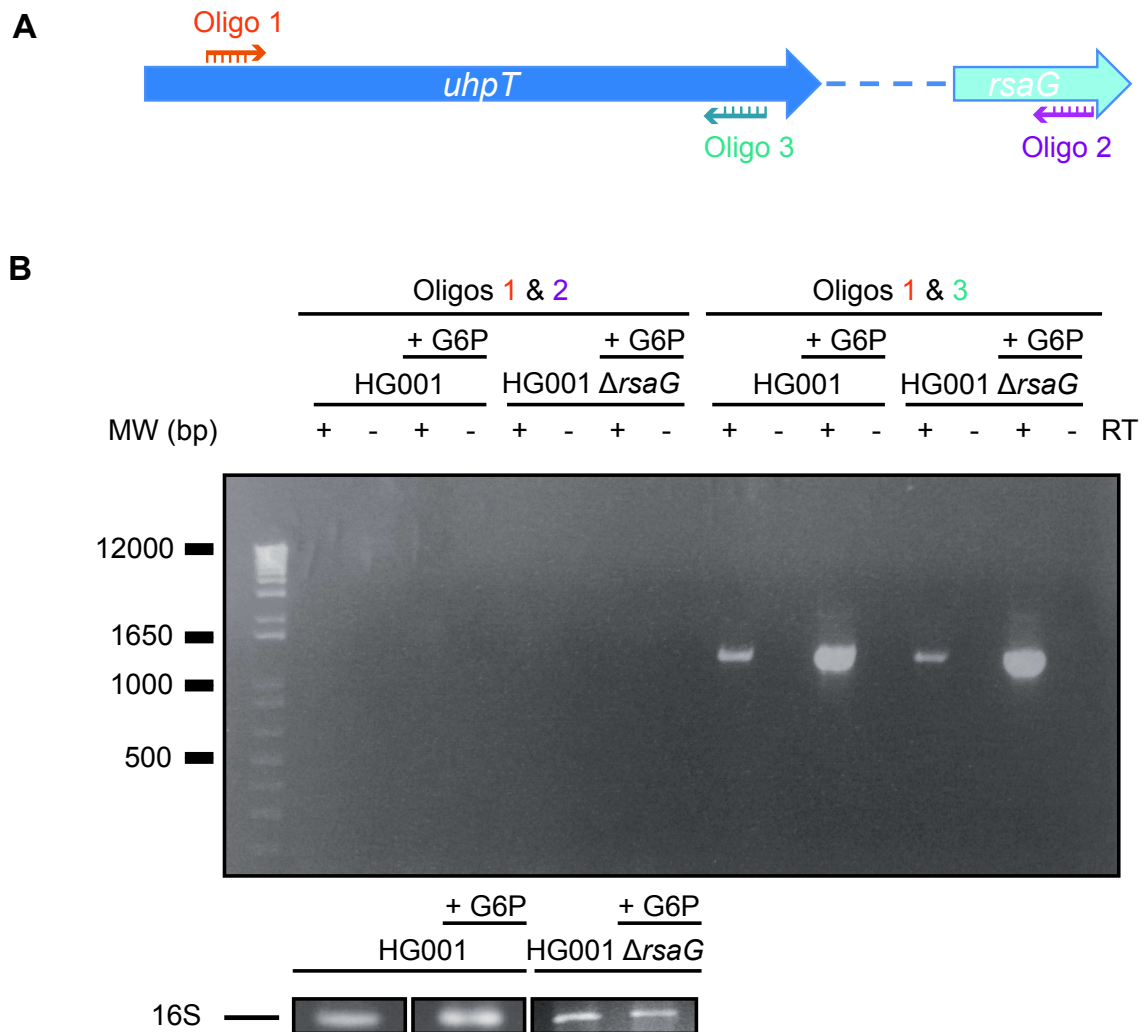


Figure S1

Figure S1: RsaG is derived from the 3'UTR of *uhpT*.

A. Genomic context of the *uhpT* locus with oligonucleotides used for RT-PCR as indicated with coloured arrows. B. RT-PCR using oligonucleotides in both *uhpT* and *rsaG* genes and total RNA extracted from HG001 or HG001 Δ *rsaG* strains after 6 h of growth in BHI medium with or without glucose-6-phosphate (+G6P). 16S rRNA was used as loading control.

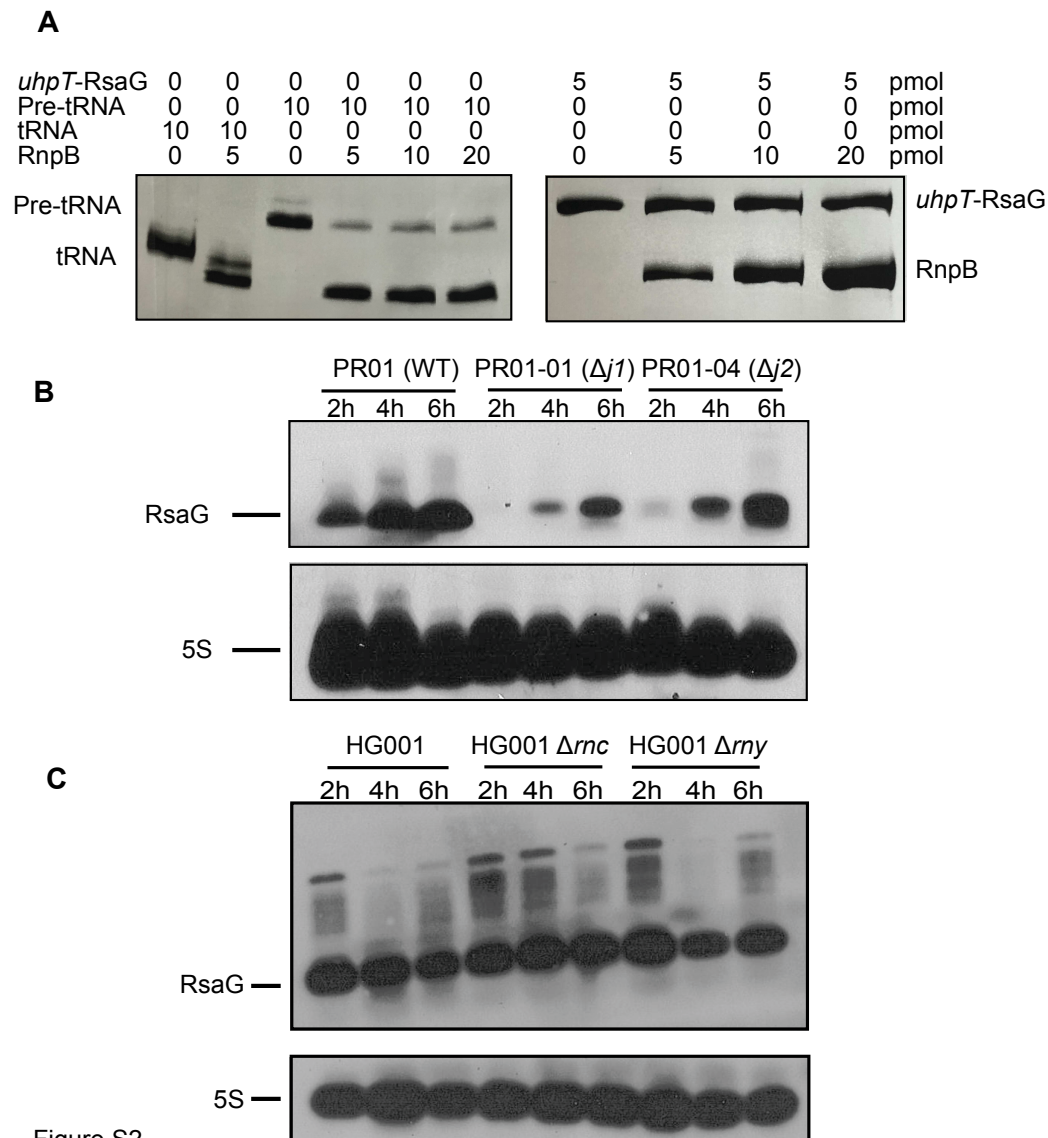


Figure S2

Figure S2: RsaG is matured from *uhpT* mRNA by 5'-3'exoribonucleases J1-J2 and not by RNase P, RNase III or RNase Y.

A. RNaseP *in vitro* assay of pretRNA^{tyr}, tRNA^{tyr} and *uhpT*-RsaG by *rnpB*. Cleavage products are run on an 10% (tRNA) or 6% (*uhpT*-RsaG) acrylamide-8 M urea gel revealed by ethidium bromide. B. Northern blot analysis of RsaG in PR01 wild type (WT), PR01-01 ($\Delta j1$) and PR01-04 ($\Delta j2$) strains. Total RNA was extracted at 2, 4 and 6 h of growth in MHB medium supplemented with 10 mg/L uracil and 0.5% glucose-6-phosphate. C. Northern blot analysis of RsaG in HG001, HG001 Δrnc and HG001 Δrny strains. Total RNA was extracted at 2, 4 and 6 h of growth in MHB medium supplemented with 0.5% glucose-6-phosphate. 5S rRNA (5S) was used as loading control. However, for this controls, we used aliquots of the same RNA preparations, but the migration of the samples was performed in parallel to the experiments on a separate agarose gel because RsaG and 5S rRNA have very similar sizes.

A

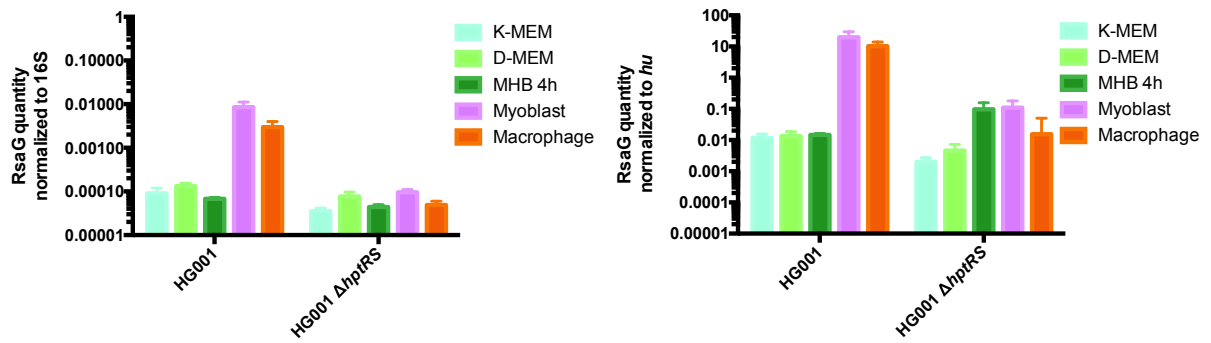


Figure S3

Figure S3: RsaG is induced when *S. aureus* is internalized into myoblasts or macrophages.

Level of RsaG in HG001 and HG001ΔhptRS strains, determined using qRT-PCR. Samples were taken after 1h of *S. aureus* internalization into RAW 264.7 macrophages or CTi400 myoblasts. As control RsaG relative amount was also quantified for HG001 wild type or HG001ΔhptRS cultured for 1 h in K-MEM or D-MEM medium. Data were normalized to 16S or *hu* and represent the mean of three independent experiments ± SD.

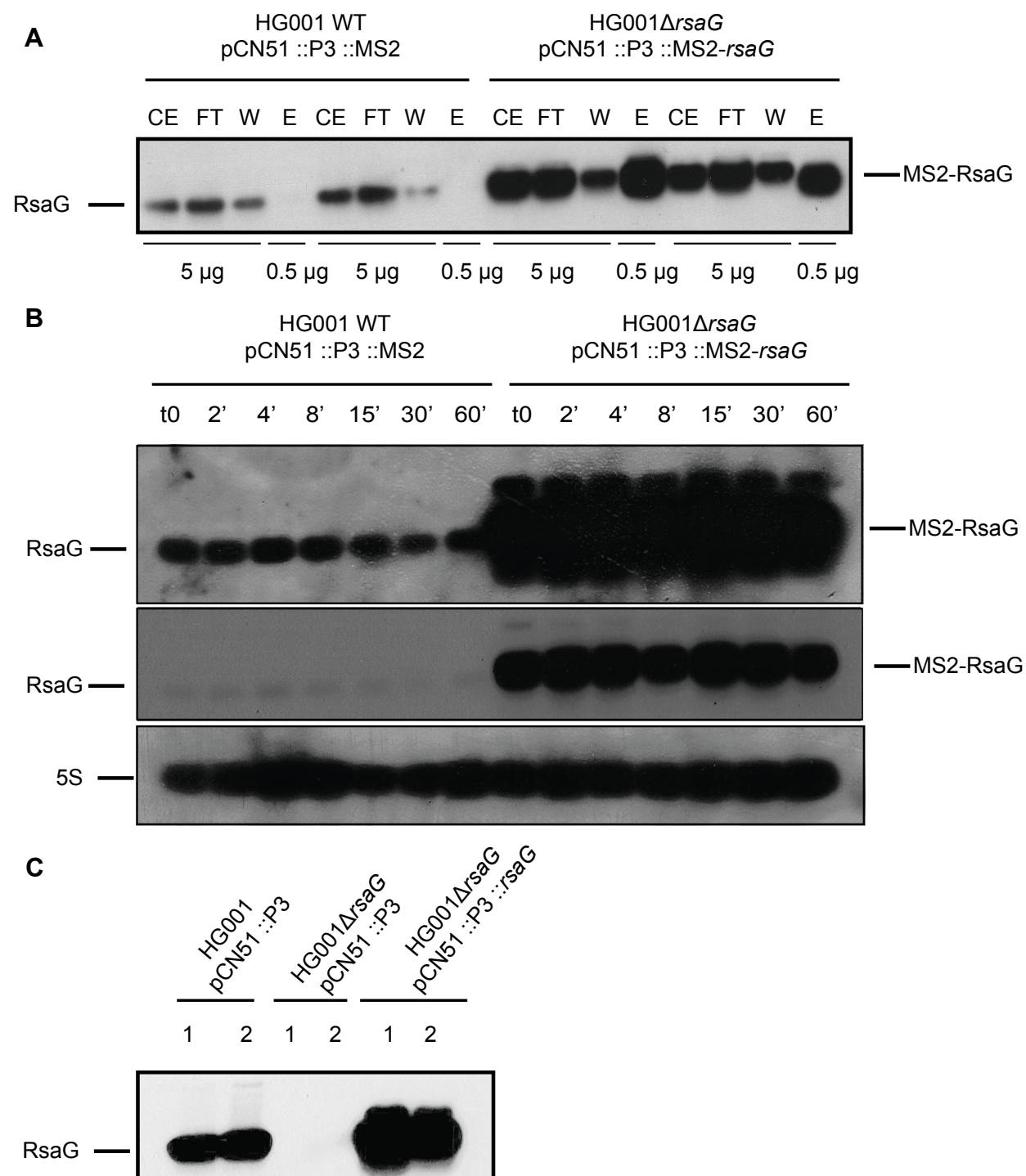


Figure S4

Figure S4: The MS2-RsaG variant is correctly retained by affinity chromatography and is as stable as RsaG.

A. Northern blot targeting RsaG and MS2-RsaG performed on RNAs purified after MS2 chromatography affinity; CE: crude extract ; FT : Flow-Through ; W : Washing ; EL : Elution. For CE/FT/W samples, 5 µg of total RNA was loaded on a 1,5% agarose gel while for E sample, only 0,5 µg of total RNA was used. The two replicates are shown for the RNA purified from the strains HG001::P3::MS2 and HG001ΔrsaG::MS2-RsaG. B. Northern blots of half-lives of RsaG and MS2-RsaG in HG001 and HG001-ΔrsaG strains. The bacterial cultures were treated with rifampicin at 4h of growth and total RNAs were extracted after 2, 4, 8, 15, 30 and 60 minutes at 37°C in BHI. Note that two exposures of the same membrane were presented. 5S rRNA was probed to quantify the yield of RNA in each lane using the same samples but run on two different gels. C. Northern blot experiments showing the expression of RsaG in HG001 and HG001ΔrsaG strains containing the empty plasmid pCN51::P3, and HG001ΔrsaG complemented with a plasmid expressing RsaG under the control of P3 promoter (HG001ΔrsaG pCN51::P3::rsaG). Total RNA was prepared from samples grown for 5 h in BHI and then 0.5% glucose-6-phosphate was added for another hour of incubation. Samples are biological duplicates.

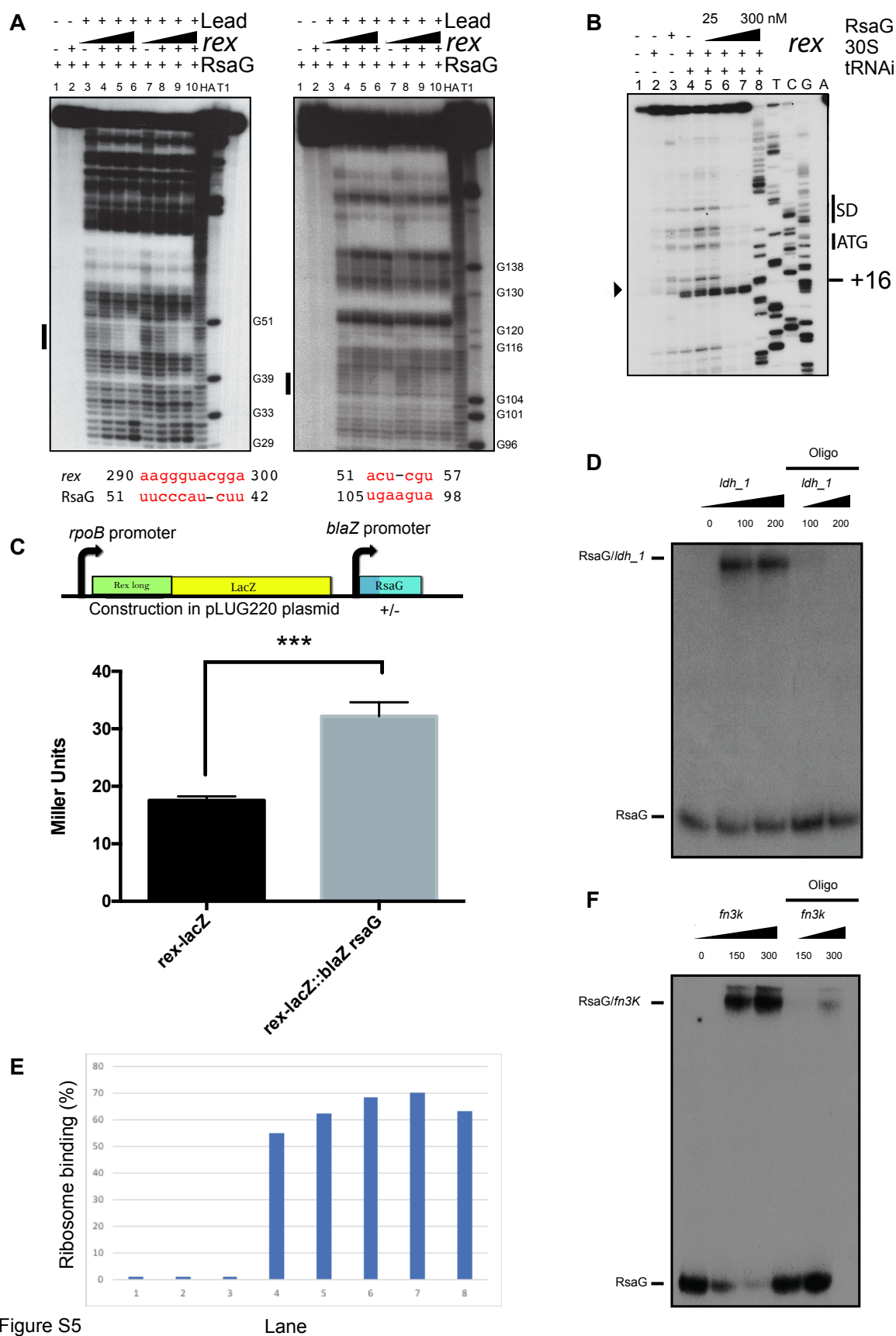


Figure S5

Figure S5: RsaG interaction with *rex*, *ldh_1* and *fn3k*.

A. Footprinting experiments of *rex* mRNA on 5'-end-labeled RsaG probed by lead(II) induced cleavages. The experiments were performed with increasing concentrations of *rex* mRNA (400 to 1600 nM) and two lead(II) concentrations (4 or 8 mM). The regions of RsaG protected by the mRNA, are shown by bars on the left side of the autoradiographies. The proposed basepairing interactions are given below the autoradiographies. B. Toe-print assays showing RsaG effect on the formation of the ribosomal initiation complex of *rex*. Lane 1 : incubation control of mRNA alone; lane 2 : incubation control of mRNA with 30S subunits; lane 3: incubation control of mRNA with RsaG; lane 4 : formation of the ribosomal initiation complex containing mRNA, 30S and the initiator tRNA^{Met} (tRNA_i); lanes 5 to 8 : formation of the initiation complex in the presence of increasing concentrations of RsaG, respectively : 25 nM (lane 5), 50 nM (lane 6), 150 nM (lane 7), 300 nM (lane 8). Lanes T, A, C, G: sequencing ladders. The Shine and Dalgarno (SD) sequence, the start site of translation (ATG) and the toe-printing signals (+16) are indicated. C. The β -galactosidase activity (Miller Units) have been measured from *PrpoB::rex::lacZ* and *PrpoB::rex::lacZ::PblaZ::rsaG* in HG001 Δ *rsaG* strain. The β -galactosidase activity was normalized for bacterial density and the results represented the mean of four independent experiments, statistical analysis with Student's *t*-test, ****p*<0,001. At the top of the graph is represented the genetic construction in plasmid pLUG220. D. Electrophoretic mobility shift assay (EMSA) showing the formation of the complex between RsaG and *ldh_1* and in competition with an oligonucleotide encompassing the region from -14 to +20 of *ldh_1*. 5' end-labeled RsaG was incubated with increasing concentration of cold mRNA (written on the top of the gel, in nM) or with 200 nM of oligonucleotide. E. Quantification of 30S ribosome binding on *ldh_1* mRNA. Relative toeprints were calculated by relating the intensity of the band corresponding to the toeprint signal to the sum of the intensities of this band and the band corresponding to the full-length transcript. F. EMSA showing the formation of the complex between RsaG and *fn3K* and in competition with an oligonucleotide encompassing the region from 268 to 285 of *fn3K*. 5' end-labeled RsaG was incubated with increasing concentration of cold mRNA or with 200 nM of oligonucleotide (written on the top of the gel for each transcript, in nM).

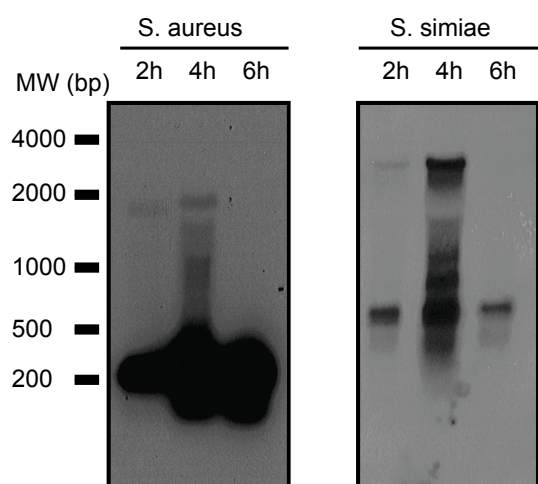


Figure S6

Figure S6: Maturation of a 3'UTR derived sRNA from *uhpT* mRNA who present size difference between *S. aureus* and *S. simiae*.

Northern blot analysis of RsaG in *S. aureus* and *S. simiae*. Total RNA was extracted at 2, 4 and 6h of growth in BHI medium containing 0.5% glucose-6-phosphate. RsaG from *S. aureus* was detected with a specific probe against *S. aureus* RsaG sequence, while for *S. simiae*, the probe was designed against RsaG like sequence.

Table S1: The list of RNAs sequenced using MS2-RsaG as the bait. The enrichment values corresponded to the fold changes (ratio MS2-RsaG/MS2 control), and the p-values have been calculated using DEseq2. We have considered in this study all RNAs, which were enriched at least 2-fold. Duplicates have been carried out for each assay.

id	gene	product	FoldChange	pvalue	padj	Kdapp	COPRA position mRNA	COPRA position sRNA	intaRNA position mRNA	intaRNA position sRNA	Common targets
HG001_02756	-	RsaG	512.328	9.63E-30	2.54E-26						
HG001_01207	mscL	Large-conductance mechanosensitive channel	33.182	1.20E-14	7.91E-12				9 - 31	123 - 144	
HG001_00556	sarA	Transcriptional regulator SarA	27.433	2.09E-11	9.17E-09	< 50	-47 -9 (SD)	87 - 118	-48 -9 (SD)	89 - 112	RsaC
HG001_02744	cspA	cspA	23.595	3.89E-19	5.12E-16	no					
HG001_00902	slyA_2	Transcriptional regulator SlyA	20.283	3.15E-06	0.00028634				-18 -5 (SD)	141 - 152	
HG001_00607	sarX	HTH-type transcriptional regulator SarX	19.302	1.75E-08	2.31E-06	< 50			-37 -22	30 - 46	
HG001_02670	arcR	HTH-type transcriptional regulator ArcR	17.366	2.68E-05	0.00138302		-26 -7	34 - 53	-27 -8 (SD)	38 - 56	
HG001_02069	rex	Redox-sensing transcriptional repressor Rex	15.478	6.17E-15	5.42E-12	< 150			-54 -48 et 184 - 195	98 - 105 et 42 - 51	
HG001_01075	rpmB	50S ribosomal protein L28	14.691	3.25E-10	9.50E-08				150 - 207	(12 - 19)	
HG001_00933	ptsH	Phosphocarrier protein HPr	13.164	2.19E-12	1.16E-09	< 50			226 - 234	18 - 24	
HG001_02671	arcC2	Carbamate kinase 2	12.351	0.00021113	0.00567693				-27 -8	35 - 55	
HG001_02394	tcaR_2	HTH-type transcriptional regulator TcaR	11.804	9.14E-11	3.44E-08	< 50	-39 -9	88 - 108	-41 -8 (SD)	89 - 112	
HG001_01639	rpsD	30S ribosomal protein S4	11.767	9.00E-09	1.48E-06				214 - 225	166 - 176	
HG001_02105	kdpA	Potassium-transporting ATPase A chain	10.478	0.00125157	0.01998718				-125 -113	145 - 156	
HG001_00932	-	hypothetical protein	10.129	1.91E-09	3.36E-07				-44 -21	132 - 150	
HG001_02293	glcU_2	putative glucose uptake protein GlcU	9.983	1.67E-08	2.31E-06	no					RsaI
HG001_02122	-	TPP	9.704	1.47E-08	2.25E-06				-39 -31	50 - 57	
HG001_00785	ndh	NADH dehydrogenase-like protein	9.675	4.77E-10	1.26E-07				-15 -1 (SD)	135 - 150	
HG001_02206	sdrM	Multidrug efflux MFS transporter	9.575	1.49E-09	3.22E-07	< 600			382 - 415	33 - 62	
HG001_00285	ydaF	Putative ribosomal N-acetyltransferase YdaF	9.545	4.94E-06	0.00038304						
HG001_01629	ald2	Alanine dehydrogenase 2	9.38	1.34E-09	3.20E-07				789 - 808	159 - 178	
HG001_01234	femA_2	Aminoacyltransferase FemA	9.299	2.52E-10	8.29E-08						
HG001_01556	-	T-box	8.845	1.59E-09	3.22E-07						
HG001_02347	RsaI	RsaOG	8.714	1.74E-09	3.27E-07	< 50			32 - 25	50 - 43	
HG001_01627	-	Metal-dependent hydrolase	8.364	1.54E-08	2.25E-06						
HG001_02107	kdpE	KDP operon transcriptional regulatory protein KdpE	8.35	9.70E-05	0.00350195				32 - 39	17 - 23	
HG001_01455	-	putative lipoprotein	8.145	2.21E-06	0.000208						
HG001_02147	rpmE2	50S ribosomal protein L31 type B	7.594	4.60E-06	0.00037882						
HG001_02246	budA1	Alpha-acetolactate decarboxylase	7.571	0.0010457	0.01789238				58 - 65	43 - 49	
HG001_02104	kdpB	Potassium-transporting ATPase B chain	7.49	0.00369526	0.03942114				789 - 808	159 - 178	
HG001_02657	p2d	Precorrin-2 dehydrogenase	7.472	0.00016451	0.00494177						
HG001_02672	arcD	Arginine/ornithine antiporter	7.422	0.0049597	0.04840302				917 - 926	166 - 174	
HG001_02118	thiE	Thiamine-phosphate synthase	7.388	0.00130406	0.02035637				540 - 550	38 - 47	
HG001_01730	-	hypothetical protein	7.178	6.18E-07	6.26E-05						
HG001_01025	-	putative N-acetyltransferase	7.007	3.90E-05	0.00177305		(-15 -6)	128 - 148			
HG001_00508	vraX	Staphylococcal complement inhibitor	6.952	0.00013683	0.0044513						
HG001_00926	thiW	Putative HMP/thiamine import ATP-binding protein	6.768	0.00281055	0.03366277						
HG001_02119	thiM	Hydroxyethylthiazole kinase	6.733	0.00129687	0.02035637				617 - 645	36 - 74	
HG001_02677	isaB	Immunodominant staphylococcal antigen B precursor	6.692	0.00098973	0.01715745						
HG001_02351	RpiRc	putative HTH-type transcriptional regulator RpiR family	6.689	1.68E-07	2.10E-05		(-196 -164)	23 - 55	-197 -164	26 - 58	
HG001_01030	ftsL	Cell division protein FtsL	6.534	2.79E-07	3.07E-05						
HG001_01014	ssl14	superantigen-like protein	6.46	1.21E-05	0.00076196						
HG001_01343	fer	Ferredoxin	6.447	8.10E-07	7.91E-05						
HG001_01152	glpD	Aerobic glycerol-3-phosphate dehydrogenase	6.406	0.00027062	0.00685646				798 - 811	35 - 48	
HG001_01151	glpK	Glycerol kinase	6.268	1.73E-05	0.00097074				76 - 88	44 - 55	
HG001_01188	hom	Homoserine dehydrogenase	6.179	0.00142826	0.02175408		(193 -180)	8-20			
HG001_02485	norB_5	Quinolone resistance protein NorB	6.179	3.13E-05	0.00152825						
HG001_01140	ymdB	2,3-cyclic-nucleotide 2-phosphodiesterase	6.162	2.62E-07	3.01E-05						
HG001_01237	-	hypothetical protein	6.114	1.90E-07	2.28E-05						
HG001_00546	-	hypothetical protein	6.093	6.66E-06	0.00046208						
HG001_00107	capH	Capsular polysaccharide biosynthesis protein, Putative acety	6.012	0.00031129	0.00732354						
HG001_02475	-	Putative MFS transporter	5.959	3.73E-05	0.00177305						
HG001_01174	-	Putative phage protein	5.943	0.00087606	0.01580584						
HG001_01309	-	hypothetical protein	5.924	7.70E-05	0.00307318						
HG001_02393	tcaR_1	HTH-type transcriptional regulator TcaR	5.892	0.00026398	0.00675335	< 50					
HG001_00617	-	hypothetical protein	5.857	5.57E-07	5.87E-05						
HG001_02639	budA2	Alpha-acetolactate decarboxylase	5.78	4.00E-05	0.00178763						

HG001_00953	-	hypothetical protein	5.773	2.09E-05	0.00114751					
HG001_00690	hpf	Ribosome hibernation promotion factor	5.74	0.00014053	0.00451591	< 300				
HG001_02120	thiD	Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase	5.737	0.00210012	0.02823381		(-117 -53)	4-56	(-118 -76)	31 - 59
HG001_00330	-	Teg76	5.702	0.00298158	0.03507348					
HG001_02421	-	hypothetical protein	5.678	0.00031061	0.00732354					
HG001_01351	-	phage protein	5.676	5.37E-05	0.00224802					
HG001_00362	-	hypothetical protein	5.662	0.00226702	0.02957229					
HG001_00491	azo1	FMN-dependent NADPH-azoreductase	5.496	0.00171781	0.02528731					
HG001_02659	goxA2	Glutathione peroxidase homolog	5.496	0.00171312	0.02528731	< 50				
HG001_01204	-	hypothetical protein	5.492	0.00091821	0.01632369					
HG001_00570	mntC	Manganese ABC transporter substrate-binding lipoprotein pre	5.462	0.00017954	0.00519884				239 - 280	133 - 175
HG001_02247	budB	Acetolactate synthase	5.437	0.00451921	0.04545088				58 - 65	43 - 49
HG001_00616	-	hypothetical protein	5.429	4.90E-06	0.00038304					
HG001_02106	kdpD	Sensor protein KdpD	5.414	0.00303192	0.03519429				573 - 612	(11 - 50)
HG001_02632	-	hypothetical protein	5.404	2.57E-05	0.00135593					
HG001_00303	-	RsaOB, T-box	5.396	0.00019574	0.00542912					
HG001_01462	-	Glycine	5.316	0.00035778	0.00806856					
HG001_00885	-	hypothetical protein	5.281	0.00068703	0.01302398					
HG001_01057	-	hypothetical protein	5.272	0.00017398	0.00509361					
HG001_00964	ctaA	Heme A synthase	5.254	1.31E-05	0.00080047					
HG001_02704	icaR	Biofilm operon icaADBC HTH-type negative transcriptional re	5.241	3.43E-06	0.00030165					
HG001_01210	-	Putative 4-hydroxybenzoyl-CoA thioesterase	5.203	5.23E-06	0.00039405					
HG001_01280	ctpA	Carboxy-terminal processing protease	5.177	5.63E-06	0.00040955					
HG001_01264	-	5-bromo-4-chloroindolyl phosphate hydrolysis protein	5.152	1.38E-05	0.0008102		(-138 -111)			
HG001_01634	-	putative transporter	5.146	2.77E-05	0.00140562		71 - 95	42 - 71	27 - 52	
HG001_00401	holB	DNA polymerase III subunit delta	5.078	6.73E-05	0.00272727					
HG001_01953	pcnA	Isochorismatase	5.05	0.00012858	0.00428878					
HG001_01965	pmtC	Phenol-soluble modulin export ABC transporter, ATP-binding	5.042	0.00030684	0.00732354				363 - 401	31 - 70
HG001_02113	csoZ	putative copper chaperone	5.037	0.00010945	0.00384519					
HG001_00318	-	Transglycosylase associated protein	4.972	4.30E-06	0.00036539					
HG001_00657	ltaS	Lipoteichoic acid synthase	4.957	1.34E-05	0.00080047					
HG001_01452	accB_1	Biotin carboxyl carrier protein of acetyl-CoA carboxylase	4.949	5.75E-06	0.00040955					
HG001_00573	mntR	Manganese-dependent transcriptional regulator, DtxR family	4.943	0.00013059	0.00430136				274 - 285	39 - 49
HG001_01002	ecb	Extracellular complement-binding protein	4.933	0.00487195	0.04790146					
HG001_02205	sepA	Multidrug resistance efflux pump SepA	4.901	9.95E-05	0.00354214	< 300			216 - 227	99 - 109
HG001_01308	-	hypothetical protein	4.882	0.0002962	0.00722664					
HG001_00279	mepB	MepB protein, unknown function	4.875	0.00192453	0.02766728					
HG001_01461	gcvT	Aminomethyltransferase	4.835	0.0002793	0.00700905					
HG001_02517	-	hypothetical protein	4.758	0.00017097	0.00506187					
HG001_00934	ptsI	Phosphoenolpyruvate-protein phosphotransferase	4.722	8.30E-06	0.00054692				226 - 234	14 - 21
HG001_01925	gatD	Lipid II isoglutaminyl synthase (glutamine-hydrolyzing) subuni	4.715	0.0004689	0.00980588					
HG001_02174	czrB	Zinc efflux protein	4.681	9.07E-05	0.00331913					
HG001_01263	acyP	Acyolphosphatase	4.669	0.00018254	0.00522805					
HG001_01945	-	hypothetical protein	4.667	0.00051375	0.01057597					
HG001_02548	-	Inner membrane protein YqjA	4.488	0.00120603	0.01949631					
HG001_02518	-	putative oxidoreductase	4.444	0.00012715	0.00428878					
HG001_00251	-	putative Lipoprotein	4.414	2.34E-05	0.00125583					
HG001_01570	rpmA	50S ribosomal protein L27	4.4	0.00086801	0.01577378		38 - 50	135 - 147		
HG001_02641	-	putative amino acid permease	4.351	0.00082115	0.01513103		-41 5	129 - 170		
HG001_01967	pmtA	Phenol-soluble modulin export ABC transporter, ATP-binding	4.344	0.00358356	0.0385416				775 - 783	15 - 22
HG001_01963	-	hypothetical protein	4.32	3.80E-05	0.00177305		-161 -147	122 - 135		
HG001_02160	hmrA	Carboxydipeptidase	4.319	0.00093822	0.01648141					
HG001_02248	-	hypothetical protein	4.318	8.40E-05	0.0031189					
HG001_02587	adaB	Methylated-DNA--protein-cysteine methyltransferase	4.308	0.00015808	0.00492438					
HG001_01832	-	Nucleoside diphosphatase	4.271	0.00115894	0.0188507					
HG001_01456	-	hypothetical protein	4.241	0.00255805	0.0317946					
HG001_00575	tagA	Putative N-acetylmannosaminyltransferase	4.237	0.00011339	0.00393123				362 - 379	38 - 54
HG001_00568	nhaK_1	Putative sodium, potassium, lithium and rubidium/H(+) antipor	4.224	0.0006412	0.01233264					
HG001_02594	rocA	1-pyrroline-5-carboxylate dehydrogenase	4.177	7.98E-05	0.00309191					
HG001_02386	-	hypothetical protein	4.175	0.00011791	0.00403499					

HG001_02631	queH	Epoxyqueuosine reductase	4.14	0.00427914	0.04416956	81 - 99	36 - 50		
HG001_00886	tarM	Poly(ribitol-phosphate) alpha-N-acetylglucosaminyltransferas	4.119	0.00239934	0.03054229				
HG001_01271	-	Putative dioxygenase	4.104	0.00014668	0.0046567				
HG001_01657	ccpA	Catabolite control protein A	4.09	4.63E-05	0.00203138	< 150	-16 19	36 - 69	
HG001_00753	-	hypothetical protein	4.086	0.00105347	0.01790891		17 - 43	30 - 61	-29 +3 (SD)
HG001_02128	-	hypothetical protein	4.077	7.85E-05	0.00308918				
HG001_01137	recA	Protein RecA	4.07	4.83E-05	0.00208465				
HG001_01782	airS	Sensor histidine kinase	4.058	0.00108801	0.0182605				
HG001_01178	cls_1	Cardiolipin synthase	4.055	0.00020383	0.00553695				
HG001_02349	spdB	Surface protein display B	4.043	0.00193198	0.02766728				
HG001_02520	-	hypothetical protein	4.034	8.29E-05	0.0031189	no			
HG001_00571	mntB	Manganese transport system membrane protein MntB	3.932	0.00325059	0.03659811				
HG001_01216	glcT	GlcA/glcB genes antiterminator	3.928	0.00102723	0.01769115				
HG001_00576	tagH_1	Teichoic acids export ATP-binding protein TagH	3.875	0.00039307	0.00855989	> 600			
HG001_00581	pbp4	ribonuclease R	3.873	0.00030461	0.00732354				
HG001_00225	-	hypothetical protein	3.871	0.00200011	0.02812416				
HG001_01269	nirQ	Denitrification regulatory protein NirQ	3.864	0.00051929	0.01060721				
HG001_01338	ansA	putative L-asparaginase	3.835	0.00055774	0.01121862				
HG001_01669	trmB	tRNA (guanine-N(7)-)-methyltransferase	3.827	0.00029618	0.00722664				
HG001_00402	ricT	Regulatory iron-sulfur containing complex, subunit T	3.782	0.00056502	0.01127905				
HG001_01964	pmtD	Phenol-soluble modulins export ABC transporter , permease cc	3.772	0.00261416	0.03233946				
HG001_00627	-	Oxidoreductase	3.767	0.00018884	0.00529344				
HG001_00869	-	Putative manganese efflux protein	3.723	0.00363211	0.03890495				
HG001_01670	-	Putative manganese efflux protein	3.71	0.00123298	0.01981038				
HG001_00720	est_2	Carboxylesterase	3.698	0.0001643	0.00494177				
HG001_01182	-	Response regulator, NarL family	3.638	0.00036132	0.00806856				
HG001_01416	srrB	Sensor protein SrrB	3.637	0.00034344	0.00786923				
HG001_00656	-	Putative 5-oxoprolinase, subunit C	3.601	0.00028918	0.00718864				
HG001_02092	sigB	RNA polymerase sigma-B factor	3.578	0.00023733	0.00613096				
HG001_02158	-	hypothetical protein	3.548	0.00349954	0.03794773				
HG001_01954	ppaC	putative manganese-dependent inorganic pyrophosphatase	3.526	0.00494395	0.04840302	(-12 7)		29 - 50	
HG001_01558	dtb	D-tyrosyl-tRNA(Tyr) deacylase	3.517	0.00145233	0.02199362				
HG001_00487	dck	Deoxyadenosine/deoxycytidine kinase	3.494	0.00228416	0.02964913				
HG001_01313	nth	Endonuclease III	3.482	0.00279998	0.03366277				
HG001_02172	-	Putative NAD(P)-dependent epimerase	3.475	0.00060212	0.011666				
HG001_01780	-	Putative transcription factor, Xre family	3.449	0.00243055	0.03079087				
HG001_01840	-	Putative cysteine protease	3.409	0.00050394	0.01045586				
HG001_00131	-	Putative inner membrane protein	3.407	0.00322355	0.03657262				
HG001_01312	-	hypothetical protein	3.404	0.00426839	0.04416956				
HG001_01217	cozEB	Putative cell division protein	3.311	0.00138173	0.02116785				
HG001_00128	-	Putative inner membrane protein	3.286	0.00059516	0.01161671				
HG001_02303	sarV	HTH-type transcriptional regulator SarV	3.256	0.0020406	0.02812416	< 50			
HG001_01314	dnaD	DNA replication protein DnaD	3.24	0.00167392	0.0249856				
HG001_02547	-	Putative MFS transporter	3.237	0.00205077	0.02812416				
HG001_01001	-	Putative metallophosphoesterase	3.22	0.00513031	0.04951779				
HG001_01612	phoP	Alkaline phosphatase synthesis transcriptional regulatory pro	3.216	0.00464932	0.04623				
HG001_01784	fumC	Fumarate hydratase class II	3.192	0.0039306	0.0414285				
HG001_02382	fni	Isopentenyl-diphosphate delta-isomerase	3.165	0.0011101	0.01851337				
HG001_01603	dnaI	Primosomal protein DnaI	3.155	0.00108483	0.0182605				
HG001_00687	fakB1	Fatty acid kinase subunit	3.122	0.00179097	0.02621784	< 150			
HG001_02428	nreC	Oxygen regulatory protein NreC	3.106	0.00131331	0.02035637				
HG001_00721	rnv	Ribonuclease R	3.079	0.00115754	0.0188507				
HG001_01543	-	ATP-dependent RecD-like DNA helicase	3.067	0.00147349	0.02218656				
HG001_01112	cdsA	Phosphatidate cytidyltransferase	3.058	0.00291229	0.0345671	-9 13		121 - 147	
HG001_01946	-	putative membrane protein	3.012	0.00204828	0.02812416				
HG001_01489	glyS	Glycine--tRNA ligase	2.978	0.00323394	0.03657262				
HG001_01836	-	Putative ABC transporter, ATP-binding component	2.978	0.00225266	0.0295312				
HG001_02444	-	Type II toxin-antitoxin system, YoeB family toxin	2.978	0.00235841	0.03026283				
HG001_01912	-	putative acyl-CoA thioester hydrolase	2.956	0.00293705	0.03470456				
HG001_00463	rae1	Ribosome-associated A site endoribonuclease	2.945	0.0020751	0.02818493				
HG001_00424	divIC	Cell division protein DivIC	2.927	0.00343356	0.03752667				
HG001_01191	-	Putative sugar phosphatase	2.898	0.00277378	0.03366277				
HG001_00422	mazG	Nucleoside triphosphate pyrophosphohydrolase/pyrophosph	2.891	0.00302678	0.03519429				
HG001_01932	dagK	Diacylglycerol kinase	2.89	0.004308	0.04416956				
HG001_00540	-	Putative NAD(P)H-dependent FMN-containing oxidoreductas	2.889	0.00310868	0.03587386				
HG001_01593	-	hypothetical protein	2.885	0.00337599	0.03737708				
HG001_01588	hemX	Conserved membrane protein	2.785	0.0047126	0.04668312				
HG001_01031	pbp1	Penicillin-binding protein 2B	2.778	0.00352136	0.03802776	-47 -0		33 - 95	640 - 653
HG001_01268	-	Putative nitric oxide reductase activation protein	2.776	0.00327786	0.03659811				133 - 145
HG001_00140	mupG	6-phospho-N-acetylmuramidase	2.739	0.00429238	0.04416956				
HG001_01587	hemC	Prophobilinogen deaminase	2.681	0.00462524	0.04616482				
								598 - 606	105 - 112

Table S2: Comparative transcriptomic analysis. The data were obtained from RNA samples prepared from the wild-type strain HG001, and the mutant HG001 Δ *rsaG* strain and from the same mutant strain transformed with a plasmid expressing RsaG under the control of P3 promoter (HG001 Δ *rsaG* pCN51::P3::*rsaG*). These experiments were carried out in duplicates. The comparison of the data were done with DSeq2: RNAs with p-adj < 0.05 differentially expressed from HG001 Δ *rsaG* and HG001 are listed in tab “DOWN_deltavsWT p-adj” and “UP_deltavsWT p-adj”, RNAs with p-adj < 0.05 differentially expressed between HG001 Δ *rsaG* and HG001 Δ *rsaG* pCN51::P3::*rsaG* are listed in tab “DOWN_deltavsComp p-adj” and “UP_deltavsComp p-adj”, and RNAs with p-adj < 0.05 differentially expressed between HG001 Δ *rsaG* pCN51::P3::*rsaG* and HG001 are listed in tab “DOWN_CompvsWT p-adj” and “UP_CompvsWT p-adj”.

