



# UNIVERSITÉ DE STRASBOURG



**ÉCOLE DOCTORALE des Sciences de la Vie et de la Santé**  
**UPR 9002-ARN, Architecture et Réactivité de l'ARN, IBMC**

## **THÈSE** présentée par : **Cédric ROMILLY**

soutenue le : **14 septembre 2012**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**  
Discipline/ Spécialité : **Aspects Moléculaire et Cellulaires de la Biologie**

**Fonction de nouveaux ARN non codant dans la  
régulation de l'expression des gènes de  
*Staphylococcus aureus* :  
Adaptation à l'environnement et virulence**

**THÈSE dirigée par :**

**Madame ROMBY Pascale**

Docteur, IBMC-Strasbourg

**RAPPORTEURS :**

**Monsieur Linder Patrick**

Professeur, CEBUG-Genève

**Monsieur MSADEK Tarek**

Professeur, Institut Pasteur-Paris

**Madame FLORENTZ Catherine**

Professeur, IBMC-Strasbourg

---

**AUTRES MEMBRES DU JURY :**

**Madame GUILLIER Maude**

Docteur, IBPC-Paris

## Remerciements

Pour être honnête, je pensais que cette partie serait la plus facile, mais comme je l'ai découvert pendant ces 4 années passées ici, beaucoup de choses paraissent faciles mais peu le sont au final.

Je tiens tout d'abord à remercier les membres de mon jury, Patrick Linder, Tarek Msadek, Catherine Florentz et Maude Guillier d'avoir accepté d'évaluer ce travail.

Merci également à Eric Westhof de m'avoir accueilli au sein de l'UPR 9002, et je tiens également à remercier la DGA, la FRM et la région Alsace pour le financement de ma thèse.

Ensuite, j'aimerais te remercier, Pascale, pour m'avoir donné ma chance. Si je peux présenter ce travail aujourd'hui, c'est grâce à toi. Merci pour ton accueil, tes conseils, ton soutien dans les moments difficiles. Merci pour ta gentillesse, ta disponibilité, ces opportunités incroyables de ~~voyages~~ congrès (et de prendre l'avion pour la première fois, d'être malade en avion pour la première fois aussi...) qui ont façonnés le doctorant que je suis.

Je remercie aussi tous les membres passés et présents de l'équipe "ARN messagers et ARN régulateurs procaryotiques" : Clément et Tom pour m'avoir encadré et aidé à avoir confiance en moi (malgré les belles bêtises que j'ai pu faire...), Pierre pour ton aide sur la protéomique qui a débloqué la situation, Isabelle pour tes conseils en microbiologie, Efi pour tout un tas de raisons scientifiques mais aussi pour la découverte du Jack Daniel's et le réveil le plus rapide de l'histoire de la science, Brian avec qui je partage un certain attrait pour les films de zombies, Anne Catherine et Marie pour votre aide et vos conseils, à nos collaborateurs de Lyon, Paris et aussi Philippe pour la protéomique et enfin Stephano pour nos parties de foot endiablées qui vont beaucoup me manquer. Enfin merci à Arnaud, Delphine et Mélo "Téqui" pour les nombreux fous rires qui participent tant à la vie du labo, pour ces discussions politiques/cinéma/jeuxvidéo/geek/etc, je vous souhaite bon vent pour l'avenir.

J'aimerais aussi remercier l'ensemble du DU-MA-NE-RO pour l'aide et la superbe ambiance de travail qui règne ici, les BBQ, la pause café et thé et le pot du vendredi soir : JC,

Santi, Red, Patrick, Valérie, Julien, Serena (et son Cup collector Will & Kate à l'origine de tant de discussions royales), Catherine, Roland, Eric, Guillaume, Philippe (les trois), Dom', Joëlle, Sondes, Sarah, Matéo, Paola, Vidjay, Flore, Daniela, Delphine et les "anciens" Gaëlle, Emilie, Séverine, Céline, Lucille, Mag (Pardon pour ceux que j'oublie...)

Merci à Joern et Fabrice pour votre accueil et vos conseils pour le monitorat, et Fred, Anne, Paul et j'en passe pour la superbe ambiance qui y régnait.

Merci à Denise, Emiko, Wei, Shuntaro, Yusuke, et Catherine, ce premier voyage au Japon a eu une telle influence sur le reste de ma thèse.

Merci aussi à Simon, ami de toujours de "Kertzfeld" pour tous les lundis soir détente et tes conseils culinaires, oenologiques et cinéphiles et merci aux Flying Racoons pour ces dimanches soir qui font si bien commencer la semaine,

Merci à Jeannine, Rémy et Hélène pour votre soutien et tous ces bons moments.

Merci à toi Anne, amie de tous les jours et pour toujours. Merci pour tes coups de pieds aux f%\$%= à la fin de la première année de fac, sans ton aide et ton soutien, je ne serais pas là où je suis. Merci pour nos vendredis jogging endiablés (j'exagère un peu là...), et nouveaux ans, d'avoir été présente dans les moments de déprime et de galères, merci pour les shifumis et la blague sur les chaussettes, bref merci pour tout. On a passé toute l'université ensemble, les semaines de révisions, les exams puis la thèse. Et cette dernière se finira la même semaine...si ce n'est pas beau ça ?

Merci aussi aux amis de la fac, Vince, Opi, Yann, Quiongyue, Titi et aux multiples Célines.

Enfin, merci à ma famille. Papa, Maman, merci de m'avoir donné la liberté de choisir ma voie, merci pour votre soutien et votre présence à mes côtés à chaque fois que j' ai eu besoin. Christelle, Lili, Séb, Tib' et surtout le petit dernier, Loris, tous ces moments passés avec vous m'ont toujours redonné de l'énergie pour continuer à aller de l'avant. Votre soutien, votre compréhension comptent tellement pour moi. Bienvenue au petit dernier qui devrait pointer le bout de son nez très rapidement après la soutenance.

Je crois que je vais m'arrêter ici, j'espère n'avoir oublié personne, et si tel est les cas je m'en excuse.

Ced 436

## Table des matières

### Introduction

<b>I. L'incroyable pouvoir d'adaptation des procaryotes.....</b>	<b>3</b>
I.A. Adaptation à l'environnement.....	3
I.B. Des liens étroits avec l'homme.....	5
<b>II. Nouvelles classes d'ARN chez les bactéries.....</b>	<b>8</b>
II.A. Découverte des sARN ("small RNA") exprimés dans les régions intergéniques et des ARN antisens, codés en <i>cis</i> .....	8
II.B. Nouveaux mécanismes de transcription des gènes.....	10
II.C. Quels sont les rôles de ces ARN?.....	12
II.D. Dégradation des ARN.....	13
<b>III. La régulation de l'expression des gènes par l'ARN.....</b>	<b>17</b>
III.A. Avantages de la régulation via l'ARN.....	17
III.A.1. Une réponse rapide aux conditions de stress.....	17
III.A.2. Circuits de régulation.....	20
III.A.3. Localisation intracellulaire des ARN et dynamisme des populations.....	23
III.B. Mécanismes d'action des ARN régulateurs.....	25
III.B.1. Les éléments agissant en <i>cis</i> .....	26
III.B.1.a. Les riboswitch.....	26
III.B.1.b. Les thermosenseurs.....	31
III.B.2. Les ARN régulateurs ciblant les ARNm.....	33
III.B.2.a. Les ARN antisens.....	34
III.B.2.b. Les sARN agissant en <i>trans</i> .....	38
III.B.3. Les sARN séquestrant des protéines.....	43
III.C. Machineries protéiques associées à la régulation par les sARN.....	46
III.C.1. La protéine Hfq.....	46
III.C.2. Autres machineries protéiques associées aux sARN.....	50
<b>IV. Staphylococcus aureus.....</b>	<b>51</b>
IV.A. Un pathogène opportuniste remarquable.....	51
IV.B. Facteurs de virulence.....	53
IV.C. Régulation de l'expression des facteurs de virulence.....	55
IV.C.1. Les facteurs $\sigma$ .....	57
IV.C.2. Facteurs de transcription.....	57
IV.C.3. Systèmes à deux composants.....	58
IV.D. ARN régulateurs chez <i>S. aureus</i> .....	60
IV.D.1. Etat des lieux.....	61
IV.D.2. L'ARNIII, effecteur intracellulaire du système de densité cellulaire agr.....	63
IV.D.3. RsaE et la régulation du métabolisme central.....	65
IV.D.4. Îlots de Pathogénie : rôle de SprD dans la virulence.....	66
IV.D.5. SprA1-AS, régulateur d'un système toxine-antitoxine.....	67

## Objectifs

## Résultats

### Mécanisme de reconnaissance des interactions boucle-boucle par l'endoribonucléase III de *Staphylococcus aureus*

<b>I. L'endoribonucléase III, une enzyme-clé dans la régulation par les ARN</b> .....	<b>76</b>
<i>I.A. La topologie des boucles est importante dans la reconnaissance par Rnt1p chez <i>Saccharomyces cerevisiae</i>.</i> .....	78
<i>I.B. La RNase III des bactéries coupe les hélices ARN sans spécificité apparente de séquence</i> .....	79
<i>I.C. La RNase III de <i>S. aureus</i> joue un rôle important dans la régulation par les ARN.</i> .....	80

### Etudes de la fonction et du mécanisme d'action d'ARN régulateurs chez *Staphylococcus aureus*

<b>II. RsaA, un ARN régulateur appartenant au régulon <math>\sigma^B</math></b> .....	<b>113</b>
<i>II.A. Identification des cibles de RsaA.</i> .....	115
<i>II.A.1. Comment identifier les cibles de RsaA ?</i> .....	117
II.A.1.a. Prédiction <i>in silico</i> d'ARNm cibles.....	117
II.A.1.b. Analyse comparative du protéome .....	120
<i>II.B. RsaA inhibe la traduction de l'ARNm mgrA.</i> .....	125
<i>II.B.1. L'ARNm mgrA est régulé au niveau post-transcriptionnel.</i> .....	126
<i>II.B.2. RsaA se fixe à l'ARNm mgrA à deux sites distants</i> .....	127
<i>II.B.3. RsaA inhibe la traduction de l'ARNm mgrA</i> .....	129
<i>II.B.4. Les deux sites interagissent avec l'ARNm mgrA de manière coopérative pour inhiber la traduction.</i> .....	131
<i>II.C. RsaA régule la formation de biofilm et de capsules in vivo</i> .....	134
<i>II.C.1. RsaA régule positivement la synthèse de biofilm.</i> .....	134
<i>II.C.2. RsaA régule négativement la synthèse des capsules</i> .....	136
<b>III. RsaE : un lien dans l'adaptation à l'anaérobiose conservé chez <i>B. subtilis</i> et <i>S. aureus</i> ?</b> <b>140</b>	
<i>III.A. Etude fonctionnelle de RsaE chez <i>B. subtilis</i></i> .....	141
<i>III.A.1. Construction d'une souche <i>B. subtilis</i> 168 délétée de <i>rsaE</i></i> .....	141
<i>III.A.1. RsaE est induit en réponse à un stress anaérobiose chez <i>B. subtilis</i></i> .....	142
<i>III.A.2. Perspectives pour ce projet.</i> .....	145

## Discussions & Perspectives

<b>I. Le réseau de régulation de RsaA est connecté au système de densité cellulaire</b> .....	<b>149</b>
<b>II. D'autres cibles régulées par RsaA ?</b> .....	<b>151</b>
<i>II.A. Les ARNm <i>yabJ-spoV</i>, <i>sa1956</i> et <i>glyS</i>.</i> .....	151
<i>II.B. Un motif riche en C et des topologies variées pour la liaison ARNm / sARN : une adaptation à l'absence d'Hfq?</i> .....	154
<i>II.C. Développements récents dans l'identification des cibles</i> .....	157
<i>II.D. RsaA interagit-il avec des protéines ?</i> .....	158
<i>II.D.1. La RNase III, co-facteur de RsaA ?</i> .....	159
<i>II.D.2. ARN régulateurs et traduction.</i> .....	160

II.D.3. *Comment identifier les protéines cibles ?*..... 162

III. **Dynamique des réseaux de régulation au sein de la population de *S. aureus***..... 163

IV. **La recherche de nouvelles molécules thérapeutiques** ..... 168

## **Annexes**

I. **Matériel & Méthodes**..... 173

II. **Publications scientifiques**..... 199

III. **Communications scientifiques**..... 213

IV. **Bibliographie** ..... 215

## Table des Figures & Tableaux

<b>Figure 1</b> : Plongée dans le trou bleu de "Dean" aux Bahamas. ....	3
<b>Figure 2</b> : Bactéries et archées extrêmophiles.....	4
<b>Figure 3</b> : Interférence transcriptionnelle chez les bactéries d'après Georg 2011.....	12
<b>Figure 4</b> : Dégradation des ARN chez <i>E. coli</i> et <i>B. subtilis</i> . ....	16
<b>Figure 5</b> : Simulations de réponses au stress d'après Shinomi <i>et al.</i> , 2007.....	20
<b>Figure 6</b> : Circuits de régulation de l'expression des gènes et interconnexion des réseaux... ..	22
<b>Figure 7</b> : Localisation des ARNm chez <i>E. coli</i> et régulation de l'ARNm <i>pstG</i> . ....	24
<b>Figure 8</b> : Résistance aux antibiotiques d'une population bactérienne d'après Lee <i>et al.</i> , 2010. .....	25
<b>Figure 9</b> : Mécanisme de régulation des <i>riboswitch</i> en tandem, d'après Breaker 2012. ....	28
<b>Figure 10</b> : Etude prédictive : fréquence d'apparition des différentes classes de <i>riboswitch</i> et proportion des mécanismes de régulation utilisés chez les bactéries (Barrick et Breaker 2007).....	29
<b>Figure 11</b> : Régulation de l'expression de la région 5'UTR de l'ARNm <i>alx</i> chez <i>E. coli</i> . (Nechooshtan <i>et al.</i> , 2009) de l'ARNm <i>glmS</i> chez <i>B. subtilis</i> (Collins <i>et al.</i> , 2007). ....	31
<b>Figure 12</b> : Dénaturation successives de l'élément ROSE, d'après Chowdhury <i>et al.</i> , 2006. ..	32
<b>Figure 13</b> : Chez <i>L. monocytogenes</i> , l'expression du facteur de transcription PrfA est contrôlée par la température et l'état métabolique. ....	33
<b>Figure 14</b> : La régulation par les ARN chez les bactéries. ....	34
<b>Figure 15</b> : Exemples de régulation par les ARNs.....	38
<b>Figure 16</b> : Répression de la traduction par les sARN. ....	40
<b>Figure 17</b> : Mécanisme de répression de la traduction par l'ARN, d'après Darfeuille <i>et al.</i> , 2007.....	41
<b>Figure 18</b> : Activation de la traduction et stabilisation des ARNm. ....	43
<b>Figure 19</b> : Régulation de l'activité de RsmA par les sARN RsmY et RsmZ d'après Bordi <i>et al.</i> , 2010.....	45
<b>Figure 20</b> : Structures 3D des protéines Hfq d' <i>E. coli</i> ( <i>Hfq<sub>ec</sub></i> ) et <i>S. aureus</i> ( <i>Hfq<sub>sa</sub></i> ) et modèle d'interaction sur la face latérale chez <i>S. typhimurium</i> .....	47
<b>Figure 21</b> : Recyclage dynamique des ARN liés à Hfq d'après Fender <i>et al.</i> , 2010.....	48
<b>Figure 22</b> : <i>Staphylococcus aureus</i> . ....	51
<b>Figure 23</b> : Evolution de la résistance aux antibiotiques de <i>S. aureus</i> .....	52
<b>Figure 24</b> : Prédictions des riboswitch et T-box dans les régions leaders des ARNm de <i>S.</i> <i>aureus</i> N315 (d'après Caldelari <i>et al.</i> , 2010).....	62
<b>Figure 25</b> : Le système de densité cellulaire <i>agr</i> .....	65
<b>Figure 26</b> : Système toxine-antitoxine SprA1/SprA1-AS.....	68
<b>Figure 27</b> : Réseaux de régulation médiés par les ARN régulateurs chez <i>S. aureus</i> : états des connaissances (D'après Felden <i>et al.</i> , 2011). ....	72
<b>Figure 28</b> : Généralités sur la famille des RNases III. ....	77
<b>Figure 29</b> : La topologie des tétra-boucles est la clé dans la reconnaissance des tiges-boucles par Rnt1p (Wang <i>et al.</i> , 2011). ....	78
<b>Figure 30</b> : Déterminants de séquence favorisant la liaison de la RNase III (d'après Pertzev <i>et</i> <i>al.</i> , 2006) et structure cristalline de la RNase III d' <i>A. aeolicus</i> (D'après Gan <i>et al.</i> , 2006 et 2008).....	80
<b>Figure 31</b> : Modèle de la régulation de l'ARNm <i>coa</i> par l'ARNIII. ....	81

<b>Figure 32</b> : La topologie des interactions boucle-boucle et des hélices guide la RNase III vers son site de coupure. ....	83
<b>Figure 33</b> : <i>RsaA</i> , nouveau ARN régulateur chez <i>S. aureus</i> . ....	114
<b>Figure 34</b> : Combinaison de stratégies <i>in silico</i> , <i>in vivo</i> et <i>in vitro</i> pour l'étude fonctionnelle des sARN. ....	116
<b>Figure 35</b> : Analyse comparative du protéome par la technique 2D-DIGE. ....	122
<b>Figure 36</b> : Prédiction des appariements <i>RsaA</i> / ARNm cible et contexte génomique. ....	125
<b>Figure 37</b> : <i>RsaA</i> régule l'expression de l'ARNm <i>mgrA</i> au niveau post-transcriptionnel <i>in vivo</i> . ....	127
<b>Figure 38</b> : <i>RsaA</i> se lie efficacement à l'ARNm <i>mgrA</i> en utilisant deux sites d'interaction distants. ....	128
<b>Figure 39</b> : La liaison de <i>RsaA</i> à l'ARNm <i>mgrA</i> inhibe la formation du complexe de pré-initiation de la traduction. ....	130
<b>Figure 40</b> : Les deux sites d'interaction agissent coopérativement dans la formation du duplexe et l'inhibition de la traduction. ....	133
<b>Figure 41</b> : Etude des effets de la délétion de <i>RsaA</i> sur la formation de biofilm. ....	136
<b>Figure 42</b> : <i>RsaA</i> régule la synthèse des capsules <i>in vivo</i> . ....	138
<b>Figure 43</b> : <i>RsaE</i> est conservé chez <i>S. aureus</i> et <i>B. subtilis</i> . ....	140
<b>Figure 44</b> : Construction de la souche de <i>B. subtilis</i> 168 $\Delta$ <i>rsaE</i> et complémentation. ....	141
<b>Figure 45</b> : Candidats pour une régulation par <i>RsaE</i> . ....	145
<b>Figure 46</b> : Modèle de régulation de l'ARNm <i>mgrA</i> par <i>RsaA</i> . ....	149
<b>Figure 47</b> : Réseau de régulation de <i>MgrA</i> et de <i>RsaA</i> avant et après activation de $\sigma^B$ . ....	150
<b>Figure 48</b> : L'ARNm <i>yabJ-spoVG</i> , une autre cible de <i>RsaA</i> ? ....	152
<b>Figure 49</b> : L'ARNm <i>glyS</i> , une cible potentielle de <i>RsaA</i> ? ....	154
<b>Figure 50</b> : Les motifs riches en C des ARN régulateurs de <i>S. aureus</i> sont retrouvés dans des régions accessibles pour permettre une interaction optimale avec la séquence Shine et Dalgarno des ARNm, d'après Geissmann <i>et al.</i> , 2009 (+ supplementary material). ....	156
<b>Figure 51</b> : Structure secondaire de <i>RsaA</i> et <i>RsaC</i> avec un RBS en vert et une succession de codons STOP en violet et orange. ....	161
<b>Figure 52</b> : Interconnectivité entre les réseaux de régulation de <i>RsaA</i> , <i>RsaE</i> et ARNIII. ....	167
<b>Tableau 1</b> : Facteurs de virulence chez <i>S. aureus</i> . ....	56
<b>Tableau 2</b> : Facteurs $\sigma$ codés par <i>S. aureus</i> . ....	57
<b>Tableau 3</b> : Facteurs de transcription chez <i>S. aureus</i> . ....	58
<b>Tableau 4</b> : Systèmes à deux composants caractérisés chez <i>S. aureus</i> . ....	60
<b>Tableau 5</b> : Liste des 10 meilleurs ARNm candidats formant potentiellement un duplexe avec <i>RsaA in silico</i> . ....	119
<b>Tableau 6</b> : Récapitulatif des protéines séquencées sur les gels 2D-DIGE. ....	124
<b>Tableau 7</b> : Résultats des analyses comparatives du protéome pour <i>RsaE</i> chez <i>B. subtilis</i> . ....	143



## Abréviations :

### Bases :

A	Adénine
U	Uracile
G	Guanine
C	Cytosine
T	Thymine
N	A, C, T, G
R	Purines (A ou G)
Y	Pyrimidines (C, T ou U)
W	A ou T

### ADN :

ADNc	ADN complémentaire
IGR	régions intergéniques

### ARN :

ARNm	ARN messenger
ARNnc	ARN non-codant
ARNr	ARN ribosomique
ARNt	ARN de transfert
sARN	small ARN
ARNas	ARN antisens
Rsa	RNA of <i>S. aureus</i>
RBS	site de fixation du ribosome
TSS	site d'initiation de la transcription
SD	Shine et Dalgarno
5' UTR	région 5' non traduite
3' UTR	région 3' non traduite
ROSE	<i>repressor of heat shock protein expression</i>

### Métabolites / ions :

Ca	calcium
Mg	magnesium
AdoCbl	adénosylcobalamine
FMN	flavine mononucléotides
GlcNP	glucosamine-6-phosphate
LPS	liposaccharide
preQ1	7-aminoéthyl 7-deazaguanine
SAM	S-adénosyl-méthionine
TCA	<i>tricarboxylic acid cycle</i>
TPP	thiamine pyrophosphate

### Pathogénie :

AsIP	ilots de pathogénie de <i>S. aureus</i>
CHIPS	<i>chemotaxis inhibitory protein of staphylococci</i>
CMI	concentration minimale inhibitrice
IL	Interleukine
LT	lymphocyte T
PVL	<i>Panton Valentin Leukocidin</i>
SAR	<i>staphylococcal accessory regulator</i>
SARM	<i>S. aureus</i> résistant à la méthicilline

Pour information, les noms des ARN régulateurs sont en violet pour mieux les distinguer des gènes cibles.

### Espèces de bactéries :

#### Gram positif :

<i>B. subtilis</i>	<i>Bacillus subtilis</i>
<i>L. monocytogenes</i>	<i>Listeria monocytogenes</i>
<i>L. innocua</i>	<i>Listeria innocua</i>
<i>S. aureus</i>	<i>Staphylococcus aureus</i>

#### Gram négatif :

<i>E. coli</i>	<i>Escherichia coli</i>
<i>H. pylori</i>	<i>Helicobacter pylori</i>
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
<i>S. typhimurium</i>	<i>Salmonella typhimurium</i>
<i>Aquifex aeolicus</i>	<i>A. aeolicus</i>

### Protéines :

DBD	domaine de liaison à l'ADN
GFP	<i>green fluorescent protein</i>
Hfq	<i>Host factor Q<math>\beta</math></i>
HK	histidine kinase
OMP	protéine de la membrane externe
Onr	oligoribonucléase
PAP I	polyA polymérase I
PNPase	polynucleotide phosphorylase
RBD	domaine de liaison à l'ARN
RNase	ribonucléase
RNAP	ARN polymérase
RR	<i>response regulator</i>
TCS	système à deux composants
TF	facteur de transcription
ORF	cadre ouvert de lecture
CDS	Région codante
C-terminale	carboxy terminale
N-terminale	amino terminale
aa	acides aminés

### Généralités :

kb	kilobase
nts	nucléotides
pb	paire base
°C	degrès Celsius
nM	nano molaire
$\mu$ M	micro molaire
pmol	pico moles
cpm	coups par minutes
DO	densité optique
nm	nano mètre
K <sub>D</sub>	constante de dissociation

### Milieu :

BHI	brain heart infusion
NZM	NZ amine medium

# Introduction



# I. L'incroyable pouvoir d'adaptation des procaryotes

## I.A. Adaptation à l'environnement

Les procaryotes (du latin "pro" signifiant *avant* et du grec "caryon" signifiant *noyau*) sont des organismes vivants unicellulaires dépourvus de noyau. Deux taxons distincts les composent, les bactéries et les archées. Ces fascinants micro-organismes sont des modèles d'étude uniques pour la compréhension des mécanismes du vivant et de l'évolution. L'étude des trous bleus de l'océan en est un parfait exemple ; ces grottes submergées à l'abri des marées n'ont pas été infiltrées par l'oxygène atmosphérique. Les bactéries qui subsistent dans ces lieux survivent grâce à l'exploitation du sulfate en lieu et place de l'oxygène, devenant des témoins privilégiés d'une époque où la vie s'est développée sans oxygène, il y a environ 4 milliards d'années (**Figure 1**).

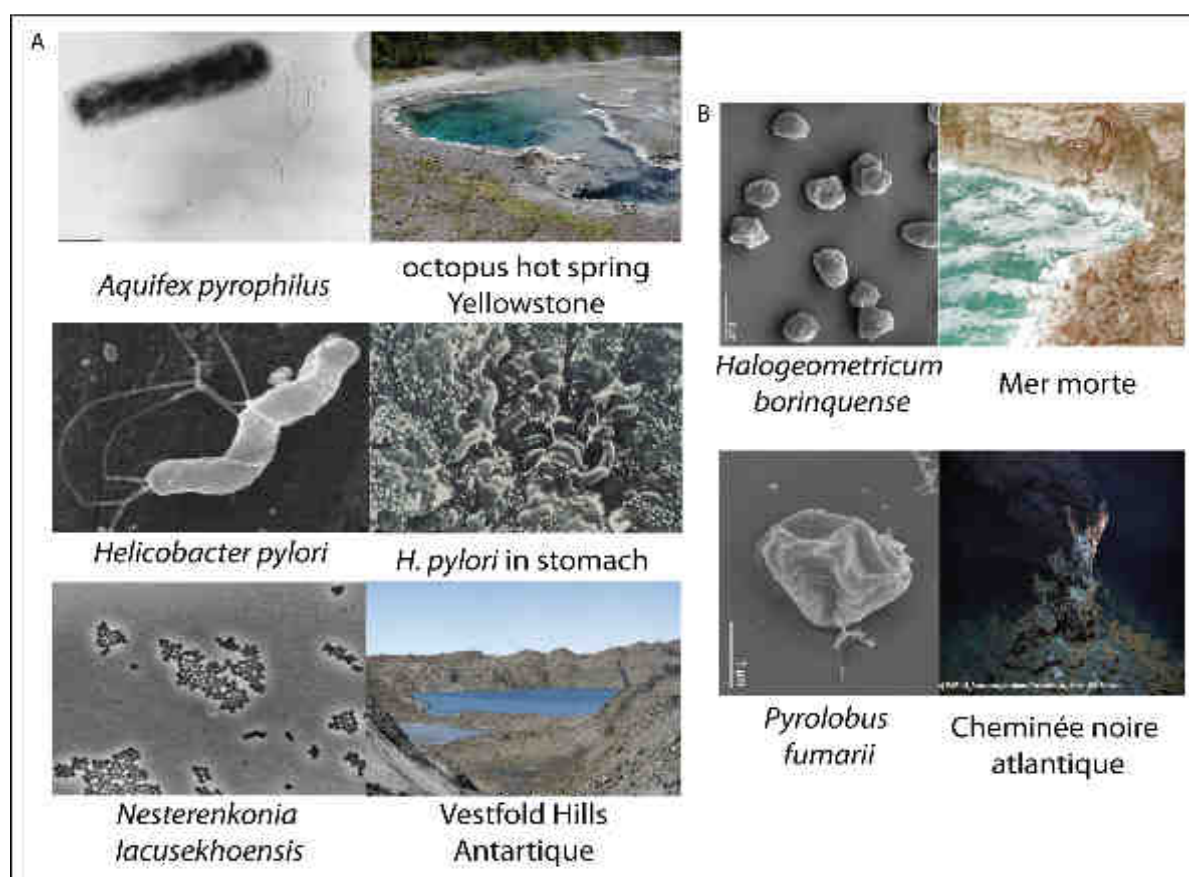


**Figure 1:** Plongée dans le trou bleu de "Dean" aux Bahamas.

Photos de Wes C. Skiles, National geographic. (<http://ngm.nationalgeographic.com/2010/08/bahamas-caves/todhunter-text>)

Les procaryotes ont colonisé toutes les niches écologiques de la planète, même les plus extrêmes, montrant ainsi leurs formidables capacités d'adaptation. Des bactéries et archées extrémophiles ont ainsi été retrouvées près des cheminées noires au fond des océans, comme *Pyrolobus fumarii* dont la souche 121 survit à une température de 113°C. *Deinococcus radiodurans* supporte des doses de radioactivité 1500 fois supérieures à ce que l'homme peut endurer. Des espèces de *Ferroplasma* se sont adaptées à la vie dans des zones

d'eau acide au pH proche de 0, alors que *Bacillus firmus* OS4 survit au sein de zones alcaline de pH 9 (Rothschild *et al.*, 2001) (Figure 2). La majorité des bactéries ne sont pas cultivables en laboratoire, soulignant que la classification actuelle de ces organismes n'est que la partie émergée de l'iceberg. Le jeu des statistiques et des estimations est fascinant avec ces êtres vivants. Ainsi, la masse de carbone correspondant à la biomasse bactérienne équivaldrait à la masse de carbone présente dans l'ensemble des plantes terrestres et aquatiques (Pedersen 1999).



**Figure 2 : Bactéries et archées extrêmophiles.**

**A.** Exemples de bactéries extrêmophiles. *Aquifex pyrophilus*, bactérie Gram négative hyperthermophile, survit dans les sources chaudes comme celles trouvées dans le parc de Yellowstone (température de 83°C). Photo *A. pyrophilus* (Behammer *et al.*, 1995) Octopus Hot Spring (Lunar and planetary institut [http://www.lpi.usra.edu/education/EPO/yellowstone2002/workshop/octopus1/octopus1\\_imgs/p7280452\\_lg.jpg](http://www.lpi.usra.edu/education/EPO/yellowstone2002/workshop/octopus1/octopus1_imgs/p7280452_lg.jpg)). *Helicobacter pylori* est une bactérie Gram négative, acidophile et pathogène de l'homme, capable de résister au pH acide de l'estomac. (gauche : O'Rourke *et al.*, 2001, droite Martin J. Blaser, 1989, New York city university). *Nesterenkonia lacusekhoensis*, halophile Gram positif, a été isolée dans le lac salin de Ekho dans l'est de l'antarctique. Ce lac salin ne gèle pas. (gauche : Collins *et al.*, 2002, droite : Doug McVeigh <http://www.antarctica.gov.au/living-and-working/stations/davis/this-week-at-davis/2011/13-may-2011/1#>). **B.** Exemples d'archées extrêmophiles. *Halogeometricum borinquense*, halophile, isolée dans la Mer Morte (275 g/L de sel) (photo Malfatti *et al.*, 2009, Mer Morte : M Disdero, licence creative commons, Wikipedia). *Pyrolobus fumarii*, hyperthermophile, capable de se reproduire lorsque la température 90°C à 113°C. (photo *P. fumarii* Göker 2011, photo cheminée noire marum (Centre des Sciences Environnementales Marines, université de Brême).

## I.B. Des liens étroits avec l'homme

La majorité des espèces de bactéries et archées introduites jusqu'ici représente des espèces saprophytes. Retrouvées dans l'environnement (eau, air et sol), elles survivent indépendamment de la présence d'autres êtres vivants. Les bactéries dites détritivores colonisent aussi l'environnement, mais ingèrent de la matière organique provenant d'êtres vivants ayant malheureusement trépassés. Enfin, de nombreuses relations se sont établies au cours de l'évolution entre les procaryotes et l'être humain.

Chez l'homme, le nombre de bactéries présentes est estimé à  $10^{12}$  sur la peau ( $10^7$  cellules par  $\text{cm}^2$  de peau ; Fredricks *et al.*, 2001),  $10^{10}$  dans la bouche et  $10^{14}$  au niveau du système gastro-intestinal, soit au total 10 fois plus que le nombre moyen de cellules eucaryotes constituant un être humain (Nicholson *et al.*, 2005).

Grâce aux nouvelles techniques de séquençage à haut débit, il est désormais possible de caractériser la flore bactérienne existante chez l'être humain. Par exemple, l'étude de Fierer *et al.*, a caractérisé le microbiome présent sur les mains de 51 étudiants en master (volontaires). Les résultats montrent qu'en moyenne, 150 espèces différentes de bactéries ont été recensées au niveau des mains. Plus intéressant encore, seules 17% des espèces sont communes entre les mains des individus testés, et seules 13% de l'ensemble des bactéries décrites (soit 4700 dans cette étude) sont retrouvées chez tous les étudiants (Fierer *et al.*, 2008). Des microbiomes au moins aussi complexes sont retrouvés dans l'estomac et la flore buccale (Cogen *et al.*, 2008).

Les relations homme/bactéries peuvent être mutualistes, c'est à dire bénéfique à la fois pour l'hôte et la bactérie, ou commensales. A l'inverse des précédentes, la relation commensale n'a pas d'effets positifs directs pour l'hôte, mais cet aspect est sujet à discussion aujourd'hui. La flore bactérienne du système digestif en est une parfaite illustration. Le système digestif est stéril à la naissance, mais à l'âge adulte, il est colonisé en moyenne par 1 kg de bactéries (Chow *et al.*, 2011). Les genres *Firmicutes* et *Bacteroidetes* représentent 98% des bactéries du système digestif (Ley *et al.*, 2006). Dans le gros intestin, un exemple de relation mutualiste est la bactérie Gram négative *Bacteroides cellulosilyticus*. Cette dernière est capable de digérer les fragments insolubles de cellulose provenant des

cellules végétales. Les fragments générés sont ensuite dégradés par d'autres bactéries comme *Bacteroides thetaiotaomicron* en mono/di saccharides assimilable à la fois par les bactéries (cross-feeding) et par les cellules humaines (Flint 2008).

Les bactéries commensales et mutualistes de la flore intestinale assistent le système immunitaire en renforçant la barrière épithéliale (inhibition de la pénétration de bactéries pathogènes) et en stimulant la production d'immunoglobulines ou d'agents antimicrobiens (Chung *et al.*, 2010). Les membres du genre *Bifidobacterium* synthétisent des molécules d'acétate qui stimulent les défenses immunitaires des cellules épithéliales de l'estomac et inhibent le passage dans la circulation sanguine de la toxine Shiga de la souche *Escherichia coli* 0157:H7 (*E. coli*) (Fukuda *et al.*, 2010). Chez la souris, le pathogène *Citrobacter rodentium* infecte de manière persistante des souris sans microbiome intestinal. Dans des souris normales, le pathogène est en compétition avec d'autres bactéries commensales pour l'utilisation des monosaccharides ou d'autres sources de nutriments, empêchant ainsi la colonisation de l'intestin (Kamada *et al.*, 2012).

Ainsi, la présence des bactéries commensales rend plus difficile la colonisation pour des bactéries pathogènes. Cet effet étant bénéfique pour l'hôte, la notion de commensale est-elle toujours d'actualité ? De plus, des animaux sans microbiome présentent des défauts de vascularisation, d'enzymes de digestion, de lymphocytes soulignant les relations structurales, métaboliques et immunologiques fortes qui se sont établies au cours de l'évolution entre l'homme et les bactéries. Aujourd'hui, ces relations se rapprochent d'avantage de la symbiose que du mutualisme (Brenchley *et al.*, 2012).

Les bactéries pathogènes ont la capacité d'envahir les tissus de l'hôte pour établir un foyer d'infection. Le pathogène doit tout d'abord s'adhérer aux tissus *via* l'expression d'adhésines. Puis l'expression de facteurs de virulence neutralise les défenses immunitaires, et la synthèse de toxines détruit des cellules de l'hôte pour permettre l'accès à de nouvelles sources de nutriments et favoriser la dissémination des bactéries dans les tissus voisins. Ces infections peuvent être temporaires, chroniques, voire mortelles et pourraient augmenter les risques de développer certains types de cancers du tube digestif et du côlon (Grivennikov *et al.*, 2010). Deux types de bactéries pathogènes se distinguent : les pathogènes obligatoires (*Shigella*) et les pathogènes opportunistes. Ces derniers se comportent la plupart du temps comme des bactéries commensales. Mais, lorsque les défenses de l'hôte

sont affaiblies (i.e. immunodépression) ou lors d'une carence en nutriments, les pathogènes opportunistes expriment leurs gènes de virulence pour induire une infection. Dans la dernière partie de cette introduction, je décrirai en détail les facteurs de virulence d'un pathogène opportuniste de l'homme, *Staphylococcus aureus* (*S. aureus*).

Les enjeux de la microbiologie moderne méritent toute notre attention. En terme de santé publique, l'émergence de souches multi-résistantes aux antibiotiques est inquiétante. L'étude de la régulation de l'expression des gènes de virulence et des mécanismes d'acquisition de gènes de résistance aux antibiotiques ouvrira de nouvelles pistes pour lutter contre ces pathogènes.

Le pouvoir d'adaptation des bactéries peut être habilement détourné et utilisé dans des applications environnementales. Par exemple, des équipes de chercheurs ont généré une souche dérivée d'*E. coli* capable de produire des ester gras à partir de sucres, ouvrant la voie à la production de fuel par les microbes (Steen *et al.*, 2010). Les bactéries sont également une piste pour la décontamination des sols pollués aux métaux lourds (Singh *et al.*, 2003).

**Ce pouvoir d'adaptation remarquable des bactéries passe par une régulation rapide et précise de l'expression des gènes bactériens appropriés selon les différents contextes environnementaux. Quels sont les mécanismes qui coordonnent l'expression des gènes ? Quels sont les mécanismes de régulation mis en place par la bactérie pour adapter l'expression de son génome en fonctions des situations rencontrées ?**



## II. Nouvelles classes d'ARN chez les bactéries

De manière intéressante, le nombre total de gènes codant des protéines n'est pas corrélé à la taille du génome. Par exemple, le génome humain est 19 fois plus grand que celui de *Drosophila melanogaster*, mais ne code qu' 1,6 fois plus de protéines. En parallèle des ARN non-codant (ARNnc) ménagers tels que les ARNr et les ARNt, existe-t-il d'autres types d'ARN exprimés dans le génome, et si oui dans quelle proportion?

### II.A. Découverte des sARN ("*small RNA*") exprimés dans les régions intergéniques et des ARN antisens, codés en *cis*

Paradoxalement, de tels ARN ont été découverts dès les années 1980. Des molécules exerçant leurs fonctions biologiques sous forme d'ARN ont été mises en évidence dès les années 1980. Le premier ARN régulateur chromosomique a été identifié en 1984. Dans cette étude, Mizuno *et al.*, ont caractérisé un fragment d'ADN qui, exprimé en *trans* à partir d'un plasmide, réprime l'expression de l'ARNm *ompF*. Ce fragment d'ADN contenait une unité de transcription codant pour l'ARN régulateur *MicF*, capable de se lier au site de fixation du ribosome (RBS ou *ribosome binding site*) de l'ARNm, composé de la séquence Shine et Dalgarno (SD) et du codon d'initiation de la traduction AUG, pour inhiber sa traduction (Mizuno *et al.*, 1984). En parallèle, les premiers ARN antisens (ARNas), codés en *cis*, ont été découverts dans les plasmides, phages et transposons (revue Wagner 2002 ; Brantl 2007).

**Découverte des sARN ("*small RNA*") dans les régions intergéniques.** La vision de la communauté scientifique a profondément évolué dans un premier temps depuis 2001 grâce à l'identification d'une nouvelle classe d'ARN exprimés dans les régions intergéniques (*intergenic region* ou IGR) de la bactérie *E. coli*. Aujourd'hui, ces ARN désignés comme "*trans-encoded small RNA (sARN)*", régulent la fonction de cibles codées à un autre locus du génome. La découverte des sARN connaît un essor fulgurant depuis 2001 grâce à l'amélioration des techniques prédictives, combinée à des tests d'expression *in vivo*

(Wassarman *et al.*, 2001 ; Argaman *et al.*, 2001 ; Rivas *et al.*, 2001). L'idée de départ de ces études était que des ARN autre que les ARN non-codants (ARNnc) dit ménagers (*housekeeping*) devaient être exprimés à partir des IGR conservées des génomes. Les critères utilisés étaient : **(1)** le ciblage des IGR d'une taille allant de 50 à 300 nucléotides (nts), **(2)** la prédiction d'éléments structuraux (terminateur Rho-indépendant) et de promoteurs orphelins, et **(3)** la conservation des IGR dans des génomes de bactéries phylogénétiquement proches. Ces études ont abouti à la validation de l'expression de nouveaux ARNnc chez *E. coli*. Depuis, l'existence de plus de 80 sARN a été validée (Waters et al 2009) chez cette bactérie. Ces ARN sont très hétérogènes en taille (de 50 à 500 nts), en structure, et sont retrouvés dans tous les génomes proches d'*E. coli*. Leurs fonctions et mécanismes de régulation seront détaillés dans la partie III.B.2.

**Découverte des ARN antisens (ARNas).** L'importance de la régulation par les ARNas a été mise à jour grâce à l'essor des techniques de séquençage à haut débit (Sharma *et al.*, 2010) et des puces à ADN couvrant les deux brins du génome (Georg *et al.*, 2009, Güell *et al.*, 2009, Toledo-Arana *et al.*, 2009). Ces avancées technologiques ont mis en évidence que les deux brins du génome codent pour des ARN, révélant l'existence d'une transcription d'ARN antisens (ARNas) chez les bactéries. En parallèle, ces techniques ont également facilité l'identification de nouveaux sARN. Par exemple, chez *H. pylori*, une carte des sites d'initiation de la transcription au sein du génome (TSS ou *transcription start site*) a été obtenue en comparant les ADN complémentaires (ADNc) obtenus à partir d'un extrait total d'ARN à ceux obtenus après un traitement à l'exonucléase "terminator<sup>TM</sup>". L'action de l'enzyme est inhibée par la présence d'un 5' triphosphate permettant un enrichissement en transcrits primaires. L'organisation des TSS peut être classée en plusieurs catégories : (i) TSS présent à moins de 500 pb d'une ORF, (ii) plusieurs TSS pour une ORF (exprimé différemment), (iii) TSS à l'intérieur d'une ORF, (iv) TSS situé dans ou à l'extrémité d'un gène, mais sur le brin antisens (ARNas), (v) Tss orphelin situé dans des IGR et codant donc pour des sARN (Sharma *et al.*, 2010). Depuis, la transcription d'ARNas a été retrouvée dans plusieurs génomes de bactéries (revue Georg *et al.*, 2011 ; Gottesman *et al.*, 2011), et des études récentes menées chez *S. aureus* (Lasa *et al.*, 2011; Lioliou *et al.*, 2012) montrent qu'une transcription antisens est retrouvée sur presque l'ensemble des deux brins d'ADN.

Ces ARNs sont-ils seulement dus à une transcription permanente, ou sont-ils induits spécifiquement par des facteurs de transcription ? Leurs fonctions et mécanismes de régulation seront détaillés dans la partie III.B.2.

## II.B. Nouveaux mécanismes de transcription des gènes

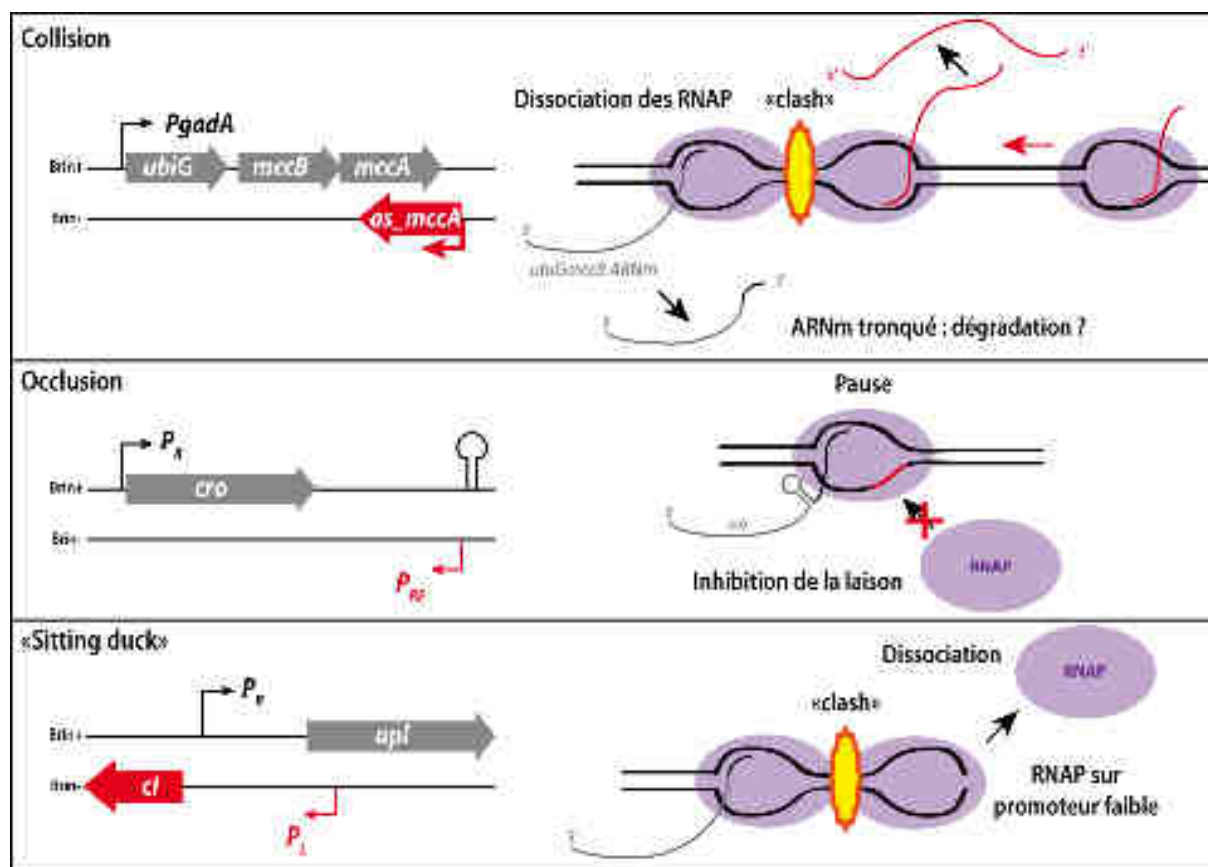
**Les nanoARN comme initiateurs de la transcription.** Cette voie alternative d'initiation de la transcription a été mise en évidence chez la bactérie Gram négative *Pseudomonas aeruginosa* (*P. aeruginosa*). S. R. Goldman *et al.*, (2011) ont montré que l'ARN polymérase (RNAP ou *RNA polymerase*) de la bactérie était capable d'initier la transcription à partir de nanoARN de 2 à 4 nts. Dans un contexte normal, ces nanoARN sont dégradés par l'enzyme Onr (oligoribonucléase) qui joue un rôle crucial dans la dégradation des ARN (II.D) chez les bactéries Gram négatives. La délétion de l'enzyme chez *P. aeruginosa* entraîne l'accumulation de nanoARN qui servent alors d'amorce à la RNAP pour initier la transcription des gènes. Ainsi, le site TSS de certains ARNm varie de 2 à 3 nts *in vivo* et la présence de nanoARN induit une altération globale de l'expression des gènes. L'expression de 586 gènes augmente tandis que celle de 572 gènes diminue. Plusieurs questions se posent : Comment ces nanoARN sont-ils générés en présence d'Onr ? Quels sont les mécanismes expliquant un tel impact sur l'expression des gènes ? Une première explication s'appuie sur la particularité que les nanoARN ne possèdent pas de groupement 5' triphosphate. Les ARNm générés seraient donc plus sensibles à la dégradation (II.D). L'ajout de nucléotides supplémentaires pourrait également modifier la structure des ARN et donc influencer la traduction ou la stabilité. Enfin, une dernière hypothèse est que dans certains cas, l'utilisation des nanoARN comme amorce influencerait l'efficacité d'initiation de la transcription, et donc le niveau des ARN transcrits (Goldman, Sharp *et al.*, 2011). Depuis, la production de nanoARN a été confirmée chez *E. coli* dans sa forme sauvage exprimant l'Onr. Leur production est dépendante de la phase de croissance. La majorité de ces ARN est détectée en phase stationnaire de croissance, porte une extrémité 5'OH, et affecte de manière notable les promoteurs de gènes possédant aux positions -1/+1 un T/A. La présence d'une extrémité 5'OH a-t-elle un lien avec la dégradation médiée par la RNaseE, qui préfère les extrémités 5' monophosphate (voir II.D)? (Vvedenskaya *et al.*, 2012)

**L'interférence transcriptionnelle, nouveau mécanisme de régulation de la transcription.** L'interférence transcriptionnelle a connu son essor depuis la découverte des ARNas. Trois types d'interférence sont décrites : la collision, l'occlusion, le "sitting-duck" (revue Georg 2011) (Figure 3).

La collision entre deux RNAP en phase d'élongation résulte en un arrêt de la transcription pour les deux gènes. De manière surprenante, l'arrêt ne semble pas être dû à une interaction stérique, mais se ferait par l'intermédiaire d'interactions longues distantes ou la présence d'ADN super-enroulé devant les RNAP. Une fois les deux RNAP proches, plusieurs cas de figures sont possibles : dissociation de l'un ou l'autre des complexes, «back-tracking», ou blocage. Chez *Clostridium acetobutylicum*, la collision entre deux RNAP est nécessaire à la synthèse d'un ARNas de 1 kb *in vivo*, *As\_mccA*, qui répond à la concentration intracellulaire du métabolite SAM (S-adenosyl méthionine) (André *et al.*, 2008).

L'occlusion de promoteurs a été décrite pour les promoteurs  $P_R$  et  $P_{RE}$  du phage  $\lambda$ . Lorsque la RNAP transcrit à partir de  $P_R$ , cette dernière stoppe au niveau du terminateur  $T_R$ , empêchant l'initiation de la transcription à partir de  $P_{RE}$  situé au voisinage proche du terminateur sur le second brin d'ADN (Georg 2011). Ce système est important dans la régulation du choix entre le cycle lytique et lysogénique du phage.

Le dernier type d'interférence utilise une RNAP élongatrice qui empêche une autre RNAP de se fixer sur un promoteur faible situé sur le brin antisens pour passer à l'étape d'élongation de la transcription. Ce mécanisme est nécessaire à l'expression des gènes lytiques chez le bactériophage 186.



**Figure 3 : Interférence transcriptionnelle chez les bactéries d'après Georg 2011.**

La collision se produit quand deux RNAP en phase d'élongation se rencontrent, aboutissant à un blocage, un relargage ou un "backtracking". L'occlusion se base sur le principe qu'une RNAP bloquée à son site de terminaison empêche l'initiation de la transcription des promoteurs au voisinage proche. Enfin le "sitting-duck" est une situation dans laquelle une RNAP en phase d'élongation empêche la transition initiation/élongation d'une autre RNAP sur un promoteur faible.

### II.C. Quels sont les rôles de ces ARN?

Les ARNs et sARN sont retrouvés dans plusieurs génomes de bactéries (revue Thomason *et al.*, 2010, Georg *et al.*, 2011, Storz *et al.*, 2011). Comme nous le verrons dans la partie III.A, la régulation par les ARN possède des propriétés distinctes à celles des facteurs de transcription, et ces derniers sont particulièrement adaptés à une réponse rapide à l'environnement ou la réponse immunitaire d'un hôte. Le rôle central des sARN régulateurs dans la régulation des gènes de virulence est parfaitement illustré chez *S. aureus* par l'ARNIII, effecteur intracellulaire du système de densité cellulaire qui contrôle la balance entre l'expression des exotoxines et des facteurs d'adhésions (Cf. IV). Un autre exemple est la myriade d'ARN régulateurs induits lors d'un stress osmotique pour réguler l'expression des protéines de la membrane d'*E. coli* et *S. typhimurium* (revue Vogel *et al.*, 2006).

D'un point de vue évolutif, l'analyse et la comparaison des répertoires d'ARN régulateurs pourraient nous aider à comprendre les différences entre des souches pathogènes ou non, ainsi que les mécanismes mis en place pour contrôler l'expression des gènes de virulence dans le cas d'infections persistantes. Par exemple, une attention particulière portée sur le génome non-codant de *Listeria monocytogenes* (*L. monocytogenes*, pathogène) et *Listeria innocua* (*L. innocua*, non-pathogène) montre que certains sARN et ARNas sont spécifiques de la bactérie virulente et que d'autres sont conservés dans chaque espèce, mais avec des profils d'expressions différents pouvant souligner la perte ou le gain de fonction. (Wurtzel *et al.*, 2012).

## II.D. Dégradation des ARN

Je souhaiterais brièvement décrire les mécanismes de dégradation des ARN chez les bactéries Gram positives et négatives. La dégradation des ARN est capitale dans la régulation de l'expression des gènes en particulier dans les phénomènes adaptatifs qui requièrent une réponse rapide. De plus, l'interaction ARN régulateur/ARN cible aboutit souvent à une dégradation de l'un (médiée par la RNase E, III.B.2.b), voire des deux partenaires (médiée par la RNase III, III.B.2 et IV.D.2)

**La dégradation des ARN chez les bactéries Gram négatives fait intervenir un clivage endonucléolytique suivi d'une dégradation exonucléolytique 3' vers 5' (Figure 4).** La dégradation des ARNm est particulièrement documentée chez le modèle *E. coli*. De manière générale, celle-ci passe par un clivage endonucléolytique par l'endoribonucléase E (RNase E), suivi d'une dégradation exonucléolytique de 3' vers 5' (Condon *et al.*, 2007). L'enzyme clé est la RNase E, qui est spécifique des régions en simple-brin avec une préférence pour les séquences riches en A/U (Arraiano *et al.*, 2010). Cette protéine agit au sein d'un complexe multiprotéique appelé dégradosome (Carpoussis 2007). C'est l'extrémité C-terminale de la RNase E qui sert de base au recrutement de l'hélicase RhlB, de la PNPase (Polynucleotide phosphorylase) et de l'énolase. L'activité de la RNase E est sensible à l'état de phosphorylation de l'extrémité 5' de l'ARNm : elle est plus active en présence d'une extrémité 5' monophosphate que triphosphate. Il a été montré récemment qu'une étape clé dans la dégradation des ARNm chez *E. coli* passe par la réduction de l'extrémité 5'

triphosphate en monophosphate par la pyrophosphatase RppH (Deana *et al.*, 2008). Après clivage par la RNase E, les extrémités simple-brin générées sont ciblées par la PNPase qui est une phosphorylase 3'-5'. Celle-ci est bloquée par des structures secondaires, comme par exemple les terminateurs de transcription Rho-indépendant (Mackie *et al.*, 2000). Pour passer outre ces structures, la PNPase forme des complexes avec d'autres protéines, dont l'hélicase RhlB au sein du dégradosome. Une autre stratégie consiste à recruter la PolyA-polymérase I (PAP I). L'addition de 7-10 adénines sur l'ARNm est suffisante pour recruter la PNPase couplée à l'hélicase RhlB (Chen *et al.*, 2005). Deux autres exoribonucléases 3'-5' sont impliquées dans la dégradation des ARNm : la RNase II et R. La première est complètement inhibée par des structures secondaires (Mackie *et al.*, 2000). La RNase R possède un domaine hélicase qui lui permet de défaire des structures secondaires (Awano *et al.*, 2010). Cette propriété pourrait expliquer sa conservation au détriment de la RNase II chez les bactéries psychotropes. Elle est globalement conservée dans l'ensemble des bactéries au cours de l'évolution (Kaberdin *et al.*, 2011). De plus, l'accumulation et la stabilité de cette enzyme varient en réponse à de multiples stress (Chen *et al.*, 2005). Les fragments générés par les exoribonucléases varient de 2 à 5 nts et sont pris en charge l'oligoribonucléase Onr pour générer des mononucléotides. Onr est enzyme essentielle chez *E. coli* (Ghosh *et al.*, 1999).

**Rôle de l'endoribonucléase Y et des exoribonucléases 5'-3' chez les bactéries Gram positives (Figure 4).** Chez les bactéries Gram positives, la RNase E n'est pas conservée (Kaberdin *et al.*, 2011 ; Jester *et al.*, 2012). La dégradation des ARNm fait intervenir deux enzymes clés qui agissent en coopération : l'endoribonucléase Y (RNase Y) et le complexe RNase J1/J2 (Condon *et al.*, 2007). La RNase Y fait partie d'un complexe similaire au dégradosome chez *Bacillus subtilis* (*B. subtilis*) (Commichau *et al.*, 2009) et *S. aureus* (Roux *et al.*, 2011). Ce dernier comprend deux protéines glycolytiques, l'énolase et la phosphofructokinase, ainsi que l'hélicase CshA, les RNases J1/J2 et la PNPase. De manière intéressante, la localisation membranaire de la RNase Y de *B. subtilis* est indispensable à sa fonction *in vivo* (Lahnik-Habrink *et al.*, 2011). Les extrémités des ARNm clivés par la RNase Y sont ciblés par deux types d'exoribonucléases : de 3' vers 5' via la PNPase (Deikus *et al.*, 2004), et de 5' vers 3' via les RNases J1 et J2. La RNase J1 possède en plus une activité endoribonucléase. Le rôle majeur de J1 et J2 dans la dégradation des extrémités 5' explique

pourquoi la présence d'éléments structuraux à cette extrémité protège efficacement les ARNm contre la dégradation chez *B. subtilis* (Sharp *et al.*, 2005). La RNase J1 est essentielle chez *B. subtilis* (Mathy *et al.*, 2007), de même que la RNase Y et la RNase III (Durand *et al.*, 2012). En revanche, chez *S. aureus*, la RNase Y et la RNase III sont requises pour la virulence, mais ne sont pas essentielles pour la croissance bactérienne. Associé au dégradosome chez *B. subtilis* et *S. aureus*, se trouve la RpnA, co-facteur de la RNase P. Cette protéine est essentielle chez *S. aureus* et dégrade les ARNm *in vitro*. Un inhibiteur de RpnA a un effet global sur la stabilité des ARNm et sur la croissance bactérienne (Olson *et al.*, 2010).

**Des similitudes entre bactéries Gram positives et négatives.** En plus de la RNase R et la PNPase, une enzyme commune et importante dans la dégradation des ARNm est la RNase III, enzyme ubiquitaire retrouvée dans l'ensemble du vivant (MacRae *et al.*, 2007). Cette endoribonucléase spécifique des régions double-brin (Gan *et al.*, 2006) mature les ARNr et ARNt (Condon *et al.*, 2002), dégrade les duplexes sARN/ARNm (III.B.2) et joue un rôle dans le contrôle qualité des ARN en dégradant les long ARN antisens issus de la transcription omniprésente du génome et des régions UTR des ARNm divergents qui sont complémentaires chez *S. aureus* (Lioliou *et al.*, 2012) (III.B.2.a).

Enfin, la présence d'enzymes glycolytiques au sein des dégradosomes est intrigante. Senseurs de l'état métabolique de la bactérie (cycle TCA et glycolytique), il est suggéré qu'elles pourraient coordonner l'action des RNases en réponse au stress (Carpoussis *et al.*, 2007).