DOWN_deltavsComp p-adj

id	gene	product	FoldChange	pvalue	padj
HG001_02756	-	RsaG	0	1.0322E-98	2.8356E-95
HG001_00732	nuc	Thermonuclease precursor	0.215	5.6502E-06	0.00220481
HG001_01298	ald1	Alanine dehydrogenase 1	0.324	1.0742E-08	7.3768E-06
HG001_01296	steT	Serine/threonine exchanger SteT	0.327	4.3006E-11	5.9069E-08
HG001_00868	-	yybP-ykoY	0.331	2.1551E-05	0.00591998
HG001_01295	norB_4	Quinolone resistance protein NorB	0.332	6.1106E-09	5.5952E-06
HG001_01297	tdcB	L-threonine dehydratase catabolic TdcB	0.353	2.0006E-06	0.00091595
HG001_01976	chp	Chemotaxis inhibitory protein precursor	0.418	0.00013775	0.02899197
HG001_01150	glpF	Glycerol uptake facilitator protein	0.453	6.421E-06	0.00220481
HG001_00082	sodM	Superoxide dismutase [Mn/Fe] 2	0.474	1.1341E-07	6.2305E-05
HG001_00502	gph_1	Phosphoglycolate phosphatase	0.475	1.4447E-05	0.00440964
HG001_00362	-	hypothetical protein	0.563	0.00017883	0.03275057
HG001_02690	asp2	Accessory Sec system protein Asp2	0.61	0.00021756	0.03735302

UP_deltavsComp p-adj

HG001_00730	ssp	Extracellular matrix protein-binding protein emp precursor	3.409	7.4037E-05	0.01848908
HG001_01548	-	hypothetical protein	1.902	0.00012977	0.02899197
HG001_00492	sdrC	Serine-aspartate repeat-containing protein C precursor	1.737	0.00014776	0.02899197

DOWN_deltavsWT p-adj

HG001_02756	-	RsaG	0.026	5.5635E-23	1.5283E-19
-------------	---	------	-------	------------	------------

UP_deltavsWT p-adj

HG001_01499	rpsU	30S ribosomal protein S21	2.096	1.1392E-07	0.00015647
HG001_00389	-	Archaea_SRP	1.816	4.6728E-06	0.00427877
HG001_01824	-	tRNA-Ile(gat)	2.044	2.2091E-05	0.01517127
HG001_00819	-	hypothetical protein	1.627	6.079E-05	0.03339798
HG001_01092	rplS	50S ribosomal protein L19	2.155	8.9433E-05	0.04094524
HG001_01475	rpmG2	50S ribosomal protein L33 2	2.257	0.00012529	0.0491692

DOWN_CompsvWT p-adj

HG001_02643	-	hypothetical protein	0.289	2.2002E-05	0.0024688
HG001_01549	-	CsbD-like protein	0.371	1.1161E-07	3.76E-05
HG001_00492	sdrC	Serine-aspartate repeat-containing protein C precursor	0.389	7.9526E-11	5.35E-08
HG001_01548	-	hypothetical protein	0.395	3.0087E-08	1.16E-05
HG001_02222	-	Alkaline shock protein 23	0.4	6.6997E-05	0.00540452
HG001_02223	-	hypothetical protein	0.436	0.00018766	0.01142216
HG001_00932	-	hypothetical protein	0.456	1.9394E-07	5.80E-05
HG001_00664	-	Putative lipid kinase	0.487	0.00053405	0.02437622
HG001_00728	clfA	Clumping factor A precursor	0.5	5.8393E-05	0.00524174
HG001_02418	ydaG	General stress protein 26	0.531	0.00017121	0.01092815
HG001_00741	-	hypothetical protein	0.537	0.00103141	0.0025481
HG001_02186	mtlD	Mannitol-1-phosphate 5-dehydrogenase	0.546	1.1514E-05	0.00141915
HG001_01659	-	hypothetical protein	0.549	6.8234E-05	0.00540452
HG001_00252	-	FtsX-like permease family protein	0.551	7.5868E-05	0.00567536
HG001_02185	mtlF	Mannitol-specific phosphotransferase enzyme IIA componen	0.559	4.9E-05	0.00488728
HG001_00125	-	hypothetical protein	0.574	5.7956E-05	0.00524174
HG001_01787	nagB_2	Glucosamine-6-phosphate deaminase	0.574	0.00025832	0.01480093
HG001_00415	spoVG	Putative septation protein SpoVG	0.621	0.0003586	0.01788346
HG001_00759	-	CsbD-like protein	0.629	0.00123869	0.04633032
HG001_01574	mreC	Cell shape-determining protein MreC precursor	0.673	0.00026613	0.01493113
HG001_00121	-	hypothetical protein	0.676	0.00030587	0.01615135

UP_CompsvWT p-adj

id	gene	product	FoldChange	pvalue	padj
HG001_02756	-	RsaG	58.448	6.8695E-83	1.85E-79
HG001_00822	ampR	HTH-type transcriptional activator AmpR	3.466	0.00035471	0.01788346
HG001_00508	-	hypothetical protein	3.151	8.9514E-06	0.00126874
HG001_02347	-	RsaOG	3.047	0.00114766	0.04353029
HG001_02666	-	hypothetical protein	2.771	0.00030476	0.01615135
HG001_02210	-	hypothetical protein	2.765	0.00013051	0.00878659
HG001_01763	-	tRNA-Met(cat)	2.714	1.0771E-05	0.00141915
HG001_01499	rpsU	30S ribosomal protein S21	2.51	3.7299E-11	3.35E-08
HG001_01293	rnhA	14,7 kDa ribonuclease H-like protein	2.509	0.00053193	0.02437622
HG001_00774	-	hypothetical protein	2.498	9.1123E-05	0.00663229
HG001_01295	norB_4	Quinolone resistance protein NorB	2.413	3.1033E-06	0.00052232
HG001_01297	tdcB	L-threonine dehydratase catabolic TdcB	2.38	7.4434E-05	0.00567536
HG001_01298	ald1	Alanine dehydrogenase 1	2.372	1.1594E-05	0.00141915
HG001_00569	-	RsaC	2.364	0.00127829	0.0471567
HG001_01296	steT	Serine/threonine exchanger SteT	2.345	4.771E-07	0.00012554
HG001_01345	-	FMN	2.342	6.8862E-07	0.00015454
HG001_01065	-	TM2 domain protein	2.332	3.0997E-06	0.00052232
HG001_02363	-	SAM	2.326	5.6278E-06	0.00084199
HG001_00887	catD	Putative oxidoreductase CatD	2.302	0.00037992	0.01860221
HG001_00164	-	Isoprenylcysteine carboxyl methyltransferase (ICMT) family protein	2.298	0.00014078	0.00924706
HG001_00389	-	Archaea_SRP	2.285	2.2614E-10	1.22E-07
HG001_01976	chp	Chemotaxis inhibitory protein precursor	2.275	0.00032688	0.01692854
HG001_00648	-	hypothetical protein	2.273	0.00140649	0.04919079
HG001_00082	sodM	Superoxide dismutase [Mn/Fe] 2	2.239	1.002E-08	4.50E-06
HG001_00819	-	hypothetical protein	2.229	2.8104E-11	3.35E-08
HG001_02176	-	rli28	2.222	0.0002276	0.01332469
HG001_00502	gph_1	Phosphoglycolate phosphatase	2.21	3.5845E-06	0.00056782
HG001_01075	rpmB	50S ribosomal protein L28	2.192	2.8879E-05	0.00311083
HG001_01186	-	hypothetical protein	2.102	0.00010053	0.00696466
HG001_01510	rpsT	30S ribosomal protein S20	2.092	0.00028517	0.01567275
HG001_01092	rplS	50S ribosomal protein L19	2.077	0.00019086	0.01142216
HG001_00723	ssrA	transfer-messenger RNA%2C SsrA	2.075	3.6296E-05	0.00375941
HG001_02640	ldh2	L-lactate dehydrogenase 2	2.064	1.882E-05	0.00220356
HG001_02338	ssaA2_3	Staphylococcal secretory antigen ssaA2 precursor	2.025	6.2535E-05	0.00540452
HG001_01150	glpF	Glycerol uptake facilitator protein	2.013	6.4652E-05	0.00540452
HG001_00954	-	hypothetical protein	2.001	5.1279E-07	0.00012554
HG001_02754	rpmH	50S ribosomal protein L34	1.903	0.00099876	0.04014404
HG001_02440	nirC	Nitrite transporter NirC	1.888	0.00095783	0.03908224
HG001_00362	-	hypothetical protein	1.855	5.3588E-05	0.00515399
HG001_00828	-	hypothetical protein	1.842	0.00051311	0.0242424
HG001_00879	-	CAAX amino terminal protease self-immunity	1.84	0.00046578	0.02239911
HG001_01208	opuD_1	Glycine betaine transporter OpuD	1.838	0.00090704	0.03757935
HG001_01262	-	hypothetical protein	1.809	0.00010086	0.00696466
HG001_00254	-	hypothetical protein	1.801	0.0006009	0.02697028
HG001_00441	-	tRNA-Arg(agg)	1.801	0.00087815	0.03695104
HG001_01824	-	tRNA-Ile(gat)	1.78	0.00062967	0.02779851
HG001_01768	traP	Signal transduction protein TRAP	1.739	1.1583E-06	0.00023995
HG001_01556	-	T-box	1.737	3.1003E-06	0.00052232
HG001_00802	yugl_2	General stress protein 13	1.734	0.00132034	0.04740903
HG001_02582	-	hypothetical protein	1.703	0.00131748	0.04740903
HG001_00370	-	Sodium:neurotransmitter symporter family protein	1.702	0.00102501	0.04025481
HG001_02732	yfiS	Putative malate transporter YfiS	1.684	0.0006471	0.02810689
HG001_02454	gpmA_2	2%2C3-bisphosphoglycerate-dependent phosphoglycerate mutase	1.665	0.00079735	0.03408358
HG001_00719	-	preprotein translocase subunit SecE	1.581	0.00017449	0.01092815
HG001_00772	lipA_1	Lipoyl synthase	1.524	0.00111384	0.042851
HG001_01647	htrA	Serine protease Do-like HtrA	1.447	0.00139396	0.04919079

Table S3: Strains and plasmids used in the study

Description		Reference
Bacterial strains		
<i>E. coli</i> strains		
DC10B	<i>dcm</i> mutant ; for direct transformation in <i>S. aureus</i>	Monk <i>et al.</i> , 2012
DH5α	Mutant <i>endA1</i> , <i>hsdR17</i> , <i>recA</i>	Taylor <i>et al.</i> , 1993
<i>S. aureus</i> strains		
HG001	Derivative of 8325-4, <i>rsbU</i> restored RN1, <i>agr</i> positive	Herbert <i>et al.</i> , 2010
RN4220	Cloning intermediate	Nair <i>et al.</i> , 2011
HG001Δ <i>hptRS</i>	In frame deletion of <i>hptRS</i> in HG001	Villanueva <i>et al.</i> , 2018
HG001Δ <i>rnc</i>	Deletion of <i>rnc</i> (RNase III)	Lalaouna <i>et al.</i> , 2019
HG001Δ <i>P_{uhpT}</i>	Deletion of the promoteur sequence of <i>uhpT</i> including the HptR BOX	This study
PR02	RN4220Δ <i>pyrFE</i>	Linder <i>et al.</i> , 2014
PR02-06	RN4220Δ <i>j1::ermC</i>	Linder <i>et al.</i> , 2014
PR01-01	Δ <i>rnjA</i> (SA0940)	redder <i>et al.</i> , 2012
PR01-04	Δ <i>rnjB</i> (SA1118)	redder <i>et al.</i> , 2012
HG001Δ <i>rsaG</i>	Δ <i>rsaG</i>	This study
<i>S. argenteus</i> strains		
MSHR11	Strain isolated from blood culture of a 55 year old indigenous Australian female	Hoit <i>et al.</i> , 2011
<i>S. epidermidis</i> strains		
ATCC35983	Strain isolated from human blood Tennessee, United States	Christensen <i>et al.</i> , 1982
<i>S. simiae</i> strains		
CCM7213T	Strain isolated from the faeces of a South American squirrel monkey	Suzuki <i>et al.</i> , 2012
<i>S. schweitzeri</i> strains		
FSA084	Strain isolated from the nares of a red-tailed monkey, Gabon	Suzuki <i>et al.</i> , 2012
Host cell lines		
CTi 400	Human myoblasts cell line	Baude <i>et al.</i> , 2019
RAW 264.7	Murin macrophages cell line	Ralph <i>et al.</i> , 1977
A549	Human tumorigenic lung epithelial cell line	Giard <i>et al.</i> , 1973
HU-H7	Human tumorigenic liver cell line	Nakabayashi <i>et al.</i> , 1982
HT-29MTX	Human colon adenocarcinoma cells line differentiated into mature goblet cells using methotrexate	Behrens <i>et al.</i> , 2001
HT-29	Human colon adenocarcinoma cells line	von Kleist <i>et al.</i> , 1975
Plasmids		
<i>in vitro</i> transcription		
pUC18-T7:: <i>rsaG</i>	Plasmid used for T7 transcription of <i>RsaG</i>	Geissmann <i>et al.</i> , 2009
pUC18-T7:: <i>rsaG</i> mut1	Plasmid used for T7 transcription of <i>RsaG</i> mut1	This study
pUC18-T7:: <i>rsaG</i> mut2	Plasmid used for T7 transcription of <i>RsaG</i> mut2	
pUC18-T7:: <i>rsaG</i> mut1/2	Plasmid used for T7 transcription of <i>RsaG</i> mut1/2	
pUC18-T7:: <i>rsal</i>	Plasmid used for T7 transcription of <i>Rsal</i>	Bronesky <i>et al.</i> , 2019
Transcriptional fusion with GFP		
pEW::GFP	Plasmid with promoterless GFP	Menendez-Gil <i>et al.</i> , 2020
pCN51-100 <i>uhpT</i> -GFP	Plasmids expressing GFP under different DNA inserts	This study
pCN51-250 <i>uhpT</i> -GFP		
pCN51-500 <i>uhpT</i> -GFP		
pCN51-CDS <i>uhpT</i> -GFP		
pCN51-FL <i>uhpT</i> -GFP		
Translational fusion with LacZ		
pLUG220:: <i>PrpoB::lacZ</i>	Plasmid to clone 5'UTs of mRNA in front of <i>lacZ</i> under the <i>rpoB</i> promoter	Romilly <i>et al.</i> , 2014
pLUG220:: <i>PrpoB::rex::lacZ</i>	Plasmid used to transform HG001 with translational fusion <i>glcU::lacZ</i>	This study
pLUG220:: <i>PrpoB::rex::lacZ::blaZ-rsaG</i>	Plasmid used to transform HG001 with translational fusion <i>glcU::lacZ</i> and expressing <i>RsaG</i> under <i>blaZ</i> promoter	
pCN40	Plasmid used to amplify the <i>blaZ</i> promoter	
MAPS		
pCN51::P3	pCN51 with P3 promoter	Tomasini <i>et al.</i> , 2017
pCN51::P3::MS2	pCN51 expressing MS2 tag under P3	
pCN51::P3::MS2- <i>rsaG</i>	pCN51 expressing MS2- <i>RsaG</i> under P3 promoter	This study
<i>RsaG</i> complementation plasmid		
pCN51::P3	pCN51 with P3 promoter	Tomasini <i>et al.</i> , 2017
pCN51::P3:: <i>rsaG</i>	pCN51 expressing <i>rsaG</i> under P3	This study
In frame deletion		
pMAD	Thermosensitive origine of replication, constitutively expressed <i>bgal</i> gene	Arnaud <i>et al.</i> , 2004
pMAD:: <i>rsaG</i>	Plasmid to delete <i>rsaG</i>	This study
pMAD-P <i>uhpT</i>	Plasmid to delete the <i>uhpT</i> promotor including the HotR box	

Table S4: List of oligonucleotides used in the study

Name	Sequence	
DIG probes		
RsaG T7 rev	TAATACGACTCACTATAGGG GACACCGCTATATGTC	
RsaG for	TGCCCCGACGGCATGTGCGA	
5S T7 rev	TAATACGACTCACTATAGGG GATTTGTCAATTGCGCTGGC	
5S for	GTAAGTTATTTTGTCTGGTGGCTATAGC	
rex T7 rev	TAATACGACTCACTATAGGG GCTGTGCAACTCTTTC	
rex for	GATTTCTTTAAATCTGAACCTAAGCG	
ldh 1 T7 rev	TAATACGACTCACTATAGG GAATTTCCATGTTGCATACG	
ldh 1 for	GGATCTAGTTGTCACTGTGCTG	
RsaI T7 rev	TAATACGACTCACTATAGGG GACGTGCTAGCCGACAAATA	
RsaI for	TAACAGGGGGAGCGATTAAA	
fn3k T7 rev	TAATACGACTCACTATAGGG TGATTATCAAAAGAAATATCGCCA	
fn3k for	TCCAACGTGGACGTAAGA	
3'UTR uhpT <i>S. epidermidis</i> T7 rev	TAATACGACTCACTATAGGG TGCAATCAATCACGAAAAAG	
3'UTR uhpT <i>S. epidermidis</i> for	GATTAAGTGAATTTAAATATATACTCCC	
3'UTR uhpT <i>S. simiae</i> T7 rev	TAATACGACTCACTATAGGG CCCCCAATTAGTGGG	
3'UTR uhpT <i>S. simiae</i> for	CGGGCATAGTTGGGC	
<i>in vitro</i> transcription		Localisation PCR product from AT
tcaR T7 for	TAATACGACTCACTATAGGG TATGGTCAATATTTTGTG	(-74/+226)
tcaR rev	TTAAGTTTAAATTCGTCGG	
sarA T7 for	TAATACGACTCACTATAGGG TAAATAATAGAGCATTAAAT	(-146/+145)
sarA rev	TCTCTTTGTTTTCGCTGATG	
ccpA T7 for	TAATACGACTCACTATAGGG CGTTTACAAGGAGGA	(-64/+515)
ccpA rev	TGTTCTCCACCTACTAAAGC	
rex T7 for	TAATACGACTCACTATAGGG TCAATTGTAATCAATTATTTG	(-103/+205)
rex rev	CTATATTATATCCGTACCT	
fn3K SD T7 for	TAATACGACTCACTATAGGG TGACGCAAGTTTGAATAACG	(-88/+81)
fn3K SD rev	TTTCGTTTACATCACCACCA	
fn3K FL T7 for	TAATACGACTCACTATAGGG AAGATTAAATCATCCTAATTAGGCAA	(-64/+867)
fn3K FL rev	CTAACTTGTGTATCTTGTAAATCTT	
ldh 1 T7 for	TAATACGACTCACTATAGGG TGAATAAATCACAACTTAAAG	(-166/+418)
ldh 1 rev	GTAACCCAGAGAATTTCCATGTTGC	
tRNA ^{tyr} T7 for	GATTACATAATACGACTCACTATAGG GGGAGGGGTAGCGAAGTGGC	
tRNA ^{tyr} rev	TGGTGGAGGGGGGCGAGATTC	
pre-tRNA ^{tyr} T7 for	GATTACATAATACGACTCACTATAGG GCACCATTTATGGAGGGGTAGCG	
RsaG short rev	AAAACAATACTCCCTACAATATTG	
uhpT short T7 for	GATTACATAATACGACTCACTATAGG GAATTGGTGTATCATTGACTG	
PE		
RsaG-like for	CGAATGCGACCACTG	
RsaG-like rev	CTATGCCCGTCTTGC	
Reverse transcription for Toe-print		
tcaR rev	GTTCTTTATTTAACA	
1146-sarA rev	GTCAGCATAAGTGAC	
rex rev	TGAATTTACAGATC	
ccpA rev	CTGCTTTAACATTTTGG	
Transcriptional fusion with GFP		
uhpT FL SphI for	GCATGCATACTATTAACTATTATTAATTTAG	
uhpT FL BamHI rev	GGATCCGATTGCCAAGTATAT	
uhpT CDS SphI for	GCATGCAACGAAAATGATTGAG	
1119-uhpT CDS+956 SphI for	GCATGCTATGGGGCTACGTAT	
uhpT CDS+1206 SphI for	GCATGCATTCGGTGACTCAAT	
uhpT CDS+1356 SphI for	GCATGCAATCGTTGCTTTCTATG	
MAPS		
PstI-MS2_RsaG for	<u>CGCCTGCAG</u> CGTACACCATCAGGGTACGTTTTTTCAGACACCATCAGGG	
RsaG-BamHI rev	<u>TCTGTTTCCCGACGGCATGTGCGTGAA</u>	
Translational fusion with LacZ	<u>CGCGGATCCAAAACAATACTCCCTACAATA</u>	
PstI-pblaZ/pES for	CGCCTGCAGTGCAGCTTACTATGCCATTA	
RsaG5'-pblaZ/pES3' rev	TGCCGTGCGGGCAGTGCAGAAATAAC	
RsaG for	CTGCCCCGACGGCATGTGC	
PstI-RsaG rev	CGCCTGCAGAAAACAATACTCCCTA	
rex for BamHI	CGCGGATCCTCATTGTAATCAATTATTTG	
rex rev BamHI	CGCGGATCCTCTATATTATATCCGTACCC	
q-RT-PCR		
RsaG for	ATGTGCGTGAAGAGATGAAAGA	
RsaG rev	CTAAGTCGGGCAAAATAAGGATAC	
gyrB for	GGTGCGGACTTTGATCTAGC	
gyrB rev	TTATACAACGGTGGCTGTGC	
hu for	TTTACGTGCAGCACGTTTAC	
hu rev	AAAAAGAAGCTGTTTCTAGCAGTAG	
16S for	CGTGAGGGTCAATTGGA	
16S rev	CGTTTACGGCGTGGACTA	
Cloning RsaG in pCN51		
BamHI RsaG rev	CGCGGATCCAAAACAATACTCCCTACAATA	
PstI RsaG for	CGCCTGCAGCTGCCCGACGGCATGTGC	
SphI P3 for	TTTGATGCATACGTGGCAAACT	
PstI P3 rev	CGCCTGCAGAGTTATATTAACATGC	
In frame deletion		
pMAD rsaG-hptR for downstream EcoRI	CGCGAATTCTATATTTAAACTCTCTAAAAAG	
pMAD rsaG-hptR rev downstream SalI	CGCGTCGACCACTTTCAAGGGGGAA	
pMAD rsaG-uhpT for upstream SmaI	CGCCCCGGGATTGAGGTGACAG	
pMAD rsaG-uhpT rev upstream EcoRI	CGGAATTCGATTGCCAATATATTGC	
pMAD PuhpT-uhpT for downstream BamHI	GGATCCTAACGAAAATGATTGAGGTGAC	
pMAD PuhpT-uhpT rev downstream EcoRI	GAAATCTCTGTAATCGCAGATAAGATAAG	
pMAD PuhpT-HG001 00164 for upstream BamHI	GGATCCCATTTTGCAAAAGCTAAATGTATCTT	
pMAD PuhpT-HG001 00164 rev upstream AscI	GGCGCGCCTTGACAAGCTTTATCTATAAAATACT	
Quickchange RsaG		
RsaG rev mut CCC 2 substitution	GCAAAATAAGGATACTTCATAACACCCAATTATGATCTGCCCGG	
RsaG for mut CCC 2 substitution	CGGGGGCAGATCATAATTTGGGTGTTATGAAGTATCCTTATTTGC	
RsaG rev mut CCC 1 substitution	CGAGACATAGAGATGATATATTTGCAACCCCTAGAAGCAGTATCTTTC	
RsaG for mut CCC 1 substitution	GAAAGATACTGCTTCTAGGGTTGCAAAATATATCATCTCTATGTCTCG	
RT-PCR		
uhpT for	GAGGAAATTGGATTATCTACATTAGA	
uhpT rev	TTTTTAAACTTCTAATTTCTTTTC	

Restriction sites sequences are underlined, the T7 promoter sequence is in bold. The MS2 sequence is highlighted in light grey

3. Chapitre 3 : Mise au point de la technique DBRI « Direct determination of bacterial RNA-RNA interactome »

Cette technique dérivée du LigR-Seq (Sharma et al., 2016, Durand et al., 2019 (bioRxiv)) a pour but de donner une vision globale de l'interactome cellulaire ARN-ARN en fonction des conditions de croissance choisies, sans l'intervention d'un ARN ou d'une protéine étiquetée. Tout d'abord, les duplexes ARN-ARN sont stabilisés *in vivo* par un pontage aux UV en présence d'un dérivé du psoralène. Suite à la précipitation des ARN, les ARN non pontés par le psoralène ou un de ces dérivés sont éliminés par digestion enzymatique et/ou par chromatographie d'affinité. La digestion des extrémités non appariées, permet de restreindre les fragments d'ARN aux régions d'interaction. Les ARN pontés subissent ensuite une ligation de proximité qui permettra après réversion du pontage de séquencer des ARN chimériques pour définir les ARN qui étaient en interaction *in vivo*, ainsi que les nucléotides engagés dans les interactions.

Dans ce chapitre, je présente un descriptif détaillé des méthodes utilisées à chaque étape (variante n° 1 à 7) du développement de la technique. Parfois, plusieurs conditions ont été testées pour une même étape. La partie « Résultats » est présentée de manière chronologique et récapitule les conditions expérimentales utilisées pour chaque variante du protocole (n° 1 à 7), les résultats obtenus, et les conclusions qui ont amené aux modifications successives. La figure 19 résume les différents protocoles testés au cours du processus d'optimisation et la figure 20A représente le protocole de manière simplifiée.

3.1. Matériel et méthodes

3.1.1. Culture bactérienne

Une pré-culture de *S. aureus* HG001 sur la nuit à 37°C sous agitation en milieu BHI, est diluée pour ensemer 50 mL de BHI frais à $DO_{600nm} = 0,05$. La culture est maintenue à 37°C sous agitation pendant 6 h jusqu'à une $DO_{600nm} \sim 3$.

3.1.2. Mise au point du protocole : variantes 1-4

3.1.2.1. Traitement au psoralène et pontage par irradiation

La culture bactérienne est répartie dans des tubes Falcon 15 mL contenant chacun 5 mL de culture et centrifugée 15 minutes à 3750 g à 4°C (le nombre de tubes préparés est adapté à chaque variante en fonction de la concentration de 4'-Aminomethyltrioxsalen hydrochloride

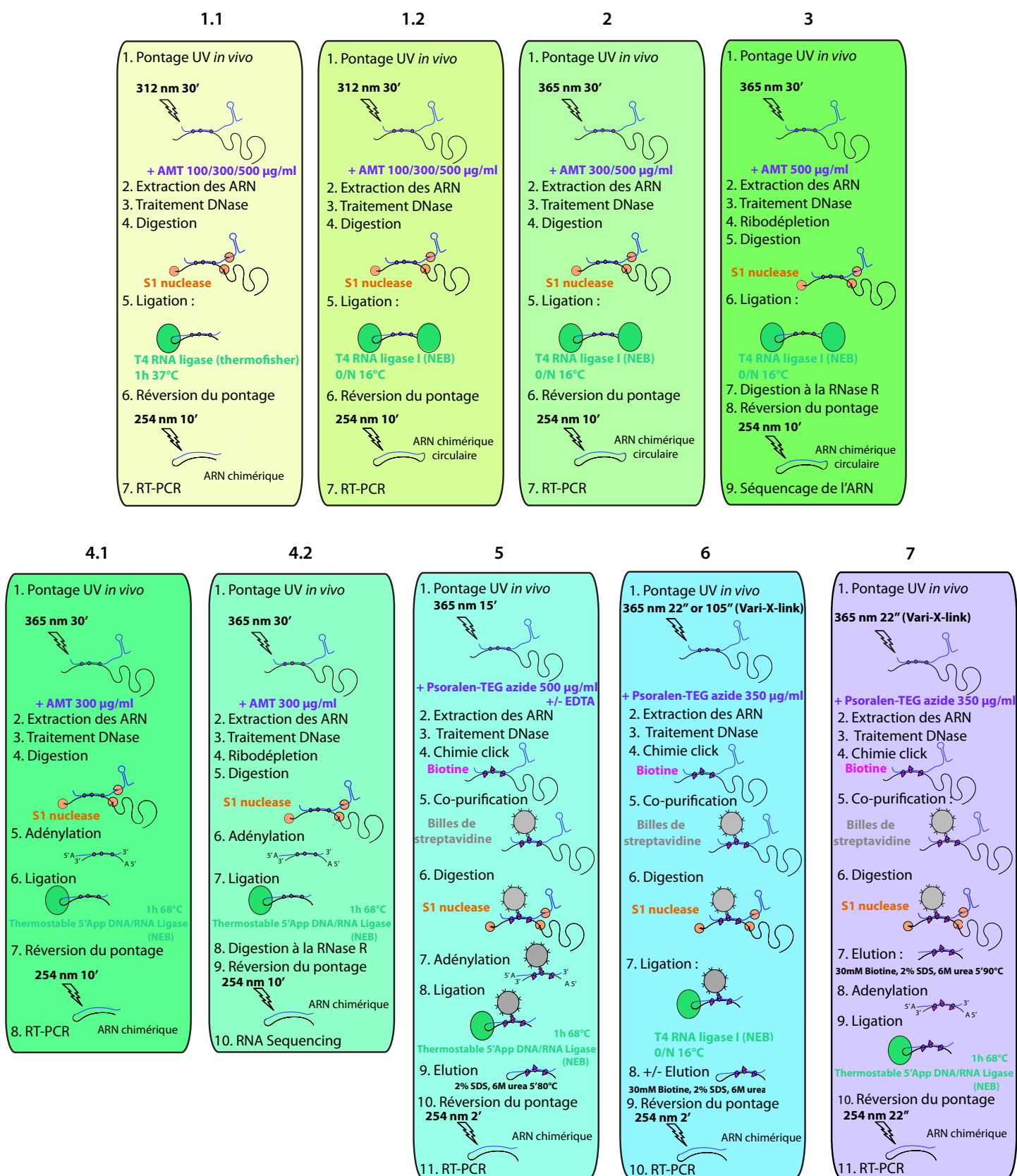


Figure 19 : Étapes chronologiques de l'optimisation de DBRI « *Direct Determination of Bacterial RNA-RNA interacom* ».

(AMT) testé). Les culots bactériens sont lavés une fois avec 5 mL de PBS 1X pH 7,4, centrifugés (15 minutes à 3750 g à 4°C), et sont ensuite resuspendus avec du PBS 1X pH 7,4, à 5.10^9 bactéries/mL. Différentes concentrations (100 µg/mL, 300 µg/mL et 500 µg/mL (variante 1), ou 300 µg/mL et 500 µg/mL (variante 2), ou 500 µg/mL (variante 3), ou 300 µg/mL (variante 4) (Figure 19)) d'AMT (Sigma, un dérivé du psoralène) sont ajoutées aux bactéries. Une condition contrôle sans AMT est également effectuée dans un des tubes. Les bactéries en présence ou en absence d'AMT sont maintenues sur glace pendant 10 minutes.

Pour l'irradiation, les bactéries en présence d'AMT sont réparties dans des boîtes de Pétri de 60 mm et soumises à un traitement UV à 312 nm (Biolink BLX Fisher Scientific, variante 1) ou 365 nm (Stratalinker 1800 Stratagene, variante 2 à 4) pendant 30 minutes sur glace à environ 15 cm de la source. Les bactéries sont transférées dans des tubes Falcon 15 mL et centrifugées 10 minutes à 3750 g à 4°C.

3.1.2.2. Extraction des ARN

Les ARN totaux sont extraits avec le kit « *Fast RNA Pro blue* » (MP Biomedicals) en suivant le protocole commercial, les tubes sont maintenus sur la glace pendant toute la procédure, afin de préserver la qualité du matériel biologique. Brièvement, les culots sont repris dans 1 mL de solution « *RNA Pro Solution* » (MP Biomedicals) et transférés dans des tubes contenant des billes de verre qui permettent la lyse mécanique des cellules. Les tubes sont placés dans l'appareil « *Fastprep* » (MP Biomedicals), agités pendant 40 secondes et centrifugés 15700 g pendant 15 minutes à 4°C. Les surnageants sont récupérés et transférés dans des tubes Eppendorf. Les ARN sont purifiés avec 300 µL de phénol/chloroforme puis avec 300 µL de chloroforme : alcool isoamylique (19 :1 vol). Finalement, la phase aqueuse contenant les ARN est précipitée par l'ajout de 3 volumes d'éthanol 100% pendant une nuit à -20°C. Après centrifugation (15700 g, 15 minutes, 4°C) les culots d'ARN sont lavés à l'éthanol 80%, séchés et repris dans 50 µL d'eau milli-Q (Merck).

3.1.2.3. Traitement DNase

10 µg d'ARN sont incubés avec 2 U de Turbo DNase (Invitrogen) et 5 µL de tampon 10X dans un volume final de 50 µL pendant 30 minutes à 37°C. Les ARN sont purifiés avec un volume de PCI (phénol, chloroforme, alcool isoamylique 25 :24 :1 pH 4,5-5, ROTH) et précipités par l'ajout de 3 volumes d'éthanol 100% et 1/10 de volumes d'AcNa 3 M, pendant 2 heures à -20°C. Après centrifugation (15700 g, 15 minutes, 4°C) les culots d'ARN sont lavés à l'éthanol 80% puis séchés. Les culots sont repris dans 10 µL d'eau milli-Q (Merck).

3.1.2.4. Ribodéplétion

Cette étape a été réalisée seulement pour les variantes 3 et 4.2 (figure 19). Les ARN totaux ont été traités avec le kit « *Ribo-zero* rRNA removal kit (Gram-positive *bacteria*) » (Illumina) en suivant le protocole. Pour cette expérience, 5 µg d'ARN, correspondant à la quantité limite supérieure du kit sont traités. Après précipitation, les culots d'ARN ribodéplétés sont repris dans 10 µL d'eau milli-Q (Merck).

3.1.2.5. Digestion à la nucléase S1

3.1.2.5.1. Digestion à la nucléase S1 des transcrits *in vivo*

La digestion effectuée avec 1 µg d'ARN est incubée avec 10 U de S1 nucléase (ThermoFischer) et 4 µL de tampon 5X dans un volume final de 20 µL pendant 30 minutes à 20°C. Les ARN sont purifiés avec un volume de PCI (phénol, chloroforme, alcool isoamylique 25 :24 :1 pH 4,5-5, ROTH) et précipités avec 3 volumes d'éthanol 100% et 1/10 de volumes d'AcNa 3 M, pendant 2 heures à -20°C. Après centrifugation (15700 g, 15 minutes, 4°C), les culots d'ARN sont lavés à l'éthanol 80%, séchés et repris dans 10 µL d'eau milli-Q (Merck).

3.1.2.5.2. Digestion à la nucléase S1 des transcrits *in vitro*

3.1.2.5.2.1. Transcription *in vitro* de *tcaR* et *RsaG*

Les transcriptions *in vitro* sont réalisées sur 10 µg de plasmide linéarisé exprimant *rsaG* sous le contrôle du promoteur T7 et sur un produit de PCR pour *tcaR* généré avec les amorces *tcaR* for et *tcaR* rev (*cf* IV.1). Le milieu réactionnel contient 40 µL de tampon T7 ARN polymérase 10X (0,4 M Tris-HCl pH 8, 150 mM MgCl₂, 0,5 M NaCl), chacun des quatre NTP à 4 mM, de la spermidine à 1 mM, du DTT à 5 mM, de la BSA à 0,1 mg/mL, 100 U de RNasin (Promega) et de la T7 ARN polymérase (40 U, produite au laboratoire) dans un volume réactionnel de 400 µL. Après une nuit d'incubation à 37°C, 10 U de DNase I et 40 µL de tampon DNase I 10X sont ajoutés et le mélange est incubé 15 minutes à 37°C afin d'éliminer la matrice d'ADN. La réaction est stoppée par l'ajout de 200 µL d'EDTA (0,1 M). Après extraction au PCI, les ARN sont précipités à l'éthanol, lavés puis séchés par évaporation sous vide. Le culot est repris dans 50 µL d'eau milliQ puis l'intégrité des transcrits est vérifiée sur mini-gel de polyacrylamide 8%, urée 8 M, TBE 1X polymérisé en présence de persulfate d'ammonium (APS) 0,1% et de TEMED 1/1000. Une aliquote contenant 500 ng d'ARN (dilué dans 5 µL d'eau milli-Q, Merck) en présence de 5 µL de tampon bleu urée 2X est chauffée 5 minutes à 70°C, avant d'être chargée sur le gel.

3.1.2.5.2.2. Digestion à la nucléase S1 des transcrit *in vitro*

Chaque ARN transcrit (30 µM) *in vitro* est mélangé puis dénaturé 1 minute à 95°C et une minute dans la glace, puis incubé avec 10 U de S1 nucléase (ThermoFisher) et 6 µL de tampon 5X dans un volume final de 30 µL pendant 30 minutes à 20°C permettant ainsi la digestion des ARN renaturés et appariés entre eux. Les ARN dilués dans un volume final de 100 µL sont extraits par l'ajout d'un volume de PCI. Après vortex et centrifugation (15700g, 15 minutes, 4°C), les ARN sont précipités dans 3 volumes d'éthanol 100% et 1/10 de volume d'AcNa 3 M, pendant 2 heures à -20°C. Après centrifugation (15700 g, 15 minutes, 4°C) les culots d'ARN sont lavés à l'éthanol 80% puis séchés. Les culots sont ensuite repris dans 10 µL d'eau milli-Q (Merck).

3.1.2.6. Ligation

3.1.2.6.1. Ligation (1) T4 RNA ligase (ThermoFischer)

Au cours de la variante 1.1 (Figure 19), les ARN (10 µL) digérés à la nucléase S1 sont incubés avec 5 U de T4 RNA ligase (ThermoFischer) et 2 µL de tampon 10X dans un volume final de 20 µL pendant 1 heure à 37°C. Les ARN sont purifiés avec un volume de PCI, et précipités avec 3 volumes d'éthanol 100% et 1/10 de volume d'AcNa 3 M, pendant 2 heures à -20°C. Après centrifugation (15700 g, 15 minutes, 4°C) les culots d'ARN sont lavés à l'éthanol 80%, séchés et repris dans 10 µL d'eau milli-Q (Merck).

3.1.2.6.2. Ligation (2) T4 RNA ligase (NEB)

Au cours des variantes 1.2, 2 et 3 (Figure 19), les ARN digérés par l'enzyme S1 (10 µL) sont incubés avec 10U de RNA ligase I (NEB), 2 µL de tampon 10X, 20 U de RNasine (Promega), 10% de PEG8000, et 50 µM d'ATP dans un volume final de 20 µL pendant une nuit à 16°C, puis deux minutes à 70°C. Les ARN sont purifiés avec un volume de PCI, et précipités avec 3 volumes d'éthanol 100% et 1/10 de volumes d'AcNa 3 M, pendant 2 heures à -20°C. Après centrifugation (15700 g, 15 minutes, 4°C) les culots d'ARN sont lavés à l'éthanol 80%, séchés et repris dans 10 µL d'eau milli-Q (Merck).