**Les bactéries régulent l'expression des gènes en utilisant des protéines et des ARN régulateurs. Quelles sont les caractéristiques qui distinguent ces régulateurs ? Existe-t-il des connections entre leurs réseaux de contrôle? Quels sont les mécanismes de régulation et quel est le poids des ARN régulateurs dans le pouvoir d'adaptation des bactéries ?**



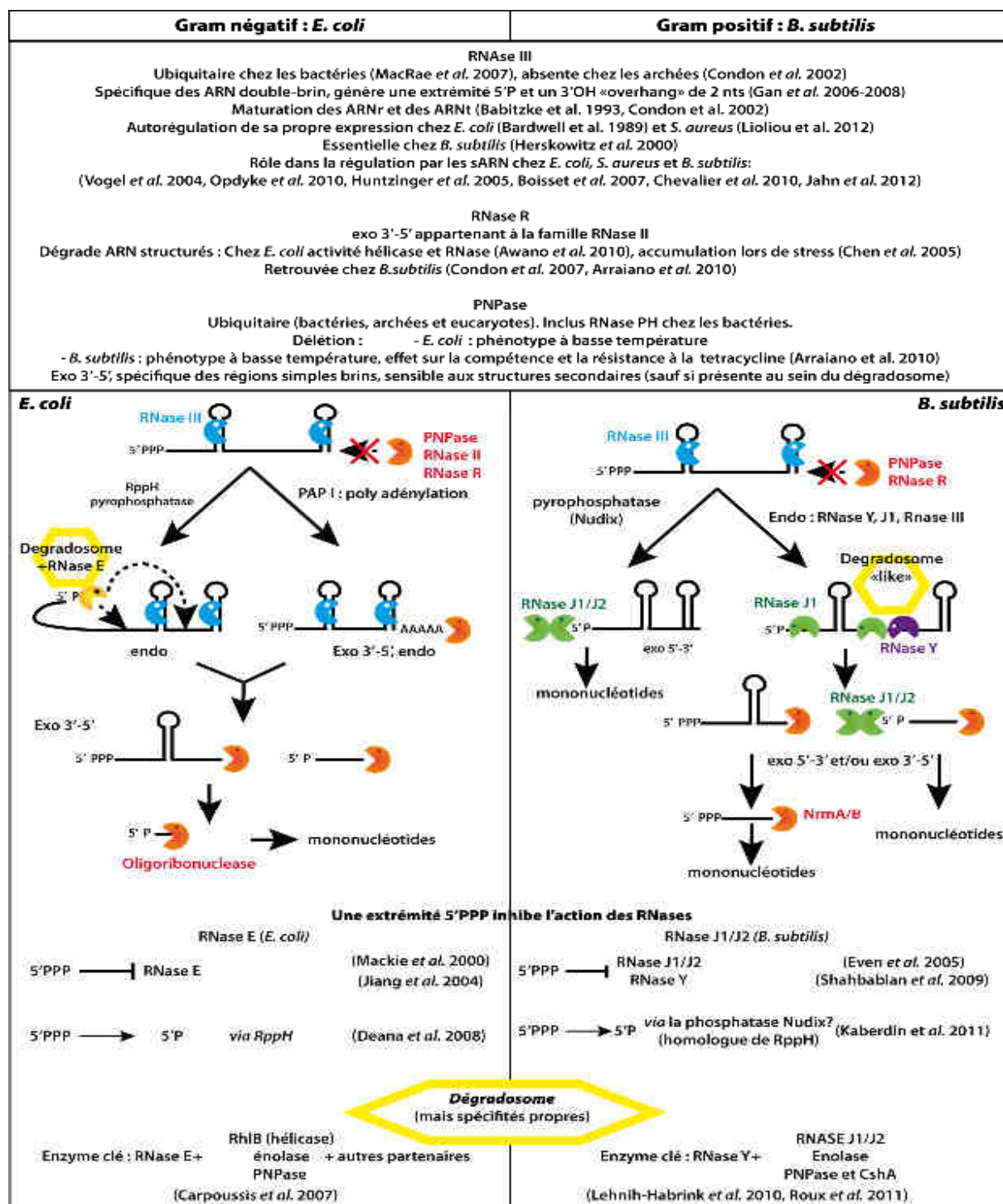


Figure 4 : Dégradation des ARN chez *E. coli* et *B. subtilis*.

Trois RNases sont retrouvées en commun chez ces deux bactéries. La RNase III, endoribonucléase spécifique des ARN double-brin, la RNase R et la PNPase, exoribonucléase 3' vers 5' spécifique des régions simple brin. Chez *E. coli*, la dégradation des ARNm passe par un clivage par la RNase E qui génère des fragments simple-brin qui sont dégradés par les exoribonucléases 3'-5' (PNPase, RNase II et R). L'action de la pyrophosphatase RppH facilite l'action de la RNase E. La polyadénylation des ARN par PAP I déstabilise les ARN. Les fragments de 2-4 nts générés sont recyclés par Onr qui libère des mononucléotides. Chez *B. subtilis*, le clivage généré par la RNase Y (endoribonucléase) génère des substrats aux RNases J1 et J2 ayant une activité exoribonucléase 5'-3'. Les enzymes Nrm A et B ont le même rôle que Onr. Un dégradosome est présent chez les deux bactéries, mais les protéines qui le composent varient.

### III. La régulation de l'expression des gènes par l'ARN

La régulation par les ARN présente des propriétés uniques en terme de réponse et de vitesse d'adaptation, par rapport aux facteurs de transcription (TF ou *transcription factor*). La bactérie utilise avantageusement les deux types de régulation au sein de réseaux complexes pour détecter et s'adapter à de brusques changements environnementaux.

#### III.A. Avantages de la régulation *via* l'ARN

##### III.A.1. Une réponse rapide aux conditions de stress

Des études de simulations de réponse au stress ont apporté une vision quantitative sur la régulation par les sARN (Shimoni *et al.*, 2007, revue Beisel *et al.*, 2010). Trois cas distincts de régulation ont été comparés : *via* un répresseur de la transcription se liant à l'ADN, *via* un sARN répresseur de la traduction d'un ARNm cible et enfin *via* des interactions protéine-protéine au niveau post-traductionnelle. L'attention portée sur les sARN répresseurs se base sur le fait que la majorité des ARN régulateurs identifiés chez les entérobactéries, les bactéries les plus étudiées étant *E. coli* et *Salmonella typhimurium* (*S. typhimurium*), inhibe la traduction d'un ARNm cible et entraîne sa dégradation (voir III.B).

**La régulation par les sARN est idéale pour répondre rapidement au stress.** Par souci de clarté, je discuterai uniquement le scénario suivant : la cible à réguler est exprimée avant la réponse au stress, les régulateurs non. Ce scénario implique que les régulateurs doivent d'abord être synthétisés *de novo* en réponse au stress. Plusieurs facteurs sont à prendre en compte. Pour les sARN, les paramètres sont la vitesse de transcription du sARN et de l'ARNm cible, leurs stabilités, leurs vitesses d'association, le nombre d'ARNm ciblés et pour finir, la vitesse de dégradation des protéines cibles présentes avant le stimulus. Pour les deux autres types de régulation, il s'agit de la vitesse de transcription des deux ARNm, leurs stabilités, leurs vitesses de traduction, l'affinité des partenaires, et le nombre de sites à bloquer sur l'ADN (dans le cas des TF) ou le nombre de protéines à lier (dans le cas des interactions protéine-protéine). Une différence fondamentale distingue ces deux types de régulation : le

taux de dégradation des protéines cibles présentes avant le stimulus. Ce dernier n'intervient pas dans la régulation protéine-protéine puisque la formation du complexe suffit à inhiber l'activité de la protéine cible. A l'inverse, les protéines cibles déjà présentes au sein de la bactérie continueront d'exercer leurs fonctions biologiques, même lorsque la transcription est éteinte. La répression ne sera totale que lorsqu'elles seront dégradées.

Dans le contexte décrit précédemment, la régulation par les sARN est la plus rapide. En effet, les gènes codant pour les sARN sont en général de petites tailles (de l'ordre d'une centaine de nts). Couplés à une transcription massive et suffisante pour avoir une molécule fonctionnelle, les sARN vont réguler rapidement la traduction de l'ARNm codant pour la cible. L'effet sur la traduction des ARNm sera optimal quand la concentration intracellulaire du sARN sera saturante. Sa vitesse de transcription est donc fondamentale. Néanmoins, la répression du système ne sera totale que lorsque les protéines présentes avant le stress sont dégradées. La régulation par les interactions protéine-protéine est plus lente à des temps proches du stimulus. Cependant à intervalle de temps plus long, cette régulation aboutit plus rapidement à une inhibition totale de la protéine cible car aucune étape de dégradation n'est nécessaire. La régulation par les TF se fixant sur les promoteurs est la plus lente. Le TF doit être transcrit, traduit en quantité suffisante et se fixer à l'ADN. Durant ce laps de temps, le gène cible continue d'être transcrit et traduit, et les protéines cibles resteront actives tant qu'elles ne sont pas dégradées. En résumé, pour répondre rapidement à un stress, la régulation par les sARN est la plus avantageuse (**Figure 5 A**), à condition qu'ils soient transcrits rapidement et massivement pour inhiber efficacement la traduction de l'ARNm cible.

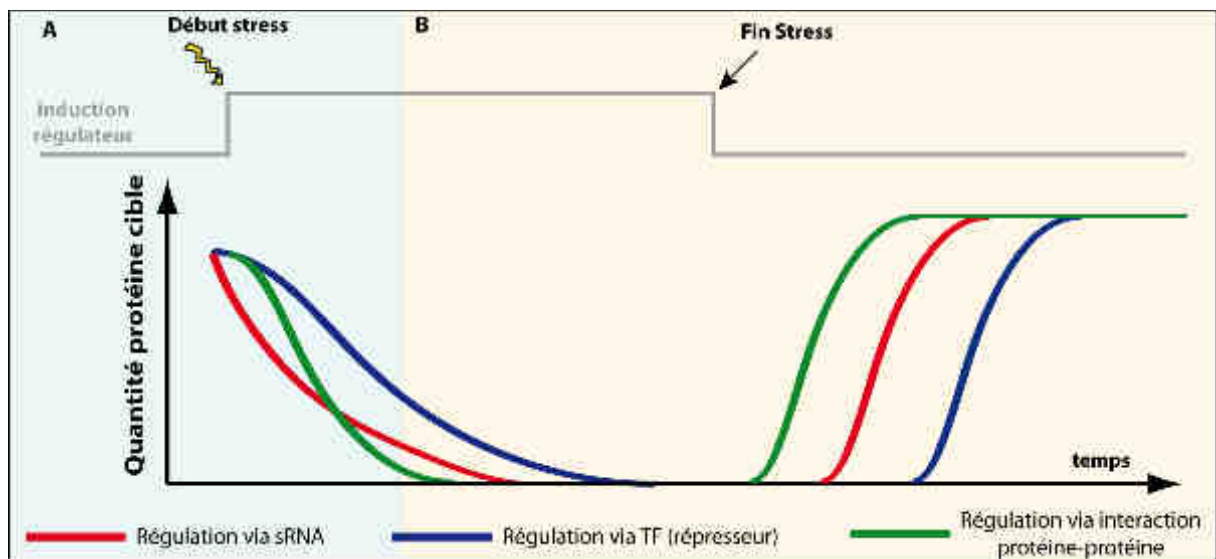
Que se passe t-il lorsque le stress est fini ? Dans le cas de nombreuses interactions protéine-protéine, la réponse au stress passe par un changement conformationnel favorisant au non la formation du complexe. Ainsi, la protéine ciblée est rapidement libérée et donc fonctionnelle après stress. Dans le cas des sARN, deux paramètres sont à prendre en compte. Premièrement, la vitesse de transcription des deux ARN, car tant que ce dernier est en quantité supérieur à l'ARNm, le système restera éteint. Ensuite, la stabilité du sARN détermine le laps de temps pendant lequel il régulera sa cible. De manière intéressante, de nombreux sARN sont stables. Chez les bactéries Gram positives ils sont généralement très structurés empêchant l'action des RNases (**Rsa A-K** chez *S. aureus*, voir IV.D.3). Chez les

bactéries Gram négatives, la liaison du sARN à Hfq favorise la formation des duplexes ARNm/sARN et protège ces derniers contre la dégradation (voir III.C.1). Le temps de latence est donc plus important pour les sARN que les interactions protéine-protéine. Ce temps de latence est intéressant dans les cas de fluctuations rapides du milieu extracellulaire. Dans une telle situation, la régulation par les sARN permettra de maintenir un niveau bas de la cible (Beisel *et al.*, 2010). Encore une fois, la régulation transcriptionnelle est la plus lente (**Figure 5 B**) (Shinomi *et al.*, 2007).

**Pourquoi la bactérie ne privilégie t-elle pas uniquement la régulation par des interactions protéine-protéine ?** Cette question mérite d'être posée puisque de nombreux système à deux composants (*Two Component System* ou TCS) et senseurs environnementaux se basent sur ce type d'interaction. En général, ces systèmes aboutissent à une activation ou répression directe de la transcription des gènes (exemple de l'opéron *agr*, voir IV.D.2). De manière intéressante, les bactéries ont astucieusement combiné la régulation protéine-protéine et les sARN pour établir des réseaux de régulation capables de répondre rapidement à un stress et intégrer de multiples signaux extracellulaires. En effet, la majorité des sARN identifiés sont sous le contrôle de senseurs environnementaux. Ces derniers sont soit des TF (OxyR, Fur), soit des régulateurs de transcription (*response regulator* ou RR) issus des systèmes à deux composant (TCS) (EnvZ-OmpR). Un TCS est constitué d'une histidine kinase membranaire (HK) qui lors d'une variation brusque dans l'environnement proche de la bactérie (pression osmotique, carence en nutriments etc.) subit une autophosphorylation. Le groupement phosphate généré est ensuite transféré sur le RR qui induit dans la plupart des cas un changement conformationnel de la protéine aboutissant à la libération du domaine de liaison à l'ADN (DBD ou *DNA binding domain*). La fixation du RR à l'ADN induit une transcription massive de gènes nécessaires à l'adaptation au stress (Stock *et al.*, 2000), dont les gènes des sARN. Cette transcription massive rend la régulation par les sARN très efficace, expliquant pourquoi les sARN ont été choisis au cours de l'évolution pour accélérer la réponse au stress. De plus, les sARN sont capables de cibler toutes les régions de l'ARNm, ouvrant de nouvelles perspectives de régulation (voir III.B) à l'inverse des RR et activateurs/répresseurs qui assurent leurs fonctions uniquement au niveau des promoteurs des gènes. L'accumulation de mutations aux niveaux des séquences

des sARN peut aboutir au ciblage de nouveaux ARNm. Un exemple est donné en III.B.2.b avec le régulon de  $\sigma^E$  qui active l'expression de plusieurs gènes lors d'un stress membranaire dont deux sARN. Le régulon de ces derniers implique des gènes contrôlant la viabilité de la bactérie et permet à l'activateur de la transcription  $\sigma^E$  de réprimer indirectement des gènes.

Un autre avantage de la régulation pourrait être l'économie d'énergie et métabolique due à l'absence de traduction. Néanmoins, cet aspect est sujet à débat de part la transcription massive d'ARN nécessaire pour avoir un système de répression efficace (Beisel *et al.*, 2010).



**Figure 5 : Simulations de réponses au stress d'après Shinomi *et al.*, 2007.**

La réponse via un sARN réprimant la traduction de sa cible est représentée en rouge. La régulation via l'interaction entre deux protéines est montrée, enfin la régulation via un répresseur de la transcription est en bleu. La cible à réguler est exprimée avant la réponse au stress, les régulateurs non. **A.** Niveaux de la protéine cible pendant la réponse au stress (début symbolisé par un éclair). **B.** Ré-expression de la cible après stress.

### III.A.2. Circuits de régulation

Les réseaux de régulation des ARN reflètent leurs fonctions au sein de la physiologie générale de la bactérie. La caractérisation de leurs fonctions passe obligatoirement par l'identification de leur(s) cible(s) (ARNm ou protéine), ainsi que des signaux qui induisent leur expression. Les circuits de régulation ont été particulièrement bien étudiés chez *E. coli* et *S. typhimurium* et ceux-ci sont incroyablement variés.

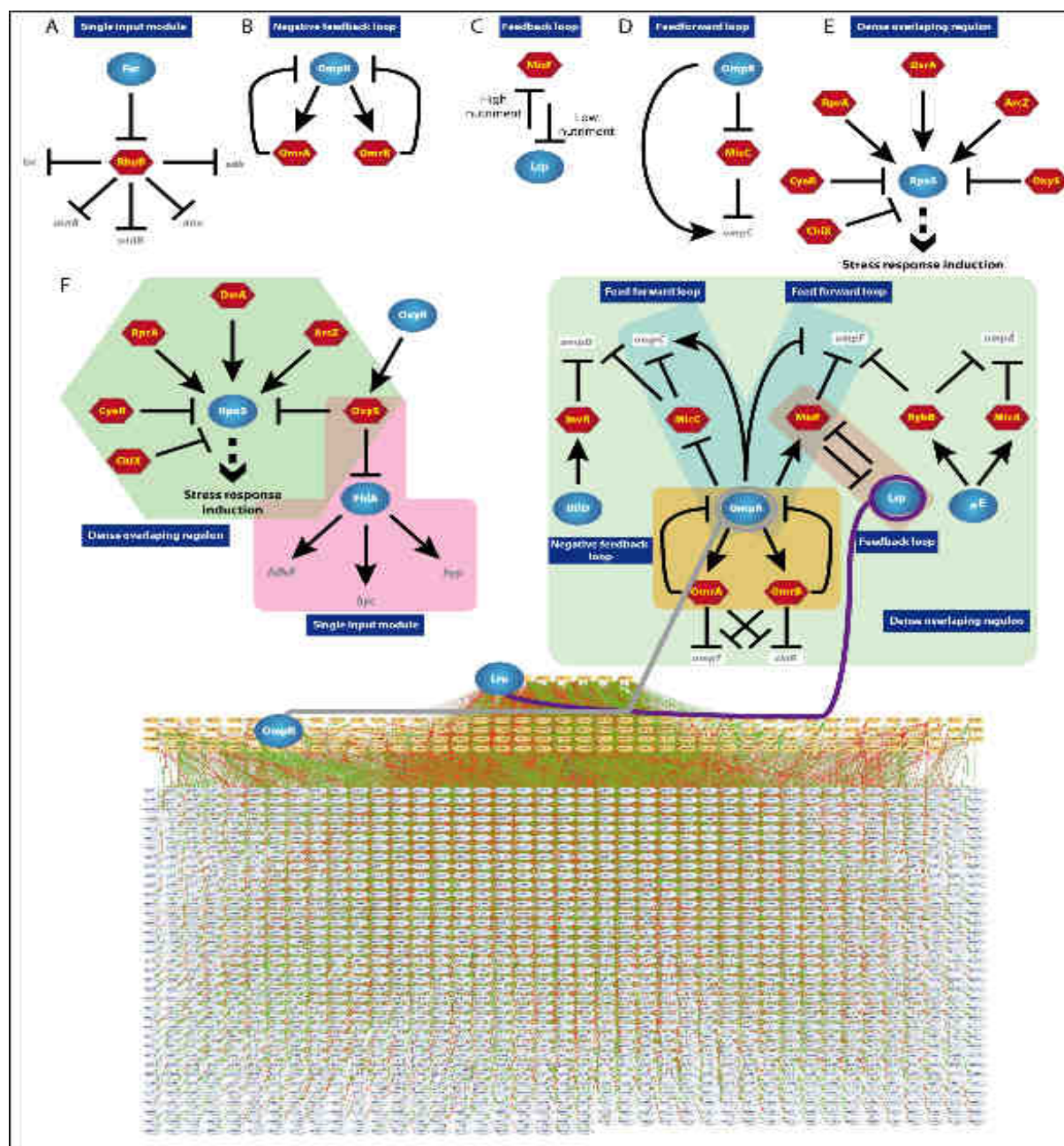
**Single input module :** L'expression d'un seul sARN peut coordonner l'expression de plusieurs gènes. Ce circuit de régulation est dit *single input module*. Par exemple, le sARN *RyhB* d'*E. coli* est uniquement induit lors d'une carence en fer. RhyB inhibe l'expression

d'ARNm codants des protéines utilisant le fer, dont l'expression est inutile à la bactérie lors d'une carence de cet élément (Massé *et al.*, 2002) (**Figure 6 A**). Dans ces conditions, RhyB est essentiel à la croissance bactérienne. En présence de concentrations optimales en fer, l'expression de *RyhB* est inhibée par le régulateur de transcription Fur pour conduire à la synthèse des protéines impliquées dans le métabolisme du fer et des mécanismes de détoxification des ions hydroxyles réactifs.

**Feedback loop** : Dans le cas des boucles de rétro-contrôle (*feedback loop*), la cible régule l'expression de son régulateur. Cette régulation peut être positive et aboutit alors à une amplification du signal, ou négative avec pour conséquence une vitesse de réponse accrue et une diminution de la variabilité de réponse intercellulaire. Un exemple de feedback loop négative est le réseau incluant les sARN *OmrA* et *OmrB*. La transcription de ces sARN est sous la dépendance de OmpR, qui est l'effecteur du TCS EnvZ-OmpR induit lors d'un choc osmotique. *OmrA* et *OmrB* inhibent la traduction d'ARNm codant des protéines membranaires (*ompT*, *ciaR*) tout en stimulant leur dégradation rapide par la RNase E. Mais *OmrA* et *OmrB* agissent également de la même manière sur l'expression d'OmpR (Guillier *et al.*, 2008) (**Figure 6 B et C**).

**Feedforward loop** : La *feedforward loop* (FFL) associe des régulateurs de nature différente qui vont contrôler l'expression d'un même gène. Ce type de circuit permet d'intégrer plusieurs signaux de régulation pour amplifier la réponse soit en suractivant ou au contraire en inhibant de manière irréversible l'expression du gène cible. A haute pression osmotique, l'ARNm *ompF* est inhibé de deux manières : OmpR est phosphorylé et inhibe sa transcription. En parallèle, OmpR active la transcription de l'ARN *MicF* qui inhibe la traduction de l'ARNm *OmpF* (Chen *et al.*, 2004) (**Figure 6 D**).

Ces circuits s'imbriquent les uns dans les autres pour former des réseaux denses et complexes capables d'intégrer simultanément plusieurs signaux provenant du milieu extracellulaire. Lorsqu'en plus ils incluent des TF, ces réseaux deviennent d'une étonnante complexité (**Figure 6 E, F et G**) (Beisel *et al.*, 2010).



**Figure 6 : Circuits de régulation de l'expression des gènes et interconnexion des réseaux.**

**A. SIM:** *single input module*. Fur réprime la transcription de RhyB qui inhibe la traduction de plusieurs ARNm impliqués dans le métabolisme du fer. **B. NFL :** *negative feedback loop*. Le TF OmpR induit la transcription des sARN OmrA et OmrB qui inhibent la traduction de protéines membranaires et d'OmpR. **C. FL :** *feedback loop*. Selon la quantité de nutriments disponible, MicF inhibe la traduction de Lrp, ou bien Lrp inhibe la transcription de MicF (Holmqvist *et al.*, 2011). **D. FFL :** *feedforward loop*. Selon la pression osmotique, OmpR inhibe la transcription de *ompC*. En parallèle, OmpR active la transcription de MicC qui inhibe la traduction de l'ARNm *ompC*. **E.** Réseau de régulation du facteur de transcription  $\sigma^S$  (codé par le gène *rpoS*), induit lors d'une réponse au stress (Beisel *et al.* 2010). **F. et G.** Interconnexion entre les circuits de régulation RpoS/OxyS (F) et OmpR/Lrp/RpoE/HliD (G). Lrp et OmpR sont des TF qui possèdent un réseau de régulation important (<http://regulondb.ccg.unam.mx/index.jsp>) qui s'ajoute au régulon décrit. Les ARN régulateurs sont représentés par des hexagones rouges, les TF par des ovales bleus. Les activations sont symbolisées par des flèches, les inhibitions par des traits se terminant par une barre perpendiculaire.

### III.A.3. Localisation intracellulaire des ARN et dynamisme des populations

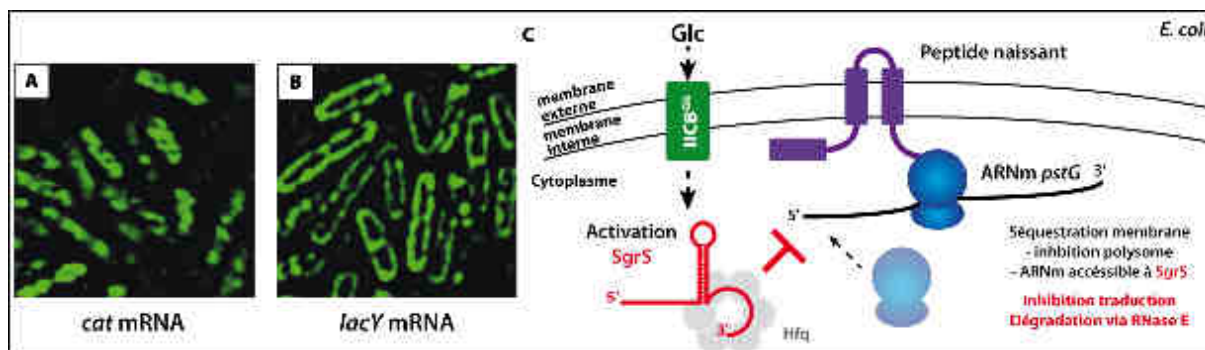
Pour terminer cette partie, j'aimerais mettre en avant des recherches récentes qui révèlent d'autres paramètres pouvant avoir une incidence sur le contrôle par les sARN.

***La localisation intracellulaire des ARNm joue-t-elle un rôle dans la régulation ?*** Le premier aspect repose sur la localisation des ARNm et des sARN dans la bactérie : certains ARNm seraient transportés pour être traduits et/ou régulés directement à l'endroit où la protéine est requise, comme par exemple, les protéines membranaires. Nevo Dinur *et al.*, 2011 ont suivi la localisation des ARNm *cat*, codant pour une acétyltransférase cytoplasmique, et *lacY*, codant pour une perméase liée à la membrane. Les ARNm ont été fusionnés en 5' avec 6 copies de l'ARN aptamère, qui est spécifiquement reconnu par la protéine MS2, elle-même fusionnée à la GFP (*Green Fluorescent Protein*). La localisation des ARNm est donc réalisée par le biais de la formation d'une ribonucléoparticule chez *E. coli*. Les résultats obtenus montrent que le signal de l'ARNm *cat* adopte une structure hélicoïdale caractéristique d'une localisation chromosomique, alors que l'ARNm *lacY* s'accumule à la membrane de la bactérie (Nevo-Dinur *et al.*, 2011) (**Figure 7 A et B**). A l'inverse, Montero-Llopis *et al.*, 2010, qui ont utilisé une approche similaire, ont montré que les ARNm sont transcrits et traduits au locus où ils sont codés. Ces résultats contradictoires montrent qu'il est encore difficile d'échaffauder des hypothèses sur la localisation des ARNm au sein des bactéries mais l'utilisation de nouvelles méthodes de marquage de l'ARN pourrait lever rapidement ces ambiguïtés permettant la localisation intracellulaire d'ARN et un suivi en temps réel de leurs déplacements (Paige *et al.*, 2011).

Néanmoins, le transport des ARNm peut avoir un impact sur la régulation par les ARN. L'équipe d'H. Aiba a montré que la traduction de l'ARNm *ptsG* est réprimée par l'ARN *SgrS*, lorsque l'ARNm est préférentiellement localisé à la membrane par l'insertion du peptide naissant à la membrane interne. Dans des conditions normales de croissance, la localisation de l'ARNm *ptsG* permet sans doute d'empêcher d'autres cycles de traduction. Localisé à la membrane, l'ARNm est moins accessible au ribosome donc moins protégé, permettant à *SgrS* d'interagir avec ce dernier. La liaison de *SgrS* à l'ARNm *ptsG* résulte en une inhibition de la traduction suivie d'une dégradation rapide de l'ARNm par la RNase E. Cette



étude souligne que les sARN doivent diffuser (passivement ou activement) rapidement à l'endroit où sont localisés les ARNm cibles, afin d'éviter un délai trop long dans la réponse au stress (Kowatano *et al.*, 2005) (Figure 7 C).

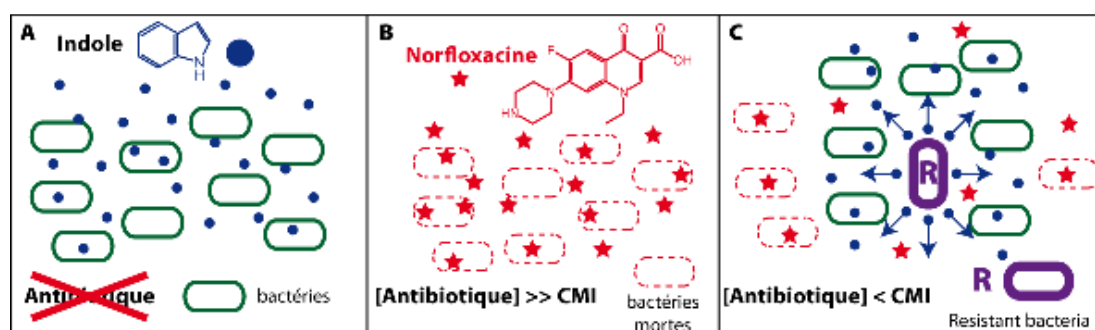


**Figure 7 : Localisation des ARNm chez *E. coli* et régulation de l'ARNm *pstG*.**

**A. et B.** L'ARNm *cat* code pour la chloramphénicol acétyltransférase, une protéine cytoplasmique. Le gène *lacY* code pour la lactose perméase, une protéine liée à la membrane. Les transcrits sont visualisés via la MS2::GFP qui se lie aux aptamères MS2 clonés dans la région 5' des ARNm (Nevo-Dinur *et al.*, 2011). **C.** L'ARNm *ptsG* est localisé à la membrane pour éviter plusieurs cycles de traduction. Séquestré à la membrane, l'ARNm est accessible au sARN *SgrS*, induit lors d'un stress au glucose. L'hybridation de *SgrS* à l'ARNm *ptsG* résulte en une inhibition de la traduction couplée à une dégradation médiée par la RNase E. La protéine Hfq est requise dans la régulation (Kowatano *et al.*, 2005).

**Mécanisme de résistance aux antibiotiques au sein d'une population.** Un autre niveau de régulation concerne le comportement des cellules individuelles au sein d'une population. Une étude récente de Lee *et al.*, 2010 décrit le comportement d'une population de bactéries *E. coli* exposée de manière cyclique à des doses croissantes d'un antibiotique. De manière inattendue, la population s'adapte dynamiquement à l'augmentation constante de la concentration en norfloxacine. La grande majorité des isolats prélevés possède une résistance moindre à l'antibiotique que la population dans son ensemble, à l'exception de quelques clones hyper-résistants. Ces derniers possèdent une CMI (concentration minimale inhibitrice) à la norfloxacine plus importante que la CMI mesurée au sein de la population. La présence de ces quelques bactéries hyper-résistantes améliore la tolérance à l'antibiotique de la population entière. De quelle manière ce message est-il transmis à la population ? Le séquençage des génomes des isolats résistants révèle une mutation conservée dans des gènes cibles de la norfloxacine. Par exemple, tous les isolats présentent des mutations récurrentes dans le gène de l'ADN gyrase *gyrB*, qui est l'une des cibles de l'antibiotique et une mutation dans le gène codant une potentielle oxydoréductase conservée. D'autres mutations sont spécifiques à chaque clone résistant. La présence commune de deux

mutations dans tous les isolats tend à montrer qu'il y a eu un "ancêtre" résistant à partir duquel d'autres mutants ont dérivés. Un autre facteur déterminant est la production d'indole par les cellules. Cette molécule, produite normalement par des bactéries non stressées, est importante dans l'adaptation au stress oxydatif et dans l'activation des pompes à efflux de médicaments. C'est la capacité des clones résistants à synthétiser en continu des molécules d'indole qui est cruciale dans la réponse à l'antibiotique : en effet, ces bactéries continueront à produire de l'indole même en présence de norfloxacine. Cet indole activera les réponses de tolérance au stress des bactéries sensibles à l'antibiotique se trouvant dans leur voisinage proche, qui pourront alors survivre en maintenant leur production de pompes à efflux de médicaments. C'est ainsi que le taux de survie général de la population est amélioré (**Figure 8**). Les différences entre bactéries au sein d'une population et les communications cellule-cellule à grande échelle permettent ainsi d'intégrer les signaux provenant du milieu extérieur pour favoriser une adaptation rapide aux conditions de stress. Il est donc fort possible que la régulation par les sARN ne se fasse pas de manière homogène dans toutes les cellules.



**Figure 8** : Résistance aux antibiotiques d'une population bactérienne d'après Lee *et al.*, 2010.

**A.** Population non stressée. Toutes les cellules sécrètent un niveau normal d'indole (rond bleu) **B.** La population est soumise à une dose massive de Norfloxacine (étoile rouge) détruisant l'ensemble de la population **C.** La population est soumise à une dose modérée d'antibiotique (60% de la CMI). Une partie de la population est détruite. L'autre est protégée grâce à l'apparition d'un mutant résistant qui continue à sécréter des molécules d'indole, ce qui va activer l'expression des pompes à efflux de médicaments et les gènes de tolérance aux antibiotiques des bactéries à son voisinage.

### III.B. Mécanismes d'action des ARN régulateurs

La fonction d'un ARN régulateur est portée par sa structure et certains éléments de séquence. Plusieurs classes d'ARN régulateurs ont été caractérisées à ce jour. Les thermosenseurs et les *riboswitch* sont des ARNm portant directement les signaux nécessaires à leur régulation. Les ARN non-codants agissant *en trans* se divisent en trois

catégories. Les deux premières regroupent des ARN régulateurs qui se lient à des ARNm cibles par complémentarité de bases. On distingue les ARN antisens, codés au même locus que leurs gènes cibles, et les ARN régulateurs codés à un autre locus. La troisième regroupe les ARN régulateurs interagissant directement avec des protéines. Enfin, il existe des ARN régulateurs codant des peptides et des toxines (Fozo *et al.*, 2008), et certains d'entre eux possèdent des domaines multiples de régulation (Romilly *et al.*, 2012). Les figures illustrant cette partie sont issues de deux revues dont je suis co-auteur : **RNA-mediated regulation in bacteria: from natural to artificial systems** (Lioliou *et al.*, 2010) et **RNA mediated regulation of virulence gene expression: another layer of complexity** (Lioliou *et al.*, 2011)

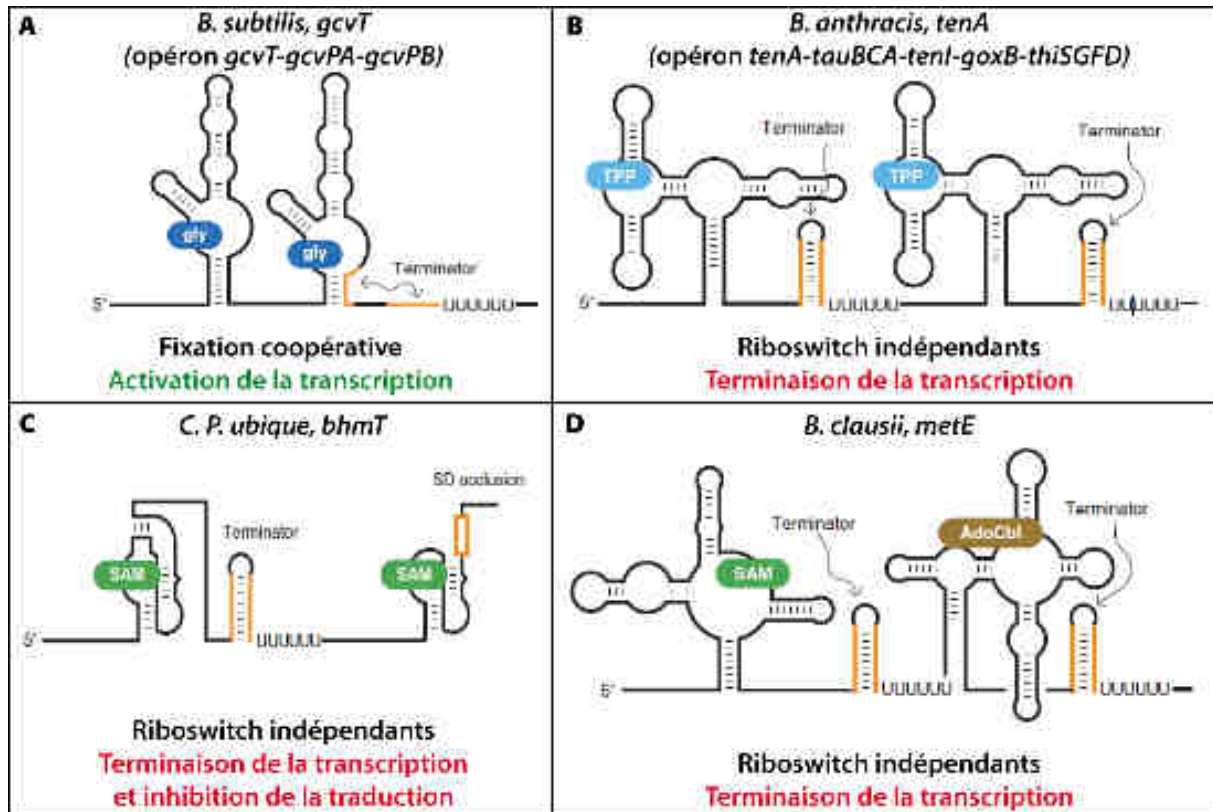
### III.B.1. Les éléments agissant en *cis*

Les régions 5'UTR des ARNm ont évolué au cours du temps pour devenir de véritables senseurs de l'état métabolique de la cellule. Ces régions 5'UTR, communément appelées "*riboswitch*", sont capables de fixer des métabolites (thiamine pyrophosphate (TPP), adénosylcobalamine (AdoCbl), flavine mononucléotide (FMN), glycine, S-adénosyl-méthionine (SAM), lysine, Glucosamine-6-phosphate (GlcNP), guanine, adénine, 7-aminoéthyl 7-deazaguanine (preQ1), vitamine B-12) sans l'aide de cofacteurs (revues : (Winkler *et al.*, 2005, Coppins *et al.*, 2007, Smith *et al.*, 2010, Breaker 2011). Aux *riboswitch* s'ajoutent les *T-box*, impliquées dans la liaison des ARNt et la modulation des besoins en amino-acyl ARNt synthétase (revue Green *et al.*, 2010), et les thermosenseurs capables de détecter des variations de température. Je détaillerai par la suite uniquement les *riboswitch*, en soulignant l'étonnante évolution qui s'est opérée entre les bactéries Gram positives et négatives, et les thermosenseurs pour leurs liens entre la température et la synthèse des facteurs de virulence.

#### III.B.1.a. Les *riboswitch*

**Caractéristiques des *riboswitch*.** Les *riboswitch* les plus simples se composent de deux parties. La région dite aptamère est responsable de la reconnaissance spécifique du métabolite. La fixation de ce dernier entraîne un changement conformationnel de la plateforme d'expression. Ce changement de structure a des conséquences au niveau

transcriptionnel, en favorisant la formation d'un terminateur ou d'un antitermineur de transcription, ou au niveau traductionnel en rendant accessible ou en piégeant le site de fixation du ribosome. De manière intéressante, la structure et la séquence des aptamères sont très conservées au cours de l'évolution en opposition à la conservation de la plateforme d'expression (Winkler *et al.*, 2005). Depuis, des situations plus complexes montrant des *riboswitch* en tandem sont décrites dans la revue de R. Breaker (Breaker 2012). Dans le cas du gène *gcvT* chez *B. subtilis*, deux aptamères lient de manière coopérative de la glycine pour réguler l'expression d'une seule plateforme de régulation au niveau transcriptionnel (Mandal *et al.*, 2004) (Figure 9 A). Chez *Bacillus anthracis*, le gène *tenA* est sous le contrôle de deux *riboswitch* indépendants. Chacun d'entre eux fixe la TPP pour favoriser la formation d'un terminateur de transcription (Welz *et al.*, 2007) (Figure 9 B). Chez une bactérie marine, *Candidatus Pelagibacter ubique*, deux *riboswitch* indépendants contrôlent le gène *bhmT* et fixent la SAM. L'originalité de ce système est que le premier inhibe la transcription alors que le second régule la traduction de l'ARNm (Pojata *et al.*, 2009) (Figure 9 C). Enfin, chez *Bacillus clausii*, le gène *metE* est sous le contrôle de deux *riboswitch* indépendants modulant la transcription, le premier fixant de la SAM, le second du AdoCbl (Sudarsan *et al.*, 2006) (Figure 9 D).

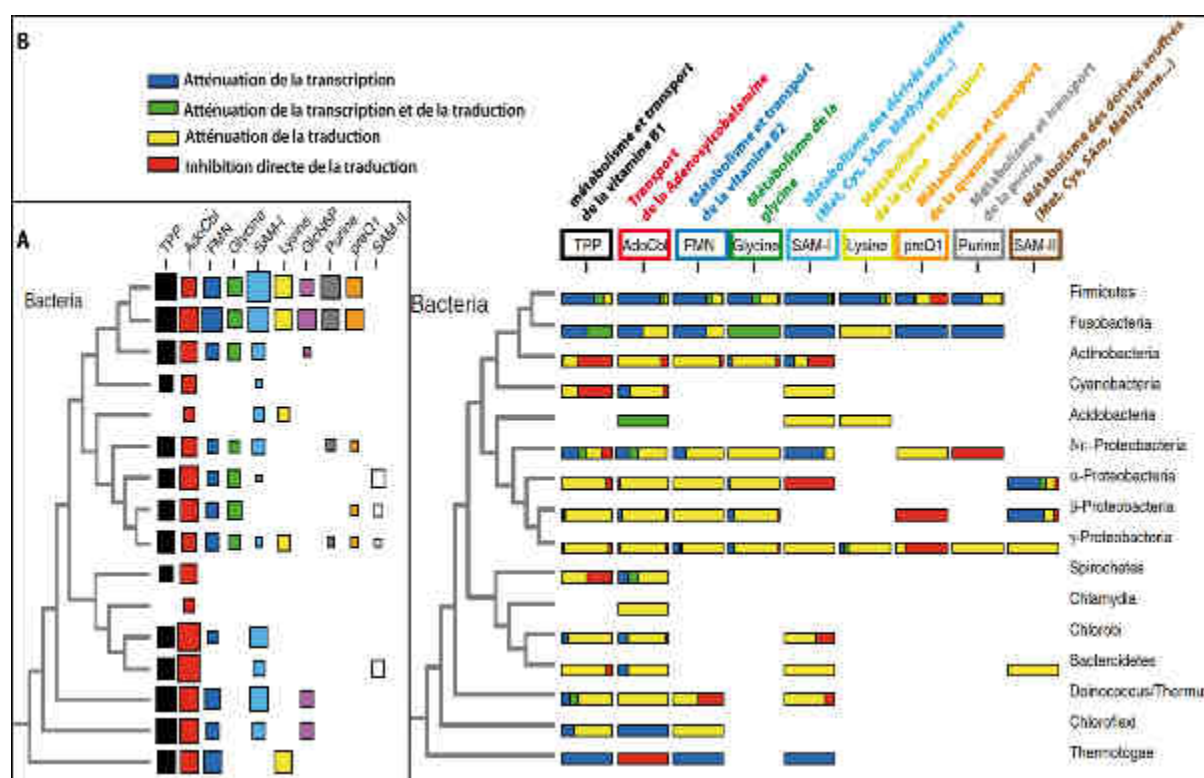


**Figure 9 : Mécanisme de régulation des riboswitch en tandem, d'après Breaker 2012.**

**A.** Chez *B. subtilis*, l'opéron *gcvT-gcvPA-gcvPB* est sous le contrôle d'un riboswitch ayant deux aptamères liant la glycine et régulent de manière coopérative une seule plateforme d'expression. La fixation du ligand empêche la formation d'un terminateur de transcription (Mandal *et al.*, 2004). **B.** Chez *B. anthracis* l'opéron *tenA-tauBCA-tenI-goxB-thiSGFD* est sous le contrôle de deux riboswitch TPP indépendants, favorisant la formation d'un terminateur de transcription en présence du ligand (Welz *et al.*, 2007). **C.** Chez *C. P. ubiqua*, le gène *bhmT* est sous le contrôle de deux riboswitch SAM. La fixation du ligand conduit à la terminaison de la transcription ou l'inhibition de la traduction (Poiata *et al.*, 2009). **D.** Chez *B. clausii*, le gène *metE* est sous le contrôle de deux riboswitch indépendants, l'un liant SAM, l'autre AdoCbl. La fixation de l'un des ligands inhibe la transcription du gène (Sudarsan *et al.*, 2006).

**La fréquence et les mécanismes de régulation des riboswitch varient entre bactéries Gram positives et négatives.** Les riboswitch sont généralement trouvés en amont d'un opéron. Ce dernier regroupe des gènes impliqués dans la biosynthèse et le transport du ligand fixé par le riboswitch. Les bactéries ont ainsi développé un mécanisme permettant de moduler en temps réel l'expression des gènes appartenant à une même voie métabolique, selon les ressources du métabolite en question (Winkler *et al.*, 2005, Coppins *et al.*, 2007, Smith *et al.*, 2010, Breaker 2011). D'un point de vue évolutif, les riboswitch fixant la TPP, FMN, SAM, et AdoCbl sont très conservés chez les bactéries (les TPP étant également conservés chez les champignons et les plantes). Chez les firmicutes (bactéries à faible pourcentage en GC comme *S. aureus* et *B. subtilis*), l'usage des riboswitch est très répandu, puisque l'ensemble des classes sont retrouvées (TPP, AdoCbl, FMN, glycine, SAN-1 lysine,

GlcNP, guanine, adénine, preQ1) à l'exception de SAM de classe II, qui reste spécifique aux proteobactéries (*E. coli*). A l'inverse, même si toutes les classes sont retrouvées chez les  $\gamma$ -portéobactéries, les espèces n'utilisent pas toutes les classes présentes. Par exemple, *E. coli* ne possède que quatre types de riboswitch (TPP, AdoCbl, FMN, lysine). Un autre aspect intéressant est que les *riboswitch* présents chez les firmicutes (Gram positif) régulent majoritairement la transcription des gènes en opposition au Gram négatif, qui utilise majoritairement le contrôle traductionnel (Barrick *et al.*, 2007). Cette observation a-t-elle un lien avec l'utilisation préférentielle de terminateurs de transcription Rho-indépendant chez les firmicutes (Hoon *et al.*, 2005) ou l'organisation des gènes en opéron plus importantes chez les bactéries Gram positives (253 opérons prédits chez *E. coli* contre 565 chez *B. subtilis*) (Dam *et al.*, 2007) ? (Figure 10)



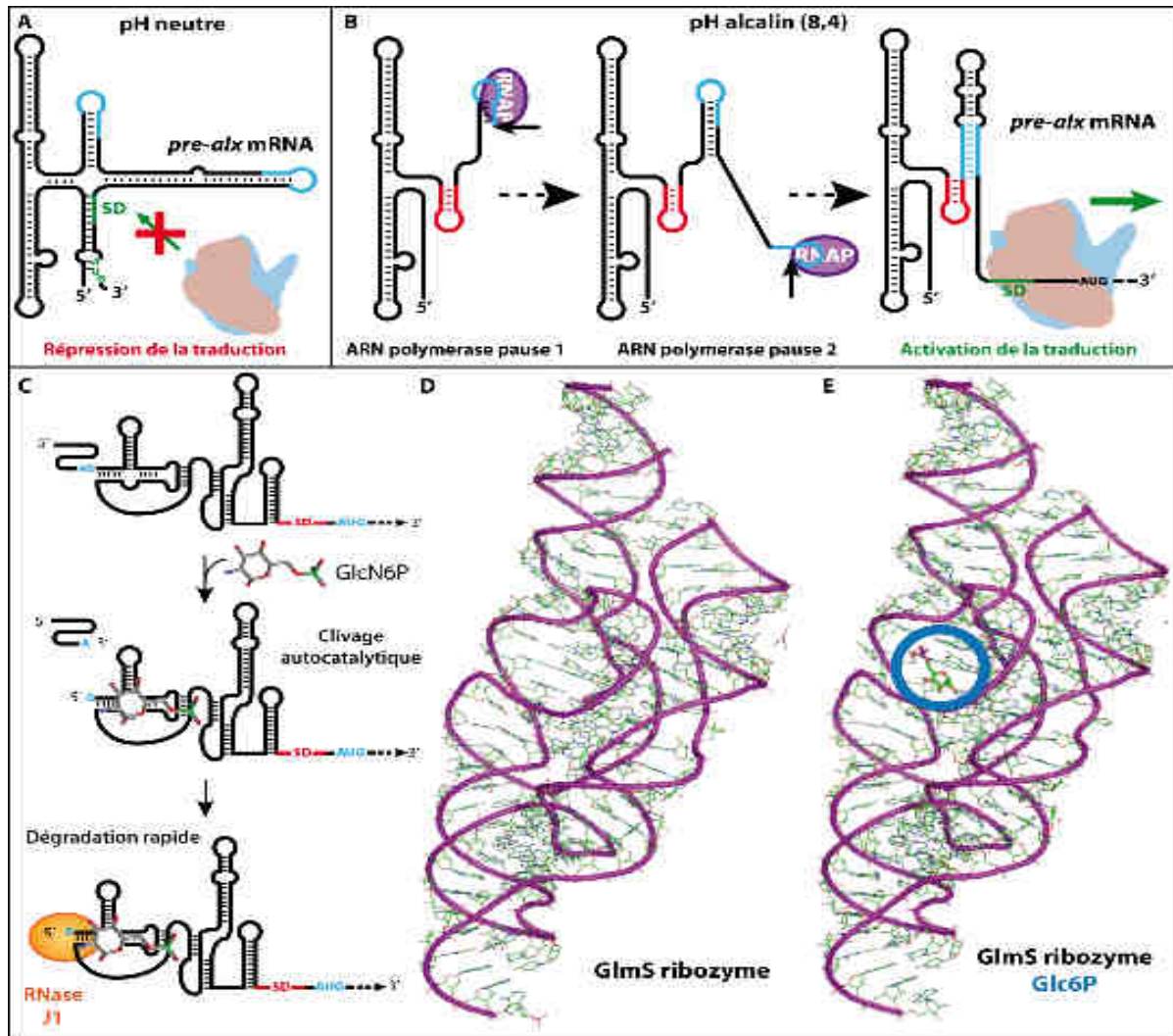
**Figure 10 : Etude prédictive : fréquence d'apparition des différentes classes de *riboswitch* et proportion des mécanismes de régulation utilisés chez les bactéries (Barrick et Breaker 2007).**

**A.** Distribution des *riboswitch* : la taille des carrés est proportionnelle à la fréquence d'apparition de la classe de *riboswitch* dans le groupe taxonomique considéré. **B.** Proportion des mécanismes de régulation utilisés par de la plateforme (régulation transcriptionnelle, traductionnelle ou les deux) parmi les *riboswitch* prédits.

**La structure du *riboswitch* de l'ARNm *alx* est sensible à la vitesse de la RNAP.** La vitesse de transcription est régulée par le taux de traduction (Proshkin *et al.*, 2010) et peut être sensible à certaines conditions de stress rencontrées par la bactérie. C'est la base de la

régulation de l'expression de l'ARNm *alx* sensible en fonction des variations du pH chez *S. typhimurium*. Pour que l'ARNm puisse être traduit, une transcription *de novo* est indispensable. A pH neutre, la transcription de la région 5' UTR aboutit à la formation d'une structure qui piège le site de liaison du ribosome (RBS ou *ribosome binding site*) dans une région d'ARN double brin, rendant impossible la traduction. A l'inverse, à pH alcalin, les pauses successives de l'ARNP favorisent la conformation active de la région 5'UTR permettant la traduction de l'ARNm *alx* (Nechooshtan *et al.*, 2009) (Figure 11 A et B).

**Le riboswitch *glmS* de *B. subtilis* ne change pas de conformation avec la liaison de *GlcN6P*.** Chez *Bacillus subtilis*, l'expression de l'ARNm *glmS* est sous le contrôle d'un ribozyme qui affecte la stabilité de l'ARNm. Une caractéristique remarquable est que sa structure est indépendante de la fixation de *GlcN6P*. La fixation de *GlcN6P* aboutit à un clivage auto catalytique du ribozyme rendant l'extrémité 5' de l'ARNm accessible à la RNase J1, ce qui entraîne une dégradation rapide de cet ARNm (Collins *et al.*, 2007) (Figure 11 C, D et E). De manière surprenante, la régulation de l'ARNm *glmS* chez les entérobactéries ne fait pas intervenir de *riboswitch*, mais deux sARN agissant en *trans* qui activent la traduction (Görke *et al.*, 2008).



**Figure 11** : Régulation de l'expression de la région 5'UTR de l'ARNm *alx* chez *E. coli*. (Nechooshtan *et al.*, 2009) de l'ARNm *glmS* chez *B. subtilis* (Collins *et al.*, 2007).

**A.** Structure secondaire de la partie 5'UTR de l'ARNm *alx* à pH neutre. L'ARNP transcrit l'ARNm sans pause, le RBS, constitué de la séquence Shine et Dalgarno (SD, en vert) et du codon d'initiation de la traduction AUG (en vert également), est piégé dans une région double-brin empêchant l'initiation de la traduction **B.** A pH alcalin, l'ARNP (RNAP, en violet) effectue deux pauses critiques (flèches oranges) durant la transcription de l'ARNm. La première permet la formation d'une tige boucle (en rouge), la seconde l'interaction de deux tige-boucle distantes (en bleu). La structure résultante présente un RBS accessible au ribosome, permettant l'initiation de la traduction. **C.** Structure secondaire du ribozyme. A haute concentration, la glucosamine-6-phosphate (GlcN6P) libre va se fixer sur la région aptamère située en 5'UTR de l'ARNm. Cette liaison n'induit pas de modification de la structure secondaire mais un clivage auto catalytique site spécifique. Le fragment généré subit une dégradation rapide *via* la RNase J1 (Collins *et al.*, 2007). **D.** Structure cristalline du ribozyme *glmS* (PDB 2HOX). **E.** Structure cristalline du ribozyme *glmS* en présence de glucose-6-phosphate (Glc6P) (PDB 2HOZ), un inhibiteur compétitif du GlcN6P ne générant pas de clivage (Klein *et al.*, 2006).

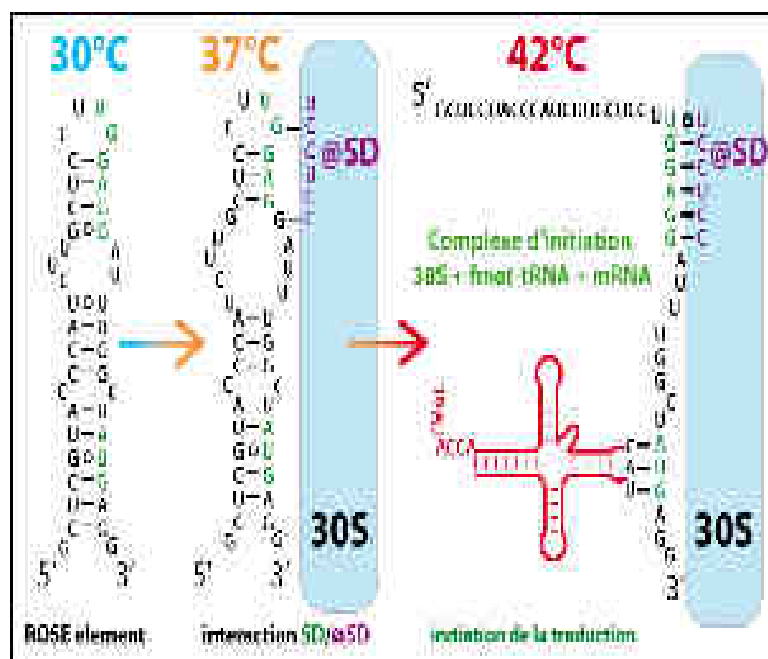
### III.B.1.b. Les thermosenseurs

**La structure des thermosenseurs répond aux variations de température.** Le mode d'action de la grande majorité des thermosenseurs (Naberhaus *et al.*, 2010) est basé sur une



dénaturation progressive de la région 5'UTR d'un ARNm en fonction de la température. En effet, le RBS est souvent piégé à basse température (30°C) dans une structure secondaire qui est instable à plus haute température. La séquence SD est alors libérée et devient accessible à la petite sous-unité du ribosome qui peut initier la traduction de l'ARNm.

**Les appariements de bases non canoniques constituent le noyau fonctionnel des thermosenseurs.** Les thermosenseurs sont associés aux réponses aux chocs thermiques et à la virulence de certains pathogènes. La conservation de séquence est relativement faible, mais certains thermosenseurs présentent une structure secondaire conservée, comme les éléments ROSE (*Repressor Of heat Shock protein Expression*) et four-U (Naberhaus *et al.*, 2010) où la séquence SD est impliquée dans des appariements de bases non canoniques. Une étude NMR (Chowdhury *et al.*, 2006) montre la dénaturation progressive de l'élément ROSE en fonction de l'augmentation de la température qui est initiée au niveau des appariements non canoniques



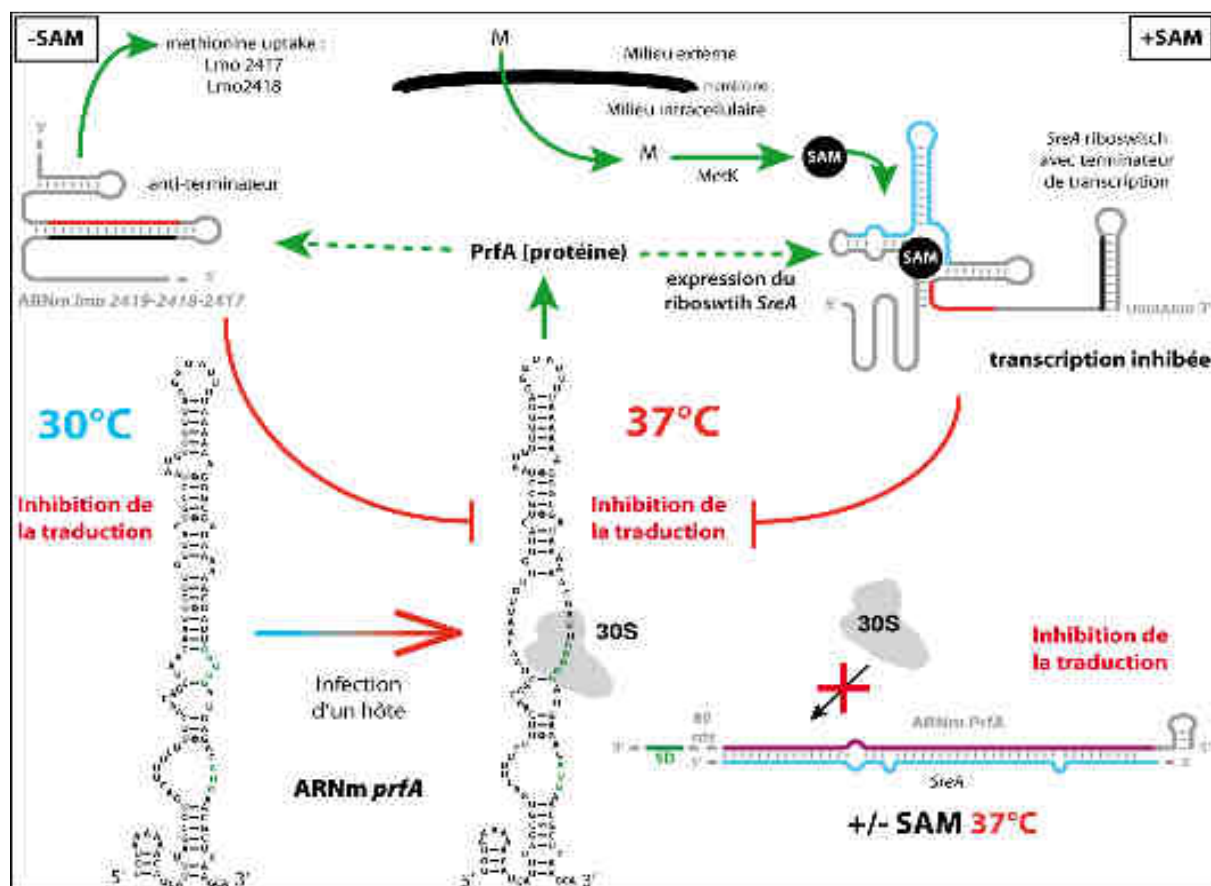
(Figure 12).

**Figure 12 : Dénaturation successive de l'élément ROSE, d'après Chowdhury *et al.*, 2006.**

Etude de la fusion du thermosenseur ROSE de *Bradyrhizobium japonicum* par RMN. A 30°C (en bleu), la séquence SD (en vert) et le codon d'initiation de la traduction AUG (en vert) sont piégés par des interactions faibles. A 37°C (en orange), les interactions faibles qui piègent la séquence SD fusionnent, permettant l'initiation de l'interaction SD/anti-SD (@SD, en violet) avec la sous-unité 30S. A 42°C (en rouge) la fusion de l'élément ROSE est totale, permettant une complète interaction SD/@SD et la venue de l'ARNt initiateur.

**Un thermosenseur est à l'origine de l'expression des facteurs de virulence chez *L. monocytogenes*.** *L. monocytogenes* ne produit des facteurs de virulence qu'à 37°C, soit la température de son hôte ; cette particularité est due à un thermosenseur situé dans la région 5'UTR de l'ARNm *prfA*, un TF régulant les gènes de virulence chez *L. monocytogenes* (Scotti *et al.*, 2007). A 37°C, la structure qui piège la séquence SD se désagrège, rendant la traduction de l'ARNm possible (Johansson *et al.*, 2002) De manière unique, ce

thermosenseur est régulé par un *riboswitch* fixant de la SAM au niveau l'ARNm *sreA*. Dans ce cas précis, le riboswitch agit *en trans* via la complémentarité de bases pour empêcher la traduction de l'ARNm *prfA* en piégeant le RBS. La formation de l'hybride riboswitch/ARNm n'est possible qu'à haute température, lorsque le site de liaison sur l'ARNm *prfA* est accessible. Ainsi, la virulence de la bactérie est doublement contrôlée par la température et la disponibilité en nutriments (Loh *et al.*, 2009) (**Figure 13**).



**Figure 13 :** Chez *L. monocytogenes*, l'expression du facteur de transcription PrfA est contrôlée par la température et l'état métabolique.

La structure secondaire du thermosenseur présent en 5'UTR de l'ARNm *prfA* piège le RBS à basse température. A plus haute température (chez un hôte), ce dernier devient accessible au ribosome. L'accumulation de la protéine PrfA induit l'expression des facteurs de virulence et de l'ARNm *sreA*. Ce dernier forme avec PrfA une boucle de retro-contrôle négative. Le riboswitch en 5' de l'ARNm *serA* peut agir *en trans* en s'appariant à la forme ouverte de l'ARNm *prfA* pour piéger le RBS et inhiber la traduction (Loh *et al.*, 2009)

### III.B.2. Les ARN régulateurs ciblant les ARNm

Deux classes peuvent être distinguées. D'une part, les ARNs, codés au même locus que leur ARNm cible, sont strictement complémentaires à leur cible. A l'inverse, les sARN codés à un autre locus que l'ARNm cible, peuvent tolérer des mésappariements dans leur

liaison. Les sARN ont donc la capacité de reconnaître plusieurs ARNm cibles. La majorité des ARN régulateurs, dont la fonction a été caractérisée inhibent la traduction d'ARNm en piégeant le RBS dans une région en double-brin. Les données accumulées montrent cependant que toute la séquence de l'ARNm peuvent être ciblées par des sARN (Figure 14). La formation des hybrides sARN-ARNm peut influencer la transcription, la traduction, ou la stabilité de l'ARNm.

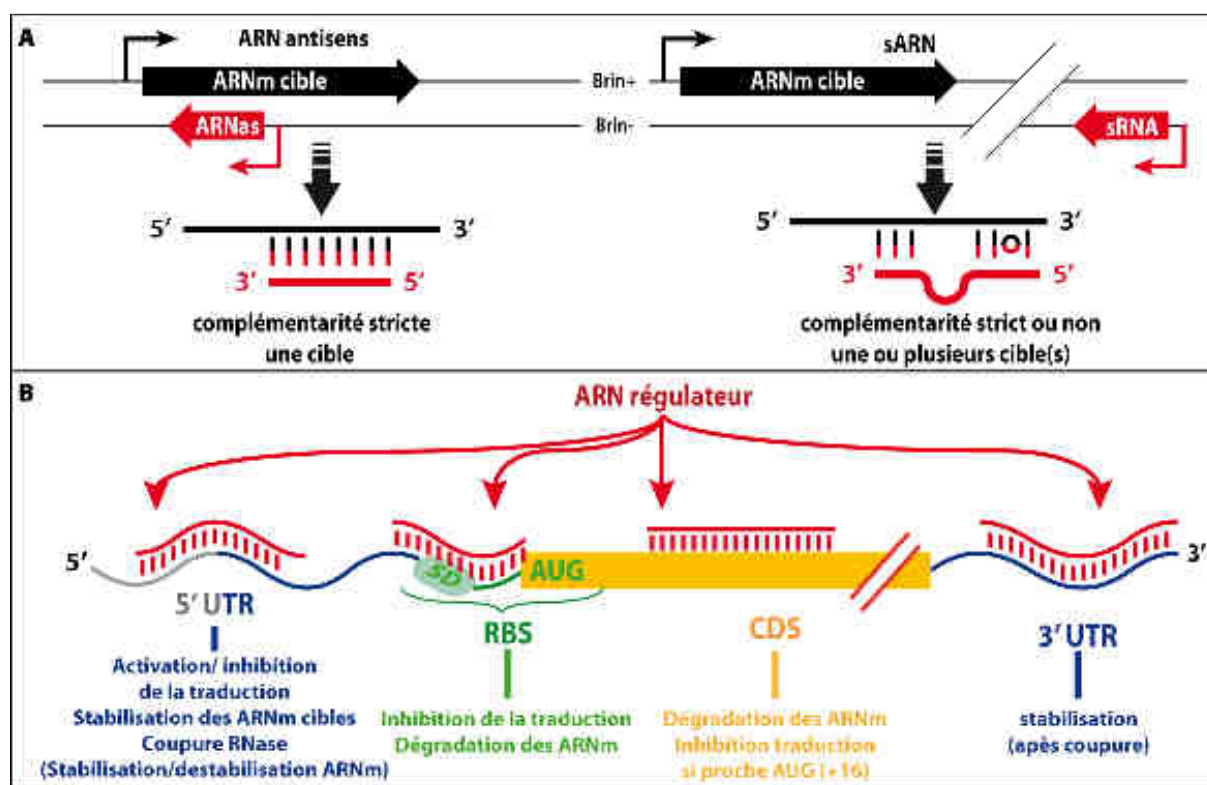


Figure 14 : La régulation par les ARN chez les bactéries.

A. Généralités sur les ARN régulateurs et les sARN. B. Régions de l'ARNm ciblées par les ARN.

### III.B.2.a. Les ARN antisens

Les ARNas ont été les premiers ARN régulateurs identifiés chez les bactéries. Jusqu'à récemment, la présence d'ARNas était confirmée principalement dans les éléments génétiques mobiles, tels que les plasmides, les transposons et les phages (Wagner *et al.*, 2002). Les ARN antisens contrôlent le taux de répllication, le taux de conjugaison ou la partition des plasmides. Chez les phages, ils permettent le passage du cycle lytique au cycle lysogénique. Certains ARNas du chromosome régulent l'expression des toxines (système toxines-antitoxines de type I) impliquées dans des voies métaboliques, la réponse au stress

ou la virulence. Les mécanismes de régulation influent sur la transcription, la traduction et la stabilité des ARN ciblés (revue Brantl *et al.*, 2007, Georg *et al.*, 2011).

**Un ARNas contrôle la réplication du plasmide R1.** La réplication du plasmide R1 est sous le contrôle de l'ARNas CopA (Nordström *et al.*, 1984). Cet ARNas est transcrit constitutivement et est rapidement dégradé. Dans des cellules en division rapide, les concentrations de CopA et du répresseur transcriptionnel CopB sont faibles, favorisant la transcription de l'ARNm codant pour RepA qui initie la réplication du plasmide. Cet ARNm est faiblement traduit et sa traduction doit être couplée à celle d'un peptide leader situé en 5' de l'ARNm, le peptide Tap. La traduction du peptide libère le site de fixation du ribosome de l'ARNm *repA*, permettant sa traduction (revue Nordström *et al.*, 2005). CopA est complémentaire à la région 5'UTR de l'ARNm *tap* appelée CopT. En interagissant avec CopT, CopA empêche la fixation du ribosome sur le RBS de *tap* et à cause du couplage traductionnel, inhibe indirectement la synthèse de RepA. L'efficacité du contrôle est directement reliée à la vitesse de fixation de CopA à CopT. Le mécanisme d'interaction est un processus multi-étapes qui commence par une reconnaissance boucle-boucle. Cette interaction s'étend de manière irréversible pour former une hélice ARN à quatre brins stabilisée par une hélice régulière de 30 pb située 2 nucléotides en amont de la séquence SD de l'ARNm *tap*. Ce long duplexe est ciblé par la RNase III qui dégradent les deux ARN (Kolb *et al.*, 2000, Kolb *et al.*, 2001) (Figure 15 A).

**Maturation de l'extrémité 3' UTR par appariement d'un ARNas.** GadY est un ARNas chromosomique impliqué dans l'adaptation à un stress acide. Le gène codant pour GadY se situe dans l'IGR de l'opéron *gadXW* et est complémentaire à la région 3'UTR de *gadX*. L'ARNm *gadXW* est instable et rapidement dégradé mais la liaison du sARN GadY induit une maturation qui génère deux ARNm distincts et plus stables, conduisant à une traduction des protéines GadX et GadW (Opdyke *et al.*, 2004). La RNase III est impliquée dans cette maturation (Opdyke *et al.*, 2010), mais un mécanisme alternatif requiert également la RNase E (Takada 2007) (Figure 15 B).

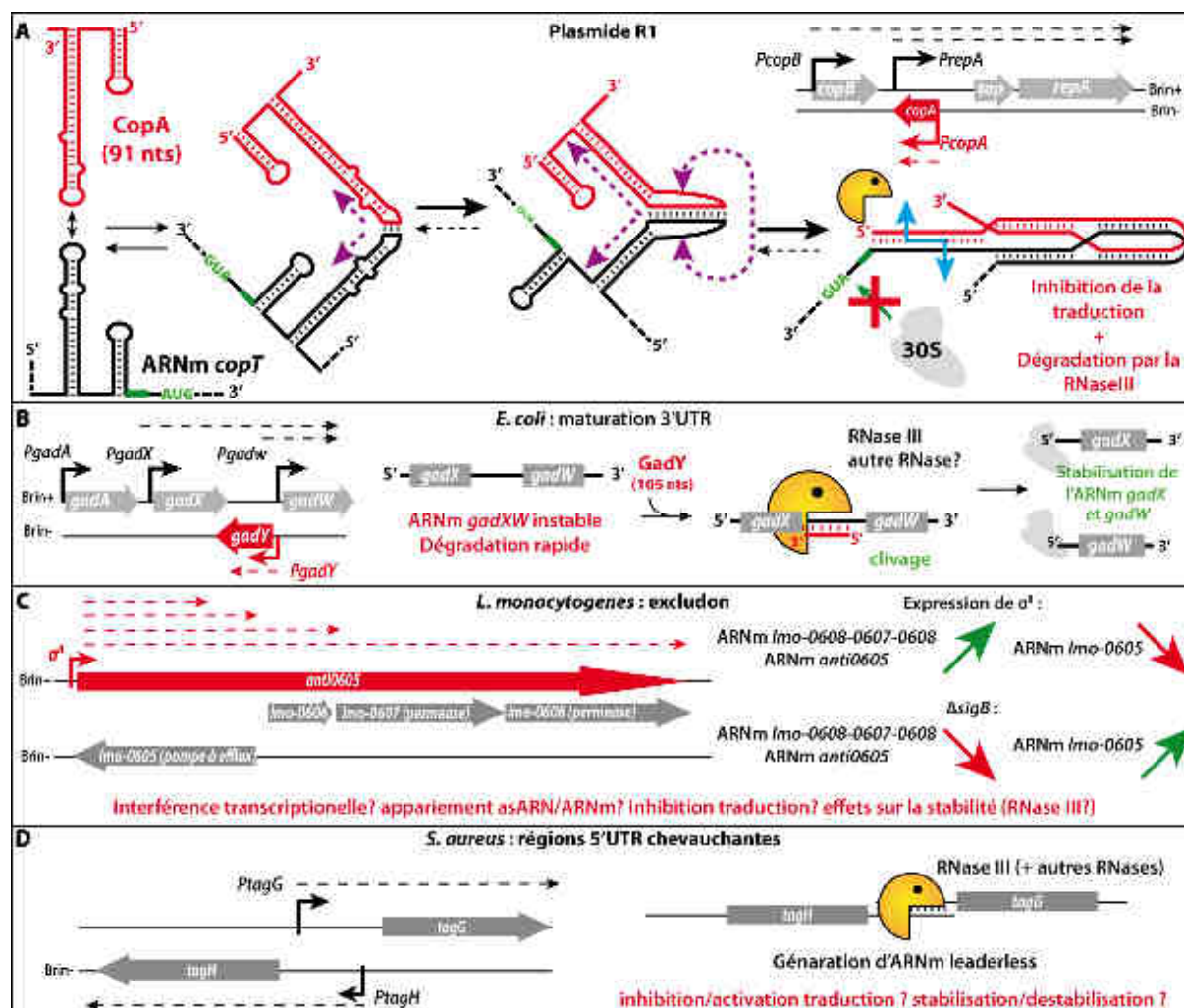
**Un long ARNas est à l'origine de l'expression différentielle des pompes à efflux chez *L. monocytogenes*.** Des études de séquençage à haut débit menées chez *L. monocytogenes* ont abouti à l'identification de long ARNas grâce à l'enrichissement en transcrits primaires des échantillons séquencés. L'un d'entre eux, *anti0605*, a une taille de 5,8 kb et est sous le

contrôle d'un promoteur  $\sigma^B$ . Sa partie 5' est complémentaire au gène *Imo0605* codé sur l'autre brin ADN alors que sa partie 3' code pour l'opéron *Imo0606-0607-0608*. Le gène *Imo0605* code pour une pompe à efflux de médicament de la famille MatE, l'opéron *Imo0606-0607-0608* pour deux pompes à efflux ATP-dépendantes et un TF. Dans une souche sauvage,  $\sigma^B$  induit la transcription l'ARNas *anti0605* qui s'accompagne d'une augmentation de l'expression de l'opéron et d'une diminution des niveaux de *Imo0605*. A l'inverse, une souche mutante n'exprimant plus  $\sigma^B$  montre des niveaux d'ARNm *Imo0605* plus important que celle de l'opéron. S'agit t-il d'un cas d'interférence transcriptionnel ? *anti0605* s'apparie t-il à l'ARNm *Imo0605* pour inhiber sa traduction ou diminuer sa stabilité, et si oui, quelles RNases seraient impliquées dans la dégradation du duplexe, avec quels effets sur l'opéron *Imo0606-0607-0608* ? Grâce à cet ARNas, *L. monocytogenes* dispose d'un système permettant de découpler l'expression de deux familles de pompes en efflux. Il a été montré que les pompes à efflux de la famille MatE sont exprimées préférentiellement lorsque la bactérie se trouve dans la circulation sanguine ; à l'inverse les pompes ATP dépendantes qui sont exprimées au niveau de l'intestin. De manière intéressante, ce système est également trouvé chez une bactérie proche mais non-pathogène, *L. innocua*, mais très faiblement exprimé. Le répertoire d'ARNas et l'étude de leur expression pourraient donc permettre de comprendre les mécanismes d'adaptation à l'humain des bactéries pathogènes (Wurtzel *et al.*, 2012) (Figure 15 C).

***L'ARNas AmgR module l'expression des facteurs de virulence pour pérenniser l'infection de Salmonella enterica.*** Définir les conditions dans lesquelles les ARNas sont synthétisés est essentiel pour appréhender leurs fonctions biologiques. Les premières étapes de l'infection par *S. enterica* requièrent la synthèse des facteurs de virulence pour coloniser l'hôte. Cependant, pour persister durablement, elle doit aussi veiller à masquer sa présence au système immunitaire de l'hôte en limitant la synthèse de ces facteurs (Lee *et al.*, 2010). Cette dynamique est en particulier régulée par un ARNas de 1.2 kb, nommé *AmgR*, qui est complémentaire à l'ARNm codant la protéine MgtC. Cette protéine est indispensable à la virulence, la survie contre les macrophages et à des concentrations faibles en magnésium ( $Mg^{++}$ ) (Alix *et al.*, 2007). De manière surprenante, le régulateur et sa cible sont sous le contrôle du même TF, PhoP, sensible à la concentration intracellulaire de  $Mg^{++}$ . Cependant, le promoteur de *amgR* est plus faible que celui de *mgtC*. Lors d'une carence en  $Mg^{++}$ , PhoP

active la transcription des deux ARNm. Mais l'ARNm *amgR* est transcrit à un taux moindre que *mgrC*, ce qui permet l'accumulation de la protéine MgtC pendant un laps de temps jusqu'à ce que la concentration de l'ARNas augmente. Il est intéressant de noter que l'inactivation de *AmgR* rend *Salmonella* hypervirulente. Ce mécanisme original met en évidence l'utilisation d'un ARN régulateur comme une horloge biologique répondant à un métabolite, ici le  $Mg^{++}$ , pour contrôler finement la virulence de *S. enterica*.

**Les régions 5'UTR chevauchantes, nouveaux mécanismes de régulation de l'expression des gènes ?** Récemment, I. Lasa *et al.*, (2011) et E. Lioliou *et al.*, (2012) ont montré que des ARNas de *S. aureus* sont exprimés de manière faible sur l'ensemble des deux brins du génome, et que ces derniers sont rapidement dégradés par la RNase III. De manière intéressante, plusieurs ARNm possèdent des régions 5'UTR chevauchantes qui sont ciblées par la RNase III pour aboutir à la formation d'ARNm leaderless ou avec une région 5'UTR très courte. C'est le cas des ARNm *tagG* et *tagH* qui possèdent une région chevauchante d'environ 180 nucléotides et qui codent chacun pour une sous-unité du transporteur ABC TagGH. La RNase III, en coopération avec d'autre(s) RNase(s) encore inconnue(s), est responsable du raccourcissement des régions 5' UTR chevauchantes des deux ARNm. Celle de l'ARNm *tagH*, après coupure, s'arrête quelques nucléotides en amont de la séquence SD. Ces ARNm sont-ils traduits ? Quel est l'impact de cette "maturation" sur leurs stabilités ? Cette transcription omniprésente a été décrite dans d'autres bactéries mais le rôle de la RNase III dans la dégradation des ARNas semblerait spécifique aux bactéries Gram-positives (Lasa *et al.*, 2011) (**Figure 15 D**).



**Figure 15 : Exemples de régulation par les ARNs.**

**A.** Régulation de la réplication du plasmide R1 par l'ARN *CopA* (en rouge), antisens à l'ARNm *copT* (en noir) codant pour RepA. Le processus d'interaction se fait en plusieurs étapes. La première consiste en une interaction boucle-boucle stable, qui se propage pour former une jonction à 4 hélices. Cette structure empêche la fixation du ribosome sur le RBS de *tap* (Kolb *et al.*, 2001). Les duplexes sont dégradés par la RNase III **B.** Stabilisation de l'ARNm *gadXW* via l'interaction de l'ARN antisens *GadY* en 3' UTR. Le duplex est clivé par une RNase, ce qui aboutit à une stabilisation des transcrits (Klein *et al.*, 2006). **C.** Chez *L. monocytogenes*, l'ARNas anti0605 est complémentaire à *Imo-0605* et possède la séquence codante pour l'opéron *Imo-0606-0607-0608*. L'ARNas est sous le contrôle d'un promoteur  $\sigma^B$ , et aboutit à une expression de l'opéron et une inhibition de l'expression de l'ARNm *Imo-0605*. Le mécanisme exact de régulation reste à déterminer (Wurtzel *et al.*, 2012). **D.** Chez *S. aureus*, les régions 5'UTR chevauchantes des ARN *tagG* et *tagH* sont résolues par la RNase III, générant des régions 5'UTR très courtes, notamment pour l'ARNm *tagH* (Lioliou *et al.*, 2012)

### III.B.2.b. Les sARN agissant en trans

La liaison d'un sARN à l'ARNm cible ouvre la voie à une pléthore de mécanismes de régulation aussi bien au niveau de la transcription, l'initiation de la traduction que sur la stabilité des ARN. Les études réalisées chez les entérobactéries ont permis d'établir certains concepts généraux sur la fonction et le mode d'action des sARN: leur transcription est

souvent sous le contrôle de signaux extérieurs (stress, système de densité cellulaire, voir III.A.1), ils coordonnent l'expression de nombreux gènes qui sont fonctionnellement reliés (métabolisme du fer et des sucres, facteurs de virulence, protéines membranaires) ou régulent la synthèse des facteurs de transcription induisant de nombreux effets indirects ; les sARN agissent en concert avec la protéine Hfq sur leurs ARNm cibles (voir III.C.1). Quelques exemples qui illustrent la diversité des mécanismes sont présentés ci-dessous.

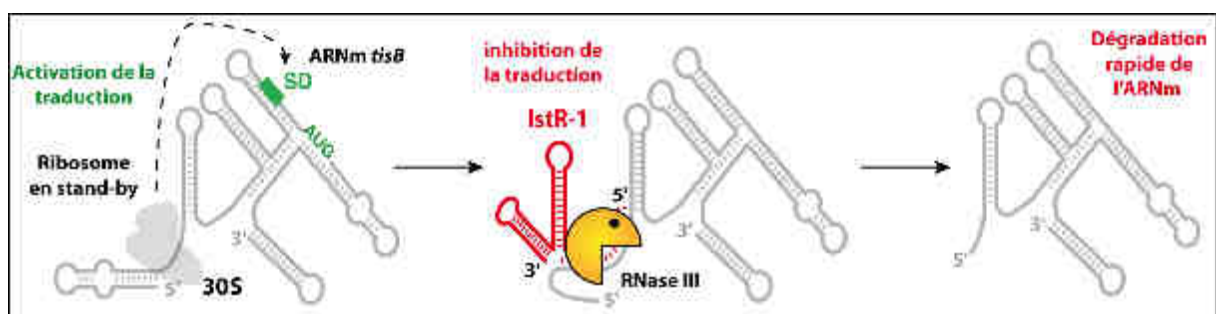
**Le sARN se fixe au RBS pour réprimer la traduction.** Les protéines membranaires (OMP ou *Outer Membrane Protein*) et les liposaccharides (LPS) assurent l'intégrité de la bactérie en formant une barrière sélective qui la protège contre des stress extérieurs ou des corps étrangers comme les peptides antimicrobiens du système immunitaire. Chez *E. coli* et *S. typhimurium*, des ARN régulateurs sont au cœur des réseaux de régulation des protéines de l'enveloppe pour adapter rapidement la croissance de la bactérie à un large panel de stress (Guillier *et al.*, 2006, Papenfort *et al.*, 2009). Des dommages de la membrane externe ou une accumulation de protéines mal repliées, conduit à l'activation de  $\sigma^E$ , séquestré au niveau de la membrane interne. L'expression de plus de 100 gènes est activée directement par  $\sigma^E$ , i.e. des protéines chaperonnes, des protéines impliquées dans la synthèse de LPS ou d'autres facteurs  $\sigma$  (comme *rpoH* ou  $\sigma^{32}$ , impliqué dans la réponse aux chocs thermiques) (Rhodius *et al.*, 2006) (Figure 16 A). Cependant, l'induction de  $\sigma^E$  entraîne aussi une diminution de l'expression de certains gènes, ce qui au premier abord peut paraître paradoxal pour un activateur de la transcription. Cette répression est en fait indirecte et implique l'activation de deux ARN régulateurs, *MicA* et *RybB*, qui sont les répresseurs du régulon de  $\sigma^E$  (Gogol *et al.* 2011). Ces sARN régulent l'expression de plusieurs ARNm cibles, dont certains sont en commun, chacun utilisant un site de liaison différent (Figure 16 B). Certains circuits semblent contradictoires. Par exemple, *MicA* et *RybB* régulent négativement l'expression de *yfeK* et *htrG*, qui sont pourtant induits par  $\sigma^E$ . Cette "feedforward" loop contrôle un mécanisme de mort programmée de la bactérie, la surexpression de ces deux protéines étant létale pour la bactérie. Lorsque la perturbation de l'homéostasie membranaire atteint un niveau trop important qui ne peut être secouru, la surexpression de  $\sigma^E$  et donc de *yfeK* et *htrG*, devient létale. La mort programmée des bactéries diminue la taille de la population et la compétition pour l'assimilation des nutriments en période de stress sera moindre. *MicA* et *RybB*, avec la protéine Hfq (III.C.1)





**Le sARN se fixe au niveau d'un site transitoire du ribosome appelé « stand-by ».**

Chez *E. coli*, le système toxine/antitoxine *tisAB* joue un rôle dans la réponse SOS déclenchée par des dommages au niveau de l'ADN. Ce système consiste en la régulation de l'expression de l'ARNm *tisAB* par le sARN *IstR1*. *TisAB* est une toxine dont l'expression est sous le contrôle du répresseur LexA. La transcription de *IstR1* est constitutive et permet de neutraliser le niveau basal de l'ARNm *tisAB* qui échappe au contrôle par LexA. L'ARNm *tisAB* est fortement structuré et son RBS est inaccessible. Pour être traduit, le ribosome se fixe sur une région éloignée en aval du RBS en 5'UTR, appelée "stand-by". Il a été suggéré que le ribosome effectue un saut depuis ce site pour se lier au RBS transitoirement accessible. L'ARN *IstR1* cible la région "stand-by" par un appariement de 30 bases sans l'aide de Hfq, empêchant la venue du ribosome. L'inhibition indirecte de la traduction de l'ARNm cible est rapidement suivie par la dégradation du duplex par la RNase III (Darfeuille *et al.*, 2007). Dès que la réponse SOS est enclenchée, la répression assurée par LexA est levée, entraînant la synthèse de l'ARNm *tisAB*, et donc de la toxine. La présence de cette toxine peut soit ralentir la croissance, permettant ainsi à *E. coli* de s'adapter au stress (Vogel *et al.*, 2004), soit être létale pour économiser les ressources et nutriments au profit du reste de la population (Bukowski *et al.*, 2011) (Figure 17). On pourrait se demander si ce type de mécanisme n'est pas plus général sachant que dans de nombreux ARNm chez *E. coli*, la séquence SD est souvent prise dans une structure secondaire. Par ailleurs, d'autres sARN se fixent dans des régions en simple brin éloignées du site RBS (Holmqvist *et al.*, 2010). Récemment, Desnoyers & Massé (2012) ont montré que *Spot 42* se fixe à l'ARNm *sdhC* dans une région éloignée du RBS et recrute la protéine Hfq pour inhiber l'initiation de la traduction (III.C.1).



**Figure 17 : Mécanisme de répression de la traduction par l'ARN, d'après Darfeuille *et al.*, 2007.**

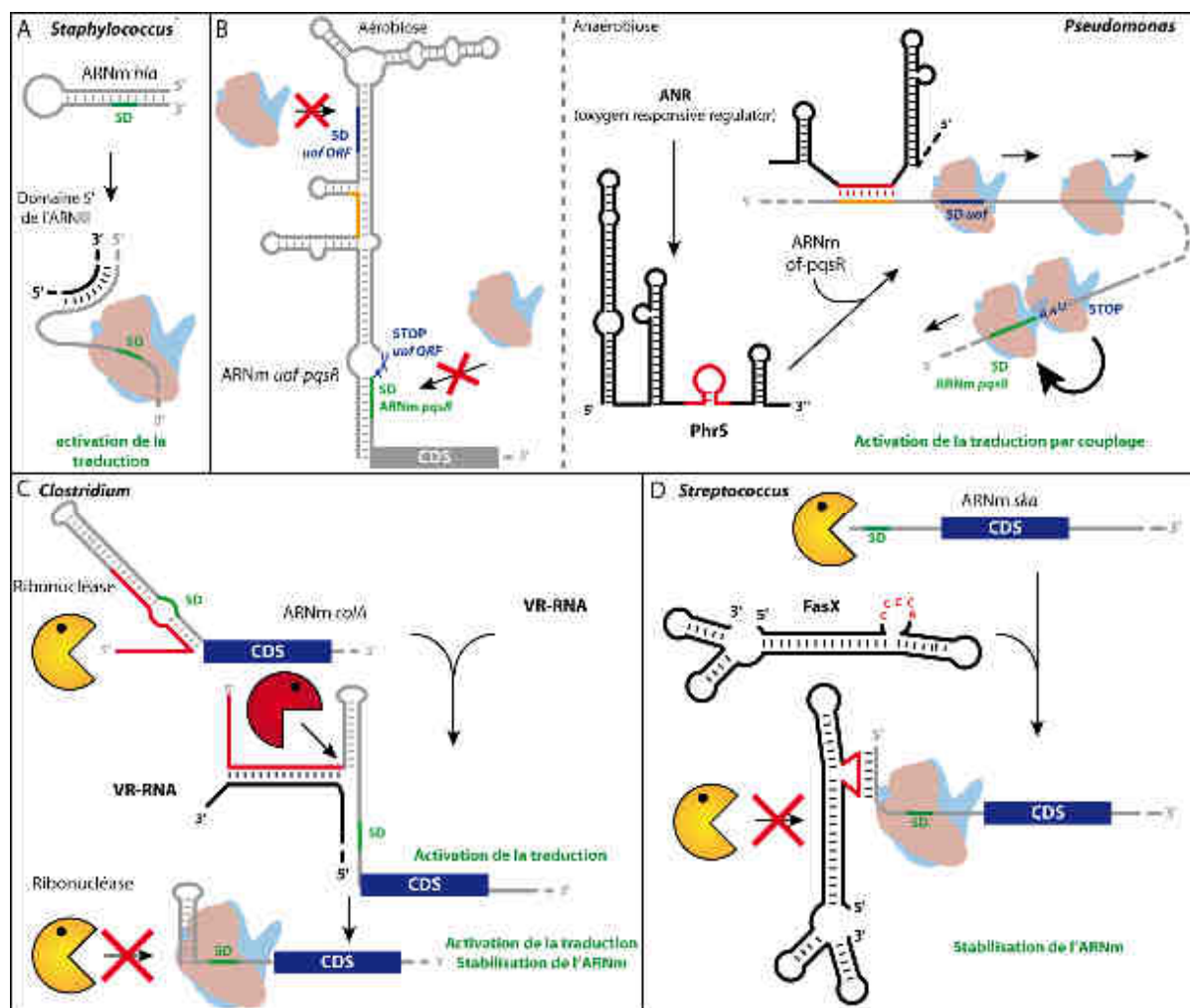
Dans des conditions normales, le ribosome se fixe sur une région en 5'UTR, puis "saute" vers le RBS. L'ARN *IstR-1* cible cette région "stand-by". Le duplex formé est ciblé par la RNase III. L'ARNm non traduit sera ensuite rapidement dégradé.

**Le sARN active la traduction par une diversité de mécanismes.** De nouveaux mécanismes originaux d'activation ont été découverts récemment. Un premier mécanisme consiste à empêcher la formation d'une structure inhibitrice séquestrant le RBS d'un ARNm. C'est le cas de l'ARNIII qui possède plusieurs domaines de régulation, dont le domaine 5' qui interagit en amont du RBS de l'ARNm *hla*, codant l'hémolysine  $\alpha$ . L'ARNIII empêche la formation d'une région en double brin pour activer la traduction de l'ARNm (Figure 18 A).

Chez *Pseudomonas aeruginosa*, le sARN *PhrS* active la traduction de l'ARNm *uof-pqsR* de manière indirecte en utilisant un couplage traductionnel. Les RBS des deux ORF sont piégés dans une structure en double-brin. En condition anaérobie, le régulateur ANR (*Oxygen response regulator*) active la traduction de *PhrS*. Par une région hautement conservée, *PhrS* se lie à l'ARNm *uof-pqsR* et libère le RBS du peptide leader *uof*. La traduction du peptide leader libère la séquence SD du gène *pqsR* situé en 3' pour activer sa traduction (Sonnleitner *et al.*, 2011) (Figure 18 B).

Un troisième mécanisme conduit à la stabilisation des ARNm. Chez *Clostridium perfringens*, le RBS de l'ARNm *colA*, codant la collagénase, est inaccessible au ribosome car piégé dans une structure en tige-boucle stable. En se fixant à l'extrémité 5' de l'ARNm, l'ARN VR induit une maturation spécifique de l'ARNm par une endoribonucléase inconnue qui permet de générer un ARNm doté d'une région 5'UTR plus courte. Cet ARNm présente une tige-boucle stable en 5' mais un RBS libre et accessible, permettant sa traduction. L'absence de résidus non appariés en 5' empêche probablement la dégradation de l'ARNm par la RNase J1 (Obana *et al.*, 2010) (Figure 18 C).

L'ARNm *ska*, qui code un facteur de virulence de *Streptococcus pyogenes*, est reconnu par le sARN *FasX*. La fixation de *FasX* n'a pas d'effet direct sur la traduction, mais protège l'extrémité 5' de l'ARNm *ska* probablement contre l'activité exoribonucléase de la RNase J1 (Ramirez-Pena *et al.*, 2010) (Figure 18 D).



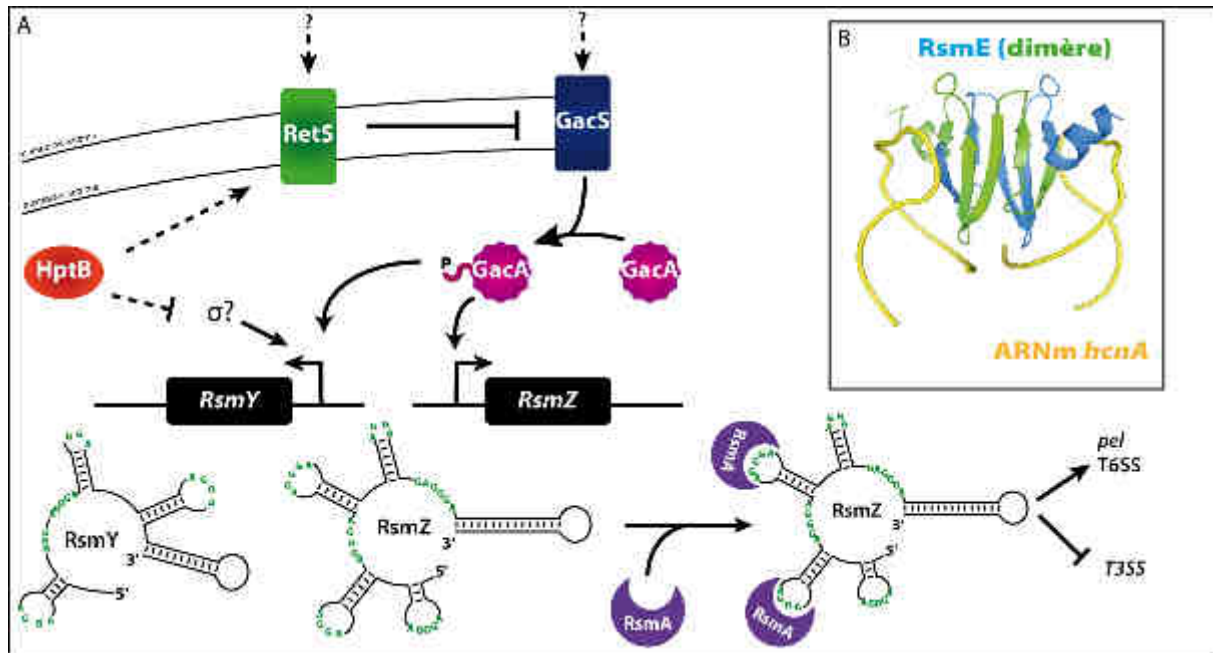
**Figure 18 : Activation de la traduction et stabilisation des ARNm.**

Les ARN régulateurs sont en noir et les ARNm cible en gris. **A.** Chez *S. aureus*, l'ARNIII cible l'ARNm *hla* et libère le RBS piégé dans une région en double-brin, activant ainsi la traduction (Morfelt *et al.*, 1996) **B.** Chez *P. aeruginosa*, le passage en condition anaérobie induit l'expression de *Phr5*. Sa liaison à l'ARNm *uof-pqsR* libère le RBS du peptide leader *uof*. La traduction de *uof* libère le RBS de *pqsR* (Sooleitner *et al.*, 2011). **C.** Chez *C. perfringens*, la liaison du sARN VR à l'ARNm *colA* va induire le clivage par une endoribonucléase (en rouge). L'ARNm généré est à la fois plus stable de par la présence d'une tige-boucle en 5' qui le protège de l'action de la RNase J1 (en jaune) et accessible au ribosome (Obana *et al.*, 2010). **D.** Chez *S. pyogenes*, le sARN *FasX* se lie à l'extrémité 5' de l'ARNm *ska*, le protégeant de la dégradation par la RNase J1 (Ramirez-Pena *et al.*, 2010).

### III.B.3. Les sARN séquestrant des protéines

Certains ARN exercent leur fonction en interagissant directement avec des protéines et en modulant leur activité. Chez *P. aeruginosa*, les ARN *RsmY* et *RsmZ* se fixent à la protéine *RsmA*. Cette interaction aboutit à une répression des gènes impliqués dans le système de sécrétion de type 3, et de manière additive, a un effet positif sur le gène *pel* et le système de sécrétion de type 6 impliqué dans la formation de biofilm. La transcription de ces sARN passe par une cascade de régulation complexe qui débute par le système TCS

GacA/GacS. L'activation de GasA *via* une phosphorylation stimule l'expression de RsmY et RsmZ. Ce système TCS est régulé soit par le senseur RetS (Ventre *et al.*, 2006), soit par la protéine HptB qui appartiennent à deux réseaux de signalisation différents. RetS inactive directement GacS, et indirectement, réprime la synthèse des deux sARN, alors que HptB ne module que l'expression de RsmY (Bordi *et al.*, 2010). Le mécanisme d'action de ces ARN a été mieux compris après l'étude structurale de la RNP RsmE (homologue de RsmA) fixée à un ARNm cible par RMN, chez *P. fluorescens* (Schubert *et al.*, 2007). Un site minimal de liaison sur l'ARN a ainsi été mis en évidence, qui est défini par la séquence A/UCANGGANGU/A. La protéine RsmE reconnaît spécifiquement la séquence GGA accessible localisée dans une boucle apicale d'une structure en tige-boucle (Figure 19). Les sARN, qui séquestrent des protéines présentent la même signature, que celle trouvée dans les ARNm ciblés par ces facteurs. Alors que ce motif est présent en une ou deux copies sur l'ARNm, RsmZ/Y possèdent 5 à 6 motifs. Ces sARN représentent un leurre pour RsmA conduisant à sa séquestration. Chez *E. coli*, CsrB (homologue de RsmY et Z de *P. aeruginosa*) contient 22 séquences contenant le motif GGA et peut interagir avec 9 protéines CsrA (homologue de RsmA) (Babitzke *et al.*, 2007, Liu *et al.*, 1997). Un système de régulation similaire a été récemment découvert pour le régulateur global de la traduction Crc qui permet à la bactérie de choisir la meilleure source carbonée. Cette protéine, qui reconnaît un motif UAAUAAUAA présent dans les sites RBS des ARNm cibles, est séquestrée par un ARN CrcZ contenant 6 motifs identiques dans des conditions moins favorables de croissance (Sonntleiner *et al.*, 2009 ; Moreno *et al.*, 2012). Ces exemples illustrent une évolution convergente de deux ARN pour réguler l'activité d'une protéine fixant l'ARN.



**Figure 19 : Régulation de l'activité de RsmA par les sARN RsmY et RsmZ d'après Bordi *et al.*, 2010.**

**A.** Chez *P. aeruginosa*, le système à deux composants GacS (en mauve)/GacA (en rose) active la synthèse de RsmY et Z. Ces ARN séquestrent la protéine RsmA via des motifs CCA en vert. Le système est modulé par le senseur RetS (en vert), et RsmY par HptB (en bleu). **B.** Structure de l'interaction de RsmE (dimère bleu et vert) avec l'ARN *hcnA*, résolue par RMN (Schubert *et al.*, 2007), PDB : 2JPP.

Chez *E. coli*, l'ARN6S est un ARNnc stable qui s'accumule en phase stationnaire de croissance. Son rôle est de diminuer la transcription de certains gènes par le facteur  $\sigma^{70}$  durant cette phase. Sa structure secondaire mime une bulle de transcription qui recrute l'ARNP lorsqu'elle est liée au facteur général de la transcription  $\sigma^{70}$  (Wassarman *et al.*, 2000, Trotaud *et al.*, 2005). Pour se dissocier de ce site, l'holoenzyme ARNP/ $\sigma^{70}$  synthétise un ARN dit ARN produit (pARN), en se servant de l'ARN6S comme matrice. Cette synthèse est indispensable à la bactérie pour quitter la phase stationnaire lorsque les conditions du milieu sont favorables, i.e. lorsque la concentration en NTP augmente. En effet, chez des mutants de l'ARN6S capables de lier l'holoenzyme tout en empêchant la synthèse des pARN, on observe une diminution de la viabilité cellulaire. Deux copies de l'ARN 6S, 6S-1 et 6S-2, sont trouvées chez *B. subtilis*. De manière intéressante, leurs fonctions sont indépendantes puisque 6S-2 ne permet pas la synthèse de pARN (Cavanagh *et al.*, 2012).

### III.C. Machineries protéiques associées à la régulation par les sARN

Les protéines associées aux sARN exercent de nombreuses fonctions en stabilisant des sARN, en facilitant l'interaction sARN-ARNm, et en modifiant les structures secondaires des ARN.

#### III.C.1. La protéine Hfq

La protéine Hfq (*Host factor Q $\beta$* ), homologue aux protéines Sm chez les eucaryotes, joue un rôle important dans la régulation par les sARN chez de nombreuses bactéries (revue Vogel *et al.*, 2011). Sous sa forme fonctionnelle homohexamérique, Hfq facilite l'interaction ARNm/sARN en favorisant les repliements intra- et inter-moléculaires, et ainsi la formation d'un duplex stable (Aiba *et al.*, 2007, Hopkins *et al.*, 2009). En recrutant la RNase E, Hfq stimule la dégradation des ARNm ciblés aux sARN (Morita *et al.*, 2005 ; Masse *et al.*, 2003). A l'inverse, la liaison de Hfq peut protéger les ARN contre la dégradation. Par exemple, le site de fixation de Hfq à RyhB coïncide avec le site reconnu par la RNase E (Moll *et al.*, 2003). La bibliographie attenante à Hfq est impossible à résumer en quelques paragraphes. C'est pourquoi je me suis concentré sur quelques études qui me semblaient apporter une vue assez précise du fonctionnement de Hfq.

**Trois faces pour trois sites de fixation des ARN.** D'un point de vue structural, Hfq forme un homohexamère dont la structure quaternaire ressemble à un anneau (Sauter *et al.*, 2003). L'une des faces de l'hexamère, appelée face proximale, lie les séquences riches en U retrouvées chez les sARN possédant à leur extrémité 5' un terminateur de transcription Rho-indépendant (Sauer *et al.*, 2011). L'autre face, dite distale, lie les ARN présentant les motifs AAYAAYAA ou ARNN (ou Y représente les pyrimidines C et U et R, les purines G et A, N pour tout nucléotide). Cependant, il est important de souligner que les faces distales et proximales ne se lient pas préférentiellement à une classe d'ARN (Vogel *et al.*, 2011). Les protéines Hfq de  $\gamma$ - et  $\beta$ -*proteobacteria* possèdent une longue extension C-terminale, à l'inverse de celles synthétisées dans les bactéries Gram positif comme *S. aureus* ou *B. subtilis*. Chez *E. coli*, le domaine C-terminal présente de nombreuses charges positives utiles pour la liaison de long ARN (Vecereck *et al.*, 2008 ; Beich-Frandsen *et al.*, 2011). Les données

cristallographiques disponibles montrent que la structure du cœur N-terminal de l'enzyme est conservée malgré des potentiels électrostatiques différents entre les bactéries Gram positif et Gram négatif (Brennan *et al.*, 2007) (Figure 20).

Récemment il a été suggéré que la surface latérale de Hfq représente un troisième site indépendant de fixation au sARN. Un modèle de l'interaction Hfq de *S. typhimurium*/sARN a été proposé dans lequel l'extrémité 3' du sARN se fixerait sur le site proximal de Hfq alors que le cœur du sARN et la séquence de fixation à l'ARNm seraient reconnus par la surface latérale. Après fixation de l'ARNm, le sARN serait relargué par un changement de conformation de Hfq (Sauer *et al.*, 2012) (Figure 20).

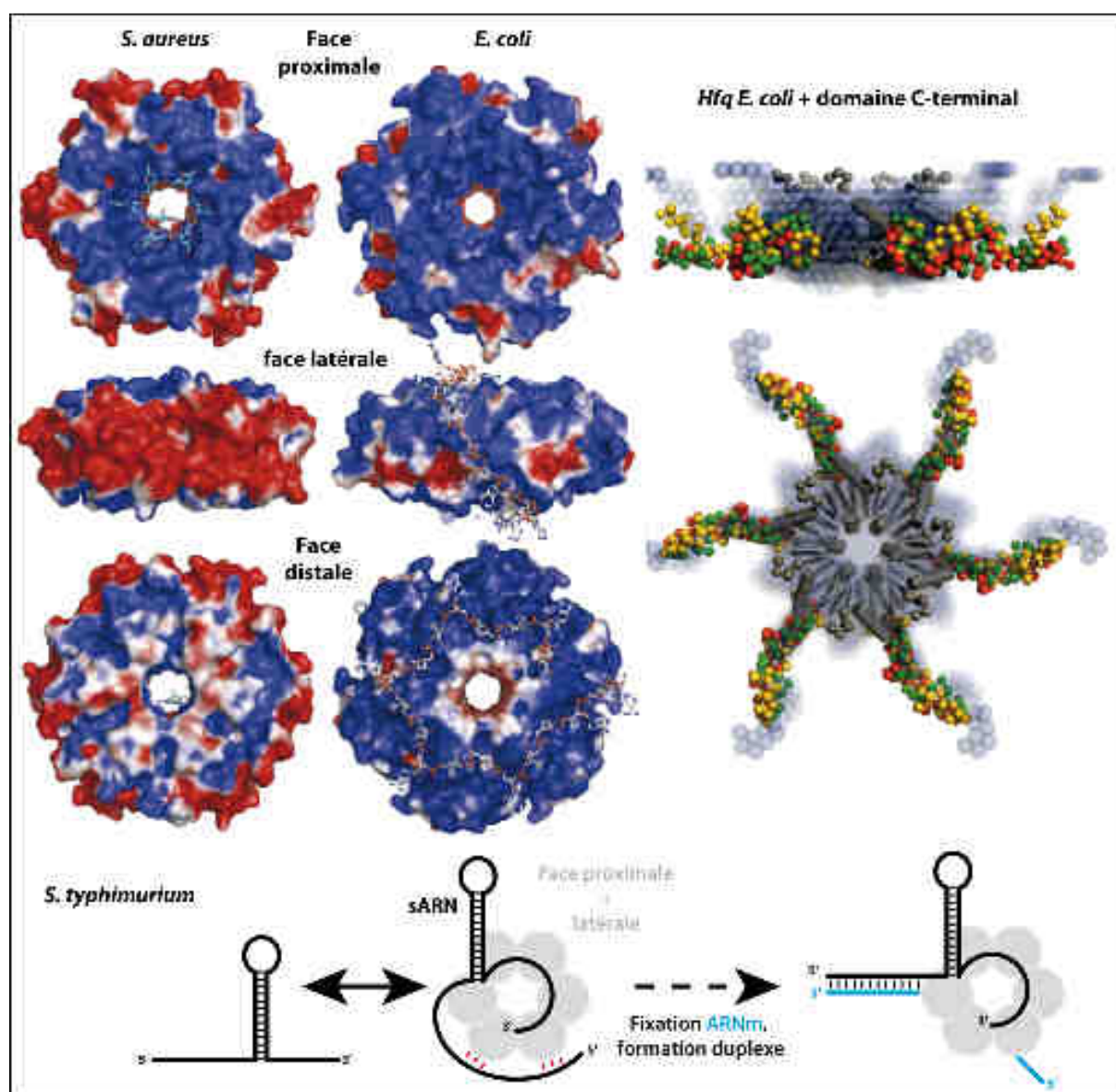
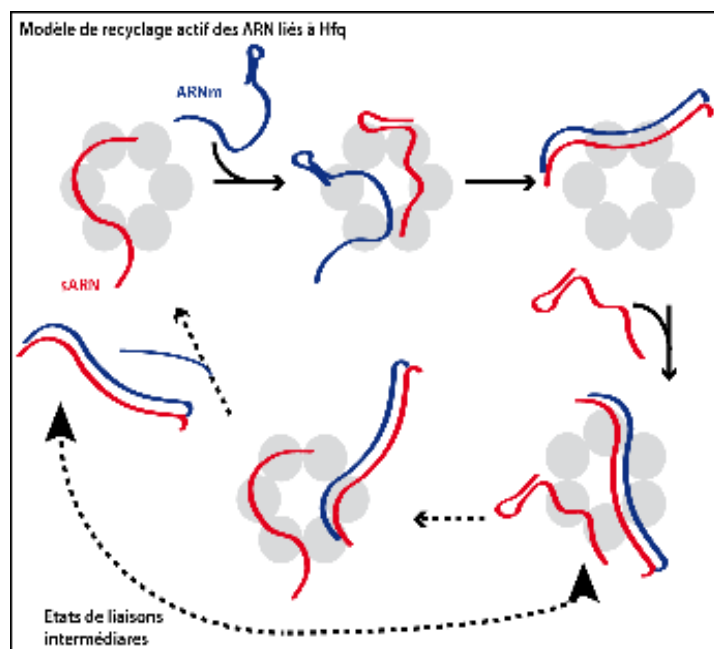


Figure 20 : Structures 3D des protéines Hfq d'*E. coli* (*Hfq<sub>ec</sub>*) et *S. aureus* (*Hfq<sub>sa</sub>*) et modèle d'interaction sur la face latérale chez *S. typhimurium*.



Gauche : surfaces électrostatiques potentielles et sites de liaison des ARN d'après la revue de Brenan *et al.*, 2007. La structure de Hfq<sub>sa</sub> montre la liaison d'un ARN AU<sub>5</sub>G sur la face proximale, celle de Hfq<sub>ec</sub> montre les sites de liaison d'ARN polyA (A<sub>27</sub> pour la vue de côté et A<sub>18</sub> pour la vue de la face distale). Les régions bleues sont électropositives, les zones rouges électro négatives. Droite : modélisation de la partie C-terminale de Hfq<sub>ec</sub> avec au centre la structure cristalline de la région N-terminale en gris (Beich-Frandsen *et al.*, 2011). En bas, modèle d'interaction entre un sARN et la protéine Hfq de *S. typhimurium*. Ici, la face latérale de l'hexamère fixe certaines régions du sARN, tandis que le terminateur se fixe via la face proximale. Après venue de l'ARNm (en bleu) sur la face distale, l'hybride ARNm/sARN se forme (Sauer *et al.*, 2012).

**Compétition des ARN pour la fixation à Hfq : recyclage dynamique des ARN.** Hfq est une protéine abondante chez *E. coli*, i.e. 400 et 1000 hexamères de Hfq/cellule. Néanmoins, le nombre de molécules d'ARNm et de sARN capables de se lier à Hfq serait largement excédentaire. De plus, la liaison de Hfq à l'ARN est stable et incompatible avec les temps de réponse au stress. Quels sont les effets d'Hfq sur la vitesse de formation des hybrides



ARNm/sARN et comment sont-ils déplacés de l'anneau? En général, Hfq augmente la vitesse d'association entre les deux ARN.

**Figure 21 : Recyclage dynamique des ARN liés à Hfq d'après Fender *et al.*, 2010.**

La liaison des ARN à Hfq favorise la formation d'un duplex sARN/ARNm cible. Le relargage du duplex se fait de façon dynamique par des cycles très rapides d'associations/dissociations d'autres sARN *en trans* au niveau de l'hexamère, qui à terme vont remplacer le duplex. sARN en rouge, ARNm en bleu, hexamère Hfq en gris.

Une étude récente a montré que la

formation du duplex est couplée à un recyclage actif des ARN par une série d'associations/dissociations au cours desquelles les sARN *en trans* déplacent graduellement les duplex stables liés à Hfq. Par exemple, dans le cas de l'interaction *MicC/ompC* chez *E. coli*, Hfq favorise la stabilisation des duplex préformés et, les molécules d'ARN *MicC* libres *en trans* déplacent activement le duplex en se liant à Hfq (Fender *et al.*, 2010) (Figure 21).

**Divergence entre bactéries Gram positif et négatif.** Le rôle d'Hfq est prépondérant chez les bactéries Gram négatif. Si la délétion du gène codant n'affecte pas la viabilité, son

impact sur la croissance et l'adaptation aux stress est considérable chez *E. coli*. Chez les bactéries Gram positif, l'importance d'Hfq soulève d'avantage de questions. En effet, l'étude comparative de plusieurs génomes montre qu'une seule espèce connue de  $\gamma$ -protéobactéries ne possède pas Hfq, alors que la proportion de bactéries n'ayant pas Hfq est plus grande chez les bactéries Gram positif (Chao *et al.*, 2010). Cependant, Hfq est requise pour l'expression des gènes de virulence chez *L. monocytogenes* (Nielsen *et al.*, 2010), alors que la délétion du gène codant Hfq dans plusieurs souches de *S. aureus* n'a pas d'influence sur la croissance, l'adaptation au stress et la virulence (Bohn *et al.*, 2007). La régulation des gènes cibles de l'ARNIII chez *S. aureus* (Boisset *et al.*, 2005) et CodY chez *B. subtilis* (Preis *et al.*, 2009) ne requiert pas la présence de Hfq. L'expression en *trans* de la protéine Hfq de *S. aureus* ne complémente pas des mutants *Hfq* d'*E. coli* (Vecerek *et al.*, 2007) et de *S. typhimurium* (Rochat *et al.*, 2012). De manière contradictoire, la délétion de Hfq dans la souche NCTC-8325 de *S. aureus* semble avoir un impact sur la pigmentation et la pathogénie de la bactérie et sur la virulence, cependant un lien direct entre la délétion de Hfq et la régulation par les ARN n'a pas été démontré (Liu *et al.* 2005). Pourquoi ne pas supposer que les différences de charge observées entre Hfq d'*E. coli* et de *S. aureus* pourraient résulter en des propriétés de fixation différente des deux protéines. Ainsi les protéines auraient peut-être divergées au cours de l'évolution pour acquérir ou perdre de nouvelles fonctions. Le cas de *S. aureus* est particulièrement intéressant, puisque la face distale, importante dans la fixation des ARNm, et la latérale, importante dans la fixation des sARN ont accumulé un potentiel électrostatique négatif fort au cours de l'évolution. Quel type d'ARN Hfq lie t'elle et comment assure-t-elle sa fonction biologique chez *S. aureus* ?

**De part son mode d'action, Hfq est douée de multiples activités.** Chez *E. coli*, Hfq est capable de fixer le facteur de terminaison Rho pour favoriser le processus d'anti-termination. Le complexe Rho-Hfq est plus stable en présence d'ARN et empêche l'hydrolyse de l'ATP par Rho (Rabhi *et al.*, 2011). Par ailleurs, le sARN Spot 42 recrute Hfq pour se lier à une séquence riche en A/U de son ARNm cible *sdhC* proche du RBS. Spot 42 n'inhibe pas directement la traduction car son site d'interaction est trop éloigné du RBS de l'ARNm *sdhC*. Néanmoins, l'encombrement stérique dû à la fixation de Hfq entraîne une compétition directe avec la sous-unité 30S, inhibant ainsi la traduction de l'ARNm cible (Desnoyer *et al.*, 2012). Hfq agit en tant que régulateur et le sARN comme facteur accessoire.

L'activité de Hfq peut être aussi modulée par diverses protéines. Chez *E. coli*, la protéine RelA stimule l'activité de RyhB en agissant sur Hfq. Cette protéine est responsable du ralentissement de la croissance de la bactérie sur un milieu pauvre en fer, et permet la synthèse de la molécule signal ppGpp. Ce ralentissement est important pour la bactérie, puisqu'il permet un redémarrage très rapide lorsque le milieu est complété en fer, à l'inverse de ce qui est observé pour une souche délétée du gène *relA*. RelA favorise l'oligomérisation de Hfq sous sa forme active hexamérique. Ainsi, RyhB est plus facilement reconnu par Hfq en présence de RelA, et régule plus efficacement la traduction de ses ARNm cibles (Argaman *et al.*, 2012). Toujours chez *E. coli*, l'activation de la traduction de l'ARNm *rpoS* par DsrA requiert Hfq, mais également l'action d'une ARN hélicase, CsdA, qui va libérer le site d'interaction de DsrA sur l'ARNm. La liaison du sARN empêche la formation d'une structure inhibitrice piégeant la SD, activant la traduction de l'ARN à basse température. Cependant, il n'y a pas d'interaction physique entre CsdA et Hfq (Resch *et al.*, 2010).

Hfq est capable de se lier à des enzymes impliqués dans la dégradation des ARN (PNPase, RNase E, PAP I), à la réplicase du phage Q $\beta$  et à l'ADN (revue : Vogel *et al.*, 2011).

### III.C.2. Autres machineries protéiques associées aux sARN

Existe-t-il une machinerie protéique, autre que Hfq, capable de contribuer à la régulation par les ARN ?

Une étude menée chez *Sinorhizobium melitoli* montre que la protéine SMC011113 (orthologue à la protéine YbeJ chez *E. coli*), une métalloprotéine ubiquitaire et conservée chez presque toutes les bactéries, a un impact sur la régulation par les sARN, de façon similaire à Hfq. Néanmoins, son implication et sa liaison directe aux sARN restent à démontrer. Du point de vue structural, SMC011113 partage une étonnante homologie avec le domaine MID des protéines AGO présentes chez les eucaryotes (Pandey *et al.*, 2011), qui sont impliquées dans la régulation par les miRNA et siRNA.

**Les travaux effectués pendant ma thèse ont porté sur une bactérie commensale et pathogène opportuniste, *S. aureus*. Quel est le répertoire des facteurs de virulence pour l'établissement d'un foyer infectieux, et comment sont-ils régulés ? En quoi cette bactérie constitue-t-elle un modèle idéal de l'étude du contrôle transcriptionnel et post-transcriptionnel par les sARN?**



toxique pour lequel il n'existe pas encore de traitement (Lappin *et al.*, 2009, Wertheim *et al.*, 2005) (Figure 22 C). En milieu hospitalier, les infections par *S. aureus* se produisent lors d'une rupture cutanée (coupures, abrasions, brûlures), et le plus souvent chez les patients immunodéprimés atteints de pathologies graves, comme le cancer, le SIDA ou certaines maladies génétiques. Une forte association a été caractérisée entre le développement du carcinome de la peau SCC et la colonisation des tissus par *S. aureus* (Kullander *et al.*, 2009). Cependant, on ignore encore si *S. aureus* pourrait être à l'origine du cancer ou si la pathologie SCC présente une susceptibilité accrue vis-à-vis de la colonisation par *S. aureus*. 30% des infections nosocomiales sont dues à cette bactérie, et cette proportion peut augmenter selon le type d'infection. La moitié des isolats à l'origine d'infections du sang dans certains pays d'Europe (Grande Bretagne, Irlande, France, Espagne, Portugal, Italie, Grèce) sont des souches de *Staphylococcus aureus* (Woodford *et al.*, 2009). L'impact de cet organisme en matière de santé publique est considérable. Aux États-Unis, environ 20 000 personnes meurent chaque année suite à une infection par *S. aureus*, et 500 000 personnes sont infectées. En France, 220 000 cas d'infections par des SARM (*S. aureus* résistants à la méthicilline) ont été répertoriés (Coignard *et al.*, Institut de veille sanitaire 2010). L'impact sur le plan financier est également important, les infections nosocomiales augmentant la durée moyenne du séjour des patients en milieu hospitalier. En 2001, le coût des infections dues aux SARM était estimé entre 42 et 59 millions de dollars dans les hôpitaux canadiens (Kim *et al.*, 2001). L'émergence et la dissémination rapide de souches multi-résistantes aux antibiotiques en font aujourd'hui un problème majeur en santé publique (Pantosti *et al.*, 2009) (Figure 23). L'évolution de la sensibilité à la vancomycine de certaines souches de SARM est inquiétante : si en 1999, seules 10 % des souches isolées avaient une CMI vancomycine supérieure à 1 mg/L, en 2008, cette proportion atteignait 40 % (Dumitrescu *et al.*, 2010).

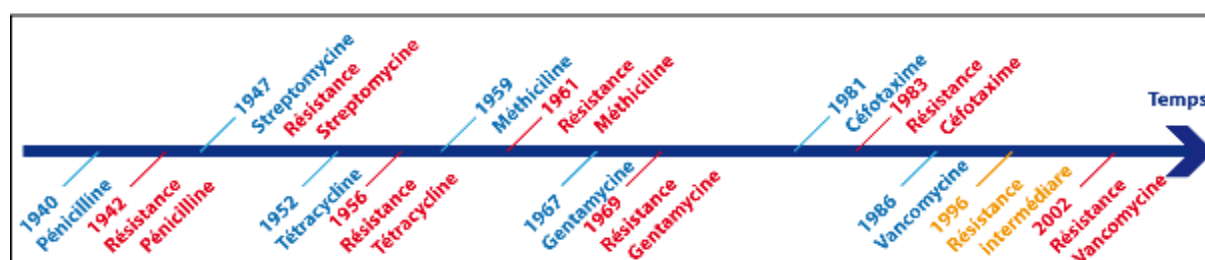


Figure 23 : Evolution de la résistance aux antibiotiques de *S. aureus*.

En bleu, année d'approbation par la FDA (Federal Drug Agency) de l'usage de l'antibiotique. En rouge, année d'apparition d'une souche de *S. aureus* résistante à l'antibiotique, en orange, résistance intermédiaire (Rammelkamp *et al.*, 1943, Jevons *et al.*, 1963, Lacey *et al.*, 1969, Deguchi *et al.*, 1982, Hiramatsu *et al.*, 1997, CDC. 2002).