3.1.2.6.3. Ligation (3) à haute température

3.1.2.6.3.1. Adénylation

Les 10 µL d'ARN digérés à la nucléase S1 sont incubés avec 100 pmol de *Mth* RNA ligase (enzyme qui catalyse l'adénylation, NEB), 2 µL de tampon 10X, et 2 mM d'ATP dans

un volume final de 20 µL pendant 1 heure à 65°C, suivi de 5 minutes à 90°C. Les ARN sont purifiés et précipités comme ci-dessus (variante 4.1 et 4.2, figure 19).

3.1.2.6.3.2. Ligation

Les 10 µL d'ARN adénylés sont incubés avec 40 pmol de Thermostable 5'App DNA/RNA Ligase (NEB), et 2 µL de tampon 10X dans un volume final de 20 µL pendant 1 heure à 68°C, puis pendant 3 minutes à 90°C. Les ARN sont purifiés et précipités comme indiqué ci-dessus (variante 4.1 et 4.2, figure 19).

3.1.2.7. Digestion des ARN par la RNase R

Cette étape a été réalisée seulement pour les variantes 3 et 4.2 (Figure 19). Après une étape de ligation de type (2) (variante 3), les ARN ne sont pas précipités, mais la RNA ligase I (NEB) est inactivée à la chaleur (2 minutes à 70°C). Les 20 µL d'ARN ligués sont directement utilisés pour la digestion par la RNase R. Après une étape de ligation de type (3) (variante 4.2), les ARN sont purifiés et précipités comme présenté précédemment (*cf* II.3.1.2.6.3.1). L'intégralité du matériel ligué est incubé avec 0,5 µL de RNase R (Epicentre), et 2,5 µL de tampon 10X dans un volume final de 25 µL, 10 minutes à 37°C. Enfin les ARN sont purifiés et précipités comme ci-dessus.

3.1.2.8. Réversion du pontage

Les tubes Eppendorf contenant 10 µL d'ARN sont placés ouvert dans la glace à environ 15 centimètres de la source d'UV et subissent une irradiation à 254 nm pendant 10 minutes. Les ARN sont dilués dans un volume final de 100 µL et précipités dans 3 volumes d'éthanol 100%, 1/10 de volume d'AcNa 3 M, et 1/100 de volume de glycogène 1 µg/µL pendant une nuit à -20°C. Après centrifugation (15700 g, 15 minutes, 4°C), les culots d'ARN sont lavés à l'éthanol 80% puis séchés. Les culots sont ensuite repris dans 10 µL d'eau milli-Q (Merck).

3.1.3. Mise au point du protocole : variantes 5 et 6

3.1.3.1. Traitement au psoralène

La culture bactérienne, divisée en trois fois 16 mL dans des tubes Falcon 50 mL, est centrifugée pendant 30 min à 3750 g à 4°C. Les bactéries sont lavées avec 10 mL de PBS 1X pH 7,4, puis resuspendues dans 1 mL du même tampon pour atteindre 5.10⁹ bactéries/mL. Puis, 500 µg/mL (variante 5) ou 350 µg/mL (variante 6) de psoralène TEG-azide (PTA, Berry & associates) sont mis dans deux tubes (Figure 19). Dans un des deux tubes, 50 mM d'EDTA est également ajouté (variante 5). Une condition contrôle sans PTA est effectuée dans le tube restant.

Le mélange est incubé pendant 5 min à 37°C sous agitation. Du milieu BHI (13,5 mL) est ensuite ajouté pour stopper le traitement EDTA (cela est effectué dans tous les tubes indépendamment de la présence d'EDTA ou de PTA afin d'assurer la même dilution des bactéries). Un contrôle de viabilité des bactéries est effectué (variante 6) en étalant 5 µL de chaque tube sur une gélose au sang (VWR), qui sera ensuite incubée une nuit à 37°C.

3.1.3.2. Irradiation aux UV

Les suspensions bactériennes sont réparties dans des boîtes de Pétri de 9 cm gardées sur glace et soumises aux UV à 365 nm pendant 15 minutes (Stratalinker 1800 Stratagene), à environ 15 cm de la source (variante 5) ou dans des boîtes de Pétri en quartz préalablement refroidies sur glace et traitées aux UV pendant 22 ou 105 secondes à 365 nm (Vari-X-link UVO3) (variante 6). Un contrôle de viabilité des bactéries est effectué (variante 5 et 6) en étalant 5 µL de chaque boîte de Pétri sur une gélose au sang (VWR) qui sera ensuite incubée une nuit à 37°C. Les bactéries sont centrifugées pendant 15 min à 3750 g à 4°C. Une mesure de la DO_{360nm} du surnageant dans une cuvette en quartz permet de vérifier si le PTA est encore abondant dans le milieu extracellulaire (variante 5). Le surnageant des bactéries non traitées au PTA est utilisé comme « blanc ».

3.1.3.3. Extraction des ARN et traitement à la DNase

Ces étapes sont réalisées comme présenté ci-dessus (respectivement partie II.3.1.2.2 et II.3.1.2.3). Cependant, après le traitement avec la DNase, les ARN sont purifiés avec le kit « *RNA clean & concentrator* » (Zymo research) et resuspendus dans un volume final de 15 µL d'eau milli-Q (Merck).

3.1.3.4. Réaction de chimie « click »

L'ARN (5µg) est incubé avec 15 mM de Biotin Dibenzocyclooctyne (DIBO) Alkyne (Fisher scientific) et 1 µL de tampon Tris-HCl 0,25 M pH 7,5, dans un volume final de 100 µL pendant 1 heure 30 à 37°C. Les ARN sont purifiés avec le kit « *RNA clean & concentrator* » (Zymo research) et resuspendus dans un volume final de 200 µL d'eau milli-Q (Merck).

3.1.3.5. Purification sur des billes magnétiques de streptavidine

Les étapes de purification sont réalisées à l'aide d'un rack magnétique (Amersham Biosciences). Les billes (50 µL par échantillon) (streptavidin mag sepharose, GE healthcare) sont équilibrées avec 500 µL de Tris-HCl 10 mM pH 7,5, pendant 5 minutes à 20°C, et bloquées avec 40 µL d'oligonucléotide à 100 pmol/µL (AGAG... 33 nucléotides, IDT) pendant 15

minutes à 20°C sous légère agitation (300 rpm, variante 5) ou en utilisant l'agitateur PTR35 Rotator (Grant Instruments), qui permet de réaliser un programme alternant entre la fonction agitation et vortex (100 rpm 10 secondes, rotation 80°C 10 secondes, vortex 5°C 5 secondes) (variante 6). Les billes sont ensuite lavées trois fois avec 500 µL de Tris-HCl 10 mM pH 7,5. Les billes sont finalement resuspendues dans 200 µL de tampon BW2X (Tris-HCl 10 mM pH 7,5, 1 mM EDTA, 2 M NaCl) et sont ajoutées aux 200 µL d'ARN pontés et biotinylés. Le mélange est incubé pendant 30 minutes à 20°C, sous légère agitation (300 rpm), avec un passage de vortex (800 rpm) de 10 secondes toutes les 5 minutes (variante 5) ou en utilisant l'agitateur PTR35 Rotator (Grant Instruments, 100 rpm 10 secondes, rotation 80°C 10 secondes, vortex 5°C 5 secondes) (variante 6). Les ARN fixés sur les billes sont ensuite lavés trois fois avec 500 µL de tampon BW1X.

3.1.3.6. Digestion des ARN à la nucléase S1 sur les billes

Les ARN fixés sur les billes sont incubés avec 25 U de nucléase S1 (ThermoFisher) et 20 µL de tampon 5X dans un volume final de 100 µL pendant 30 minutes à 20°C, en alternant 5 minutes avec une agitation au vortex (800 rpm) et 5 minutes avec une agitation légère (300 rpm) (variante 5) ou en utilisant l'agitateur PTR35 Rotator (Grant Instruments, 100 rpm 10 secondes, rotation 80°C 10 secondes, vortex 5°C 5 secondes) (variante 6). À la fin, 3 lavages sont effectués avec 500 µL de tampon Tris-HCl 10 mM pH 7,5.

3.1.3.7. Ligation sur les billes (1) : variante 5

3.1.3.7.1. Adénylation

Les ARN fixés sur les billes sont incubés avec 200 pmol de *Mth* RNA ligase (NEB), 10 µL de tampon 10X, et 10 mM d'ATP dans un volume final de 100 µL pendant 1 heure à 65°C, avec une agitation légère (300 rpm). Les billes sont ensuite lavées trois fois avec 500 µL de tampon Tris-HCl 10 mM pH 7,5.

3.1.3.7.2. Ligation

Les ARN fixés sur les billes sont incubés avec 80 pmol de Thermostable 5'App DNA/RNA Ligase (NEB), et 10 µL de tampon 10X dans un volume final de 100 µL pendant 1 heure à 68°C, avec une agitation légère (300 rpm). Les ARN fixés sur les billes sont ensuite lavés trois fois avec 500 µL de tampon Tris-HCl 10 mM pH 7,5.

3.1.3.8. Ligation sur les billes (2) : variante 6

Les ARN fixés sur les billes sont incubés avec 40 U de RNA ligase I (NEB), 10 µL de tampon 10X, 40 U de RNasine (Promega), 1 mM d'ATP dans un volume final de 100 µL pendant une nuit à 16°C sur le PTR35 Rotator (Grant Instruments, 100 rpm 10 secondes, rotation 80°C 10 secondes, vortex 5°C 5 secondes). Les ARN fixés sur les billes sont lavés trois fois avec 500 µL de tampon Tris-HCl 10 mM pH 7,5.

3.1.3.9. Réversion du pontage

3.1.3.9.1. Réversion directe du pontage

Les ARN fixés sur les billes sont resuspendus dans 400 µL d'eau milli-Q (Merck) et déposés par gouttes de 30 µL sur une couche de parafilm elle-même fixée sur un support froid. Les gouttes sont placées à 15 cm de la source d'UV et subissent une irradiation à 254 nm pendant 2 minutes (Stratalinker 1800, Stratagene) (variante 5). Dans la variante 6, les gouttes sont déposées dans une boîte de Pétri en quartz préalablement refroidie, puis placée dans le Vari-X-link pour une irradiation à 254 nm pendant 22 secondes.

Les gouttes sont ensuite collectées dans un tube Eppendorf placé sur le rack magnétique permettant de retenir les billes et de récupérer les ARN qui ne sont plus intercalés par le PTA dans le surnageant. Les ARN sont concentrés avec le kit « *RNA clean & concentrator* » (Zymo research) dans un volume final de 10 µL d'eau milli-Q (Merck).

3.1.3.9.2. Elution suivie de la réversion du pontage

Les ARN fixés sur les billes sont resuspendus dans 400 µL de tampon d'élution (2% SDS, 6 M urée) et incubés pendant 5 min à 90°C. Les ARN détachés des billes magnétiques sont récupérés dans le surnageant, après que le tube Eppendorf soit placé sur le rack magnétique. Les ARN sont concentrés avec le kit « *RNA clean & concentrator* » (Zymo research) dans un volume final de 60 µL d'eau milli-Q (Merck). Les ARN sont déposés en gouttes de 30 µL sur une couche de parafilm fixée sur un support froid pour la variante 5, puis la réversion du pontage est effectuée comme ci-dessus. Dans la variante 6, le tampon d'élution contient également de la biotine (30 mM biotine) et la réversion du pontage a lieu dans le Vari-X-link avec une irradiation à 254 nm pendant 22 ou 105 secondes (Figure 19). Les gouttes sont ensuite collectées dans un tube Eppendorf et les ARN sont purifiés avec le kit « *RNA clean & concentrator* » (Zymo research), dans un volume final de 10 µL d'eau milli-Q (Merck).

3.1.4. Séquençage

3.1.4.1. Séquençage d'une chimère connue résultante du couple d'ARN *tcaR*-RsaG

3.1.4.1.1. RT-PCR

L'ARN (10 µL) est incubé en présence de 50 µM d'amorce aléatoire (random hexaprimer, ThermoFisher Scientific) et 0,25 mM de chaque désoxyribonucléotide (GE Healthcare) dans un volume final de 16 µL pendant 5 minutes à 70°C. Ce mélange est ensuite incubé avec 5 U d'AMV-RT (Life Science), 2 µL de tampon 10X, et 8 U de RNasin (Promega) dans un volume final de 20 µL pendant 10 minutes à 25°C, 1 heure à 42°C et 5 minutes à 90°C.

L'ADNc (1 µL) synthétisé est mis en présence de 0,1 µM de chaque amorce spécifique correspondant à un couple d'ARN (IDT), 2,5 U de Taq polymérase (produite au laboratoire), 5 µL de tampon 10X, et 0,25 mM de chaque désoxyribonucléotide dans un volume final de 50 µL. La première étape de la PCR consiste en une étape de dénaturation (3 minutes à 95°C), suivie d'une étape d'amplification réalisée en 49 cycles (30 secondes à 95°C, 30 secondes à 49°C et 30 secondes à 72°C), et d'une étape de terminaison pendant 5 minutes à 72°C. Les produits de PCR ont été purifiés avec le kit nucleospin gel & PCR clean-up (Macherey-Nagel) et resuspendus dans un volume final de 30 µL d'eau milli-Q (Merck). Le résultat de la PCR est ensuite visualisé sur gel d'agarose 1%, TBE 1X, contenant du BET 0,001%. Une aliquote des échantillons (7,2 µL) supplémentée de 1,8 µL de bleu glycérol 5X est déposée dans les puits. Après migration à 135 V pendant 20 minutes, le gel est ensuite révélé par exposition aux UV (Gel Doc EZ, BioRad).

3.1.4.1.2. Clonage des fragments de PCR

Les produits de PCR purifiés sont clonés dans le vecteur pJET1.2/blunt (ThermoFisher Scientific) (variante 1.2 à 6) selon le protocole du fournisseur. Étant issu d'une PCR avec la Taq polymérase, 0,15 pmol d'ADN sont débarrassés des extrémités libres avant ligation. Enfin, les produits de ligation sont transformés dans les bactéries *E. coli* DH5α chimio-compétentes (cf III.3).

3.1.4.2. Séquençage à haut débit

Pour la variante 3, les banques d'ADNc ont été réalisées à la plateforme de séquençage de l'IBMP avec le kit NEBNext Small RNA library prep (NEB), puis séquencées à l'IGBMC avec le séquenceur illumina HiSeq 4000. Pour la variante 4.2, la société Novogene a construit les banques d'ADNc avec ou sans fragmentation (80-85°C pendant 1 min dans un tampon Tris-

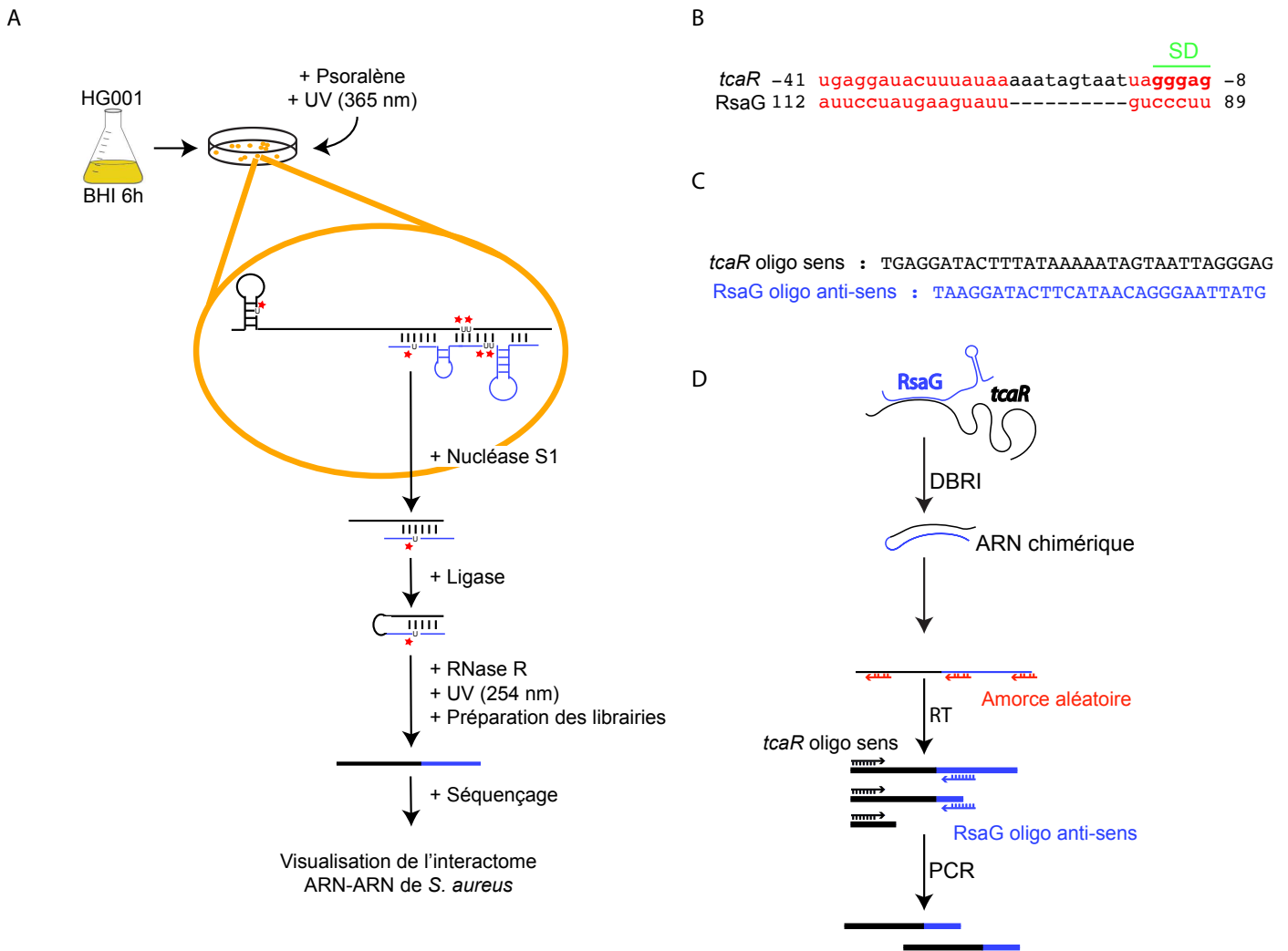


Figure 20 : Protocol schématique de DBRI et de la RT-PCR de chimères *tcaR*/*RsaG* : A) Représentation schématique simplifiée des étapes du protocole de DBRI. B) Site d'appariement entre l'ARNm *tcaR* et *RsaG*. En rouge sont présentés les nucléotides impliqués dans l'interaction, en gras la séquence Shine et Dalgarno (SD) de *tcaR* (prédiction d'appariement, IntaRNA). C) Séquence des oligonucléotides (oligo) sens et anti-sens utilisés pour l'amplification des chimères *tcaR*/*RsaG*. D) Représentation schématique des étapes du protocole de RT-PCR. Le sARN *RsaG* est représenté en violet, l'ARNm *tcaR* en noir, les mêmes couleurs sont utilisées pour leurs oligonucléotides spécifiques. Les amorces aléatoires (random primer) sont représentées en orange.

HCl pH 7,5) avec le kit Multiplex Small RNA Library Prep (NEB). Novagene a également réalisé le séquençage à haut débit avec le séquenceur illumina NovaSeq 6000.

3.2. Analyse des résultats

3.2.1. Explication du choix des conditions de culture

Une densité optique à 600 nm d'environ 3 dans le milieu BHI a été choisie pour réaliser la mise au point du protocole de DBRI (Figure 20A). Elle correspond à la fin de la phase exponentielle de croissance chez *S. aureus*. Un grand nombre d'ARN régulateurs sont transcrits à ce moment là (Geissmann et al., 2009). Une expérience de MAPS (résultats non présentés) a montré que *tcaR* était l'ARNm le plus enrichi lors de la chromatographie d'affinité avec MS2-RsaG réalisée dans les mêmes conditions de culture que celles utilisées pour DBRI. Cela a conforté notre choix du couple d'ARN RsaG-*tcaR* comme preuve de concept.

Ainsi, dans toutes les étapes de mise au point de la technique DBRI, une RT-PCR a été établie afin de mettre en évidence la formation d'une chimère entre le sARN RsaG et l'ARNm *tcaR* pour valider l'efficacité du protocole (*cf* ci-dessous et Figure 20B, chapitre II.2). La transcription inverse a eu lieu avec des amorces aléatoires permettant une amplification de tous les ARN en ADNc. Ensuite, la PCR a été réalisée avec un couple d'oligonucléotides spécifiques localisés dans la région d'interaction entre les deux ARN. L'oligonucléotide sens correspond à la séquence du site d'interaction localisé dans le gène *tcaR*, et l'oligonucléotide antisens à la séquence complémentaire inverse dans le gène *rsaG* (Figure 20). Cette stratégie permet d'amplifier les chimères circulaires (ligué aux deux extrémités) ainsi que les chimères en lasso (ligué à une seule extrémité) qui ont été ligués entre l'extrémité 3' d'un fragment d'ARN *tcaR* à l'extrémité 5' d'un fragment d'ARN RsaG.

3.2.2. Variantes 1.1 et 1.2

Dans ces deux tentatives, trois concentrations d'AMT ont été testées, et seule l'étape de ligation diffère. Le protocole de la variante 1 n'a pas permis l'amplification de chimère par RT-PCR. Différentes raisons sont possibles :

Premièrement, le psoralène et ces dérivés, s'intercalent au niveau des duplexes d'ARN faisant intervenir des bases pyrimidiques. Cette étape permet de créer des liaisons covalentes entre l'AMT et l'ARN. Cette réaction est optimale à une longueur d'onde de 365 nm. Après irradiation nous obtenons une « photographie » des interactions ARN-ARN *in vivo* à un temps donné et dans les conditions de culture choisies. Cependant, la longueur d'onde utilisée pour le pontage à l'AMT a été de 312 nm contrairement à la longueur d'onde recommandée de 365 nm

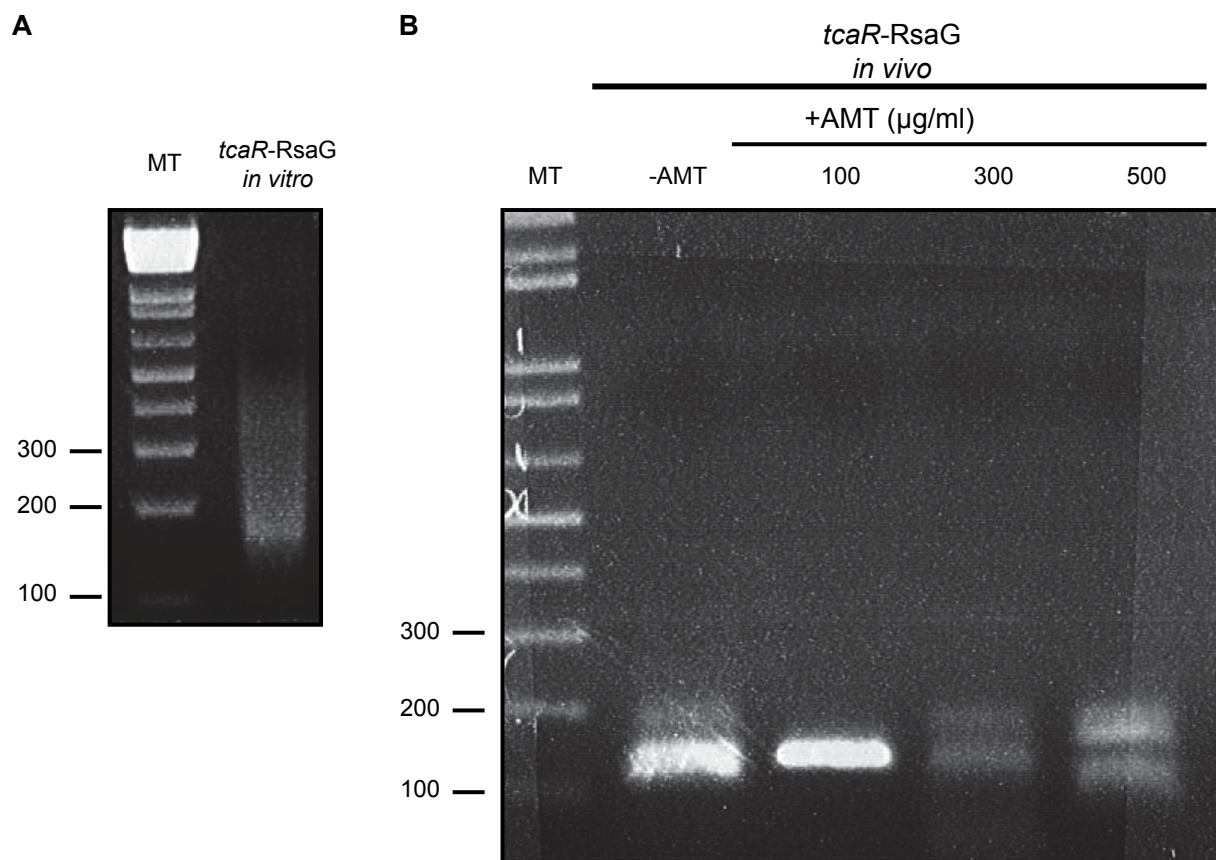


Figure 21 : Amplification des chimères *tcaR-RsaG* par RT-PCR : Sur gel d'agarose 1% TBE 1X, BET 0,001%, révélé par exposition aux UV. A gauche de chaque gel un marqueur de taille (MT) est migré en parallèle des échantillons, les graduations d'intérêt (100, 200, 300) sont indiquées et correspondent à un nombre en paires de bases. A) Résultat de l'amplification par RT-PCR des chimères *tcaR-RsaG* réalisé après un protocole DBRI variante 1.2 sur des transcrits *in vitro*. B) Résultat de l'amplification par RT-PCR des chimères *tcaR-RsaG* réalisé après un protocole DBRI variante 1.2 sur des transcrits *in vivo*. De gauche à droite les pistes correspondent aux résultats de RT-PCR pour une culture non traitée à l'AMT (-AMT), une culture traitée avec 100 µg/mL d'AMT (100), 300 µg/mL d'AMT (300), et 500 µg/mL d'AMT (500).

(en raison du matériel à disposition). Or, nous savons que la photo-addition du psoralène (et certainement de ses dérivés) se produit à une longueur d'onde comprise entre 320-400nm (Cimino et al., 1985). L'utilisation d'une longueur d'onde non optimale et en dehors des longueurs d'ondes décrites pour permettre la photo-addition a peut-être réduit l'efficacité du pontage, menant à un nombre trop faible de chimères ne pouvant pas être amplifiées par RT-PCR.

Deuxièmement, nous avons réalisé la ligation avec la T4 RNA ligase (Thermofisher) que nous avons à disposition au laboratoire. Néanmoins elle n'est certainement pas la plus adaptée pour permettre une ligation de proximité. Des contraintes structurelles pourraient en partie réduire son efficacité de ligation.

Ainsi j'ai optimisé le protocole de ligation dans la variante 1.2 pour permettre la formation de chimères circulaires notamment favorisée dans la procédure de la technique LigR (Sharma et al., 2016) par l'utilisation de la circRNA ligase à 60°C. Cette enzyme n'étant plus commercialisée, nous avons utilisé la T4 RNA ligase I (NEB) avec un protocole favorisant la circularisation des chimères à 16°C. En effet, la circularisation semble une option avantageuse sachant que les ligations de proximité ne sont pas toujours optimales. Elle améliore les chances d'obtenir au moins des chimères en lasso, et dans le meilleur des cas des chimères circulaires. Comme contrôle, seulement au cours de cette variante, j'ai effectué la ligation d'un couple d'ARN transcrit *in vitro* afin de définir l'efficacité de cette étape indépendamment des autres.

Ce protocole a permis une amplification par RT-PCR de chimères à partir des transcrits *in vitro* de *tcaR* et *RsaG* (en absence d'AMT et de pontage) indiquant que les conditions de ligation avec la T4 RNA ligase I sont adéquates (Figure 21A). L'expérience avec des ARN totaux a également permis la formation de chimères *RsaG-tcaR*. Cependant, des chimères ont pu être amplifiées *in vivo* avec ou sans pontage à l'AMT suggérant que leur formation est indépendante de la présence de ce composé. En effet la figure 21B montre une amplification en présence d'AMT, et ce avec toutes les concentrations utilisées (nous avons choisi plusieurs concentrations car nous ne connaissions pas la concentration nécessaire pour permettre le pontage des ARN *in vivo* chez les bactéries), mais également dans la condition sans AMT.

En plus d'une amplification des chimères par RT-PCR, les produits PCR sont également clonés dans le plasmide pJET afin de permettre leur séquençage et ainsi de caractériser la nature des chimères. L'analyse de leur séquence a montré qu'ils correspondent à des chimères entre les ARN *tcaR* et *RsaG*, aussi bien pour l'expérience avec les transcrits *in vitro* qu'*in vivo* avec ou sans AMT (Figure 22). La formation des chimères en absence d'AMT est certainement possible grâce à l'interaction très stable formée entre ces deux ARN, comme c'est le cas dans

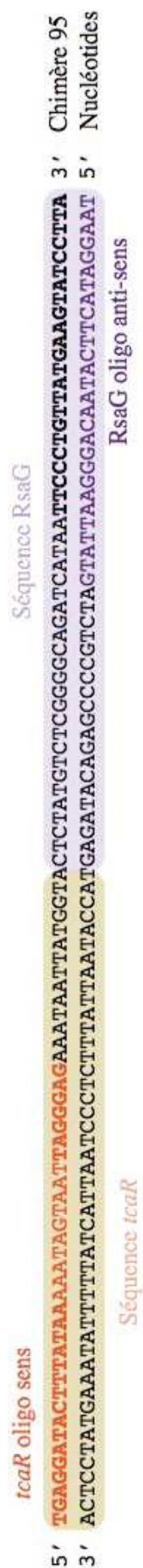


Figure 22 : Résultats du séquençage des fragments amplifiés par RT-PCR d'une chimère *tcaR*-RsaG : Les nucléotides surlignés en orange correspondent à la séquence de *tcaR*, ceux surlignés en violet à la séquence de RsaG. L'oligonucléotide utilisé pour l'amplification spécifique de *tcaR* est écrit en orange, celui écrit en violet pour RsaG. En gras les nucléotides de *tcaR* et RsaG prédits pour être en interaction.

l'expérience *in vitro*. Une hypothèse est que les appariements chez *S. aureus* pourraient être plus stables que ceux observés chez les entérobactéries qui requièrent souvent la présence d'Hfq. *S. aureus*, à l'inverse d'*E. coli*, n'a pas de protéine chaperonne qui facilite les interactions ARN-ARN. Bien que la protéine Hfq soit présente dans le génome de *S. aureus*, celle-ci ne semble pas remplir ce rôle (Bohn et al., 2007). De plus, aucun homologue de ProQ n'a été identifié (Olejniczak & Storz, 2017). Pour renforcer cette hypothèse, lorsque nous effectuons les expériences de MAPS (*cf* I.4.4.2 et Lalaouna et al., 2018) chez *S. aureus*, nous détectons, après extraction, un grand nombre d'interactions stables entre sARN et ARNm sans pontage aux UV, dont l'interaction *tcaR*-RsaG. L'avantage de l'AMT serait donc de permettre un enrichissement des chimères en stabilisant des interactions moins stables et/ou nécessitant une protéine chaperonne encore inconnue.

3.2.3. Variante 2

La procédure expérimentale s'est déroulée exactement comme lors de la variante 1.2, exceptée l'étape d'irradiation qui a été effectuée à la longueur d'onde de 365 nm, plus adaptée. Le séquençage des produits issus de la RT-PCR a confirmé la présence de chimères entre RsaG et *tcaR* avec les deux concentrations d'AMT utilisées et en absence d'AMT (Figure 23). Pour vérifier l'efficacité du pontage à 312 nm versus 365 nm, j'ai effectué une expérience de *Northern blot* avec les ARN totaux traités ou non à l'AMT aux deux longueurs d'onde testées en utilisant une sonde spécifique de l'ARN RsaG. Ainsi, j'ai regardé le profil de migration de RsaG suite aux différents traitements (Figure 24). Plus la migration de RsaG est retardée (dans un gel dénaturant), plus j'ai considéré que l'ARN forme des interactions covalentes (grâce à l'AMT) intra ou intermoléculaires. Les résultats indiquent que le pontage à 365 nm est effectivement plus efficace (Figure 24).

Pour conclure sur les 3 premières étapes, j'estime que les conditions de ligation avec la T4 RNA ligase I sont efficaces pour permettre une ligation de proximité, qu'un pontage à 365 nm est préférable qu'à 312 nm, et que la concentration d'AMT, même la plus faible, semblent suffisantes pour permettre un pontage entre ARN (Figure 24). Enfin, il semblerait que chez *S. aureus*, certaines interactions sont suffisamment stables pour former des chimères indépendamment d'un pontage à l'AMT.

3.2.4. Variante 3

3.2.4.1. Approche expérimentale utilisée

Forte des résultats obtenus en utilisant les variantes 1 et 2, j'ai reproduit la variante 2 afin de séquencer à haut débit l'ensemble des chimères formées. Afin de permettre un séquençage à haut débit, deux étapes ont été ajoutées au protocole de la variante 2.

Premièrement, une étape de ribodéplétion a été effectuée après l'extraction des ARN totaux. Cette étape permet d'éliminer les ARN ribosomiques de la préparation d'ARN, qui représentent au moins 85% des ARN totaux (Scott et al., 2010). Cette procédure permet d'augmenter la possibilité de séquencer en plus grande proportion les ARN chimériques.

Deuxièmement, une étape de digestion à la RNase R après la ligation des ARN a été ajoutée. La RNase R digère tous les ARN linéaires et permet d'enrichir les ARN chimériques sous forme de lasso et circulaires. Cette procédure permet également de diminuer la profondeur du séquençage nécessaire, en éliminant les ARN qui n'ont pas subi de ligation.

L'AMT à 500 µg/mL a été privilégié pour cette expérience, puisque les résultats précédents n'ont pas montré de différences entre les concentrations testées. Comme contrôle, une condition sans ligase subissant exactement le même traitement que la condition expérimentale avec ligation, a été effectuée.

3.2.4.2. Analyse du séquençage à haut débit

L'analyse du séquençage à haut-débit a été réalisée en collaboration avec Béatrice Chane Woon Ming. Après le séquençage, les lectures brutes (« *raw reads* ») ont été obtenues. Le type de séquençage que nous avons utilisé s'effectue à partir d'une extrémité (*single-end*) sur 50 nucléotides. La première étape de l'analyse bio-informatique appelée prétraitement (réalisé avec le programme *cutadapt*) consiste à éliminer la séquence des adaptateurs en 3'. Elle permet de retirer les nucléotides en 3' qui auraient été séquencés avec une faible qualité (seuil de 15). A la fin de cette étape, seules les séquences nucléotidiques de bonne qualité dépourvues d'adaptateur sont conservées. Afin de pouvoir aligner les séquences sur le génome une étape consiste ensuite à éliminer du pool de lecture celles qui ont une taille inférieure à 15 nucléotides (en effet les séquences inférieures à 15 nucléotides ont plus de probabilité d'être alignées sur plusieurs emplacements génomiques, ce qui pourrait biaiser les résultats). La distribution de la taille des séquences supérieures ou égales à 15 nucléotides, de bonne qualité et dépourvues d'adaptateur est semblable entre les quatre échantillons (avec/sans (+/-) AMT et avec/sans (+/-) ligation) (Figure 25). La majorité des séquences comprend entre 30 et 40 nucléotides. Malgré la présence d'AMT ou de ligation, la distribution révèle des fragments de tailles équivalentes

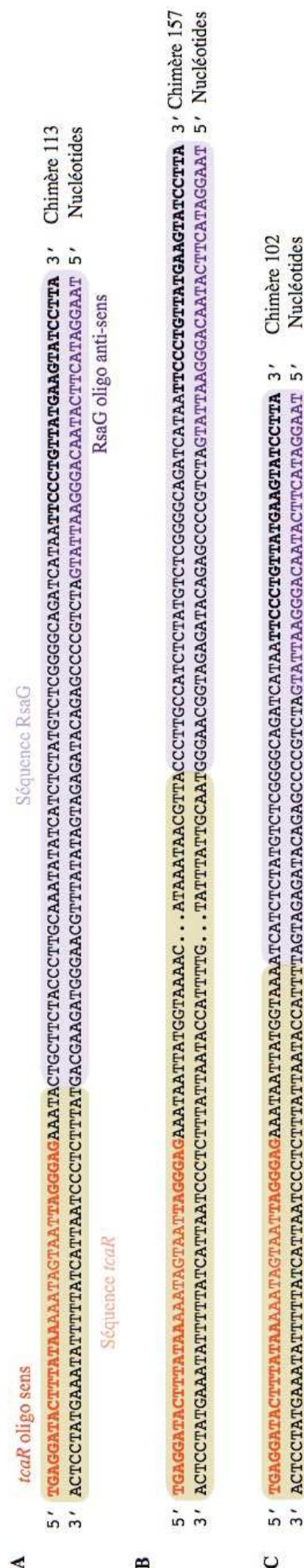


Figure 23 : Résultats du séquençage des fragments amplifiés par RT-PCR des chimères *tcaR*-RsaG en fonction de la présence d'AMT : Les nucléotides surlignés en orange correspondent à la séquence de *tcaR*, ceux surlignés en violet à la séquence de RsaG. L'oligonucléotide utilisé pour l'amplification spécifique de *tcaR* est écrit en orange, celui écrit en violet pour RsaG. En gras les nucléotides de *tcaR* et RsaG prédits pour être en interaction. A) La chimère est obtenue avec le protocole de la variante 2 en absence d'AMT, B) en présence de 300 µg/mL d'AMT, C) présence de 500 µg/mL d'AMT.

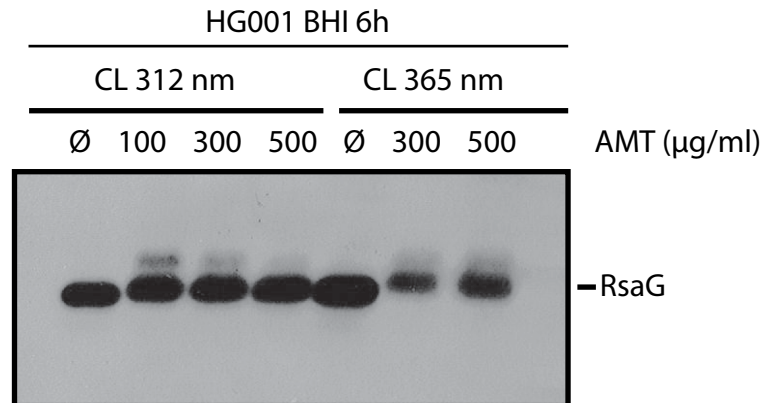


Figure 24 : Effet de la longueur d'onde du pontage et de la présence d'AMT sur la migration de RsaG en *Northern blot* : Les ARN sont extraits de cultures de *S. aureus* HG001 à 6h de croissance, traités à différentes concentrations d'AMT (0 (Ø), 100, 300 µg/mL, ou 500 µg/mL) et irradiées aux UV à 312 nm ou 365 nm (CL 312 nm ou CL365 nm), séparés sur un gel d'agarose 1%, puis transférés sur une membrane de nitrocellulose. RsaG est détecté avec une sonde ARN anti-sens marquée à la DIG et révélé par autoradiographie.

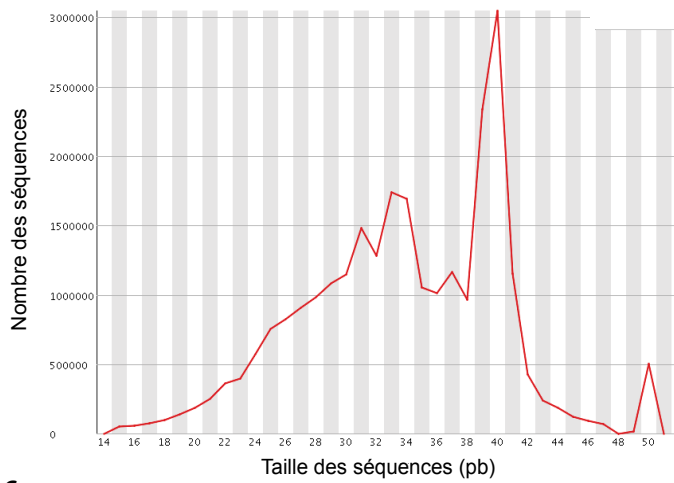
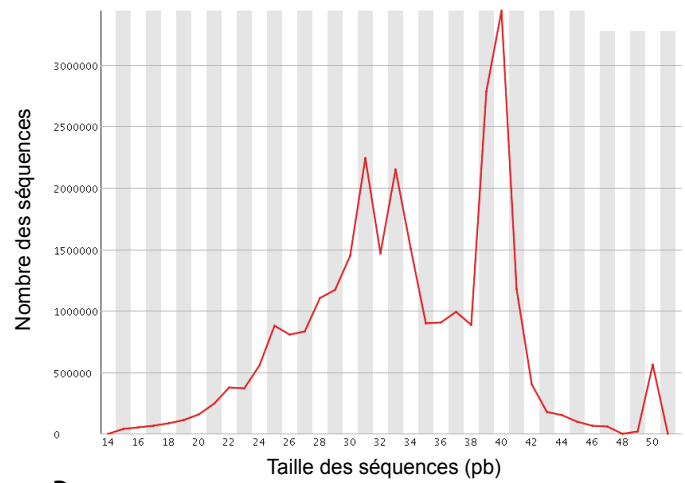
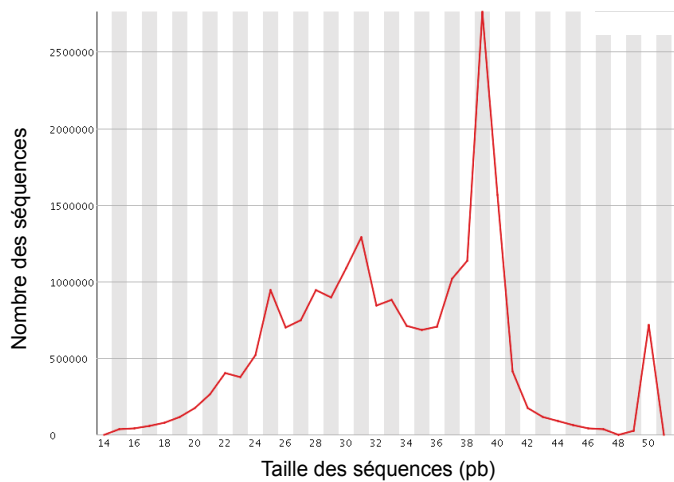
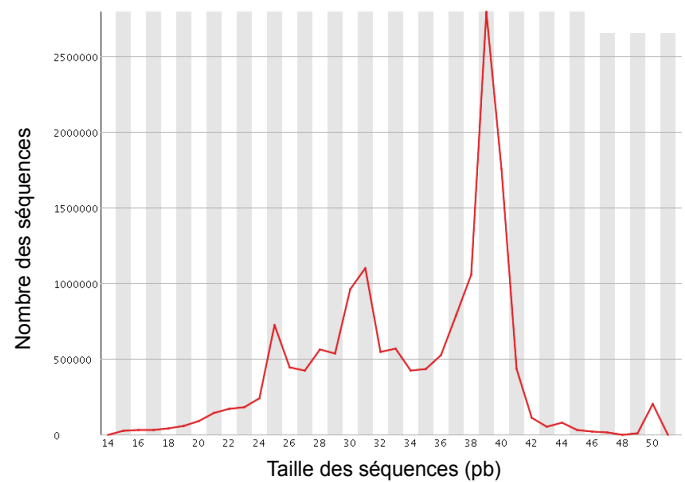
A**Distribution de la longueur des séquences : - AMT - ligation****B****Distribution de la longueur des séquences : - AMT + ligation****C****Distribution de la longueur des séquences : + AMT - ligation****D****Distribution de la longueur des séquences : + AMT + ligation**

Figure 25 : Représentation graphique de la distribution des lectures au sein des différents échantillons : Le nombre de lectures dans les différentes bibliothèques sont représentés en fonction de leur taille en paires de bases (pb). A) Echantillon traité sans AMT (-AMT) sans ligation (-ligation). B) Echantillon traité sans AMT (-AMT) avec ligation (+ligation). C) Echantillon traité avec AMT (+AMT) sans ligation (-ligation). D) Echantillon traité avec AMT (+AMT) avec ligation (+ligation).

entre échantillons. Les séquençages des chimères amplifiées par RT-PCR (Figure 23) avait révélé des chimères de tailles comparables avec ou sans AMT, il n'est donc pas étonnant qu'après séquençage à haut débit les tailles soient également similaires entre les conditions avec ou sans AMT. En revanche, un enrichissement du nombre de fragments était attendu dans la condition +AMT. Au contraire, la taille des pics majoritaires (autour de 30 à 40 nucléotides) est moins importante dans les conditions +AMT que -AMT, suggérant un nombre moins important de séquences lues dans les conditions +AMT. En revanche, en comparant les conditions +/- ligation, la distribution aurait dû être différente, reflétant l'absence ou la présence de ligation, ce qui n'a pas été le cas.

Un grand nombre de séquences a été aligné sur le génome (environ 46%), mais très peu d'entre elles étaient chimériques (environ 0,6%) (Figure 26A, colonnes 4 et 6). Il se peut que la digestion à la RNase R qui aurait dû enrichir les ARN ligués en lasso ou circulaire n'a pas vraiment fonctionné. Néanmoins, nous ne pouvons pas exclure que parmi les 46% de séquences alignées sur le génome, certaines résultent de chimères intramoléculaires. En effet si une chimère est formée intramoléculairement entre deux régions très proches (par exemple une tige-boucle), celle-ci pourrait être reconnue comme une séquence non chimérique. Dans la suite de l'analyse bio-informatique, parmi les fragments chimériques identifiés seront sélectionnées les interactions séquencées au moins 5 fois, appelées chimères redondantes (R-chimère). Les R-chimères représentent de nouveau un très faible pourcentage des séquences chimériques. Leur nombre varie de 142 (pour l'échantillon AMT – ligation +) à 45 (pour l'échantillon AMT + ligation -) (Figure 26A, colonne 7).

Les séquences chimériques ne s'alignant pas sur des gènes ou s'alignant sur les gènes des ARNr sont éliminées afin de définir la nature des ARN impliqués dans les chimères formés (Figure 26A, colonne 8). Des séquences d'ARNr sont tout de même présentes dans nos échantillons, bien qu'une étape de ribodéplétion ait été réalisée. Une représentation en diagramme du pourcentage de chaque type d'interactions retrouvées par rapport au nombre de R-chimères annotées par condition est présentée dans la figure 26B. Les chimères sont de cinq types : interactions entre ARNm, entre ARNt, entre ARNt et ARNm, entre ARNt et sARN et enfin entre ARNm et sARN. Les interactions faisant intervenir au moins un ARNt sont très récurrentes. Les interactions ARNt-mRNA sont retrouvées en grand nombre avec notre approche. Les interactions ARNt-ARNt et ARNt-sARN sont moins facilement explicables mais ont également été retrouvées dans des études précédentes (Liu et al., 2011; Waters et al., 2017). Les interactions auxquelles nous portons le plus d'intérêt, c'est-à-dire entre ARNm et sARN, ont été retrouvées seulement dans les conditions sans AMT et celles-ci ne représentent qu'un

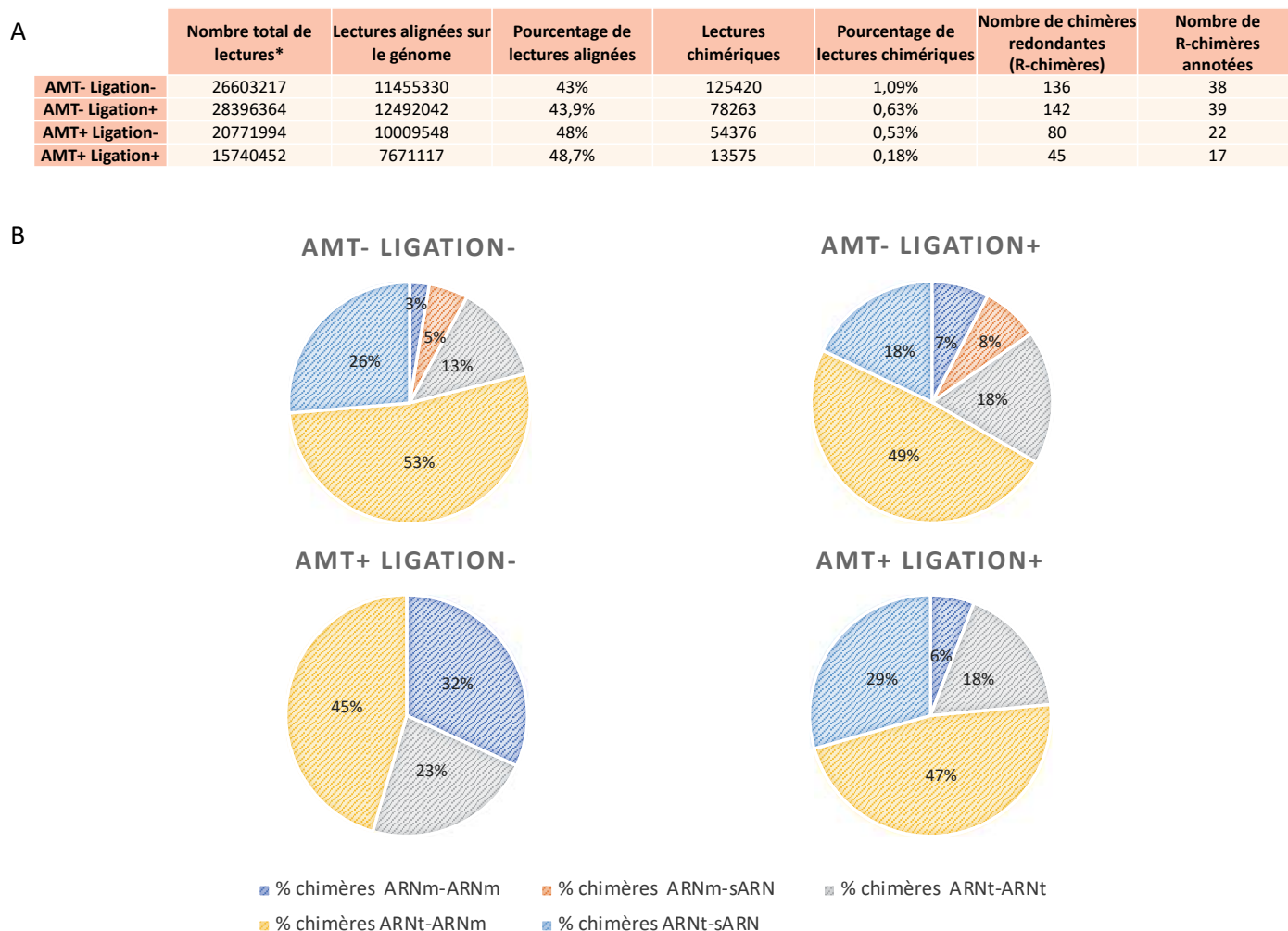


Figure 26 : Représentation de la nature des chimères obtenues après séquençage à haut débit de la variante 3 du protocole DBRI : A) Tableau des nombres de lectures ou de chimères en fonction des différentes étapes de l'analyse bio-informatique. *Après l'étape de pré traitement (*cf* texte pour les détails). B) Diagrammes pour chaque échantillon, avec ou sans AMT (+/-AMT), avec ou sans ligation (+/-LIGATION) de la répartition de la nature des chimères trouvées parmi les chimères redondantes (R-chimère) annotées. Les couleurs correspondent à celles de la légende sous les histogrammes. ARNm correspond à ARN messenger, ARNt à ARN de transfert, et sARN à ARN régulateur.

faible pourcentage de l'ensemble des interactions. De plus un pourcentage similaire de chimères ARNm-sARN a été retrouvé dans les conditions ligation + (8%) et ligation - (5%) suggérant que ces interactions pourraient relever d'un artéfact.

3.2.4.3. Conclusion

Malheureusement, nous n'avons pas obtenu de résultats exploitables et ce pour deux raisons majeures. Premièrement, nous avons pratiqué un séquençage à partir d'une extrémité sur 50 nucléotides, sachant que les chimères amplifiées par RT-PCR étaient de l'ordre de 100 nucléotides (*cf* figure 22 et 23). Nous pouvons imaginer que ce type de séquençage ne permet pas d'obtenir la séquence entière des chimères, rendant ainsi impossible leur assignation (Figure 26). On ne peut pas exclure que le protocole ait fonctionné et permette un enrichissement du nombre de chimères dans la condition +AMT +ligation au regard des 3 autres conditions. Pour nous permettre de recouvrir la séquence complète des chimères du même type que celles présentées en Figure 23, un séquençage des deux extrémités (*paired-end*) aurait été préférable, et cette stratégie sera donc employée à l'avenir. Le deuxième problème est que le protocole de ligation utilisé favorise la ligation circulaire des chimères (*cf* II.3.1.2.6.2 et II.3.2.3). Cependant nous n'avons pas effectué de fragmentation avant la construction des bibliothèques car celle-ci a l'inconvénient de raccourcir les chimères et de les cliver aléatoirement. Cela complexifie fortement l'analyse bio-informatique et l'identification des partenaires d'interactions. La construction de nos bibliothèques repose sur la ligation d'adaptateurs en 3'. Ainsi, toutes les chimères circulaires n'ayant pas d'extrémité 3' libre n'ont pas pu être liguées avec les adaptateurs et n'ont ainsi pas pu être séquencées. Cela explique certainement le faible nombre de chimères (environ 0,6%) retrouvées dans les différentes conditions ainsi que l'absence de différences dans la distribution de nos échantillons entre les conditions avec et sans ligation (Figure 25). Nous imaginons que notre protocole optimisé pour une ligation circulaire est efficace, en résulte une faible proportion de chimères en lasso, qui représentent les seules ayant pu être identifiées dans cette analyse.

Nous avons décidé de pratiquer un séquençage à haut débit pour cette variante suite aux résultats encourageants de RT-PCR. Bien que l'amplification par RT-PCR ne soit pas impactée par la circularisation des chimères, celle-ci empêche l'obtention de bibliothèque d'ADNc avec un protocole de ligation d'adaptateurs. Ces résultats bien que négatifs nous ont donc fait prendre conscience que des modifications devaient encore être apportées au protocole.

3.2.5. Variante 4.1

Suite aux conclusions apportées par le séquençage haut débit de la variante 3, nous avons voulu repenser l'étape de ligation pour qu'elle ne nécessite aucune fragmentation des ARN avant la création des banques de séquençage. Comme mentionné précédemment, la technique LigR (Sharma et al., 2016) utilise la circRNA ligase permettant une ligation circulaire des chimères à 60°C. Nous avons estimé qu'outre la circularisation, cette ligase était une des rares sur le marché à être thermostable et donc à pouvoir fonctionner dans des conditions dénaturantes pour l'ARN. La ligation en conditions dénaturantes est très avantageuse car elle permet de réduire le nombre de faux positifs qui représentent les interactions aspécifiques entre ARN qui auraient pu être ligués non spécifiquement. En effet, en réalisant la ligation à basse température des interactions entre ARN peuvent se produire dans le tube alors qu'elles n'auraient pas eu lieu *in vivo*. Nous avons donc cherché une ligase thermostable que nous pourrions utiliser dans notre protocole, et qui ne favoriserait pas la circularisation. La ligase Thermostable 5'App DNA/RNA Ligase (NEB) permet de réaliser l'étape de ligation à 68°C. Toutefois, la ligation avec cette enzyme nécessite une première étape d'adénylation qui est indispensable pour son fonctionnement optimal. Cette étape permet de convertir les extrémités 5' des ARN en adénosine triphosphate (5' adénosine triphosphate (5'App)). La procédure expérimentale est conduite comme lors de la variante 2, mais en modifiant l'étape de ligation (*cf* II.3.1.2.6.3). Une seule concentration d'AMT a été utilisée pour cette expérience, 300 µg/mL (*cf* II.3.2.4). Ce protocole, avec ligation à haute température a également permis une amplification par RT-PCR du couple RsaG-*tcaR* dans toutes les conditions (+/-AMT, +/- ligation). Cependant, c'est la première fois que nous réalisons une RT-PCR sur un échantillon sans ligation.

La ligation à haute température avait été choisie pour permettre une ligation plus spécifique des interactions intercalées par l'AMT, car normalement à cette température les interactions non pontées devraient être déstabilisées. Par ailleurs, le fait qu'il y ait une amplification en absence de ligation indiquerait qu'il y a un problème de spécificité avec la RT-PCR, ce qui expliquerait aussi l'amplification en absence d'AMT. La réalisation d'une PCR sur de l'ADN génomique de *S. aureus* ainsi qu'un contrôle sans RT devrait nous permettre de mieux comprendre où réside le problème. Cela démontre que bien que nous observions une amplification sur gel, il est indispensable de recourir à un séquençage pour déterminer la nature des produits PCR.

Les fragments PCR ont été ensuite clonés afin d'en étudier la nature après séquençage. En fonction des conditions, nous observons différents profils. Pour la condition avec AMT et

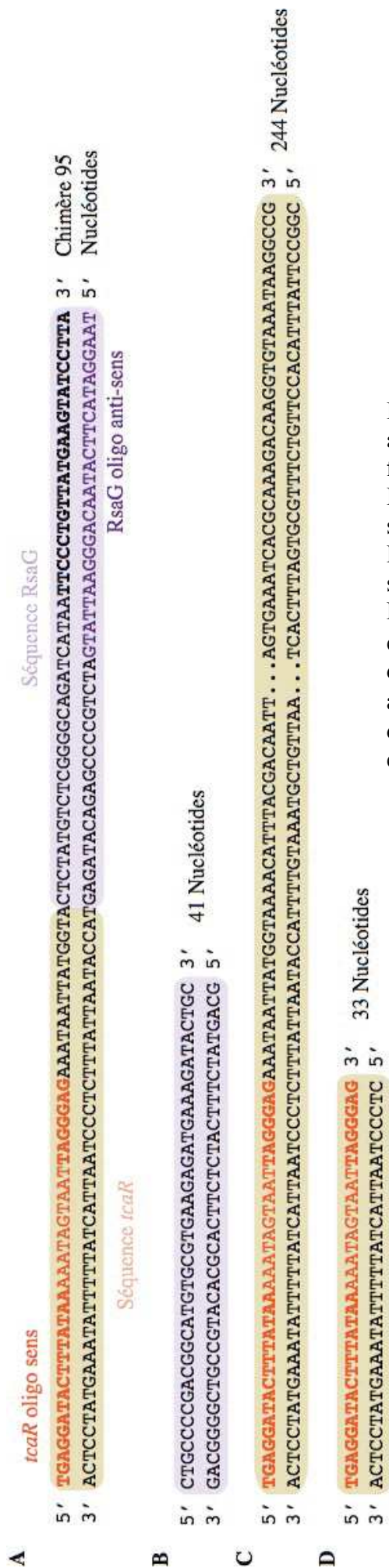


Figure 27 : Résultats du séquençage des fragments amplifiés par RT-PCR des chimères *tcaR*-RsaG en fonction de la présence d'AMT et de ligase (variante 4.1) : Les nucléotides surlignés en orange correspondent à la séquence de *tcaR*, ceux surlignés en violet à la séquence de RsaG. L'oligonucléotide utilisé pour l'amplification spécifique de *tcaR* est écrit en orange, celui écrit en violet pour RsaG, en gras les nucléotides de *tcaR* et RsaG prédits pour être en interaction. A) Chimère obtenue en présence de 300 µg/mL d'AMT et de ligase. B) Chimère obtenue en absence d'AMT avec ligase. C) Chimère obtenue en présence de 300 µg/mL d'AMT sans ligase. D) Chimère obtenue en absence d'AMT et de ligase.

ligation, nous observons des chimères issues de la ligation depuis l'extrémité 3' d'un fragment d'ARN *tcaR* à l'extrémité 5' d'un fragment d'ARN RsaG (Figure 27A). Pour les conditions sans AMT et sans ligation, avec AMT et sans ligation, ou bien sans AMT et sans ligation, nous n'observons pas la présence de chimères (Figure 27BCD). Ces observations nous montrent que la RT-PCR entraîne une amplification aspécifique en plus de l'amplification spécifique de chimère *tcaR*-RsaG. Puisqu'on ne retrouve pas de chimère spécifique de *tcaR*-RsaG dans la condition sans AMT, nous pouvons confirmer que la ligation à haute température a réduit le nombre de faux positifs lors de cette étape. Enfin, étant donné qu'une amplification aspécifique est observée par RT-PCR, il est indispensable lors du séquençage de nos échantillons d'ajouter un contrôle négatif sans ligation en parallèle.

3.2.6. Variante 4.2

3.2.6.1. Approche experimental utilisée

Dans la mesure où le protocole utilisé pour la variante 4.1 a fait ses preuves, nous avons mise en place la variante 4.2 permettant le séquençage à haut débit des chimères formées. Comme décrit dans la variante 3, les étapes de ribodéplétion après l'extraction des ARN et de digestion à la RNase R après la ligation des ARN ont été ajoutées.

Tous les échantillons soumis au séquençage ont été traités avec 300 µg/mL d'AMT. Dans cette expérience, nous avons comparé des échantillons avec ligation avec des échantillons sans ligation subissant exactement le même traitement que les échantillons avec ligation. Les échantillons « Minus » (3 ou 7) correspondent aux conditions sans ligation et les échantillons « Plus » (3 ou 7) aux conditions avec ligation. Ensuite, lors de la préparation des librairies, les échantillons 7 (Minus ou Plus) n'ont pas été fragmentés (80-85°C pendant 1 min) alors que les échantillons 3 (Minus ou Plus) l'ont été. La comparaison des échantillons avec ou sans fragmentation permettra d'estimer le pourcentage d'ARN ligué en lasso et d'ARN ligué de manière circulaire. Bien que le protocole de ligation ne favorise pas la circularisation, la possibilité d'un tel évènement ne peut pas être exclue. Il existe certainement une variabilité entre les échantillons, mais que nous n'avons pas quantifié car nous n'avons pas réalisé de réplicats des conditions avec et sans fragmentation. La construction des librairies repose sur la ligation d'adaptateurs en 5' et en 3' des ARN. Cette stratégie nous permet de séquencer les fragments par leurs deux extrémités en pratiquant un séquençage de type « *paired-end* » sur 150 nts.

3.2.6.2. Analyse du séquençage à haut débit

L'analyse du séquençage à haut-débit a été réalisée en collaboration avec l'équipe du Pr. Bjoern Voss (Université de Stuttgart, Allemagne). Comme lors de la variante 3 (cf II.3.2.5.2), la première étape de l'analyse bio-informatique appelée prétraitement (réalisé avec le programme RNA_{NUE}), consiste à éliminer la séquence des adaptateurs (ici présents en 5' et 3') et les séquences de mauvaise qualité. Le nombre de séquences nucléotidiques totales après séquençage et le nombre de séquences après l'étape de prétraitement pour chaque échantillon sont présentés dans le tableau 4 (colonnes 3 et 4). Les étapes suivantes de l'analyse bio-informatique permettent d'aligner sur le génome les différentes séquences individuelles générées par le séquençage de type « *paired-end* » afin de reconstituer des fragments (c'est-à-dire deux fragments de séquences provenant d'un même ADNc). La distribution de ces séquences lues est représentée pour chaque échantillon en figure 28. Nous observons que la distribution est différente entre les quatre échantillons. Les échantillons avec ligation (Plus_3, Plus_7) sont similaires entre eux mais différents des échantillons sans ligation (Minus_3, Minus_7), eux-mêmes similaires entre eux. Un pic du nombre de séquences est présent autour de 60 nucléotides dans les échantillons avec ligation et absent des échantillons sans ligation. En revanche, la distribution est similaire en présence ou en absence de fragmentation (3, 7) suggérant que la fragmentation n'a pas eu d'effet drastique sur la taille et le nombre de fragments dans nos échantillons. La distribution est différente en fonction de la présence de ligation, En effet, le pic majoritaire à 60 nts est absent des échantillons sans ligation suggérant que ce pic résulte de l'application de cette étape de ligation. De plus, le nombre de séquences lues est largement supérieur dans les conditions avec ligation (Plus_3, Plus_7), ce qui suggère que cette étape a été efficace.

Parmi les séquences alignées sur le génome, on retrouve très peu de séquences chimériques dans les conditions sans ligation (Minus) (inférieur à 1%), cela reflète qu'en absence de ligation, la formation de chimères est rendue impossible (Tableau 4, 6ème colonne). En revanche, dans les conditions avec ligation, on observe un enrichissement significatif du nombre de séquences chimériques (50,44% et 53,02%), confirmant que le protocole de ligation est essentiel pour former les chimères (Tableau 4, 6ème colonne).

Pour exploiter nos résultats, nous avons tout d'abord écarté les chimères faisant intervenir une ou deux séquences alignées sur des séquences non annotées de notre génome de référence (Tableau 4, 7ème colonne). Parmi les chimères annotées, un grand nombre était identiques reflétant sans doute la présence d'une expression assez forte de ces ARN (un séquençage du transcriptome en parallèle de l'expérience DBRI devrait nous permettre à

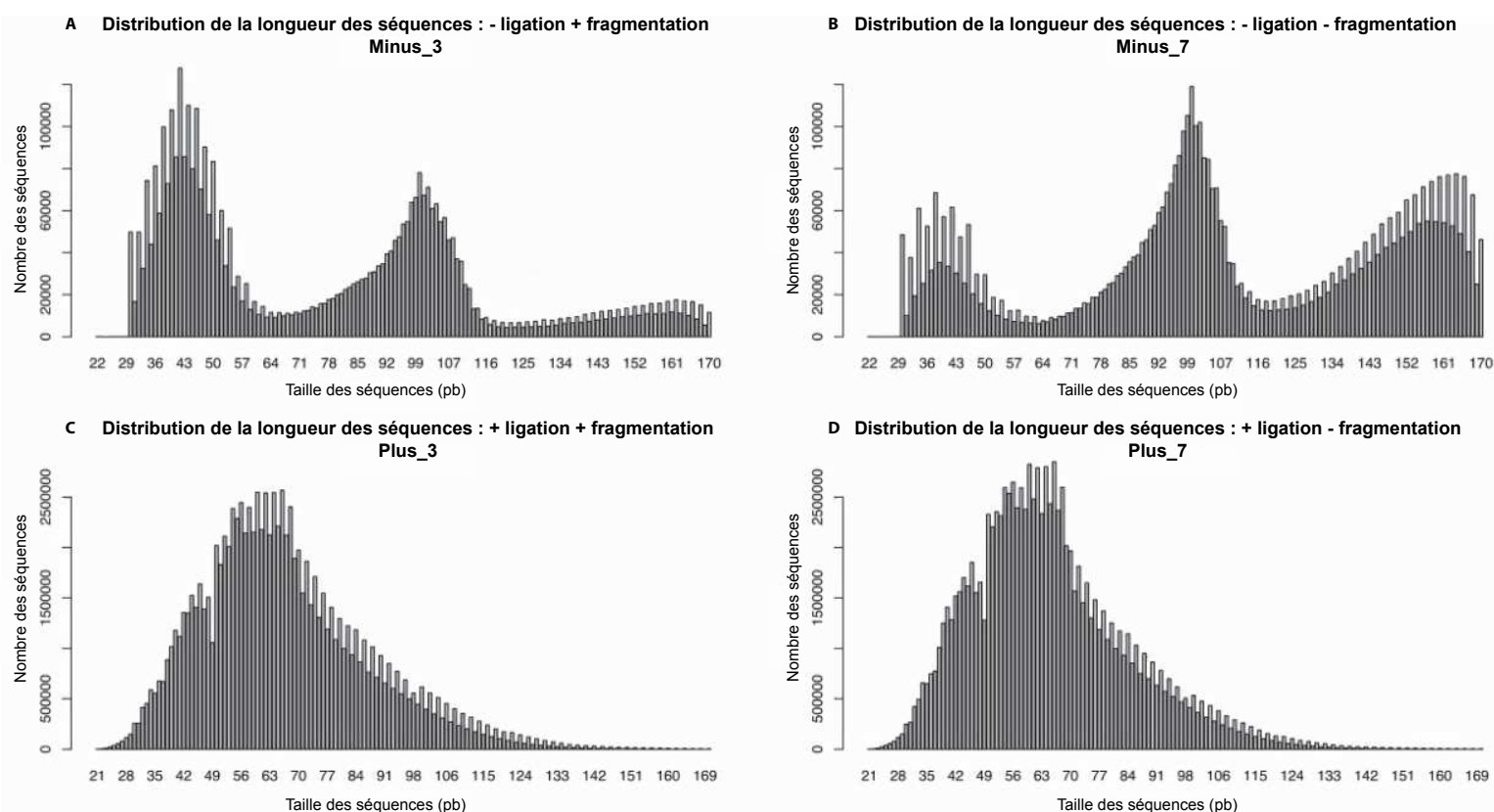


Figure 28 : Représentation graphique de la distribution des lectures au sein des différents échantillons : Le nombre de lectures dans les différentes librairies sont représentées en fonction de leur taille en paires de base (pb). A) Echantillon traité sans ligation (-ligation) avec fragmentation (+fragmentation). B) Echantillon traité sans ligation (-ligation) et sans fragmentation (-fragmentation). C) Echantillon traité avec ligation (+ligation) et avec fragmentation (+fragmentation). D) Echantillon traité avec ligation (+ligation) et sans fragmentation (-fragmentation).

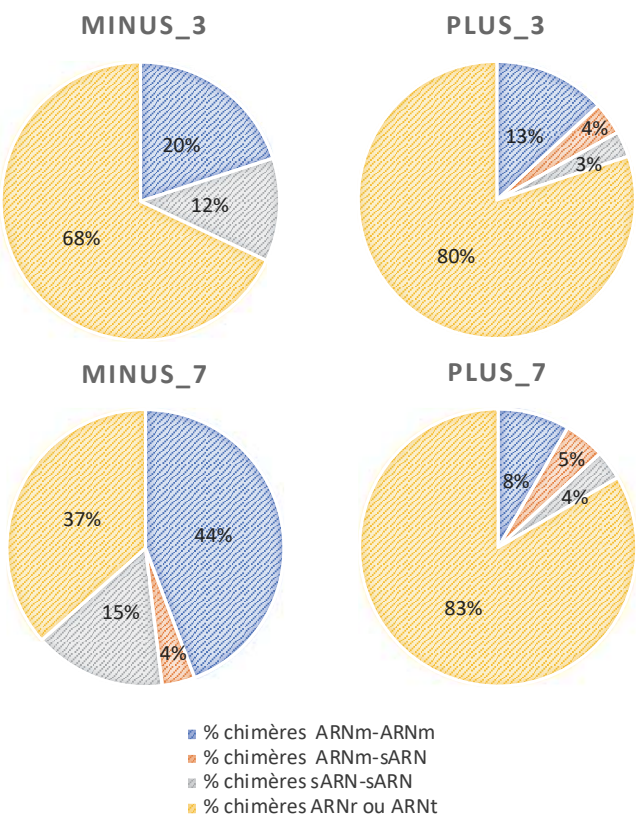
l'avenir d'estimer si l'abondance du nombre de chimères pour un couple d'ARN est en corrélation avec l'abondance respective des transcrits impliqués dans l'interaction). Si l'on regroupe toutes les chimères identiques, on obtient alors le nombre de chimères individuelles annotées, présentées dans le tableau 4 (colonne 8). Parmi les séquences chimériques individuelles annotées, nous n'avons retenu que les interactions redondantes (R-chimère, retrouvées au moins 5 fois dans un échantillon). Les R-chimères sont présentées dans le tableau 4 (colonne 9). C'est avec ces R-chimères que l'analyse de la composition en ARN des différents échantillons a été réalisée (Figure 29, interprétation ci-après). Utilisant l'ensemble de ces R-chimères, nous avons quantifié l'incidence d'une étape de fragmentation. D'une part, la fragmentation a l'avantage de permettre le séquençage des chimères circulaires. En revanche, elle a le désavantage d'entraîner une diminution de la taille des chimères rendant impossible l'assignation de certaines d'entre elles au cours de l'analyse bio-informatique (au moins 15 nucléotides par fragment doivent pouvoir être alignés pour considérer le fragment). Cet aspect ne semble pas avoir été pris en compte dans les études précédentes concernant l'analyse des interactomes ARN-ARN (Aw et al., 2016; Lu et al., 2016; Sharma et al., 2016). En analysant les R-chimères communes aux échantillons ligués avec fragmentation (Plus_3) et aux échantillons ligués sans fragmentation (Plus_7), nous avons dénombré 1507 R-chimères. Un calcul du taux d'enrichissement (ou fold change (FC)) de chaque R-chimère entre les deux échantillons nous permet d'estimer le rapport de R-chimères retrouvées en plus grand nombre grâce à la fragmentation, et à l'inverse, le rapport de R-chimères retrouvées en plus grand nombre en absence de fragmentation. Nous avons retenu les R-chimères qui étaient différentiellement enrichies dans l'un ou l'autre des échantillons si le FC est supérieur ou égal à 2. Outre le fait qu'il existe une variabilité entre échantillons, les résultats montrent que 24,42% des R-chimères sont plus enrichies dans l'échantillon avec fragmentation (pour chaque R-chimère, $FC = (Plus_3/Plus_7) \geq 2$). Ceci indique que ces 24,42% de chimères ont majoritairement été liguées de manière circulaire et sont donc retrouvées en plus faible quantité dans l'échantillon sans fragmentation que dans l'échantillon avec fragmentation. En revanche, 11,94% des R-chimères sont enrichies dans l'échantillon sans fragmentation ($FC = (Plus_7/Plus_3) \geq 2$ ou $(Plus_3/Plus_7) \leq 0,5$), suggérant que parmi les chimères liguées en lasso, l'étape de fragmentation a empêché une assignation correcte des deux partenaires (dans l'échantillon Plus_3). Cela suggère que notre protocole a généré environ un quart de chimères circulaires qui n'ont pas pu être séquencées dans l'échantillon non fragmenté (Plus_7, ~24%). De même, seulement un faible pourcentage de chimères n'a pas pu être assigné en tant

A

Nom	Echantillons	Nombre total de lectures	Nombre total de lectures après prétraitement	Lectures chimériques	Pourcentage de lectures chimériques	Nombres total de chimères annotées	Nombre de chimères individuelles annotées	Nombre de chimères redondantes (R-chimère) individuelles annotées
Minus_3	AMT+ Ligation-Fragmentation+	154178530	10183724	38151	0,37%	798	139	25
Minus_7	AMT+ Ligation-Fragmentation-	149834777	12567044	44769	0,36%	702	333	52
Plus_3	AMT+ Ligation+ Fragmentation+	156958645	129507877	65319882	50,44%	409865	11690	3859
Plus_7	AMT+ Ligation+ Fragmentation-	159301612	140540414	74514608	53,02%	303615	11289	3475

Tableau 4 : Analyse quantitative des lectures et des chimères après séquençage à haut débit de la variante 4.2 du protocole DBRI aux différentes étapes de l’analyse bio-informatique.

A



B

	Minus_3	Plus_3	Minus_7	Plus_7
% chimères ARNr-ARNr	94%	5%	100%	5%
% chimères ARNr-ARNm	/	15%	/	9%
% chimères sARN-ARNr	6%	4%	/	3%
% chimères ARNt-ARNt	/	35%	/	42%
% chimères ARNm-ARNt	/	25%	/	24%
% chimères sARN-ARNt	/	5%	/	5%
% chimères ARNr-ARNt	/	11%	/	12%

Figure 29 : Représentation schématique de la nature des chimères obtenues après séquençage à haut débit de la variante 4.2 du protocole DBRI : A) Diagramme représentant la répartition de la nature des R-chimères annotées, pour les échantillons, avec ou sans ligation (MINUS ou PLUS), avec ou sans fragmentation (3 ou 7). ARNm correspond à ARN messager, ARNt à ARN de transfert, ARNr à ARN ribosomique, et sARN à ARN régulateur. B) Tableau représentant parmi la part des chimères avec ARNr ou ARNt (représenté en jaune dans le panel (A)) la proportion en pourcentage de chaque type de chimères dans chaque échantillon.

que chimère à cause de la réalisation d'une étape de fragmentation (Plus_3, ~11%). Il sera donc important pour la suite des expériences de garder à l'esprit que bien que le protocole de ligation utilisé ne favorise pas la circularisation, ce phénomène existe. Notre étude montre qu'il sera important de réaliser systématiquement des échantillons avec et sans fragmentation. En effet, outre les chimères qui sont différentiellement enrichies en fonction de la réalisation d'une étape de fragmentation, il existe également des chimères qui ne sont retrouvées qu'au sein d'un seul des deux échantillons et qui seraient donc perdues si nous faisons le choix de réaliser l'échantillon soit avec fragmentation ou soit sans fragmentation. En effet, nos résultats montrent que 3481 R-chimères sont retrouvées au sein d'un seul des deux échantillons (1848 dans Plus_3 ; 1633 dans Plus_7) sur 7334 R-chimères totales, cela représente 25,2% de R-chimères présentes seulement dans l'échantillon fragmenté (Plus_3), et 22,27% dans l'échantillon non fragmenté (Plus_7).

En étudiant la répartition des classes d'ARN impliquées dans les chimères, nous observons que les échantillons « Plus » sont très homogènes dans la répartition des différents types de chimères, contrairement aux échantillons « Minus » où la répartition n'est pas équivalente (Figure 29). Un grand nombre de chimères fait intervenir au moins un ARNr ou un ARNt (Figure 29). Ce type d'interaction représente la majorité des interactions (80% et 83%) dans les échantillons ligués (Plus_3 et Plus_7). Cette donnée n'est pas étonnante dans la mesure où nous savons que les ARNr et les ARNt à eux seuls représentent au moins 95-97% des ARN totaux (Rosenow et al., 2001). La composition exacte en ARN des interactions faisant intervenir au moins un ARNr ou un ARNt est présentée dans la figure 29B. Nous observons une bonne homogénéité au sein des échantillons ayant subi une ligation, de même qu'au sein de ceux n'ayant pas subi ce traitement. Dans les échantillons ligués (Plus_3 et Plus_7), nous remarquons que les interactions ARNt-ARNt sont majoritaires, suivies par les interactions ARNm-ARNt. Les plus petites fractions correspondent dans les deux échantillons aux interactions sARN-ARNr, ARNr-ARNr (qui représentent les chimères les plus représentatives dans les échantillons « Minus »), et sARN-ARNt. Malheureusement, nous n'avons pas de stratégies disponibles au laboratoire pour éliminer les séquences d'ARNt avant séquençage. Cependant, une stratégie basée sur l'utilisation de sondes spécifiques des ARNt couplées à des billes magnétiques, pourrait être utilisée. Concernant les ARNr, nous avons effectué une ribodéplétion (Ribo-zero plus, illumina). Bien que cela ait éliminé en partie les ARNr, nous pouvons tout de même observer qu'il persiste un grand nombre d'interactions faisant intervenir des ARNr dans nos échantillons (Figure 29B). L'efficacité de ce traitement est certainement réduite à cause du

pontage des ARNr, empêchant l'élimination d'une partie d'entre eux. Nous trouvons également dans une plus faible quantité des chimères entre ARNm, sARN, et des chimères ARNm-sARN (Figure 29A). Parmi les 4 et 5% de chimères ARNm-sARN, respectivement au sein des échantillons Plus_3 et Plus_7, aucune interaction connue n'a été retrouvée, comme celles faisant intervenir les sARN RsaA, RsaC, RsaI, et RsaG (Bronesky et al., 2019; Lalaouna et al., 2019; Tomasini et al., 2017, Desgranges et al., en préparation, cf II.2) (Tableau 4, colonne 8). Cela est d'autant plus étonnant que les contrôles réalisés dans les variantes précédentes avaient permis d'amplifier par RT-PCR des chimères entre le sARN RsaG et sa cible ARNm *tcaR*. Il a été décrit que la ligation des adaptateurs au cours de la construction des bibliothèques peut créer un biais qui se reflète par une inconsistance entre les expressions relatives obtenues par qRT-PCR et par séquençage haut débit (Fuchs et al., 2015). De plus, avant séquençage à haut débit nous réalisons une étape de digestion à la RNase R (absente lors des RT-PCR) qui pourrait entraîner une réduction trop importante des extrémités 3' de nos chimères rendant impossible leur assignation. Nous avons tout de même vérifié parmi toutes les séquences alignées sur le génome, indépendamment du fait qu'elles soient assignées en tant que chimères, si nous étions capables de retrouver des séquences alignées dans les gènes des sARN RsaA, RsaG, et RsaI (Figure 30). Aucun fragment n'a été détecté pour RsaG et RsaI dans les échantillons sans ligation et très peu pour RsaA (Miuns_3 et Miuns_7) alors qu'un grand nombre de fragments est observé dans les échantillons avec ligation (Plus_3 et Plus_7) pour tous les sARN (Figure 30A). Le fait qu'il y ait une telle différence d'abondance des séquences pour ces ARN entre les conditions +/- ligation semble indiquer que c'est l'étape de ligation qui permet d'expliquer leur présence. De manière intéressante, les séquences alignées sur les gènes des sARN ne sont pas uniformément réparties sur toute la longueur de leur gène (Figure 30B). Ces séquences sont présentes et pourraient donc faire partie de chimères qui n'ont pas pu être assignées lors de l'analyse. Un exemple de la distribution des lectures dans l'échantillon Plus_7 sur la séquence génomique des sARN RsaG, RsaI et RsaA est montré dans la figure 30. Chacun de ces sARN comporte des motifs définis comme importants pour la régulation de leurs cibles (cf II.1 ; II.2 ; Tomasini et al., 2017). Pour RsaG et RsaA, il s'agit de motifs riches en C, connus pour interagir avec la séquence SD de leurs ARNm cibles. Pour RsaA, les séquences sont largement réparties au niveau du motif riche en C et aux alentours. Cela paraît logique car si ces séquences sont chimériques, malgré qu'elles n'aient pas été assignées, nous nous attendons à ce que le motif riche en C de RsaA directement impliqué dans les interactions avec ces ARNm cibles, soit au cœur des chimères formées. À l'inverse pour RsaG, nous observons que les lectures sont réparties majoritairement de part et d'autre des deux motifs riches en C. Une explication

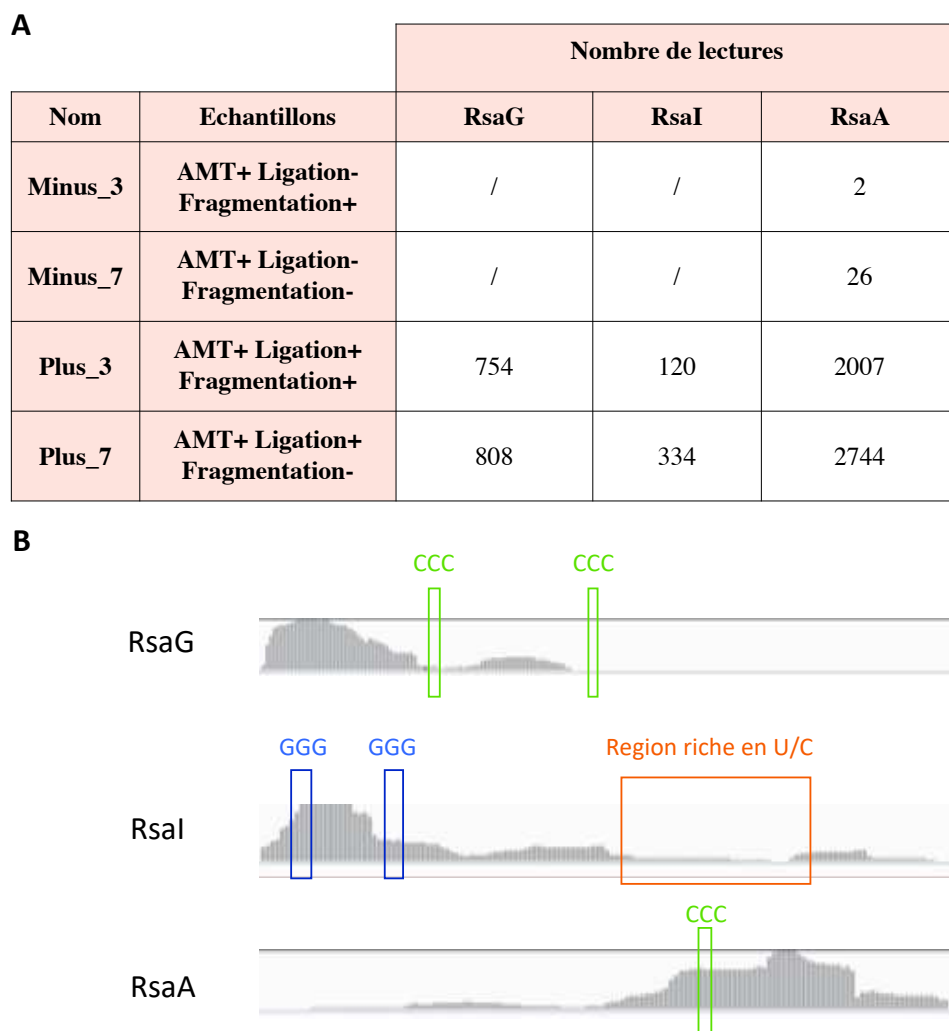


Figure 30 : Représentation des lectures alignées dans la séquence des ARN régulateurs (sARN) RsaG, RsaI et RsaA : A) Tableau représentant pour chaque échantillon, le nombre de lectures alignées sur les sARN B) Représentation de la distribution des lectures pour l'échantillon Plus_7, sur les séquences des sARN RsaG, RsaI, et RsaA. La courbe en gris représente les lectures le long des séquences. Les motifs importants pour les mécanismes de régulation de ces sARN sont représentés en vert pour les motifs riches en C (CCC), en bleu pour les motifs riche en G (GGG) et en orange un motif riche en U et C (Région riche en U/C) (*cf* texte pour les détails).

possible est que les séquences lues retrouvées correspondent majoritairement à des interactions intramoléculaires de RsaG, car celles-ci se répartissent sur une tige-boucle de RsaG. En effet, cette tige-boucle peut produire une chimère après le protocole DBRI, cependant au cours de l'analyse bio-informatique des lectures « *paired-end* », ces séquences seront reconstituées comme des séquences uniques non chimériques. De manière intéressante, ce phénomène n'est pas retrouvé pour RsaA qui présente également une tige-boucle à son extrémité 5'. Enfin, pour RsaI, ces motifs fonctionnels consistent en deux motifs riches en G, qui interagissent avec d'autres sARN et d'une région riche en U/C qui s'apparie avec la séquence SD des ARNm cibles de RsaI. Peu de séquences lues ont été retrouvées au niveau de la région riche en U/C, mais un grand nombre est présent au niveau des motifs riches en G. Cette région est également structurée en tige-boucle. Ce résultat pourrait refléter comme pour RsaG la formation de chimères intramoléculaires, ou bien encore des chimères intermoléculaires entre RsaI et ces cibles sARN (RsaG, RsaD et RsaE), mais qui n'ont pas été assignées au cours de l'analyse bio-informatique. Il sera maintenant intéressant pour chaque sARN de regarder si des séquences se retrouvent au niveau de leurs cibles primaires et la distribution de celles-ci.