Une des raisons de l'émergence si rapide de résistances est l'utilisation massive d'antibiotiques, le plus souvent à des doses inférieures à la CMI. Dans plusieurs pays industrialisés, des mesures d'hygiène et de contrôle dans l'utilisation des antibiotiques ont été élaborées pour tenter d'enrayer ces phénomènes de résistance. Ces mesures sont importantes puisqu'elles ont permis de limiter le taux de mortalité aux SARM dans certains hôpitaux (Ibrahim *et al.*, 2000). En France, le plan antibiotique du Haut Conseil de la Santé Publique commence à porter ses fruits, mais les infections à SARM restent élevées (Colloque SARM 2010, Institut de veille sanitaire). L'une des raisons de la persistance de *S. aureus* en milieu hospitalier est la capacité de la bactérie à former des structures adhésives multicellulaires appelées biofilms. En effet, les biofilms peuvent se former sur les cathéters intraveineux, les tubes endo-trachéaux, les prothèses ou les *pacemakers* (e.g. Lopez *et al.*, 2010) (**Figure 22 D et E**). Les biofilms offrent une protection efficace contre les mécanismes de défense de l'hôte et les antibiotiques. De plus, les bactéries présentes au sein du biofilm supportent d'importantes variations de leurs conditions de survie (carence en nutriments, stress, pH, oxygène) générant une concentration locale de bactéries capable de s'adapter rapidement aux environnements divers rencontrés chez le patient (Stoodley *et al.*, 2009). Un autre aspect inquiétant est l'apparition de souches résistantes aux antibiotiques au sein de la communauté, personnes sans passif hospitalier.

#### **IV.B. Facteurs de virulence**

Les facteurs de virulence de *S. aureus* sont à l'origine de l'établissement de l'infection chez l'homme. Ces facteurs exercent des fonctions multiples, ils contribuent à la neutralisation du système immunitaire de l'hôte, à l'adhésion de la matrice cellulaire, à la dissémination et à l'acquisition de nouveaux nutriments (**Tableau 1**) (Foster 2005 ; Gordon *et al.*, 2008).

**Les facteurs d'adhésion.** Les facteurs d'adhésion, qui sont généralement exprimés durant les phases précoces d'infection, se répartissent en deux catégories. La première, les

MSCRAMM (*Microbial surface components recognizing molecules*), regroupent des facteurs associés à la paroi bactérienne comme les "*Fibronectin Binding Protein*" (FnBP) ou les "*clumping factors*" (ClfA/ClfB). Les protéines FnBPA et B (Jönsson *et al.*, 1991) se lient spécifiquement à la fibronectine qui est une protéine clé dans l'adhésion cellulaire et dans l'organisation de la matrice extracellulaire. Ces facteurs d'adhésion interviennent dans les phases précoces de l'infection et seront dégradés ensuite par la protéase V8 (McGavin *et al.*, 1997). Les FnBP sont importantes pour l'adhésion de la bactérie aux tissus de l'hôte (Grundmeier *et al.*, 2004), l'internalisation de *S. aureus* dans les cellules endothéliales (Jett *et al.*, 2002) et contribuent à l'initialisation des infections sur corps étrangers. Les *Clumping factors* A et B, synthétisés durant la phase stationnaire de croissance, protègent la bactérie contre la phagocytose chez la souris (Palmqvist *et al.*, 2004) en augmentant l'adhésion à la matrice extracellulaire. La deuxième grande classe de facteurs d'adhésion comprend les protéines SERAM (*Secretable Expanded Repertoire Adhesive Molecules*). La protéine Efb (*Extracellular FnBP*) cible la protéine du complément C3, empêchant l'opsonisation du pathogène médiée par le fragment C<sub>3b</sub> (Lee *et al.*, 2004). La protéine Eap (*extracellular adherence protein*) est quant à elle impliquée dans la réduction de la migration des leucocytes vers le site d'infection (Chavakis *et al.*, 2002). Ainsi, les adhésines permettent à *S. aureus* de fixer des molécules plasmatiques (fibrinogène, fibronectine) ou tissulaires (collagène) et jouent un rôle primordial dans la colonisation des tissus.

**Facteurs de virulence sécrétés.** D'autres facteurs de virulence permettent également à la bactérie d'échapper au système immunitaire. Les CHIPS (*Chemotaxis inhibitory protein of staphylococci*) inhibent les récepteurs du facteur du complément C<sub>5a</sub>. Ce système de transduction de signal est responsable de l'activation des neutrophiles et de leur recrutement au site d'infection. Les CHIPS ciblent également les récepteurs des peptides formylés (Wright *et al.*, 2007). Ces peptides sécrétés par la bactérie sont reconnus par les neutrophiles et conduisent à la phagocytose. La protéine A est une protéine ancrée à la paroi de *S. aureus*. Elle possède des rôles multiples. Elle se lie aux IgG *via* leur fragment constant empêchant l'activation des cellules du système immunitaire (Foster *et al.*, 2005). La protéine A cible également les régions variables des IgM situés sur les lymphocytes B (LB), entraînant une prolifération lymphocytaire anarchique sans sécrétion d'anticorps dirigés contre la bactérie (Bekeredjian-Ding *et al.*, 2007). Cette prolifération anarchique mène à

l'anergie (inactivation) des LB, ou bien à l'apoptose selon la concentration en protéine A (Silverman *et al.*, 2006). Les superantigènes modulent l'action du système immunitaire de manière similaire à la protéine A. Ils se lient (TSST-1, entérotoxines ABC ou D) aux récepteurs du CMH de classe II des lymphocytes T (LT), entraînant la prolifération incontrôlée de LT non spécialisés (Foster *et al.*, 2005). La suractivation des LT couplée à la libération massive de cytokines est à l'origine du syndrome foudroyant du choc toxique (Lappin *et al.*, 2009).

Le dernier type de facteurs de virulence regroupe les exotoxines et leucotoxines. Leur rôle consiste à créer des pores dans la membrane des neutrophiles et/ou des érythrocytes. Par exemple, l'hémolysine  $\alpha$  forme un homoheptamère dans la membrane de nombreuses cellules (Song *et al.*, 1996 ; Bhakdi *et al.*, 1991). La leucocidine PVL (*Panton Valentin Leukocidin*) forme un homohexamère qui, à forte concentration, induit la nécrose des cellules par activation des canaux calciques ou, à faible concentration, génère des pores dans la membrane des mitochondries (Genestier *et al.*, 2005). Cette toxine est maintenant retrouvée dans la majorité des souches communautaires de SARM (Lo *et al.*, 2011). Une étude récente montre que la toxine PVL est le facteur majeur de virulence qui cause une activation de l'inflammation par IL-1 $\beta$  (Perret *et al.*, 2012). D'autres toxines (hémolysines  $\delta/\gamma/\beta$ , LuxDE) coopèrent avec la PVL pour initier l'inflammation.

### **IV.C. Régulation de l'expression des facteurs de virulence**

La régulation des gènes impliqués dans la virulence de *S. aureus* est assurée par un réseau complexe faisant intervenir des facteurs de transcription, des systèmes à deux composants, des facteurs  $\sigma$  alternatifs et des ARN régulateurs. Ces réseaux sont eux-mêmes finement régulés en fonction du temps et de la densité de la population et sont capables d'intégrer divers signaux provenant du milieu extracellulaire (Novick *et al.*, 2003).



	facteurs de virulence	Gène	Fonctions	Références
Super-antigènes	Entérotoxine A Entérotoxine B Entérotoxine C Entérotoxine D Entérotoxine G Entérotoxine H	<i>sea</i> <i>sab</i> <i>sec</i> <i>sed</i> <i>seg</i> <i>seh</i>	Suractivation des lymphocytes T <i>via</i> ciblage des récepteurs CMH (Complexe Majeur d'histocompatibilité) de type II	Foster <i>et al.</i> , 2005
	Toxic shock syndrome toxin-1	<i>TSST-1</i>	Idem, à l'origine du syndrome foudroyant du choc toxique	Lappin <i>et al.</i> , 2009
	Pro. d'adhérence extracellulaire	<i>Eap</i> <i>/Map</i>	Rôle dans la liaison aux protéines du plasma Modulation du système immunitaire ( <i>via</i> liaison au récepteur ICAM-1)	Harraghy <i>et al.</i> , 2003
	Exfoliatine A Exfoliatine B Exfoliatine D	<i>eta</i> <i>etb</i> <i>etd</i>	Suractivation des LT dans le cas de SSSS (Staphylococcal Scaled Skin Syndrom) Activité protéase au niveau des desmosomes des cellules épithéliales (perte d'adhésion)	Plano <i>et al.</i> , 2004
Hémolysines Leucocidines	Hémolysine $\alpha$	<i>hla</i>	Homo-heptamère, formation de pores chez les leucocytes	Menestrina <i>et al.</i> , 2003
	Hémolysine $\delta$ Hémolysine $\gamma$	<i>hld</i> <i>hlg</i>	Formation de pores chez les leucocytes /érythrocytes	
	Leucocidine AB Leucocidine ED	<i>lukAB</i> <i>lukED</i>	Formation de pores chez les leucocytes	Kaneko <i>et al.</i> , 2004 Lo <i>et al.</i> , 2011
	Panton Valentine leucocidine	<i>pvl</i>	Formation de pores chez les leucocytes, CA et HA-MRSA avec cette toxine	
Facteurs d'adhésion	Protéine A	<i>spa</i>	Inactivation du système immunitaire (liaison IgG/IgM) Induit sécrétion d'histamine par les basophiles	King <i>et al.</i> , 1981 Gomez <i>et al.</i> , 2004
	Collagène Binding protein	<i>cna</i>	Formation de pores chez les leucocytes /érythrocytes	Kaneko <i>et al.</i> , 2004 Lo <i>et al.</i> , 2011
	Clumping factor A Clumping factor B	<i>clfA</i> <i>clfB</i>	Liaison au fibrinogène, inhibition des opsonines	Foster <i>et al.</i> , 2005
	Fibrinogen binding protein A/B	<i>fnbpA</i> <i>fnbpB</i>	Adhésion, invasion de phages non professionnels	Menzies <i>et al.</i> , 2003
	Extracellular FnBP	<i>efb</i>	Bloque l'adhésion des neutrophiles, diapédèse	Ko <i>et al.</i> , 2011
	Coagulase	<i>coa</i>	Adhésion	Giraud <i>et al.</i> , 1997
	Bone sialoprotein binding protein	<i>sdr</i>	Rôle dans la liaison de <i>S. aureus</i> au niveau des ostéoblastes	Tung <i>et al.</i> , 2010
Enzymes	Staphylokinase	<i>ska</i>	Activation du plasminogène, neutralisation des défenses	Bokarewa <i>et al.</i> , 2006
	Chemotaxis inhibitory protein of staphylococci	<i>CHIPS</i>	Inactivation des récepteurs FP (formyle peptide) et des récepteurs C5a. modulation du système immunitaire	Wright <i>et al.</i> , 2007
	Protéines capsulaires 5 et 8 Protéase	<i>cps</i> <i>sspA</i> <i>sspB</i> <i>V8</i>	Résistance à la phagocytose, persistance sanguine	O'Riordan <i>et al.</i> , 2004
	Auréolysine	<i>aur</i>	Dégradation des Ig Inactivation des inhibiteurs de protéinase du plasma	Gustafsson <i>et al.</i> , 2008
	Lipase Glycerol ester hydrolase	<i>lip</i> <i>geh</i>	Métabolisme des lipides	Rosenstein 2000

Tableau 1 : Facteurs de virulence chez *S. aureus*.

### IV.C.1. Les facteurs $\sigma$

Les facteurs  $\sigma$  sont responsables de la liaison de l'ARNP au niveau de promoteurs spécifiques. Comme les autres bactéries, *S. aureus* possède un facteur général de la transcription :  $\sigma^A$ . En parallèle, un panel de trois autres facteurs  $\sigma$  :  $\sigma^B$ ,  $\sigma^H$  et  $\sigma^S$  ont été identifiés (**Tableau 2**) qui sont requis pour l'adaptation rapide au stress et l'expression de facteurs de virulence. Le facteur  $\sigma^B$  et son réseau de régulation, dans lequel se trouve l'ARN régulateur *RsaA*, seront détaillés dans la partie (**résultat**).

**Tableau 2 : Facteurs  $\sigma$  codés par *S. aureus*.**

Facteurs $\sigma$	Gène	Fonction(s) générale(s)	Références
$\sigma^A$	<i>sigA</i>	facteur général	Mittenhuber <i>et al.</i> , 2002
$\sigma^S$	<i>sigS</i>	réponse au stress, virulence	Shaw <i>et al.</i> , 2008
$\sigma^H$	<i>sigH</i>	excision/intégration des prophages	Morikawa <i>et al.</i> , 2003, Tao <i>et al.</i> , 2010
$\sigma^B$	<i>sigB</i>	Stress de l'enveloppe, carence, virulence	Wu <i>et al.</i> , 1996

### IV.C.2. Facteurs de transcription

L'expression des facteurs de virulence est sous le contrôle de TF appartenant à la famille Sar. SarA (*Staphylococcal accessory regulator A*) a été le premier régulateur de la transcription identifié. Il est exprimé constitutivement durant la croissance cellulaire grâce à trois promoteurs. Les promoteurs P2 et P3 sont sous la dépendance de  $\sigma^A$ , et P1 sous celle du facteur de stress  $\sigma^S$  (Bayer *et al.*, 1996). SarA se lie sous forme de dimère à des séquences riches en résidus A et T dans les régions promotrices de gènes. Néanmoins, aucune séquence consensus de liaison de SarA n'a encore été établie à ce jour (Novick *et al.*, 2003). L'expression de SarA est intimement liée à celle du système *agr*. En effet, l'expression de nombreux gènes est commune aux deux systèmes comme celle de l'hémolysine  $\alpha$  ou la répression de la protéine A. SarA active également l'expression de l'opéron *agr* (Cheung *et al.*, 1997, 2004 et 2008). Cependant, les mécanismes de régulation de SarA restent encore mal caractérisés. Des données *in vitro* montrent que son état de phosphorylation influence sa fixation aux promoteurs de gènes cibles (Didier *et al.*, 2010). Une autre hypothèse est que cette protéine agirait au niveau de la transcription de répresseurs pour activer l'expression de gènes de manière indirecte. Par exemple, SarA activerait l'expression de *hla* en réprimant la synthèse du répresseur SarT (Schmidt *et al.*, 2001). SarA pourrait également induire la

**Tableau 3 : Facteurs de transcription chez *S. aureus*.**

TF	Fonction	Références
SarA	Activateur /répresseur de la transcription Réseaux de régulation dépendants et indépendants de l'expression d' <i>agr</i>	Manna <i>et al.</i> , 1998 Chakraborti <i>et al.</i> , 2000 Schmidt <i>et al.</i> , 2001
SarR	Active <i>agr</i> et SarS Inhibe SarA	Cheung <i>et al.</i> , 2004 Mana <i>et al.</i> , 2006
SarS	Active l'expression des gènes de surface et de la protéine A Est réprimé par SarA/ <i>agr</i>	Tegmark <i>et al.</i> , 2000 Oscarsson <i>et al.</i> , 2006
SarT	Active SarS Répresseur de hla et d'autres exotoxines Est réprimé par SarA/ <i>agr</i>	Schmidt <i>et al.</i> , 2001 Cheung <i>et al.</i> , 2004
SarU	Active <i>agr</i> Est réprimé par SarT	Cheung <i>et al.</i> , 2004
SarV	Régulateur de l'autolyse Est réprimé par SarA/MgrA	Manna <i>et al.</i> , 2004
SarX	Active sa propre expression Inhibe <i>agr</i> Est activé par MgrA	Mana <i>et al.</i> , 2006
SarY	Fonction inconnue	
SarZ	Active hla, SsaA ( ?)et les gènes du biofilm Réprime SarS Est activé par MgrA	Kaito <i>et al.</i> , 2006 Ballal <i>et al.</i> , 2009 Tamber <i>et al.</i> , 2009
Rot	Répresseur des toxines Est réprimé par <i>agr</i> (via l'ARNIII)	Said-Salim <i>et al.</i> , 2003 Boisset <i>et al.</i> , 2007
MgrA	Active les gènes impliqués dans l'autolyse, <i>agr</i> et la capsule, réprime SarS, active MgrA	Ingavale <i>et al.</i> , 2003

transcription d'ARN régulateurs, mais aucune donnée expérimentale ne confirme cette hypothèse à ce jour. Des données récentes suggèrent aussi que SarA aurait une incidence sur la stabilité de nombreux ARNm bien qu'aucune évidence n'existe sur la possibilité que SarA puisse se fixer à l'ARN (Roberts et al 2006). Plusieurs autres facteurs Sar ont été identifiés à ce jour, principalement grâce à leur homologie de séquence avec SarA et leurs réseaux de régulation s'entrecroisent avec celui de SarA. Les principaux membres de la famille Sar sont présentés dans la **Tableau 3**.

### IV.C.3. Systèmes à deux composants

17 TCS différents sont prédits chez *S. aureus* N315 SARM (Kuroda *et al.*, 2001). Ces systèmes TCS sont les senseurs de divers stress (pH, température, anaérobiose, antibiotiques et peptides antimicrobiens) (**Tableau 4**) et sont impliqués dans la régulation de la synthèse de facteurs de virulence et l'adaptation de la bactérie. Un seul TCS est essentiel chez *S. aureus*, le système WalkR impliqué dans le métabolisme du peptidoglycane et la dégradation de la paroi (Dubrac *et al.*, 2007). Une délétion du système aboutit à un

épaississement de la paroi, perturbant la division cellulaire. Deux enzymes importantes dans la dégradation du peptidoglycane sont sous la dépendance de ce système : LytM (glycyl-glycyl endopeptidase) et SsaA (hydrolase à domaine CHAP) (Delaune *et al.*, 2011).

Parmi les TCS impliqués dans la régulation des facteurs de virulence, on retrouve :

Le TCS codé par l'opéron *saePQRS* (*S. aureus exoprotein expression*) est activé lors d'une variation de pH, mais aussi en fonction de la concentration en sels et en défensines (phagosomes) (Novick *et al.*, 2003 ; Geiger *et al.*, 2008). L'histidine kinase (HK) membranaire est codée par le gène *saeS* qui phosphoryle, et active le régulateur de réponse (RR) SaeR. La phosphorylation a lieu au niveau du domaine N-terminal, entraînant un changement de conformation libérant le DBD (*DNA Binding Domain*) situé dans la région C-terminale. Ce TCS contrôle d'une manière générale l'expression de nombreuses exotoxines (Rogash *et al.*, 2006) et intervient également dans la régulation des processus d'adhérence (Liang *et al.*, 2006). Plusieurs cibles de ce TCS sont communes au système de densité cellulaire *agr* qui sera décrit plus loin.

Le TCS *srrAB* (Staphylococcal respiratory response) est impliqué dans l'adaptation de *S. aureus* à un milieu anaérobie (Yarwood *et al.*, 2001). Il est homologue au TCS *resDE* exprimé chez *B. subtilis* (Nakano *et al.*, 1996). Le RR du système est SrrA, l'HK est srrB. L'induction de SrrA inhibe l'expression des facteurs de virulence du système *agr* (exotoxines) (Pragman *et al.*, 2004), et active l'expression des gènes d'adhésion *ica* et *pia*. L'expression de ces adhésines favorise l'ancrage de la bactérie au niveau de la matrice extracellulaire, augmentant la résistance contre la phagocytose par les macrophages (Ulrich *et al.*, 2007).

Le système *arlRS* (*autolysis regulated locus*) est, quant à lui, impliqué dans le remodelage de la membrane bactérienne et la virulence. Un mutant *arlSR* montre une autolyse plus importante accompagnée de défaillances dans l'hydrolase de peptidoglycane (Fournier *et al.*, 2000). Encore une fois, l'expression de ArlR est liée au système *agr*, puisque ArgA est surexprimé en présence de ce dernier (Liang *et al.*, 2005). Parmi les facteurs de virulence induits par ce TCS, on retrouve les protéines liant la sialoprotéine osseuse. Ces protéines pourraient être importantes dans la localisation de *S. aureus* au niveau des tissus osseux et dans les pathologies associées (Tung *et al.*, 2000).

Tableau 4 : Systèmes à deux composants caractérisés chez *S. aureus*.

TCS	HK	RR	Fonction	Références
<i>agrBCDA</i>	AgrC	AgrA	Quorum sensing, régulation des gènes d'exotoxines et d'adhésion	Novick <i>et al.</i> , 1993
<i>saePQRS</i>	SaeS	SaeR	Sensible aux variations de pH et à la concentration en sels, régulation des gènes d'exotoxines et d'adhésion	Jeong <i>et al.</i> , 2011 Geiger <i>et al.</i> , 2008
<i>arlRS</i>	ArlS	ArlR	Remodelage de la membrane, autolyse. Induction de la sialoprotéine osseuse	Fournier <i>et al.</i> , 2000 Tung <i>et al.</i> , 2000
<i>srrAB</i>	SrrB	SrrA	Homologue à <i>resDE</i> chez <i>B. subtilis</i> , adaptation anaérobie Activation des adhésines <i>ica</i> et <i>pia</i> , inhibition des gènes de virulence	Yarwood <i>et al.</i> , 2001 Pragman <i>et al.</i> , 2004
<i>lytRS</i>	LytS	LytR	Régulation de l'autolyse	Brunskill <i>et al.</i> , 1996
<i>vraSR</i>	VraS	VraR	Induit par des dommages au niveau de la paroi Résistance aux antibiotiques et synthèse de la paroi	Kuroda <i>al.</i> 2003 Yin <i>et al.</i> , 2006
<i>walkR</i>	Walk	WalR	Essentiel, conservé chez les Gram + à faible % en GC Métabolisme du peptidoglycane	Dubrac <i>al.</i> 2007 Delaune <i>et al.</i> , 2011 Türck <i>et al.</i> , 2012
<i>braSR</i>	BraS	BraR	Induit en réponse à la bacitracine et à la nisine : activation de BraD/BraE, impliqués dans la détection de l'antibiotique et de VraD/VraE impliqués dans la détoxification	Hiron <i>et al.</i> , 2011 Kolar 2011
<i>graSR</i>	GraS	GraR	Impliqué dans la résistance aux antimicrobiens cationiques	Herbert <i>et al.</i> , 2007 Farold <i>et al.</i> , 2012
<i>AirSR</i>	AirS	AirR	Impliqué dans la régulation de l'expression des gènes en réponse à un stress oxydatif (rôle d'un cluster 2Fe-2S) Rôle dans l'adaptation à l'anaérobiose	Sun <i>et al.</i> , 2012
<i>KdpDE</i>	KdpD	KdpE	Induit en réponse à des variations de concentration de K extracellulaire Régulateur des gènes de virulence	Xue <i>et al.</i> , 2011
<i>HssRS</i>	HssS	HssR	Induit en réponse à de fortes concentrations d'hèmes Rôle dans la modulation de l'expression des facteurs de virulence chez la souris	Stauff <i>et al.</i> 2007 Torres <i>et al.</i> , 2007

Enfin, le dernier TCS impliqué dans la régulation des gènes de virulence est codé par l'opéron *agr*, induit en fonction de la densité cellulaire. Ce dernier est détaillé en IV.D.2, ainsi que la fonction de ses effecteurs intracellulaires, AgrA et l'ARNIII.

#### IV.D. ARN régulateurs chez *S. aureus*

Dans cette sous-partie, un état des lieux de la régulation des gènes par les sARN chez *S. aureus* sera présenté. Les données et figures résumées sont reprises des revues auxquelles j'ai participé : **Current knowledge on regulatory RNAs and their machineries in *Staphylococcus aureus*** (Romilly *et al.*, 2012) et **A current overview of regulatory RNAs in *Staphylococcus aureus*** (Caldelari *et al.*, 2011).

### IV.D.1. Etat des lieux

Au début de ma thèse, peu d'ARN régulateurs avaient été caractérisés dans le génome de *S. aureus* : 12 sARN localisés dans les îlots de pathogénie (Pichon *et al.*, 2005), 10 dans les régions intergéniques (IGR) (Geissmann *et al.*, 2009) et le plus étudié, l'ARNIII. L'identification de ces ARN s'est faite par des prédictions bio-informatiques combinant plusieurs critères (recherche de motifs structurés, conservation des séquences IGR). Seul l'ARNIII, effecteur intracellulaire du système de densité cellulaire, a été caractérisé par des expériences de complémentation dans des souches mutantes pour *agr* (Novick *et al.*, 1993).

Aujourd'hui, *S. aureus* est un des organismes modèles pour l'étude de la régulation de l'expression des gènes par les ARN. En effet, plus de 250 sARN ont été découverts *via* l'utilisation d'approches bio-informatiques prédictives couplées à des tests d'expression (Livny *et al.*, 2008 ; Marchais *et al.*, 2009), de puces à ADN (Anderson *et al.*, 2006, Roberts *et al.*, 2006), de techniques de clonage et de séquençage d'ADNc (Abu-Quatouseh *et al.*, 2010), et de techniques de séquençage à haut débit 454 (Bohn *et al.*, 2010) et Illumina (Beaume *et al.*, 2010).

La plupart de ces ARN sont codés dans le chromosome bactérien, certains sont localisés dans les éléments mobiles tels que les îlots de pathogénie et les plasmides (revues : Felden *et al.*, 2011 ; Caldelari *et al.*, 2011). Deux grandes classes d'ARN sont retrouvées chez *S. aureus*. La première décrit l'ensemble des régions leaders conservées au niveau des ARNm regroupant les riboswitch et les T-box. 33 gènes (dont 7 opérons) sont prédits pour être régulés par 15 types de riboswitch, fixant de la SAM, TPP, FMN, Lys, Gly, Purine, preQ1 ou Glc-6P. En parallèle, l'expression de 31 gènes est sous le contrôle de 12 T-box. De manière similaire à ce qui est prédit chez d'autres bactéries Gram positives (voir III.B.1.a), ces régions régulatrices en 5'UTR régulent majoritairement la terminaison de la transcription lorsqu'elles fixent un métabolite ou un ARNt non aminoacylé (**A current overview of regulatory RNAs in *Staphylococcus aureus***, Caldelari *et al.*, 2011) (voir Figure ).

Des ARN antisens sont exprimés dans les plasmides, les transposons (revue Brantl *et al.* 2007, Wagner *et al.*, 2002), ou encore les systèmes toxine / @toxine de type I (revue, Fozo *et al.*, 2008). Enfin, l'expression de plus de 50 sARN a été validée à la fois dans le cœur du génome ou dans les îlots de pathogénie. Néanmoins, la fonction de peu d'entre eux a été

caractérisée à l'heure actuelle. Dans la suite de cette partie, je décrirai les sARN de *S. aureus* dont les fonctions ont été caractérisées. Deux sont codés pour le cœur du génome, l'ARNIII et *RsaE* ; les deux autres sont exprimés au sein de îlots de pathogénie, *SprD* et *SprA1-AS*.

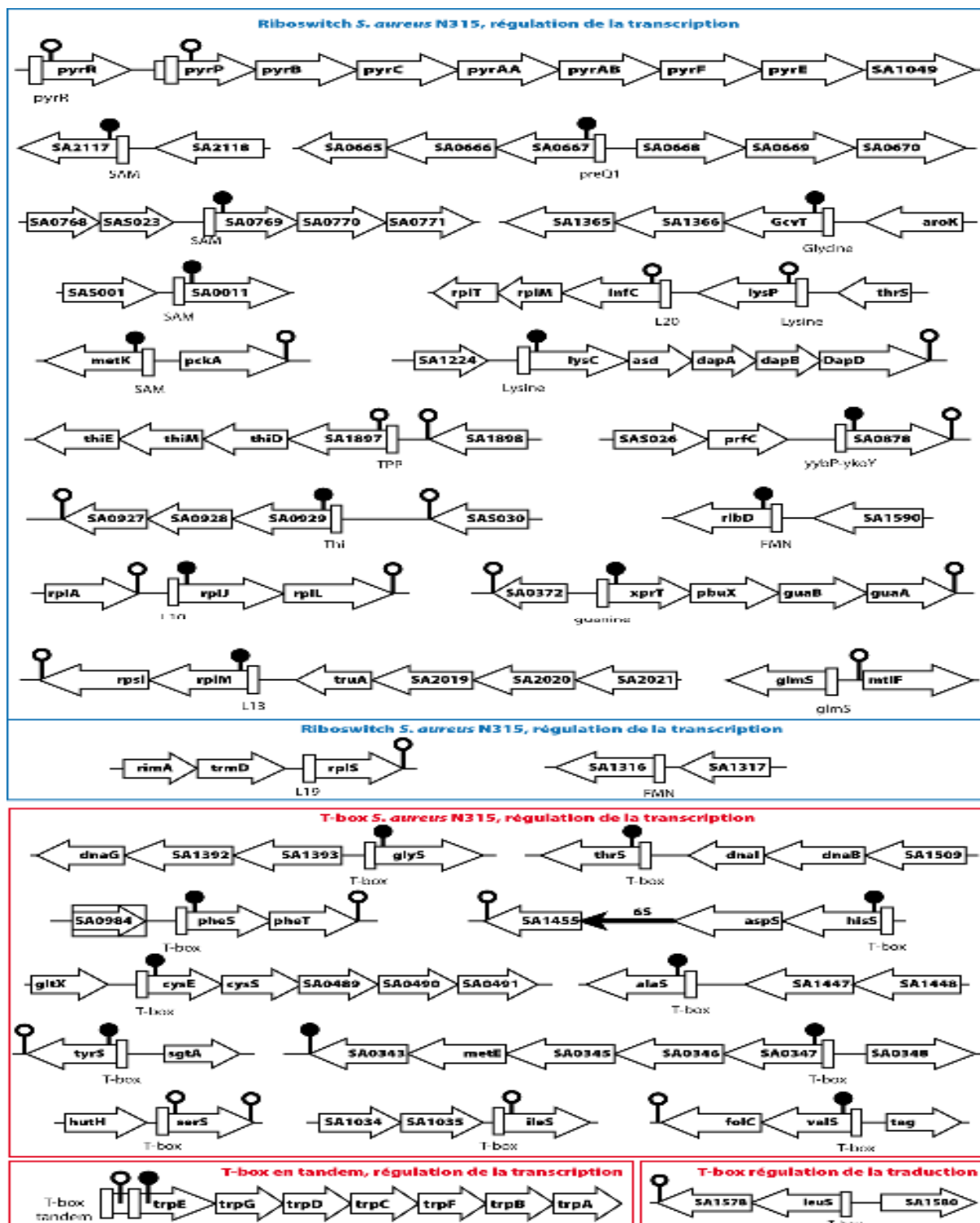


Figure 24 : Prédictions des riboswitch et T-box dans les régions leaders des ARNm de *S. aureus* N315 (d'après Caldelari et al., 2010).

Les riboswitch (cadres bleu) fixent spécifiquement un métabolite aboutissant à un terminaison de la transcription. Les riboswitch fixant la protéine ribosomique L19 ou la FMN, devant les gènes *rplS* et *sa1316* respectivement, régulent quant à eux l'initiation de la traduction. Les T-box (cadres rouge) entraînent également une terminaison de la transcription dans la plupart des cas. Seul la T-box régulant l'expression de l'ARNm *leuS* est prédite comme agissant au niveau traductionnel. Les éléments régulateurs sont représentés par un rectangle blanc, les terminateurs de transcription sont schématisés par une tige-boucle. Les terminateurs en noirs représentent des éléments en *cis* dans les régions 5'UTR ressorties par l'approche de séquençage à haut débit de Beaume et al., 2010.

#### IV.D.2. L'ARNIII, effecteur intracellulaire du système de densité cellulaire *agr*

Le système de densité cellulaire est à la base d'une communication cellule-cellule dépendant de la densité cellulaire de *S. aureus*. En produisant, sécrétant et détectant des molécules-sigaux, les bactéries coordonnent l'expression des gènes de virulence, impliqués dans la formation de biofilm et capsule, et du métabolisme lorsque la densité cellulaire devient trop élevée. Chez *S. aureus*, l'adaptation à haute densité cellulaire se fait *via* le système *agr*, dont les deux effecteurs intracellulaires sont AgrA et un ARN régulateur multifonctionnel, l'ARNIII (Figure 24). A haute densité cellulaire, la transcription de l'opéron *agrBDCA* est activée. Le pro-peptide codé par *agrD* est sécrété *via* la protéine membranaire AgrB et mûré en peptide auto inducteur AIP. Au delà d'une concentration seuil, ce dernier est reconnu par l'HK AgrC. L'autophosphorylation de AgrC active AgrA qui active alors sa propre transcription à partir du promoteur P2, et également induit l'expression de l'ARNIII à partir du promoteur P3. AgrA active également l'expression de gènes codant pour des peptides PSM (*Phenol soluble modulines*), et réprime l'expression de gènes du métabolisme et la synthèse de staphyloxanthine par un mécanisme encore inconnu (Queck *et al.*, 2008). Le second effecteur, l'ARNIII code un petit peptide de 26 aa, l'hémolysine  $\delta$ , et agit également comme régulateur par ses régions non codantes. L'ARNIII possède plusieurs domaines fonctionnels. Le domaine 5' est nécessaire à l'activation de la traduction de l'ARNm *hla* (voir introduction III.B.2.b) Les domaines central et 3' sont caractérisés par la présence de trois tiges-boucles possédant dans leur boucle apicale une séquence conservée et redondante UCCC. Ce motif constitue un site d'interaction avec le RBS des ARNm cibles codant pour des protéines impliquées dans les phases précoces de l'infection, pour réprimer leur traduction. Les facteurs d'adhésion contrôlés par l'ARNIII sont la coagulase, la protéine A et la protéine se liant au fibrinogène SA1000. L'ARNIII cible également l'ARNm *rot* codant



pour le répresseur de toxine Rot ainsi que l'enzyme LytM impliqué dans le métabolisme du peptidoglycane (Boisset *et al.*, 2007 ; Geisinger *et al.*, 2006). L'ARNIII inhibe la traduction en ciblant le RBS des ARNm sans l'aide de Hfq, et le duplexe formé est rapidement dégradé par la RNase III. Les structures formées par l'interaction ARNIII/ARN cible montrent une étonnante variété de topologies. Dans le cas de l'ARNm *coa*, l'ARNIII forme un duplex étendu au niveau du RBS, qui est stabilisé par une interaction boucle-boucle ciblant la région codante de l'ARNm. Pour l'ARNm *rot*, deux interactions boucle-boucle impliquant 6 nucléotides chacune sont suffisantes pour inhiber la traduction et induire la dégradation par la RNase III (Figure 24). Ainsi, l'expression de l'ARNIII confère à la bactérie d'activer l'expression des gènes codant pour les exotoxines au détriment de ceux codant pour les protéines d'adhésion. Cette stratégie favoriserait ainsi la dissémination de la bactérie.

Il semblerait que l'opéron *agr* ait évolué en deux réseaux de régulation, l'un affectant le métabolisme et les peptides PSM est indépendant de l'ARNIII, l'autre dépendant de l'ARNIII régulant les facteurs de virulence. L'interaction entre ces deux régulons a permis à la bactérie d'acquérir de nouvelles fonctions et sans doute de coloniser de nouveaux tissus de l'hôte. Ainsi, il se pourrait que la fonction primaire de *S. aureus* fût d'être une bactérie commensale non pathogène des fosses nasales comme *S. xylosus* et *S. carnosus*. Ce ne serait qu'après l'acquisition de facteurs de virulence via les éléments génétiques mobiles (Rosenstein *et al.*, 2010) et d'un répertoire d'ARN régulateurs capables de contrôler leurs synthèses dans des conditions spécifiques que *S. aureus* serait devenu pathogène. Il serait donc très intéressant de comparer l'ensemble des ARN transcrits entre espèces de staphylocoques pathogènes et non-pathogènes et de combiner ces données aux analyses classiques de comparaisons des séquences codantes.

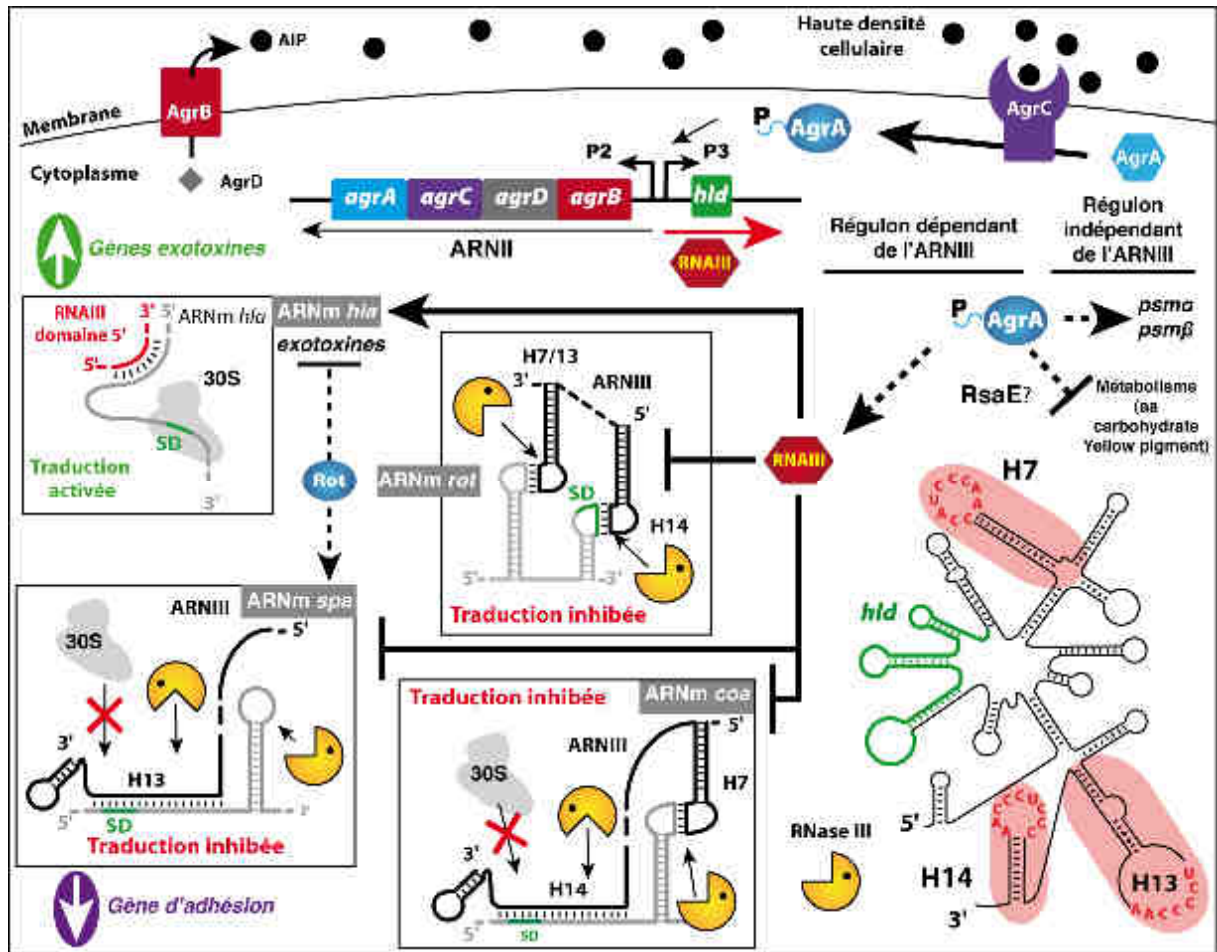


Figure 25 : Le système de densité cellulaire *agr*.

A haute densité cellulaire, le peptide AgrD (en gris) est maturé en peptide auto inducteur AIP et sécrété dans le milieu extracellulaire *via* la protéine AgrB (en bordeaux). Le peptide se fixe à l'HK ArgC (en violet) qui phosphoryle AgrA (en bleu). AgrA régule l'expression de gènes de manière dépendante et indépendante de l'ARNIII. AgrA active sa propre transcription (*feedback loop*) et la transcription des gènes codant pour des PSM en se fixant directement à leurs promoteurs. En parallèle, AgrA régule négativement des gènes de métabolisme peut-être *via* l'activation de l'expression de *RsaE*, qui est requis dans la régulation des voies métaboliques du TCA et du folate (Geissmann *et al.*, 2009). L'induction d'autres ARN régulateurs par AgrA reste à prouver. AgrA active également la transcription de l'ARNIII. L'ARNIII est un ARN régulateur multifonctionnel codant pour l'hémolysine  $\delta$  (en vert). Sa structure secondaire présente plusieurs domaines de régulation dont trois tiges-boucles présentant un motif conservé riche en C (en rouge), ciblant la séquence SD (en vert) des ARNm cibles. L'ARNIII est au cœur d'un réseau complexe de régulation qui implique une double *feedforward loop*, aboutissant à l'induction des exotoxines au détriment des facteurs d'adhésion. L'ARNIII réprime la traduction des ARNm *spa*, *coa* et *rot* en séquestrant la SD dans une région double-brin, et recrute la RNase III pour dégrader les ARNm cibles.

#### IV.D.3. *RsaE* et la régulation du métabolisme central

*RsaE* est l'un des 10 ARN non codant de *S. aureus* identifiés peu avant mon arrivée au laboratoire (*RsaA* à *RsaK*) (Geissmann *et al.*, 2009). *RsaE* a la particularité d'être la seule région intergénique trouvée conservée chez les bactéries des genres *bacillaceae* et *staphylococcaceae*. Ce sARN de 96 nts possède deux motifs riches en C exposés dans des

régions en simple brin. Cette séquence **UCCC** est également retrouvée dans les tiges-boucles 7, 13 et 14 de l'ARNIII et pourrait constituer les régions fonctionnelles de **RsaE**. Il est également à noter que son expression est sous la dépendance du système *agr* car son expression est fortement diminuée dans une souche *agr* mutante (Geissmann *et al.*, 2009). Utilisant des approches protéomique et transcriptomique, il a été montré que **RsaE** réprime la synthèse des protéines OppB, les deux sous-unités de la succinyl-CoA synthase impliquée dans le cycle TCA, les deux protéines FhS et FOLD intervenant dans le cycle du folate et la protéine SA0873 (de fonction inconnue). Les RBS de ces ARNm peuvent former des appariements stables avec les régions riches en C de **RsaE** (Geissmann *et al.*, 2009). Il a été ainsi proposé que RsaE coordonnerait la répression du métabolisme central lorsque les sources carbonées deviennent limitantes, i.e. en phase stationnaire. Il est intéressant de signaler que cette régulation pourrait être reliée au système de densité cellulaire.

Cependant, l'expression de **RsaE** est différente dans les souches cliniques puisqu'il serait principalement exprimé en phase exponentielle, comme observé chez *B. subtilis* (Song *et al.*, 2012, voir **Résultats** III). Est-ce que **RsaE** exercerait des fonctions différentes selon le stress induit? Quelles sont donc les cibles régulées par **RsaE** chez *B. subtilis*? Comment les fonctions de RsaE ont-elles évoluées dans les bactéries distantes telles que *B. subtilis* et *S. aureus*?

#### **IV.D.4. Îlots de Pathogénie : rôle de SprD dans la virulence**

Le génome de *S. aureus* contient de nombreux éléments mobiles qui exercent des rôles importants dans le transfert horizontal et l'évolution du génome. Ils confèrent à la bactérie des fonctions adaptatives pour survivre au sein de l'hôte. Ces éléments contiennent des gènes de résistance aux antibiotiques comme la méticilline, codés par exemple par la cassette SCC-mec, ou sont issus de plasmides ou de phages. Les îlots de pathogénie de *S. aureus* (SaPI) sont d'autres éléments génétiques mobiles de 14 à 17 kb qui sont les dépositaires de nombreux gènes codant pour des facteurs de virulence comme par exemple certains superantigènes, des entérotoxines et des toxines liées au système de sécrétion TSST (Malachowa *et al.*, 2010), et des sARN, appelés **SprA-G** (Pichon & Felden, 2005). Certains de ces sARN sont présents en plusieurs copies réparties sur le génome bactérien. Alors que les

sARN sont supposés réguler les gènes codés par le SaPI, ils peuvent aussi établir des liens fonctionnels entre l'îlot de pathogénie PI et le génome de base, ce qui met en relief les étroites connexions établies au cours de l'évolution entre les éléments génétiques mobiles et les gènes du chromosome bactérien.

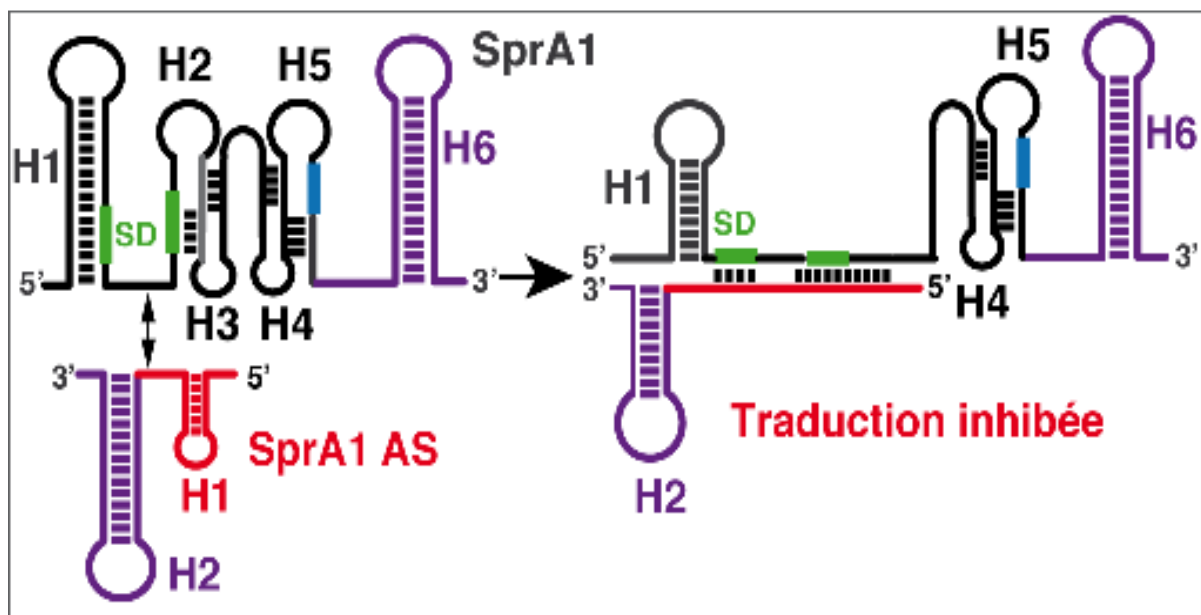
C'est ainsi le cas de *SprD* qui régule l'expression de l'ARNm *sbi*, codant pour le facteur d'adhésion, la protéine G. Un duplex de 40 pb est formé entre les deux ARN indépendamment de Hfq, inhibant la traduction de l'ARNm *sbi* sans affecter sa stabilité (Chabelskaya *et al.*, 2010). Cependant, l'expression de *SprD* semble être constitutive, i.e. transcrit à toutes les phases de croissance, alors que le taux de l'ARNm *sbi* diminue de manière significative en phase stationnaire. Ces contradictions suggèrent qu'un niveau supplémentaire de régulation a lieu en phase stationnaire soit au niveau transcriptionnel, soit au niveau de la stabilité de l'ARNm. On ne peut exclure l'intervention d'un autre ARN tel que l'ARNIII en phase stationnaire pour réprimer de manière irréversible la synthèse de la protéine G à haute densité cellulaire. Chez un modèle murin, l'expression de *SprD* module la virulence de *S. aureus*, mais indépendamment de l'expression de la protéine Sbi. Ces résultats suggèrent que *SprD* pourrait réguler d'autres ARNm impliqués dans la virulence (Chabelskaya *et al.*, 2010).

#### IV.D.5. *SprA1-AS*, régulateur d'un système toxine-antitoxine

Au niveau des SaIP, l'ARN *SprA1* est transcrit en même temps que son ARNas, *SprA1-AS*. Ce système ARNm/ARNas est prédit comme un système toxine/antitoxine de type I où *SprA1* code pour un peptide de composition hydrophobe (Fozo *et al.*, 2010). Des expériences réalisées *in vitro* ont montré que ce peptide possède une activité lytique vis-à-vis des érythrocytes, ainsi que des propriétés antimicrobiennes contre des bactéries Gram négatif et Gram positif (Sayed *et al.*, 2012). C'est dans des conditions normales et riches de croissance que l'expression de *SprA1-AS* est la plus importante alors que l'ARN *SprA1* serait exprimé en conditions de stress oxydatif.

*SprA1-AS* se lie à *SprA1* au niveau du RBS, inhibant en conséquence l'initiation de la traduction. De manière inattendue, l'hybride *SprA1/SprA1-AS* est imparfait, alors même que les parties 3' de ces ARN sont parfaitement complémentaires, puisque les deux ARN sont

codés au même locus. Pourquoi un tel site de fixation est-il privilégié ? L'explication réside dans le fait que les régions parfaitement complémentaires sont les terminateurs de transcription Rho-indépendant des deux ARN. Ces tiges-boucles riches en appariements G-C à leur base, sont probablement trop stables pour initier un appariement entre les deux ARN. A l'inverse, la région 5' de *SprA1-AS*, qui n'est pas structurée, sera plus à même de se lier à la boucle de connexion H2 du pseudo-nœud de *SprA1*. De nouveau, la régulation ne fait pas intervenir Hfq. Le point le plus intéressant de cette étude met en évidence la nécessité de tenir compte non seulement de la structure de l'ARN régulateur mais aussi de sa cible pour mieux comprendre les mécanismes de régulation en jeu (Sayed *et al.*, 2011) (Figure 25).



**Figure 26 : Système toxine-antitoxine *SprA1/SprA1-AS*.**

*SprA1* code pour un peptide cytolitique dont la synthèse est régulée par un ARN antisens, *SprA1-AS*. Les régions 5' des deux ARN s'apparient en séquestrant le RBS (SD en vert) de *SprA1* dans une région double-brin, ce qui se traduit par une inhibition de la traduction. Les régions parfaitement complémentaires (en violet) de chaque ARN n'interviennent pas dans l'appariement (Codon STOP en bleu).

# Objectifs



Le système de densité cellulaire *agr* fait intervenir deux effecteurs : l'ARNIII et AgrA. Alors qu'AgrA régule plus particulièrement les gènes impliqués dans le métabolisme et les peptides PSM, l'ARNIII régule de manière temporelle l'expression des facteurs de virulence : à haute densité, les exotoxines sont activées alors que les protéines d'adhésion sont inhibées. Ceci permet à la bactérie de passer d'un mode défensif à un mode d'attaque/dissémination. L'ARNIII, ARN structuré multifonctionnel représente donc un modèle d'étude idéal pour caractériser les mécanismes moléculaires de la régulation et le rôle de la structure des ARN dans ces processus. J'ai contribué à établir les bases structurales de la répression de l'ARNm *coa*, cible de l'ARNIII. Nous avons mis en évidence une structure inhabituelle d'un complexe régulateur impliquant deux régions éloignées de l'ARN. Les résultats obtenus à l'issue de ce travail nous ont poussés à regarder de plus près les propriétés de reconnaissance des interactions boucles boucles par la RNase III de *S. aureus*. Ce co-facteur de l'ARNIII reconnaît des topologies ARN-ARN variées, suggérant que cette enzyme doit exercer un rôle global dans la régulation par les ARN.

Par définition, le système de densité cellulaire *agr* s'auto-induit (voir [introduction IV.D.2](#)), mais son expression est aussi modulée par divers régulateurs intracellulaires dont SarA (Chien *et al.*, 1998). Les études fonctionnelles portées sur ce réseau se basent essentiellement sur des souches dérivées de la souche sauvage NCTC8325 de *S. aureus* portant une mutation au niveau du gène *rsbU* (Boisset *et al.*, 2007, Huntzinger *et al.*, 2005, Chevalier *et al.*, 2010). Cette protéine est responsable de l'activation du facteur de transcription  $\sigma^B$  en phase stationnaire de croissance ou en réponse au stress (Pané-Farré *et al.*, 2006, revue Hecker *et al.*, 2007). De manière intéressante, l'expression de  $\sigma^B$  s'accompagne d'une diminution de l'expression du système *agr* et donc de l'ARNIII (Bischoff *et al.*, 2001, Horsburgh *et al.*, 2002, Bischoff *et al.*, 2004, Herbert *et al.*, 2010). Cette inhibition est indépendante de SarA dont l'expression reste stable en présence de  $\sigma^B$  (Horsburgh *et al.*, 2002). Mon travail de thèse a eu pour but principal de déterminer la fonction et le mécanisme d'action de deux sARN *S. aureus*, dont l'un est sous la dépendance d'un promoteur  $\sigma^B$ , *RsaA*. Le second, *RsaE*, impliqué dans le cycle de Krebs, a la particularité d'être conservé chez les bacillaceae et les staphylococeae (voir [introduction IV.D.3](#)).



Quel est le mécanisme de régulation de l'ARNm *coa* par l'ARNIII ? Comment l'endoribonucléase III de *Staphylococcus aureus* reconnaît-elle spécifiquement les interactions boucles-boucles entre l'ARNIII et ses cibles ?

**RsaA** est-il le répresseur du régulon  $\sigma^B$  chez *Staphylococcus aureus* ? L'expression et la régulation des cibles de **RsaA** peut-elle relier la réponse au stress à la régulation des gènes de virulence ?

Les cibles et la fonction de **RsaE** sont-elles conservées au cours de l'évolution chez *Staphylococcus aureus* et *Bacillus subtilis* ?

De manière plus générale, la caractérisation des réseaux de régulation de ces ARN régulateurs peut-elle relier les différents réseaux déjà caractérisés chez *S. aureus*, si oui à quels niveaux ? (Figure 26)

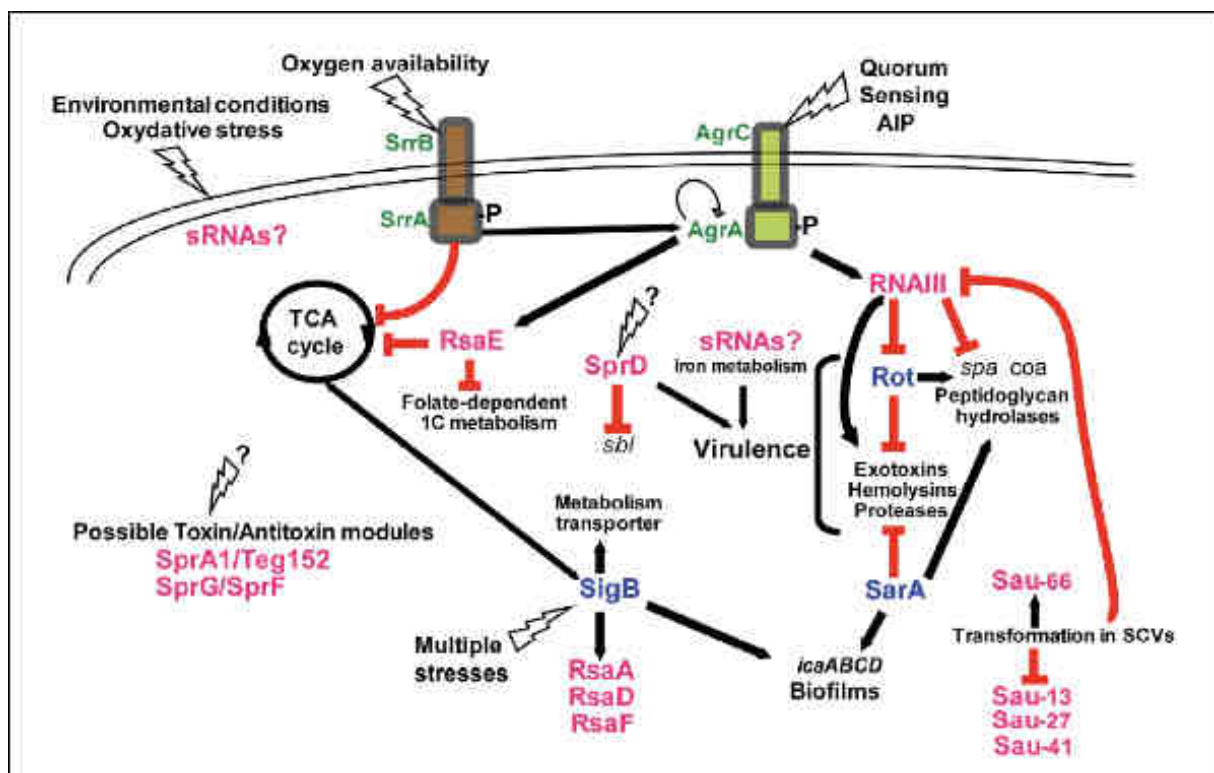


Figure 27 : Réseaux de régulation médiés par les ARN régulateurs chez *S. aureus* : états des connaissances (D'après Felden *et al.*, 2011).

# Résultats



# Mécanisme de reconnaissance des interactions boucle-boucle par l'endoribonucléase III de *Staphylococcus aureus*

Chevalier C., Boisset S., **Romilly C.**, Masquida B., Fechter P., Geissmann T., Vandenesch F. & Romby P. (2010). *Staphylococcus aureus* RNAIII binds to two distant regions of *coa* mRNA to arrest translation and promote mRNA degradation. *PLoS Pathog* 6, e1000809.

**Romilly, C.**, Chevalier, C., Lioliou, E., Marzi, S., Masquida, B., Geissmann, T., Vandenesch, F., Westhof, E. and Romby, P. **Loop-loop interactions in mRNA regulation are processed by the endoribonuclease III in *Staphylococcus aureus*. Manuscrit.**

## I. L'endoribonucléase III, une enzyme-clé dans la régulation par les ARN

La famille des RNases III regroupe des enzymes qui reconnaissent et dégradent spécifiquement les régions en double brin des ARN. Toutes les protéines de cette famille portent un domaine caractéristique appelé "domaine RNase III". Ce dernier catalyse la réaction de coupure en présence d'ions magnésium des ARN doubles brins pour générer des extrémités 5' monophosphate et des extrémités 3' OH avec deux nucléotides sortants (**Figure 27 A**). La famille des RNases III est retrouvée dans tous les royaumes du vivant et se divise en trois sous-familles (revue McRae *et al.*, 2007) (**Figure 27 B**).

Les protéines de la première sous-famille, caractéristiques des bactéries, sont en général d'une taille inférieure à 500 aa et possèdent un domaine de liaison à l'ARN double brin (*dsRBD* ou *double stranded RNA binding domain*) et un domaine RNase III. Pour couper les deux brins d'une hélice ARN, ces enzymes agissent sous forme d'homodimère où le domaine RNase III de chaque partenaire est responsable de la coupure d'un des deux brins d'ARN (voir **introduction II.D**).

La seconde famille comporte les enzymes Drosha, impliquées dans la maturation des pre-ARNr et des pri-miARN. Drosha possède deux domaines RNases III pour un RBD, ainsi qu'un domaine riche en proline, pour une longueur d'environ 1500 aa. Cette enzyme fonctionne sous forme monomérique où chaque domaine RNase III est responsable de la coupure d'un des brins de l'hélice ARN. Néanmoins, Drosha seule est incapable de maturer efficacement les pri-miARN. Elle est guidée vers son site de coupure via le recrutement de la protéine DGR-8 avec laquelle elle forme le complexe "microprocesseur" (*microprocessor complex*). DGR-8 est capable de positionner Drosha à 11 pb des jonctions simples brins et doubles brins d'ARN (Han *et al.*, 2004).

Enfin, la sous-famille la plus complexe est représentée par les protéines Dicer d'une taille de 2000 aa et qui portent 2 domaines RNase III, un RBD, un domaine PAZ (Piwi-Argonaute-Zwilli), un domaine DUF (*domain of unknown function*) et 2 domaines hélicases. De manière similaire à la famille Drosha, Dicer est active sous forme monomérique et

chaque domaine RNase III est responsable de la coupure d'un des deux brins d'ARN (Zhang *et al.*, 2004). A partir des pre-miARN, Dicer génère des duplexes miARN / miARN\* de 21 à 27 nts, dont l'un des deux brins, le brin guide, sera chargé au sein du complexe RISC (*RNA induced silencing complex*). Chez l'homme, Dicer reconnaît l'extrémité des pré-miARN via le domaine PAZ en utilisant deux poches qui fixent les extrémités 3' et 5' respectivement, pour positionner correctement l'ARN dans le site catalytique de l'enzyme (Park *et al.*, 2011).

L'ensemble des membres de la famille possèdent quatre résidus strictement conservés au cours de l'évolution qui coordonnent la liaison de deux ions magnésium indispensables à la réaction de coupe : E41, D45, D114, and E117 d'après la nomenclature d'*E. coli* (Figure 27 C) (Zhang *et al.*, 2004, Han *et al.*, 2004).