3.2.6.3. Conclusion

En résumé, ce protocole DBRI nous a permis de découvrir des chimères entre sARN et des chimères sARN-ARNm, qui sont les chimères qui nous intéressent pour définir l'interactome de l'ensemble des sARN de *S. aureus*. Cependant, nous sommes encore loin d'avoir répertorié l'ensemble des interactions réalisées par les sARN, dans la mesure où nous ne retrouvons pas les interactions connues impliquant les sARN étudiés au laboratoire et leurs cibles validées. Bien qu'elles soient en faible quantité, cette étude préliminaire a permis de révéler de nouvelles interactions sARN-ARNm qui devront être validées expérimentalement. Nous avons montré qu'avec la ligation, nos échantillons sont sujets à un faible pourcentage de circularisation. Pourtant il est préférable de continuer à comparer pour une même expérience des échantillons avec et sans fragmentation. De plus, nous avons établi qu'un grand nombre d'interactions fait intervenir des ARNt ou des ARNr, il conviendra donc d'essayer d'améliorer la ribodéplétion de nos échantillons pour mieux interpréter les données reliées aux sARN. Cependant il n'est pas à négliger que certaines des chimères obtenues entre sRNA et tRNA ou sRNA et rRNA puissent refléter une nouvelle fonction des sRNA pour le contrôle de qualité des ARNt et rRNA, ou bien encore pour faciliter l'assemblage des ribosomes.

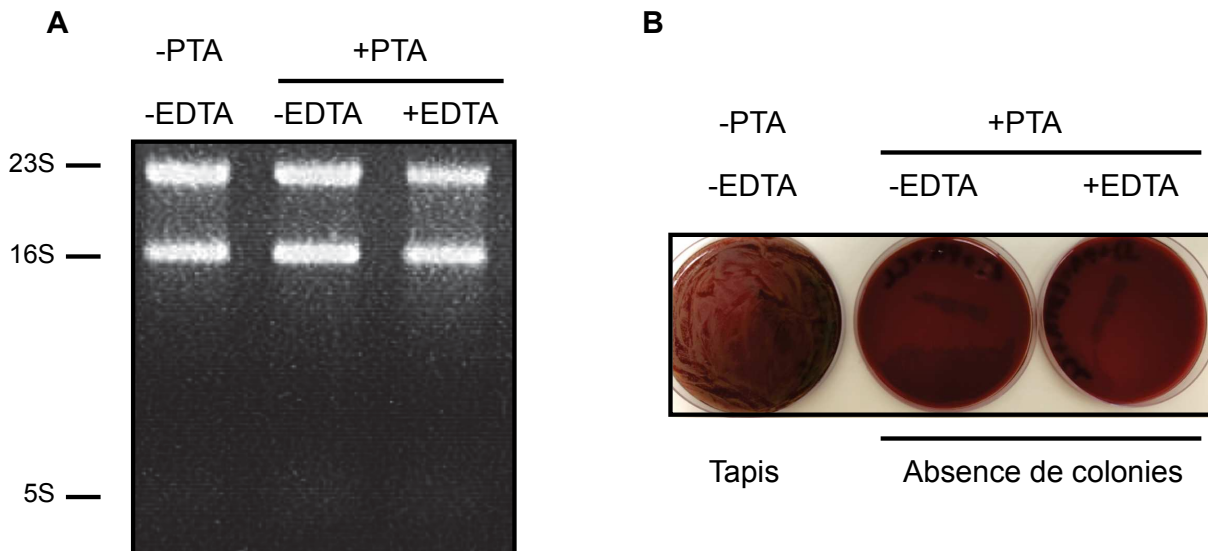


Figure 31 : Contrôle de la qualité des ARN ribosomiques et de la survie des bactéries en fonction des différents échantillons : A) Les ARN sont extraits de différentes cultures de *S. aureus* HG001 à 6h de croissance, sont séparés sur un gel d'agarose 1% TBE 1X, BET 0,001%, 20 mM guanidium thiocyanate, puis révélés par expositions aux UV. De gauche à droite les pistes correspondent à une culture non traitée au PTA et sans EDTA (-EDTA), une culture traitée avec 500 µg/mL de PTA (+PTA) sans EDTA (-EDTA), et une culture traitée avec 500 µg/mL de PTA (+PTA) avec 50mM d'EDTA (+EDTA). À gauche sont nommés les trois ARN ribosomiques 23S, 16S et 5S. B) Photo de géloses au sang (VWR) étalées avec les cultures bactériennes suite aux différents traitements.

3.2.7. Variante 5

Pour diminuer la non-spécificité de l'étape de ligation, j'ai ajouté une étape de pêche au protocole DBRI. Un dérivé du psoralène, le psoralène TEG-azide (PTA) a la particularité de permettre grâce à une réaction chimique (chimie « click »), la liaison covalente d'un résidu biotine DIBO alkyne par une réaction entre les résidus azide et alkyne de chaque composé. Une fois le PTA biotinylé, les ARN pontés *in vivo* sont pêchés grâce à des billes magnétiques couplées à la streptavidine. Seules les interactions intercalées avec le PTA sont récupérées par chromatographie d'affinité. Par ailleurs, la ligation à haute température avait montrée être plus spécifique (*cf* II.3.2.6), c'est pourquoi nous l'avons maintenue dans ce nouveau protocole. Le temps d'irradiation a été raccourci à 15 minutes.

Des contrôles additionnels ont été ajoutés au cours de cette optimisation :

- L'efficacité de pénétration du PTA dans les cellules a été quantifiée. Ainsi 1 mL de chaque surnageant après centrifugation a été récupéré pour effectuer une mesure de la DO_{360nm} dans une cuvette en quartz, afin de définir si le PTA est encore abondant dans le milieu extracellulaire. La condition non traitée au PTA est utilisée comme témoin négatif. J'ai alors comparé une condition avec PTA seul ou avec PTA supplémenté d'EDTA (l'EDTA a été montré comme facilitant l'entrée du psoralène, chez *E. coli*, communication personnelle de Pr. Bjoern Voss). La mesure de l'absorbance à DO_{360nm} a montré que le PTA pénètre de manière équivalente dans les *staphylocoques* en présence ou en absence d'EDTA suggérant que le PTA pénètre différemment dans *S. aureus* et *E. coli*, sans doute résultant de la structure différente de la paroi entre les bactéries à Gram positif et Gram négatif.
- Un contrôle de viabilité des bactéries a été effectué, pour comparer les conditions sans PTA et sans irradiation avec la condition avec PTA et irradiation (+/-EDTA). Cette expérience a révélé que le traitement au PTA suivie de l'irradiation bien que réduite à 15 minutes entraînait la mort des bactéries. La capacité des bactéries à former des colonies n'ayant pas été testée entre le traitement PTA et l'irradiation, nous ne pouvons pas attribuer à l'une ou l'autre des étapes, la mortalité observée (Figure 31B).
- La qualité des ARN a été contrôlée sur gel d'agarose. La visualisation des ARNr a montré que l'irradiation ne semble pas altérer leur qualité et n'entraîne pas de dégradation (Figure 31A).

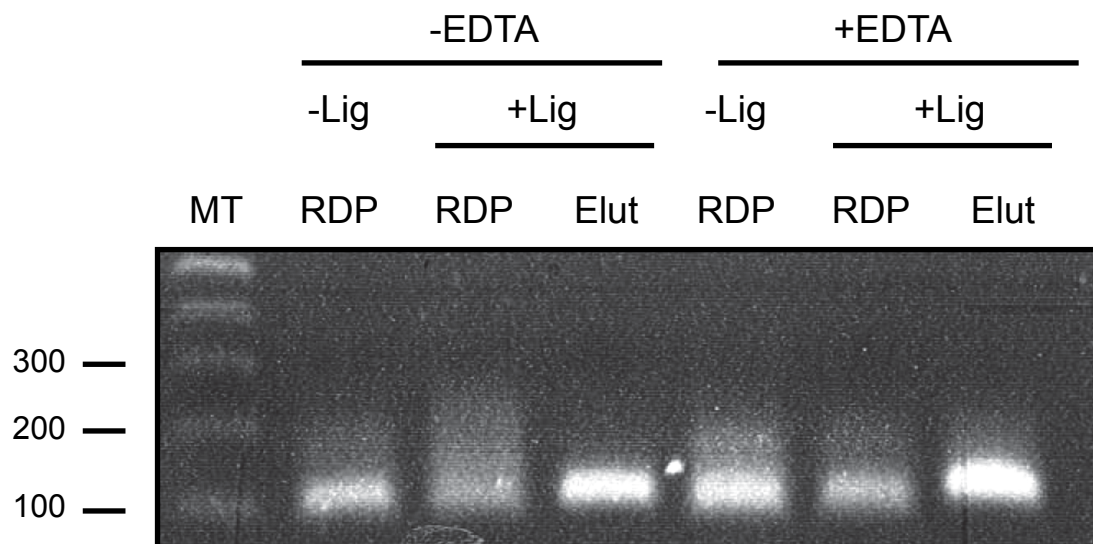


Figure 32 : Amplification des chimères *tcaR-RsaG* par RT-PCR : après un protocole DBRI variante 5 sur gel d'agarose 1% TBE 1X, BET 0,001%, révélé par expositions aux UV. À gauche du gel un marqueur de taille (MT) est migré en parallèle des échantillons, les graduations d'intérêt (100, 200, 300) sont indiquées et correspondent à un nombre en paires de bases. Les trois premières pistes n'ont pas été traitées à l'EDTA (-EDTA), les trois dernières sont traitées à l'EDTA (+EDTA). Les pistes 2, 3, 5, et 6 correspondent à des échantillons ligués (+Lig), les pistes 1 et 4 sont un contrôle sans ligation (-Lig). Les pistes 1, 2, 4, 5 correspondent à des échantillons qui ont subi une réversion directe du pontage (RDP). Les pistes 3 et 6 correspondent à des échantillons qui ont subi une élution avant la réversion du pontage (Elut).

A la fin de la réalisation de cette variante, la quantité d'ARN total après réversion du pontage était très faible (environ 1 ng ou inférieur, comparé à 10-20 ng lors de toutes les variantes précédentes). Cependant, grâce aux faibles quantités nécessaires pour réaliser une RT-PCR, l'expérience a tout de même permis une amplification (Figure 32). Pourtant, la présence de chimères entre *tcaR* et *RsaG* n'a pas été détectée par séquençage des produits PCR testés. La ligation a été réalisée à haute température sur des ARN fixés à des billes via une interaction streptavidine-biotine. Une explication serait que la ligation à haute température a décroché certains complexes des billes magnétiques, menant à une perte majeure de matériel. De plus comme précédemment (*cf* II.3.2.6), l'amplification visible sur gel après RT-PCR est certainement due à une amplification aspécifique n'impliquant pas le couple *tcaR*/*RsaG*.

3.2.8. Variante 6

Pour cette variante, j'ai voulu remédier à l'incompatibilité d'une réaction de ligation à haute température avec une chromatographie d'affinité biotine-streptavidine en pratiquant une ligation à faible température avec la T4 RNA ligase 1 (NEB) à 16°C. Bien que celle-ci soit effectuée dans des conditions moins stringentes permettant potentiellement la ligation d'ARN-ARN sans PTA, ces ARN ne seront pas présents ou seulement en faible quantité suite à la purification sur les billes. De plus, les irradiations de pontage et de réversion du pontage ont été réalisées avec le Vari-X-link, qui permet une diminution drastique du temps d'exposition aux UV : 22 ou 105 secondes à 365 nm contre 15 ou 30 minutes avec le stratalinker.

Les contrôles ont permis de confirmer que l'irradiation n'entraîne pas une dégradation des ARN en utilisant le Vary-X-link quelle que soit la quantité d'énergie de l'irradiation (Figure 33A). En revanche, l'irradiation est responsable de la mort d'une partie de la population bactérienne et non le traitement au PTA. Cette information est rassurante, car premièrement en comparant les figures 33B et 31B, nous observons que l'irradiation au Vary-X-link semble moins nocive que celle avec le stratalinker. De plus, cela nous permet de conclure que le PTA ne semble pas affecter la viabilité de la population bactérienne (Figure 33B).

Considérant les faibles quantités de matériel récupérées à la fin de la variante précédente, j'ai également voulu définir la meilleure condition de réversion du pontage. Nous avons pu définir que la condition de réversion du pontage avec une élution suivie d'une irradiation à 254 nm permet de récupérer une plus grande quantité de matériel (quantité d'ARN 6 à 11 fois supérieure) que la condition avec une réversion du pontage directe (*cf* II.3.1.3.9).

Ce protocole a permis une amplification par RT-PCR dans toutes les conditions, également en absence de ligation et quelle que soit la longueur d'onde (Figure 34). Sont

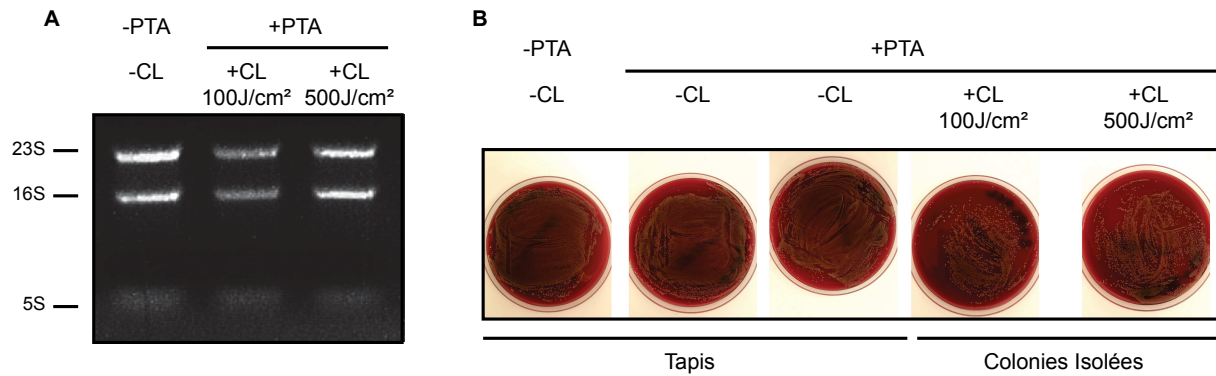


Figure 33 : Contrôle de la qualité des ARN ribosomiques et de la survie des bactéries en fonction des différents échantillons : A) Les ARN totaux sont extraits à partir de différentes cultures de *S. aureus* HG001 à 6h de croissance, sont séparés sur un gel d'agarose 1% TBE 1X, BET 0,001%, 20 mM guanidium thiocyanate, puis révélés par exposition aux UV. De gauche à droite les pistes correspondent à des ARN d'une culture bactérienne non traitée au PTA (-PTA) et sans irradiation (-CL), une culture traitée avec 350 µg/mL de PTA (+PTA) et une irradiation à 100 J/cm² (+CL 100 J/cm²), et une culture traitée avec 350 µg/mL de PTA (+PTA) et avec une irradiation à 500 J/cm² (+CL 500 J/cm²). À gauche sont nommés les trois ARN ribosomiques 23S, 16S et 5S. B) Photo de géloses au sang (VWR) étalées avec les cultures bactériennes suite aux différents traitements : sans PTA (-PTA) ou +PTA (350 µg/mL) (+PTA), sans irradiation (-CL), avec une irradiation à 100 J/cm² (+CL 100 J/cm²), à 500 J/cm² (+CL 500 J/cm²).

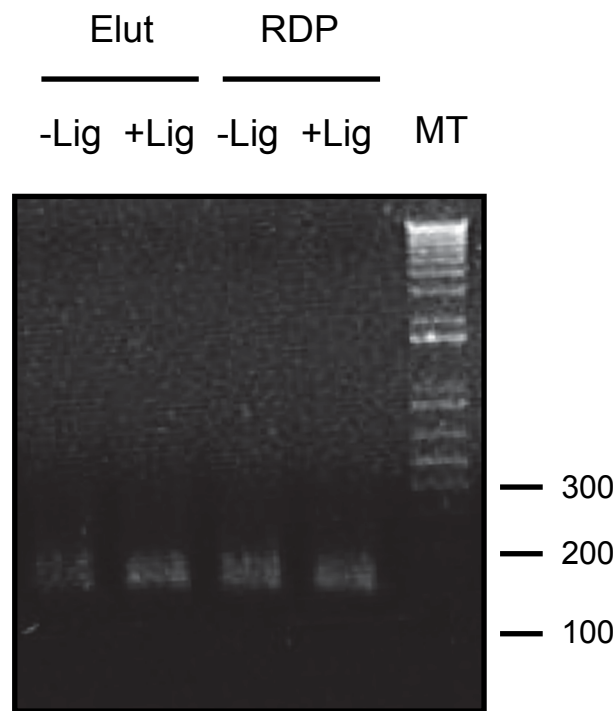


Figure 34 : Amplification des chimères *tcaR-RsaG* : par RT-PCR après un protocole DBRI variante 6 sur gel d'agarose 1% TBE 1X, BET 0,001%, révélées par expositions aux UV. À droite du gel un marqueur de taille (MT) est migré en parallèle des échantillons, les graduations d'intérêt (100, 200, 300) sont indiquées et correspondent à un nombre en paires de bases. Les deux premières pistes correspondent à des échantillons qui ont subi une élution avant la réversion du pontage (Elut). Les deux dernières correspondent à des échantillons qui ont subi une réversion directe du pontage (RDP). Les pistes 2 et 4 correspondent à des échantillons ligués (+Lig), les pistes 1 et 3 sont un contrôle sans ligase (-Lig).

présentés dans la figure 34, seulement les résultats de RT-PCR obtenus après irradiation à 100 J/cm² (22 secondes), les résultats à 500 J/cm² (105 secondes) étant identiques. L'analyse des séquences a montré de manière très surprenante des chimères *tcaR*/RsaG avec ou sans ligation (de même les profils des chimères sont identiques aux deux irradiations testées). Cependant les chimères amplifiées en absence de ligation correspondent seulement à la séquence des oligonucléotides utilisés dans la réaction de PCR (+/- 1 à 3 nucléotides de la séquence des ARN après l'oligonucléotide) (Figure 35A et C). En revanche les chimères amplifiées en présence de ligation font intervenir pour *tcaR* et RsaG des séquences plus longues que la séquence des oligonucléotides utilisés lors de la PCR (Figure 35B et D).

Ce protocole n'a donc pas donné de résultats probants, contrairement aux anciens protocoles utilisés qui en absence de ligation, après séquençage ne révélaient pas de chimères *tcaR*/RsaG, même courtes.

3.2.9. Variante 7

Cette variante résume toutes les améliorations testées au cours de mes expériences et sera testée au laboratoire. Dans la mesure où la variante 6 n'a pas donné les résultats escomptés, j'ai décidé de revenir à la ligation à haute température (*cf* variantes 4.1 et 4.2) couplée à l'enrichissement des ARN pontés grâce à la chromatographie biotine-streptavidine. L'ordre des différentes étapes a dû être modifié. Ainsi les ARN seront purifiés puis élués avant de procéder à l'étape de ligation. Ci-dessous, je présente la suite des étapes avec le numéro de la variante dans laquelle le protocole a été précédemment optimisé et entre parenthèses, le paragraphe dans lequel la méthode est décrite :

- 1) Culture bactérienne : 1.1 (II.3.1.1)
- 2) Traitement au psoralène : 6 (II.3.1.3.1)
- 3) Pontage par irradiation : 6 (II.3.1.3.2)
- 4) Extraction des ARN : 1.1 (II.3.1.2.2)
- 5) Traitement DNase : 5 (II.3.1.3.3)
- 6) Réaction de chimie « click » : 5 (II.3.1.3.4)
- 7) Purification sur les billes magnétiques de streptavidine : 6 (II.3.1.3.5)
- 8) Digestion par la nucléase S1 sur les billes : 6 (II.3.1.3.6)
- 9) Éluion : 6 (II.3.1.3.9.2)
- 10) Adénylation : 4.1 (II.3.1.2.6.3.1)
- 11) Ligation : 4.1 (II.3.1.2.6.3.2)

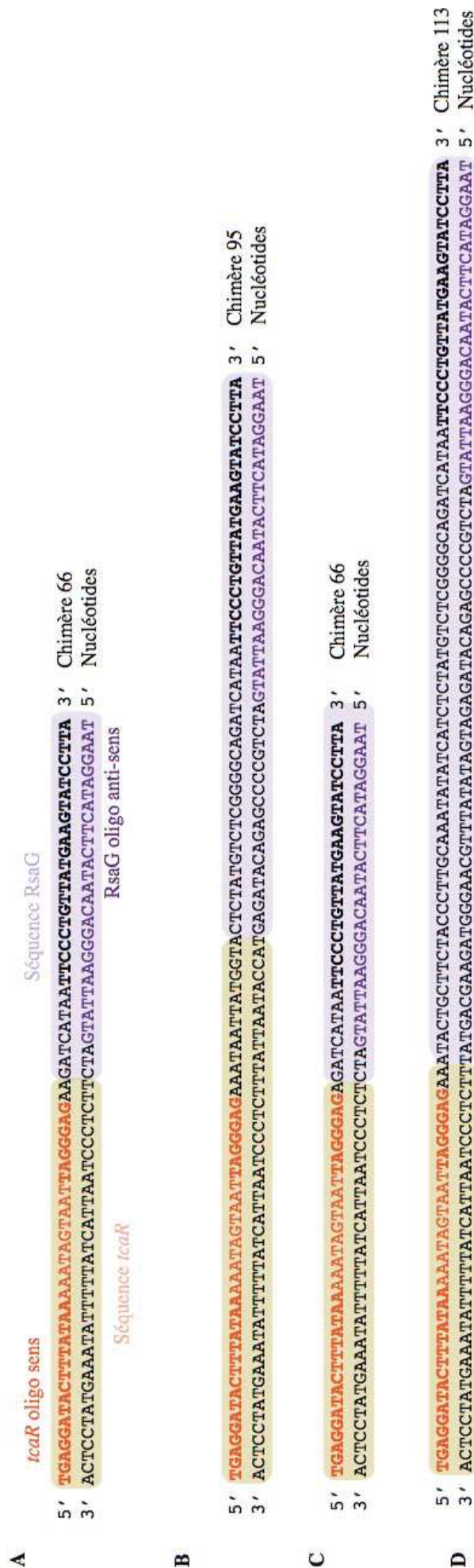


Figure 35 : Résultats du séquençage des fragments amplifiés par RT-PCR des chimères *tcaR*-*RsaG* en fonction de la présence de ligase et de la réversion du pontage effectué (variante 6) : Les nucléotides surlignés en orange correspondent à la séquence de *tcaR*, ceux surlignés en violet à la séquence de *RsaG*. L'oligonucléotide utilisé pour l'amplification spécifique de *tcaR* est écrit en orange, celui écrit en violet pour *RsaG*, en gras les nucléotides de *tcaR* et *RsaG* prédits pour être en interaction. A) Chimère obtenue avec le protocole de la variante 6 sans ligase et avec élution avant la réversion du pontage. B) Chimère obtenue avec ligase et avec l'élution avant la réversion du pontage. C) Chimère obtenue sans ligase et avec une réversion directe du pontage. D) Chimère obtenue avec ligase et avec une réversion directe du pontage.

- 12) Réversion du pontage : 6 (II.3.1.3.9.2)
- 13) RT-PCR : 1.1 (II.3.1.4.1.1)
- 14) Clonage des fragments de PCR : 1.2 (II.3.1.4.1.2)

3.3. Conclusion

L'optimisation du protocole de DBRI a permis de définir, suite au séquençage à haut débit de la variante 4.2, que des conditions de ligation à haute température permettent de former des chimères entre des ARN qui interagissent de manière spécifique *in vivo*. Ces couples d'ARN devront être validés *in vitro* par la technique de gel retard afin de prouver que les interactions qui en résultent sont réelles. Par la suite, il sera intéressant de déterminer la fonction des nouvelles interactions découvertes par DBRI. L'étape de chromatographie d'affinité permet d'enrichir les interactions ARN-ARN fixées par pontage. Ainsi la variante 7 qui concilie toutes ces améliorations est prometteuse, bien que certains paramètres tels que la ribodépletion et l'analyse bio-informatique pourraient encore être améliorés.

III. Discussion

1. Coexistence des sARN dans les réseaux de régulation

1.1. Construction d'un modèle global de régulation impliquant plusieurs sARN dépendant de l'environnement de *S. aureus*

L'application de la technique du MAPS au laboratoire a permis d'établir les targetomes de plusieurs sARN, à savoir RsaA, RsaD RsaC, ainsi que RsaI et RsaG, réalisés au cours de ma thèse (Tomasini et al., 2017, thèse de Delphine Bronesky, Bronesky et al., 2019; Lalaouna et al., 2019). Ceci me permet de proposer un modèle global de régulation présenté dans la Figure 36 qui est centré sur RsaG, mais qui intègre aussi d'autres données permettant de relier la fonction de plusieurs sRNAs.

Au cours de ma thèse j'ai ainsi montré que, contrairement à ce qui avait été proposé, RsaG n'est pas un sARN classique, mais est issu de la dégradation de l'ARNm *uhpT* en réponse à la présence de G6P extracellulaire et suite à l'activation par le TCS HptRS. Cet ARN étend le régulon de HptRS et permet, entre autres, d'activer le métabolisme des acides aminés, d'inhiber la fermentation lactique directement en ciblant *ldh1*, et indirectement en activant la traduction de *Rex* et en réprimant celle de *CcpA* (Crooke et al., 2013; Pagels et al., 2010; Desgranges et al., en préparation). De manière intéressante, il a été montré que *CcpA* et HptRS partagent la régulation de nombreux gènes, mais de manière antagoniste, en particulier *CcpA* réprime la transcription de *uhpT* (Reed et al., 2018). Ainsi, en réprimant la traduction de *CcpA*, RsaG crée un « *feedback loop* » contribuant à l'amplification de *uhpT* lorsque le G6P est l'une des seules sources carbonées disponibles. RsaG jouerait aussi un rôle indirect dans la détoxification des ROS en inhibant *SarA*, qui réprime la synthèse de *SodM* (Ballal & Manna, 2009; Desgranges et al., en préparation). RsaG forme aussi des complexes stables avec les ARNm des FT *SarX* et *SarV*, tous deux impliqués dans la régulation de la virulence de *S. aureus*. Cependant, les conséquences de l'interaction de RsaG avec ces deux cibles n'ont pas encore été étudiées. L'analyse comparative de la transcriptomique de la souche *S. aureus* sauvage, délétée de RsaG, ou complémentée par un plasmide exprimant *rsaG*, a révélé peu de changements. En particulier, la synthèse de deux ARNm codant des facteurs de virulence de *S. aureus*, *SdrC* et *Ssp*, serait réprimée par RsaG, cependant de manière indirecte, puisque ces deux ARNm n'ont pas été identifiés par l'approche MAPS. *SdrC* est une protéine de surface avec des régions riches en résidus serine et aspartate répétés. Ce type de protéine contribue à l'adhésion et à la formation de biofilm (Speziale et al., 2014). *Ssp* (ou *Emp*) est une protéine de liaison à la matrice extracellulaire (fibrinogène, fibronectine) et pourrait avoir un rôle dans l'infection des tissus profonds (Geraci et al., 2017). L'effet sur ces facteurs de virulence est

certainement indirect et pourrait résulter de la répression RsaG-dépendante des facteurs de la famille Sar (SarA, SarX, ...). En effet, il a été montré que SarA et le système *agr* activent *ssp* (Harraghy et al., 2005). L'impact de RsaG sur ces régulateurs et sur la virulence de *S. aureus* reste à déterminer. Quoi qu'il en soit, RsaG régule plusieurs régulateurs majeurs de *S. aureus*, qui interviennent dans différents domaines de sa physiologie. Il est donc surprenant que nous n'ayons pas pu mettre en évidence un phénotype particulier de la souche mutante Δ *rsaG*, comme par exemple un effet sur l'hémolyse liée à la régulation de l'ARNIII par SarA (Cheung & Zhang, 2002), un effet sur la formation des biofilms dû à la répression de *tcaR* (Jefferson et al., 2004), un phénotype de croissance en conditions normales ou en présence de stress NO...). Le fait qu'aucun phénotype ne soit lié à l'expression de RsaG peut s'expliquer de plusieurs manières. Premièrement, nous n'avons peut-être pas encore trouvé les bonnes conditions ou les bons phénotypes à évaluer. Les expériences réalisées en culture liquide ne miment pas les conditions complexes du milieu infectieux provoqué par *S. aureus*. Bien que les conditions de culture soient utiles pour définir les mécanismes moléculaires de la régulation, il reste cependant primordial de réaliser des expériences plus proches du contexte infectieux, par exemple par l'utilisation de modèles cellulaires ou animal. Ainsi, nous avons montré que UhpT et RsaG sont fortement exprimés dans deux contextes infectieux : lorsque *S. aureus* est internalisé dans des cellules humaines et lorsque *S. aureus* est en contact avec des cellules épithéliales de poumon ou de colon, toutes deux productrices de mucus. Nous avons par ailleurs montré, en culture liquide, que RsaG, à l'inverse de HptRS, n'est pas essentiel pour la croissance cellulaire en présence de G6P comme source de carbone, suggérant que RsaG n'est pas requis pour le catabolisme du G6P. Il reste à déterminer l'effet de la délétion de RsaG, en particulier sur l'internalisation de *S. aureus* dans divers modèles cellulaires (cf. II.2 Desgranges et al., en préparation). Nous savons d'ores et déjà qu'un mutant Δ *hptRS* survit et se réplique moins bien qu'une souche sauvage (WT) dans des macrophages murins (BMDM), des monocytes humains (THP-1) et des cellules pulmonaires d'origine humaine (A549) (Park et al., 2015). Ces expériences seront réalisées en comparant la survie du mutant Δ *hptRS*, d'une souche délétée du promoteur de *uhpT* (Δ *PuhpT*), et du mutant Δ *rsaG*. Nous pourrions alors distinguer si ce défaut de survie résulte de l'absence d'expression de UhpT, empêchant l'entrée du G6P, ou si l'absence de RsaG dérégulerait certaines fonctions nécessaires à la survie intracellulaire. Sachant que *S. aureus*, en contexte d'internalisation, peut utiliser des sources de carbone alternatives telles que les acides aminés, mon hypothèse irait plutôt dans le sens que *S. aureus* devrait assurer sa croissance avec d'autres nutriments disponibles que le G6P (Palma Medina

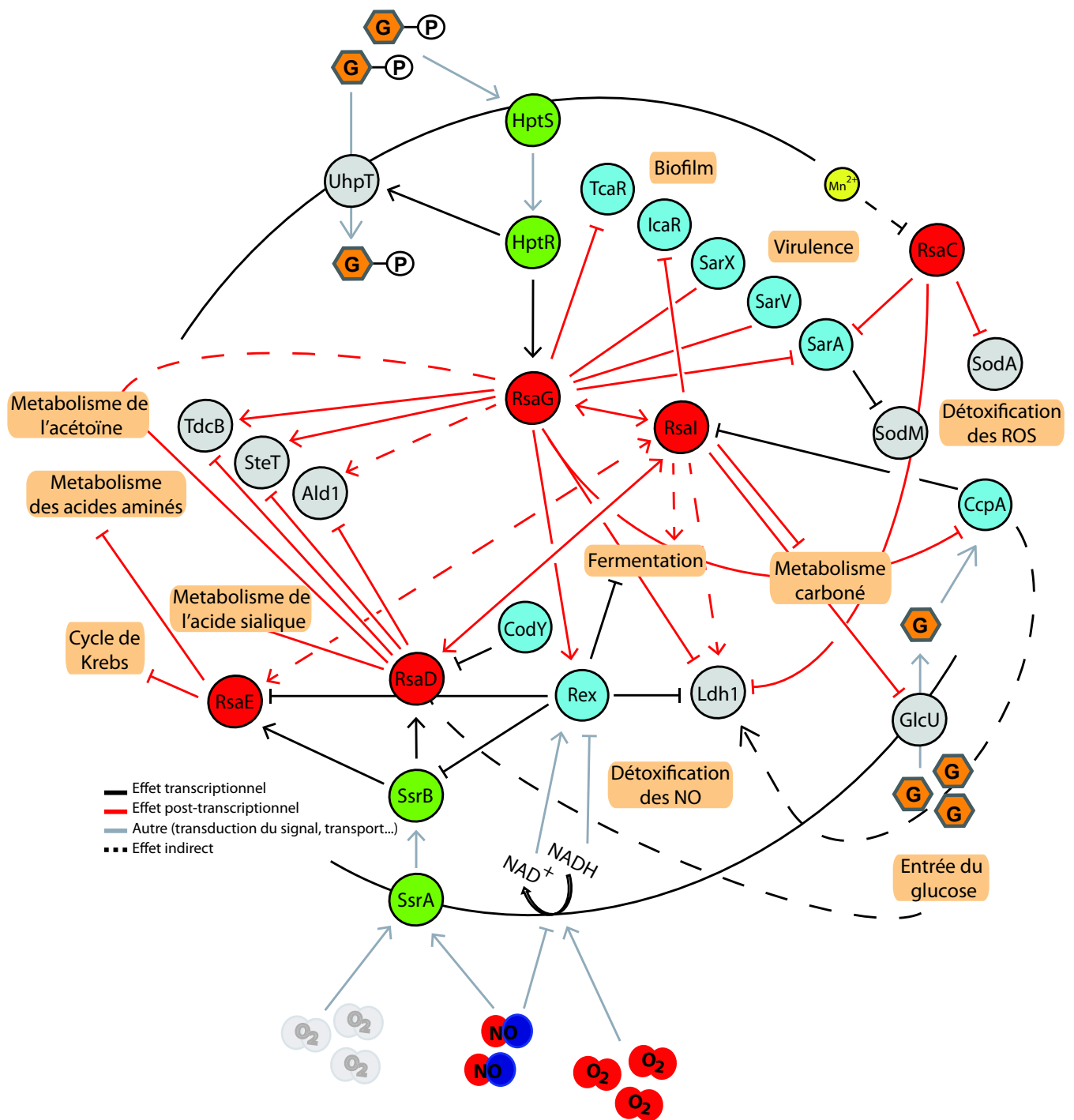


Figure 36 : Modèle de régulation entre quatre ARN régulateurs connectant le métabolisme et la réponse aux stress NO et ROS (*Reactive oxygen species*) : En rouge les sARN, en bleu les facteurs de transcription (FT), en vert les systèmes à deux composants (TCS) et en gris les enzymes et transporteurs. Dans un souci de simplification du modèle, les différents noms de gènes sont uniformément écrits, bien qu'en fonctions du type de régulation subi, ils peuvent faire référence à un ARNm, une protéine ou les deux. L'expression des différents ARN, FT et TCS est régulée en fonction de la présence de manganèse (Mn^{2+}), de glucose (hexagones orange), de glucose-6-phosphate (G6P, hexagone orange avec rond), des radicaux d'oxyde nitrique (NO), et de la présence ou absence d'oxygène (O_2 respectivement rouge ou gris). Les détails sont expliqués dans le texte.

et al., 2019). Des expériences préliminaires ont été réalisées en collaboration avec l'équipe du Prof. Vandenesch (CIRI, Lyon) pour comparer la survie en internalisation des souches WT, Δ rsaG, Δ hptRS et Δ rsaG-CC (souche délétée de RsaG complémentée par un plasmide exprimant RsaG en trans). Ces expériences ont été réalisées en suivant strictement le protocole décrit dans Park et al. (2015) et appliquées à des macrophages murins (RAW) et des cellules pulmonaires humaines (A549). De manière surprenante, les résultats n'ont montré aucune différence significative de survie entre les quatre souches de *S. aureus*, et ce sur les deux types de cellules étudiées. Comme nous ne sommes pas parvenus à reproduire le phénotype observé par Park et al. (2015), il se pourrait qu'une des différences majeures provienne des souches utilisées. En effet, nous utilisons la souche de laboratoire HG001, alors que l'étude de Park et al. (2015) utilise la souche SARM USA300 qui est hautement pathogène. Nous ne pouvons donc pas exclure le fait que HptRS présente des cibles additionnelles dans la souche USA300, et que la régulation de ces cibles supplémentaires ait un effet sur la survie intracellulaire de *S. aureus*. Il sera donc important, dans un premier temps, de renouveler les expériences d'internalisation dans le contexte de la souche HG001 en y ajoutant la souche mutante Δ Puht. Si nous ne parvenons toujours pas à reproduire les résultats de Park et al. (2015), nous envisagerons d'analyser la fonction de UhpT/RsaG dans le contexte de la souche MRSA USA300. Une deuxième raison, qui pourrait expliquer notre incapacité à trouver un phénotype lié à l'expression de RsaG, est que comme le montre la figure 36, RsaG est impliqué dans un réseau de régulation complexe faisant intervenir un grand nombre de régulateurs transcriptionnels et post-transcriptionnels. Nous pouvons alors imaginer que l'impact de RsaG dans ce réseau serait minime. En effet les sARN sont impliqués dans des régulons sophistiqués, et la perte d'un sARN individuel peut être compensé en raison de la redondance fonctionnelle entre certains régulateurs. RsaG pourrait permettre une régulation subtile masquée par d'autres régulateurs dans les conditions étudiées. Cette hypothèse rejoint la conclusion apportée par l'étude de Gao et al. (2020). Ils avaient montré que bien que certains sARN soient surexprimés dans une condition donnée (en présence de linezolide), cela ne se manifeste pas toujours par un phénotype observable (un impact sur la CMI par exemple). En effet, les sARN apportent une régulation fine avec un impact rare sur la croissance cellulaire (Gao et al., 2020). Le réseau complexe dans lequel est impliqué RsaG fait intervenir d'autres sARN, tels que RsaC, RsaD, RsaE et RsaI (Figure 36). RsaG régulerait conjointement avec RsaC la détoxification des ROS, alors qu'avec RsaD il régulerait le métabolisme des acides aminés (cf. ci-après III.1.1.2 et III.1.1.3). Le réseau de RsaG fait également intervenir RsaI, avec qui il interagit de manière

directe, et avec lequel il partage des cibles communes (cf. ci-après III.1.1.1). Ce modèle nous permet de voir que les sARN qui y figurent sont exprimés de manière dynamique en fonction de différents stimuli et ne sont probablement pas tous exprimés en même temps. La redondance des cibles entre ARN régulateurs permet un contrôle fin de l'expression des gènes, mais surtout d'intégrer des signaux multiples. Ce type de circuit est communément retrouvé dans les réseaux de régulation, qu'on appelle « *overlapping regulon* » (cf. I.4.5, figure 14). Dans ce modèle, nous observons également pour chaque sARN, des circuits « *single-input module* », qui résultent du fait que nous avons identifié plusieurs cibles pour chacun de ces sARN. Les cibles dans ces circuits voient leur expression coordonnée dans les conditions d'expression du sARN qui les régule. C'est ainsi par exemple, que RsaG permet de faire concorder l'activation du catabolisme des acides aminés et l'inhibition de la fermentation lactique dans un environnement riche en G6P (cf. III.1.1.2 et III.1.1.3 ci-après). Enfin, nous trouvons également un circuit « *feedforward loop* » composé de RsaG, Rex et Ldh1. Ce type de circuit permet d'intégrer deux signaux différents, le G6P et un potentiel rédox élevé, nécessaires à l'activation de RsaG et Rex, respectivement (cf. II.2 ; Pagels et al., 2010). Ainsi l'expression de Ldh1 est régulée dans deux conditions différentes, et si ces conditions sont réunies, sa régulation est effectuée à deux niveaux, transcriptionnel et post-transcriptionnel. De plus, l'activation de Ldh1 a été montrée comme nécessitant d'une part l'inhibition de *rex* et d'autre part l'activation de CcpA par le glucose (Crooke et al., 2013). Il est donc intéressant que RsaG, qui inhibe *ldh1* de manière directe, ainsi que de manière indirecte par l'activation de *rex*, l'inhibe également potentiellement par l'inhibition de *ccpA* de manière indirecte, puisque nous avons montré que RsaG interagit avec *ccpA* au niveau de sa séquence SD, ce qui mène probablement à l'inhibition de sa traduction.

Très récemment, une étude a regroupé les données de vingt-deux transcriptomes précédemment publiés, en ajoutant l'annotation de trois cent trois sARN connus chez *S. aureus* USA300, afin de révéler les informations non prises en compte sur l'expression de ces sARN. Ces données montrent par exemple une diminution de l'expression de RsaG en conditions anaérobies (Sorensen et al., 2020). Cette étude a également analysé des données de ribosome profiling afin d'identifier les sARN qui auraient la capacité de coder des peptides. Les résultats indiquent que RsaG pourrait être un sARN bifonctionnel en codant un peptide de 25 acides aminés, comme c'est le cas pour l'ARNIII qui code pour l'hémolysine δ (Sorensen et al., 2020). Ces données corroborent une expérience de ribosome profiling utilisant la retapamuline comme antibiotique réalisé au laboratoire (communication personnelle, Stefano

Marzi) pour analyser tous les sites d'initiation de la traduction. La capacité de RsaG à coder un peptide demande des investigations supplémentaires car aucune séquence Shine et Dalgarno évidente n'est présente en amont du codon d'initiation de la traduction de cette petite ORF. Ceci sera d'autant plus important que la région de RsaG qui contient cet ORF potentielle englobe le site de fixation de RsaI, ainsi que le site d'autres ARNm comme *rex*, ce qui voudrait dire que contrairement à l'ARNIII, la traduction de RsaG pourrait potentiellement interférer avec ses fonctions de régulateur.

1.1.1. Interaction entre RsaI et RsaG

RsaI est un sARN inhibé par CcpA lorsque la concentration en glucose est élevée. En revanche, lorsque le glucose est épuisé du milieu, RsaI réprime la traduction de ses cibles ARNm, ce qui s'accompagne d'une diminution de la transcription de gènes impliqués dans le catabolisme du carbone, et une activation de gènes impliqués dans la production d'énergie et la fermentation. RsaI interagit également avec d'autres sARN : RsaG, RsaD (Bronesky et al., 2019 ; cf. II.1) et RsaE (dont la validation est en cours au laboratoire).

Les expériences de gel retard montrent que RsaG n'affecte pas l'interaction de RsaI avec ses cibles (Bronesky et al., 2019). En revanche, RsaI est capable d'empêcher l'interaction entre RsaG et certaines de ses cibles lorsqu'elles interagissent au niveau de son premier motif riche en C (cf. II.2). Ainsi, une prédiction d'appariements entre RsaG et ses ARNm a été réalisée via IntaRNA, afin de déterminer pour chacun la séquence de RsaG impliquée dans l'interaction. Les sites d'interaction de RsaG avec les ARNm *rex*, *fn3k*, *ldh1*, *tdcB* et *treB* ont été validés *in vitro* (cf. II.2, et données personnelles, figure 37). Les cibles de RsaG peuvent ainsi être classées en différentes catégories :

- *Les cibles régulées par RsaG indépendamment de la présence de RsaI.*

RsaG exprimé à un niveau basal ou surexprimé par le G6P régule ses cibles, qui interagissent à son deuxième motif riche en C, ou d'autres sites ne faisant pas intervenir plusieurs C (Figure 37AB, deuxième et troisième colonne). Parmi les cibles du second motif riche en C, la visualisation d'un complexe ternaire par gel retard (interaction de RsaG à la fois avec RsaI et une cible ARNm) a pu être observée avec *sarA*, *tcaR*, *treB* et *ldh1*.

- *Les cibles qui ne devraient pas être régulées par RsaG dans les conditions d'expression de RsaI.*

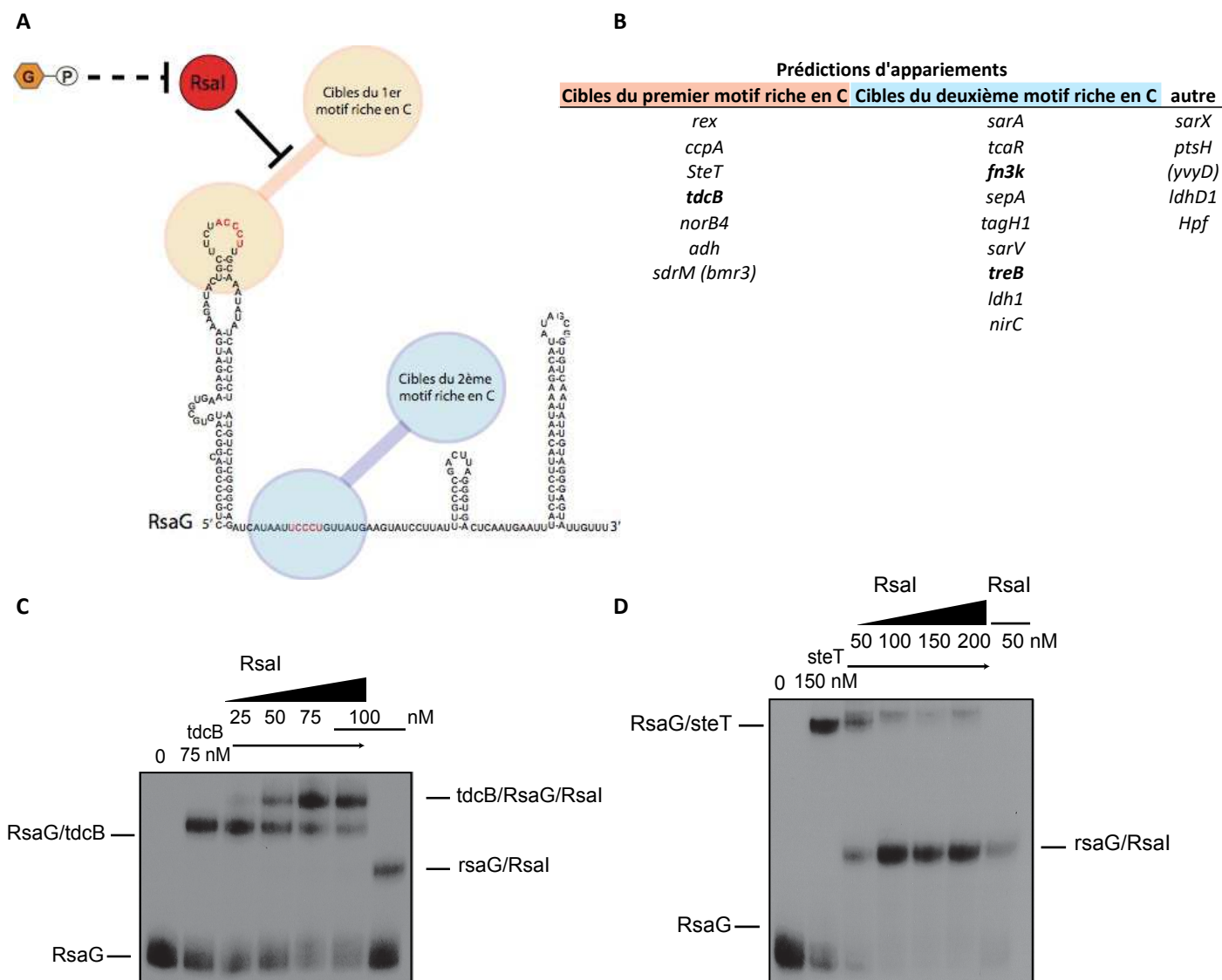


Figure 37 : Hiérarchisation des cibles de RsaG : A) Les cibles de RsaG qui interagissent avec le premier motif riche en C sont représentées en orange , en bleu celles qui interagissent avec le deuxième motif riche en C. RsaI dans un milieu carencé en glucose empêche l'interaction de RsaG et ses cibles se lient au premier motif riche en C alors qu'en présence de glucose-6-phosphate (G6P) RsaG est libre de réguler les cibles se fixant à l'un ou l'autre des deux motifs riches en C. B) Prédictions des sites d'interactions entre RsaG et ses ARNm cibles (IntaRNA). Sont indiqués en gras les cibles communes avec RsaI (*fn3k* et *treB*, validé, *tdcB* non validé). C et D) Gel retard par compétition montrant (C) la formation du complexe ternaire entre RsaG *tdcB* et RsaI, et (D) la compétition entre RsaI et *steT* pour RsaG. RsaG radio marqué en 5' est incubé avec des concentrations croissantes de RsaI (écrites au-dessus du gel en nM) et/ou *tdcB* (75 nM (C)) ou *steT* (150 nM (D)). Les complexes formés dans les différentes pistes sont indiqués sur les bords du gel.

RsaG exprimé à un niveau basal est dans l'incapacité de réguler les cibles qui interagissent à son premier motif riche en C, dans la mesure où RsaI le séquestre (Figure 36AB, première colonne). La réalisation de gel retard a permis de révéler, parmi les cibles du premier motif riche en C de RsaG, que RsaI est capable de rentrer en compétition avec *rex* et *steT*, appuyant ainsi pour ces deux ARNm, le site d'interaction prédit (*cf* II.2, Figure 37D).

- *Les cibles communes à RsaG et RsaI.*

RsaI et RsaG ont pour cible commune *fn3k*, codant pour une fructose amine kinase, une enzyme responsable de la déglycation des protéines (Gemayel et al., 2007). La glycation des protéines est une réaction qui se produit de manière spontanée, alors que la déglycation permet de détoxifier la cellule de ses protéines non fonctionnelles. RsaI inhibe la traduction de l'ARNm *fn3k* par séquestration de la séquence SD, et RsaG s'hybride à la séquence codante déstabilisant l'ARNm (ce qui doit certainement conduire au recrutement de la RNase III, une hypothèse qui reste encore à confirmer). Les sites de liaison des deux sARN étant distincts, RsaI et RsaG se fixent de manière indépendante (Figure 37B). L'ARNm *treB*, codant pour un transporteur du tréhalose, est également une cible commune aux deux sARN. Alors que RsaI inhibe la traduction de *treB*, l'effet de RsaG sur cet ARNm n'a pas encore été étudié en détail. Enfin, *tdcB* pourrait être une cible commune aux deux sARN. Cependant par gel retard, alors que *tdcB* interagit avec RsaG au niveau de son premier motif riche en C (données personnelles), nous observons la formation d'un complexe ternaire RsaG-RsaI-*tdcB* (Figure 37C). L'interaction directe entre *tdcB* et RsaI n'a pas été validée, bien que l'ARNm *tdcB* a été retrouvé co-purifié avec MS2-RsaI. Ainsi, le complexe ternaire pourrait représenter l'interaction directe de RsaI et ses deux cibles, *tdcB* et RsaG.

Mon étude suggère que RsaI est un ARN bi-fonctionnel qui à la fois régule ses propres cibles ARNm, mais également des ARN régulateurs en entravant leurs fonctions régulatrices, et ce par compétition directe lorsque le glucose est consommé. RsaG et RsaI sont tous deux exprimés dans un milieu pauvre en glucose, bien que RsaG soit très faiblement exprimé dans ces conditions en comparaison de RsaI, et en comparaison de son expression dans un milieu contenant du G6P. La différence de niveau d'expression entre ces deux ARN nous laisse à penser que RsaI serait nécessaire pour séquestrer la quantité basale de RsaG présente lorsque RsaI est actif. De plus, le fait que RsaI ne déstabilise pas RsaG (Bronesky et al., 2019), mais

bloque seulement certaines de ses fonctions régulatrices, permet en cas de changement de conditions plus favorables à l'expression de RsaG (G6P), qu'un pool de RsaG soit déjà présent dans la cellule pour induire une réponse rapide. Dans les conditions propices à l'expression de RsaG, en présence de G6P, RsaG peut réguler l'ensemble de ses cibles car le taux de RsaI a fortement diminué (comme le montre le transcriptome *cf.* II.2). Si nous arrivons à établir que RsaI modifie les fonctions régulatrices de RsaG *in vivo*, cela mènera à la découverte d'un nouvel exemple d'ARN éponge (ARNe), qui a la particularité de modifier la régulation de seulement une partie des cibles du sARN ciblé en fonction des conditions environnementales ; ce qui n'a pas encore été décrit à ma connaissance. Chez *E. coli*, un exemple proche existe, il s'agit de la séquestration du sARN GcvB par l'ARNe SroC maturé depuis la région 3'UTR de *gltI* de l'opéron *gltIJKL* codant pour un transporteur ABC d'acides aminés (Miyakoshi et al., 2015). GcvB possède trois sites d'interaction avec ses cibles (R1, R2, R3) et SroC interagit seulement au site R3, pour en théorie rentrer en compétition avec seulement une partie des cibles de GcvB. En revanche, l'interaction entre SroC et GcvB a pour effet de déstabiliser GcvB l'empêchant ainsi de réguler toute ses cibles (Lalaouna et al., 2018). À l'inverse, l'interaction entre RsaI et RsaG n'affecte pas la stabilité de RsaG, permettant à RsaG de réguler une partie de ses cibles bien qu'il soit en complexe avec RsaI (Bronesky et al, 2019). En revanche, nous avons pu établir que RsaG, exprimé de manière basale (en absence de G6P), stabilise le sARN RsaI. En effet, le temps de demi-vie de RsaI chute de 24 à 15 minutes dans une souche Δ *rsaG* (Bronesky et al., 2019, *cf.* II.1). L'interaction entre les deux sARN pourrait modifier la structure de RsaI le rendant plus stable, ou masquer le site de clivage d'une RNase. La formation d'une structure en pseudo-nœud a été prédite dans RsaI (RsaOG) (Marchais et al., 2010). Le deuxième motif en G de RsaI (impliqué dans l'interaction avec RsaG), ainsi que les trois nucléotides en amont, seraient en interaction avec sept nucléotides de la région riche en U/C de RsaI (montré comme impliquée dans la régulation de ces cibles ARNm) pour permettre la formation de ce pseudo-nœud. Cependant, RsaG, en interagissant avec RsaI, devrait avoir pour conséquence d'empêcher la formation du pseudo-nœud, permettant de rendre accessible la région riche en U/C de RsaI pour la régulation de ses cibles ARNm. La structure pseudo-nœud avait été proposée comme pouvant contraindre la structure de RsaI afin de protéger sa région inter-hélicoïdale non appariée. Cependant nos résultats contredisent cette hypothèse dans la mesure où RsaI est plus stable en présence de RsaG, où cette structure ne devrait plus pouvoir se former (Marchais et al., 2010). Nous pouvons émettre l'hypothèse que RsaG, lorsque son expression est concomitante à celle de RsaI (en absence de glucose), permet d'amplifier la régulation

exercée par RsaI sur ses cibles. En effet, nous pouvons imaginer que la formation d'un complexe ternaire entre RsaG, RsaI et les cibles de RsaI, permet, premièrement d'entraver certaines fonctions régulatrices de RsaG, et deuxièmement de stabiliser RsaI, lui permettant de maintenir une concentration cellulaire élevée pour la régulation de ses cibles ARNm, et une structure ouverte sans pseudo-nœud.

Pour définir l'impact de RsaI sur la régulation des cibles de RsaG *in vivo*, l'approche suivante a été testée au cours de ma thèse. Je voulais suivre *in vivo* l'effet de la compétition entre RsaI et l'ARNm *rex*, cibles se liant au premier motif riche en C de RsaG. Une fusion traductionnelle entre la région 5' de *rex* et le gène rapporteur *LacZ* a été insérée dans les mutants Δ *rsaI*, Δ *rsaG* ou le double mutant Δ *rsaI-rsaG*, et l'activité β -galactosidase mesurée. Cependant, aucune différence n'a été observée entre les différentes souches. Néanmoins, la fusion traductionnelle *rex-LacZ* est peu active, rendant difficile l'interprétation des données (*cf.* II.2).

En perspective, la technique DBRI (*cf.* II.3) permettrait de comparer les chimères résultantes d'une culture de *S. aureus* sauvage (WT) et d'une culture de *S. aureus* Δ *rsaI*. Ainsi, dans la souche Δ *rsaI*, nous devrions retrouver en chimère avec RsaG, toutes ses cibles ARN. Nous aurons également l'information du motif riche en C de RsaG engagé dans chaque interaction. À l'inverse, en utilisant la souche WT, RsaI empêchera l'interaction de RsaG avec certaines de ses cibles, entraînant une absence des chimères résultant de ces interactions. Normalement, nous devrions observer une corrélation entre les cibles du premier motif riche en C de RsaG, obtenu avec la souche Δ *rsaI*, et les cibles de RsaG non retrouvées avec la souche WT. Pour contrôler le fait que RsaG n'interfère pas dans les mécanismes de régulation de RsaI (Bronesky et al., 2019), nous devrions retrouver les mêmes ARNm en chimère avec RsaI en utilisant l'expérience DBRI réalisée avec la souche mutante Δ *rsaG*, en comparaison de la souche WT. Ceci devrait confirmer que les interactions entre RsaI et ses cibles ARNm ne sont pas affectées par RsaG.

La construction de fusions traductionnelles à partir des ARNm cibles, confirmées au premier motif riche en C de RsaG, permettra de valider l'effet de RsaG. De plus, en exprimant les fusions traductionnelles dans les souches WT, Δ *rsaI*, Δ *rsaG* ou le double mutant Δ *rsaI-rsaG*, nous devrions pouvoir définir si RsaI entrave les fonctions régulatrices de RsaG. Ces approches confirmeront si RsaI joue un rôle d'éponge atypique en dérégulant une partie du régulon de RsaG en fonction des conditions de l'environnement (Denham, 2020).

Enfin, non seulement RsaG présente des cibles communes avec RsaI, mais aussi une cible impliquée dans la régulation du même processus, *i.e.* la formation de biofilm. En effet, RsaI régule le FT IcaR en se liant dans la région 3'UTR et, en conséquence, induit la production de PIA-PNAG par un mécanisme encore inconnu (*cf.* II.1 ; Bronesky et al., 2019). De manière intéressante, RsaG régule également un FT impliqué dans la formation de biofilm, *tcaR* (*cf.* II.2). IcaR et TcaR sont tous deux des répresseurs de la transcription de l'opéron *icaADBC*, responsable de la synthèse de PIA-PNAG (Figure 36). Cependant, IcaR est un régulateur plus fort et masque l'effet de TcaR dans une souche WT, l'effet de TcaR n'est donc observable que dans un mutant *icaR*, où il amplifie son phénotype (*cf.* I.3.1.2.1.2 ; Jefferson et al., 2004). RsaG inhibe la formation du complexe d'initiation de la traduction sur l'ARNm *tcaR* en séquestrant la séquence SD avec son deuxième motif riche en C. RsaI n'est donc pas capable d'empêcher l'interaction entre RsaG et *tcaR* (Figure 37B, *cf.* III.1.1.1). Il serait donc possible dans les conditions d'expression conjointe de RsaI et RsaG de pouvoir observer l'effet amplificateur de *tcaR* dans le contexte où *icaR* est inhibé par RsaI. Il serait ainsi intéressant pour confirmer cette hypothèse de quantifier la production de PNAG dans les souches WT, Δ *rsaI*, Δ *rsaG* et double mutante Δ *rsaI-rsaG* (déjà disponible au laboratoire).

1.1.2. Redondance possible des cibles entre RsaG et RsaD

RsaD est régulé à différents niveaux, entre autres, par le système à deux composants SrrAB (Bronesky et al, 2019) et par le facteur de transcription CodY (Augagneur et al., 2020) (*cf.* I.5.3.1.2). Comme RsaG, RsaD est capable de former une interaction directe avec le sARN RsaI (Bronesky et al., 2019). La fonction de cette interaction n'est pas élucidée.

Des résultats non publiés de MAPS ont révélé que RsaD est co-purifié avec les transcrits de l'opéron *aldI-tdcB-steT-norB4* (manuscrit de thèse Delphine Bronesky). Le gène *tdcB* code pour une enzyme qui catabolise la thréonine, SteT est impliqué dans le transport de la thréonine et de la sérine, ainsi que des acides aminés aromatiques (mais de manière moins efficace), et *aldI* code pour une enzyme qui catabolise l'alanine pour produire du pyruvate. Enfin *norB4* code pour une pompe à efflux qui apporte la résistance aux quinolones. De plus, une analyse différentielle du transcriptome de *S. aureus* HG001 sauvage et délété de *rsaD*, qui requiert une validation, a montré que RsaD diminue l'abondance des transcrits *aldI*, *steT* et *tdcB* (résultats non présentés et Figure 38). Comme décrit précédemment (*cf.* I.5.3.1.2), le répresseur de RsaD, CodY, régule un grand nombre de gènes impliqués dans l'import et la biosynthèse des acides aminés, mais l'opéron *aldI-tdcB-steT-norB4* ne fait pas partie des cibles

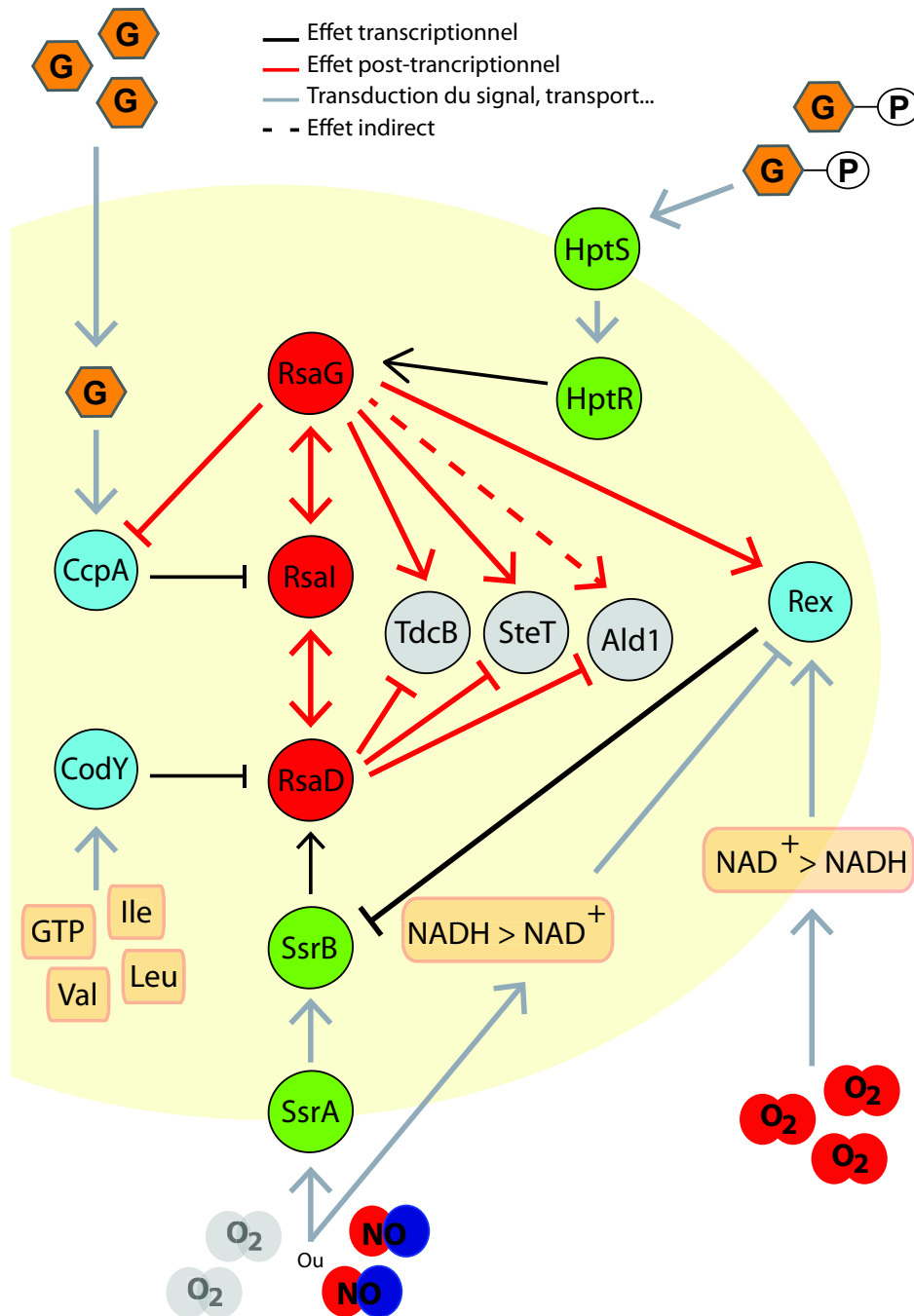


Figure 38 : Réseau de régulation du catabolisme des acides aminés par RsaG, RsaD et RsaI en réponse à divers stimuli (NO, anaérobiose, glucose (G) et glucose-6-phosphate (G6P)) : Les facteurs de transcription Rex et CcpA sont représentés en bleu, les systèmes à deux composantes HptRS et SsrAB en vert, les ARN régulateurs RsaG, RsaD, et RsaI en rouge, enfin les cibles *tdcB*, *steT* et *ald1* en gris. Dans un souci de simplification du modèle, les différents noms de gènes sont uniformément écrits, bien qu'en fonctions du type de régulation subi, ils peuvent faire référence à un ARNm, une protéine ou les deux. Les abréviations Ile, Val, et Leu correspondent respectivement aux acides aminés isoleucine, valine et leucine. Les détails sont expliqués dans le texte.

de CodY (Majerczyk et al., 2010). En inhibant RsaD, CodY devrait avoir un effet indirect sur les cibles régulées par RsaD, en empêchant leur inhibition.

J'ai montré que RsaG a également été co-purifié avec cet opéron en MAPS et des expériences de gel retard ont montré que RsaG est capable d'interagir directement avec tous les transcrits séparément, sauf *aldI*. Une comparaison du transcriptome en présence de G6P entre HG001 sauvage, délété de RsaG ou surexprimant RsaG (par un plasmide qui augmente la quantité de RsaG de presque cinquante fois) a montré que l'opéron *norb4-steT-tdcB-aldI* est significativement enrichi lorsque RsaG est surexprimé (cf II.2, Figure 38). De manière intéressante, l'interaction de RsaG est prédite dans la séquence codante de ses cibles, il sera important de valider ces sites d'interaction et de comprendre le mécanisme permettant la stabilisation de ces ARNm.

Ces résultats permettent d'intégrer RsaG et RsaD dans le réseau de régulation avec RsaI présenté dans la figure 38. En effet, les deux sARN activés par des stimuli différents ont un effet opposé sur la régulation de l'opéron *aldI-tdcB-steT-norb4*. En effet, en présence d'oxygène (et de G6P), Rex activé par RsaG inhibe le TCS SsrAB empêchant l'expression de RsaD. Ainsi RsaG assurerait la synthèse de l'opéron *aldI-tdcB-steT-norb4*. À l'inverse, en absence de G6P ou en présence de NO qui entrave la respiration, RsaD réprime ce même opéron quand RsaG n'est pas ou est faiblement exprimé. Enfin, en absence de G6P et en présence d'oxygène, RsaG et RsaD faiblement exprimés auront un effet minime sur l'expression de cet opéron. Cela révèle que le catabolisme des acides aminés pourrait avoir lieu en présence de G6P. Ces conditions ressemblent à celles retrouvées dans le contexte d'internalisation de *S. aureus*, où il a effectivement été remarqué que le métabolisme des acides aminés est activé (Palma Medina et al., 2019; cf I.5.2). Par exemple, l'alanine déshydrogénase Ald1 est surexprimée en internalisation, montrant que l'alanine peut être utilisée comme source de carbone alternative dans cette niche (Palma Medina et al., 2019).

Des expériences de gel retard par compétition ont montré que l'interaction de RsaI avec RsaG affecte l'interaction *steT*-RsaG, alors qu'elle ne déstabilise pas l'interaction *tdcB*-RsaG (Figure 37CD), bien que l'interaction se produise au niveau du premier motif riche en C de RsaG (cf III.1.1.1). Pour définir le rôle de RsaI dans ce réseau de régulation complexe, il conviendra de répéter le même type d'expérience, afin d'établir si l'interaction RsaI/RsaD empêche l'interaction de RsaD avec *tdcB* et *aldI*. Dans la mesure où RsaG et RsaD ciblent ces gènes au niveau de leurs séquences codantes (prédiction), nous pouvons présumer un impact sur la stabilité de cet opéron. La stabilité de l'ARNm *aldI-tdcB-steT-norb4* sera testée sur

diverses souches WT, Δ *rsaI*, Δ *rsaG*, Δ *rsaD*, Δ *rsaIrsaG* et Δ *rsaIrsaD*, en présence de rifampicine, suivie d'une visualisation par northern blot (ou qRT-PCR en raison de la taille de l'opéron) pour définir le temps de demi-vie des ARNm dans les conditions de stress où les deux ARN sont exprimés.

Si RsaI a une activité d'éponge sur RsaG et RsaD, empêchant la régulation de certaines de leurs cibles lorsque le glucose (et le G6P) est absent (*cf* III.1.2), il favorisera alors un métabolisme fermentaire à la place du catabolisme des acides aminés privilégié par RsaG. Il sera intéressant d'évaluer le niveau d'expression de RsaI et RsaG dans des conditions de stress NO ou d'anaérobiose, sachant que RsaD est induit dans ces conditions par SrrAB. Lorsque *S. aureus* est internalisé, le G6P est une source de carbone largement présente dans les cellules de l'hôte (Chico-Calero et al., 2002), ainsi RsaI ne devrait pas être exprimé au contraire de RsaG (Desgranges et al, en préparation). Dans ce contexte intracellulaire, il a été montré que les régulateurs de CodY et de CcpA sont dérégulés (Michalik et al., 2017). Cela suggère que RsaD devrait être présent en conditions d'internalisation, puisqu'il n'est plus inhibé par CodY. La répression de CcpA pourrait s'expliquer par l'action de RsaG qui bloquerait la traduction de *ccpA* (Desgranges et al, en préparation). En conséquence, RsaI devrait être dérégulé, ce qui est paradoxal si la concentration en G6P est élevée intracellulairement. RsaG étant nettement surexprimé dans ces conditions, nous émettons l'hypothèse que RsaI ne serait pas en quantité suffisante pour séquestrer en partie RsaG. Il est aussi fort possible que d'autres régulateurs de RsaI prennent le relais de CcpA dans ces conditions particulières. La concentration relative de ces divers ARN est certainement un des critères déterminants pour adapter le métabolisme de la bactérie.

L'étude de Augagneur et al. (2020) a révélé que RsaD régule l'ARNm *alsD* en interagissant au niveau de sa région 5'UTR (incluant le RBS), ainsi qu'au niveau d'un deuxième site localisé dans sa région codante, ayant ainsi pour effet d'inhiber la traduction et de déstabiliser cet ARNm. Le gène *alsD* code pour une enzyme impliquée dans le métabolisme carboné transformant le pyruvate en acétoïne. L'opéron *alsSD* est exprimé en phase stationnaire de croissance et est positivement régulé par CodY. Son expression contribue à la réponse au stress acide (*cf* I.4.3.1.2 ; Augagneur et al., 2020). Le MAPS réalisé avec RsaD n'a pas révélé *alsD* comme cible directe de RsaD, cependant *alsS* est co-purifié avec RsaD. De manière intéressante, l'opéron *alsSD*, est co-purifié avec RsaG en MAPS. Cela pourrait représenter un autre exemple d'opéron régulé de manière redondante entre RsaG et RsaD, si l'interaction directe de *alsS* et *alsD* est validée avec RsaG.

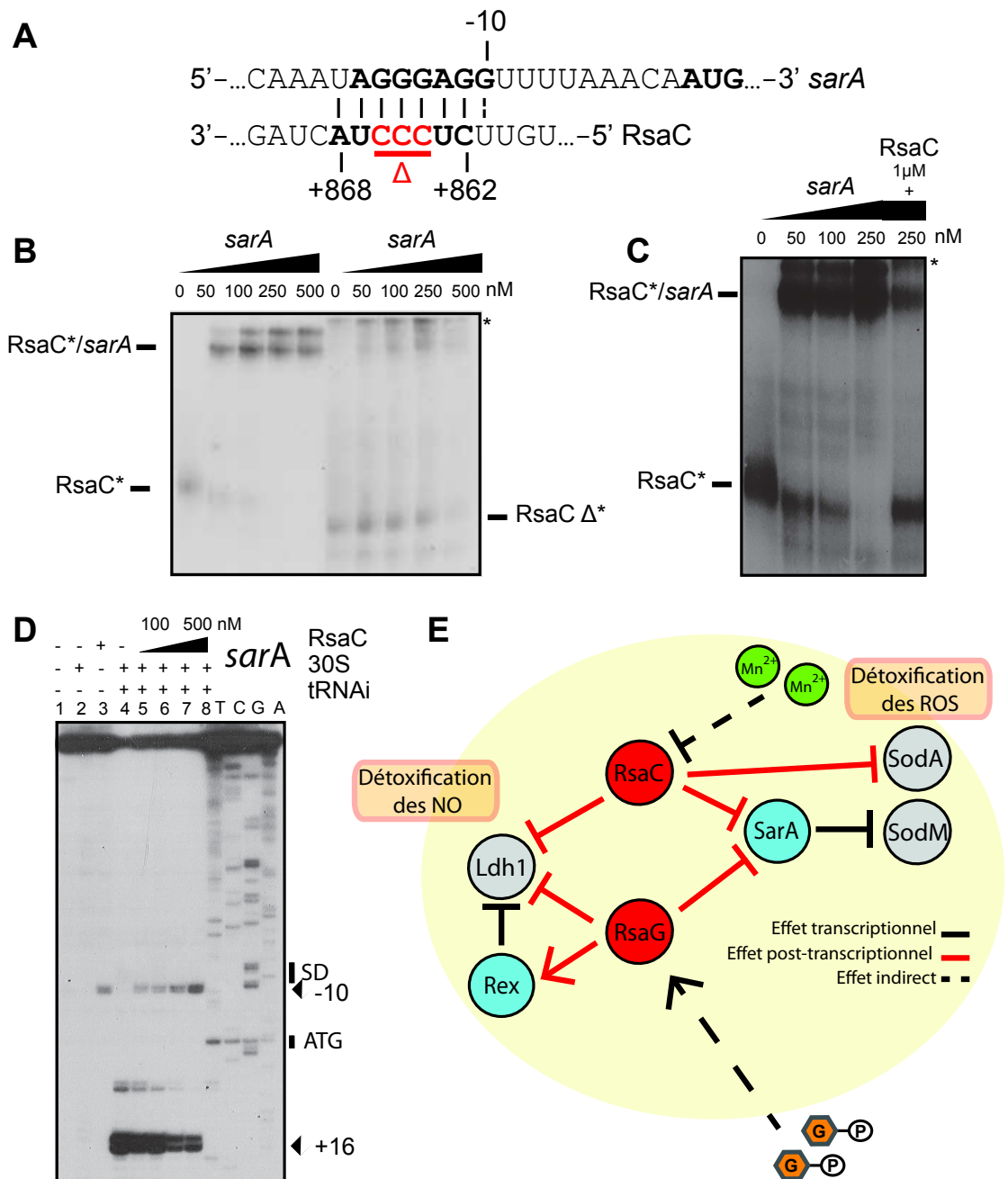


Figure 39 : Régulation de SarA par RsaC, une cible commune avec RsaG : A) Site d'appariement entre l'ARNm *sarA* et RsaC. En rouge sont présentés les nucléotides délétés dans l'ARN RsaC Δ . B) Gel retard à gauche entre l'ARN RsaC radiomarqué (RsaC*) et *sarA*, à droite entre l'ARN RsaC Δ radiomarqué (RsaC Δ *) et *sarA*. C) Gel retard entre l'ARN RsaC radiomarqué et *sarA*. Dans le dernier puits, 1 μ M de RsaC froid est ajouté pour rentrer en compétition avec RsaC radiomarqué pour l'interaction avec *sarA*. D) Expérience de Toe-print avec différentes concentrations de RsaC (piste 3, 5, 6, 7 et 8, respectivement 300 nM, 100 nM, 200 nM, 400 nM, et 500 nM) mise en présence de l'ARNm *sarA* (50 nM) et du complexe d'initiation de la traduction (sous unité 30S du ribosome de *S. aureus* (300 nM, piste 2, 4-8), ARNt initiateur (0,1 μ g/ μ L piste 4-8). Les quatre dernières pistes annotées T, C, G, et A, correspondent à un séquençage de type Sanger et permettent de définir le +16 correspondant au signal de toe-print et le -10, site d'interaction de RsaC avec *sarA*. SD indique la séquence Shine et Dalgarno et ATG le codon d'initiation de la traduction. E) Réseau de régulation impliquant les ARN régulateurs (en rouge) RsaC et RsaG, les facteurs de transcription en bleu et les cibles en gris, en fonction de la présence de manganèse (Mn²⁺) et de glucose-6-phosphate (hexagones orange). Dans un souci de simplification du modèle, les différents noms de gènes sont uniformément écrits, bien qu'en fonctions du type de régulation subi, ils peuvent faire référence à un ARNm, une protéine ou les deux.

Enfin, les résultats non publiés de MAPS réalisés avec RsaD ont révélé d'autres cibles directes de RsaD qui requièrent une validation. Parmi ces dernières, *sglT* et *PtsG_2* sont, respectivement, un co-transporteur sodium/glucose et une sous unité du système PTS spécifique du glucose, ainsi RsaD pourrait avoir un impact sur l'entrée du glucose. Sont retrouvés également *nanA* et *nanE*. Les gènes susmentionnés font partie du locus *nanE nanR nanK nanA nanT* qui code les protéines requises pour le transport et le catabolisme de l'acide sialique en pyruvate (l'acide N-acétylneuraminique). L'acide sialique est une source de carbone et d'azote très abondante au niveau des surfaces muqueuses et dans leurs productions (Olson et al., 2013). Il est donc intéressant de noter que RsaD pourrait réguler le catabolisme de l'acide sialique dans des conditions où RsaG est fortement exprimé (en présence de mucus).

1.1.3. Redondance des cibles entre RsaG et RsaC

RsaC est aussi un sARN dérivé d'une région 3'UTR, comme RsaG. Celui-ci provient de l'ARNm polycistronique *mntABC* qui code pour un transporteur ABC du manganèse. Les ARNm cibles régulés par RsaC sont impliqués dans la réponse au stress oxydatif, ainsi que dans l'homéostasie des métaux (cf I.4.3.1.1 Lalaoua et al., 2019).

Comme RsaD et RsaI, RsaC partage des cibles communes avec RsaG. En effet, l'approche de MAPS réalisée avec RsaC a révélé la superoxide dismutase SodA (cf. I.4.3.1.1), mais aussi le FT SarA et la lactate déshydrogénase Ldh1 (Lalaoua et al., 2019). Au cours de mon travail, j'ai validé la liaison directe entre RsaC et *sarA* par gel retard (Figure 39B). Cette interaction fait intervenir le cinquième motif riche en C de RsaC sur les huit qu'il comporte. Une mutation au niveau de ce site (RsaCΔ) perturbe fortement l'interaction avec l'ARNm *sarA* (Figures 39A et 39B). Par ailleurs, les ARN radiomarqués RsaC* et RsaCΔ* ne migrent pas de la même façon sur un gel en conditions non dénaturantes. Il est possible que la délétion du motif riche en C (RsaCΔ*) entraîne un changement de conformation qui pourrait modifier la migration. Afin de s'assurer que l'interaction entre RsaC et l'ARNm *sarA* est spécifique, nous avons réalisé une expérience de gel retard par compétition (Figure 39C). L'ajout d'une concentration élevée de RsaC sauvage non marqué à l'ARN radiomarké RsaC* sauvage (en quantité négligeable) montre que celui-ci est entré en compétition pour interagir avec *sarA*. Par *Toe-print*, j'ai montré que RsaC, en interagissant au niveau de la séquence SD de l'ARNm *sarA*, empêche la formation du complexe d'initiation de la traduction simplifiée *in vitro*, formé de la sous-unité 30S, de l'ARNt initiateur et de l'ARNm (Figure 39D). Une souche de *S. aureus* avec une version chromosomique de SarA étiquetée par l'épitope FLAG a été construite et sera

utilisée pour évaluer l'expression de la protéine SarA-FLAG, en comparant des conditions de culture avec et sans manganèse (Mn) (respectivement sans et avec RsaC). Ainsi, en carence de manganèse, RsaC réprime la traduction de *sodA*-dépendante du Mn (*cf.* I.4.3.1.1). La bactérie s'adapte en activant SodM, qui utilise le fer comme cofacteur pour se protéger du stress oxydatif. La transcription de *sodM* étant réprimée par SarA (Ballal & Manna, 2009), RsaC, en inhibant la traduction de l'ARNm *sarA*, active de manière indirecte *sodM*. RsaG agit de façon similaire en présence de G6P (*cf.* II.2). Ainsi, en absence de manganèse et/ou en présence de G6P, *S. aureus* est capable de détoxifier les ROS suite à l'activation de *sodM*, indirectement par RsaC ou RsaG. De plus, Ldh1, qui permet la résistance aux radicaux d'oxyde nitrique NO en orientant le métabolisme vers une fermentation lactique (De Groote & Fang, 1995; Richardson et al., 2006), est aussi une cible communément réprimée par RsaC et RsaG (Lalaouna et al, 2019, Desgranges et al, en préparation). RsaC interagit directement avec l'ARNm *ldh1* et une analyse comparative du protéome de *S. aureus* HG001 sauvage ou délétée de RsaC a montré que RsaC inhiberait la synthèse de Ldh1 (Lalaouna et al., 2019). À la fois, RsaG déstabilise l'ARNm *ldh1* et active Rex, le répresseur de *ldh1*. Ces résultats suggèrent que RsaG et RsaC doivent être inactifs en conditions de stress NO pour permettre à *ldh1* de réaliser la détoxification du NO (Grosser et al., 2016; Pagels et al., 2010) ; *cf* figure 7, et I.3.1.3). En revanche, en présence de G6P ou en absence de Mn, la bactérie adapterait son métabolisme en défaveur de la fermentation du pyruvate en lactate.

1.1.4. Conclusion

Les études individuelles des sARN conduisent à l'établissement de réseaux de régulation complexes, faisant intervenir un grand nombre de protéines régulatrices et ARN, afin d'intégrer divers stimuli, comme celui présenté Figure 36. Cependant, un grand nombre de questions quant à la fonction des régulations exercées reste en suspens. Premièrement, nous ne connaissons pas la dynamique de ces régulations. Deuxièmement, notre vision des réseaux n'est pas complète. Ainsi, en utilisant des approches globales pour définir dans une même expérience les régulations mises en place en réponse à une condition donnée, nous devrions mieux comprendre les réseaux de régulation et leur dynamique qui dictent la physiologie de *S. aureus*. Cependant, dans une même population, les bactéries présentent un comportement hétérogène. Il serait donc intéressant d'étudier l'expression des sARN, membres du réseau Figure 36, au sein de cellules uniques. En effet, l'hétérogénéité entre cellules entraîne une perte d'information car l'expression mesurée pour chaque ARN est moyennée au sein de la population. Par exemple,

dans le contexte de l'internalisation, où il est actuellement compliqué d'intégrer la balance qui existe entre l'expression de RsaG, RsaI et RsaD, une technique appelée scDual-seq pourrait être utilisée. Cette approche, basée sur le séquençage de l'ARN à la fois des bactéries internalisées et de leurs cellules hôte, a été adaptée récemment à l'échelle de la cellule unique (Avital et al., 2017).