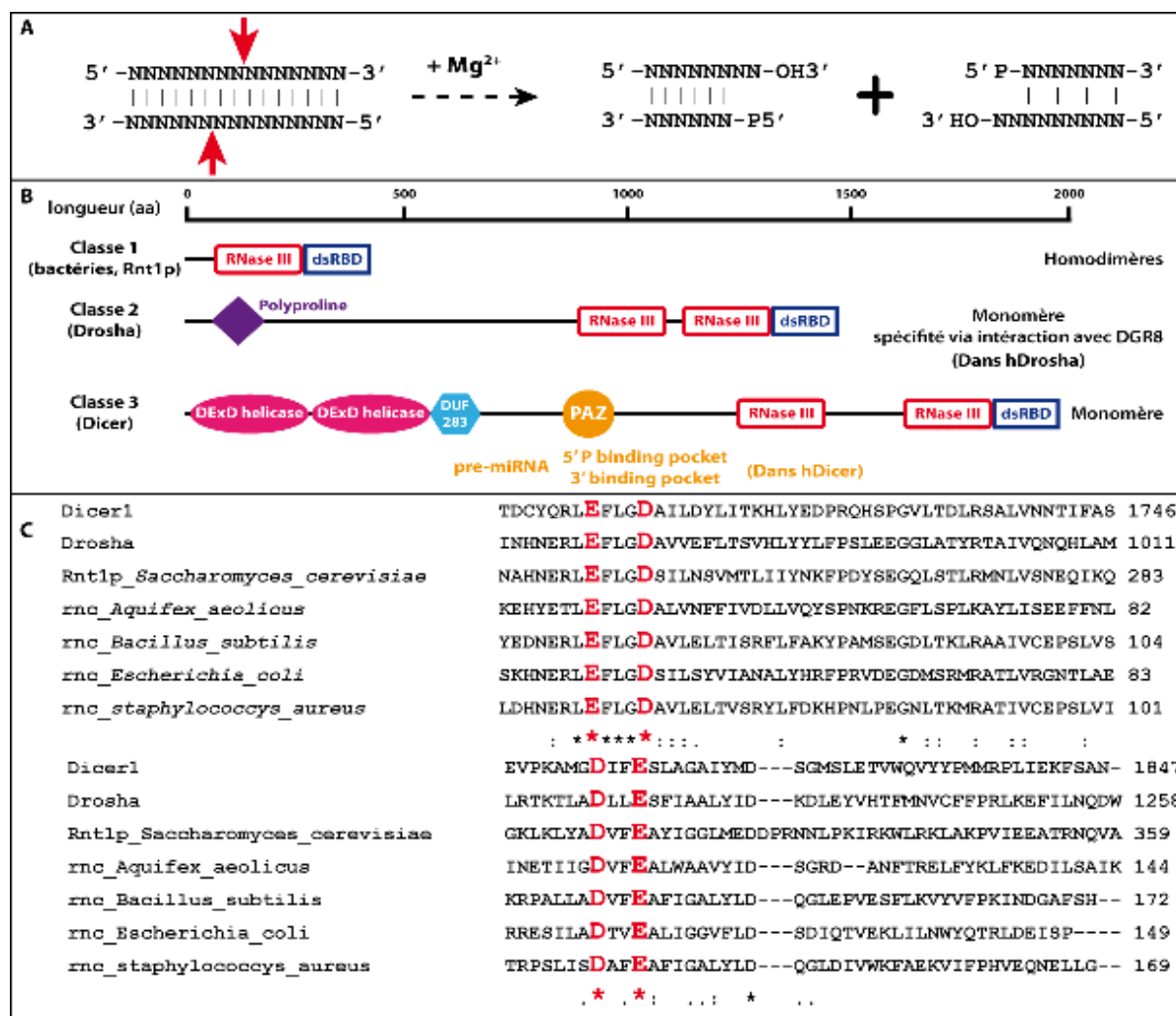


Figure 28 : Généralités sur la famille des RNases III.

**A.** Réaction de coupe générant deux extrémités 5' monophosphate et deux extrémités 3'OH avec deux nucléotides sortants. **B.** Présentation des domaines caractérisant les trois sous-familles de RNase III avec la longueur relative en aa (d'après MacRae *et al.*, 2007). **C.** Alignement local des séquences protéiques des

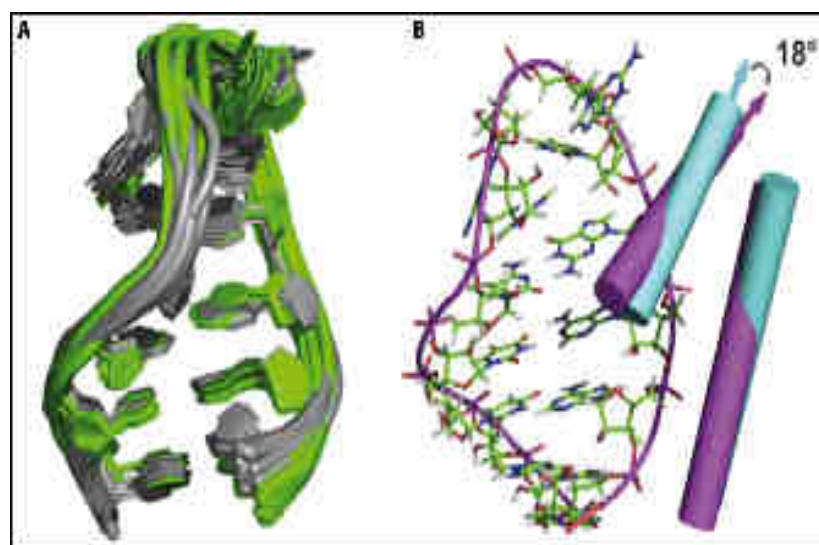
différents membres de la famille des RNases III. Les acides aminés strictement conservés (en rouge) coordonnent la liaison d'ions magnésium.

### I.A. La topologie des boucles est importante dans la reconnaissance par Rnt1p chez *Saccharomyces cerevisiae*.

Chez *S. cerevisiae*, la RNase III Rnt1p (voir Résultats I) est impliquée dans la maturation des ARNr, snARN, snoARN et des séquences introniques des pré-ARNm. En tant que membre de la première sous famille des RNase III, Rnt1P possède un domaine RNase III et un dsRBD. A l'inverse des enzymes présentes chez les bactéries, Rnt1p possède une extension N-ter qui semble être importante dans la stabilisation des homodimères.

De manière surprenante, Rnt1p est capable de reconnaître deux types de boucles différentes au sein des substrats. Le premier type arbore une tétra-boucle AGNN, alors que le second type possède une tétra-boucle AAGU. Comment le dsRBD de Rnt1p parvient à reconnaître spécifiquement ces deux types de substrats ? Des études de spectroscopie par RMN par Wang *et al.*, ont montré que la topologie générale des deux types de tétra-boucles est la même. Le RBD de Rnt1p est capable de se lier à chaque tige-boucle, et la spécificité de substrat est obtenue via un changement conformationnel de l'hélice  $\alpha 1$  du RBD. Ce changement d'orientation de  $18^\circ$  permet à l'hélice de s'insérer parfaitement dans le sillon mineur des tétra-boucles au sommet de l'hélice, sans changer la structure de cette dernière

(Wang *et al.*, 2011) (Figure 28).



**Figure 29 : La topologie des tétra-boucles est la clé dans la reconnaissance des tiges-boucles par Rnt1p (Wang *et al.*, 2011).**

**A.** Superposition des structures des tiges-boucles AGAA (gris) et AAGU (vert) obtenues par NMR.  
**B.** Comparaison de l'angle de l'hélice  $\alpha 1$  seule (bleu) et en présence de la tige-boucle AGUU (violet).

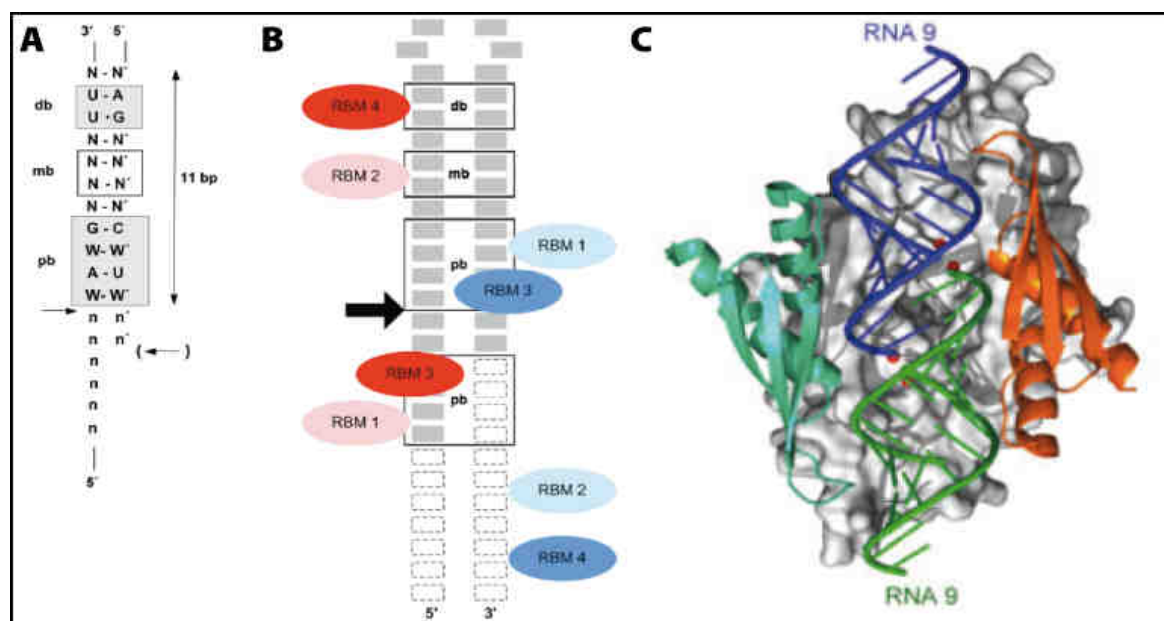
## I.B. La RNase III des bactéries coupe les hélices ARN sans spécificité apparente de séquence

Les RNases III eucaryotes utilisent différentes stratégies pour s'ancrer à leurs substrats ARN : Drosha interagit avec un partenaire protéique, Dicer possède un domaine PAZ et Rnt1p chez la levure s'ancre spécifiquement à certaines boucles apicales. Qu'en est-il chez les bactéries ?

Les RNases III bactériennes sont les enzymes de plus petite taille retrouvées dans l'ensemble de la famille des endoribonucléases III. Des études sur la spécificité de reconnaissance des duplexes ARN-ARN montre que ces dernières ne reconnaissent pas de séquences spécifiques dans les long duplexes. Néanmoins, des déterminants (Pertzev *et al.*, 2006) et des anti-déterminants (Zhang *et al.*, 1997) de séquences capables de moduler la liaison de la RNase III ont été mis en évidence (**Figure 29 A**). Ces éléments sont des séries d'interactions W-C au niveau de boîte dites proximales, médianes et distales qui agissent de manière coopérative pour établir une hiérarchie dans les substrats ARN au sein de la cellule (**Figure 29 B**). Les données cristallographiques obtenues sur la RNase III de *Aquifex aeolicus* (*A. aeolicus*) montre que le substrat minimum clivé par l'enzyme est de 11 pb. En conséquence, une hélice ARN double-brin sera coupée par un dimère de RNase III si elle est d'une taille minimum de 22 pb. Quatre RBM (*RNA binding motif* ou motif de liaison à l'ARN) sont impliqués dans la liaison du substrat. Le RBM 1, constitué de la première hélice  $\alpha$  de RBD, forme des liaisons hydrogènes avec le O2' des riboses des bases de la boîte proximale. Cette boîte se situant aux abords du site de coupure, le RBM 1 joue un rôle capital dans le positionnement de la RNase III. Le RBM 2 se situe entre les brins  $\beta$ 1 et  $\beta$ 2 du RBD et effectue des liaisons hydrogènes avec les O2' des bases au niveau de la boîte médiane. De manière intéressante, le résidu Q157 interagit directement avec la base située 5 nucléotides avant le site de coupure. Néanmoins, ce résidu est capable de lier un A ou un U suggérant que le RBM 2 n'est pas impliqué dans la reconnaissance d'une séquence spécifique (Gan *et al.*, 2008). Le RBM 3, porté par l'hélice  $\alpha$ 4 du RBD, interagit avec le O2' du résidu au niveau du site de coupure et en forme une liaison hydrogène avec le résidu en amont de ce dernier. Pour finir, le RBM 4 se situe dans la boucle entre les hélices  $\alpha$ 5 et  $\alpha$ 6 du RBD. Ce dernier est



impliqué dans la reconnaissance des appariements de bases de la boîte distale en effectuant des interactions hydrogènes avec les O2' des riboses. Pour résumer, le RBD forme 7 interactions hydrogènes avec les groupements hydroxyles en O2' des riboses des bases (Gan *et al.*, 2006) (Figure 29 C). La réaction de coupure fait obligatoirement intervenir deux ions magnésiums, et se fait en plusieurs étapes allant du transfert du groupement phosphoryle reliant les deux bases à séparer jusqu'à la libération du produit (Gan *et al.*, 2008). De manière surprenante, l'absence de ces ions ou la présence d'une mutation dans le site catalytique empêche un ARN double-brin de s'accommoder dans la vallée catalytique du domaine RNase III. A l'inverse, l'ARN en double brin reste lié exclusivement au dsRBD (Gan *et al.*, 2008; Blaszczyk *et al.*, 2004).



**Figure 30 : Déterminants de séquence favorisant la liaison de la RNase III (d'après Pertzev *et al.*, 2006) et structure cristalline de la RNase III d'A. *aeolicus* (D'après Gan *et al.*, 2006 et 2008).**

**A. Pertzev *et al.*, 2006 :** Déterminants de séquences pour la liaison de la RNase III. db : boîte distale, mb : boîte médiane et pb : boîte proximale. **B. Gan *et al.*, 2006 :** représentation des appariements de bases reconnus par les RBM du RBD de la RNase III. **C. Gan *et al.*, 2008 :** Structure cristalline d'un dimère de RNase III d'A. *aeolicus* en présence de deux hélices d'ARN de 9 paires de bases avec 2 nucléotides 3' sortants.

### I.C. La RNase III de *S. aureus* joue un rôle important dans la régulation par les ARN

Plusieurs cibles de l'ARNIII, effecteur intracellulaire du système de densité cellulaire *agr*, ont été validées avant mon arrivée dans le laboratoire, dont les ARNm codant pour la protéine *rot*, un répresseur des toxines et la protéine A (Boisset *et al.*, 2007 ; Huntzinger et

al., 2005). En collaboration avec Clément Chevalier, j'ai déterminé le mécanisme de régulation d'une des cibles de l'ARNIII, l'ARNm *coa*, codant pour la coagulase. Ma contribution personnelle a été de définir les régions d'interaction entre les deux ARN et l'apport de chaque site dans la formation du complexe et la régulation. Utilisant la cartographie de l'ARN en solution et le suivi des complexes par gel retard, nous avons montré que deux régions distantes de l'ARNm *coa* interagissent avec deux tiges-boucles de l'ARNIII contenant les motifs riches en C (H13 et H7). Pour déterminer les régions d'interactions, nous avons utilisé divers ARNIII variants (délétés soit des tiges-boucles 7, 13 ou 14) et analysés leur fixation à l'ARNm par gel retard ou par cartographie de l'ARN en solution (voir [Annexes](#), Helfer Romilly *et al.*, 2011). Nous avons ainsi montré que l'interaction entre la tige-boucle 13 de l'ARNIII et le site de fixation du ribosome est cruciale pour la stabilité du duplexe et pour une vitesse d'association rapide entre les deux ARN (de l'ordre de  $10^6 \text{ M}^{-1}\text{s}^{-1}$ ). En revanche, l'interaction entre la tige-boucle 7 de l'ARNIII et une tige-boucle de la région codante de l'ARNm contribue à la stabilité du complexe, mais la perte de l'interaction boucle-boucle n'a pas d'incidence majeure sur la formation du premier appariement. La tige-boucle 13 de l'ARNIII doit probablement initier la fixation à l'ARNm par une interaction de type boucle-boucle mais la cartographie en solution indiquent que ces appariements sont étendus rapidement pour former un duplex irrégulier. Le complexe ARNIII-ARNm *coa* ainsi formé empêche la formation du complexe d'initiation et recrute la RNase III pour induire une dégradation de l'ARNm réprimé. Nous avons montré que la région qui séquestre la séquence SD est en grande partie responsable de la répression *in vivo* et *in vitro* alors que la deuxième interaction améliore l'efficacité du contrôle *in vivo* ([Figure 30](#)) (Chevalier, Boisset, Romilly *et al.*, 2010, voir [Annexes II](#)).

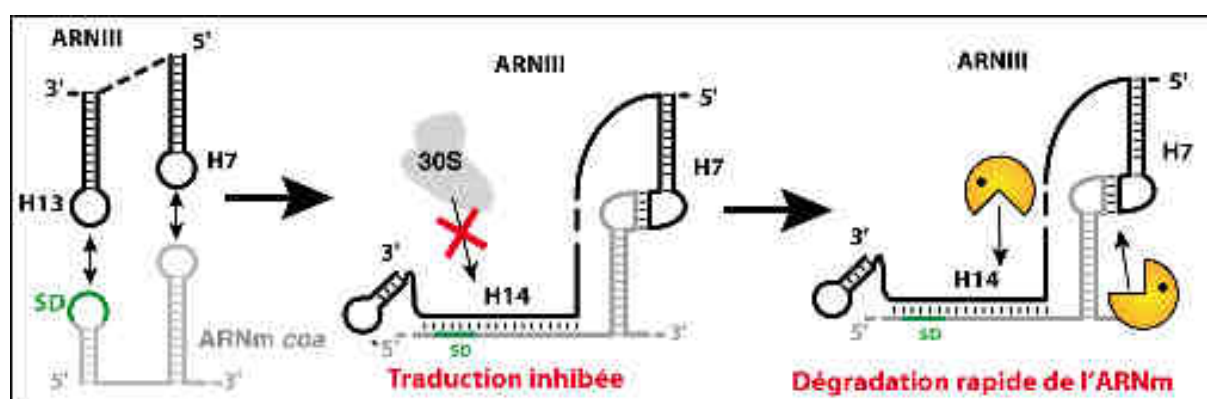


Figure 31 : Modèle de la régulation de l'ARNm *coa* par l'ARNIII.

Ainsi la RNase III est un co-facteur de l'ARNIII qui réprime de manière irréversible la synthèse des facteurs d'adhésion. Cette étude apporte un nouvel exemple qui montre que la RNase III reconnaît une grande variété de topologies ARNIII/ARNm. De manière surprenante, l'enzyme est capable de reconnaître et couper les interactions boucles-boucles à une position unique dans les deux ARN, bien que ces dernières ne comprennent que 7 à 8 paires de bases en moyenne et ceci indépendamment de la longueur des hélices des tiges-boucles. De manière encore plus surprenante, les séquences impliquées dans les interactions boucle-boucle entre l'ARNm *rot/coa* et l'ARNIII sont identiques. La séquence de type SD UGGGAU, retrouvée au niveau de la région codante de l'ARNm *coa* et de la région 5'UTR de l'ARNm *rot*, s'apparie au motif AUCCCA des tiges-boucles 7, 13 et 14 de l'ARNIII. Une modélisation graphique des interactions boucles-boucles des deux systèmes montre une topologie commune où l'hélice intermoléculaire et les deux hélices intramoléculaires s'empilent de manière co-axiale pour former un long domaine hélicoïdal. En revanche, on observe une légère courbure au niveau des séquences appariées lorsque l'on superpose ce modèle avec une hélice régulière de même taille (**Figure 31 A et B**).

**Chez *S. aureus*, quels sont les éléments structuraux et/ou de séquences présents au niveau des interactions boucles-boucles capable de guider la RNase III vers son site de coupure spécifique ? (Romilly, Chevalier *et al.*, 2012)**

Dans l'article soumis ci-dessous, nous avons modifié la séquence de l'interaction boucle-boucle et avons analysé le profil de coupures par la RNase III (**Figure 31 D**). Une étude comparative a été réalisée sur une longue hélice régulière de même taille que le complexe boucle-boucle et sur une longue tige-boucle. Les résultats suggèrent que l'interaction boucle-boucle adopte une topologie appropriée pour créer un site de fixation à la RNase III alors que l'hélice régulière ou la tige-boucle sont coupées à plusieurs endroits révélant l'existence de plusieurs sites d'interaction. Il est intéressant de signaler que l'introduction d'une boucle qui interrompt l'hélice restreint le nombre de coupures par la RNase III (**Figure 31 E**). Tout type d'interaction boucle-boucle n'est pas coupé par l'enzyme, par exemple si les boucles qui traversent les sillons l'interaction boucle-boucle sont trop longues, la RNase ne coupe plus l'ARN probablement à cause d'un empêchement stérique

(Figure 31 C). Ainsi, chaque ARN contient des éléments positifs et négatifs qui dirigent ou qui empêchent le positionnement de la RNase III.

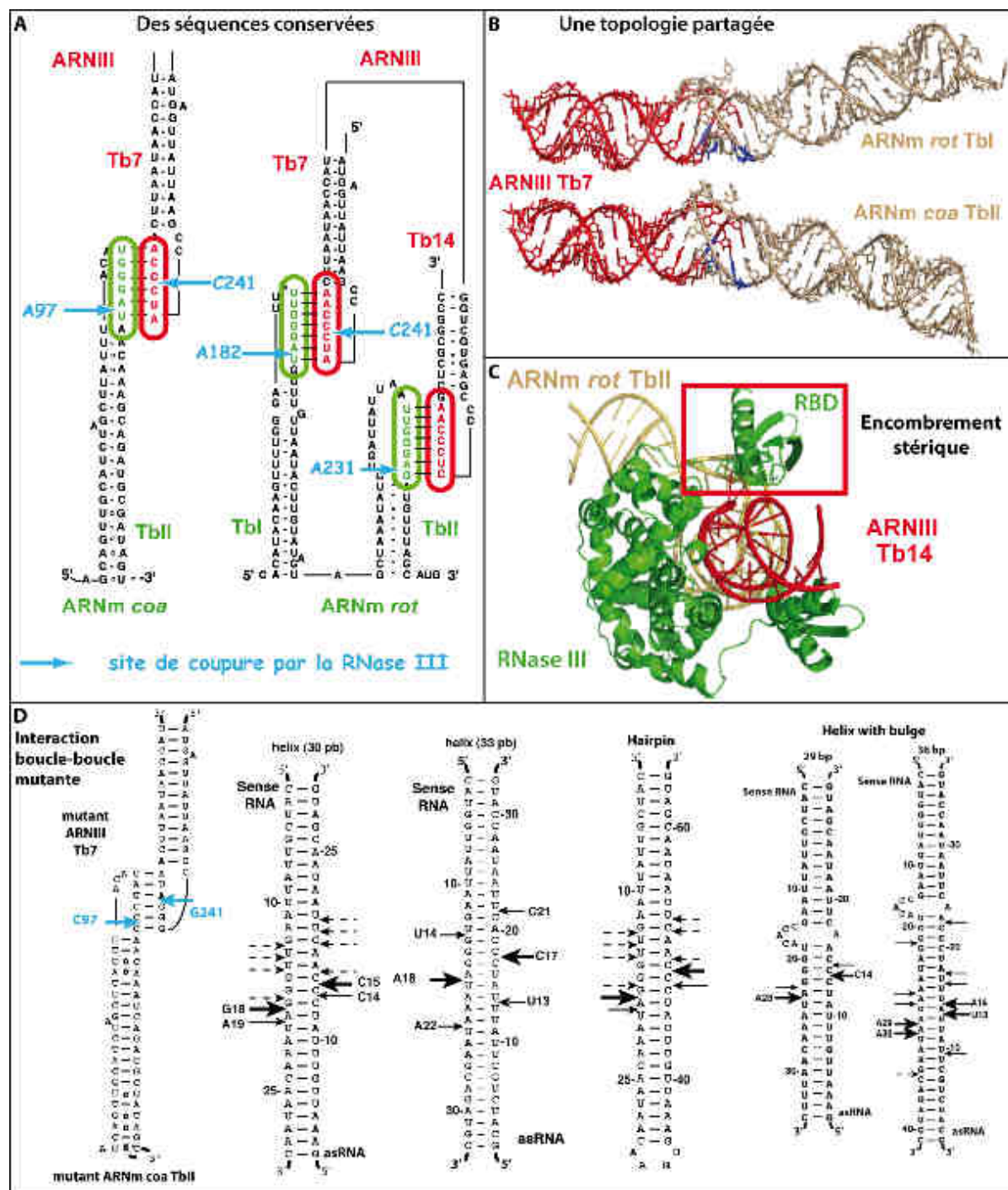


Figure 32 : La topologie des interactions boucle-boucle et des hélices guide la RNase III vers son site de coupure.

A. Structure secondaire des interactions boucle-boucle entre l'ARNIII et ses ARNm cibles *coa* et *rot*. Les séquences SD retrouvées au niveau des ARNm sont indiquées en vert, le motif riche en C de l'ARNIII en rouge, les sites de coupures de la RNase III en bleu. B. Modélisation des interactions boucle-boucle (S. Marzi, B. Masquida). C. Modélisation de l'interaction RNase III / interaction boucle-boucle ARNIII-ARNm *rot* TbII. La longue boucle de la TbII de l'ARNm montre un encombrement stérique avec le RBD d'un des dimères de la

RNase III. **D.** Coupure de la RNase III sur une interaction boucle-boucle mutante et diverses hélices (voir manuscrit 2 pour plus de détails).

**Loop-loop interactions involved in antisense regulation are processed by the endoribonuclease III in *Staphylococcus aureus***

Cédric Romilly<sup>1§</sup>, Clément Chevalier<sup>2§</sup>, Stefano Marzi<sup>1</sup>, Benoit Masquida<sup>1</sup>, Thomas Geissmann<sup>3,4</sup>, François Vandenesch<sup>3,4,5</sup>, Eric Westhof<sup>1</sup>, & Pascale Romby<sup>1</sup>

<sup>1</sup>Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, 15 rue René Descartes, F-67084 Strasbourg, France; <sup>2</sup>Department of Biology, ETH Zurich, CH-8092 Zurich, Switzerland; <sup>3</sup>Université de Lyon, France; <sup>4</sup>Inserm U851, 21 Avenue Tony Garnier, Lyon, F-69007 Lyon, France; <sup>5</sup>Hospices Civils de Lyon, France.

§ co-first authors

\*To whom correspondence should be addressed: [p.romby@ibmc-cnrs.unistra.fr](mailto:p.romby@ibmc-cnrs.unistra.fr)

**Running title:** Loop-loop interaction as a substrate of RNase III

**Keywords:** endoribonuclease, gene regulation, processing, loop-loop interaction

## **Abstract**

The endoribonuclease III (RNase III) belongs to the enzyme family known to process double-stranded RNAs. *Staphylococcus aureus* RNase III was shown to regulate, in concert with the quorum sensing induced RNAIII, the degradation of several mRNAs encoding virulence factors and the transcriptional repressor of toxins Rot. Two of the mRNA-RNAIII complexes involve fully base paired loop-loop interactions with similar sequences that are cleaved by RNase III at a unique position. We show here that the sequence of the base pairs within the loop-loop interaction was not critical for RNase III cleavage but that the co-axial stacking of three consecutive helices provides an ideal topology for RNase III recognition. In contrast, RNase III induces several strong cleavages in a regular helix, which carries a sequence similar to the loop-loop interaction. The introduction of a bulged loop that interrupts the regular helix restrains the number of cleavages. This work shows that *S. aureus* RNase III is able to bind and cleave a variety of RNA-mRNA substrates, and that specific structure elements direct the action of RNase III.

## Introduction

The regulation of mRNA decay plays essential roles in the regulation of gene expression enabling the bacteria to selectively activate or repress appropriate genes in response to a specific physiological state. This is particularly true for pathogenic bacteria such as *Staphylococcus aureus*, which can survive under extreme circumstances in the human host as well as in the environment. *S. aureus* is both a commensal bacterium and an opportunistic pathogen, which causes a wide variety of infections, owing in part to the coordinated expression of a large repertoire of virulence factors<sup>1-3</sup>. Genome-wide studies have shown that the decay of mRNAs encoding virulence factors is highly regulated during growth<sup>4; 5</sup>. Furthermore, the stability of many transcripts is modified upon the entry into the stationary phase of growth and in response to various stresses<sup>6; 7</sup>. Detailed knowledge about the RNA decay machinery in *S. aureus* was recently gained<sup>5; 8</sup>. Indeed the genome encodes all the orthologs of *Bacillus subtilis* exoribonucleases and endoribonucleases suggesting that the mechanisms of RNA decay are conserved in Gram-positive bacteria<sup>9; 10</sup>. A two-hybrid system identified a degradosome-like complex composed of the two bi-functional RNases J1 and J2, the endoribonuclease Y, the 3'-5' exoribonuclease PNPase, enolase, phosphofructokinase, the DEAD-box RNA helicase CdhA, and the protein subunit of RNase P, RnpA<sup>9</sup>. Surprisingly, the essential protein RpnA is endowed with a ribonuclease activity and a small molecule inhibitor of RnpA-mediated RNA degradation has been recently selected, which exhibited antimicrobial activity against hypervirulent and multi-drug resistant *S. aureus* strains<sup>11</sup>. Thus, RNA degradation appears to be a cellular process that can be targeted by antimicrobial drugs. Two other endoribonucleases of *S. aureus*, RNase Y and RNase III were shown to be essential for the full virulence of *S. aureus* in several murine models including bacteraemia<sup>12</sup>, intraperitoneal<sup>13</sup>, and peritonitis<sup>14</sup> although they are not essential for growth<sup>12; 14; 15</sup>. Interestingly, both enzymes are important for the processing and stabilization of specific mRNAs, and contribute to the turnover of small non-coding RNAs<sup>12; 16</sup>.

RNase III is one of the extensively *studied RNases* in *S. aureus*. It belongs to the family of Mg<sup>2+</sup>-dependent endoribonucleases in eukaryotes and bacteria, which recognize and cleave double-stranded RNA (dsRNA) to generate a dsRNA with a 5' phosphate group and a two nucleotides 3'-overhang<sup>17</sup>. Bacterial RNase III acts as an active homodimeric enzyme



with each subunit carrying the catalytic activity and substrate recognition <sup>18</sup>. We have previously shown that the biochemical properties of *S. aureus* RNase III are very similar to those of *Escherichia coli* RNase III. As for *E. coli* <sup>19; 20</sup>, conserved acidic residues and Mg<sup>2+</sup> are essential for the cleavage reaction <sup>16; 21</sup>. *S. aureus* RNase III was first shown to act as a co-factor of the quorum-sensing regulatory RNAIII to irreversibly repress the synthesis of several adhesins and the transcriptional repressor of toxins, Rot <sup>15; 22-24</sup>. Moreover, a recent study discovered an antisense transcription all over the genome, which is hidden in *S. aureus* strains due to RNase III processing of sense/antisense transcripts <sup>25</sup>. Finally, a wide range of structured RNAs recognized by RNase III have been identified by *in vivo* immunoprecipitation of wild type or cleavage-defective mutant RNase III followed by deep sequencing <sup>16</sup>. This study revealed that besides its universal function in rRNA maturation and RNA turnover, RNase III affects processing of mRNAs carrying overlapping 5'UTRs. Moreover, it directs maturation of mRNAs resulting in their stabilization and antisense regulation mediated by small RNAs.

We have previously shown that RNase III unexpectedly cleaved several RNAIII-mRNA complexes forming loop-loop interactions <sup>23; 24</sup>. Intriguingly, these short intermolecular base paired helices (7 to 8 nucleotides) share structure and sequence similarities and are *a priori* not considered as RNase III substrates. We demonstrate here that the sequence of these base pairs is not critical for cleavage but that the co-axial stacking of the newly formed intermolecular helices provides an ideal topology for RNase III recognition and cleavage. This work also shows that RNA-RNA complexes contain both positive and negative elements to limit and/or direct the action of RNase III, as it was previously demonstrated for *E. coli* RNase III <sup>26; 27</sup>.

## Results

### Design of the RNA substrates to be probed by RNase III cleavage assays

*S. aureus* RNAIII is a multi-functional RNA that encodes  $\delta$ -hemolysin and regulates the synthesis of many virulence factors and enzymes involved in peptidoglycan metabolism at high cell density <sup>3; 24; 28</sup>. We have previously shown that the large 3' UTR of RNAIII is acting as an antisense RNA (asRNA) to repress the translation of multiple mRNAs <sup>15; 23; 24</sup>. This

region contains three hairpin structures with a conserved C-rich sequence located in the apical loops, which serve as the seed sequence to initiate fast binding with G-rich sequences of mRNA targets<sup>24</sup>. Determination of the structures of several regulatory RNAIII-mRNA complexes revealed bi-partite binding sites that differed according to the mRNA structures: an imperfect duplex that sequestered the Shine and Dalgarno sequence (SD), was stabilized by a loop-loop interaction in the coding region of *coa* mRNA encoding coagulase (Figure 1A), or two loop-loop interactions involving the 5' untranslated region and the SD sequence of *rot* mRNA that encoded the transcriptional repressor of toxins (Figure 1A). These complexes prevent the formation of the initiation ribosomal complex and recruit RNase III to initiate the degradation of the repressed mRNAs. Interestingly, previous RNase III cleavage assays performed on 5' end-labeled mRNAs bound to RNAIII showed a single strong cleavage in *coa* and *rot* mRNAs located at an equivalent position of the loop-loop pairings<sup>23; 24</sup> (Figure 1A). In both systems, the sequence of the loop-loop interactions is highly similar although the length and sequences of the intramolecular helices are different (Figure 1A). The same reaction was performed with 5' end-labeled RNAIII alone and in the presence of increasing concentrations of cold *coa* mRNA or *rot* mRNA (Figure 1B). The data showed that binding of *coa* or of *rot* mRNA similarly induced an RNase III cleavage located in the hairpin loop H7 of RNAIII at C241 (Figure 1). However, the two loop-loop interactions found in RNAIII-*rot* mRNA complex are not strictly equivalent because no cleavage was observed in the hairpin loop H14 of RNAIII (Figure 1B).

Because the sequences involved in the kissing loops were very similar, we have analyzed whether the RNase III cleavage pattern resulted from a specific recognition of the base pairs or from the topology of the kissing complex. To distinguish between these two hypotheses, we have constructed several RNA substrates based on the kissing interactions as found in *coa* mRNA-RNAIII complex and have compared their RNase III cleavage patterns. The sequences of the kissing loops, i.e. loop III of *coa* mRNA and loop 7 of RNAIII, were changed so that the stability of the base pairs was not perturbed (Figure 1A). Structure modeling of the kissing interaction has been previously completed revealing a co-axial stacking of the intramolecular and intermolecular helices to form a long helical structure<sup>24</sup>. Based on this model, we have superimposed a regular helix on the kissing complex to define the appropriate length of the helix for comparative RNase III probing. Two regular helices (30

and 33 base pairs) and a stem-loop structure of 30 base pairs (bp), mimicking the size of the long helical structure promoted by the kissing interactions, were chemically synthesized. Two last RNA constructs introducing a bulged loop were made to interrupt regular helices (Figure 1C).

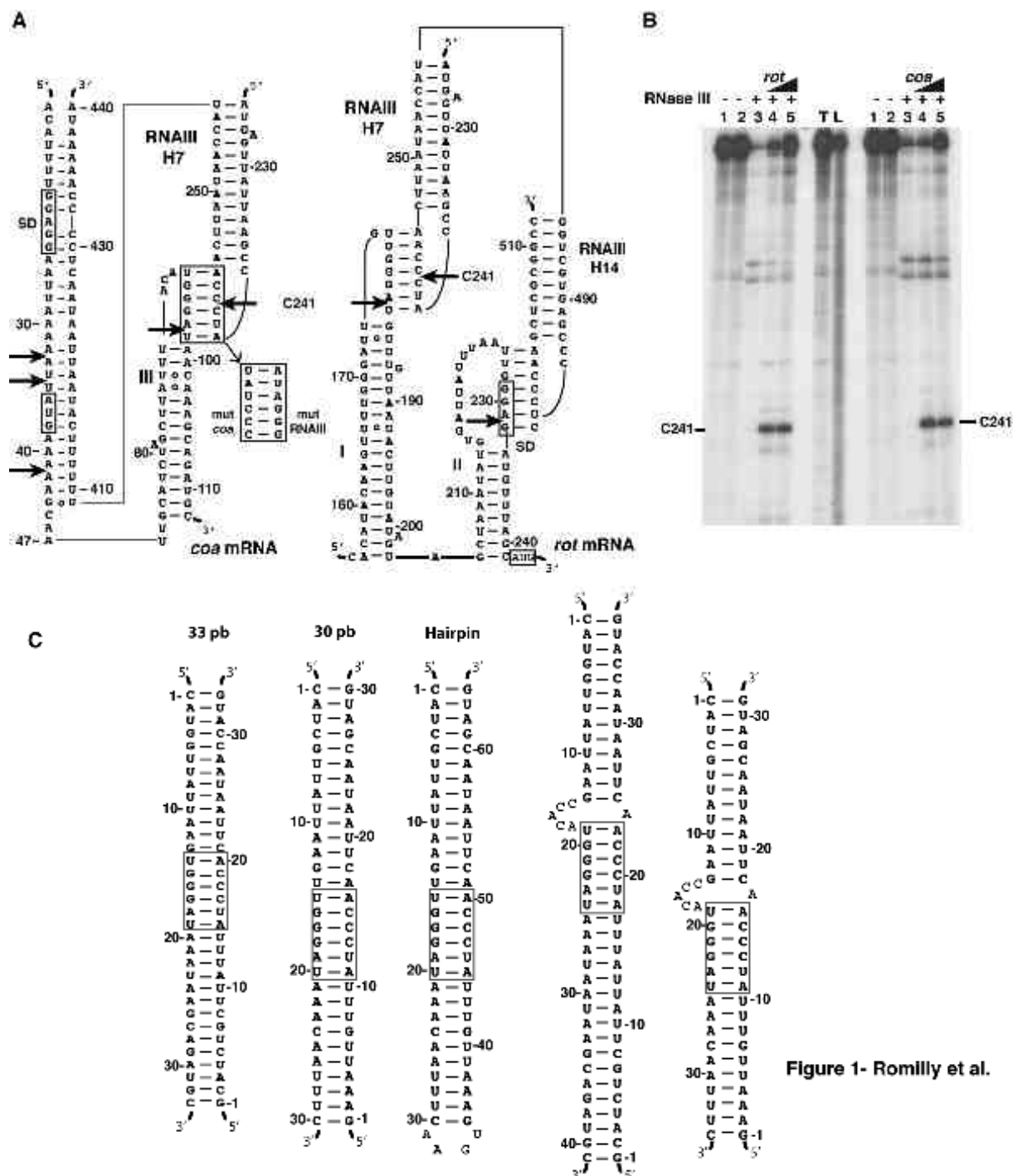


Figure 1- Romilly et al.

**Figure 1: RNA substrates of *S. aureus* RNase III used in this study.** (A) Comparison of the kissing interactions which are part of the RNAIII-*coa* mRNA<sup>24</sup> and RNAIII-*rot* mRNA<sup>23</sup> complexes. Base pair substitutions at the loop-loop interactions, resulting from the pairings of mutant *coa* mRNA and mutant RNAIII, are shown in the insert. Arrows denote the RNase III cleavage sites. (B) RNase III hydrolysis performed on 5' end-labeled RNAIII

free (lane 3) or in complex with either *coa* mRNA or *rot* mRNA (lane 4, 50 nM; lane 5: 100 nM). Lanes 1, 2: Incubation controls of free RNAIII or bound to the mRNA, respectively; lane T: RNase T1 ladder performed under denaturing conditions; lane L: alkaline ladder. (C) Various RNA helices designed in the present study. The boxed sequences correspond to the base pairings found in the kissing interactions of RNAIII-*coa* mRNA complex.

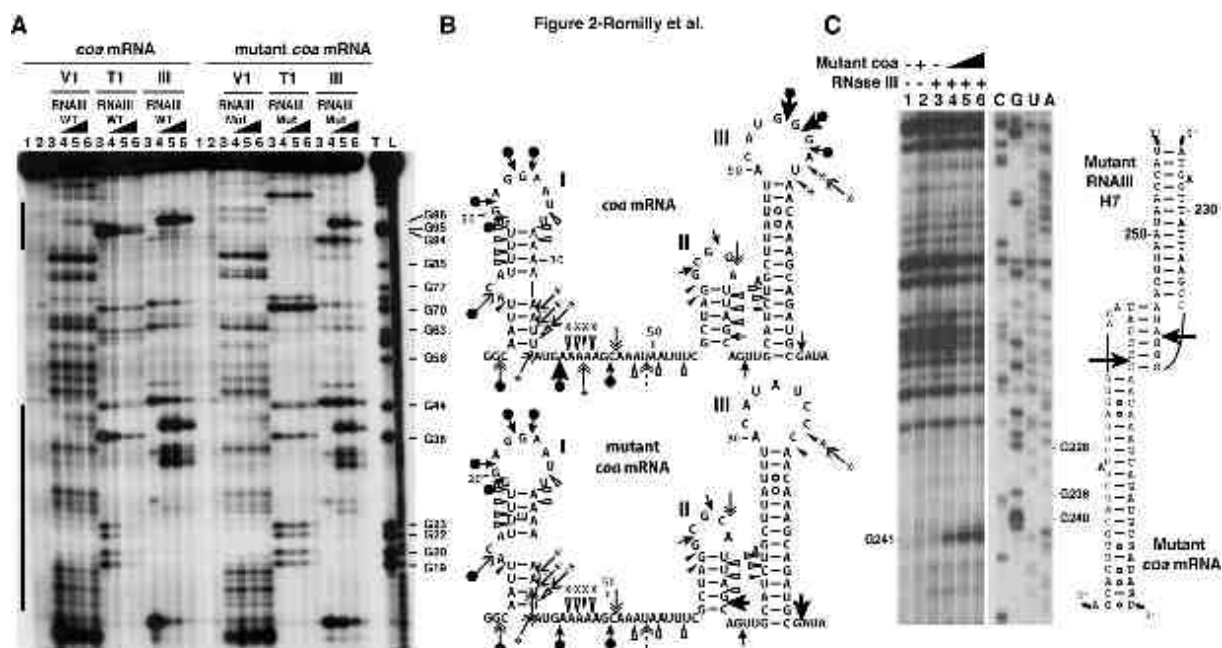
### **Changing the sequence of the loop-loop interaction has no effect on RNase III cleavage**

We first verified that the nucleotide substitution in RNAIII and *coa* mRNA mutants (Figure 1A) did not alter the RNA structure and/or binding. The secondary structures of wild type (WT) and mutant *coa* mRNAs were probed using RNase T1 (specific for unpaired guanines) and RNase V1 (specific for helical regions). The cleavages were mapped using 5' end-labeled mRNA and sized on a denaturing polyacrylamide gel (Figure 2A). The data correlated well with the secondary structure of *coa* mRNA which was previously published<sup>24</sup> (Figure 2A). Indeed, the main RNase T1 cleavages are located in the apical loops I and III and strong RNase V1 cuts occurred in helical regions of *coa* mRNA (Figure 2B). In the mutant mRNA, the major changes occurred in loop III because the three guanines 94-96 have been substituted by cytosines. Binding of mutant RNAIII induced changes in the region encompassing the ribosome binding site (RBS, nucleotides U10 to A48) of the mutant *coa* mRNA, as it was shown for the WT RNAIII-mRNA complex. Mutant RNAIII protected the guanines (G19-20, G22-23) of the SD sequence against RNase T1, while new RNase V1 cleavages were observed at positions 39-41 of the mRNA (Figure 2A). These data show that the mutant RNAIII binds to the RBS of the mutant *coa* mRNA in a manner similar to that of the WT RNAs. In addition, the WT RNAIII binding induced strong protection of G94-G96 in the apical loop III of WT *coa* mRNA towards RNase T1 because of the kissing interaction.

We then compared the RNase III-dependent cleavages using 5'-end labeled WT or mutant mRNAs. The reactions were performed on the free mRNAs as well as on the native WT and mutant RNAIII-*coa* mRNA complexes using a purified His-tagged RNase III from *S. aureus* (Figure 2A). When the 5' end-labeled mutant mRNA was incubated with mutant RNAIII, four major cleavages occurred at positions 32, 34, 41, and 97 in the mRNA as observed for the WT complex (Figures 1A, 2A). These cleavages were specific because the addition of Ca<sup>2+</sup> instead of Mg<sup>2+</sup> in the buffer inactivated RNase III activity (result not shown). Previous experiments have shown that the hairpin 7 of RNAIII was responsible for the specific RNase III-cleavage at position 97 of *coa* mRNA while the hairpin 13 of RNAIII promoted major cleavages at positions 32, 34 and 41 in the RBS of *coa* mRNA<sup>24</sup> (Figures 1A).

We have also monitored the RNase III cleavage sites in RNAIII using primer extension (Figure 2C). The data show that binding of the mutant *coa* mRNA to mutant RNAIII induced a major cleavage at G241 in the hairpin loop 7 of RNAIII (Figure 2C).

In summary, mutations in the sequences of the complementary hairpin loops of RNAIII and *coa* mRNA involved in the kissing interactions do not affect the structure of the mRNA-RNAIII complex nor the recognition by RNase III.



**Figure 2: Comparative enzymatic probing with the WT and mutant RNAIII-*coa* mRNA complexes.** (A) Gel fractionation of enzymatic cleavages of 5'-end-labeled wild-type (WT) and mutant *coa* mRNAs. Lanes 1, 2: Incubation controls of mRNA free or bound to RNAIII, respectively; lane 3: RNase hydrolysis on free WT or mutant *coa* mRNAs; lanes 4-6: RNase hydrolysis on WT or mutant *coa* mRNA in the presence of increasing concentrations of WT or mutant RNAIII (lane 4, 50 nM; lane 5, 100 nM; lane 6, 300 nM), respectively; lanes T, L: RNase T1 under denaturing conditions and alkaline ladders, respectively. T1, V1, III: RNase T1, RNase V1, and RNase III hydrolysis, respectively. (B) Enzymatic cleavages are reported on the schematic representation of the secondary structures of WT and mutant *coa* mRNAs. Black and white triangles are for strong and moderate RNase V1 cuts, respectively. Plain and dashed arrows are for strong and moderate RNase T1 cuts. Double arrows are for RNase III cleavages. Effect of RNAIII binding: black and white circles are for strong and moderate protection, respectively; stars are for enhancement and new RNase III cuts. (C) RNase III cleavage sites were mapped on *in vitro* transcribed mutant RNAIII using primer extension with a 5' end-labeled oligonucleotide (see Material and Methods). Lanes 1, 2: Incubation controls of mutant RNAIII free or bound to mutant *coa* mRNA, respectively; lane 3: RNase hydrolysis on free mutant RNAIII; lanes 4-6: RNase hydrolysis on RNAIII in the presence of increasing concentrations of mutant *coa* mRNA (lane 4, 50 nM; lane 5, 100 nM; lane 6, 300 nM). The cleavage sites (arrows) are shown on the secondary structure model of the loop-loop interactions involving the mutated hairpin loop 7 of RNAIII and the mutated hairpin loop III of *coa* mRNA.

### ***S. aureus* RNase III cleaves regular RNA-RNA helices and RNA hairpins in a similar manner**

We then analyzed the RNase III-dependent cleavages of regular RNA helices of 30 and 33 bp with free ends or terminated by an apical loop (Figure 3). These helices contained the UGGGAU/AUCCCA pairings at an equivalent position to that found in the kissing interaction of the *coa* mRNA-RNAIII complex (Figure 1A). The RNAs have been chemically synthesized and purified on a denaturing polyacrylamide gel. Prior to RNase III probing, we verified by gel retardation assays that the conditions of hybridization between the two complementary RNAs led to efficient duplex formation at a stoichiometry of 1:1 (result not shown). Either of the two RNA fragments was then labeled at the 5' end, and the RNase III cleavage reactions were performed after annealing of the two complementary RNAs, the labeled and the unlabeled strands (Figures 3A, 3C). The same RNase III reaction was performed on a 5' end-labeled RNA hairpin structure (Figure 3B).

Within the 30 bp RNA duplex, two strong RNase III-dependent cleavages were mapped at G18 and A19 of the sense RNA strand, and at C14 and C15 of the asRNA strand, leading to two nucleotides 3'-overhang (Figure 3A). Several additional weak but typical RNase III cleavages were located at G13/U14/U15 on the sense RNA strand and at A17/C18/U19 on the asRNA strand, particularly at high RNase III concentrations (Figure 3A). These cleavages were specific of the duplex formation because they were not observed in the RNA strands alone (Figure 3A, lane 3). The same cleavages were found in the hairpin structure, namely at G18, A19, C47 and C48, revealing that the apical loop did not alter the recognition of the RNA by the enzyme. Interestingly, the strong RNase III cleavages located at A19/C14 in the RNA-RNA duplex and at A19/C48 in the RNA hairpin corresponded to the RNase III cleavages found in the loop-loop interaction. We also analyzed the cleavage pattern on an RNA duplex containing 33 bp. Three main cleavages were obtained at positions U14, A18, and A22 of the sense RNA strand and at positions U13, C17 and C21 of the asRNA strand (Figure 3C). The strongest cleavages (A18/C17) are located in the middle of the helix while the other cleavages (U14/C21, A22/U13) occurred 12 and 11 bp from the 3' end. Thus, the cleavage pattern of the 33 bp helix is somehow different to that obtained for the 30 bp helix although the main cleavages at A18/C17 take place in the other constructs (Figure 3).

Taken together these data showed that the RNA helix structure, i.e. number of base pairs, rather than the sequence direct the RNase III cleavages. In addition, an apical loop

does not affect the RNase III cleavage location in the regular helix. In contrast to the loop-loop complex, the regular helices are cleaved at several positions leading to 11 bp fragments.

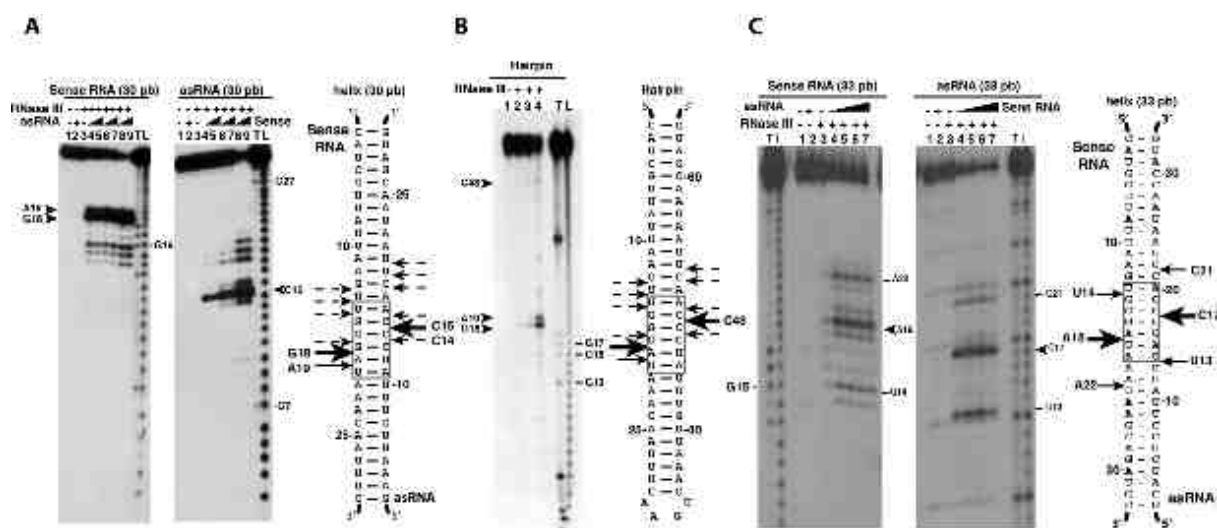


Figure 3- Romilly et al.

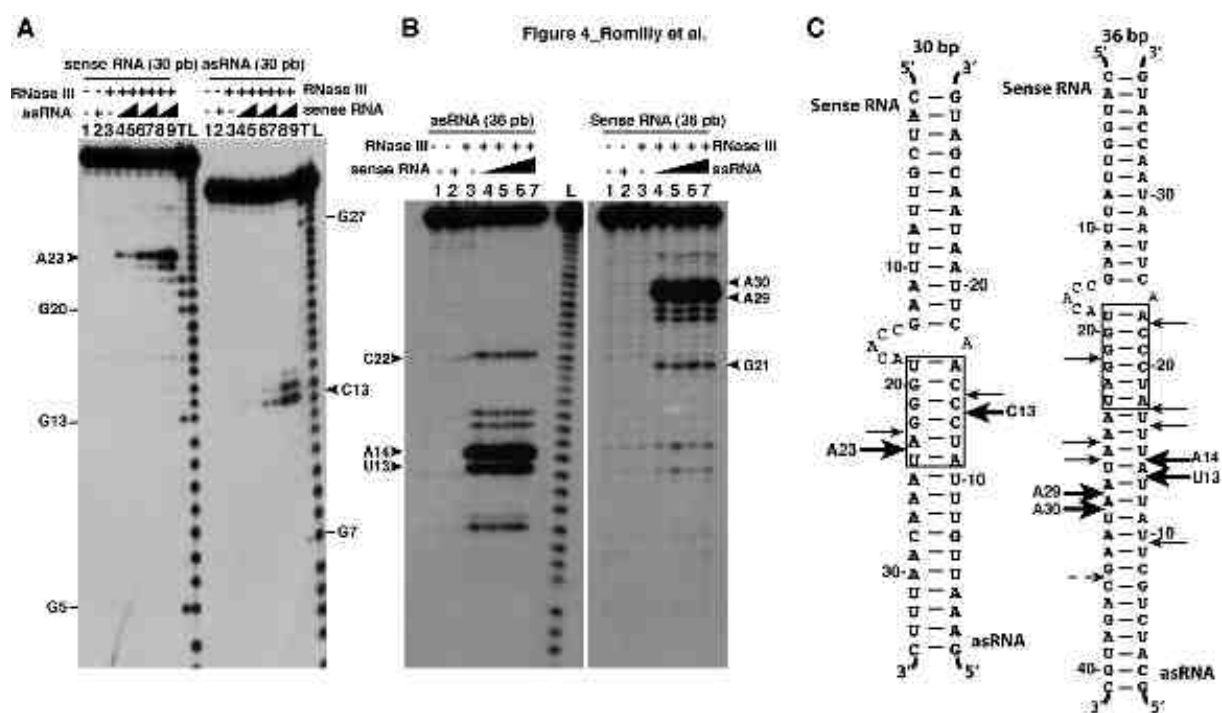
**Figure 3: RNase III cleavage assays with various RNA helices.** (A) RNase III assays with the 30 bp regular helix. Gel fractionation of RNase III cleavages of 5' end labeled sense RNA strand or antisense RNA (asRNA) strand. Duplex formation was done as described in Material and Methods. Lanes 1, 2: Incubation controls of the 5' end labeled RNA free or in duplex, respectively; lane 3: RNase III (3  $\mu$ M) hydrolysis on the 5' end labeled RNA alone; lanes 4-9: RNase III assays performed on the RNA duplex in the presence of increasing concentrations of the complementary strand (lanes 4, 6, 8: 100 nM; lanes 5, 7, 9: 250 nM) and with various concentrations of RNase III (lanes 4, 5: 0,5  $\mu$ M; lanes 6, 7: 1  $\mu$ M; lanes 8, 9: 3  $\mu$ M); lanes T, L: RNase T1 under denaturing conditions and alkaline ladders, respectively. Main RNase III cleavage sites are shown by arrows on one side of the autoradiography. The RNase III cleavages are reported on the RNA secondary structure of the 30 bp helix. Thick, thin and dashed arrows are for strong, moderate, and weak cleavages, respectively. The sense and antisense (asRNA) RNA strands are specified. (B) RNase III assays on the 30 bp helix terminated by a tetraloop. Lane 1: Incubation control of the 5' end labeled RNA free; lanes 2-4: RNase III hydrolysis on the 5' end labeled RNA using increasing concentrations of RNase III (lane 2: 0,5  $\mu$ M ; lane 3: 1  $\mu$ M ; lane 4: 3  $\mu$ M); lanes T, L: RNase T1 under denaturing conditions and alkaline ladders, respectively. The RNase III cleavages are reported on the RNA secondary structure, same legend as above. (C) RNase III assays with the 33 bp regular helix. Lanes 1, 2: Incubation controls of the 5' end labeled RNA free or in duplex, respectively; lane 3: RNase III (1  $\mu$ M) hydrolysis with the 5' end labeled RNA; lanes 4-7: RNase III hydrolysis (1  $\mu$ M) with the RNA duplex in the presence of increasing concentrations of the complementary strand (lane 4: 50 nM; lane 5: 100 nM; lane 6: 200 nM; lane 7: 300 nM). The RNase III cleavages are reported on the RNA secondary structure of the 33 bp helix using the same legend as above.

### A bulged loop restrained the number of RNase III cleavages in the helix

In order to create an irregularity into the regular 30 bp helix, we have introduced a bulged loop 5' to the UGGGAU/AUCCCA base pairings (Figure 4). The RNAs were chemically synthesized and purified on a denaturing polyacrylamide gel and the conditions of hybridization between two complementary RNA strands were optimized (results not shown). After 5' end labeling, the RNase III cleavage reactions were performed on the free RNAs and

the pre-formed RNA duplexes (Figures 4A, 4B). The data revealed that only one strong cleavage occurred at C13 of the asRNA strand and at A23 of the sense RNA strand when the RNA duplex was formed (Figures 4A, 4C). Hence, the addition of the bulged loop restrained the number of RNase III binding sites in comparison to the 30 bp regular helix (Figure 4C).

The same experiment was performed with the 36 bp helix interrupted by the same internal loop (Figure 4B). Two major RNase III cleavages were found at A29 and A30 of the sense RNA strand and at U13 and A14 of the asRNA strand, giving a two nucleotides 3'-overhang (Figure 4C). Other weaker but typical RNase III cleavages were also detected (Figure 4B). The RNase III cleavage patterns obtained with the interrupted 30 bp and 36 bp helices are very different. Thus, increasing the number of consecutive Watson-Crick (WC) base pairings created other preferred RNase III binding sites centered in the helix.



**Figure 4: RNase III cleavage assays with the 30 bp and 36 bp regular helices interrupted by an internal loop.**

(A) Gel fractionation of RNase III cleavages of 5'-end-labeled sense RNA strand or asRNA strand (30 bp irregular helix). Duplex formation was done as described in Material and Methods. Lanes 1, 2: Incubation controls of the 5' end labeled RNA free or in duplex, respectively; lane 3: RNase III hydrolysis (1  $\mu$ M) on the 5' end labeled RNA; lanes 4-9: RNase III hydrolysis on the RNA duplex in the presence of increasing concentrations of the complementary strand (lanes 4, 6, 8: 100 nM; lanes 5, 7, 9: 250 nM) with various concentrations of RNase III (lanes 4, 5: 0.5  $\mu$ M; lanes 6, 7: 1  $\mu$ M; lanes 8, 9: 3  $\mu$ M); lanes T, L: RNase T1 under denaturing conditions and alkaline ladders, respectively. (B) Gel fractionation of RNase III cleavages of 5'-end-labeled sense RNA strand or antisense RNA (asRNA) strand (36 bp irregular helix). Lanes 1, 2: Incubation controls of 5' end labeled RNA free or in duplex, respectively; lane 3: RNase III hydrolysis (1,5  $\mu$ M) on free 5' end labeled RNA; lanes 4-7: RNase III hydrolysis (1,5  $\mu$ M) was performed on the RNA duplex in the presence of increasing concentrations of one of the complementary strand (lane 4: 50 nM; lane 5: 100 nM; lane 6: 200 nM; lane 9: 300 nM); lane L: alkaline



ladder. (C) The RNase III cleavages are reported on the 30 and 36 bp helices interrupted by an internal loop. Thick, thin and dashed arrows are for strong, moderate, and weak cleavages, respectively. The sense and antisense strands are specified. The sequences corresponding to the kissing interactions as found in RNAIII-*coa* mRNA complex are boxed.

### **RNase III does not improve the repressor activity of RNAIII on *coa* mRNA translation**

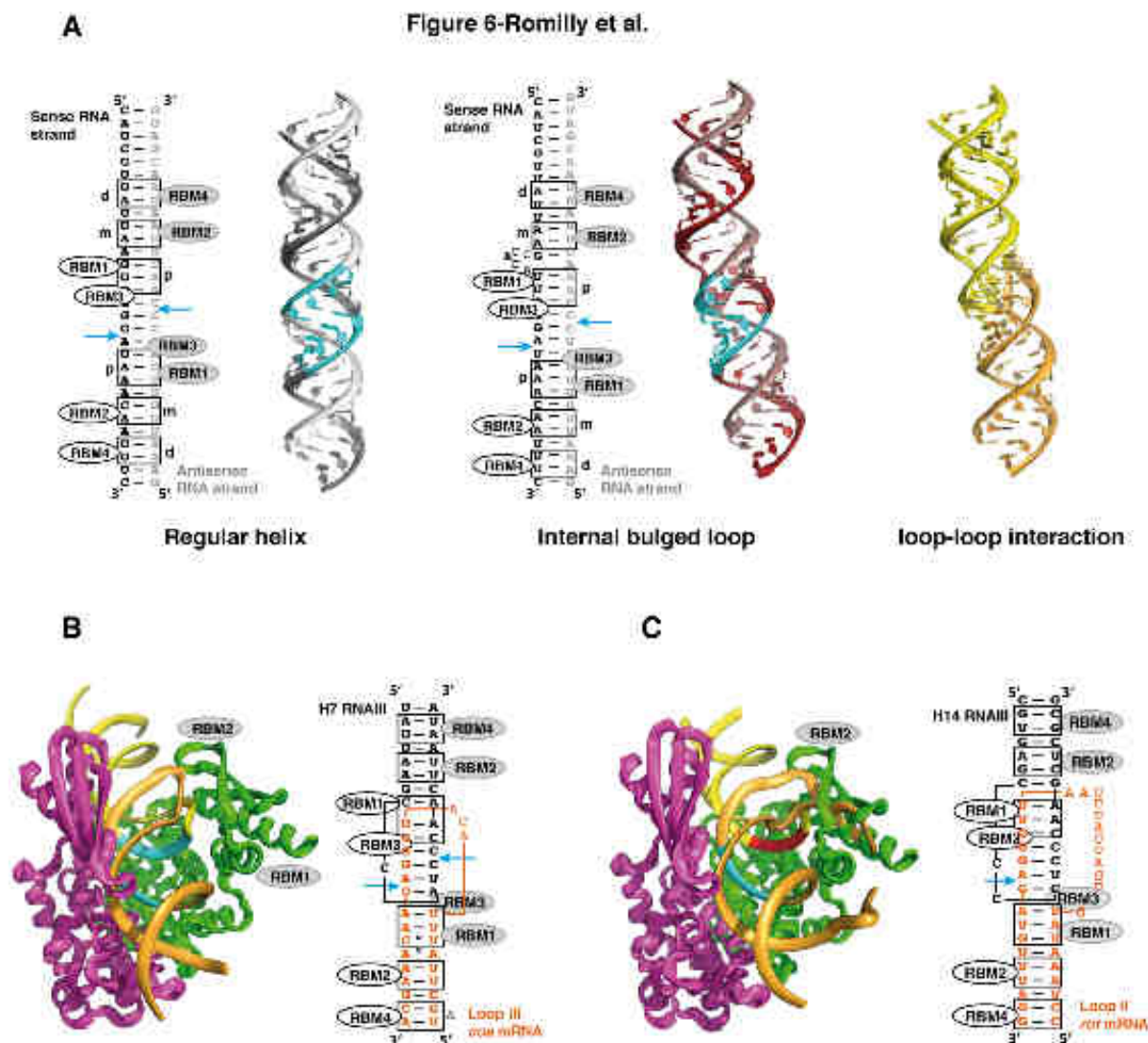
Knowing that RNase III contributed to the efficiency of the RNAIII-dependent repression of *coa* mRNA<sup>24</sup>, we monitored whether the enzyme is able to stabilize the binding of RNAIII to *coa* mRNA and/or to contribute to the efficiency of translation repression. Recent biochemical analysis performed on *S. aureus* RNase III showed that the two conserved acidic residues, E135 and D63, are essential for catalysis, as their substitution by alanine compromised cleavages without affecting RNA binding<sup>16</sup>. These acidic residues are part of the catalytic site, which together with two Mg<sup>2+</sup> ions perform the stepwise hydrolysis of the phosphodiester bonds<sup>29</sup>. The catalytically inactive E135A variant of RNase III carrying a histidine epitope tag at its N-terminus was purified to homogeneity following expression in *E. coli*<sup>16; 21</sup>. We first used gel retardation assays to monitor the effect of E135A mutant enzyme on complex formation between terminally labeled *coa* mRNA and RNAIII. The data showed that the mutant E135A enzyme was able to bind the mRNA-RNAIII complex but had no significant effect on the duplex stability (results not shown).

We then analyzed whether the binding of E135A RNase III to the mRNA-RNAIII complex contributes to the occlusion of the ribosome recruitment to its loading site. Toeprinting assays were performed to monitor the formation of the ternary complex between the *S. aureus* 30S subunits, *coa* mRNA and the initiator tRNA<sup>Met</sup>. A 5' end-labeled primer was first annealed to *coa* mRNA, followed by the addition of increasing concentrations of E135A RNase III in the presence of RNAIII, the 30S subunits and the tRNA. The experiment revealed that RNAIII binding alone was sufficient to prevent efficiently the ribosome binding. However, we have observed the appearance of a reverse transcriptase stop at position G+3 as the result of the interaction between RNAIII and E135A RNase III (Figure 5, lane 4). Because E135A RNase III has lost its catalytic activity<sup>16</sup>, these data suggest that RNase III binding to the RNAIII-*coa* mRNA regulatory complex arrest reverse transcriptase elongation.



4 bp adjacent to the cleavage site and the distal box located 8 bp after the cut. Interestingly, the proximal box contains both positive and negative determinants, which cooperate to define the cleavage efficiency and specificity<sup>26; 27</sup>. The crystal structure of *Aquifex aeolicus* RNase III in complex with different dsRNAs<sup>29; 31</sup> identified a large RNA-binding surface cleft called the catalytic valley, which accommodates the dsRNA. The contacts between protein and RNA span almost 11 bp from the cleavage site, and the hydrolysis of each strand requires the two subunits<sup>26; 29; 31; 32</sup>. Furthermore, the dsRNA and the protein undergo significant conformational changes during catalysis mediated through two Mg<sup>2+</sup> ions<sup>29</sup>. Each subunit of the protein contains four RNA binding motifs (RBM) where RBM1, RBM2, and RBM4 interact with the proximal, middle and distal sites of the helix, respectively, and RBM3 binds closely to the cleavage site (Figure 5A). Emerging from these studies, a model was proposed where the RNA substrate is first selected by RBM1 and RBM2 and then transferred to the catalytic site for scissile bond selection, mainly through RBM3<sup>26; 29; 31; 32</sup>. All these important findings provided some rules explaining how bacterial RNase III selects RNA substrates for binding and cleavage.

Besides regular helices, the *E. coli* enzyme was shown to bind and cleave a variety of structures such as coaxially stacked short helices<sup>33</sup> or dsRNA interrupted by internal loops<sup>16; 34</sup>. We show here that *S. aureus* RNase III cleaves loop-loop interactions at a specific site, which is not strictly dependent on the nature of the WC base pairs (Figures 1, 6). The base pair substitutions introduced at the kissing interactions formed between *coa* mRNA and the quorum sensing regulatory RNAIII had no significant effect on the cleavage reaction and specificity. Therefore, we propose that the loop-loop interaction provides a unique local structure to direct the RNase III cleavages at single positions of the kissing interactions. This was visualized on a three-dimensional model of the kissing interactions, which took into account the chemical probing data<sup>24</sup>. The model revealed a co-axial stacking of the intermolecular and intramolecular helices to form a long helical structure harboring a slight bent (Figure 6A). This RNA structure model was then docked into the crystal structure of the dsRNA-bound form of *Aquifex* RNase III<sup>31</sup>. It shows that the loop-loop interaction almost perfectly fitted into the catalytic valley with the two cleavage sites located on the same face close to the Mg<sup>2+</sup> ion and the conserved acidic residues (Figure 6B).



**Figure 6: Three-dimensional models of various RNA-RNA duplexes and interaction with RNase III.** (A) Models represented the regular 30 bp helix (in grey), the 30 bp helix interrupted by a bulged loop (in red), and the kissing interactions as found in RNAIII (yellow)-*coa* mRNA (orange). The III-D representations were drawn with PYMOL program. The schematic views of the 30 bp RNA duplexes (regular helix, interrupted helix) with the positioning of the RNA binding domains (RBM) of RNase III are based on the crystal structure of *Aquifex aeolicus* RNase III bound to dsRNA<sup>29; 31</sup>: the boxed nucleotides represent the proximal (p), middle (m) and distal (d) boxes from the major cleavage sites that are recognized by RBM1 to RBM4 of RNase III. The RBMs for each RNase III subunit are represented by different colors. (B, C) Docking of the kissing complexes of *coa* mRNA-RNAIII (B) and of *rot* mRNA-RNAIII (C) on the crystal structure of *Aquifex aeolicus* RNase III containing Mg<sup>2+</sup> (from<sup>29</sup>; pdb coordinates 2NUG). The two subunits of RNase III are colored in magenta and green. The hairpin loop of RNAIII is colored in yellow and the mRNA in orange. The cleavage sites are in blue. The site that is not cleaved by RNase III in *rot*-RNAIII kissing interaction is in red. A steric clash is visualized between the connecting loop of the kissing interactions and one of the double stranded RNA binding domain (dsRBD) domain of RNase III. Schematic views represented the kissing interactions as it is found in (B) *coa* mRNA-RNAIII and (C) *rot*-RNAIII duplexes. Same legend as in Fig. 6A. Nucleotides of the hairpin H7 of RNAIII are colored in black and nucleotides of the mRNA hairpin are in orange.

Using the catalytic-defective mutant E135A RNase III<sup>16</sup>, we showed that the mutant enzyme did not improve either the binding of *S. aureus* RNAIII to *coa* mRNA nor the capacity

of RNAIII to prevent the formation of the initiation complex (Figure 5). These *in vitro* data support the hypothesis that *S. aureus* RNase III is required to irreversibly arrest the synthesis of coagulase by initiating the degradation of the repressed mRNA<sup>23</sup>. The loop-loop interaction provides an ideal RNase III binding site to initiate the degradation of structured elements in the mRNA. *S. aureus* RNAIII and RNase III were also shown to conjointly repress the translation of *rot* mRNA<sup>22; 23</sup>. In this system, two stem-loop structures of the regulatory RNA base pair to two complementary hairpin loops within the 5'UTR of *rot* mRNA. A single loop-loop interaction is not sufficient for repression and both kissing interactions contribute to regulation: one contact occludes the RBS while the second one stabilizes the duplex<sup>23</sup>. In addition, *rot* mRNA was processed by RNase III *in vivo* when RNAIII was produced<sup>23</sup>. The sequences of the two loop-loop interactions between *S. aureus* RNAIII and *rot* mRNA are almost identical to that found in RNAIII-*coa* mRNA<sup>22; 23</sup> (Figure 1A). While the two loops of *rot* mRNA were cleaved by RNase III similarly to *coa* mRNA, only the loop of RNAIII which was bound to the SD of *rot* mRNA was efficiently cleaved<sup>23</sup> (Figure 1). What could be the molecular explanation for the different behavior of RNase III on the two kissing interactions? The crystal structure of *Aquifex* RNase III showed that the double-stranded RNA binding domains (dsRBDs) are connected to the ribonuclease domains by flexible linkers<sup>32</sup>. In the structure of the protein bound to an RNA-RNA duplex, the RBM1 and RBM2 interact on the opposite side of the RNA helix making contacts with the sugar-phosphate backbone of the minor groove<sup>29; 31</sup>. The three-dimensional model of the kissing interaction sequestering the SD sequence of *rot* mRNA was built and docked on the crystal structure of the dsRNA-bound form of *Aquifex* RNase III (Figure 6B). The model shows that the large connecting loop crossing the deep major groove of the kissing interactions makes a steric clash with RBM1 of one monomer (Figure 6B). Interestingly, it was proposed that RBM1, which binds 1 bp from the scissile bond, should contribute to the organization of the catalytic site structure<sup>29</sup>. Therefore, we suggest that the large connecting loop acts as a steric anti-determinant for catalysis by altering the positioning of RBM1 and RBM2. Because several small asRNAs and their target RNAs initiate fast binding by defined loop-loop contacts, they create potential substrates for RNase III<sup>35; 36</sup>. One characteristic example is *E. coli* OxyS sRNA, which represses translation of *fhlA* mRNA through base-pairings between two non-contiguous regions involving loop-loop interactions<sup>37</sup>. The sequence of one of the kissing interactions

(AGGGUU/AACCCUU), which sequestered the SD sequence of *fhlA*, is very similar to the base pairings found in *coa*-RNAIII and *rot*-RNAIII complexes suggesting that OxyS-*fhlA* mRNA complex might represent another novel target for *E. coli* RNase III cleavage.

To analyze the influence of the UGGGAU/AUCCCA interaction on RNase III recognition and cleavage, this module has been introduced into regular helices of different sizes, in helices interrupted by bulged loops, and in a stem-loop structure (Figure 1). As described above, the homodimeric RNase III enzyme covers 22 bp of RNA-RNA duplex leading to cleavage staggered by two base pairs after 11 bp on each strand. Statistically, the enzyme can bind at several positions but will cleave only if the helical length requirements are satisfied. With this ruler, all cleavages obtained on the different helices and hairpins can be readily explained. For instance, two major and several weak RNase III cleavage sites on each strand were found in the 30 bp long RNA duplex, leading to RNA fragments of 10-11 bp. However, the major cleavages were obtained within the G-C rich sequence motif indicating that the stacking of nucleotides might locally favor the binding of RNase III. Closing the 30 bp helix with a tetraloop did not modify the reactivity of the enzyme and the cleavage positions were identical to the parental helix. This indicated that the presence of an apical loop did not determine the cleavage site selection as it was shown for the homologous yeast enzyme where the cut is determined by the distance from a tetraloop motif<sup>38</sup>. In contrast, the introduction of an asymmetrical internal loop in the middle of the 30 bp helix did not create a new cleavage site but restrained the number of *S. aureus* RNase III cuts (Figures 1 and 4). Structure modeling suggested that the bulged loop behaves as a negative signature and prevents the RNase III recognition of the upper helix due to steric hindrance. The addition of 3 bp to the 30 bp regular helix or 6 bp to the interrupted helix drastically changed the RNase III cleavage patterns (Figures 3, 4). In both cases, typical and strong RNase III cleavages occurred on both strands located at the middle of the helix. Recent studies based on co-immunoprecipitation assays and deep sequencing identified novel RNA substrates of *S. aureus* RNase III<sup>16</sup>. Several of the RNAs are characterized by hairpin loops that are cleaved on both sides of 22-25 bp long helices at unique positions leading to RNA fragments of 9 to 12 bp with 2 nt 3'-overhang (Figure S1). We did not observe any particular sequence motifs although the stems are A-U/U-A rich and carry several non-canonical base pairs.

In summary, this study shows that loop-loop interactions connected by short unpaired regions represent a unique topology to create a single binding site for *S. aureus* RNase III. This work also stresses the role of RNase III as a co-factor of regulatory RNAs acting as asRNA. As it was shown for the *E. coli* enzyme, the sequence of the helix, the number of consecutive WC including non-canonical bp, and the presence of bulged nucleotides contribute in a positive or negative manner to induce or prevent RNase III cleavages. We anticipate that careful analysis of the whole set of RNA substrates of *S. aureus* RNase III should unravel unexpected topologies recognized by the enzyme that might play important roles in RNA decay, processing or gene regulation.

## Methods and Materials

**RNA preparation.** RNAIII, fragments of *coa* mRNA (including the whole 5'UTR and 126 nt of the coding sequence), and *rot* mRNA (including the whole 5'UTR and 100 nt of the coding sequence) were transcribed *in vitro* using home-made T7 RNA polymerase<sup>39</sup>. Mutations have been introduced in the hairpin loop 7 of RNAIII (AUCCCA245 changed by GGGUA245) and in the coding region of *coa* mRNA (UGGGAU98 changed by UAUCCC98) using the Quickchange XL Site-directed mutagenesis protocol (Stratagene). The oligonucleotides used for the mutagenesis of *coa* mRNA were AGCAGTTGCATCTAGCTTATTTACATATCCCAACAAAGCAGATGCGTAGT and its antisense, and for RNAIII, CTAAAGTATGAGTTATTAAGCCGGGATAACTTAATAACCATGTAAAATT and its antisense; nucleotides in italic represented the mutation sites. The transcribed RNAs were purified by 8% polyacrylamide-8 M urea gel electrophoresis. After elution in 0.5 M ammonium acetate/1 mM EDTA buffer, the RNAs were precipitated twice with ethanol. Before use, the pellets were dissolved in sterile bi-distilled water and the concentration was measured accurately. The various RNA fragments were chemically synthesized by Integrated DNA Technologies (Belgium). The RNAs were purified on HPLC and their quality was verified on a 15% polyacrylamide-8 M urea gel electrophoresis.

The 5' end-labeling of dephosphorylated transcribed RNA or of the chemically synthesized RNA was performed with T4 polynucleotide kinase and [ $\gamma$ -<sup>32</sup>P]ATP. Before use, RNAs were renatured by incubation at 90°C for 2 min in water, 1 min on ice, followed by an

incubation step at 20°C for 15 min in TMN buffer (20 mM Tris-acetate pH 7.5, 10 mM magnesium-acetate, 150 mM Na-acetate).

Complex formation between sense and asRNAs was formed under denaturing and native conditions. The 5' end-labeled or unlabeled RNA ( $3 \times 10^{-8}$  M) was mixed with a fivefold excess of unlabeled complementary RNA and duplexes were formed by incubation of the two RNAs at 90°C for 2 min followed by slow cooling to 37°C in TMN buffer. Duplex formation was monitored on gel retardation assays using 8% polyacrylamide gel electrophoresis under non-denaturing conditions.

**RNase III purification.** Wild-type and mutant E135A RNase III were cloned in vectors pQE30 derivatives and the plasmids were transformed into *E. coli* strain M15[pREP4]. Overexpression and purification of the two enzymes carrying 6 histidines at their N-terminus were done as described previously<sup>16</sup>. After cell growth and grinding, the soluble fraction of the protein was precipitated by gentle addition of 4.58 g of ammonium sulfate (2 M final concentration) for 1 h on ice, and centrifuged (30 min, 10000 rpm, 4°C). The pellet was dissolved in 10 ml of 25 mM Tris-HCl (pH 8.0), 0.1 mM EDTA, and was incubated with the Ni<sup>2+</sup>-beads (Qiagen) for 1 h at 4°C. After successive washing steps, the enzymes were eluted using 500 mM of imidazole in 25 mM of Tris-HCl (pH 8), 1 M NH<sub>4</sub>Cl, 1 mM DTT. The fractions containing the protein were then directly loaded on a polyI-polyC chromatography (Pharmacia) pre-equilibrated with 25 mM of Tris-HCl (pH 8), 1 M NH<sub>4</sub>Cl, 1 mM DTT and the elution is done in the same buffer containing 2 M NH<sub>4</sub>Cl. The purified WT and E135A RNase III were then dialyzed at 4°C against 30 mM Tris-HCl pH 8.0, 500 mM KCl, 0.1 mM DTT, 0.1 mM EDTA, 50% bidistilled glycerol. Mass spectrometry analysis was carried out on the purified enzyme.

**RNase III cleavage assays.** The RNase III cleavage assay was done either on 5' end-labeled (50000 cpm) or unlabeled RNA (1-2 pmol). Reaction was carried out with purified RNase III (0.5-2 μM), 1 μg of yeast total tRNA in 10 mM Tris-HCl, pH 8, 10 mM MgCl<sub>2</sub>, 100 mM KCl, DTT 1 mM for 5 min at 37°C.

Primer extension was used to assign the RNase III cleavages on unlabeled RNA. The cleaved RNA was hybridized with a 5'-labeled oligonucleotide complementary to the hairpin



loop of H7 of RNAIII (0.4 pmol; ATTATCTGTAATGATAATTAAGAA) as follows: 90°C for 1 min, 1 min on ice and then incubated at 20°C for 15 min in 50 mM Tris-HCl pH 7.5, 20 mM MgCl<sub>2</sub>, 50 mM KCl. Reverse transcription was performed with 0.3 mM dNTPs and 0.1 U/μl RT (AMV) at 37°C for 30 min. After destruction of the RNA template, samples were run on 8% polyacrylamide-7 M urea gels. Sequencing reactions were done in parallel on the RNA transcripts. When 5' end-labeled RNA was used, the cleavages were assigned using RNase T1 hydrolysis performed under denaturing conditions and an alkaline ladder.

**Enzymatic structure probing.** The secondary structure of 5' end-labeled WT or mutant *coa* mRNA free (1-2 pmol) or bound to WT or mutant RNAIII (50-250 nM) was probed by RNase T1 and RNase V1. Complex formation was carried out on previously renatured RNAs at 37°C for 15 min in 10 μl of buffer containing 10 mM Tris-HCl, pH 8, 10 mM MgCl<sub>2</sub>, 100 mM KCl, 1 mM DTT. Enzymatic hydrolysis were performed at 37°C for 5 min in the presence of 1 μg carrier tRNA with RNase T1 (0.0025 U) or RNase V1 (0.1 U). Reactions were stopped by phenol/chloroform extraction followed by ethanol precipitation. Incubation controls were done in parallel to detect nicks in the RNA and the cleavages were assigned using RNase T1 hydrolysis under denaturing conditions and an alkaline ladder.

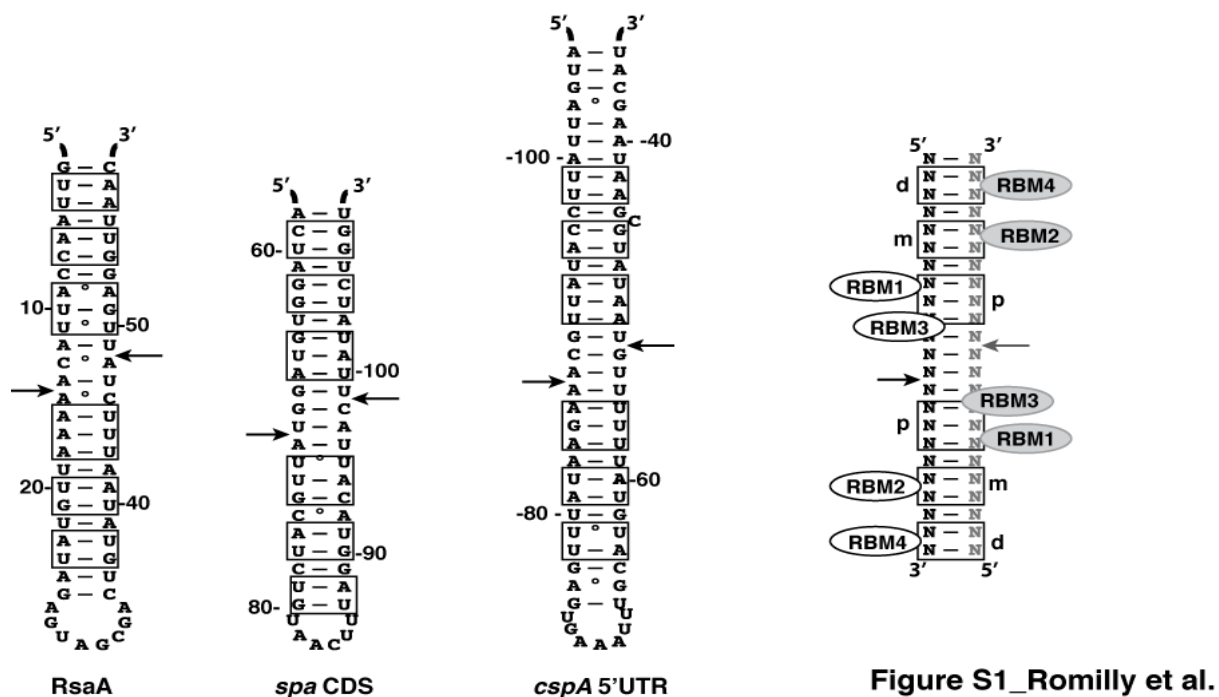
**Toeprinting assays.** The formation of a simplified translational initiation complex with *coa* mRNA was done using published procedure<sup>40</sup>. *Coa* mRNA was annealed to the 5' end-labeled oligonucleotide (complementary to nucleotides 99 to 117) in 20 mM Tris-acetate pH 7.5, 60 mM NH<sub>4</sub>Cl, 3 mM β-mercaptoethanol, heated at 90°C for 1 min and cooled on ice for 1 min. Mg-acetate was added at 10 mM final concentration and incubation was continued at 20°C for 20 min. Initiation complex formation was performed at 37°C for 15 min in 10 mM Tris-acetate pH 7.4, 60 mM NH<sub>4</sub>Cl, 10 mM Mg-acetate, 6 mM β-mercaptoethanol in the presence of *coa* mRNA annealed to the labeled primer, 0.1-0.5 μM 30S subunits and 1 μM initiator tRNA<sup>fMet</sup>. Primer extension reactions were subsequently performed by adding 2-4 units of AMV reverse transcriptase at 37°C for 15 min. Reactions were stopped by phenol extraction followed by ethanol precipitation, and samples were loaded on 8% polyacrylamide-8 M urea slab gels.

**Molecular modeling.** Modeling of the kissing interactions of the *coa* mRNA (residues U73 to A114) with the hairpin loop 7 of RNAlII (residues A223 to U256), of the hairpin loop containing the SD sequence of *rot* mRNA (residues G205 to C240) bound to the hairpin loop 14 of RNAlII (residues G484 to C512), of a 30 bp regular helix, and of a 30 bp helix interrupted by a bulged loop were carried out as previously described<sup>41; 42</sup>. After the interactive assembly step, several cycles of geometrical least-square refinements were performed until a satisfactory solution was reached. The III-D models of the kissing interactions were docked on the crystal structure of *Aquifex aeolicus* RNase III containing Mg<sup>2+</sup>. The pdb coordinates (2NUG) corresponded to the active form of the protein complexed to a cleaved dsRNA<sup>29</sup>. The pictures were prepared using the PYMOL program (DeLano WL, The PyMOL Molecular Graphics System 2002; <http://www.pymol.org>).

**Acknowledgments.** We are thankful to E. Lioliou, I. Caldelari, P. Fechter, C. Condon and S. Boisset for helpful discussions and careful reading of the manuscript. We are grateful to E. Lioliou who has prepared the E135A RNase III mutant. This work was supported by the Centre National de la Recherche Scientifique (CNRS, PR), Université de Strasbourg (UdS, EW), the Institut National pour la Recherche Médicale (INSERM, FV, TG), the Agence Nationale pour la Recherche (ANR: ANR09-BLAN-0024-01; ANR10-Pathogenomics-ARMSA; PR & FV), and the labex NetRNA (ANR-10-LABX-36, EW & PR). CR was supported by fellowships from Délégation Générale de l'Armement (DGA), Région Alsace, and the Fondation pour la Recherche Médicale.

The authors declare that they have no competing interests.

### Supplementary material



**Figure S1: Examples of various RNA substrates of *S. aureus* RNase III** (from Lioliou et al. <sup>16</sup>). RNase III affected the turnover of the non coding RNA RsaA and of *spa* mRNA encoding protein A by inducing cleavages in the hairpin structures of both RNAs. RNase III-dependent cleavages in the 5' untranslated region (5'UTR) of *cspA* mRNA, encoding the major cold shock protein CspA, led to a more stable mRNA. For comparison, a regular helix is drawn with the three boxes of nucleotides (P for proximal, m for middle, and d for distal) that are recognized by specific RNA binding motif of RNase III (RBM1 to RBM4). These three boxes are also drawn on the three hairpin motifs. Arrows denote the RNase III cleavages.

## References

1. Dinges MM, Orwin PM and Schlievert PM. Exotoxins of *Staphylococcus aureus*. Clinical Microbiology Reviews 2000; 13: 16-34.
2. Foster TJ and Höök M. Surface Protein Adhesins of *Staphylococcus aureus*. Trends Microbiol. 1998; 6: 484-88.
3. Novick RP. Autoinduction and Signal Transduction in the Regulation of Staphylococcal Virulence. Mol. Microbiol. 2003; 48: 1429-49.
4. Novick RP and Geisinger E. Quorum Sensing in Staphylococci. Annu. Rev. Genet. 2008; 42: 541-64.
5. Anderson KL and Dunman PM. Messenger RNA Turnover Processes in *Escherichia Coli*, *Bacillus Subtilis*, and Emerging Studies in *Staphylococcus aureus*. Int J Microbiol 2009; 2009: 525491.

6. Anderson KL, Roberts C, Disz T, Vonstein V, Hwang K, Overbeek R, Olson PD, Projan SJ and Dunman PM. Characterization of the *Staphylococcus aureus* Heat Shock, Cold Shock, Stringent, and Sos Responses and Their Effects on Log-Phase mRNA Turnover. *J. Bacteriol.* 2006; 188: 6739-56.
7. Anderson KL, Roux CM, Olson MW, Luong TT, Lee CY, Olson R and Dunman PM. Characterizing the Effects of Inorganic Acid and Alkaline Shock on the *Staphylococcus aureus* Transcriptome and Messenger RNA Turnover. *FEMS Immunol Med Microbiol.* 2010; 60: 208-50.
8. Jester B, Romby P and Lioliou E. When Ribonucleases Come Into Play in Pathogens: A Survey of Gram-Positive Bacteria. *Int J Microbiol* 2012; in press.
9. Roux CM, DeMuth JP and Dunman PM. Characterization of Components of the *Staphylococcus aureus* mRNA Degradosome Holoenzyme-Like Complex. *J. Bacteriol.* 2011; 193: 5520-26.
10. Kaberdin VR, Singh D and Lin-Chao S. Composition and Conservation of the mRNA-Degrading Machinery in Bacteria. *J. Biomed. Sci.* 2012; 18: 23-35.
11. Olson PD, Kuechenmeister LJ, Anderson KL, Daily S, Beenken KE, Roux CM, Reniere ML, Lewis TL, Weiss WJ, Pulse M, Nguyen P, Simecka JW, Morrison JM, Sayood K, Asojo OA, Smeltzer MS, Skaar EP and Dunman PM. Small Molecule Inhibitors of *Staphylococcus aureus* RnpA Alter Cellular mRNA Turnover, Exhibit Antimicrobial Activity, and Attenuate Pathogenesis. *PLoS Pathog* 2011; 7: e1001287.
12. Marincola G, Schäfer T, Behler J, Bernhardt J, Ohlsen K, Goerke C and Wolz C. RNase Y of *Staphylococcus aureus* and Its Role in the Activation of Virulence Gene. *Mol. Microbiol.* 2012; doi: 10.1111/j.1365-2958.2012.08144:
13. Kaito C, Kurokawa K, Matsumoto Y, Terao Y, Kawabata S, Hamada S and Sekimizu K. Silkworm pathogenic bacteria infection model for identification of novel virulence genes. *Mol. Microbiol.* 2005; 56: 934-944.
14. Liu Y, Dong J, Wu N, Gao Y, Zhang X, Mu C, Shao N, Fan M and Yang G. The Production of Extracellular Proteins is Regulated By Ribonuclease III Via Two Different Pathways in *Staphylococcus Aureus*. *PLoS One* 2011; 6: e20554.
15. Huntzinger E, Boisset S, Saveanu C, Benito Y, Geissmann T, Namane A, Lina G, Etienne J, Ehresmann B, Ehresmann C, Jacquier A, Vandenesch F and Romby P. *Staphylococcus*

- aureus* RNAIII and the Endoribonuclease III Coordinately Regulate *spa* Gene Expression. EMBO J. 2005; 24: 824-35.
16. Lioliou E, Sharma CM, Caldelari I, Helfer AC, Fechter P, Vandenesch F, Vogel J and Romby P. Global Regulatory Functions of the *Staphylococcus aureus* Endoribonuclease III in Gene Expression. PLoS Genet 2012; 8: e1002782.
  17. MacRae IJ and Doudna JA. Ribonuclease Revisited: Structural Insights Into Ribonuclease III Family Enzymes. Curr Opin Struct Biol 2007; 17: 138-45.
  18. Conrad C, Schmitt JG, Evguenieva-Hackenberg E and Klug G. One Functional Subunit is Sufficient for Catalytic Activity and Substrate Specificity of *Escherichia Coli* Endoribonuclease III Artificial Heterodimers. FEBS Lett. 2002; 518: 93-96.
  19. Dasgupta S, Fernandez L, Kameyama L, Inada T, Nakamura Y, Pappas A and Court DL. Genetic Uncoupling of the Dsrna-Binding and RNA Cleavage Activities of the *Escherichia Coli* Endoribonuclease RNase III--the Effect of dsRNA Binding on Gene Expression. Mol. Microbiol. 28: 629-40.
  20. Sun W, Jun E and Nicholson AW. Intrinsic Double-Stranded-RNA Processing Activity of *Escherichia Coli* Ribonuclease III Lacking the DsRNA-Binding Domain. Biochemistry 40: 14976-84.
  21. Chevalier C, Huntzinger E, Fechter P, Boisset S, Vandenesch F, Romby P and Geissmann T. *Staphylococcus aureus* Endoribonuclease III Purification and Properties. Methods Enzymol. 2008; 447: 309-27.
  22. Geisinger E, Adhikari RP, Jin R, Ross HF and Novick RP. Inhibition of Rot Translation By RNAIII, a Key Feature of Agr Function. Mol. Microbiol. 2006; 61: 1038-48.
  23. Boisset S, Geissmann T, Huntzinger E, Fechter P, Bendridi N, Possedko M, Chevalier C, Helfer AC, Benito Y, Jacquier A, Gaspin C, Vandenesch F and Romby P. *Staphylococcus aureus* RNAIII Coordinately Represses the Synthesis of Virulence Factors and the Transcription Regulator Rot By an Antisense Mechanism. Genes Dev 2007; 21: 1353-66.
  24. Chevalier C, Boisset S, Romilly C, Masquida B, Fechter P, Geissmann T, Vandenesch F and Romby P. *Staphylococcus aureus* RNAIII Binds to Two Distant Regions of Coa mRNA to Arrest Translation and Promote Mrna Degradation. PLoS Pathog 2010; 6: e1000809.
  25. Lasa I, Toledo-Arana A, Dobin A, Villanueva M, de Los Mozos IR, Vergara-Irigaray M, Segura V, Fagegaltier D, Penades JR, Valle J, Solano C and Gingeras TR. Genome-Wide

- Antisense Transcription Drives mRNA Processing in Bacteria. *Proc Natl Acad Sci U S A* 2011; 108: 20172-77.
26. Pertzev AV and Nicholson AW. Characterization of RNA Sequence Determinants and Antideterminants of Processing Reactivity for a Minimal Substrate of *Escherichia Coli* Ribonuclease III. *Nucleic Acids Res.* 34: 3708-21.
  27. Zhang K and Nicholson AW. Regulation of Ribonuclease III Processing By Double-Helical Sequence Antideterminants. *Proc Natl Acad Sci U S A* 1997; 94: 13437-41.
  28. Felden B, Vandenesch F, Bouloc P and Romby P. The *Staphylococcus aureus* Rnome and Its Commitment to Virulence. *PLoS Pathog* 2011; 7: e1002006.
  29. Gan J, Shaw G, Tropea JE, Waugh DS, Court DL and Ji X. A Stepwise Model for Double-Stranded RNA Processing By Ribonuclease III. *Mol. Microbiol.* 2008; 67: 143-54.
  30. Robertson HD. *Escherichia Coli* Ribonuclease III Cleavage Sites. *Cell* 1982; 30: 669-72.
  31. Gan J, Tropea JE, Austin BP, Court DL, Waugh DS and Ji X. Structural Insight Into the Mechanism of Double-Stranded RNA Processing By Ribonuclease III. *Cell* 2006; 124: 355-66.
  32. Blaszczyk J, Gan J, Tropea JE, Court DL, Waugh DS and Ji X. Noncatalytic Assembly of Ribonuclease III With Double-Stranded RNA. *Structure* 2004; 12: 457-66.
  33. Franch T, Thisted T and Gerdes K. Ribonuclease III Processing of Coaxially Stacked RNA Helices. *J. Biol. Chem.* 1999; 274: 26572-78.
  34. Calin-Jageman I and Nicholson AW. Mutational Analysis of an RNA Internal Loop as a Reactivity Epitope for *Escherichia Coli* Ribonuclease III Substrates. *Biochemistry* 2003; 42: 5025-34.
  35. Wagner EG, Altuvia S and Romby P. Antisense RNAs in Bacteria and Their Genetic Elements. *Adv. Genet.* 2002; 46: 361-98.
  36. Brantl S. Regulatory Mechanisms Employed By Cis-Encoded Antisense RNAs. *Curr. Opin. Microbiol.* 10: 102-09.
  37. Agarman L and Altuvia S. *FhlA* Repression By Oxys RNA: Kissing Complex Formation At Two Sites Results in a Stable Antisense-Target RNA Complex. *J. Mol. Biol.* 300: 1101-12.
  38. Chanfreau G, Buckle M and Jacquier A. Recognition of a Conserved Class of RNA Tetraloops By *Saccharomyces Cerevisiae* RNase III. *Proc Natl Acad Sci U S A* 97: 3142-47.

39. Benito Y, Kolb FA, Romby P, Lina G, Etienne J and Vandenesch F. Probing the Structure of RNAlII, the *Staphylococcus aureus* Agr Regulatory RNA, and Identification of the RNA Domain Involved in Repression of Protein A Expression. *RNA* 2000; 6: 668-79.
40. Fechter P, Chevalier C, Yusupova G, Yusupov M, Romby P and Marzi S. Ribosomal Initiation Complexes Probed By Toeprinting and Effect of Trans-Acting Translational Regulators in Bacteria. *Methods Mol. Biol.* 2009; 540: 247-63.
41. Westhof E, Masquida B and Jossinet F. Prediction and Modeling RNA Architecture. *Cold Spring Harb Perspect Biol* 2011; 3: pii: a003632. doi: 10.1101/cshperspect.a003632.
42. Jossinet F, Ludwig TE and Westhof E. Assemble: An Interactive Graphical Tool to Analyze and Build RNA Architectures At the 2d and 3d Levels. *Bioinformatics* 2010; 26: 2057-59.

# Etudes de la fonction et du mécanisme d'action d'ARN régulateurs chez *Staphylococcus aureus*

**Romilly C., Lays C., Lioliou E., Boisset S., Vandenesch F. Geissmann T. & Romby P. *Staphylococcus aureus* RsaA represses translation of the global regulator MgrA, in preparation.**



Avant mon arrivée au laboratoire, une étude prédictive, réalisée par Tom Geissmann en collaboration avec Christine Gaspin, a conduit à l'identification de nouveaux ARN régulateurs exprimés au sein des IGR de *S. aureus*. Les paramètres utilisés dans cette étude se distinguent en trois catégories : (i) analyse du degré de conservation des IGR d'une taille supérieure à 100 bases chez les firmicutes ou chez les Staphylococceae (genre *Macrococcus* et *Staphylococcus*) ; (ii) prédiction des régions structurées et riches en GC en discriminant les ARN déjà connus (ARNt, ARNr) issus des banques de données RFAM et GeneBank et (iii) prédiction de motifs structuraux tels que des terminateurs de transcription orphelins et recherche des séquences consensus (boîte -10 et -35) des promoteurs  $\sigma^A$  et  $\sigma^B$ .

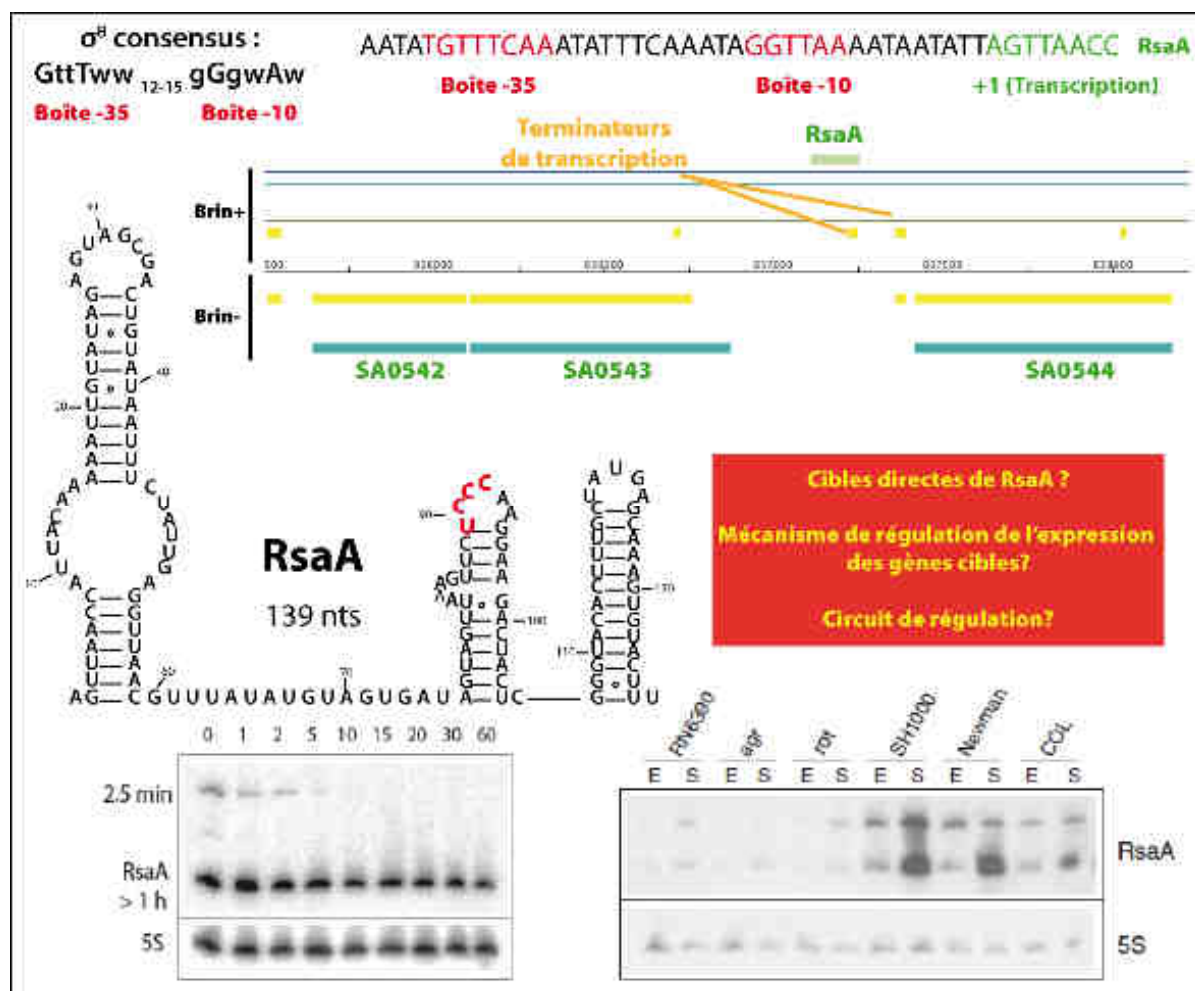
L'attention s'est portée sur les candidats présentant au moins deux des trois critères décrits ci-dessus. L'expression de 36 d'entre eux a ensuite été testée par northern dans différentes conditions de culture (BHI (milieu riche), NZM (milieu minimum), NaCl 15% (choc osmotique), paraquat ou H<sub>2</sub>O<sub>2</sub> (stress oxydatif), dipyrindole (chélateur du fer), pH) et dans diverses souches de *S. aureus* (RN6390, SH1000, RN1HG, Newman et COL) pour aboutir à la validation de 11 nouveaux ARN régulateurs.

## II. RsaA, un ARN régulateur appartenant au régulon

### $\sigma^B$

RsaA est exprimé abondamment chez les bactéries exprimant  $\sigma^B$  (Newman, COL, SH1000 et RN1HG) et s'accumule majoritairement en phase stationnaire de croissance et en réponse à un choc osmotique (15% de NaCl). La séquence consensus du promoteur  $\sigma^B$  est retrouvée en amont du TSS de l'ARN (Figure 32). RsaA fait donc partie du régulon  $\sigma^B$  et pourrait être son bras répresseur. Sa transcription aboutit à deux produits : un ARN majoritaire et stable de 139 nts (demi-vie supérieure à une heure) et un ARN de 276 nts beaucoup moins exprimé et plus instable, résultant de l'arrêt de la transcription à un second terminateur (demi-vie de l'ordre de la minute). Des expériences de cartographie en solution montrent que la version courte de RsaA est très structurée et possède trois tiges-boucles stables, dont l'une correspond au terminateur de transcription Rho-indépendant. Cette structure secondaire protège RsaA contre la dégradation par les exoribonucléases (Figure 32). En revanche, des prédictions de structures secondaires (contraFOLD et mFOLD) sur la version longue de RsaA (276 nts) montrent que l'extension de 137 nts est peu structurée en dehors du terminateur de transcription (données non montrées). La RNase III (Lioliou *et al.*, 2012) et la RNase Y (Marincola *et al.*, 2012) sont toutes deux impliqués dans la dégradation de RsaA. Le temps de demi vie de cet ARN est fortement perturbé dans les souches n'exprimant pas soit la RNase III, soit la RNase Y. Une seconde caractéristique particulière de RsaA est que ce dernier possède un motif riche en C, conservé dans plusieurs souches de *S. aureus*, similaire à celui porté par les tiges boucles 7, 13 et 14 de l'ARNIII. De manière intéressante, ce motif est retrouvé dans plusieurs ARN régulateurs de *S. aureus* et pourrait distinguer une classe d'ARN dédiée à la répression de la traduction des ARNm cibles.

**Le motif riche en C est-il le domaine fonctionnel de RsaA ? Quels sont ses cibles directes et ses réseaux de régulation ?**



**Figure 33 : RsaA, nouveau ARN régulateur chez *S. aureus*.**

RsaA est localisé dans la région intergénique entre les gènes SA0543 et SA0544 (image d'une capture d'écran du logiciel de visualisation génomique Apollo). La séquence consensus du promoteur  $\sigma^B$  (boîte -10 et -35 en rouge) est trouvée en amont du TSS (+1 en vert). La structure secondaire de RsaA a été obtenue par des expériences de cartographie en solution (Clément Chevalier) et le motif riche en C est en rouge. RsaA est stable (demi-vie supérieure à 1h) et s'accumule majoritairement en phase stationnaire de croissance (Geissmann *et al.*, 2009)

$\sigma^B$  est un facteur de transcription qui forme une holoenzyme avec la RNAP pour la guider vers des promoteurs spécifiques. La séquence consensus des promoteurs  $\sigma^B$  DGWKTNDN<sub>12-15</sub>GGRWAW (D = A, G ou T ; W = A ou T ; K = G ou T et R = A ou G, N = A, C, G ou T) est partagée entre *B. subtilis* et *S. aureus* (Petersohn *et al.*, 1999). Chez *B. subtilis* et *S. aureus*,  $\sigma^B$  est séquestré par RsbW pendant la croissance bactérienne. RsbW possède une activité kinase maintenant RsbV, son antagoniste, dans un état inactif phosphorylé. Les voies d'activation divergent entre les deux bactéries. Chez *B. subtilis*, RsbV est déphosphorylée par l'action de phosphatases induites dans deux cas de figure. Lors d'un stress énergétique (diminution des ressources en ATP), RsbQ active la phosphatase RsbP. La seconde

phosphatase, RsbU, est activée lors d'un stress environnemental (stress alcalin) *via* un complexe multi-protéique, le stressosome (Delumeau *et al.*, 2006). RsbU et RsbP, sous leurs formes actives, déphosphorylent RsbV qui se liera à RsbW, permettant ainsi la libération de  $\sigma^B$  et en conséquence la transcription des gènes appartenant à son régulon (revue Marles-Wright *et al.*, 2007). Chez *S. aureus*, l'activation de  $\sigma^B$  a évolué différemment. Seule la voie d'activation par *rsbU* est retrouvée (revue Hecker *et al.*, 2007). De plus, la délétion de ce dernier n'inactive pas totalement l'expression de  $\sigma^B$  dans la souche FDA486, laissant supposer que d'autres facteurs seraient impliqués dans l'activation de  $\sigma^B$  (Palma *et al.*, 2001). Chez *S. aureus*, l'ARNm codant pour  $\sigma^B$  s'accumule en phase stationnaire de croissance ou en réponse à un choc thermique, salin et alcalin (Pané-Farré *et al.*, 2006).

L'analyse du transcriptome des souches Newman, COL et GP268 et de leurs mutants isogéniques ( $\Delta$ *rsbUVW-sigB*) montre que  $\sigma^B$  régule l'expression de gènes impliqués dans de nombreux de processus cellulaires : métabolisme, virulence, transport et paroi (Bischoff *et al.*, 2004). De manière générale, l'expression de  $\sigma^B$  est antagoniste au système *agr* : elle résulte en une inhibition des exoprotéines au profit de la synthèse des gènes d'adhésion. En effet, l'expression de  $\sigma^B$  modulerait les niveaux d'expression du système *agr* et de l'ARNIII : la comparaison entre la souche NCTC8325 (*rsbU*) et la souche RN1HG (*rsbU* réparé) montre que l'expression de l'ARNIII est moins importante lorsque  $\sigma^B$  est actif (Herbert *et al.*, 2010).

**RsaA faisant parti du régulon de  $\sigma^B$ , existent-t-il des liens permettant de relier l'adaptation au stress médié par  $\sigma^B$  et la régulation de l'expression des gènes de virulence médiée par *agr* ?**

## II.A. Identification des cibles de RsaA

L'identification des cibles d'un ARN régulateur reste toujours à l'heure actuelle une question à laquelle il est difficile de répondre. En effet, aucune stratégie unique n'existe pour identifier directement toutes les cibles directes d'un sARN (Vogel et Wagner 2007). Les raisons se trouvent dans la grande variété des mécanismes de régulation utilisés par les sARN qui peuvent moduler aussi bien l'expression d'ARNm que de protéines (voir Introduction III.B). L'identification des cibles d'un ARN repose donc sur l'utilisation combinée d'approches *in silico* et *in vivo* pour établir une liste de candidats potentiellement régulés par

un sARN. *In silico*, l'utilisation de programmes bioinformatiques prédit les duplexes les plus stables entre le sARN et les ARNm cibles. *In vivo*, l'utilisation de techniques globales telles que l'analyse du transcriptome et du protéome identifient les fonctions régulées par le sARN. C'est souvent la combinaison des résultats qui conduit à établir une liste de cibles potentielles. La validation des cibles directes de l'ARN régulateur nécessite l'utilisation de fusion transcriptionnelle et traductionnelle pour confirmer le niveau de régulation *in vivo*. La détermination des sites d'interaction, l'affinité des deux partenaires et le mécanisme de régulation de la cible permettent de comprendre les relations entre structure et régulation (Figure 33).

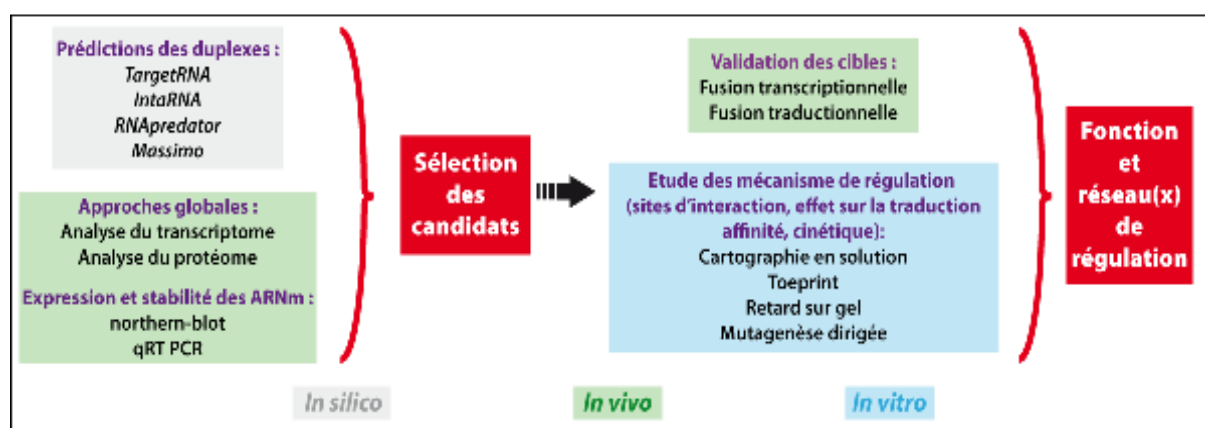


Figure 34 : Combinaison de stratégies *in silico*, *in vivo* et *in vitro* pour l'étude fonctionnelle des sARN.

En commençant l'étude fonctionnelle de *RsaA*, mon attention s'est portée tout d'abord sur les prédictions déjà disponibles à mon arrivée au laboratoire. Elles utilisent la même méthode que les prédictions des cibles de l'ARNIII, et consiste à rechercher les appariements stables potentiels entre la région 5'UTR allant de la position -50 à + 50 des ARNm (la position +1 marquant l'adénine du codon d'initiation de la traduction AUG) avec la séquence du sARN. Cette stratégie a permis l'identification de la quasi-totalité des cibles de l'ARNIII (à l'exception de l'ARNm *lytM*). Parmi les duplexes les plus stables, deux impliquaient les ARNm *citM* et *pycA* et le motif riche en C de *RsaA*. Ces deux ARNm appartiennent au régulon de  $\sigma^B$  et sont réprimés par  $\sigma^B$  (Bischoff *et al.*, 2004). Il est donc tentant d'émettre l'hypothèse que la répression observée résulte de l'action de *RsaA*. Comme les ARNm réprimés au niveau traductionnel sont en général rapidement dégradés, j'ai analysé l'expression de ces ARNm par northern. Cependant, aucune différence d'expression n'a été observée entre la souche RN1HG et une souche isogénique délétée du gène *rsaA* dans un

milieu riche. D'autres cibles ont été testées comme l'ARNm *set11* (codant pour l'exotoxine 11) et SA2226 (codant pour une perméase) sans donner de résultats probants (données non montrées). Bien que l'étude de ces cibles potentielles n'ait pas été poursuivie, je n'exclue pas la possibilité que *RsaA* puisse réguler la traduction de ces ARNm.

### **II.A.1. Comment identifier les cibles de *RsaA* ?**

La majorité des ARN régulateurs dont la fonction a été caractérisée à ce jour agissent par appariement de bases avec un ARNm cible. Pour identifier les cibles potentielles de *RsaA*, j'ai combiné deux stratégies. Plusieurs programmes bioinformatique ont été utilisés pour rechercher les appariements entre les régions RBS des ARNm et *RsaA*. Les résultats obtenus ont été comparés à ceux d'une analyse comparative globale du protéome sur des extraits totaux de protéines.

#### **II.A.1.a. Prédiction *in silico* d'ARNm cibles**

La plupart des programmes bioinformatiques utilisés dans la recherche d'ARNm cibles potentiels demandent de préciser une région de l'ARNm à prendre en considération dans la recherche de duplexes stables avec un sARN. Comme nous l'avons vu dans l'**introduction** (voir III.B.2.b), la région 5'UTR des ARNm au voisinage du RBS et de l'AUG constitue une région privilégiée de fixation des sARN. Chez les bactéries gram positives, les ARNm possèdent une séquence SD forte qui permet de recruter rapidement le ribosome pour une expression efficace (Ma *et al.*, 2002). Le recouvrement de l'ARNm par des ribosomes le protège contre la dégradation par les RNases et la fixation du sARN régulateur. Pour être efficace, la fixation du sARN à sa cible doit se faire avant l'initiation de la traduction pour empêcher la synthèse de protéines. La présence d'un motif riche en C, capable de s'apparier rapidement et fortement à la séquence Shine-Dalgarno riche en G de l'ARNm est une stratégie adaptée pour une répression efficace de l'ARNm et empêcher la fixation du ribosome. Durant mon doctorat, plusieurs outils dédiés à l'identification des appariements complémentaires entre deux ARN ont été développés. Tous disposent d'une interface WEB les rendant accessibles et faciles à utiliser. Ces approches sont très rapides et permettent d'avoir une liste importante de candidats potentiels. Néanmoins, elles résultent

également dans la détection d'un grand nombre de faux positifs, d'où l'idée de comparer les résultats de plusieurs programmes utilisant des critères pour voir si certaines cibles sont prédites plusieurs fois.

Le programme **TargetRNA** (<http://snowwhite.wellesley.edu/TargetRNA/>) se base sur le principe qu'un appariement sARN/ARNm cible constitue un module conservé et identifiable dans des génomes de bactéries proches. Ce programme retrouve des cibles validées expérimentalement pour certains sARN principalement issus des bactéries Gram négatives, comme par exemple [spot42](#). Une limitation de ce programme est qu'il ne tient pas compte de la structure secondaire des partenaires et donc de l'accessibilité des séquences prédites dans les appariements (Tjaden 2008). L'interface permet de tenir compte ou non d'un terminateur de transcription (enlevé par défaut), de choisir un organisme proche pour la recherche de duplexes orthologues, de définir la région à prendre en compte au niveau des ARNm (-30 à +20 par défaut) et de choisir la longueur minimale du duplexe (Seed) en autorisant ou non les appariements de bases non canoniques G-U. Dans le [tableau 5](#), la recherche des cibles de *RsaA* s'est faite dans le génome de la souche N315 avec comme organisme orthologue la souche Mu50.

**RNApredator** ([http://rna.tbi.univie.ac.at/RNApredator2/target\\_search.cgi](http://rna.tbi.univie.ac.at/RNApredator2/target_search.cgi)) prend en considération l'accessibilité des séquences dans les appariements sARN / ARNm et les annotations de la base de données GO pour identifier des gènes appartenant à une même voie métabolique ou cellulaire. Les auteurs mettent également en avant la rapidité du logiciel. L'ensemble de la séquence de l'ARNm est utilisé dans ce programme. Ainsi, dans le [tableau 5](#), les 10 meilleurs candidats présentés montrent des duplexes impliquant soit la région codante, soit la région 5' UTR, mais aucun impliquant le RBS des ARNm. L'interface web invite l'utilisateur à rentrer un génome via son numéro d'accès Ref-seq (NC\_002745 pour *S. aureus* N315) et une séquence d'ARN régulateur. Une des limites de ce programme est qu'il ne permet pas de modification de paramètres comme par exemple la définition de la longueur minimale d'interaction et ne tient pas compte de la conservation des cibles d'autres bactéries (Eggenhofer *et al.*, 2011).

De manière similaire, le programme **IntaRNA** (<http://rna.informatik.uni-freiburg.de:8080/IntaRNA.jsp>) tient lui aussi compte de l'accessibilité des régions

d'interactions des deux ARN. A l'inverse de **RNApredator**, il laisse la possibilité à l'utilisateur de choisir la région de l'ARNm à cibler, la longueur minimale de l'interaction, et il permet à l'utilisateur de déterminer la séquence à privilégier dans les appariements au niveau du sARN étudié. Cependant, le programme ne tient pas compte de la conservation de l'hybride dans les bactéries proches et nécessite plus de temps pour afficher les résultats (entre 20 et 30 minutes). Dans le **tableau 5**, **IntaRNA-1** représente les cibles identifiées en utilisant les paramètres par défaut et **IntaRNA-2** les cibles obtenues en orientant la recherche avec la séquence riche en C de **RsaA** (seed 87 à 93, avec au minimum 5 appariements) (Busch *et al.*, 2008).

<b>TargetRNA</b>		<b>RNApredator</b>			<b>IntaRNA-1</b>			<b>IntaRNA-2</b>			<b>Massimo</b>	
cible	$\sigma^B$	cible	$\sigma^B$	$\Delta G$ (kJ/mol)	cible	$\sigma^B$	$\Delta G$ (kCal/mol)	cible	$\sigma^B$	$\Delta G$ (kCal/mol)	cible	$\sigma^B$
SA0641	non	SA1654	non	-23,27	SA0641	non	-17.8	SA0641	non	-17.801	SA1435	non
SA1990	non	SA0305	non	-17,53	SA1956	non	-17.78	SA1001	non	-13.64	SA0688	non
SA1956	non	SA2168	non	-16,66	SA1001	non	-13.64	SA2353	non	-13.48	SA2047	non
SA0387	non	SA0523	non	-16,58	SA2353	non	-13.48	SA1833	non	-13.48	SA2181	non
SA1140	non	SA0221	non	-15,98	SA1833	non	-13.35	SA0424	non	-13.06	SA1270	non
SA0815	non	SA0522	non	-15,95	SA0424	non	-13.06	SA2485	up	-13.00	SA1256	non
SA0957	non	SA2328	non	-15,95	SA2485	up	-13.00	SA2332	non	-12.92	SA1327	non
SA0266	non	SA1945	non	-15,22	SA2332	up	-12.92	SA1735	non	-12.84	SA2051	non
SA0498	non	SA0825	non	-15,16	SA2271	non	-12.80	SA1586	non	-12.77	SA0602	non
SA2074	non	SA1020	non	-14,92	SA1586	non	-12.77	SA1394	non	-12.71	SA0726	non

**Tableau 5:** Liste des 10 meilleurs ARNm candidats formant potentiellement un duplex avec **RsaA in silico**.

En orange et bleu, candidats retrouvés dans des approches différentes. En violet, protéine identifiée par les analyses du protéome. Les stabilités estimées des duplexes sont indiquées pour **RNApredator** (kJ/mol) et **IntaRNA** (kCal/mol) (les deux autres approches se basent sur l'attribution de scores) (**Résultats II.A.1.b**).

Le tableau récapitulatif ci-dessus montre que seulement deux cibles, SA0641 et SA1956, sont retrouvées plusieurs fois. Les autres candidats sont uniques soulignant sans doute un grand nombre de faux-positifs, et les données bibliographiques en ma possession ne me permettent pas de les relier au régulon de  $\sigma^B$  (Bishoff *et al.*, 2004). Néanmoins, il est intéressant de noter que le motif riche en C de **RsaA** est reconnu comme accessible, puisque 8 des 10 candidats prédits par **IntaRNA** sont communs entre la recherche par défaut et celle utilisant le motif riche en C. De manière intéressante, même si seulement deux cibles sont communes, les critères utilisés entre **TargetRNA** et **IntaRNA** se complètent de manière



élégante. La conservation dans les génomes proches, en association avec des régions accessibles à la fois sur le sARN et l'ARNm rend les cibles SA0641 et SA1956 très intéressantes à étudier.