2. Étude globale de l'interactome ARN/ARN chez *S. aureus*

2.1. Intérêt du développement d'une nouvelle technique basée sur le pontage des ARN

La technique DBRI est une amélioration des techniques déjà existantes, telles que le LigR (Durand et al., 2019 (bioRxiv); Sharma et al., 2016), le SPLASH (Aw et al., 2016), le PARIS (Lu et al., 2016) ou le « *modified CLASH* » (Liu et al., 2017). En effet, le psoralène et ses dérivés constituent la classe majeure des composés photochimiques réagissant avec les acides nucléiques. Ils ont très largement été utilisés depuis leur première utilisation en 1976, dans le but de déterminer la structure des acides nucléiques (Cimino et al., 1985; Lustig et al., 2010). Notre but était d'appliquer la technique LigR, dont le protocole était optimisé pour les cellules eucaryotes (Sharma et al., 2016), dans le modèle d'étude *S. aureus*. Pendant nos investigations, une équipe a appliqué cette technique dans un autre Firmicute, *B. subtilis* (Durand et al., 2019 (bioRxiv)). Bien que l'article soit sous révision, les principaux résultats qui se dégagent sont la découverte d'un ARN éponge (ARNe). L'ARNe RosA de *B. subtilis* a été montré comme interagissant avec au moins deux sARN, RoxS (RsaE chez *S. aureus*) et FsrA. RosA séquestre ces deux sARN, et par conséquent dérégule la synthèse des protéines codées par les ARNm cibles de RoxS et FsrA. De plus, RosA a un impact sur la stabilité de RoxS. De manière intéressante, à l'image de RsaI de *S. aureus*, la synthèse de RosA est réprimée par CcpA. Chez *S. aureus*, RsaI interagit également avec RsaE, ces résultats suggèrent donc que RsaI et RosA pourraient remplir des fonctions communes dans *S. aureus* et *B. subtilis*. Néanmoins, RosA ne présente pas de région riche en U/C lui permettant de réguler des ARNm cibles, ce qui restreindrait certainement sa fonction à celle d'ARNe (Durand et al., 2019 (bioRxiv)).

Nos expériences ont montré qu'un grand nombre de modifications au protocole de base du ligR étaient nécessaires, et nous l'avons renommé DBRI pour « *Direct determination of bacterial RNA-RNA intercatome* ». L'intérêt de cette nouvelle technique est qu'elle regroupe différents points forts trouvés dans les techniques LigR et SPLASH. En effet, notre version

utilise le Vari-X-link permettant des irradiations UV plus courtes, en raison de la puissance supérieure des lampes. Cet appareil permet d'augmenter la dose d'UV dans un temps d'irradiation fortement diminué, ce qui contribue également à augmenter l'efficacité de pontage, mais aussi à préserver l'intégrité du matériel biologique en minimisant la mortalité cellulaire et en atténuant la dégradation des ARN. En nous basant sur la technique SPLASH, nous avons introduit une chromatographie d'affinité biotine-streptavidine pour purifier les ARN pontés avec du psoralène biotinylé. Contrairement à la technique SPLASH qui utilise directement du psoralène biotinylé sur des cellules eucaryotes, nous avons utilisé un psoralène modifié qui est plus petit et qui traverse la paroi bactérienne de manière plus efficace. Le psoralène « cliquable » est ensuite lié de manière covalente à la biotine grâce à une réaction de chimie « clic ». La fixation des duplexes d'ARN pontés sur les billes permet d'éliminer les ARN non pontés pouvant former des interactions aspécifiques au cours de la procédure. Enfin comme dans la technique LigR, nous avons opté pour une ligation à haute température. Cette étape permet de réduire la non-spécificité car seuls les duplexes d'ARN covalamment liés par le psoralène peuvent être ligués dans des conditions dénaturantes. Cependant, contrairement au LigR, nous n'utilisons pas la circRNA ligase (Sharma et al., 2016) ou la circDNA ligase (Durand et al., 2019 (bioRxiv)), mais la ligase Thermostable 5' App DNA/RNA (NEB), qui a l'avantage de ne pas favoriser la circularisation des chimères. En cas de chimères circulaires et lorsque la construction des bibliothèques est basée sur la ligation d'adaptateurs, une étape de fragmentation est nécessaire. Lors de la fragmentation, les chimères sont alors coupées de manière aléatoire. Or les chimères étant courtes, les séquences respectives de chacun des deux ARN engagés pourraient être encore raccourcies suite à la fragmentation, rendant impossible l'assignation des ARN chimères sur le génome. En raison de tous ces paramètres, il sera capital d'introduire des échantillons contrôles sans ligation, ainsi que des échantillons avec et sans fragmentation.

2.2. Les prochaines étapes nécessaires pour améliorer DBRI

Les différents tests du protocole que j'ai effectués au cours de ma thèse étaient nécessaires car chaque optimisation nous a beaucoup appris sur les limites et prérequis pour cette technique. En effet, l'optimisation n°7 sera testée en tenant compte de tous les paramètres modifiés jusqu'à présent. Son efficacité sera évaluée en amplifiant par RT-PCR des chimères validées pour le couple d'ARN, *RsaG-tcaR*. Si ce protocole donne des résultats encourageants, il sera répété au minimum en duplicat dans le but de séquencer à haut-débit l'ensemble des chimères formées.

Comme nous l'avons vu avec l'analyse des données de séquençage de la variante 4.2, bien que ce protocole ait permis l'amplification de chimères RsaG-*tcaR* par RT-PCR (variante 4.1), cette chimère n'a pas été retrouvée après séquençage (variante 4.2). En effet, nous savons que dans les données de séquençage, il y a des séquences provenant de *rsaG*, cependant l'analyse ne permet pas d'attribuer ces séquences en chimère avec des ARNm cibles de RsaG. Des séquences ont également été trouvées pour les sARN RsaI et RsaA, mais à nouveau elles n'ont pas été assignées en tant que chimères avec leurs ARNm cibles respectifs. C'est pourquoi les paramètres d'analyse des données de séquençage devront être modifiés en collaboration avec le Prof. Bjoern Voss, expert dans l'analyse bio-informatique de RNA-seq. Par ailleurs, l'analyse effectuée par le Prof. B. Voss ne tient pas compte des régions non codantes des ARNm. Du fait qu'un grand nombre d'appariements a souvent lieu dans la région 5'UTR des ARNm, un grand nombre de ces chimères n'a tout simplement pas encore été trouvé.

Il sera également important d'essayer d'améliorer la ribodépletion de nos échantillons. En effet, bien qu'une étape de ribodépletion soit effectuée après l'extraction des ARN, nous retrouvons dans nos données de séquençage un grand nombre de séquences correspondant à de l'ARN ribosomiques (ARNr). Une hypothèse pour expliquer que la ribodépletion ne soit pas optimale est que les ARNr très structurés seront ciblés par le psoralène et donc pontés. La stratégie que nous utilisons pour éliminer les ARNr repose sur l'utilisation de sondes ARN complémentaires qui reconnaissent spécifiquement des séquences d'ARNr. Nous pouvons aisément imaginer que si une séquence d'ARNr ciblée par une sonde ARN est engagée dans une interaction covalente (à cause du pontage), celle-ci ne pourra pas s'hybrider avec la sonde, rendant impossible son élimination du pool d'ARN total. La solution serait d'effectuer la ribodépletion à la fin du protocole de DBRI lorsque le pontage est réversé. Malheureusement, le kit que nous utilisons est conçu pour éliminer les ARNr entiers. Le protocole de DBRI entraîne un raccourcissement majeur de la taille des ARN pour les restreindre à leur site d'interaction, nous pouvons donc supposer que ces séquences ne seront pas suffisantes pour une ribodépletion efficace. Cependant, même si les ARNr et les ARNt doivent être éliminés pour mieux apprécier les interactions impliquant les sARN, nous ne devons pas exclure non plus la possibilité que les ARNt et les ARNr puissent aussi être des cibles de certains sARN ou ARN antisens.

Mon travail aura permis d'optimiser les étapes de pontage, ligation et enrichissement des chimères. Même si je n'ai pas réussi à terminer ce travail, j'ai néanmoins bon espoir que la technique DBRI sera suffisamment puissante pour obtenir une vue assez complète des interactions ARN-ARN de *S. aureus*.

2.3. Intérêt et critique des études à l'échelle globale pour l'étude des cibles ARN

Dans tous les domaines du vivant, l'intérêt pour l'étude des ARN non codants depuis leur découverte n'a fait que s'intensifier. En effet, le monde scientifique a compris qu'à l'instar des protéines régulatrices, les ARN sont des acteurs clefs de la régulation de l'expression des gènes et particulièrement dans les processus adaptatifs. Cependant, jusqu'à très récemment, peu de méthodes étaient disponibles pour étudier l'ensemble des sARN d'un organisme donné. C'est pourquoi des études ciblées sur un ARN ont été réalisées afin de définir individuellement ses caractéristiques (origine, taille...) et ses cibles (*cf* I.4.5.1; Desgranges et al., 2020). La détermination des cibles d'un sARN permet d'en déduire sa fonction. Par exemple, une des stratégies a été souvent de comparer le transcriptome et le protéome d'une souche délétée de l'ARN à une souche sauvage ou surexprimant le sARN. Ces approches, bien que puissantes, ne différencient pas les cibles directes des cibles indirectes. L'utilisation de techniques de chromatographie d'affinité de l'ARN (comme le MAPS, *cf* I.4.5.1; Desgranges et al., 2020), permet de purifier les complexes régulateurs et privilégie l'identification des cibles directes. Cependant, si deux sARN interagissent ensemble, on peut enrichir des ARNm cibles du sARN qui n'est pas étiqueté (*cf* II.2 ; Desgranges et al., en préparation). Le MAPS nécessite l'ajout de l'étiquette MS2 à une des extrémités du sARN, ce qui pourrait empêcher sa fonction, des contrôles sont donc nécessaires pour évaluer le phénotype des souches exprimant l'ARN modifié et vérifier la taille et stabilité des ARN générés. Cette technique lourde d'un point de vue expérimental et financier (coût des séquençages) doit être répétée pour chaque sARN.

Les techniques de CLASH et Ril-seq, permettent d'avoir une vision globale des interactions ARN-ARN en fonction de la présence d'une protéine d'intérêt. Les approches LigR, SPLASH, PARIS ou « modified CLASH », quant à elles, ne se limitent pas à l'identification des interactions dépendantes d'une protéine donnée, mais elles ne reflètent pas non plus toutes les interactions ARN-ARN au sein de la cellule, dû à diverses limitations expérimentales ou informatiques. Cependant, elles permettent de révéler un nombre conséquent d'interactions avec de moindres biais. Ces nouvelles techniques ont l'avantage de donner accès à une grande quantité de données et, potentiellement, d'obtenir l'ensemble des interactions de tous les sARN avec leurs cibles. Cette force peut, cependant, être également un inconvénient car la quantité de données générées demande des pipelines robustes pour leur analyse, ainsi qu'une validation expérimentale rigoureuse. Lorsque nous réalisons le MAPS, la validation expérimentale des cibles est une étape indispensable, qui permet d'écarter les faux positifs et de déterminer le mode de régulation (répression, activation...). Cependant, dans une étude globale, la quantité de données est telle que nous pouvons nous demander s'il est vraiment

possible de valider toutes les cibles et de faire le choix approprié des cibles à valider. Les retards sur gel, les empreintes du ribosome, la cartographie en solution ou encore l'utilisation de fusion utilisant un gène rapporteur fusionné à la région de régulation, *in vivo*, sont des méthodes classiques de validation des cibles, mais leur utilisation dans ce contexte serait très fastidieuse et limiterait le nombre d'interactions à valider.

Pour permettre une validation rapide, la technique DBRI pourra être combinée à d'autres approches globales, mais complémentaires (*cf* Desgranges et al., 2020, partie I.4.4.1). Par exemple, lorsqu'un sARN est retrouvé dans plusieurs chimères, le MAPS pourra être réalisé pour mieux définir son targetome *in vivo* (Lalaouna et al., 2018). Un transcriptome effectué dans les mêmes conditions permettra la normalisation des niveaux des ARN engagés dans les chimères. Afin de définir si les interactions ARN-ARN ont un effet sur leur stabilité, la méthode 4sU-seq détermine la vitesse de dégradation et de synthèse de l'ensemble du transcriptome en comparant une souche sauvage avec une souche mutante pour un sARN (Wachutka & Gagneur, 2017). Cette approche permet d'analyser l'effet du sARN sur la stabilité de l'ensemble du transcriptome. Cette méthode est plus puissante que l'utilisation de la rifampicine pour bloquer la transcription des ARN et mesurer ensuite individuellement leurs temps de demi-vie par Northern blot. Afin de définir si les interactions ARN-ARN ont un effet sur la traduction, le ribosome profiling apparaît comme la méthode de choix (Ingolia et al., 2009) en comparant une souche sauvage (ou surexprimant le sARN) et une souche mutante pour un sARN. Les alternatives, plus fastidieuses, consisteraient à construire des fusions traductionnelles pour chaque ARNm ou d'étiqueter individuellement chaque protéine pour regarder leur expression par western blot.

Outre le fait que les méthodes globales génèrent une quantité non négligeable de faux positifs, une interaction entre deux molécules ne conduit pas toujours à un phénotype clair et à une fonction. Dans le cas des ARN régulateurs eucaryotes, les miARN, il est communément admis qu'un miARN contrôle une large diversité de processus physiologiques, parce qu'ils interagissent avec des dizaines, voire des centaines d'ARNm cibles. Cependant, des travaux récents en particulier de l'équipe de H. Seitz (Mockly & Seitz, 2019 ; Seitz, 2017) ont montré que les méthodes expérimentales et prédictives génèrent de nombreux faux-positifs, et qu'une répression inférieure à 2 ne parvient que très rarement à déclencher un phénotype majeur. Ils considèrent que ces interactions correspondraient ainsi à un bruit de fond (Mockly & Seitz, 2019; Seitz, 2009, 2017).

La caractérisation des cibles de RsaG (*cf* II.2) par MAPS et son analyse bio-informatique ont révélé un grand nombre de cibles potentielles, plus de 180. Cependant, dans

cette analyse, nous n'avons pas tenu compte du transcriptome qui permet d'évaluer les quantités relatives des ARN produits dans les conditions utilisées. Cela apporte une information supplémentaire pour mieux évaluer le taux d'enrichissement des ARN co-purifiés avec RsaG. Vingt-sept candidats ont cependant été validés au niveau de leur interaction directe avec RsaG, parmi trente-deux testés. Les régions 5'UTR de certains des ARNm contenant le site d'interaction ont ensuite été fusionnées à un gène rapporteur pour analyser l'effet de l'expression de RsaG. Malheureusement, nous n'avons pas réussi à observer des effets significatifs de RsaG sur cette fusion pour les gènes testés (*tcaR*, et *treB*). Nous n'excluons pas que le système utilisé n'est peut-être pas approprié car le promoteur utilisé pour exprimer la fusion est constitutif et fort (*cf* III.1.1). Nous envisageons ainsi d'intégrer la fusion dans le chromosome et d'utiliser systématiquement le promoteur du gène étudié pour se rapprocher des conditions de la bactérie (et de tester également les gènes *ccpA* et *sarA*). Par ailleurs, des résultats récents indiquent que RsaG pourrait être un sARN bifonctionnel en codant un peptide, l'ORF de ce potentiel peptide comprend les régions régulatrices de RsaG (Sorensen et al., 2020). Ainsi, nous pouvons imaginer que les ribosomes en cours de traduction peuvent interférer avec les fonctions régulatrices de RsaG. Bien que certaines des cibles identifiées par le MAPS ne soient peut-être pas fonctionnelles, d'autres explications peuvent aussi être prises en considération, comme par exemple la hiérarchie des cibles (dépendant de la stabilité des interactions et du taux respectif de RsaG). Chez les bactéries, les sARN ont une expression forte pour que ceux-ci ne soient pas limitant lorsque le stress est activé. C'est pourquoi nous pensons que RsaG exerce son action principale en présence de G6P et lorsque *S. aureus* est internalisé. Par exemple chez *E. coli*, SgrS est également stimulé par le G6P, il permet une réponse au stress métabolique associé à l'accumulation intracellulaire de sucres phosphorylés. SgrS présente un régulon restreint et les mécanismes de régulation de ses cibles sont caractérisés ; ce qui a motivé Bobrovskyy et al. (2019) à utiliser SgrS comme modèle pour l'étude de la hiérarchisation de ses cibles. Il en ressort que certaines cibles de SgrS voient leur régulation priorisée, même lorsque SgrS est exprimé à un faible niveau, alors que d'autres cibles sont régulées seulement lorsque SgrS est fortement exprimé. Il en résulte une régulation temporelle des cibles de SgrS, qui suit l'augmentation de sa concentration. De plus, d'autres facteurs, tels que la RNase E et Hfq, jouent un rôle dans la hiérarchisation des cibles de SgrS en augmentant ou en diminuant l'efficacité de régulation de certaines d'entre elles (Bobrovskyy et al., 2019). Il serait maintenant intéressant, mais plus fastidieux, d'examiner la hiérarchisation des cibles au sein du régulon très large de RsaG.

Pour conclure, les nouvelles technologies pour l'étude des interactions ARN-ARN deviennent de plus en plus sensibles, rendant la validation indispensable pour distinguer les faux positifs des cibles réelles responsables de l'adaptation de la bactérie à son nouvel environnement.

2.4. Perspectives d'utilisation

La technique DBRI requiert encore des améliorations, cependant, les perspectives d'utilisations sont nombreuses pour comprendre la régulation de la physiologie et de la virulence de *S. aureus*. Grâce à DBRI, nous pourrions varier les conditions de culture (milieux carencés en sources de carbone ou métaux), appliquer des stress pour mimer l'environnement de *S. aureus* au contact du système immunitaire de l'hôte, comme un stress oxydatif (Beavers & Skaar, 2016) ou NO, ou encore moduler la concentration en oxygène équivalente à celle retrouvée dans les différents tissus infectés. La technique DBRI pourrait également être réalisée sur des cultures liquides de *S. aureus* en présence d'antibiotiques. Ce type d'expérience pourrait déceler de nouveaux mécanismes d'adaptation ou de résistance aux antibiotiques, par exemple en comparant des souches sensibles et résistantes. DBRI pourrait être employé avec des co-cultures de *S. aureus* incubés en présence de divers types cellulaires cibles (neutrophiles, macrophages, ostéoblastes, etc.). Si la technique est suffisamment sensible, il serait envisageable que DBRI soit utilisée sur des bactéries récupérées d'organes ou de tissus suite à des infections en modèle animal (souris ou cobayes). L'ensemble des différentes conditions que nous pourrions tester nous donnera une meilleure vision de la place des sARN dans les réseaux complexes de la régulation chez *S. aureus*.

3. Conclusion générale

L'étude individuelle des ARN régulateurs reste essentielle pour comprendre l'ensemble des acteurs impliqués dans les phénomènes adaptatifs des microorganismes. Par exemple, l'introduction du MAPS a été déterminante pour aborder la fonction de plusieurs sARN découverts il y a maintenant plus de 10 ans. J'ai ainsi contribué à la caractérisation des réseaux impliquant RsaI et RsaG en réponse aux sources carbonées et mis en lumière de nouveaux mécanismes de régulation et de la genèse de ces sARN (depuis un promoteur *cf.* II.1/depuis une maturation *cf.* II.2). Ces travaux ont permis de comprendre au niveau moléculaire les mécanismes mis en jeu leur permettant d'avoir une activité régulatrice

(stabilisation/déstabilisation des ARNm, inhibition de l'initiation de la traduction... cf. II.1 et II.2). Cependant, de nombreuses questions subsistent. À part les RNases, y auraient-ils d'autres protéines associées aux fonctions des sARN chez *S. aureus* ? Les riboswitch, largement impliqués dans la régulation de voies métaboliques, contribuent-ils à l'établissement de la virulence ? Quelle est l'étendue des fonctions régulatrices des régions 3'UTR ? Existe-t-il une hiérarchisation des cibles au sein de tous les régulons ?...

Une nouvelle ère est certainement arrivée avec le développement des méthodes à l'échelle des génomes. Nous avons maintenant les moyens d'obtenir une vue d'ensemble intégrant un grand nombre de sARN et leurs cibles. Ces nouvelles technologies seront très certainement à l'origine d'une augmentation rapide de notre connaissance de la biologie des ARN. Ces données pourront bien évidemment être utilisées pour alimenter la biologie des systèmes. La biologie des systèmes s'oppose à la biologie moléculaire considérée comme réductionniste, elle permet de réaliser une description globale et dynamique des phénomènes biologiques en intégrant de multiples données, qui seules seraient insuffisantes pour modéliser et prédire un comportement cellulaire. La notion de réseau est au cœur de la biologie des systèmes. Par exemple chez *S. aureus* le système de *quorum sensing* (QS, cf I.4.2.2.1 et I.5.6), qui est très important pour l'établissement de la virulence, permet de coordonner l'expression d'un grand nombre de gènes au sein de la population bactérienne. Une modélisation mathématique du QS a été réalisée pour intégrer les réseaux denses essentiellement restreints aux protéines régulatrices décrits dans la littérature, afin de comprendre la capacité de *S. aureus* à former un biofilm. Cette approche peut maintenant être utilisée pour formuler plus facilement des hypothèses à tester *in vivo*, en manipulant les nœuds des réseaux simulés *in silico* (Audretsch et al., 2013). Cette étude s'est focalisée sur le QS, mais nous pouvons imaginer que l'incorporation des données issues de la cartographie des interactions ARN-ARN dans la biologie des systèmes devrait permettre de mieux établir les liens existants entre les réponses à divers stress, les changements métaboliques, ainsi que les conséquences sur le QS et la synthèse des facteurs de virulence. Autrement dit, de mieux modéliser le comportement et la physiologie de *S. aureus* en fonction de ses divers environnements.

IV. Matériel & méthodes

La majeure partie des expériences réalisées sont décrites dans les matériel et méthodes des deux articles présentés en résultat ainsi que dans la partie II.3.1 Ce chapitre présente les techniques correspondant aux expériences montrées dans la partie III.1.3.

1. Liste des oligonucléotides

Nom	Séquence	Expérience
RsaC-T7 for	TAATACGACTCACTATAGGGGAAAAGCTTTATGTGGTTTTG	Amplification PCR
RsaC rev	AAAATAGCCACACTCATATG	Amplification PCR
SarA-T7 for	TAATACGACTCACTATAGGGTAAATAATAGAGCATTAAAT	Amplification PCR
SarA rev1	TCTCTTTGTTTTTCGCTGATG	Amplification PCR
SarA rev2	GTCAGCATAAGTGAC	Toe-print
RsaG-T7 rev	TAATACGACTCACTATAGGGGACACCGCTATATGTC	Sonde DIG
RsaG for	TGCCCCGACGGCATGTGCGA	Sonde DIG
TcaR for	TAATACGACTCACTATAGGGTATGGGTCAATATTTTTGTG	Transcription <i>in vitro</i>
TcaR rev	TTAACTTTTTTAATTCGTCGG	Transcription <i>in vitro</i>
RsaG rev	TAAGGATACTTCATAACAGGGAATTATG	RT-PCR (DBRI)
TcaR for2	TGAGGATACTTTATAAAAATAGTAATTAGGGAG	RT-PCR (DBRI)

2. Liste des plasmides

Nom	Description	Référence
pCN51-P3 RsaC ₁₁₁₆	pCN51 exprimant RsaC (1116 nucléotides) sous promoteur P3	Lalaouna et al., 2019
pCN51-P3 RsaCΔ	pCN51 exprimant RsaC délété des nucléotides 864 à 866, sous promoteur P3	Résultat non publié (Lalaouna D)
pUC18-T7 RsaG	RsaG sous promoteur T7 utilisé pour la transcription <i>in vitro</i> de RsaG	Geissmann et al., 2009

3. Amplification des plasmides Pjet::RsaG-tcaR (DBRI)

Dans le but de pouvoir amplifier et séquencer les vecteurs (Pjet) comportant les fragments de RT-PCR des chimères RsaG-tcaR obtenues au cours des différentes variantes du protocole DBRI (*cf* II.3.1.4.1.2), des cellules compétentes ont été préparées dans le but d'être transformées par ces plasmides.

3.1. Préparation des cellules compétentes

Un volume de 200 mL de LB (Roth) a été inoculé au 1/100 par une pré-culture d'*E. coli* DH5 α . Les cellules ont été cultivées à 37°C sous agitation jusqu'à une DO_{600nm} de 0,3-0,4. Après centrifugation (780 g, 15 minutes, 4°C), les culots ont été lavés une fois dans 20 mL de CaCl₂ 100 mM stérile. Les cellules ont été incubées 1 heure sur glace, puis centrifugées (780 g, 15 minutes, 4°C). Les culots ont ensuite été repris dans 1 mL de CaCl₂ 100 mM et 1 mL de glycérol 50%. Enfin des aliquotes de 100 μ L ont été congelées à -80°C.

3.2. Transformation des bactéries chimio-compétentes

La réaction de ligation (*cf* II.3.1.4.1.2) (10 μ L) a été ajoutée à 100 μ L de cellules compétentes DH5 α (*cf* II.3.1.4.1.3). Les cellules ont ensuite été incubées 20 minutes sur la glace avant de recevoir un choc thermique d'une minute à 42°C. Puis 900 μ L de LB liquide stérile ont été ajoutés et les cellules ont été incubées 30 minutes à 37°C sous agitation. Les bactéries ont ensuite été centrifugées, le surnageant éliminé et les culots repris pour être étalées sur un milieu LB, 1,5% agar-agar (Roth) contenant de l'ampicilline 100 μ g/mL. Les bactéries transformées ont été incubées à 37°C sur la nuit.

3.3. Séquençage

Les plasmides des colonies résistantes à l'ampicilline sont amplifiés dans une culture de 2 mL de LB, ampicilline (100 μ g/mL), sur la nuit à 37°C sous agitation. Les plasmides sont ensuite purifiés avec le kit nucleospin plasmide (Macherey-Nagel) et séquencés (GATC Biotech).

4. Northern blot (DBRI)

4.1. Préparation des ARN totaux

Les conditions de cultures sont présentées dans la partie II.3.1.1, ainsi que les conditions de traitement au psoralène (*cf* II.3.1.2.1), et le protocole d'extraction des ARN totaux (*cf* II.3.1.2.2).

4.2. Séparation des ARN totaux et transfert

Les ARN totaux sont séparés en fonction de leur taille par électrophorèse sur gel d'agarose 1% (PanReac AppliChem) préparés dans du TBE 1X en présence de 20 mM de guanidium thiocyanate. Pour chaque condition, 10 μ g d'ARN sont déposés dans un volume de 10 μ L auquel on ajoute 10 μ L de bleu urée 2X. Les échantillons sont chauffés pendant 3 minutes à 90°C afin de dénaturer les ARN puis placés immédiatement dans la glace avant d'être déposés sur gel. La migration se déroule à 4°C pendant 3 heures avec un voltage de 130 V. Le gel est

immergé dans une cuve à électrophorèse remplie de TBE 1X. Après la migration, le gel est incubé sous agitation 10 minutes dans un tampon SSC 2X (0,3 mM NaCl, 30 mM citrate-NaOH, pH 7). Les ARN sont ensuite transférés sur une membrane de nitrocellulose Hybond N+ (GE healthcare) pré-humidifiée dans du tampon SSC 2X. Le transfert est réalisé par aspiration sous vide pendant 1 heure avec l'appareil Vacuum-Blot system (Whatman Biometra). Après le transfert, la membrane subit un traitement aux UV ($\lambda = 254$ nm) avec une énergie de 120 mJ (UV Stratalinker 1800, Stratagene) dans le but de lier les ARN de manière covalente à la membrane.

4.3. Détection des ARN

La membrane est ensuite placée dans un tube d'hybridation en verre, en présence de tampon d'hybridation pendant 30 minutes à 68°C dans un four rotatif (Hybrigen). Une sonde ARN marquée à la digoxygénine (DIG-11-UTP) (2 μ L, 200 ng) spécifique de l'ARN recherché est dénaturée à 68°C et ajoutée à la solution d'hybridation, qui se déroule sur la nuit à 68°C. La membrane est ensuite lavée deux fois pendant 5 minutes à 20°C avec un tampon SSC 2X, SDS 0,1%, puis deux fois pendant 15 minutes à 68°C avec du tampon SSC 0,2X, SDS 0,1%. Les étapes à suivre se déroulent à 20°C, la membrane est saturée avec du tampon de blocage (blocking reagent, Roche) pendant 30 minutes afin de masquer les sites de fixation aspécifiques. La solution est renouvelée avec une concentration au 1/1000 d'anticorps spécifiques des groupements DIG couplés à la phosphatase alcaline (Roche). Le marquage s'effectue pendant 30 minutes, puis la membrane est lavée deux fois 15 minutes avec le tampon de lavage (100 mM acide maléique pH 7,5, 150 mM NaCl, 0,3% Tween 20 (sigma)), puis incubée 5 minutes avec le tampon de détection (100 mM Tris, 100 mM NaCl, pH 9,5). Après ajout du substrat chimioluminescent de la phosphatase alcaline (CDP star, Roche), la membrane est scellée dans un film plastique, puis placée dans une cassette. Enfin, la révélation des résultats se fait par autoradiographie.

4.4. Construction des sondes ARN modifiées par la digoxygénine

La sonde ARN RsaG DIG a été obtenue par transcription *in vitro*, à partir d'un produit PCR. Dans un premier temps, environ 300 nucléotides de la région reverse et complémentaire à l'ORF de RsaG ont été amplifiés par PCR avec les amorces décrites paragraphe IV.1. L'oligonucléotide sens porte la séquence du promoteur de la T7 ARN polymérase à son extrémité 5' pour permettre la transcription *in vitro*. Les PCR ont été réalisées avec l'ADN polymérase à haute-fidélité Phusion polymérase (Fermentas) et de l'ADN chromosomique

comme matrice selon les recommandations du fournisseur. Les produits de PCR ont été purifiés avec le kit nucleospin gel & PCR clean-up (Macherey-Nagel). 500 ng de produits PCR a été transcrit *in vitro* dans un volume final de 20 µL contenant 2 µL de DIG-RNA *labelling mix* 10x (Roche), 0,1 mg/mL BSA, 5 mM DTT, 1 mM spermidine, 40 mM Tris-HCl pH 8, 15 mM MgCl₂, 50 mM NaCl, et 5 µL de T7 ARN polymérase (purifié au laboratoire). La réaction de polymérisation a été réalisée à 37°C pendant 3 heures. Trois unités de DNase I, et 3 µL de tampon DNase I 10X, sont ajoutés dans un volume final de 20 µL, et le mélange est incubé 15 minutes à 37°C. Les ARN sont extraits, précipités et leur intégrité est vérifiée sur mini-gel (*cf* II.3.1.2.5.2.1).

5. PCR

RsaC et RsaCΔ ont été amplifiés à partir de 10 ng des plasmides (respectivement pCN51-P3 RsaC₁₁₁₆ et pCN51-P3 RsaCΔ, *cf* IV.1), *sarA* a été amplifié directement sur le génome de *S. aureus* HG001 à partir de 50 ng d'ADN génomique. Les différents ADN ont été mis en présence de 0,1 µM de chaque amorce spécifique (RsaC-T7 for et RsaC rev pour RsaC et RsaCΔ, SarA-T7 for et SarA rev1 pour SarA, IDT, *cf* IV.1), 1 U de *Phusion High-Fidelity DNA polymerase* (ThermoFisher), 10 µL de tampon HF 5X, et 0,25 mM de chaque désoxyribonucléotide dans un volume final de 50 µL. La première étape de la PCR consiste en une étape de dénaturation (3 minutes à 98°C), puis une étape d'amplification a été réalisée en 34 cycles (30 secondes à 98°C, 30 secondes 51°C et 45 secondes à 72°C), enfin une étape de terminaison a été effectuée 5 minutes à 72°C. Les produits de PCR ont été purifiés avec le kit nucleospin gel & PCR clean-up (Macherey-Nagel) resuspendus dans un volume final de 30 µL d'eau milliQ.

6. Transcription *in vitro*

Les transcriptions *in vitro* de RsaC, RsaCΔ et *sarA* sont réalisées sur 5 à 10 µg de produits PCR de la même manière que dans la partie II.3.1.1.1 avec les amorces décrites partie IV.1.

7. Déphosphorylation de l'extrémité 5' des ARN

Préalablement aux marquages des ARN en 5', 10 µg d'ARN sont ajoutés au tampon FastAP (Fermentas 1X final) dans un volume final de 200 µL, contenant 50 U d'alcaline phosphatase FastAP (ThermoFischer). La déphosphorylation est réalisée 10 minutes à 37°C. Après extraction au phénol/chloroforme, les ARN sont précipités à l'éthanol, lavés puis séchés par évaporation sous vide. Le culot est repris dans 10 µL d'eau milliQ.

8. Marquage d'ARN ou d'oligonucléotides en 5'

Les ARN déphosphorylés RsaC/ RsaCΔ (5 µg) ou l'oligonucléotide SarA rev2 (25µM) (*cf* IV.1) est incubé avec 50 µCi d'ATPγP32 (Hartmann Analytic, 50 mCi/mL), 2 µL de tampon A T4 PNK 10X (Fermentas) et 20 U de T4 PNK (Fermentas) dans un volume final de 20 µL pendant 1 heure à 37°C. Après ajout d'un volume de bleu urée, les ARN après incubation 5 minutes à 70°C, sont purifiés par électrophorèse sur gel de polyacrylamide 6%, urée 8 M, TBE 1X, polymérisé en présence de persulfate d'ammonium (APS) 0,1% et de TEMED 1/1000. Le gel est préalablement soumis à une pré-migration de 20 minutes à 300 V. La migration dure 4 heures à 300 V. Après migration, le gel est transféré sur un film autoradiographique usagé, entouré de film « saran wrap », et placé dans une cassette avec un écran amplificateur de signal au contact d'un film autoradiographique. La bande de gel contenant l'ARN radiomarké est découpée puis éluée pendant une nuit à 4°C sous agitation dans un tampon d'élution (500 mM ammonium acétate, 0,5 M EDTA, 10% SDS) supplémenté de phénol. Le lendemain, les ARN sont extraits par un traitement PCI (phénol, chloroforme, alcool isoamylique 25 :24 :1 pH 4,5-5, ROTH) puis précipités à l'éthanol, lavés et séchés. Le culot est repris dans 50 µL d'eau milliQ. La radioactivité est mesurée en Cerenkov avec un compteur de radioactivité (Multi-purpose scintillation counter, Beckman).

9. Retard sur gel

Le suivi de la formation des complexes ARN-ARN demande à ce que l'un des deux ARN soit marqué. L'ARN radioactif (RsaC* ou RsaCΔ* 10000 cpm/µL final) et l'autre ARN froid (SarA 1 µM final) sont dilués dans un tampon (100 mM Tris-HCl pH 7,5, 300 mM KCl, 200 mM NH₄Cl) séparément pour être tout d'abord dénaturés 1 minute à 90°C, suivi par 1 minute sur la glace. Un volume de MgCl₂ est ensuite ajouté à une concentration de 10 mM final pour renaturer les ARN 10 minutes à 20°C. L'ARN marqué est ensuite incubé seul ou en présence de concentrations croissantes de l'ARN partenaire froid dans du tampon (20 mM Tris-HCl pH 7,5, 60 mM KCl, 40 mM NH₄Cl, 10 mM MgCl₂) dans un volume final de 10 µL. Les différentes réactions sont incubées 15 minutes à 37°C, puis 10 µL de bleu glycérol sont ajoutés aux échantillons avant d'être déposés sur un gel natif de polyacrylamide 10% (rapport 1 :30 acrylamide :bis-acrylamide) dans le tampon Tris-Borate 1X contenant 1 mM de MgCl₂. Le gel est polymérisé en présence de persulfate d'ammonium (APS) 0,1% et de TEMED 1/1000, et est soumis à une pré-migration de 20 minutes à 300 V dans le tampon Tris-Borate 1X. Après dépôt des échantillons, la migration se déroule pendant 7 heures à 4°C. Après migration, le gel

est démoulé comme précédemment puis l'autoradiographie est laissée à -80°C pendant une nuit, et est révélée le lendemain matin.

10. Toe-print

Les sous unités ribosomiques 30S de *S. aureus* sont activées 5 minutes dans la glace puis 10 minutes à 37°C, puis 1,5 µM de cette solution stock est incubé 15 minutes à 37°C dans un tampon Toe+ 1X (20 mM Tris-HCl pH 7,5, 60 mM KCl, 1 mM DTT, 10mM MgCl₂). 250 nM de l'ARNm *sarA*, et l'équivalent de 1 million de coups par minute (cpm) d'oligonucléotides reverse complément de *sarA* sont dénaturés 1 minute à 90°C, suivi par 1 minute sur la glace. Le tampon toe+, contenant du magnésium, est ensuite ajouté pour renaturer les ARN 10 minutes à 20°C. La même procédure est réalisée avec 1µM du sARN RsaC seul. Ensuite deux 20 nM du mélange d'ARNm *sarA*/oligonucléotide est incubé 10 minutes à 37°C avec une concentration variable de RsaC (0 à 500 nM en fonction des tubes) dans un tampon Toe- 1X (20 mM Tris-HCl pH 7,5, 60 mM KCl, 1 mM DTT) ainsi que le volume nécessaire de MgCl₂ 100 mM pour atteindre une concentration finale de 1 mM dans l'ensemble des tubes (calculé en prenant compte de l'ajout de 30S activé à l'étape suivante) La sous unité 30S activée (0,25 µM) est ajouté dans les tubes (sauf les tubes contrôle sans ribosome) et le milieu est incubé 10 minutes à 37°C. Enfin 0,1 µg/µL d'ARNt initiateur est ajouté dans chaque tubes (sauf les tubes contrôle sans tRNAi) puis incubé 5 minutes à 37°C. Une étape de transcription inverse est réalisée pour révéler la formation d'un complexe d'initiation de la traduction sur l'ARNm *sarA*. 4 U d'AMV-RT, 4 mM d'un mélange équimolaire de chaque dNTP, et le tampon AMV commercial à une concentration finale de 1X sont ajoutés aux différents tubes, et incubés 30 minutes à 37°C pour permettre la rétrotranscription. Les ADN complémentaires (ADNc) néosynthétisés sont extraits par un traitement PCI (phénol, chloroforme, alcool isoamylique 25 :24 :1 pH 7,5-8, ROTH). Après précipitation des ADNc à l'éthanol, les culots sont directement repris dans du bleu urée 2X à raison de 5 µL pour 10000 cpm. En parallèle des séquences sont préparées afin de pouvoir lire la séquence de l'ARNm *sarA* et définir la position du signal de toe-print, du codon d'initiation de la traduction et de la séquence Shine et Dalgarno. L'ARNm *sarA* (2,5 µM), et l'équivalent de 100000 cpm d'oligonucléotide reverse complément de *sarA* sont dénaturés 1 minute à 90°C, suivi par 1 minute sur la glace. Le tampon commercial de l'AMV-RT, à une concentration finale 1X, est ensuite ajouté pour renaturer les ARN 10 minutes à 20°C. Un mélange différent pour chaque nucléotide, constitué de dNTP et de ddNTP, est ajouté dans des tubes différents (par exemple le mélange ajouté au tube pour réaliser la

séquence de l'adénine se compose de 100 μM de dCTP, dTTP, et dGTP, 25 μM de dATP et 1,67 μM de ddATP). Quatre unités d'AMV-RT et le tampon commercial de l'AMV-RT, à une concentration finale 1X, sont également ajoutés. Le mélange réactionnel est ensuite incubé 30 minutes à 37°C. Enfin la matrice ARN est détruite par l'ajout de 0,6 M de KOH et 20 μL de tampon destroy (50 mM Tris-HCl pH 8, 0,5% SDS, 7,5 mM EDTA) incubé 3 minutes à 90°C, puis une heure à 37°C. Comme précédemment les ADNc sont extraits, précipités, et repris dans du bleu urée 2X. Enfin les ADNc de la séquence et de l'expérience de toe-print sont chauffés 3 minutes à 90°C, et l'équivalent de 10000 cpm de chaque tube est déposé sur un gel d'acrylamide 10% dans le tampon Tris-Borate-EDTA 1X, pour une migration d'environ 3 heures à 75 W. Après migration, le gel est démoulé (*cf* IV.6) puis l'autoradiographie est laissée à -80°C pendant une nuit, et est révélée le lendemain matin.

V. Références Bibliographiques

Altuvia, S., and Wagner, E.G.H. (2000). Switching on and off with RNA. *Proceedings of the National Academy of Sciences* 97, 9824–9826.

Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E.G.H., Margalit, H., and Altuvia, S. (2001). Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Current Biology* 11, 941–950.

Argudín, M.Á., Mendoza, M.C., and Rodicio, M.R. (2010). Food Poisoning and *Staphylococcus aureus* Enterotoxins. *Toxins* 2, 1751–1773.

Audretsch, C., Lopez, D., Srivastava, M., Wolz, C., and Dandekar, T. (2013). A semi-quantitative model of Quorum-Sensing in *Staphylococcus aureus*, approved by microarray meta-analyses and tested by mutation studies. *Mol. BioSyst.* 9, 2665.

Augagneur, Y., King, A.N., Germain-Amiot, N., Sassi, M., Fitzgerald, J.W., Sahukhal, G.S., Elasri, M.O., Felden, B., and Brinsmade, S.R. (2020). Analysis of the CodY RNome reveals RsaD as a stress-responsive riboregulator of overflow metabolism in *Staphylococcus aureus*. *Mol Microbiol* 113, 309–325.

Avital, G., Avraham, R., Fan, A., Hashimshony, T., Hung, D.T., and Yanai, I. (2017). scDual-Seq: mapping the gene regulatory program of *Salmonella* infection by host and pathogen single-cell RNA-sequencing. *Genome Biol* 18, 200.

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.

Ballal, A., and Manna, A.C. (2009). Regulation of Superoxide Dismutase (sod) Genes by SarA in *Staphylococcus aureus*. *JB* 191, 3301–3310.

Beavers, W.N., and Skaar, E.P. (2016). Neutrophil-generated oxidative stress and protein damage in *Staphylococcus aureus*. *Pathogens and Disease* 74, ftw060.

Beenken, K.E., Blevins, J.S., and Smeltzer, M.S. (2003). Mutation of sarA in *Staphylococcus aureus* Limits Biofilm Formation. *Infect Immun* 71, 4206–4211.

Beisel, C.L., and Storz, G. (2010). Base pairing small RNAs and their roles in global regulatory networks. *FEMS Microbiol Rev* 34, 866–882.

Bischoff, M., Wonnenberg, B., Nippe, N., Nyffenegger-Jann, N.J., Voss, M., Beisswenger, C., Sunderkötter, C., Molle, V., Dinh, Q.T., Lammert, F., et al. (2017). CcpA Affects Infectivity of *Staphylococcus aureus* in a Hyperglycemic Environment. *Front. Cell. Infect. Microbiol.* 7, 172.