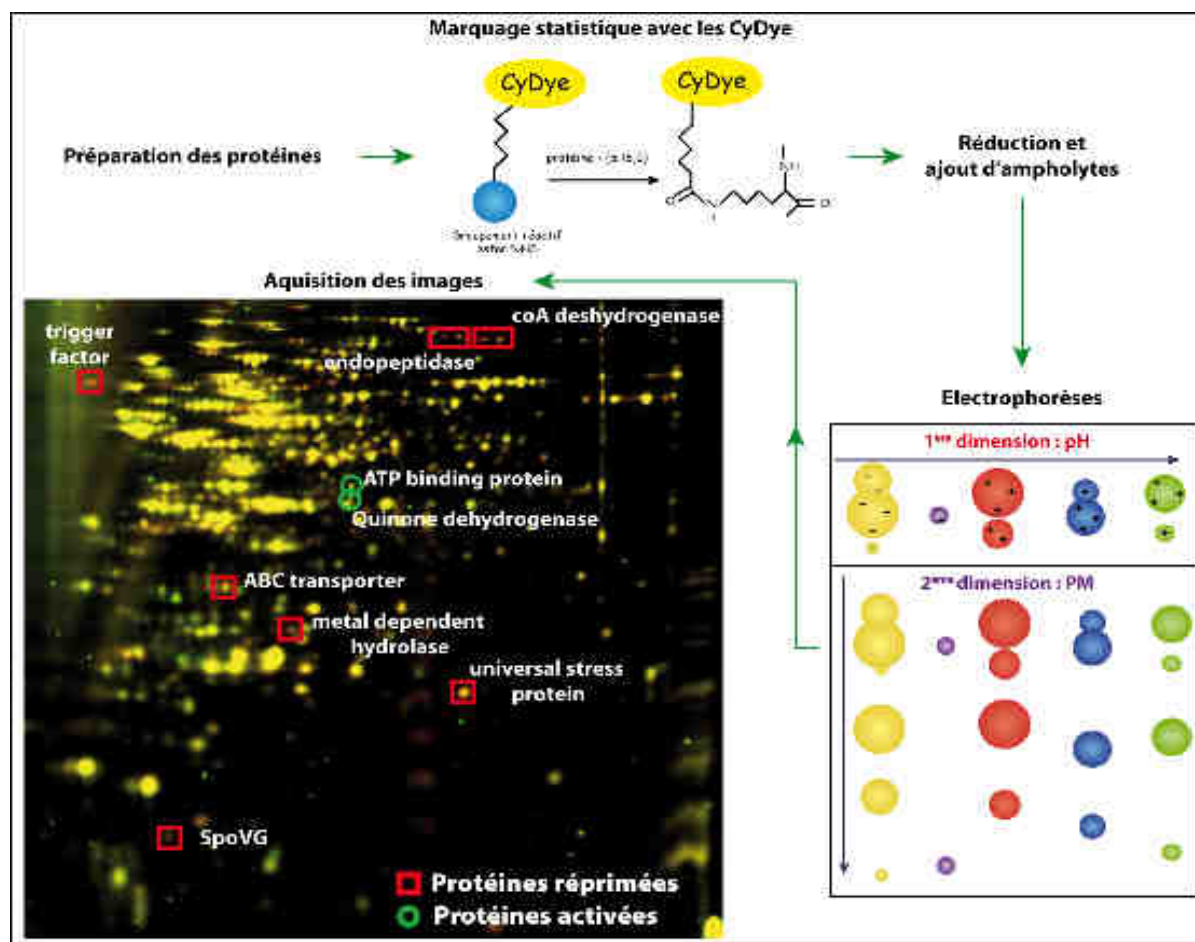
Une recherche sur les bases de données STRING et NCBI indique que SA0641 code pour le facteur de transcription MgrA (Luong *et al.*, 2003), connu également sous le nom de RAT (Ingavale *et al.*, 2003) ou de NorR (Truong *et al.*, 2003). MgrA régule la formation de biofilm (Trotonda *et al.*, 2008), de capsules (Luong *et al.*, 2006), l'autolyse (Ingavale *et al.*, 2003) et l'expression de certains facteurs de virulence. MgrA a aussi un effet positif sur le système *agrA* (Ingavale *et al.*, 2005). C'est pour les fonctions multiples qu'il exerce dans la physiologie globale de *S. aureus* qu'il est aujourd'hui considéré comme une cible dans la mise au point de nouvelles molécules thérapeutiques (Sun *et al.*, 2011).

### II.A.1.b. Analyse comparative du protéome

Pour compléter ces prédictions, j'ai utilisé en parallèle une analyse comparative du protéome pour identifier d'autres cibles et définir les processus biologiques régulés par RsaA. Cette approche possède également l'avantage d'être insensible au mécanisme de régulation (activation ou répression à un site différent du RBS) et à la stabilité de l'ARNm. Au laboratoire, l'analyse comparative du protéome a été développée par Pierre Fechter en collaboration avec Phillippe Hammann (plateforme protéomique, IBMC). Le principe est de séparer les protéines sur un gel en deux dimensions : la première en fonction de leur point isoélectrique (pI), la seconde en fonction de leur masse moléculaire. La technique que j'ai utilisée est la 2-D DIGE (*fluorescence difference gel electrophoresis*). Le principe de séparation des protéines reste le même, mais ces dernières sont préalablement marquées avec un fluorophore "CyDye". Le marquage des protéines est statistique et s'opère via la fixation d'un groupement N-hydroxysuccinimidyl sur les lysines des protéines à pH 8,5. Au final, seul 1% de l'ensemble des protéines sont marquées. Trois CyDye sont disponibles : Cy2 ( $\lambda_{\text{ex}}$  488 nm,  $\lambda_{\text{em}}$  520 nm), Cy3 ( $\lambda_{\text{ex}}$  532 nm,  $\lambda_{\text{em}}$  580 nm) et Cy5 ( $\lambda_{\text{ex}}$  633 nm,  $\lambda_{\text{em}}$  670 nm). L'ajout du fluorophore à la protéine ne change pas la charge globale car la charge positive perdue au niveau des lysines est compensée par la charge positive propre aux réactifs CyDye. Néanmoins, la fixation ajoute une masse moléculaire de 500 Da à la protéine, entraînant un

décalage durant la seconde dimension, amplifié pour les protéines de faible PM. Le séquençage des protéines à la fin de l'expérience se fait en conséquence à partir des gels colorés au bleu de coomassie puisque la majorité des protéines n'est pas marquée (**Figure 34 A**).

Dans mes expériences, un extrait des protéines cytoplasmiques et membranaires est préparé à partir des souches RN1HG et RN1HG  $\Delta$ *rsaA* mises en culture dans le milieu riche BHI jusqu'en phase stationnaire de croissance, permettant une expression maximale de RsaA. Les fluorophores Cy3 et Cy5 servent à marquer les extraits des souches sauvage et mutante, respectivement. En parallèle, un mélange des deux extraits est marqué à la Cy2 qui sert de contrôle interne dans la quantification des signaux de fluorescence. L'ensemble des réactions de marquage est séparé sur le même gel 2D. En conséquence, la superposition des images obtenues est automatique (**Figure 34 B**). Les expériences ont été réalisées sur deux échantillons de protéines, chaque expérience ayant fait l'objet d'un duplicat expérimental. Les spots présentant une différence d'expression de deux fois ou plus entre la souche sauvage et mutante et retrouvés dans tous les gels sont extraits puis soumis à une digestion à la trypsine. Les fragments ont ensuite été identifiés par spectrométrie de masse.



**Figure 35 : Analyse comparative du protéome par la technique 2D-DIGE.**

Les lysines des protéines sont marquées statistiquement avec un fluorophore de type CyDye. Après une étape de réduction pour neutraliser les lysines et les groupements ester n'ayant pas réagi, les extraits protéiques marqués avec la Cy2, Cy3 et Cy5 sont séparés sur le même gel en fonction du pI (1<sup>ère</sup> dimension : isofocalisation électrique avec un gradient de pH allant de 4 à 7) puis de la masse moléculaire (2<sup>ème</sup> dimension : gel SDS-PAGE 12%). Les images des gels sont obtenues grâce à un scanner capable d'exciter chaque sonde fluorescente et d'enregistrer les signaux émis. L'image finale est une superposition du protéome de la souche RN1HG en vert et de la souche RN1HG  $\Delta$ *rsaA* en rouge. Les protéines, dont la synthèse ne varie pas, apparaissent en jaune.

De manière surprenante, la délétion du gène *rsaA* induit des effets multiples sur le protéome de la souche RN1HG de *S. aureus*. Pour chaque protéine séquencée et identifiée, dont l'expression est dépendante de *RsaA* et pour lesquelles aucun appariements n'ont été prédits avec les programmes *RNAPredator*, *IntaRNA* et *TargetRNA*, j'ai cherché des duplexes stables éventuels entre la région 5'UTR de l'ARNm (de -50 à +50) et le motif riche en C de *RsaA* en utilisant des programmes de repliement d'ARN comme *mFOLD* (Zuker 2003) et *contraFOLD* (Do *et al.*, 2006) (Tableau 6). Pour l'ensemble des candidats testés, seul un duplexe stable ressort avec l'ARNm *yabJ-spoVG* (Figure 34). Cet ARNm est sous le contrôle d'un promoteur  $\sigma^B$  (Bishoff *et al.*, 2004) et est impliqué dans la synthèse de capsules. Ces

dernières protègent la bactérie contre la phagocytose par les macrophages et contre l'action des antibiotiques de la paroi, comme la méticilline et les glycopeptides (Meier *et al.*, 2007, Schulthess *et al.*, 2009). Pourquoi ce duplex n'est-il pas prédit avec les programmes de recherche de cibles (voir [Résultats II.A.1.a](#)) ? D'un point de vue structural, une prédiction de l'ensemble de la région 5'UTR de *yabJ* montre une région relativement structurée et présente le SD et l'AUG au niveau d'une boucle apicale, expliquant sans doute pourquoi cette cible n'a pas été détectée par **RNApredator** et **IntaRNA** qui, pour rappel, prennent en compte l'accessibilité des régions dans la recherche d'appariements stables avec un sARN. Le fait que cette cible ne soit pas non plus identifiée par **TargetRNA** suggère-t-il que le duplexe n'est retrouvé que dans la souche RN1HG? Une analyse de la région 5'UTR de l'ARNm *yabJ* montre qu'elle est très conservée dans plusieurs espèces de *S. aureus*.

Existe-t-il d'autres candidats pour une régulation par **RsaA** capable d'expliquer les résultats de l'analyse comparative du protéome ? Pour rappel, les analyses *in silico* font ressortir l'ARNm codant pour le facteur de transcription MgrA comme cible potentielle. Réguler l'expression d'un facteur global de la transcription permettrait à la bactérie de moduler simultanément l'expression de nombreux gènes et expliquerait les résultats obtenus. Une étude de Luong *et al.*, en 2006 a établi le transcriptome de MgrA dans la souche Newman de *S. aureus*. En comparant ces données avec les résultats du protéome comparatif, 10 protéines sont apparues communes ([tableau 6](#)) dont 7 vont dans le sens d'une répression de l'ARNm *mgrA* par **RsaA**. MgrA est une protéine de 17089 Da et de pI 6,84. Sur les gels 2D, cette zone montre des spots non résolus, probablement à cause des fortes concentrations de sels s'accumulant à l'extrémité du STRIP à la fin de la première dimension. De manière générale, le pI élevé et le faible PM des facteurs de transcription les rendent difficile à détecter sur des gels 2D.

Gène	Fonction	Régulation RsaA	Régulation MgrA
<i>hla</i>	hémolysine alpha	ACTIVATION	
SA2155	malate:quinone oxidoreductase	ACTIVATION	
<i>prfA</i> (SA1920)	peptide chain release factor 1	ACTIVATION	
<i>pdhC</i> (SA0945)	branched-chain alpha-keto acid dehydrogenase subunit E2	ACTIVATION	
<i>spoVG</i> (SA0456)	protéine régulatrice SpoVG	RÉPRESSION	
SA1990	quinone oxidoreductase	RÉPRESSION	
SA1969	ATP-binding Mrp/Nbp35 family protein	RÉPRESSION	
SA0224	3-hydroxyacyl-CoA dehydrogenase FadB	RÉPRESSION	
<i>infB</i> (SA1112)	Facteur d'initiation de la traduction IF-2	RÉPRESSION	
<i>ClpC</i> (SA0483)	endopeptidase	RÉPRESSION	
SA1409	DnaK	RÉPRESSION	
<i>fus</i> (SA0505)	Facteur d'élongation G	RÉPRESSION	
<i>ClpL</i> (SA2336)	ATP-dependent Clp proteinase chain clpL	RÉPRESSION	
SA1523	acetyl-CoA carboxylase (sous-unité β)	RÉPRESSION	
<i>lysS</i> (SA0475)	lysyl-ANRt synthetase	RÉPRESSION	
<i>tig</i> (SA1499)	trigger factor	RÉPRESSION	
SA0226	long-chain-fatty-acid-CoA ligase	RÉPRESSION	
SA1946	protéine hypothétique	RÉPRESSION	
<i>cysE</i> (SA0487)	serine acetyltransferase-like protein	RÉPRESSION	
<i>pdp</i> (SA1938)	pyrimidine-nucleoside phosphorylase	RÉPRESSION	
<i>glyS</i> (SA1934)	glycyl-tRNA synthetase	RÉPRESSION	
<i>Ddh</i> (SA2312)	D-lactate dehydrogenase	ACTIVATION	RÉPRESSION
<i>guaA</i> (SA0376)	bifunctional GMP synthase/glutamine amidotransferase	RÉPRESSION	ACTIVATION
SA1532	universal stress protein	RÉPRESSION	ACTIVATION
<i>lip*</i> (SA2463)	Lipase	RÉPRESSION	ACTIVATION
<i>pyrG</i> (SA1929)	CTP synthetase	RÉPRESSION	ACTIVATION
SA1529	metal-dependent hydrolase (protéine hypothétique)	RÉPRESSION	ACTIVATION
SA0774	ABC transporter ATP-binding protein	RÉPRESSION	ACTIVATION
<i>HlgB</i> (SA2209)	Chain A, Leukocidin F (HlgB)	ACTIVATION	ACTIVATION
<i>HlgC</i> (SA2208)	(composant de la leukocidin LukS)	ACTIVATION	ACTIVATION
SA0375	inositol-monophosphate dehydrogenase	ACTIVATION	ACTIVATION

**Tableau 6 : Récapitulatif des protéines séquencées sur les gels 2D-DIGE.**

Régulation par *RsaA* d'après les gels 2D. Régulation par *MgrA* d'après les données transcriptomique de Luong *et al.*, 2006. Violet : protéine prédite *in silico*. Bleu : protéines appartenant au régulon *MgrA* et régulées négativement par *RsaA*. En orange, protéines appartenant au régulon *MgrA* et régulées positivement par *RsaA*.

En résumé, quatre candidats ressortent de l'ensemble des analyses (**Figure 35**):

- L'ARNm *mgrA*, codant pour un facteur de transcription, est le meilleur candidat prédit *in silico* et peut être relié indirectement aux résultats du protéome comparatif. C'est pour cette raison que j'ai décidé de me concentrer sur cette cible dans un premier temps. De plus, hormis l'ARNIII, aucun ARN régulateur de *S. aureus* ne cible de facteur de transcription parmi ceux dont la fonction a été caractérisée à ce jour.

- L'ARNm SA1990 codant pour une quinone oxydoréductase et l'ARNm SA1394 codant pour la glycyL ARNt synthétase, sont identifiés à la fois *in silico* et *in vivo*. Les duplexes prédits font soit intervenir la base de la tige-boucle en 5' de RsaA, région non accessible au premier abord pour l'ARNm SA1990, soit la région centrale sans le motif riche en C pour l'ARNm SA1394.

- L'ARNm *yabJ-spoVG* est ressorti uniquement par l'analyse de protéome. Le duplexe prédit fait intervenir le motif riche en C de RsaA.

Dans ce qui va suivre, je me suis consacrée essentiellement à valider expérimentalement l'ARNm *mgrA* comme cible de RsaA.

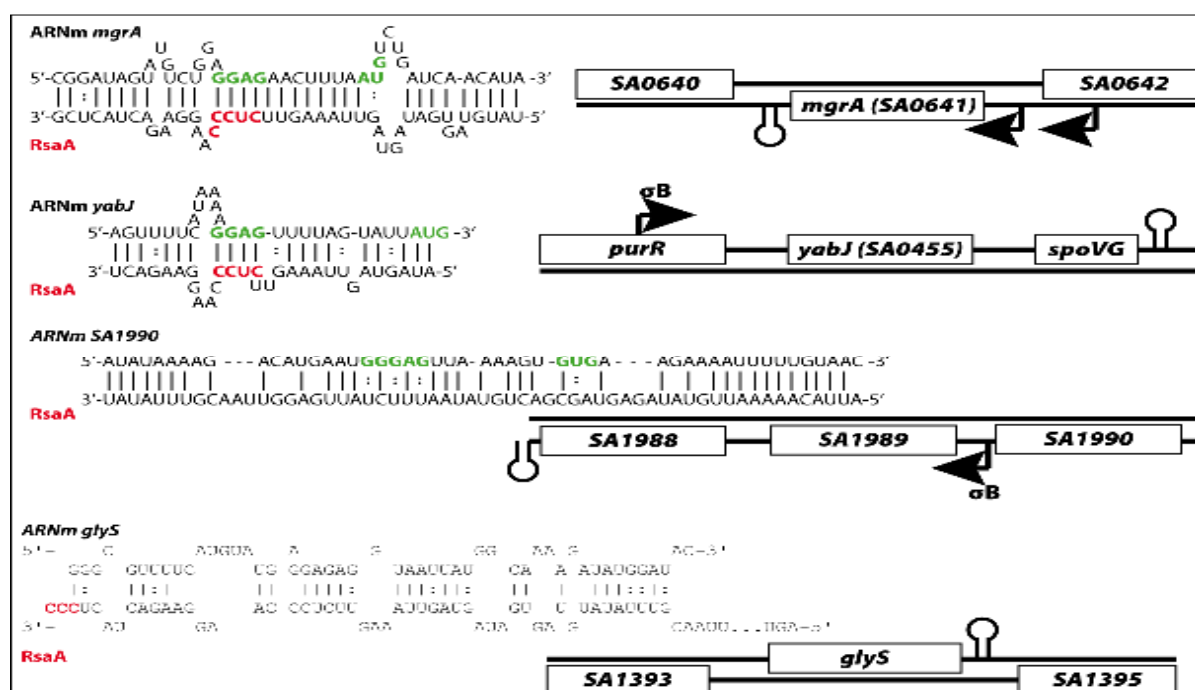


Figure 36 : Prédiction des appariements RsaA / ARNm cible et contexte génomique.

La séquence SD et le codon d'initiation de la traduction sont en vert, le motif riche en C de RsaA est en rouge.

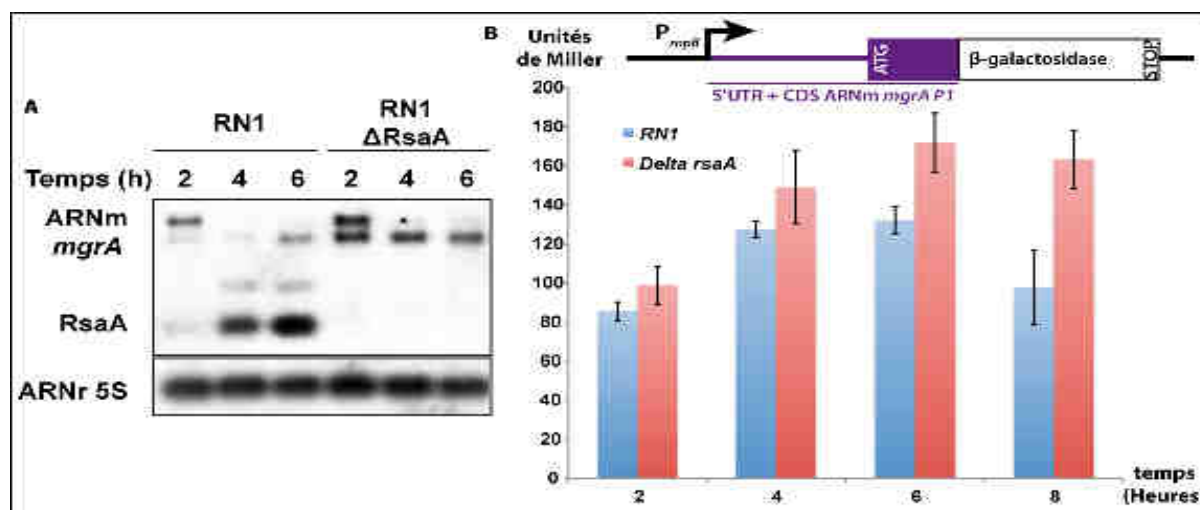
## II.B. RsaA inhibe la traduction de l'ARNm mgrA

Quel est le mécanisme et quel est le niveau de régulation de RsaA sur sa cible ? Pour répondre à ces questions, j'ai tout d'abord validé *in vivo* que l'ARNm *mgrA* est une cible directe de RsaA puis j'ai élucidé le mécanisme de régulation *in vitro*.

### II.B.1. L'ARNm *mgrA* est régulé au niveau post-transcriptionnel

Tout d'abord, j'ai regardé le profil d'expression de l'ARNm *mgrA* à 2, 4 et 6 h de croissance dans le milieu riche BHI par des expériences de northern, dans la souche sauvage RN1HG et la souche RN1HG  $\Delta$ *rsaA* (utilisées dans l'analyse comparative du protéome). Deux bandes sont détectées pour l'ARNm *mgrA* en début de phase exponentielle (2 h de culture), l'une correspondante à un ARN de 750 nts, l'autre de 560 nts. Ces résultats sont attendus, puisque l'ARNm *mgrA* est transcrit à partir de deux promoteurs (**Figure 36 A**) (Ingavale *et al.*, 2003). Dans la souche sauvage, l'expression de l'ARNm de 750 nts disparaît en phase exponentielle et stationnaire alors que l'ARNm de 560 nts est exprimé faiblement durant toutes les phases de croissance. Dans la souche mutante RN1HG  $\Delta$ *rsaA*, la bande de 560 nts s'accumule en phase exponentielle et stationnaire de croissance par rapport à son homologue sauvage, signifiant que *RsaA* exercerait effet sur la stabilité de l'ARNm. Pour déterminer si *RsaA* régule l'expression de l'ARNm *mgrA* au niveau post-transcriptionnel, j'ai construit une fusion comprenant la région 5'UTR (141 nts), les 39 premiers aa de MgrA (117 nts) en phase avec le gène rapporteur codant pour la  $\beta$ -galactosidase. Cette fusion est portée par le plasmide pLUG220, sous le contrôle du promoteur constitutif P<sub>rpoB</sub> (Boisset *et al.*, 2007). Ce plasmide a été introduit dans les souches RN1HG et RN1HG  $\Delta$ *rsaA*, puis l'activité du gène rapporteur a été mesurée en fonction de la croissance. Aucune différence significative dans l'activité de la  $\beta$ -galactosidase n'est observée entre les deux souches après 2 et 4 h de culture. Ce résultat peut être mis en corrélation avec le faible taux d'expression de *RsaA* à ces temps de croissance. A l'inverse, en phase stationnaire, l'activité  $\beta$ -galactosidase diminue dans la souche sauvage qui exprime fortement *RsaA*. Dans la souche RN1HG  $\Delta$ *rsaA*, l'activité mesurée continue d'augmenter pour atteindre un palier (**Figure 36 B**). Ainsi, les résultats obtenus montre que dans la souche RN1HG, *RsaA* régule l'expression de l'ARNm *mgrA* au niveau post-transcriptionnel.

**Cette régulation est-elle la conséquence d'une inhibition de la traduction et/ou d'une déstabilisation de l'ARNm ?**



**Figure 37 : RsaA régule l'expression de l'ARNm *mgrA* au niveau post-transcriptionnel *in vivo*.**

**A.** Analyse de l'expression de l'ARNm *mgrA* par northern blot (gel agarose 1%) dans les souches RN1HG et RN1HG  $\Delta$ *rsaA*. Les ARN totaux ont été prélevés après 2, 4 et 6 h de culture à 37°C dans le milieu riche BHI. **B.** Activité  $\beta$ -galactosidase de la fusion  $P_{rpoB} P1$  *mgrA* (5'UTR et +1/+258)::*lacZ* dans la souche RN1HG (bleu) et la souche mutante RN1HG  $\Delta$ *rsaA* (rouge). Les différents extraits ont été préparés après 2, 4, 6 et 8 h de culture dans le milieu riche BHI. L'activité est donnée en Unité de Miller et est rapportée à la concentration en protéines des différents échantillons.

### II.B.2. RsaA se fixe à l'ARNm *mgrA* à deux sites distants

Pour déterminer le mécanisme de régulation de *RsaA*, j'ai tout d'abord montré que ce dernier se fixe à l'ARNm *mgrA*. Des expériences de retard sur gel montrent que *RsaA* forme un duplexe stable avec l'ARNm *mgrA* dans sa forme longue (données non montrées) et sa forme courte (Figure 37 A). Cette interaction est stable ( $K_{d\text{ app}} \approx 20$  nM) et ne nécessite par l'intervention de la protéine Hfq. En utilisant des expériences de cartographie en solution, j'ai démontré que *RsaA* s'apparie à l'ARNm *mgrA* au niveau de deux sites d'interaction distants. Le premier correspond à l'hybride prédit par les programmes *TargetRNA* et *IntaRNA*, et implique le motif riche en C et le site de fixation du ribosome de l'ARNm, avec notamment trois appariements de bases G-C impliquant le SD. Cette interaction s'étend jusqu'au codon d'initiation de la traduction AUG (13 appariements de bases canoniques consécutifs). Ce long duplexe est reconnu et clivé à deux endroits entre la séquence SD et le codon d'initiation par la RNase III (Figure 37 B, C et D). De manière intéressante, une seconde interaction a lieu entre la tige-boucle 5' de *RsaA* et une tige-boucle localisée dans la région codante de l'ARNm *mgrA*. Là aussi, trois appariements de bases G-C sont impliqués.



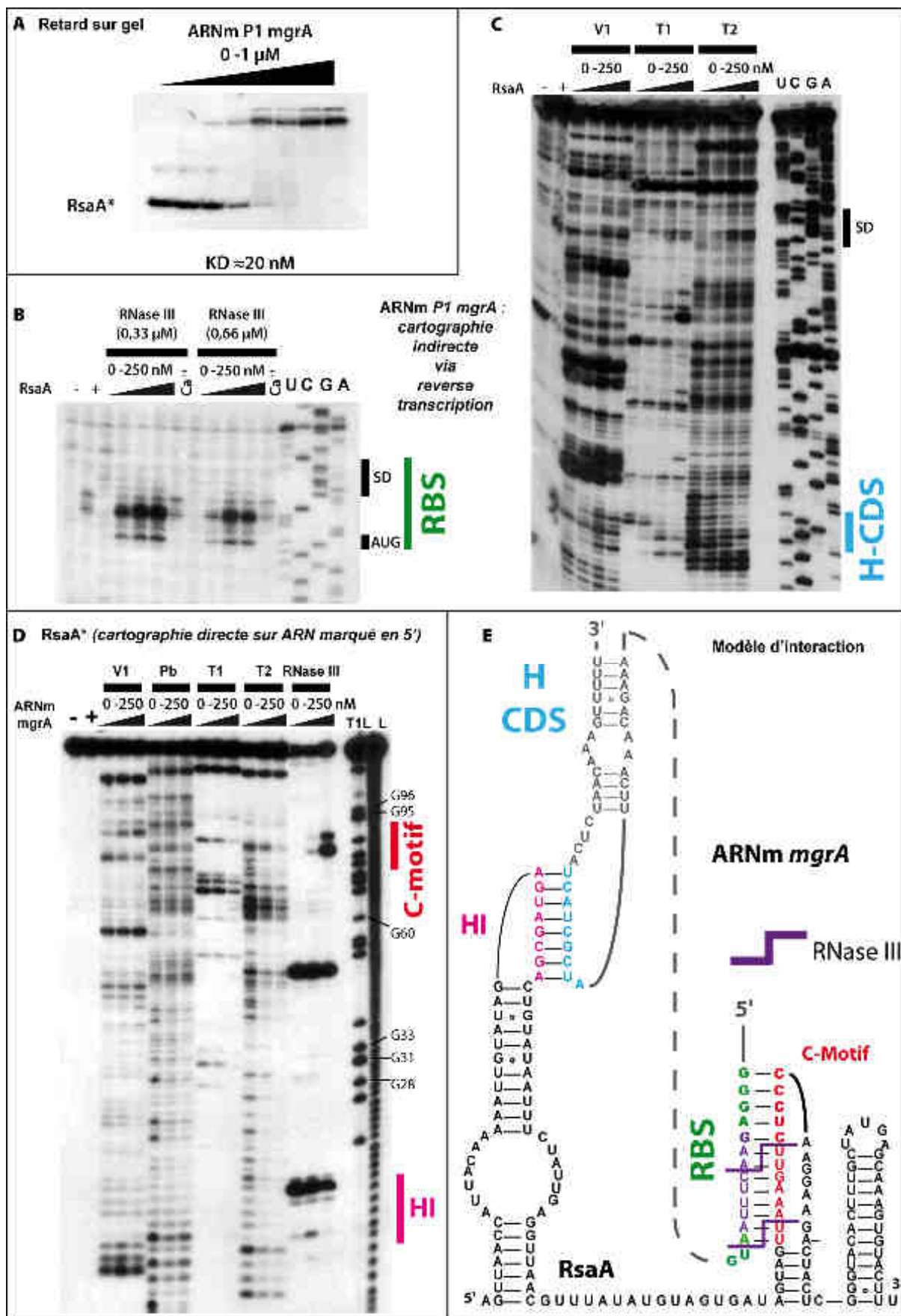


Figure 38 : RsaA se lie efficacement à l'ARNm *mgrA* en utilisant deux sites d'interaction distants.

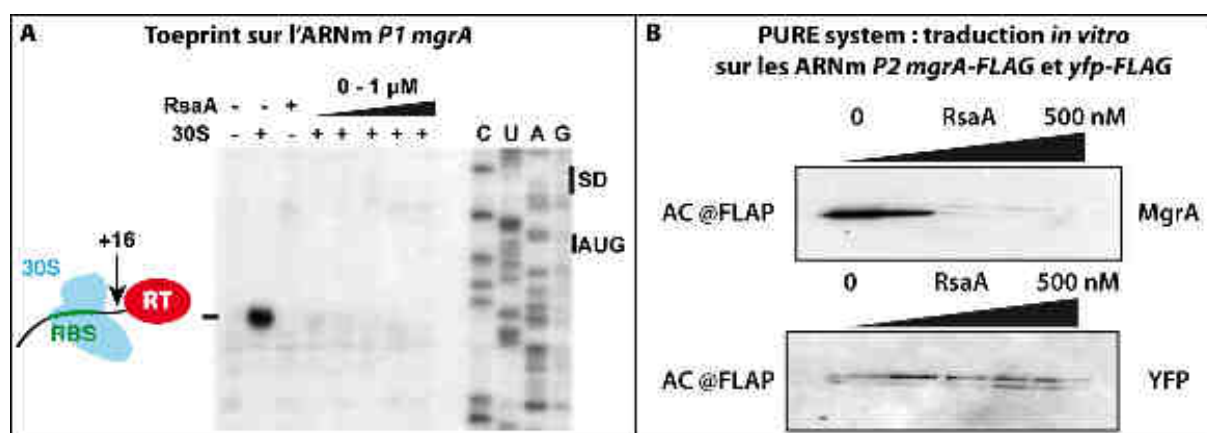
**A. Expérience de retard sur gel.** *RsaA*, marqué à son extrémité 5' au  $Y-P^{32}$  (5000 cpm/ligne) est incubé seul ou en présence de concentration croissante d'ARNm *P1 mgrA* (1, 5, 10, 20, 100, 500 et 1000 nM) puis séparé sur gel natif 6%. La fraction de *RsaA* lié à l'ARNm a été estimée en comparant l'intensité de la bande correspondante à l'intensité de toute la piste. La  $K_D \approx 20$  nM correspond à la concentration pour laquelle la moitié de *RsaA* forme un duplex avec l'ARNm *mgrA*. **B. Cartographie en solution de la structure du duplexe *RsaA*-ARNm *mgrA* par la RNase III, coté ARNm.** L'ARNm *P1 mgrA* seul (piste 3 et 8) ou en présence de concentration croissante de *RsaA* (25 nM (piste 4 et 9), 50 nM (5 et 10) et 250 nM (piste 6 et 11)) est incubé en présence de deux concentration de RNase III (0,33 et 0,66  $\mu$ M). La première piste (ARNm seul) et la deuxième (ARNm + *RsaA*) correspondent aux contrôles d'incubation. Les pistes marquées  $Ca^{++}$  correspondent aux contrôles négatifs. Les réactions de séquençage de l'ARNm via l'utilisation de ddNTP sont symbolisées par les lignes A, U, C et G. Les ARN ont été séparés sur gel acrylamide 10% en conditions dénaturantes (Urée 8M). SD : Shine et Dalgarno, RBS : Ribosome Binding Site **C. Cartographie en solution de la structure du duplexe *RsaA*-ARNm *mgrA*, coté ARNm.** L'ARNm *P1 mgrA* seul (piste 3, 7 et 11) ou en présence de concentration croissante de *RsaA* (25 nM (piste 4, 8 et 12), 50 nM (5, 9 et 13) et 250 nM (piste 6, 10 et 14)) est incubé en présence des RNases T1 ou T2 (spécifique des régions simples brins d'ARN) ou V1 (spécifique des régions doubles-brins) . La première piste (ARNm seul) et la deuxième (ARNm + *RsaA*) correspondent aux contrôles d'incubation. Les réactions de séquençage de l'ARNm via l'utilisation de ddNTP sont symbolisées par les lignes A, U, C et G. Les ARN ont été séparés sur gel de poly-acrylamide 10% en conditions dénaturantes (urée 8M) H-CDS : Hairpin in Coding Sequence. **D. Cartographie en solution de la structure du duplexe *RsaA*-ARNm *mgrA* , coté *RsaA*.** L'ARN *RsaA* seul (piste 3, 6, 9, 12 et 15) ou en présence de concentration croissante de l'ARNm *mgrA* (50 nM (piste 4, 7, 10, 13 et 16), 250 nM (5, 8, 11, 14 et 17) est incubé en présence de Plomb (sonde chimique spécifique des résidus non-appariés), des RNases T1 ou T2 (spécifique des régions simples brins d'ARN) et des RNases III ou V1 (spécifique des régions doubles-brins) . La première piste (ARNm seul) et la deuxième (ARNm + *RsaA*) correspondent aux contrôles d'incubation. Les pistes T1 et L correspondant aux échelles générées par une dégradation à la RNase T1 et une lyse alcaline en condition dénaturantes respectivement. Les ARN ont été séparés sur gel de poly-acrylamide 10% en conditions dénaturantes (urée 8M) HI : Hairpin I **E. Modèle détaillé de la structure secondaire du duplexe *RsaA* / ARNm *mgrA*.**

***RsaA* piège le site de fixation du ribosome dans une région en double brin : cet appariement implique-t-il une inhibition de la traduction de l'ARNm *mgrA* ?**

### II.B.3. *RsaA* inhibe la traduction de l'ARNm *mgrA*

Pour tester l'effet de la fixation de *RsaA* sur l'initiation de la traduction, j'ai utilisé des expériences de *toeprint* ou empreinte du ribosome qui repose sur l'utilisation d'un oligonucléotides qui s'hybride au niveau de la CDS de l'ARNm pour servir d'amorce à la RT. L'élongation est bloquée à 16 nts du codon d'initiation de la traduction si le ribosome forme un complexe ternaire avec l'ARNm et l'ARNt initiateur (Hartz *et al.*, 1988). Le complexe de d'initiation de la traduction réalisé avec des ribosomes de *S. aureus* est efficacement formé avec l'ARNm *mgrA* (Figure 38 A). A l'inverse, dès que ce dernier est incubé avec des concentrations croissantes de *RsaA*, le signal à la position + 16 disparaît. Ce résultat suggère fortement que la liaison de *RsaA* à l'ARNm *mgrA* empêche la formation du complexe ternaire (Figure 38 A). Une approche complémentaire réside dans l'utilisation d'un système de traduction *in vitro*, le *PURE System*<sup>TM</sup>. Ce système permet de transcrire un ARNm et de le

traduire en protéine à partir d'un fragment d'ADN obtenu par PCR. Pour contrôler précisément les quantités d'ARNm et de *RsaA*, j'ai réalisé les expériences en incubant directement les composants du kit avec ces derniers, en lieu et place d'ADN. La protéine *MgrA* est détectée par western blot *via* l'ajout d'une étiquette FLAG, reconnue spécifiquement par un anticorps monoclonal. L'ARNm *mgrA* seul est traduit efficacement. A l'inverse, dès que *RsaA* est en quantité équivalente ou saturante, l'inhibition de la traduction est totale. En contrôle, nous avons montré que *RsaA* n'a aucun effet sur un ARNm témoin codant pour la YFP (*Yellow fluorescent protein* ou protéine fluorescente jaune) (Figure 38 B).



**Figure 39 :** La liaison de *RsaA* à l'ARNm *mgrA* inhibe la formation du complexe de pré-initiation de la traduction.

**A. Expérience de toeprint.** L'ARNm *mgrA* (0,2 pmol) est incubé en présence d'un oligonucléotide marqué radioactivement à son extrémité 5' P au  $\gamma$ - $P^{32}$  (25000 cpm/réaction) et s'hybridant au niveau de la phase codante de l'ARNm (position +118 à +96). Ce dernier sert d'amorce à la RT pour la synthèse d'un brin d'ADNc. La formation du complexe ternaire entre l'ARNm, le ribosome de *S. aureus* (200 nM) et l'ARNt initiateur (1  $\mu$ M) est visualisé par une pause de la RT à la position +16 par rapport à l'AUG (deuxième piste). Lorsque l'ARNm est incubé en présence de concentration croissante de *RsaA* (20, 50, 150, 500 et 1000 nM, piste 4-8), la pause disparaît, signifiant l'absence de PIC. La première piste correspond à l'ARNm *mgrA* seul et la troisième piste à l'ARNm en présence de *RsaA* (contrôles d'incubation). **B. Pure system.** La protéine *MgrA* est détectée par l'ajout d'une séquence codante pour étiquette FLAG à l'extrémité C-terminale sur l'ARNm *mgrA*. 1 pmol de l'ARNm sont incubés seul (première piste) ou en présence de *RsaA* (1, 2, 5 et 10 pmol) (piste 2, 3, 4 et 5) avec les composants du kit. Les protéines sont ensuite séparées sur gel SDS-PAGE 10% puis révélées par un anticorps primaire spécifique de l'étiquette FLAG. Ce dernier est lié par un anticorps secondaire couplé à la peroxydase, qui en présence de substrat ECL émet de la lumière. En contrôle, la même expérience a été reproduite sur l'ARNm *yfp*.

Les résultats obtenus montrent que *RsaA* inhibe l'initiation de la traduction de l'ARNm *mgrA*. Le long hybride impliquant la liaison de la séquence SD et le motif riche en C est reconnu par la RNase III *in vitro*, et pourrait expliquer l'accumulation de l'ARNm dans la souche mutante RN1HG  $\Delta$ *rsaA*. Néanmoins, cet effet doit être confirmé *in vivo* en utilisant par exemple une souche délétée de l'enzyme. Récemment, E. Lioliou a co-immunoprécipité

avec un mutant de la RNase III qui ne coupe plus mais qui se fixe à l'ARN, RsaA et l'ARNm *mgrA* (Lioliou *et al.*, 2012).

**RsaA lie l'ARNm en utilisant deux sites de fixation distants : quels est le poids des deux sites de fixation dans la stabilité de l'interaction et/ou l'inhibition de la traduction ?**

#### **II.B.4. Les deux sites interagissent avec l'ARNm *mgrA* de manière coopérative pour inhiber la traduction.**

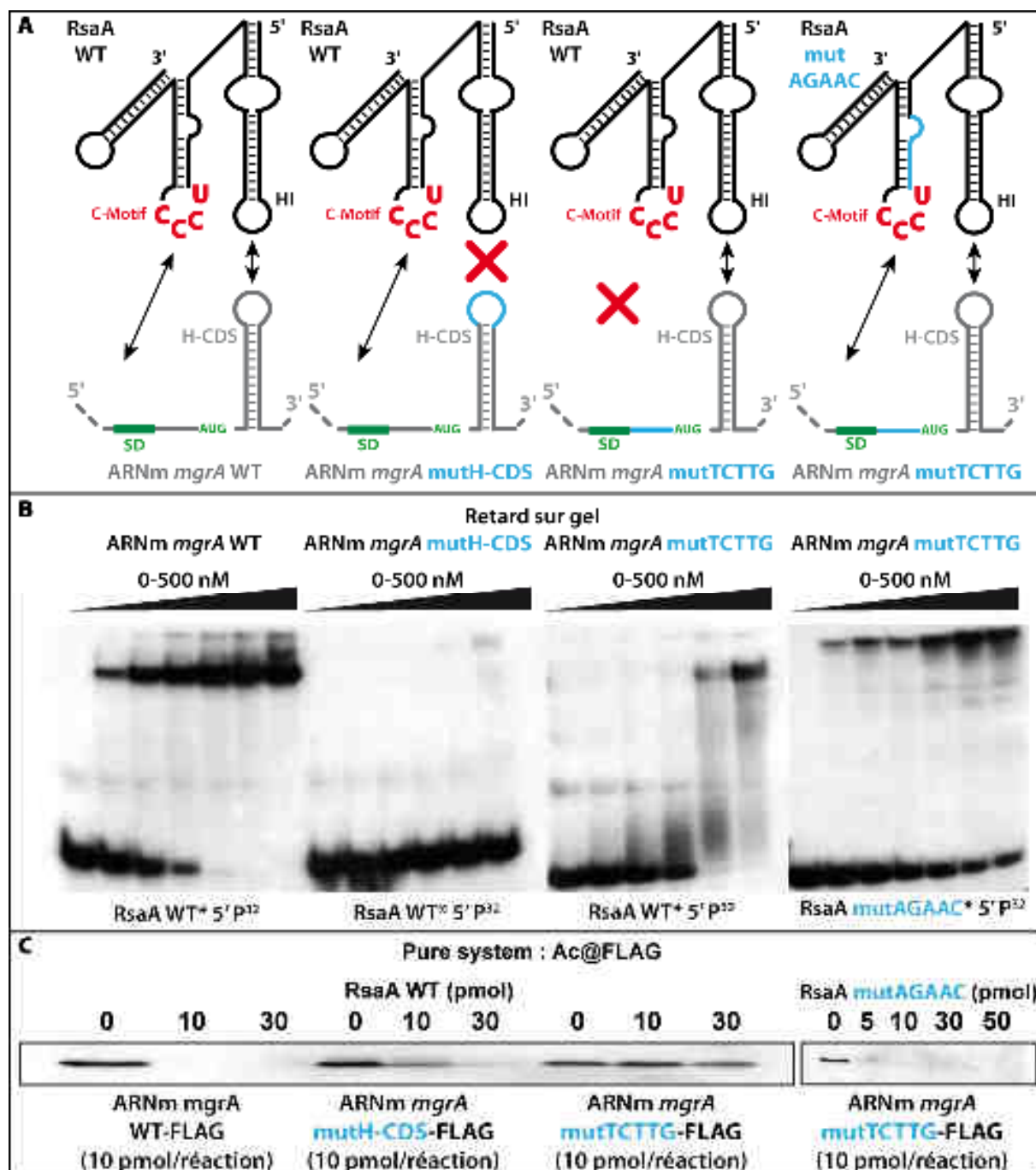
Pour évaluer le rôle de chaque région d'interaction dans la formation de l'hybride et l'inhibition de la traduction, j'ai tout d'abord effectué des expériences de mutagenèse dirigée sur RsaA et l'ARNm *P1 mgrA*. Pour chaque site d'interaction, les séquences ont été inversées entre l'ARNm et RsaA. Par exemple, la séquence portée par la tige-boucle HI de RsaA a été remplacée par la séquence de la tige-boucle localisée dans le CDS de l'ARNm *mgrA*. Le mutant ainsi obtenu, l'ARNm *mgrA* mutH-CDS est donc incapable de former l'interaction boucle-boucle avec RsaA puisque les deux ARN portent une séquence identiques ne pouvant former d'appariement de bases complémentaires. De plus, le nombre de résidus GC reste identique pour chaque mutation, et l'effet du changement de séquence sur la structure des ARN a été évalué *in silico* via **mfold** et **CONTRAFold** (données non montrées). Ces prédictions suggèrent que la structure secondaire des ARN mutants est comparable à leurs homologues sauvages. Un dernier point important concerne les mutations au niveau du site d'interaction impliquant le SD de l'ARNm *mgrA* : les nucléotides modifiés qui ont été changés ne perturbent pas la séquence SD pour éviter de modifier l'efficacité de la traduction.

L'effet des mutations sur la formation des duplexes est estimé par des expériences de retard sur gel. Une mutation au niveau de l'appariement boucle-boucle perturbe fortement la formation du duplexe. De même, le mutant TCTTG de l'ARNm *mgrA* est incapable de former un duplexe stable avec le motif riche en C de RsaA. Ce dernier semble se dissocier en cours de migration (**Figure 39 B**). La mutation simultanée des deux sites inhibe complètement la formation du duplexe (données non montrées). L'ARN RsaA mutAGAAC, portant la mutation complémentaire à l'ARNm *mgrA* mutTCTTG, restaure la formation du duplexe, cependant avec une efficacité moindre ( $K_d \approx 50$  nM). De manière surprenante, RsaA

mutTb1, portant la mutation compensatoire à l'ARNm *mgrA* mutH-CDS, ne restaure pas la formation du duplexe (données non montrées). Ceci est d'autant plus surprenant que la restauration de l'interaction implique les mêmes séquences que dans l'interaction boucle-boucle sauvage, ces dernières étant simplement inversées. Une explication serait que les mutations introduites au niveau des tiges-boucles modifieraient la structure des ARN. Des expériences de cartographie en solution sur les ARN mutants seuls seront nécessaires pour vérifier ces hypothèses.

J'ai ensuite testé l'effet de ces mutations sur la traduction *in vitro* de l'ARNm *mgrA* (Figure 39 C). Le couple RsaA / ARNm *mgrA* mutH-CDS montre une accumulation de la protéine MgrA en comparaison à l'inhibition totale observée dans la situation sauvage. Une faible inhibition est observée à la plus haute concentration en RsaA. Les concentrations en ARN sont relativement élevées, et pourraient forcer la formation d'appariements entre les ARN sauvages et mutants qui ne sont pas observées dans les expériences de retard sur gel. RsaA, dans sa forme sauvage, n'inhibe pas la traduction de l'ARNm *mgrA* mutTCTTG ; la protéine MgrA s'accumule même en présence d'excès de ce dernier. A l'inverse, RsaA mutAGAAC, qui est complémentaire à l'ARNm *mgrA* mutTCTTG, restaure une inhibition aussi efficace que celle observée pour l'hybride sauvage (Figure 39 C). L'effet de ces mutations sera par ailleurs étudié sur l'expression de la fusion *mgrA-lacZ* *in vivo*.

Les résultats obtenus montrent que les deux sites d'interaction agissent de manière coopérative pour interagir avec l'ARNm *mgrA*, et former un complexe stable qui est efficace pour ainsi inhiber sa traduction. Contrairement au complexe ARNm *coa*-ARNIII, l'interaction boucle-boucle est importante pour stabiliser le duplexe entre le RBS et le motif riche en C.



**Figure 40** : Les deux sites d'interaction agissent coopérativement dans la formation du duplexe et l'inhibition de la traduction.

**A.** Descriptif des différentes combinaisons d'ARN utilisées dans la suite de l'étude. **B.** L'effet des mutations sur la formation de l'hybride est analysé par des expériences de retard sur gel natif 6%. *RsaA* et *RsaA mut*AGAAC sont marqués radioactivement (5000 cpm/piste) et incubés seul ou avec des concentrations croissantes d'ARNm *mgrA*, *mgrA mutH-CDS* ou *mgrA mutTCTTG* (5, 10, 20, 50, 100, 500 nM). **C.** L'effet des mutations sur la traduction de l'ARNm est évalué par PURE system (Biocomber). 10 pmol d'ARNm sont mis en contact seul ou avec 10 ou 30 pmol de *RsaA* WT ou *RsaA mutTCTTG*. Dans le cas des mutations compensatoires, 10 pmol de l'ARNm *mgrA mutTCTTG* sont incubées seules ou avec 5, 10, 30 ou 50 pmol de *RsaA mut*AGAAC. Les protéines sont séparées sur gel SDS-PAGE 10% et révélées par western avec un anticorps anti-FLAG

## **II.C. RsaA régule la formation de biofilm et de capsules *in vivo***

La délétion de **RsaA** provoque des effets importants sur le protéome de *S. aureus*. En parallèle, j'ai caractérisé une des cibles de ce dernier : l'ARNm codant pour le facteur de transcription MgrA. En plus de ses fonctions dans la régulation de l'autolyse (Ingavale *et al.*, 2003), l'analyse du transcriptome de MgrA montre que de nombreux facteurs de virulence sont sous le contrôle de ce dernier, comme par exemple les gènes impliqués dans la synthèse des capsules polysaccharidiques, les pompes à efflux de médicaments NorA et NorB et plusieurs leucocidines.

*In vivo*, Claire Lays, doctorante dans l'équipe de François Vandenesch à Lyon (INSERM 851) avec qui nous collaborons, a mesuré les effets de la délétion de **RsaA** sur la formation de biofilm et de capsules. Les conséquences sur ces deux processus ont ensuite été évaluées sur l'efficacité de la phagocytose par les macrophages.

### **II.C.1. RsaA régule positivement la synthèse de biofilm.**

La production de biofilm est l'une des causes de la persistance des bactéries en milieu hospitalier. Les cas d'ostéomyélose et de périodontites résultent de l'attachement par biofilm de *S. aureus* au niveau des os et des ligaments. La formation de biofilm peut aussi favoriser la coopération de plusieurs organismes lors d'une infection. Par exemple, le champignon *Candida albicans* est retrouvé dans les biofilms formés par *S. aureus*. Ce dernier peut pénétrer dans les cellules épithéliales de l'hôte et pourrait ainsi favoriser l'entrée de *S. aureus* (Revue Archer *et al.*, 2011).

Les effets de la délétion de *rsaA* ont été étudiés dans trois souches différentes de *S. aureus* : la souche RN1HG (utilisée dans les analyses comparatives du protéome) et deux souches cliniques, Newman et Becker. Ces deux dernières souches ont la particularité de produire beaucoup de capsules polysaccharidiques. La souche Newman possède de surcroît une mutation dans l'opéron *saePQRS* codant pour un TCS et résulte en une expression constitutive de ce dernier, qui est impliqué dans la régulation des gènes d'exotoxines et d'adhésion (**introduction** IV.C.3). Deux autres mutations sont retrouvées au niveau des protéines liant le fibrinogène (FnBP A et B, **introduction** IV.B) (Herbert *et al.*, 2010).

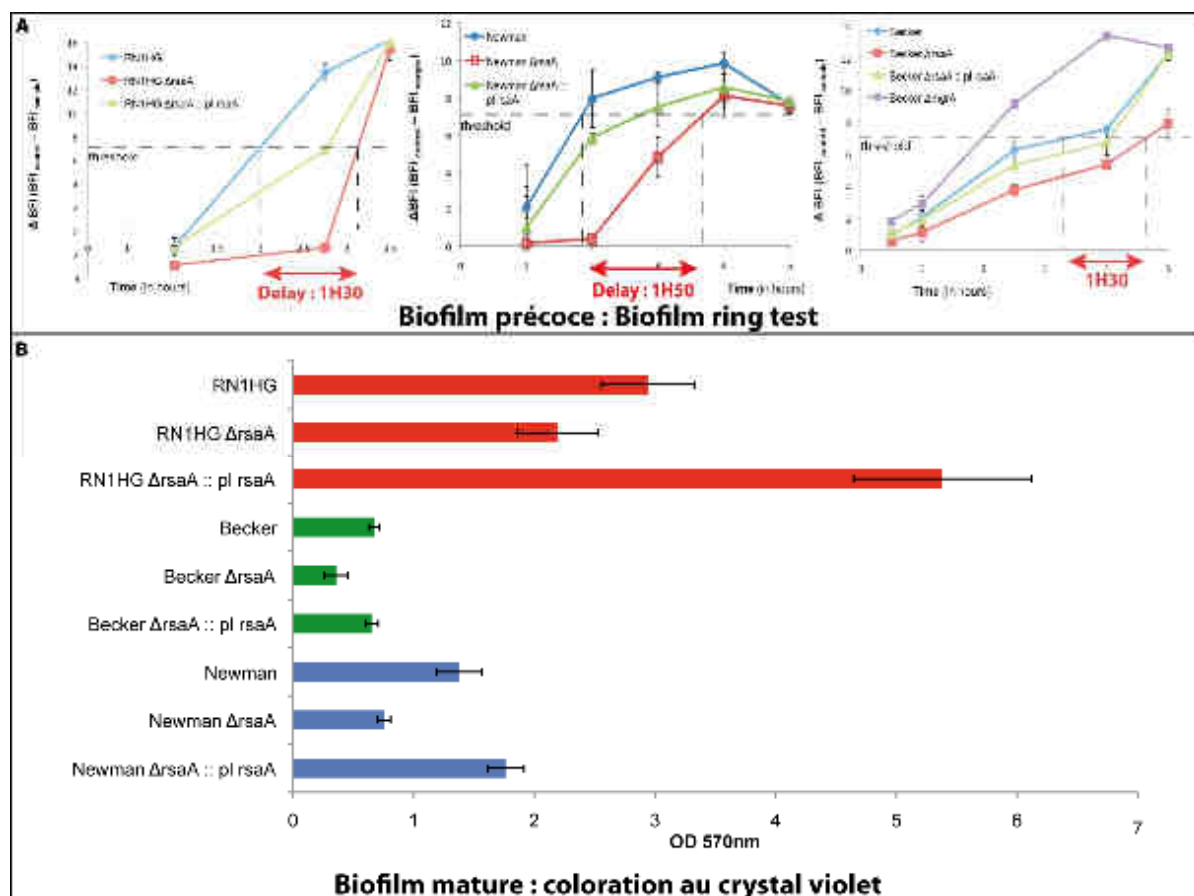
La cinétique de formation du biofilm précoce a été mesurée par Biofilm ring test. Le principe de cette expérience consiste à évaluer la mobilité de billes magnétiques au sein d'une culture de *S. aureus* en fonction du temps. Si du biofilm est formé efficacement, la mobilité des billes sera amoindrie au cours du temps. Les résultats obtenus montrent que pour l'ensemble des souches étudiées, la délétion de *rsaA* se traduit par un délai d'une heure et demi dans la mise en place de biofilm précoce (**Figure 40 A**).

Ensuite, la formation de biofilm mature a été mesurée via la technique de coloration au cristal violet. Là encore, l'expression de biofilm est moins importante pour les souches RN1HG  $\Delta$ *rsaA*, Becker  $\Delta$ *rsaA* et Newman  $\Delta$ *rsaA*. La complémentation des mutants à l'aide d'un plasmide exprimant *RsaA* restaure voire augmente la production de biofilm précoce et mature (**Figure 40 B**).

La nature des biofilms a ensuite été analysée, il semblerait que ceux-ci soit d'origine mixtes, c'est à dire qu'ils sont formés à partir d'ADN, de polysaccharides mais aussi de protéines. Cette diversité est conservée au sein des mutants, néanmoins, il semblerait que la part apportée par les protéines et l'ADN diminue pour les trois fonds génétiques mutants. L'apport des polysaccharides est également moindre dans la souche mutante Newman.

L'ensemble des données collectées *in vivo* montrent que *RsaA* est un acteur positif de la synthèse de biofilm. De manière intéressante, MgrA est décrit comme un répresseur de biofilm (Trotonda *et al.*, 2008). Le phénotype observé pourrait donc être relié à la répression de l'ARNm *mgrA* par *RsaA*.





**Figure 41 : Etude des effets de la délétion de RsaA sur la formation de biofilm.**

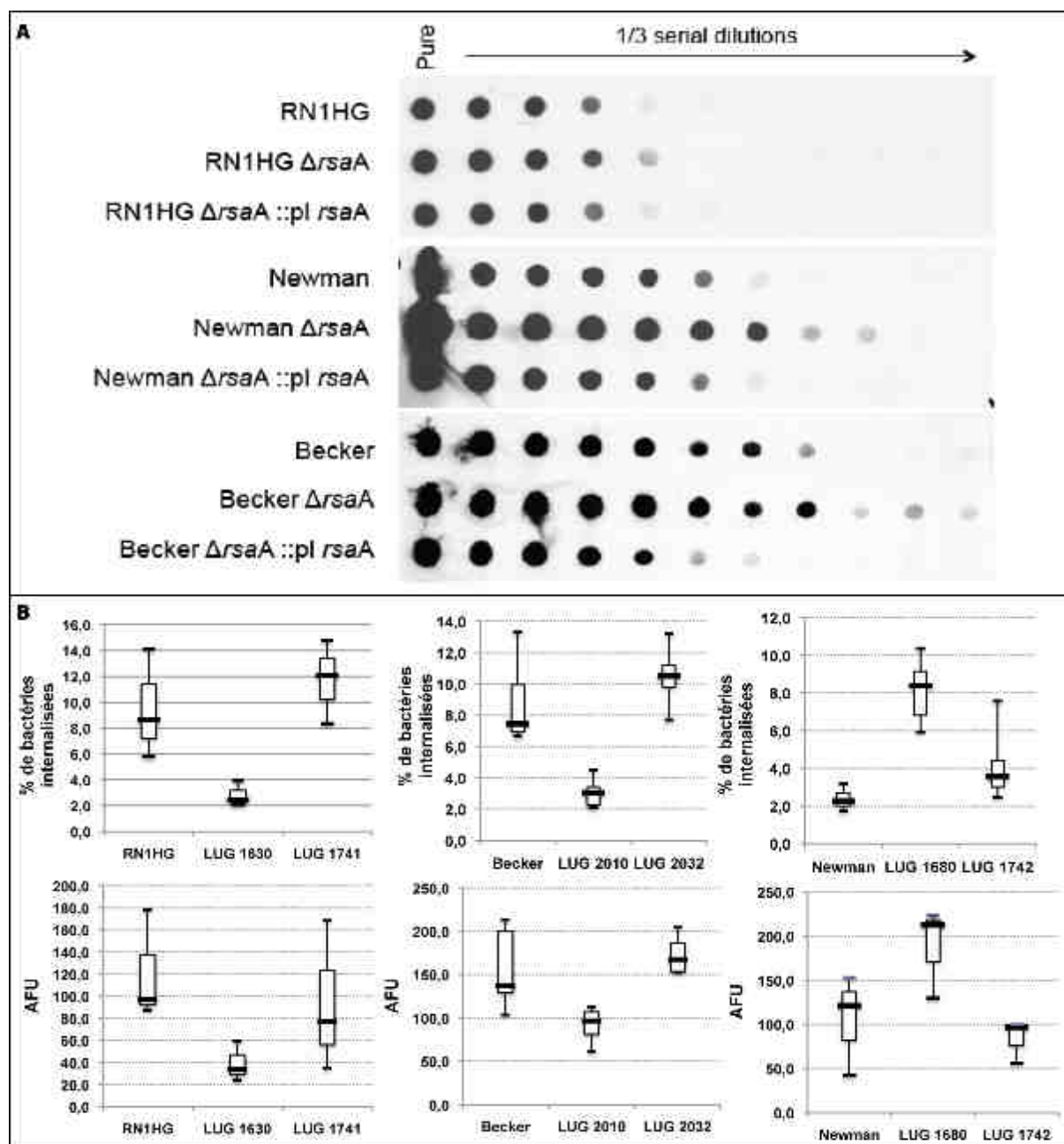
**A. Biofilm ring test.** Les bactéries sont incubées en présence de billes magnétiques dans du milieu riche à 37°C. Aux différents temps de prélèvements, une image est prise avant et après application du champ magnétique. La distance parcourue par les billes donne un indice de biofilm qui est comparé à un contrôle expérimental où les billes sont incubées dans le milieu de culture seulement. Un indice de biofilm supérieur ou égal à 7 indique la formation de biofilm. **B. Coloration au cristal violet.** Des cultures sur la nuit de *S. aureus* normalisées sont incubées 24 h dans du milieu riche supplémenté avec 0,25 % de glucose. Le biofilm mature est coloré au cristal violet (0,1 %) puis subit une étape de lavage à l'acide acétique (33%), et la densité optique à 570 nm est mesurée. Les résultats présentés sur l'histogramme montre la moyenne et l'écart type. L'écart statistique entre la souche parentale et les souches mutantes a été calculé via un "test" de student avec une p-value de 5%.