Boakes, E., Marbach, H., Lynham, S., Ward, M., Edgeworth, J.D., and Otter, J.A. (2016). Comparative analysis of phenol-soluble modulin production and *Galleria mellonella* killing by community-associated and healthcare-associated methicillin-resistant *Staphylococcus aureus* strains. *Journal of Medical Microbiology* 65, 1429–1433.

Bobrovskyy, M., Azam, M.S., Frandsen, J.K., Zhang, J., Poddar, A., Ma, X., Henkin, T.M., Ha, T., and Vanderpool, C.K. (2019). Determinants of target prioritization and regulatory hierarchy for the bacterial small RNA SgrS. *Mol Microbiol* 112, 1199–1218.

Bohn, C., Rigoulay, C., and 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., et al. (2010). Experimental discovery of small RNAs in *Staphylococcus aureus* reveals a riboregulator of central metabolism. *Nucl. Acids Res.* 38, 6620–6636.

Brandenberger, M., Tschierske, M., Giachino, P., Wada, A., and Berger-Bächi, B. (2000). Inactivation of a novel three-cistronic operon *tcaR-tcaA-tcaB* increases teicoplanin resistance in *Staphylococcus aureus*. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1523, 135–139.

Braun, F., Durand, S., and Condon, C. (2017). Initiating ribosomes and a 5'/3'-UTR interaction control ribonuclease action to tightly couple *B. subtilis* hbs mRNA stability with translation. *Nucleic Acids Research* 45, 11386–11400.

Bronesky, D., Wu, Z., Marzi, S., Walter, P., Geissmann, T., Moreau, K., Vandenesch, F., Caldelari, I., and Romby, P. (2016). *Staphylococcus aureus* RNAIII and Its Regulon Link Quorum Sensing, Stress Responses, Metabolic Adaptation, and Regulation of Virulence Gene Expression. *Annual Review of Microbiology* 70, 299–316.

Bronesky, D., Desgranges, E., Corvaglia, A., François, P., Caballero, C.J., Prado, L., Toledo-Arana, A., Lasa, I., Moreau, K., Vandenesch, F., et al. (2019). A multifaceted small RNA modulates gene expression upon glucose limitation in *Staphylococcus aureus*. *The EMBO Journal* e99363.

Caldelari, I., Chao, Y., Romby, P., and Vogel, J. (2013). RNA-Mediated Regulation in Pathogenic Bacteria. *Cold Spring Harbor Perspectives in Medicine* 3, a010298–a010298.

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 Pathog.* 6, e1000927.

Chao, Y., and Vogel, J. (2016). A 3' UTR-Derived Small RNA Provides the Regulatory Noncoding Arm of the Inner Membrane Stress Response. *Molecular Cell* 61, 352–363.

Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C.M., and Vogel, J. (2012). An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs: Hfq-dependent small RNAs from 3' UTRs. *The EMBO Journal* 31, 4005–4019.

Cheung, A.L., and Zhang, G. (2002). Global regulation of virulence determinants in *Staphylococcus aureus* by the SarA protein family. *Front. Biosci.* 7, d1825-1842.

Cheung, A.L., Bayer, A.S., Zhang, G., Gresham, H., and Xiong, Y.-Q. (2004). Regulation of virulence determinants in vitro and in vivo in *Staphylococcus aureus*. *FEMS Immunology & Medical Microbiology* 40, 1–9.

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

Chico-Calero, I., Suarez, M., Gonzalez-Zorn, B., Scotti, M., Slaghuis, J., Goebel, W., The European Listeria Genome Consortium, and Vazquez-Boland, J.A. (2002). Hpt, a bacterial homolog of the microsomal glucose- 6-phosphate translocase, mediates rapid intracellular proliferation in *Listeria*. *Proceedings of the National Academy of Sciences* 99, 431–436.

Christmas, B.A.F., Rolfe, M.D., Rose, M., and Green, J. (2019). *Staphylococcus aureus* adaptation to aerobic low-redox-potential environments: implications for an intracellular lifestyle. *Microbiology* 165, 779–791.

Cimino, G.D., Gamper, H.B., Isaacs, S.T., and Hearst, J.E. (1985). Psoralens as Photoactive Probes of Nucleic Acid Structure and Function: Organic Chemistry, Photochemistry, and Biochemistry. *Annu. Rev. Biochem.* 54, 1151–1193.

Crooke, A.K., Fuller, J.R., Obrist, M.W., Tomkovich, S.E., Vitko, N.P., and Richardson, A.R. (2013). CcpA-Independent Glucose Regulation of Lactate Dehydrogenase 1 in *Staphylococcus aureus*. *PLoS ONE* 8, e54293.

De Groote, M.A., and Fang, F.C. (1995). NO Inhibitions: Antimicrobial Properties of Nitric Oxide. *Clinical Infectious Diseases* 21, S162–S165.

De Mets, F., Van Melderden, L., and Gottesman, S. (2019). Regulation of acetate metabolism and coordination with the TCA cycle via a processed small RNA. *Proc Natl Acad Sci USA* 116, 1043–1052.

Deigan, K.E., and FerrÉ-D’AmarÉ, A.R. (2011). Riboswitches: Discovery of Drugs That Target Bacterial Gene-Regulatory RNAs. *Acc. Chem. Res.* 44, 1329–1338.

Denham, E.L. (2020). The Sponge RNAs of bacteria – How to find them and their role in regulating the post-transcriptional network. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1863, 194565.

Desgranges, E., Marzi, S., Moreau, K., Romby, P., and Caldelari, I. (2019). Noncoding RNA. *Microbiology Spectrum* 7(2) GPP3-0038-2018.

Desgranges, E., Caldelari, I., Marzi, S., and Lalaouna, D. (2020). Navigation through the twists and turns of RNA sequencing technologies: Application to bacterial regulatory RNAs. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1863, 194506.

Dumitrescu, O., Dauwalder, O., Boisset, S., Reverdy, M.-É., Tristan, A., and Vandenesch, F. (2010). Résistance aux antibiotiques chez *Staphylococcus aureus*: Les points-clés en 2010. *Med Sci (Paris)* 26, 943–949.

Durand, S., Braun, F., Lioliou, E., Romilly, C., Helfer, A.-C., Kuhn, L., Quittot, N., Nicolas, P., Romby, P., and Condon, C. (2015a). A Nitric Oxide Regulated Small RNA Controls Expression of Genes Involved in Redox Homeostasis in *Bacillus subtilis*. *PLOS Genetics* 11, e1004957.

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

Durand, S., Braun, F., Helfer, A.-C., Romby, P., and Condon, C. (2017). sRNA-mediated activation of gene expression by inhibition of 5'-3' exonucleolytic mRNA degradation. *ELife* 6.

Durand, S., Callan-Sidat, A., McKeown, J., Li, S., Kostova, G., Hernandez-Fernaund, J.R., Alam, M.T., Millard, A., Constantinidou, C., Condon, C., Denham, E.L. (2019). Novel regulation from novel interactions: Identification of an RNA sponge that controls the levels, processing and efficacy of the RoxS riboregulator of central metabolism in *Bacillus subtilis* (BioRxiv).

Dutcher, H.A., and Raghavan, R. (2018). Origin, Evolution, and Loss of Bacterial Small RNAs. *Microbiology Spectrum* 6.

Esposito, S., Purrello, S.M., Bonnet, E., Novelli, A., Tripodi, F., Pascale, R., Unal, S., and Milkovich, G. (2013). Central venous catheter-related biofilm infections: An up-to-date focus on meticillin-resistant *Staphylococcus aureus*. *Journal of Global Antimicrobial Resistance* 1, 71–78.

Ferens, W.A., and Bohach, G.A. (2000). Persistence of *Staphylococcus aureus* on mucosal membranes: Superantigens and internalization by host cells. *Journal of Laboratory and Clinical Medicine* 135, 225–230.

Figuerola-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 Dev* 23, 2004–2015.

Foster, T.J. (2005). Immune evasion by staphylococci. *Nature Reviews Microbiology* 3, 948–958.

Frantz, R., Teubner, L., Schultze, T., La Pietra, L., Müller, C., Gwozdziński, K., Pillich, H., Hain, T., Weber-Gerlach, M., Panagiotidis, G.-D., et al. (2019). The secRNome of *Listeria monocytogenes* Harbors Small Noncoding RNAs That Are Potent Inducers of Beta Interferon. *MBio* 10, e01223-19.

Fuchs, R.T., Sun, Z., Zhuang, F., and Robb, G.B. (2015). Bias in Ligation-Based Small RNA Sequencing Library Construction Is Determined by Adaptor and RNA Structure. *PLoS ONE* 10, e0126049.

Fuda, C.C.S., Fisher, J.F., and Mobashery, S. (2005). β -Lactam resistance in *Staphylococcus aureus*: the adaptive resistance of a plastic genome. *Cell. Mol. Life Sci.* 62, 2617–2633.

Gao, W., Guérillot, R., Lin, Y.H., Tree, J., Beaume, M., François, P., Monk, I.R., Seemann, T., Schrenzel, J., Howden, B.P., et al. (2020). Comparative Transcriptomic and Functional Assessments of Linezolid-Responsive Small RNA Genes in *Staphylococcus aureus*. *MSystems* 5, e00665-19.

Garzoni, C., Francois, P., Huyghe, A., Couzinet, S., Tapparel, C., Charbonnier, Y., Renzoni, A., Lucchini, S., Lew, D.P., Vaudaux, P., et al. (2007). A global view of *Staphylococcus aureus* whole genome expression upon internalization in human epithelial cells. *BMC Genomics* 8, 171.

- 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. *Nucl. Acids Res.* 37, 7239–7257.
- Gemayel, R., Fortpied, J., Rzem, R., Vertommen, D., Veiga-da-Cunha, M., and Van Schaftingen, E. (2007). Many fructosamine 3-kinase homologues in bacteria are ribulosamine/erythrulosamine 3-kinases potentially involved in protein deglycation: Bacterial fructosamine 3-kinase homologues. *FEBS Journal* 274, 4360–4374.
- Geraci, J., Neubauer, S., Pöllath, C., Hansen, U., Rizzo, F., Krafft, C., Westermann, M., Hussain, M., Peters, G., Pletz, M.W., et al. (2017). The *Staphylococcus aureus* extracellular matrix protein (Emp) has a fibrous structure and binds to different extracellular matrices. *Sci Rep* 7, 13665.
- Grosser, M.R., Weiss, A., Shaw, L.N., and Richardson, A.R. (2016). Regulatory Requirements for *Staphylococcus aureus* Nitric Oxide Resistance. *J. Bacteriol.* 198, 2043–2055.
- Guerra, F.E., Borgogna, T.R., Patel, D.M., Sward, E.W., and Voyich, J.M. (2017). Epic Immune Battles of History: Neutrophils vs. *Staphylococcus aureus*. *Front. Cell. Infect. Microbiol.* 7, 286.
- Guerrero-Mandujano, A., Hernández-Cortez, C., Ibarra, J.A., and Castro-Escarpulli, G. (2017). The outer membrane vesicles: Secretion system type zero. *Traffic* 18, 425–432.
- Halsey, C.R., Lei, S., Wax, J.K., Lehman, M.K., Nuxoll, A.S., Steinke, L., Sadykov, M., Powers, R., and Fey, P.D. (2017). Amino Acid Catabolism in *Staphylococcus aureus* and the Function of Carbon Catabolite Repression. *MBio* 8, mBio.01434-16, e01434-16.
- Hamza, T., and Li, B. (2014). Differential responses of osteoblasts and macrophages upon *Staphylococcus aureus* infection. *BMC Microbiol* 14, 207.
- Harraghy, N., Kormanec, J., Wolz, C., Homeroval, D., Goerke, C., Ohlsen, K., Qazi, S., Hill, P., and Herrmann, M. (2005). *sae* is essential for expression of the staphylococcal adhesins Eap and Emp. *Microbiology* 151, 1789–1800.
- Hecker, M., Reder, A., Fuchs, S., Pagels, M., and Engelmann, S. (2009). Physiological proteomics and stress/starvation responses in *Bacillus subtilis* and *Staphylococcus aureus*. *Research in Microbiology* 160, 245–258.
- Hindley, J. (1967). Fractionation of ³²P-labelled ribonucleic acids on polyacrylamide gels and their characterization by fingerprinting. *Journal of Molecular Biology* 30, 125–136.
- Hong, S.-W., Choi, E.-B., Min, T.-K., Kim, J.-H., Kim, M.-H., Jeon, S.G., Lee, B.-J., Gho, Y.S., Jee, Y.-K., Pyun, B.-Y., et al. (2014). An Important Role of α -Hemolysin in Extracellular Vesicles on the Development of Atopic Dermatitis Induced by *Staphylococcus aureus*. *PLoS ONE* 9, e100499.
- Howden, B.P., Beaume, M., Harrison, P.F., Hernandez, D., Schrenzel, J., Seemann, T., François, P., and 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.

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.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science* 324, 218–223.

Iosub, I.A., van Nues, R.W., McKellar, S.W., Nieken, K.J., Marchioreto, M., Sy, B., Tree, J.J., Viero, G., and Granneman, S. (2020). Hfq CLASH uncovers sRNA-target interaction networks linked to nutrient availability adaptation. *ELife* 9, e54655.

Jefferson, K.K., Cramton, S.E., Götz, F., and Pier, G.B. (2003). Identification of a 5-nucleotide sequence that controls expression of the *ica* locus in *Staphylococcus aureus* and characterization of the DNA-binding properties of IcaR: Regulation of *ica* in *S. aureus*. *Mol Microbiol* 48, 889–899.

Jefferson, K.K., Pier, D.B., Goldmann, D.A., and Pier, G.B. (2004). The Teicoplanin-Associated Locus Regulator (TcaR) and the Intercellular Adhesin Locus Regulator (IcaR) Are Transcriptional Inhibitors of the *ica* Locus in *Staphylococcus aureus*. *Journal of Bacteriology* 186, 2449–2456.

Kim, H.M., Shin, J.-H., Cho, Y.-B., and Roe, J.-H. (2014). Inverse regulation of Fe- and Ni-containing SOD genes by a Fur family regulator Nur through small RNA processed from 3'UTR of the *sodF* mRNA. *Nucleic Acids Research* 42, 2003–2014.

King, A.N., de Mets, F., and Brinsmade, S.R. (2020). Who's in control? Regulation of metabolism and pathogenesis in space and time. *Current Opinion in Microbiology* 55, 88–96.

Kirby, W.M.M. (1944). Extraction of a highly potent penicillin inactivator from penicillin resistant *Staphylococci*. *Science* 99, 452–453.

Koeppen, K., Hampton, T.H., Jarek, M., Scharfe, M., Gerber, S.A., Mielcarz, D.W., Demers, E.G., Dolben, E.L., Hammond, J.H., Hogan, D.A., et al. (2016). A Novel Mechanism of Host-Pathogen Interaction through sRNA in Bacterial Outer Membrane Vesicles. *PLoS Pathog* 12, e1005672.

Kroin, J.S., Buvanendran, A., Li, J., Moric, M., Im, H.-J., Tuman, K.J., and Shafikhani, S.H. (2015). Short-Term Glycemic Control Is Effective in Reducing Surgical Site Infection in Diabetic Rats: *Anesthesia & Analgesia* 120, 1289–1296.

Kulhankova, K., King, J., and Salgado-Pabón, W. (2014). Staphylococcal toxic shock syndrome: superantigen-mediated enhancement of endotoxin shock and adaptive immune suppression. *Immunol Res* 59, 182–187.

Lagos-Quintana, M. (2001). Identification of Novel Genes Coding for Small Expressed RNAs. *Science* 294, 853–858.

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. *Genom Data* 5, 136–138.

- Lalaouna, D., Carrier, M.-C., Semsey, S., Brouard, J.-S., Wang, J., Wade, J.T., and Massé, E. (2015). A 3' External Transcribed Spacer in a tRNA Transcript Acts as a Sponge for Small RNAs to Prevent Transcriptional Noise. *Molecular Cell* 58, 393–405.
- Lalaouna, D., Desgranges, E., Caldelari, I., and Marzi, S. (2018a). MS2-Affinity Purification Coupled With RNA Sequencing Approach in the Human Pathogen *Staphylococcus aureus*. In *Methods in Enzymology*, (Elsevier), pp. 393–411.
- Lalaouna, D., Eyraud, A., Devinck, A., Prévost, K., and Massé, E. (2018b). GcvB small RNA uses two distinct seed regions to regulate an extensive targetome. *Mol Microbiol* mmi.14168.
- Lalaouna, D., Baude, J., Wu, Z., Tomasini, A., Chicher, J., Marzi, S., Vandenesch, F., Romby, P., Caldelari, I., and Moreau, K. (2019). RsaC sRNA modulates the oxidative stress response of *Staphylococcus aureus* during manganese starvation. *Nucleic Acids Research* 47, 9871–9887.
- Lasa, I., Toledo-Arana, A., Dobin, A., Villanueva, M., Mozos, I.R. de los, Vergara-Irigaray, M., Segura, V., Fagegaltier, D., Penadés, J.R., Valle, J., et al. (2011). Genome-wide antisense transcription drives mRNA processing in bacteria. *PNAS* 108, 20172–20177.
- Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843–854.
- Lee, S.-R., Kim, S.-H., Jeong, K.-J., Kim, K.-S., Kim, Y.-H., Kim, S.-J., Kim, E., Kim, J.-W., and Chang, K.-T. (2009). Multi-immunogenic outer membrane vesicles derived from an MsbB-deficient *Salmonella enterica* serovar typhimurium mutant. *J. Microbiol. Biotechnol.* 19, 1271–1279.
- Le Pabic, H., Germain-Amiot, N., Bordeau, V., and Felden, B. (2015). A bacterial regulatory RNA attenuates virulence, spread and human host cell phagocytosis. *Nucleic Acids Research* 43, 9232–9248.
- 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 Genet* 8, e1002782.
- Liu, T., Zhang, K., Xu, S., Wang, Z., Fu, H., Tian, B., Zheng, X., and Li, W. (2017). Detecting RNA-RNA interactions in *E. coli* using a modified CLASH method. *BMC Genomics* 18.
- Liu, Y., Mu, C., Ying, X., Li, W., Wu, N., Dong, J., Gao, Y., Shao, N., Fan, M., and Yang, G. (2011). RNAlIII activates map expression by forming an RNA-RNA complex in *Staphylococcus aureus*. *FEBS Lett.* 585, 899–905.
- López-Garrido, J., Puerta-Fernández, E., and Casadesús, J. (2014). A eukaryotic-like 3' untranslated region in *Salmonella enterica* *hilD* mRNA. *Nucleic Acids Research* 42, 5894–5906.
- Lowy, F.D. (1998). *Staphylococcus aureus* infections. *N. Engl. J. Med.* 339, 520–532.
- Lu, Z., Zhang, Q.C., Lee, B., Flynn, R.A., Smith, M.A., Robinson, J.T., Davidovich, C., Gooding, A.R., Goodrich, K.J., Mattick, J.S., et al. (2016). RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure. *Cell* 165, 1267–1279.

- Lustig, Y., Wachtel, C., Safro, M., Liu, L., and Michaeli, S. (2010). ‘RNA walk’ a novel approach to study RNA–RNA interactions between a small RNA and its target. *Nucleic Acids Research* 38, e5–e5.
- Majerczyk, C.D., Dunman, P.M., Luong, T.T., Lee, C.Y., Sadykov, M.R., Somerville, G.A., Bodi, K., and Sonenshein, A.L. (2010). Direct Targets of CodY in *Staphylococcus aureus*. *JB* 192, 2861–2877.
- Maki, H., McCallum, N., Bischoff, M., Wada, A., and Berger-Bächi, B. (2004). *tcaA* Inactivation Increases Glycopeptide Resistance in *Staphylococcus aureus*. *AAC* 48, 1953–1959.
- Marchais, A., Bohn, C., Boulloc, P., and Gautheret, D. (2010). RsaOG, a new staphylococcal family of highly transcribed non-coding RNA. *RNA Biol* 7, 116–119.
- Marincola, G., Wencker, F.D.R., and Ziebuhr, W. (2019). The Many Facets of the Small Non-coding RNA RsaE (RoxS) in Metabolic Niche Adaptation of Gram-Positive Bacteria. *Journal of Molecular Biology* 431, 4684–4698.
- McCallum, N., Bischoff, M., Maki, H., Wada, A., and Berger-Bächi, B. (2004). TcaR, a Putative MarR-Like Regulator of *sarS* Expression. *JB* 186, 2966–2972.
- McGuinness, W.A., Malachowa, N., and DeLeo, F.R. Vancomycin Resistance in *Staphylococcus aureus*. *Yale J Biol Med* 90(2), 269–281.
- Mei, J., Nourbakhsh, F., Ford, C.W., and Holden, D.W. (1997). Identification of *Staphylococcus aureus* virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. *Mol Microbiol* 26, 399–407.
- Melamed, S., Peer, A., Faigenbaum-Romm, R., Gatt, Y.E., Reiss, N., Bar, A., Altuvia, Y., Argaman, L., and Margalit, H. (2016). Global Mapping of Small RNA-Target Interactions in Bacteria. *Molecular Cell* 63, 884–897.
- Melter, O., and Radojevič, B. (2010). Small colony variants of *Staphylococcus aureus* — review. *Folia Microbiol* 55, 548–558.
- Menendez-Gil, P., Caballero, C.J., Catalan-Moreno, A., Irurzun, N., Barrio-Hernandez, I., Caldelari, I., and Toledo-Arana, A. (2020). Differential evolution in 3’UTRs leads to specific gene expression in *Staphylococcus*. *Nucleic Acids Research* 48, 2544–2563.
- Michalik, S., Depke, M., Murr, A., Gesell Salazar, M., Kusebauch, U., Sun, Z., Meyer, T.C., Surmann, K., Pförtner, H., Hildebrandt, P., et al. (2017). A global *Staphylococcus aureus* proteome resource applied to the in vivo characterization of host-pathogen interactions. *Sci Rep* 7, 9718.
- 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. *EMBO J.* 34, 1478–1492.
- Mockly, S., and Seitz, H. (2019). Inconsistencies and Limitations of Current MicroRNA Target Identification Methods. *Methods Mol Biol* 1970, 291–314.

- Montgomery, C.P., Boyle-Vavra, S., Adem, P.V., Lee, J.C., Husain, A.N., Clasen, J., and Daum, R.S. (2008). Comparison of Virulence in Community-Associated Methicillin-Resistant *Staphylococcus aureus* Pulsotypes USA300 and USA400 in a Rat Model of Pneumonia. *The Journal of Infectious Diseases* 198, 561–570.
- Morrison, J.M., Miller, E.W., Benson, M.A., Alonzo, F., 3rd, Yoong, P., Torres, V.J., Hinrichs, S.H., and Dunman, P.M. (2012). 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.
- Nitzan, M., Fechter, P., Peer, A., Altuvia, Y., Bronesky, D., Vandenesch, F., Romby, P., Biham, O., and Margalit, H. (2015a). A defense-offense multi-layered regulatory switch in a pathogenic bacterium. *Nucleic Acids Research* 43, 1357–1369.
- Nitzan, M., Fechter, P., Peer, A., Altuvia, Y., Bronesky, D., Vandenesch, F., Romby, P., Biham, O., and Margalit, H. (2015b). A defense-offense multi-layered regulatory switch in a pathogenic bacterium. *Nucleic Acids Research* 43, 1357–1369.
- Nitzan, M., Rehani, R., and Margalit, H. (2017). Integration of Bacterial Small RNAs in Regulatory Networks. *Annu. Rev. Biophys.* 46, 131–148.
- Novick, R.P., and Geisinger, E. (2008). Quorum Sensing in *Staphylococci*. *Annual Review of Genetics* 42, 541–564.
- Olejniczak, M., and Storz, G. (2017). ProQ/FinO-domain proteins: another ubiquitous family of RNA matchmakers?. *Mol Microbiol* 104, 905–915.
- Olson, M.E., King, J.M., Yahr, T.L., and Horswill, A.R. (2013). Sialic Acid Catabolism in *Staphylococcus aureus*. *Journal of Bacteriology* 195, 1779–1788.
- Otto, M.P., Martin, E., Badiou, C., Lebrun, S., Bes, M., Vandenesch, F., Etienne, J., Lina, G., and Dumitrescu, O. (2013). Effects of subinhibitory concentrations of antibiotics on virulence factor expression by community-acquired methicillin-resistant *Staphylococcus aureus*. *Journal of Antimicrobial Chemotherapy* 68, 1524–1532.
- Pagels, M., Fuchs, S., Pané-Farré, J., Kohler, C., Menschner, L., Hecker, M., McNamarra, P.J., Bauer, M.C., Von Wachenfeldt, C., Liebeke, M., et al. (2010). Redox sensing by a Rex-family repressor is involved in the regulation of anaerobic gene expression in *Staphylococcus aureus*: Redox sensitive gene regulation in *S. aureus*. *Mol Microbiol* 76, 1142–1161.
- Palma Medina, L.M., Becker, A.-K., Michalik, S., Yedavally, H., Raineri, E.J.M., Hildebrandt, P., Gesell Salazar, M., Surmann, K., Pförtner, H., Mekonnen, S.A., et al. (2019). Metabolic Cross-talk Between Human Bronchial Epithelial Cells and Internalized *Staphylococcus aureus* as a Driver for Infection. *Mol Cell Proteomics* 18, 892–908.
- Park, J.Y., Kim, J.W., Moon, B.Y., Lee, J., Fortin, Y.J., Austin, F.W., Yang, S.-J., and Seo, K.S. (2015). Characterization of a Novel Two-Component Regulatory System, HptRS, the Regulator for the Hexose Phosphate Transport System in *Staphylococcus aureus*. *Infect Immun* 83, 1620–1628.
- Périchon, B., and Courvalin, P. (2009). VanA-Type Vancomycin-Resistant *Staphylococcus aureus*. *AAC* 53, 4580–4587.

Peyrusson, F., Varet, H., Nguyen, T.K., Legendre, R., Sismeiro, O., Coppée, J.-Y., Wolz, C., Tenson, T., and Van Bambeke, F. (2020). Intracellular *Staphylococcus aureus* persists upon antibiotic exposure. *Nat Commun* 11, 2200.

Potter, A.D., Butrico, C.E., Ford, C.A., Curry, J.M., Trenary, I.A., Tummarakota, S.S., Hendrix, A.S., Young, J.D., and Cassat, J.E. (2020). Host nutrient milieu drives an essential role for aspartate biosynthesis during invasive *Staphylococcus aureus* infection. *Proc Natl Acad Sci USA* 117, 12394–12401.

Proctor, R.A., Kahl, B., von Eiff, C., Vaudaux, P.E., Lew, D.P., and Peters, G. (1998). Staphylococcal Small Colony Variants Have Novel Mechanisms for Antibiotic Resistance. *Clin Infect Dis* 27, S68–S74.

Qin, L., McCausland, J.W., Cheung, G.Y.C., and Otto, M. (2016). PSM-Mec—A Virulence Determinant that Connects Transcriptional Regulation, Virulence, and Antibiotic Resistance in *Staphylococci*. *Frontiers in Microbiology* 7.

Raghavan, R., Kacharia, F.R., Millar, J.A., Sislak, C.D., and Ochman, H. (2015). Genome Rearrangements Can Make and Break Small RNA Genes. *Genome Biology and Evolution* 7, 557–566.

Reed, J.M., Olson, S., Brees, D.F., Griffin, C.E., Grove, R.A., Davis, P.J., Kachman, S.D., Adamec, J., and Somerville, G.A. (2018). Coordinated regulation of transcription by CcpA and the *Staphylococcus aureus* two-component system HptRS. *PLoS ONE* 13, e0207161.

Rich, J., and Lee, J.C. (2005). The Pathogenesis of *Staphylococcus aureus* Infection in the Diabetic NOD Mouse. *Diabetes* 54, 2904–2910.

Richards, G.R., and Vanderpool, C.K. (2011). Molecular call and response: The physiology of bacterial small RNAs. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1809, 525–531.

Richardson, A.R. (2019). Virulence and Metabolism. *Microbiology Spectrum* 7(2) GPP3-011-2018.

Richardson, A.R., Dunman, P.M., and Fang, F.C. (2006). The nitrosative stress response of *Staphylococcus aureus* is required for resistance to innate immunity. *Mol Microbiol* 61, 927–939.

Richardson, A.R., Libby, S.J., and Fang, F.C. (2008a). A Nitric Oxide-Inducible Lactate Dehydrogenase Enables *Staphylococcus aureus* to Resist Innate Immunity. *Science* 319, 1672–1676.

Richardson, A.R., Libby, S.J., and Fang, F.C. (2008b). A Nitric Oxide-Inducible Lactate Dehydrogenase Enables *Staphylococcus aureus* to Resist Innate Immunity. *Science* 319, 1672–1676.

Richardson, A.R., Somerville, G.A., and Sonenshein, A.L. (2015). Regulating the Intersection of Metabolism and Pathogenesis in Gram-positive Bacteria. *Microbiol Spectr*, 3(3), 129–165.

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

Rivas, E., Klein, R.J., Jones, T.A., and Eddy, S.R. (2001). Computational identification of noncoding RNAs in *E. coli* by comparative genomics. *Current Biology* 11, 1369–1373.

Roberts, C., Anderson, K.L., Murphy, E., Projan, S.J., Mounts, W., Hurlburt, B., Smeltzer, M., Overbeek, R., Disz, T., and 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., Bohn, C., Morvan, C., Le Lam, T.N., Razvi, F., Pain, A., Toffano-Nioche, C., Ponien, P., Jacq, A., Jacquet, E., et al. (2018). The conserved regulatory RNA RsaE down-regulates the arginine degradation pathway in *Staphylococcus aureus*. *Nucleic Acids Research* 46, 8803–8816.

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 Pathog* 10.

Rosenow, C. (2001). Prokaryotic RNA preparation methods useful for high density array analysis: comparison of two approaches. *Nucleic Acids Research* 29, 112e–1112.

Rowe, S.E., Wagner, N.J., Li, L., Beam, J.E., Wilkinson, A.D., Radlinski, L.C., Zhang, Q., Miao, E.A., and Conlon, B.P. (2020). Reactive oxygen species induce antibiotic tolerance during systemic *Staphylococcus aureus* infection. *Nat Microbiol* 5, 282–290.

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.

Runyen-Janecky, L.J., and Payne, S.M. (2002). Identification of Chromosomal *Shigella flexneri* Genes Induced by the Eukaryotic Intracellular Environment. *Infect Immun* 70, 4379–4388.

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 Res* 37, e133.

Schoenfelder, S.M.K., Lange, C., Prakash, S.A., Marincola, G., Lerch, M.F., Wencker, F.D.R., Förstner, K.U., Sharma, C.M., and Ziebuhr, W. (2019). The small non-coding RNA RsaE influences extracellular matrix composition in *Staphylococcus epidermidis* biofilm communities. *PLoS Pathog* 15, e1007618.

Scott, M., Gunderson, C.W., Mateescu, E.M., Zhang, Z., and Hwa, T. (2010). Interdependence of Cell Growth and Gene Expression: Origins and Consequences. *Science* 330, 1099–1102.

Seidl, K., Stucki, M., Ruegg, M., Goerke, C., Wolz, C., Harris, L., Berger-Bächi, B., and Bischoff, M. (2006). *Staphylococcus aureus* CcpA Affects Virulence Determinant Production and Antibiotic Resistance. *Antimicrobial Agents and Chemotherapy* 50, 1183–1194.

Seidl, K., Bischoff, M., and Berger-Bächi, B. (2008a). CcpA Mediates the Catabolite Repression of *tst* in *Staphylococcus aureus*. *Infect Immun* 76, 5093–5099.

- Seidl, K., Goerke, C., Wolz, C., Mack, D., Berger-Bächi, B., and Bischoff, M. (2008b). *Staphylococcus aureus* CcpA Affects Biofilm Formation. *Infect Immun* 76, 2044–2050.
- Seidl, K., Müller, S., François, P., Kriebitzsch, C., Schrenzel, J., Engelmann, S., Bischoff, M., and Berger-Bächi, B. (2009). Effect of a glucose impulse on the CcpA regulon in *Staphylococcus aureus*. *BMC Microbiol* 9, 95.
- Seitz, H. (2009). Redefining MicroRNA Targets. *Current Biology* 19, 870–873.
- Seitz, H. (2017). Issues in current microRNA target identification methods. *RNA Biology* 14, 831–834.
- 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.
- Shimoni, Y., Friedlander, G., Hetzroni, G., Niv, G., Altuvia, S., Biham, O., and Margalit, H. (2007). Regulation of gene expression by small non-coding RNAs: a quantitative view. *Mol. Syst. Biol.* 3, 138.
- Sit, B., Crowley, S.M., Bhullar, K., Lai, C.C.-L., Tang, C., Hooda, Y., Calmettes, C., Khambati, H., Ma, C., Brumell, J.H., et al. (2015). Active Transport of Phosphorylated Carbohydrates Promotes Intestinal Colonization and Transmission of a Bacterial Pathogen. *PLoS Pathog* 11, e1005107.
- Sjöström, A.E., Sandblad, L., Uhlin, B.E., and Wai, S.N. (2015). Membrane vesicle-mediated release of bacterial RNA. *Sci Rep* 5, 15329.
- Sorensen, H.M., Keogh, R.A., Wittekind, M.A., Caillet, A.R., Wiemels, R.E., Laner, E.A., and Carroll, R.K. (2020). Reading between the Lines: Utilizing RNA-Seq Data for Global Analysis of sRNAs in *Staphylococcus aureus*. *MSphere* 5.
- Speziale, P., Pietrocola, G., Foster, T.J., and Geoghegan, J.A. (2014). Protein-based biofilm matrices in *Staphylococci*. *Front. Cell. Infect. Microbiol.* 4.
- Sun, J., Zheng, L., Landwehr, C., Yang, J., and Ji, Y. (2005). Identification of a Novel Essential Two-Component Signal Transduction System, YhcSR, in *Staphylococcus aureus*. *JB* 187, 7876–7880.
- Tomasini, A., Moreau, K., Chicher, J., Geissmann, T., Vandenesch, F., Romby, P., Marzi, S., and Caldelari, I. (2017). The RNA targetome of *Staphylococcus aureus* non-coding RNA RsaA: impact on cell surface properties and defense mechanisms. *Nucleic Acids Research* 45, 6746–6760.
- Tomizawa, J., and Itoh, T. (1981). Plasmid ColE1 incompatibility determined by interaction of RNA I with primer transcript. *Proceedings of the National Academy of Sciences* 78, 6096–6100.
- Ubeda, C., Djukovic, A., and Isaac, S. (2017). Roles of the intestinal microbiota in pathogen protection. *Clin Trans Immunol* 6, e128.

Valle, J., Toledo-Arana, A., Berasain, C., Ghigo, J.-M., Amorena, B., Penadés, J.R., and Lasa, I. (2003). SarA and not σ^B is essential for biofilm development by *Staphylococcus aureus*: *Staphylococcus aureus* development of biofilm. *Mol Microbiol* 48, 1075–1087.

Villanueva, M., García, B., Valle, J., Rapún, B., Ruiz de los Mozos, I., Solano, C., Martí, M., Penadés, J.R., Toledo-Arana, A., and Lasa, I. (2018). Sensory deprivation in *Staphylococcus aureus*. *Nat Commun* 9, 523.

Vitko, N.P., Grosser, M.R., Khatri, D., Lance, T.R., and Richardson, A.R. (2016). Expanded Glucose Import Capability Affords *Staphylococcus aureus* Optimized Glycolytic Flux during Infection. *MBio* 7.

Vogel, J. (2020). An RNA biology perspective on species-specific programmable RNA antibiotics. *Mol Microbiol* 113, 550–559.

Wachutka, L., and Gagneur, J. (2017). Measures of RNA metabolism rates: Toward a definition at the level of single bonds. *Transcription* 8, 75–80.

Wang, B., Yao, M., Lv, L., Ling, Z., and Li, L. (2017). The Human Microbiota in Health and Disease. *Engineering* 3, 71–82.

Wang, C., Chao, Y., Matera, G., Gao, Q., and Vogel, J. (2020). The conserved 3' UTR-derived small RNA NarS mediates mRNA crossregulation during nitrate respiration. *Nucleic Acids Research* 48, 2126–2143.

Wang, E., Bauer, M.C., Rogstam, A., Linse, S., Logan, D.T., and von Wachenfeldt, C. (2008). Structure and functional properties of the *Bacillus subtilis* transcriptional repressor Rex. *Mol Microbiol* 69, 466–478.

Wang, Y., Bojer, M.S., George, S.E., Wang, Z., Jensen, P.R., Wolz, C., and Ingmer, H. (2018). Inactivation of TCA cycle enhances *Staphylococcus aureus* persister cell formation in stationary phase. *Sci Rep* 8, 10849.

Wassarman, K.M. (2001). Identification of novel small RNAs using comparative genomics and microarrays. *Genes & Development* 15, 1637–1651.

Wassarman, K.M., and Storz, G. (2000). 6S RNA Regulates *E. coli* RNA Polymerase Activity. *Cell* 101, 613–623.

Waters, S.A., McAteer, S.P., Kudla, G., Pang, I., Deshpande, N.P., Amos, T.G., Leong, K.W., Wilkins, M.R., Strugnell, R., Gally, D.L., et al. (2017). Small RNA interactome of pathogenic *E. coli* revealed through crosslinking of RNase E. *The EMBO Journal* 36, 374–387.

Wertheim, H.F., Melles, D.C., Vos, M.C., van Leeuwen, W., van Belkum, A., Verbrugh, H.A., and Nouwen, J.L. (2005). The role of nasal carriage in *Staphylococcus aureus* infections. *The Lancet Infectious Diseases* 5, 751–762.

Wu, S., Lin, K., Liu, Y., Zhang, H., and Lei, L. (2020). Two-component signaling pathways modulate drug resistance of. *Biom Rep*.

Yanagihara, K., Tashiro, M., Fukuda, Y., Ohno, H., Higashiyama, Y., Miyazaki, Y., Hirakata, Y., Tomono, K., Mizuta, Y., Tsukamoto, K., et al. (2006). Effects of short interfering RNA

against methicillin-resistant *Staphylococcus aureus* coagulase in vitro and in vivo. *Journal of Antimicrobial Chemotherapy* 57, 122–126.

Yang, Y., Sun, H., Liu, X., Wang, M., Xue, T., and Sun, B. (2016). Regulatory mechanism of the three-component system HptRSA in glucose-6-phosphate uptake in *Staphylococcus aureus*. *Med Microbiol Immunol* 205, 241–253.

Zapf, R.L., Wiemels, R.E., Keogh, R.A., Holzschu, D.L., Howell, K.M., Trzeciak, E., Caillet, A.R., King, K.A., Selhorst, S.A., Naldrett, M.J., et al. (2019). The Small RNA Teg41 Regulates Expression of the Alpha Phenol-Soluble Modulins and Is Required for Virulence in *Staphylococcus aureus*. *MBio* 10.

Étude de l'interactome ARN-ARN chez *Staphylococcus aureus* : caractérisation, fonction et impact sur les réseaux de régulation

Staphylococcus aureus est un pathogène opportuniste de l'homme, responsable de multiples infections en colonisant un grand nombre de tissus et d'organes. Cette bactérie a développé un large éventail de stratégies qui lui permettent de s'adapter rapidement en réponse à divers stress et changements de l'environnement. En plus des facteurs de transcription et des systèmes à deux composants, diverses études révèlent que les ARN régulateurs principalement non-codants (sARN) régulent de manière dynamique l'expression de gènes codant pour des facteurs de virulence, des régulateurs transcriptionnels, et des enzymes du métabolisme. Ils agissent bien souvent par le biais d'interaction entre paires de bases avec leurs ARN messagers (ARNm) cibles, en utilisant un motif riche en C.

Le but premier de mon projet de thèse est de caractériser la fonction de deux sARN, appelés RsaG et RsaI. Premièrement, RsaG est un sARN dérivé de la région 3' non traduite de l'ARNm *uhpT*, codant pour un transporteur du glucose-6-phosphate (G6P) exprimé en réponse à la concentration extracellulaire de G6P, une source de carbone majeur. Ensuite, l'expression de RsaI, dépend également de la présence d'un autre sucre d'importance. En effet, RsaI est inhibé par CcpA (régulateur impliqué dans le contrôle catabolique) en présence de fortes concentrations de glucose. Le targetome des deux sARN a été défini par MAPS et transcriptomique. Les résultats révèlent une interaction originale entre les deux sARN. L'étude dans son entièreté semble indiquer que l'expression de RsaI est une signature d'un changement métabolique qui s'opère chez *S. aureus* lorsque la source de carbone préféré est épuisée du milieu. RsaG quant à lui, est apparu comme connectant l'homéostasie rédox avec l'adaptation métabolique.

Deuxièmement, le projet avait aussi comme objectif d'établir une méthode pour cartographier l'ensemble des interactions ARN-ARN à l'échelle du génome de *S. aureus*, dans le but d'obtenir une vue globale de la physiologie de cet organisme.

Mots clés : *Staphylococcus aureus*, ARN régulateurs, métabolisme, interactome.

Staphylococcus aureus is an opportunistic pathogen that causes multiple infections by colonizing a large number of tissues and organs. This bacterium has developed numerous strategies that allow its quick adaptation in response to various stresses and changes in the environment. In addition to transcriptional factors and two-component systems, several studies revealed that non-coding RNAs (sRNAs) regulate in a dynamic manner the expression of numerous genes encoding virulence factors, transcriptional regulators, and metabolic enzymes. They often act through direct base-pairing with their target messenger RNAs (mRNAs) using a C-rich motif as a seed sequence.

The primary aim of the project is to characterize the function of two peculiar sRNAs, called RsaG and RsaI. First, RsaG is a so-called 3'UTR derived sRNA because it is located in the 3' untranslated region of the *uhpT* gene encoding a glucose-6-phosphate (G6P) transporter expressed in response to extracellular G6P, a major carbon source. Then the expression of RsaI, also depends on the presence of another important sugar. Indeed, RsaI is inhibited by the carbon catabolite protein A (CcpA) in the presence of high concentrations of glucose. The targetome of both sRNA was defined by MAPS and transcriptomic analysis. The results revealed an original interaction between the two sRNAs. In addition, the study seems to indicate that RsaI expression provides a signature for the bacterial metabolic switch when the preferred carbon source is depleted. RsaG on the other hand appears to connect redox homeostasis and metabolism adaptation.

Secondly, the project also attempts to establish an approach to map all RNA-RNA genome-wide interactions of *S. aureus* in order to obtain a global view of the physiology of this organism.

Keywords: *Staphylococcus aureus*, regulatory RNA, metabolism, interactome.