### II.C.2. RsaA régule négativement la synthèse des capsules

Les capsules polysaccharidiques sont produites par presque toutes les souches de *S. aureus*. Il en existe 8 types différents, les plus représentés étant le type 5 (CP-5, 25% des souches de *S. aureus*) et le type 8 (CP-8, 50% des souches de *S. aureus*). Les CP-8 sont détectées dans le sérum de rat avec un modèle d'endocardite par cathéter. A l'inverse, les CP-5 semblent plus exprimées dans des modèles d'infections nasales chez la souris (revue O'Riordan 2004). Deux cibles de RsaA, l'ARNm *mgrA* et *yabJ-spoVG*, activent la production de capsules (Luong *et al.*, 2006 ; Meier *et al.*, 2007). Cette dernière est-elle perturbée dans des souches de *S. aureus* n'exprimant pas le sARN ?

Dans un premier temps, Claire a estimé la quantité de capsules produites par les souches RN1HG, Newman et Becker (souches testées au niveau du biofilm) et les mutants respectifs délétés de *rsaA* par western "dot blot" en utilisant des anticorps contre la capsule. Il est important de préciser que les souches RN1HG et Newman sont productrices de CP-5 et la souche Becker de CP-8. La production de capsules varie beaucoup entre les différentes souches. Notamment, la souche RN1HG produit peu de capsules en comparaison des deux autres (l'inverse est observé au niveau du biofilm). Néanmoins, de manière globale, les résultats obtenus montrent que tous les mutants produisent plus de capsules (gain d'une dilution entre RN1HG et RN1HG  $\Delta$ *rsaA*, gain de 3 dilutions entre Newman et Newman  $\Delta$ *rsaA* et gain de 4 dilutions entre Becker et Becker  $\Delta$ *rsaA*) (Figure 41 A). Ainsi, *RsaA* réprime la formation des capsules, en accord avec son rôle possible dans l'inhibition de la traduction de *mgrA* et *spoVG*.

L'expression des capsules confère la résistance à la phagocytose. Pour tester cet effet, Claire Lays a testé l'internalisation des bactéries par les macrophages. La comparaison entre les souches parentales et mutantes est très surprenante. En effet, l'augmentation de capsules dans les souches RN1HG  $\Delta$ *rsaA* (CP-5) et Becker  $\Delta$ *rsaA* (CP-8) se traduit comme attendu par une diminution de l'internalisation des bactéries. A l'inverse, la souche Newman délétée de *rsaA*, qui produit plus de capsule, est mieux internalisée par les macrophages. L'expression de *RsaA* et ses conséquences sur la phagocytose varient selon le fond génétique considéré (Figure 41 B).



**Figure 42 : RsaA régule la synthèse des capsules *in vivo*.**

**A. Analyse de l'expression des capsules à l'aide d'anticorps anti-CP-5 (Newman et RN1HG) et anti-CP-8 (Becker).** Après lecture de la  $DO_{600nm}$ , des cultures sur la nuit sont lysées successivement avec de la lysostaphine, puis traitées avec de la DNase I et de la protéinase K. Les échantillons sont ensuite dilués en cascade (dilution de trois fois) puis transférés sur une membrane de nitrocellulose avant d'être incubés avec les anticorps spécifiques. **B. Internalisation par les macrophages.** Les macrophages sont incubés avec les bactéries pendant 2 h à 37°C et 5% de  $CO_2$ . Les bactéries non-internalisées sont lysées. Les macrophages sont ensuite perméabilisés et les bactéries récoltées sont marquées au niveau de leurs membranes avec de la vancomycine couplée à un fluorochrome. Après fixation des cellules, la fluorescence est mesurée avec l'aide d'un cytomètre de flux. Les unités de fluorescence arbitraires (AFU) sont exprimées en pourcentage par rapport aux données obtenues pour la souche sauvage (RN1HG, Newman et Becker respectivement).

Comment expliquer ce résultat contradictoire selon les souches ? Une réponse pourrait venir des différences entre le répertoire des gènes de virulence et des protéines de surface codés par les différentes souches, en particulier au niveau des îlots de pathogénie.

En collaboration avec Isabelle Caldelari et Jean-Marc Reichhart, nous avons analysé le rôle de RsaA dans la virulence sur le modèle *Drosophila melanogaster*. Des infections par voie orale ou par injection ont été réalisées mais aucun effet sur la viabilité des mouches n'a été observé. Récemment, il a été montré que la drosophile est un bon modèle pour suivre la formation des biofilms au niveau de l'intestin (Mulcahy *et al.*, 2011), il sera alors important par la suite de vérifier la présence des bactéries au niveau des voies digestives et analyser la présence ou non de biofilms. Par ailleurs, l'équipe de François Vandenesch est actuellement entrain de tester les effets de la délétion de *rsaA* sur un modèle d'infection par cathéter chez la souris. Ce modèle permet de suivre une infection locale (abcès cutané) lié à la formation de biofilm. L'énumération des bactéries dans le cathéter et aux niveaux des tissus au voisinage proche donnera une idée du rôle de RsaA dans la pathologie chez cet organisme.

**Pour résumer, RsaA régule l'expression de l'ARNm *mgrA* au niveau post-transcriptionnel. Via deux interactions distantes et coopératives, RsaA séquestre le site de fixation de ribosome de l'ARNm dans une région en double-brin, ce qui inhibe la traduction de l'ARNm. *In vivo*, RsaA est un acteur positif dans la régulation du biofilm et un inhibiteur de la synthèse de capsules. Ces phénotypes peuvent être reliés à la régulation de l'expression de MgrA. De manière plus générale, RsaA semble être un ARN régulateur impliqué dans la régulation de la paroi.**

### III. RsaE : un lien dans l'adaptation à l'anaérobiose conservé chez *B. subtilis* et *S. aureus* ?

Parmi les sARN identifiés au laboratoire, *RsaE* est le seul ayant la particularité d'être conservé chez les Bacillaceae comme *B. subtilis*. La structure secondaire est conservée au cours de l'évolution : chez *B. subtilis* et *S. aureus*, elle montre deux séquences répétées contenant un motif riche en C, l'une exposée au niveau d'une boucle apicale, la seconde dans une région en simple brin. La structure secondaire présente également des tiges-boucles aux extrémités 5' et 3'. Comme *RsaA*, *RsaE* est un ARN stable chez *S. aureus* puisque sa demi-vie est estimée à 24 minutes. De manière surprenante, le profil d'expression de *RsaE* dans un milieu riche est inversé entre la souche RN6390 de *S. aureus* (adapté au laboratoire) et *B. subtilis* en phase exponentielle et stationnaire de croissance (Figure 42). En revanche, *RsaE* apparaît exprimé principalement en phase exponentielle dans les souches de *S. aureus* cliniques (Song *et al.*, 2012). Grâce à des analyses comparatives du transcriptome et du protéome des souches RN6390 et RN6390  $\Delta$ *rsaE*, plusieurs cibles de *RsaE* ont été révélées chez *S. aureus* (voir introduction IV.D.3) : l'ARNm *oppB*, et les ARNm *sucC* et *sucD*. *RsaE* cible le RBS de ces ARNm pour inhiber la traduction. (Geissmann *et al.*, 2009). Ainsi *RsaE* régulerait le métabolisme central lorsque les sources carbonées viennent à manquer à l'entrée de la phase stationnaire (Bohn *et al.*, 2010 ; Geissmann *et al.*, 2009).

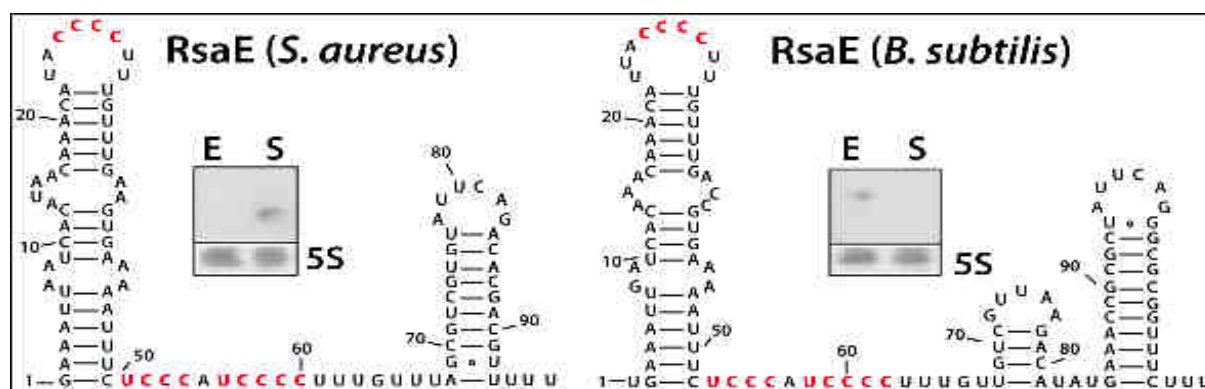


Figure 43 : *RsaE* est conservé chez *S. aureus* et *B. subtilis*.

Structures secondaires des deux ARN. Les motifs riches en C retrouvés dans l'ARNIII et d'autres ARN régulateurs de *S. aureus* sont indiqués en rouge. L'expression de *RsaE* a été analysée par northern blot sur des extraits d'ARN totaux de *S. aureus* (souche RN6390) et *B. subtilis* (souche 168) mises en cultures dans le milieu riche BHI

pendant 2 h (phase exponentielle de croissance E) et 6 h (phase stationnaire de croissance S) (Geissmann *et al.*, 2009).

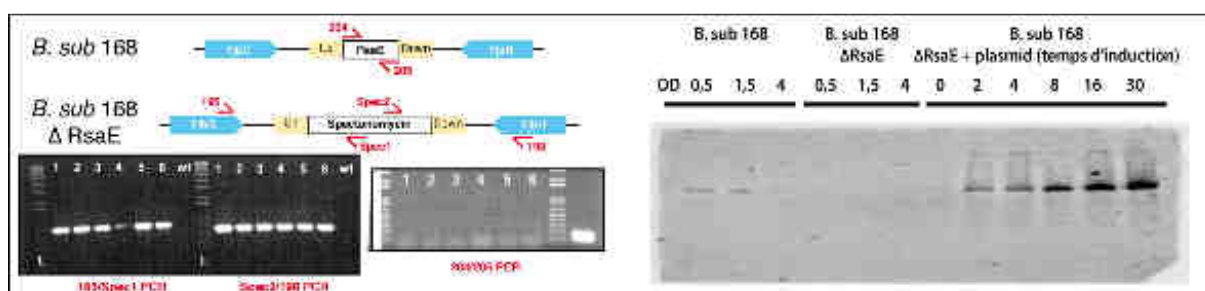
**Comment les fonctions ont évoluées entre les bactéries distantes dans l'évolution comme *S. aureus* et *B. subtilis* qui existent dans des niches écologiques bien différentes ? Quelles sont les ARNm régulés par *RsaE* chez *B. subtilis*?**

### III.A. Etude fonctionnelle de *RsaE* chez *B. subtilis*

De manière similaire à *RsaA*, l'étude fonctionnelle de *RsaE* a été basée sur l'analyse comparative du protéome et du transcriptome. Pour se faire, la première étape indispensable a été de construire un mutant de *RsaE* chez cette bactérie.

#### III.A.1. Construction d'une souche *B. subtilis* 168 délétée de *rsaE*

En collaboration avec Efthimia Lioliou, nous avons entrepris de construire une souche de *B. subtilis* 168 délétée de *rsaE*. La mutation consiste en un remplacement allélique du gène *rsaE* par une cassette de résistance à la spectinomycine. Le changement de séquence a été validé par PCR en utilisant des couples d'amorces appropriés et l'inhibition de l'expression de *RsaE* a été visualisée par northern. En parallèle, nous avons également construit un plasmide portant le gène *rsaE* sous le contrôle d'un promoteur inducible à la tétracycline pour compléter la souche mutante (Figure 43).



**Figure 44 : Construction de la souche de *B. subtilis* 168  $\Delta$ rsaE et complémentation.**

Le remplacement de *rsaE* par la cassette de résistance à la spectinomycine est vérifié par PCR dans 6 clones différents en utilisant des couples d'amorces capables de discriminer le contexte sauvage du mutant. L'expression de *RsaE* dans les souches sauvage, mutante et complétement a été analysée par northern à partir d'un gel polyacrylamide 5% en conditions dénaturantes. Les échantillons ont été prélevés à une  $DO_{600nm}$  de 0,5 1,5 et 4 dans du milieu YT de Difco. Pour les pistes correspondant à l'induction de plasmide, ce dernier est induit par l'ajout de tétracycline-HCl pendant 2, 4, 8, 16 et 30 minutes à une  $DO_{600nm}$  de 0,5.

De manière intéressante, les prédictions des cibles de *RsaE* chez *B. subtilis*, effectuées avant mon arrivée au laboratoire, montrent des cibles liées au métabolisme des carbohydrates, dont l'ARNm *sucC* qui est déjà validée chez *S. aureus*, et l'ARNm *cstA* (Geissmann *et al.*, 2009). Des prédictions plus récente avec les programmes **TargetRNA** et **IntaRNA** montrent également la cible codant pour l'ARNm *cstA*, une protéine induite en réponse à un stress en carbone (données non montrées). Néanmoins, ce duplexe n'est prédit que chez *B. subtilis*. De manière générale, aucun des 10 meilleurs candidats prédits n'est retrouvé simultanément chez les deux bactéries.

Par northern blot, j'ai testé l'expression de l'ARNm *cstA* et *sucC* entre les souches 168, 168  $\Delta$ *rsaE* et 168  $\Delta$ *rsaE* complémentée. Les résultats obtenus montre que l'expression de ces ARNm ne varie pas entre les souches. Néanmoins, les signaux obtenus sont faibles. Ces expériences gagneraient à être refaites en utilisant des sondes plus sensibles (technique DIG) en cultivant les bactéries dans un milieu pauvre en glucose pour tenter d'établir un stress lié au carbone.

### III.A.1. *RsaE* est induit en réponse à un stress anaérobie chez *B. subtilis*

De manière similaire à *RsaA*, nous avons entrepris de croiser les résultats de plusieurs approches globales pour établir une liste de candidats potentiels régulés par *RsaE* chez *B. subtilis*. L'analyse comparative du transcriptome, effectuée par Sylvain Durand de l'équipe de Ciaran Condon, et l'analyse comparative du protéome effectuée avec Phillippe Hammann ont été effectuées en parallèle.

Les extraits de protéines totales des souches de *B. subtilis* 168  $\Delta$ *rasE*, 168 $\Delta$ *rasE* complémentée sans et après 5 min d'induction à la tétracycline (le plasmide portant *RsaE* est sous le contrôle d'un promoteur inductible) ont été analysés par la technique DIGE. De manière surprenante, les intensités des signaux de fluorescence obtenus à partir des extraits de protéines réalisés avant et après induction montrent des résultats similaires. L'analyse de l'expression par northern montre un niveau très faible de l'ARN lorsque le plasmide n'est pas induit (**Figure 43**). Néanmoins, il n'égale pas les niveaux observés dans la souche sauvage ou dans la souche complément après induction du plasmide pendant 5 min. Une seconde

analyse a été menée en séparant les mêmes extraits de protéines sur des gels différents, colorés au bleu de coomassie. Les résultats montrent peu de différences entre la souche mutante et la souche exprimant le plasmide non induit. A l'inverse, la synthèse de nombreuses protéines varie entre la souche mutante et la souche qui surproduit *RsaE* pendant 5 min.

**Tableau 7 : Résultats des analyses comparatives du protéome pour *RsaE* chez *B. subtilis*.**

Gène	Fonction	Régulation RsaA	Duplexe <i>RsaE</i>
<b>Analyses DIGE</b>			
<i>yjcG</i>	2'-5' RNA Ligase hypothétique	RÉPRESSION	oui*
<i>dhbA</i>	2,3-dihydroxybenzoate-2,3-dehydrogenase	RÉPRESSION	oui
<i>resD</i>	ORFX17, régulateur réponse anaérobie	RÉPRESSION	oui
<i>sigB</i>	Facteur $\sigma^B$	ACTIVATION	
<i>ssbAB</i>	Protéine de liaison à l'ADN simple brin	RÉPRESSION	
<i>ywfl</i>	Hème peroxidase	RÉPRESSION	
<i>tpx</i>	Peroxi-redoxine hypothétiques	RÉPRESSION	
<b>Analyses coomassie</b>			
<i>ycgT</i>	Thioredoxin reductase hypothétique	RÉPRESSION	oui
<i>dhbA</i>	2,3-dihydroxybenzoate-2,3-dehydrogenase	RÉPRESSION	oui
<i>tpi</i>	triose phosphate isomerase	RÉPRESSION	
<i>hprT</i>	hypoxanthine-guanine phosphoribosyltransferase	RÉPRESSION	
<i>yceD</i>	protéine d'adaptation au stress	RÉPRESSION	
<i>rplD</i>	50S ribosomal protein L4	RÉPRESSION	
<i>yvyD</i>	ribosome-associated sigma 54 modulation protein	RÉPRESSION	
<i>ahpC</i>	alkyl hydroperoxide reductase	RÉPRESSION	
<i>ytpR</i>	protéine hypothétique	ACTIVATION	
<i>pyrB</i>	aspartate transcarbamoylase	ACTIVATION	
<i>pyrAA</i>	carbamoyl phosphate synthase	ACTIVATION	
<i>ytpR</i>	protéine hypothétique	ACTIVATION	
<i>ctc</i>	50S ribosomal protein L25/general stress protein Ctc	ACTIVATION	
<i>ppaC</i>	manganese-dependent inorganic pyrophosphatase	ACTIVATION	

Résultats des analyses comparatives du protéome de *B. subtilis* par DIGE (trois souches comparées : 168  $\Delta$ rasE, 168  $\Delta$ rasE complétée (sans induction) et 168  $\Delta$ rasE complétée (avec induction de 5' avec de la tétracycline)) et par coloration des gels au bleu de coomassie (souche 168  $\Delta$ rasE et 168  $\Delta$ rasE complétée après induction avec de la tétracycline). Les prédictions de duplexes ont été réalisées manuellement avec des programmes de repliements *mFOLD* et *contraFOLD*. Les cibles en violet sont détectées dans les analyses comparatives du transcriptome.

De manière surprenante, seulement une protéine est retrouvée entre les deux analyses. Les résultats présentés ici sont préliminaires et nécessitent d'être validés avec un second duplicat biologique. J'ai décidé de les présenter car plusieurs correspondances ont été trouvées avec les résultats de l'analyse comparative du transcriptome effectuée par Sylvain Durand (Tableau 7). Les résultats obtenus par bleu de coomassie entre la souche mutante et



celle complémentée sans induction de plamide ne sont pas montrés, puisque seules les protéines portées par le plasmide sont détectées,

Je n'ai pas réussi à montrer de convergence entre les résultats de protéomique et les prédictions de cibles par les programmes *TargetRNA* et *RNA Predator*. Néanmoins, plusieurs duplexes ont été prédits manuellement via l'utilisation des programmes de repliements *mfold* et *contraFOLD* entre *RsaE* et la région 5'UTR de certains ARNm (**Figure 44**).

Parmi l'ensemble des candidats, l'ARNm *resD* codant pour le système à deux composants *resDE* chez *B. subtilis* a retenu notre attention. En effet, une analyse de la région promotrice de *RsaE* montre un site de fixation potentiel du régulateur de la transcription ResD, qui répond à l'anaérobiose, situé à environ 70 nts en aval de TSS chez *B. subtilis*. Il serait donc possible que *RsaE* soit induit en réponse à une faible concentration en oxygène et qu'il forme une boucle de rétrocontrôle négative avec son inducteur. Effectivement, Sylvain Durand a montré que *RsaE*, dans la souche de *B. subtilis* 168, s'accumule lorsque les bactéries sont cultivées en anaérobiose. De manière intéressante, le TCS *resDE* est homologue au TCS *srrAB* lui-même impliqué dans l'adaptation à l'anaérobiose et l'expression des facteurs de virulence par l'activation d'*agr* chez *S. aureus* (Pragman *et al.*, 2004). Le site de fixation du régulateur de la transcription est également conservé, de même que le duplexe entre *RsaE* et le gène *srrA* (**Figure 44**).

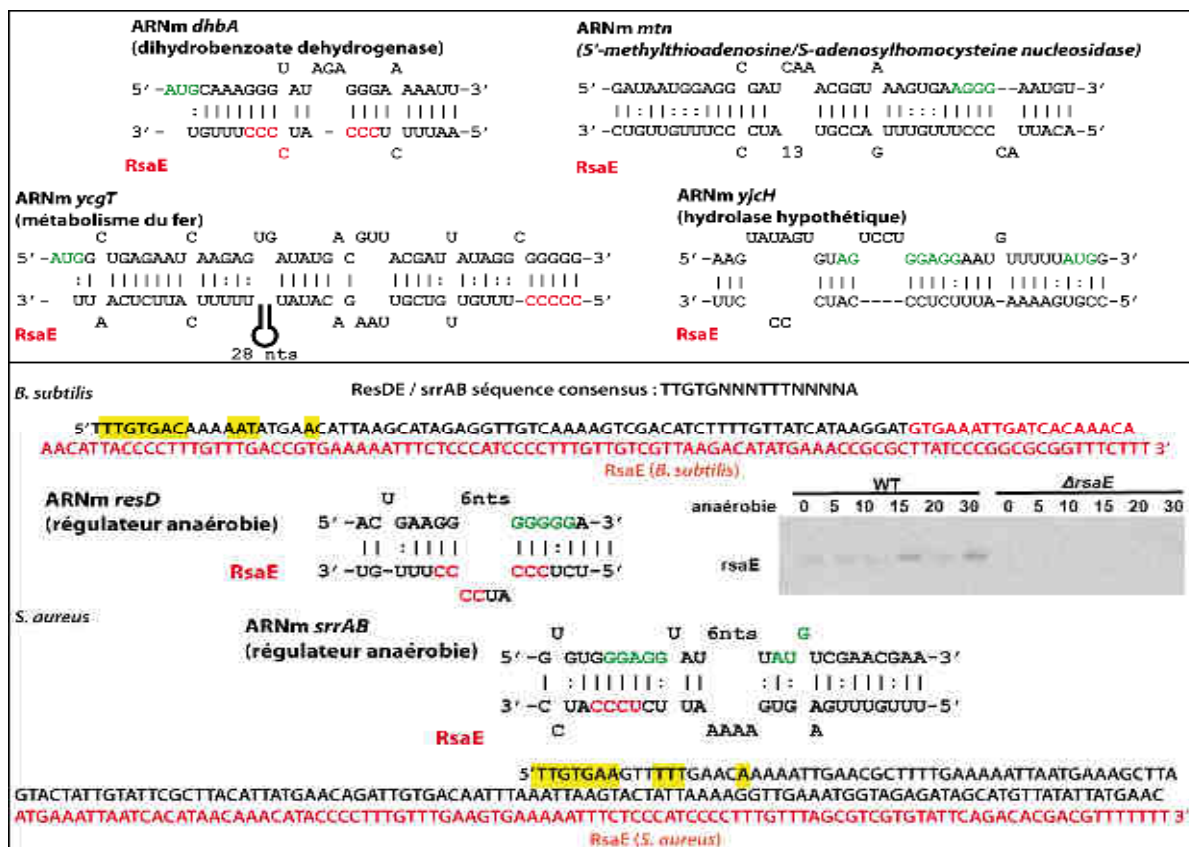


Figure 45 : Candidats pour une régulation par RsaE.

Les motifs riche en C de RsaE sont indiqués en rouge, la SD et l'AUG des ARNm en vert. En bas, localisation du site de fixation du régulateur de réponse à l'anaérobiose ResD ou SrrA chez *S. aureus* et *B. subtilis* d'après Geng *et al.*, 2007, et les duplexes prédits chez *S. aureus* et *B. subtilis* respectivement. L'induction de RsaE en réponse au stress a été mesurée chez *B. subtilis* pendant 0, 5, 10, 15, 20 et 30 min lorsque la culture de bactérie a atteint une DO<sub>600nm</sub> de 0,5 dans le milieu riche YT.

### III.A.2. Perspectives pour ce projet

La caractérisation des cibles de RsaE identifiées à partir des approches globales est en cours chez *B. subtilis*. Pour se faire, de manière similaire au travail effectué sur RsaA, ces dernières devront être validées à la fois *in vivo* et *in vitro*.

Des premiers résultats encourageant sont obtenus avec l'ARNm *ycgT*. En effet, l'analyse de l'expression de cet ARNm par northern blot montre une nette accumulation de ce dernier dans la souche de *B. subtilis* délétée du gène *rsaE* en comparaison de la souche sauvage. En parallèle, chez *S. aureus*, des expériences préliminaires n'ont pas mis en évidence une accumulation de RsaE en réponse à un stress anaérobie (données non montrées). Néanmoins, ces expériences ont été menées dans le milieu riche BHI. La réponse à l'anaérobiose devra être testée dans différents milieux et en utilisant différents protocoles (volume et durée du stress) pour confirmer ou infirmer ces premiers résultats.

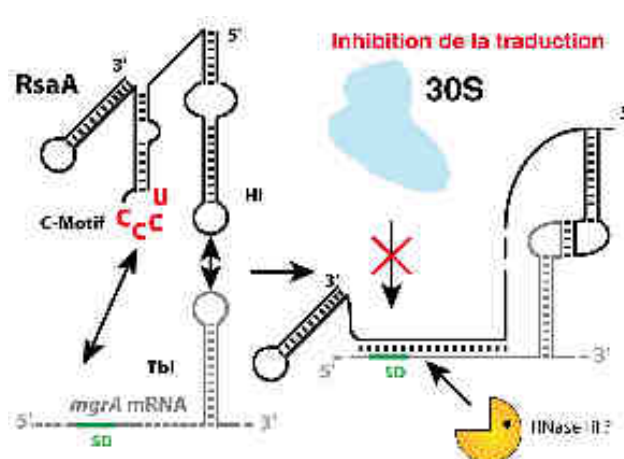


# Discussions & Perspectives



## I. Le réseau de régulation de RsaA est connecté au système de densité cellulaire

Durant ce travail de doctorat, j'ai caractérisé le mécanisme de régulation de RsaA sur l'ARNm *mgrA* : RsaA s'apparie à l'ARNm en utilisant deux sites d'interaction distants. Le motif riche en C de RsaA piège la séquence SD dans une région en double brin, empêchant ainsi



l'initiation de la traduction. Le second site d'interaction fait intervenir un appariement boucle-boucle entre la tige-boucle en 5' de RsaA et une tige-boucle localisée dans la séquence codante de l'ARNm. Les deux interactions agissent de manière coopérative pour une liaison et

une inhibition optimale de la traduction (Figure 45).

Figure 46 : Modèle de régulation de l'ARNm *mgrA* par RsaA.

D'un point de vue structural, il est intéressant de souligner que l'interaction boucle-boucle n'est pas reconnue *in vitro* par la RNase III de *S. aureus*. Pourtant cette enzyme a été décrite comme très efficace dans la reconnaissance et le clivage spécifique des interactions boucle-boucle entre l'ARNIII et les ARNm *rot* et *coa* (voir Résultats I.C). Une explication pourrait venir des boucles internes qui interrompent les hélices intramoléculaires et pourraient gêner la reconnaissance et la fixation du RBD de la RNase III. Une autre différence notable par rapport à la régulation ARNIII / ARNm *coa* est que l'interaction boucle-boucle est essentielle dans l'efficacité de la régulation.

La régulation de l'ARNm *mgrA* peut être corrélée aux observations *in vivo* sur la synthèse de biofilm et des capsules polysaccharidiques (Luong *et al.*, 2006). De manière intéressante, le régulon de MgrA est lié à celui de l'opéron *agr*. En effet, MgrA active l'expression du système *agr* (Ingavale *et al.*, 2005) et se lie au promoteur de l'ARNm *hla* pour activer sa transcription. MgrA inhibe également SarS, un activateur de l'expression de la

protéine A (Ingavale *et al.*, 2005). De manière contradictoire, MgrA active l'expression de SarZ (Ballal *et al.*, 2009), effecteur positif de l'opéron *agr* (Tamber *et al.*, 2009), ainsi que SarX qui est un répresseur de ce système (Manna *et al.*, 2006). Le régulon de MgrA est donc complémentaire au système *agr* et permet l'expression des exotoxines au profit des facteurs d'adhésion et du biofilm. En régulant l'expression de l'ARNm *mgrA*, RsaA régulerait aussi l'expression de l'ARNIII et du système *agr* indépendamment de la densité cellulaire, expliquant ainsi les niveaux d'expression plus faible de l'ARNIII dans les souches exprimant  $\sigma^B$  (Figure 46) (Ballal *et al.*, 2009).

Néanmoins, deux cas de figure doivent être distingués à mon sens : l'effet de l'expression de  $\sigma^B$  en phase exponentielle et stationnaire de croissance, et l'effet de l'induction rapide de  $\sigma^B$  en réponse à un stress environnemental. Il serait donc intéressant d'évaluer le niveau d'expression de RsaA, de l'ARNIII, de l'ARNm *mgrA* et *agr* par northern pour confirmer ou infirmer ces hypothèses.

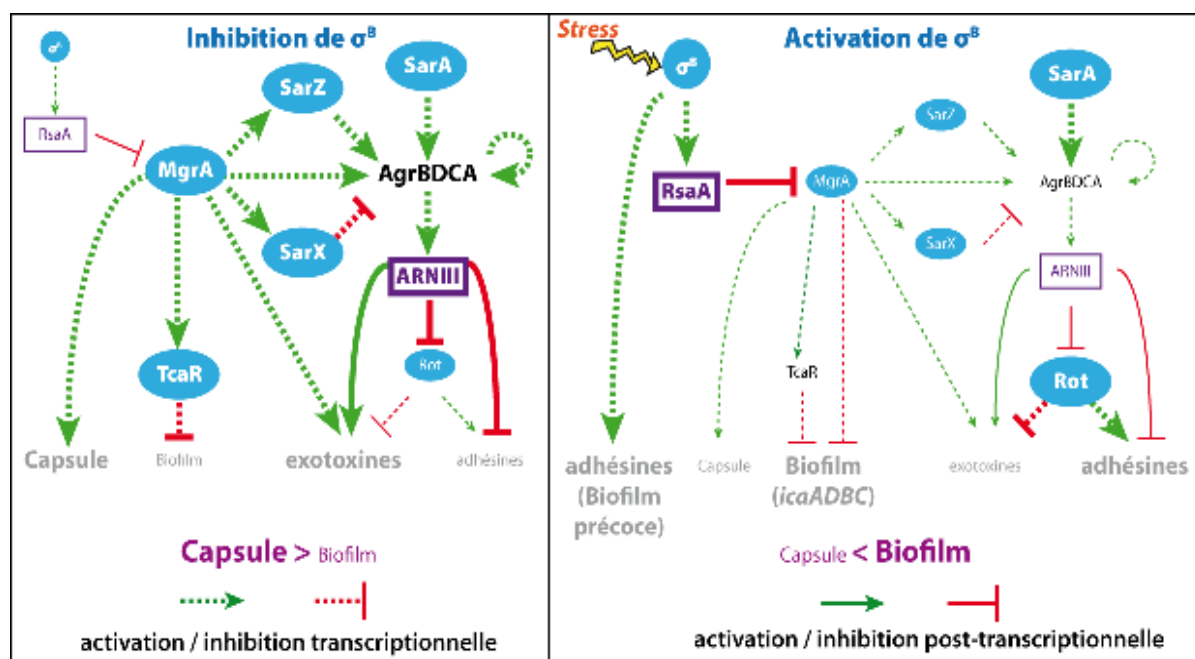


Figure 47 : Réseau de régulation de MgrA et de RsaA avant et après activation de  $\sigma^B$ .

Les activations sont représentées par des flèches (en vert), les inhibitions par des traits finissant par une barre (en rouge). Une régulation au niveau transcriptionnel est symbolisée par des pointillés, une régulation post-transcriptionnelle par des traits pleins. Les mots en gras et les flèches plus épaisses correspondent aux régulations dominantes selon l'expression de  $\sigma^B$ .

## II. D'autres cibles régulées par RsaA ?

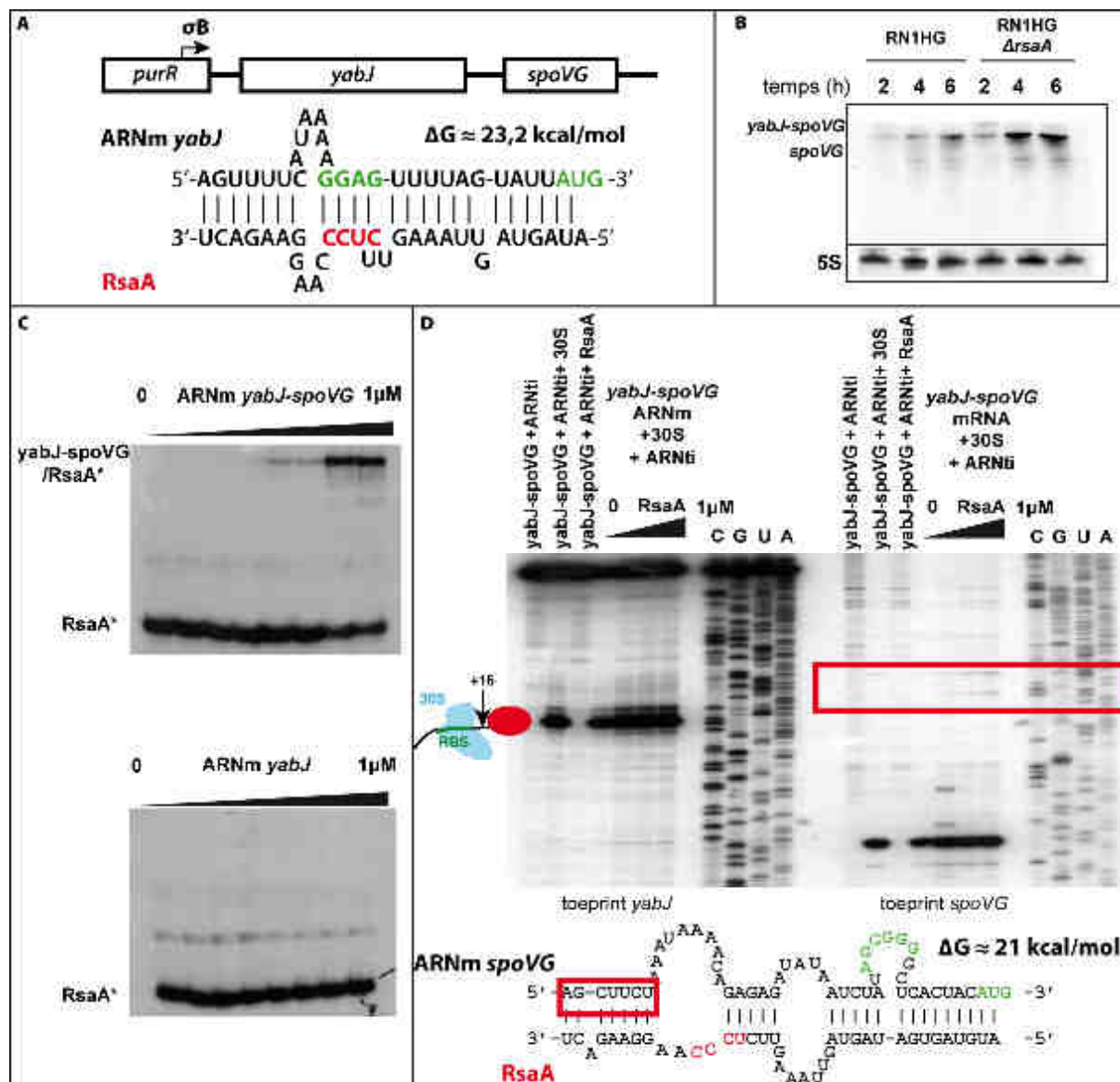
Si les données expérimentales ont conduit à la validation de l'ARNm *mgrA* comme cible directe de RsaA, les analyses prédictives et globales que j'ai effectuées montrent aussi d'autres cibles potentielles. Plusieurs pistes restent aussi à explorer : RsaA régule-t-il l'expression d'ARNm uniquement via le ciblage du RBS ? RsaA régule-t-il directement l'activité de protéines ?

### II.A. Les ARNm *yabJ-spoV*, *sa1956* et *glyS*

Parmi les cibles identifiées par protéomique comme étant réprimées par RsaA, l'ARNm *yabJ-spoVG* avec RsaA forme un duplexe stable, où le motif riche en C se fixe au RBS de l'ARNm *yabJ* (Figure 47 A). De manière intéressante, l'analyse de l'expression de l'ARNm par northern montre une accumulation de l'ARNm en phase stationnaire de croissance dans la souche RN1HG  $\Delta$ *rsaA* (6 h de culture en milieu riche BHI) (Figure 47 B). Des expériences de retard sur gel montrent que RsaA est capable de se fixer à l'ARNm. Néanmoins, la stabilité de cette interaction est faible avec une constante de dissociation de l'ordre de 1  $\mu$ M. De plus, dans ces expériences, RsaA n'est pas capable de se lier à un fragment d'ARN comprenant uniquement la région codante pour l'ARNm *yabJ*, région pour laquelle le duplexe le plus stable est prédit (Figure 47 C). Des expériences de toeprint suggèrent également que RsaA n'inhibe pas la formation du complexe d'initiation de la traduction au niveau des RBS des deux CDS de l'opéron (Figure 47 D). Contrairement aux résultats des approches prédictives, il semblerait que RsaA se lie à l'ARNm *yabJ-spoVG* en ciblant une région différente du RBS des deux CDS. Une analyse de la séquence de l'ensemble de l'opéron n'a pas mis en lumière d'autres motifs riches en G pouvant s'apparier au motif riche en C de RsaA. Les expériences de toeprint montrent une pause dépendante de RsaA dans la région 5' UTR de l'ARNm *spoVG*. Un duplexe est prédit pour cette région, mais il ne fait pas intervenir le motif riche en C de RsaA et laisse la séquence Shine et Dalgarno de l'ARNm accessible au ribosome. Il sera important de définir par cartographie en solution les régions protégées par RsaA de l'ARNm pour ensuite effectuer des mutations spécifiques *in vivo* et étudier leur effet sur le contrôle.



Quel est le mécanisme de régulation de *RsaA* sur l'ARNm *yabJ-spoVG* ? *In vivo*, existe-il des partenaires protéiques capables de favoriser l'interaction entre les deux ARN et d'assister *RsaA* dans la répression de l'ARNm ?



**Figure 48 : L'ARNm *yabJ-spoVG*, une autre cible de *RsaA* ?**

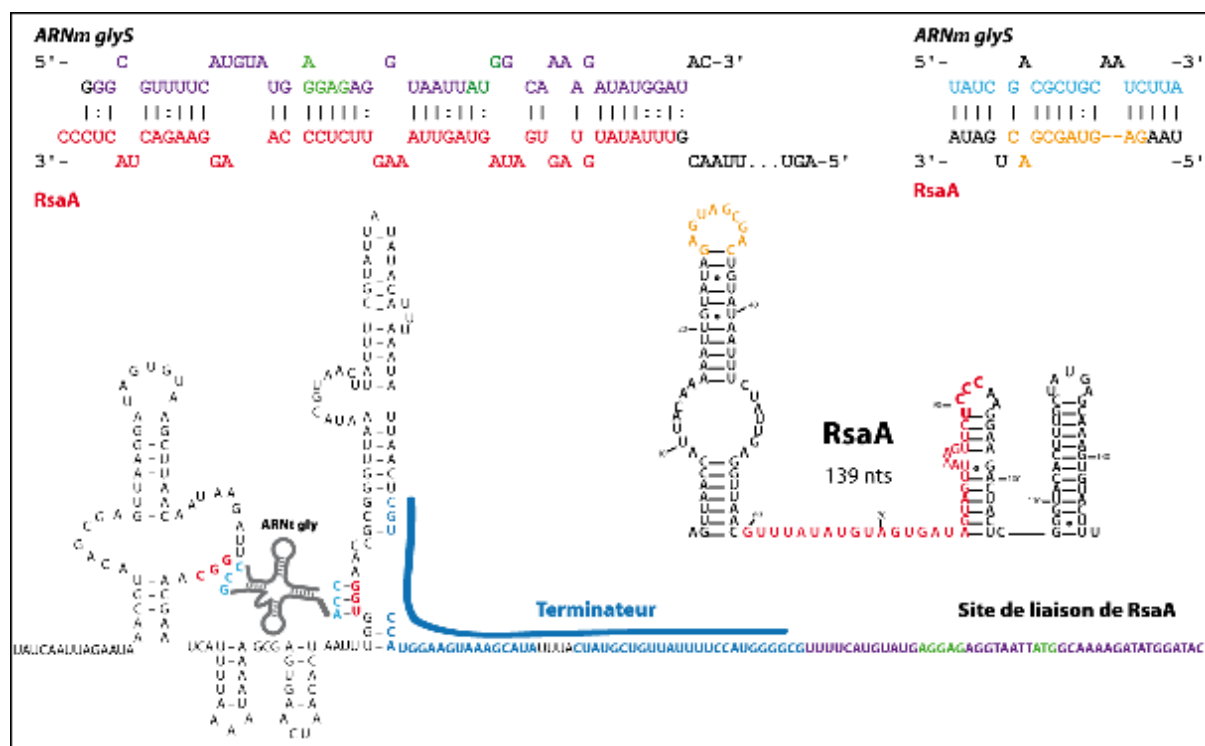
**A.** Contexte génomique de l'opéron. **B.** Analyse de l'expression de l'ARNm *yabJ-spoVG* par northern-blot (gel agarose 1%) dans les souches RN1HG et RN1HG  $\Delta$ *RsaA*. Les ARN totaux ont été prélevés après 2, 4 et 6 heures de culture à 37°C en milieu riche BHI. **C. Expérience de retard sur gel.** *RsaA*, marqué à son extrémité 5' au  $Y$ - $P^{32}$  (5000 cpm/ligne) est incubé seul ou en présence de concentrations croissantes d'ARNm *yabJ-spoVG* ou d'ARNm *yabJ* (1, 5, 10, 20, 100, 500 et 1000 nM) puis séparé sur gel natif 6%. La fraction de *RsaA* liée à l'ARNm a été estimée en comparant l'intensité de la bande correspondante à l'intensité de toute la piste. La  $K_D \approx 1 \mu$ M correspond à la concentration pour laquelle la moitié de *RsaA* forme un duplex avec l'ARNm *yabJ-spoVG*. **D. Expérience de toeprint.** L'ARNm *yabJ-spoVG* (0,2 pmol) est incubé en présence d'un oligonucléotide marqué radioactivement à son extrémité 5' P au  $Y$ - $P^{32}$  (25000 cpm/réaction) et s'hybride soit au niveau de la phase codante de l'ARNm *yabJ* (toeprint *yabJ* position +116 à +132), soit au niveau de la phase codante de l'ARNm *spoVG* (toeprint *spoVG* position +50 à +66). Ce dernier sert d'amorce à la RT pour la synthèse d'un brin d'ADNc. La formation du complexe ternaire entre l'ARNm, le ribosome de *S. aureus* (200 nM) et l'ARNt initiateur est visualisée par une pause de la RT à la position +16 par rapport à l'AUG (deuxième piste). Lorsque l'ARNm est

incubé en présence de concentrations croissantes de *RsaA* (100, 200, 500 et 1000 nM, piste 4-8), la pause persiste. La première piste correspond à l'ARNm *yabJ-spoVG* seul et la troisième piste à l'ARNm en présence de *RsaA* (contrôles d'incubation). Une pause de la RT, spécifique de *RsaA* apparaît en amont de la SD de l'ARNm *spoVG*. Un duplex potentiel est possible à cette région entre les deux ARN. Ce dernier ne fait intervenir ni le motif riche en C de *RsaA*, ni la séquence Shine et Dalgarno de l'ARNm.

Deux autres cibles potentielles de *RsaA*, que j'avais écartées jusqu'à présent, seraient reliées au métabolisme de la paroi (voir [Discussion et perspectives III](#)). Tout d'abord, des appariements entre l'ARNm *sa1956* et *RsaA* ont été prédits à la fois par les approches *IntaRNA* et *TargetRNA*. *SA1956* serait un régulateur de l'autolyse et fait partie du régulon de *MgrA* (Luong *et al.*, 2006). Dans plusieurs souches de *S. aureus* (N315, ECT-R2, Mu50 etc...), le locus *sa1956* est interrompu par le transposon Tn554. Mais de manière intéressante, ce locus semble être intact dans les souches Newman et NCTC8325 que nous utilisons au laboratoire.

Enfin, de manière surprenante, la synthèse de la glycyl-ARNt synthétase (GlyRS) codée par le gène *glyS* apparaît sous la dépendance de *RsaA* (voir [Résultats II.A.1](#) et [Figure 48](#)). L'expression de cet ARNm est sous la dépendance de l'anti-terminaison médiée par l'interaction de l'ARNt<sup>Gly</sup> non aminoacylé avec la région 5'UTR de l'ARNm au niveau de la boîte T (T-box) (Yusef *et al.*, 2005 ; Grundy *et al.*, 2005). Le site d'interaction avec *RsaA*, faisant intervenir la région juste en 3' de la boîte T, suggère que la régulation potentielle de cet ARNm serait indépendante de la régulation par anti-terminaison de la transcription. De plus, un motif complémentaire à la tige-boucle en 5' de *RsaA* est retrouvé au niveau d'une boucle interne de la séquence codante de l'ARNm. Cette boucle expose trois résidus, CGC, qui sont capables d'interagir avec la tige-boucle en 5' de *RsaA* pour former un duplex court contenant 7 interactions GC ([Figure 48](#)).

Les implications biologiques et les réseaux de régulation qui pourraient découler de la régulation de ces cibles seront discutés plus loin (Voir [Discussions & Perspectives III](#)).



**Figure 49 : L'ARNm *glyS*, une cible potentielle de *RsaA* ?**

Représentation des duplexes impliquant le RBS de l'ARNm *glyS* (en violet) avec la région située en 3' du motif riche en C de *RsaA* (en rouge) (d'après les données obtenues par le programme *intaRNA* (voir Résultats)). Un appariement entre la tige-boucle en 5' de *RsaA* (en orange) et une boucle localisée dans la CDS de l'ARNm *glyS* (en bleu). La structure de la boîte T, liant un ARNt Gly non aminoacylé, empêche le formation d'un terminateur de la transcription (en bleu). Le potentiel site de liaison de *RsaA* se situerait juste après, indiquant que la régulation serait indépendante de l'anti-terminaison.

## II.B. Un motif riche en C et des topologies variées pour la liaison ARNm / sARN : une adaptation à l'absence d'Hfq?

Les ARN régulateurs de *S. aureus* présentent deux caractéristiques particulières : ils sont généralement très structurés et la plupart portent des motifs riches en C, accessibles au niveau de régions en simple brin (Figure 49). De plus, aucun des ARN régulateurs de *S. aureus*, dont la fonction a été caractérisée à ce jour, ne nécessite l'aide de la protéine Hfq pour accoupler ses fonctions régulatrices (voir Introduction IV.D). **Comment les ARN régulateurs exercent-ils leurs fonctions sans l'aide de cette protéine ?**

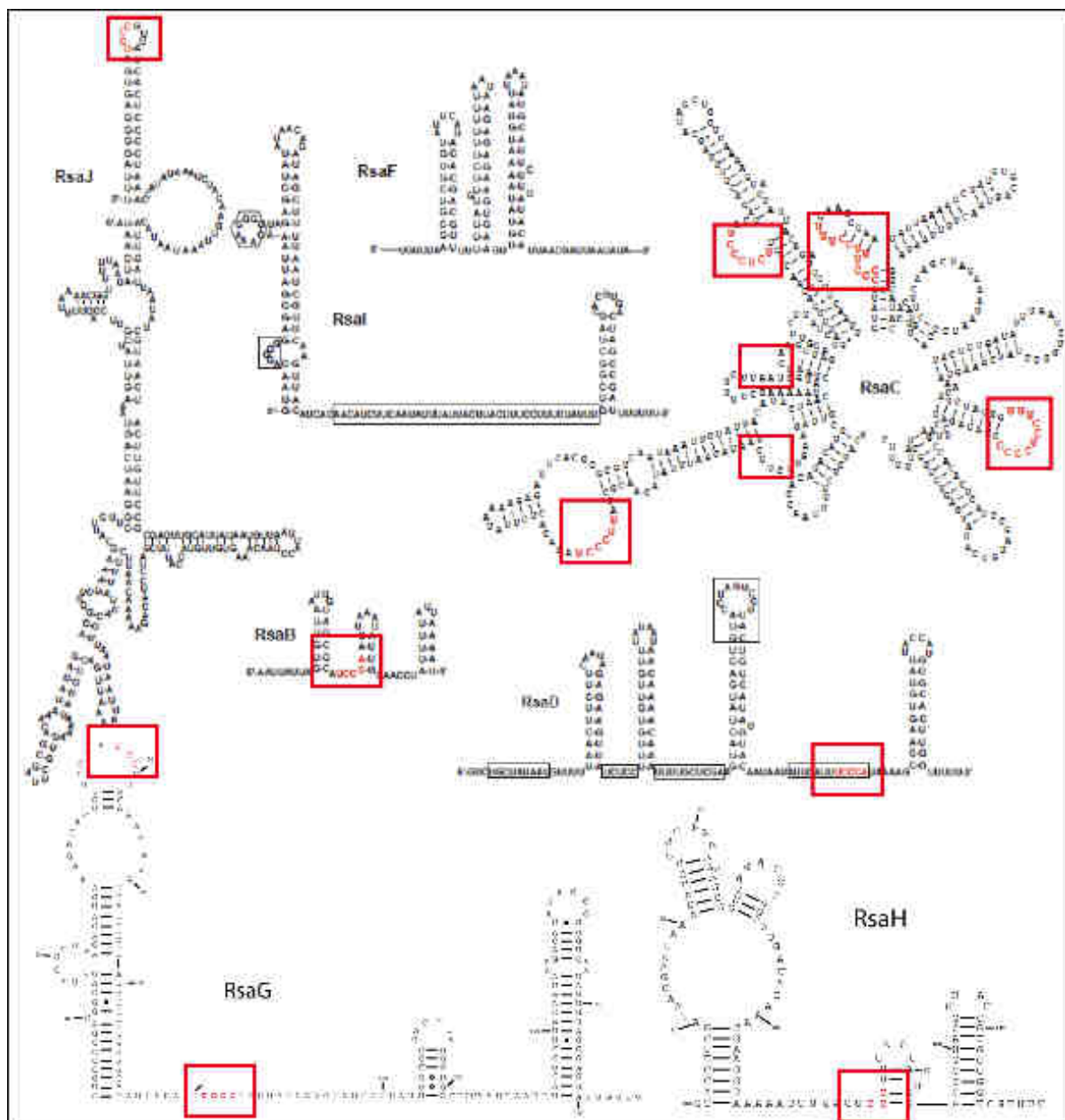
**Des ARN structurés résistant à l'assaut des ribonucléases.** Pour rappel, l'une des actions de Hfq chez *E. coli* est la protection des ARN contre la dégradation par les ribonucléases (voir Introduction III.C.1). Chez *S. aureus*, les ARN régulateurs possèdent une tige-boucle en 5' et un terminateur de transcription Rho-indépendant en 3'. Ces structures

protègent les sARN contre l'action des exoribonucléases 5' vers 3' ou 3' vers 5', et possèdent généralement des boucles internes pour empêcher une dégradation par la RNase III (Brantl *et al.*, 2007) (voir [Introduction](#) II.D). Cependant, certains sARN restent sujets à une dégradation spécifique par certaines RNases, comme *RsaA* qui est reconnu par la RNase III et la RNase Y (voir [Résultats](#) II). Malgré cela, la stabilité de cet ARN reste élevée dans une souche sauvage de *S. aureus*, puisqu'elle est estimée entre 30 min et 1 h (Geissmann *et al.*, 2009 ; Lioliou *et al.*, 2012)

**Un motif riche en C idéal pour cibler la séquence Shine et Dalgarno du RBS.** Un second rôle de Hfq chez *E. coli* est d'assister la formation des duplexes sARN / ARNm cibles via des cycles rapides d'association-dissociation, qui augmentent la concentration locale d'ARN pouvant interagir (voir [introduction](#) III.C.1). *S. aureus* est une bactérie à faible pourcentage en GC, et les séquences Shine et Dalgarno sont en majorité fortes chez cet organisme (Ma *et al.*, 2002). Les motifs riches en C, conservés au sein des ARN régulateurs de *S. aureus*, sont donc particulièrement adaptés à la reconnaissance de la séquence Shine et Dalgarno riche en G. De manière intéressante, ces motifs riches en GC sont souvent situés dans des boucles de 7 à 9 nucléotides qui possèdent des structures orientant les nucléotides pour une interaction initiale optimale (Wagner *et al.*, 2002, Brantl *et al.*, 2007) ([Figure 49](#)). Couplée à une transcription massive des ARN régulateurs, assurant une concentration saturante (voir [introduction](#) III.A.1), cette caractéristique des interactions boucle-boucle permet une vitesse d'association rapide, de l'ordre de  $10^6 \text{ M}^{-1}\text{s}^{-1}$ . Celle-ci a été mesurée entre les ARNs naturels de la bactérie et leurs cibles (Wagner *et al.*, 2002) et l'ARNIII et l'ARNm *coa* (Chevalier *et al.*, 2010). Une fois l'interaction boucle-boucle établie, cette dernière peut rester stable ou s'étendre pour former de longs hybrides entre les deux ARN. Les nombreuses topologies formées, impliquant des appariements longs et courts, sont souvent la cible de la RNase III, qui par son action rend la répression médiée par les sARN irréversible.

**En résumé, la présence de motifs riches en C au niveau des tiges-boucles des ARN régulateurs de *S. aureus* assure l'initiation forte et rapide d'une liaison avec les ARNm cibles. Ainsi, l'ARN régulateur est capable de réguler l'expression de sa cible avant le recrutement du ribosome sur l'ARNm. La prise en compte simultanée des interactions**

boucle-boucle potentielles en plus de la séquestration de la séquence Shine et Dalgarno améliorera la prédiction des ARNm cibles des sARN de *S. aureus*.



**Figure 50 :** Les motifs riches en C des ARN régulateurs de *S. aureus* sont retrouvés dans des régions accessibles pour permettre une interaction optimale avec la séquence Shine et Dalgarno des ARNm, d'après Geissmann *et al.*, 2009 (+ supplementary material).

Pour finir, il est intéressant de noter que *RsaI* possède les propriétés inverses des autres ARN régulateurs en arborant deux motifs riches en G dans des régions en simple brin. *RsaI* est-il un ARN régulateur d'ARN régulateur (Figure 49)?

## II.C. Développements récents dans l'identification des cibles

Des études récentes montrent que plusieurs critères peuvent améliorer la prédiction d'ARNm cibles par voie bioinformatique. Ainsi, l'analyse des couples ARN régulateur/ARNm cibles, validés expérimentalement, montre que les séquences impliquées dans les appariements sont sélectionnées au cours de l'évolution pour être dans des régions accessibles (Peer *et al.*, 2011). L'utilisation des régions en simple brin seulement améliorent l'identification des cibles potentielles de *spot42* par le programme **TargetRNA**. De plus, la prise en compte des sites de liaison à Hfq montre que ces derniers ne doivent pas être trop proches des séquences appariées pour permettre la formation du duplex (Beisel *et al.*, 2012).

La prise en compte des sites de liaison à Hfq chez *S. aureus* n'est pas, selon moi, un facteur déterminant dans la prédiction des cibles à cause de son rôle controversé dans la régulation par les sARN chez cette bactérie (voir **Introduction** III.C.1). Néanmoins, l'accessibilité des régions d'interaction pourrait s'avérer intéressante dans l'identification des cibles des ARN régulateurs de *S. aureus*. Plus encore, l'utilisation combinée des programmes **TargetRNA** et **IntaRNA** est avantageuse pour identifier des hybrides à la fois conservés dans des génomes de bactéries phylogénétiquement proches et possédant des régions accessibles favorisant la formation des duplex.

Enfin, le développement de la technique du « *ribosome profiling* » chez les bactéries pourrait être utile pour détecter des ARNm ciblés au niveau de la traduction. Cette technique consiste à séquencer les régions des ARNm protégées par le ribosome. Dans une étude récente de Gene-Wei Li, *et al.*, (2012), la technique « *ribosome profiling* » a montré le rôle des séquences *Shine et Dalgarno like* dans la vitesse de traduction des ARNm. Si cette séquence est nécessaire à l'ancrage du ribosome au niveau du site d'initiation, quel est le rôle des séquences Shine et Dalgarno retrouvées au sein des régions codantes ? Les travaux montrent que la présence de séquences *Shine et Dalgarno like* à l'intérieur des CDS (*Coding sequence*) induit des pauses du ribosome au cours de la traduction. Ces pauses pourraient favoriser le repliement des peptides néo-synthétisés, ou induire une régulation

transcriptionnelle en favorisant la formation de terminateurs. De manière intéressante, les duos de codons mimant une séquence Shine et Dalgarno forte (proche du consensus AGGAGG) à l'intérieur des CDS ne semblent pas être sélectionnés au cours de l'évolution. Ainsi, les paires de codons Gly-Gly GGCGGC avec la plus faible affinité pour l'anti-SD de l'ARN16S sont positivement sélectionnées au sein du génome, par opposition aux codons Gly-Gly GGAGGU qui sont parfaitement complémentaires à la séquence anti-SD (Li *et al.*, 2012).

Ces séquences *Shine et Dalgarno like* sont-elles ciblées par des ARN régulateurs comportant le motif riche en C ? Existe-t-il des sites de liaison d'ARN régulateurs à proximité de ces séquences, qui permettraient à ces derniers de profiter du ralentissement du ribosome pour se lier à un ARNm en cours en traduction?

**Pourquoi cette technique pourrait-elle s'avérer utile dans l'identification des ARNm cibles ?** La majorité des ARN inhibe la traduction d'ARNm cibles en empêchant l'initiation de la traduction. La comparaison des profils de séquençage à haut-débit obtenus entre une souche sauvage et une souche mutante permettrait de mettre en évidence un enrichissement des empreintes du ribosome au niveau de certains ARNm dans une des deux souches. L'avantage de cette technique est qu'elle pourrait donner une idée du site ciblé au niveau de l'ARNm, et révéler par exemple les cas d'activation de la traduction par les sARN. Par ailleurs, cette stratégie pourrait révéler des cas d'initiation atypique de la traduction comme le "stand-by". Enfin, les analyses pourraient être effectuées sur des temps courts d'induction du sARN pour enrichir les profils en ARNm "primaires". Ce dernier point est également envisageable dans les analyses comparatives du protéome.

## **II.D. RsaA interagit-il avec des protéines ?**

La protéine Hfq ne semble pas impliquée dans la régulation médiée par les ARN régulateurs chez *S. aureus* (voir [introduction III.C.1](#)). *RsaA* interagit-il avec d'autres protéines?

### II.D.1. La RNase III, co-facteur de RsaA ?

Au sein de notre laboratoire, Efi Lioliou a mis au point un mutant de la RNase III au niveau des résidus E117 et D45 d'après la nomenclature d'*E. coli*. L'enzyme résultante est grandement affectée dans son activité catalytique mais garde sa capacité à lier les ARN double brin. Après immuno précipitation de l'enzyme, les ARN liés ont été convertis en ADNc identifiés par pyro-séquençage à haut débit. Les résultats obtenus montrent des rôles multiples de la RNase III chez *S. aureus* ; ainsi cette enzyme remplit des fonctions importantes dans la maturation des opérons d'ARNr, la maturation de son propre ARNm ou d'ARNm structurés et la régulation des duplex formés par la transcription persuasive d'ARNas dans tout le génome. Cette stratégie pourrait servir dans l'identification des cibles d'ARN régulateurs. En effet, *RsaA*, dans sa forme longue et courte est retrouvé dans le pool d'ARN liés à la RNase mutante. De plus, la stabilité des deux formes est plus importante dans une souche portant la version mutante de la RNase III. Le site de reconnaissance se situe dans la région 5' de *RsaA*, c'est pourquoi les deux formes sont stabilisées chez le mutant. Ces coupures sont visibles dans la **Figure 37** de la partie **Résultats II.B.2**. De manière générale, de nombreux ARN régulateurs sont retrouvés par cette approche, comme l'*ARNIII*, mais également les régions 5'UTR de certaines de ses cibles, comme les ARNm *rot* et *spa*. En conséquence, cette approche permet d'isoler les sARN et leurs cibles. Ainsi, les données obtenues par cette approche pourraient servir de base dans la prédiction de duplex sARN / ARNm. Il serait donc intéressant de réaliser des études similaires avec d'autres nucléases de *S. aureus* pour mettre en évidence d'autres enzymes impliquées dans la régulation médiée par les sARN. De manière intéressante, l'ARNm *mgrA*, cible de *RsaA*, est retrouvé par cette approche.

D'une manière plus générale, les données de séquençage à haut-débit et de puces à ADN permettront de compléter l'annotation des génomes (promoteurs, ARNas, régulon des facteurs de transcription...) et aideront à mieux appréhender les réseaux de régulation des ARN régulateurs.

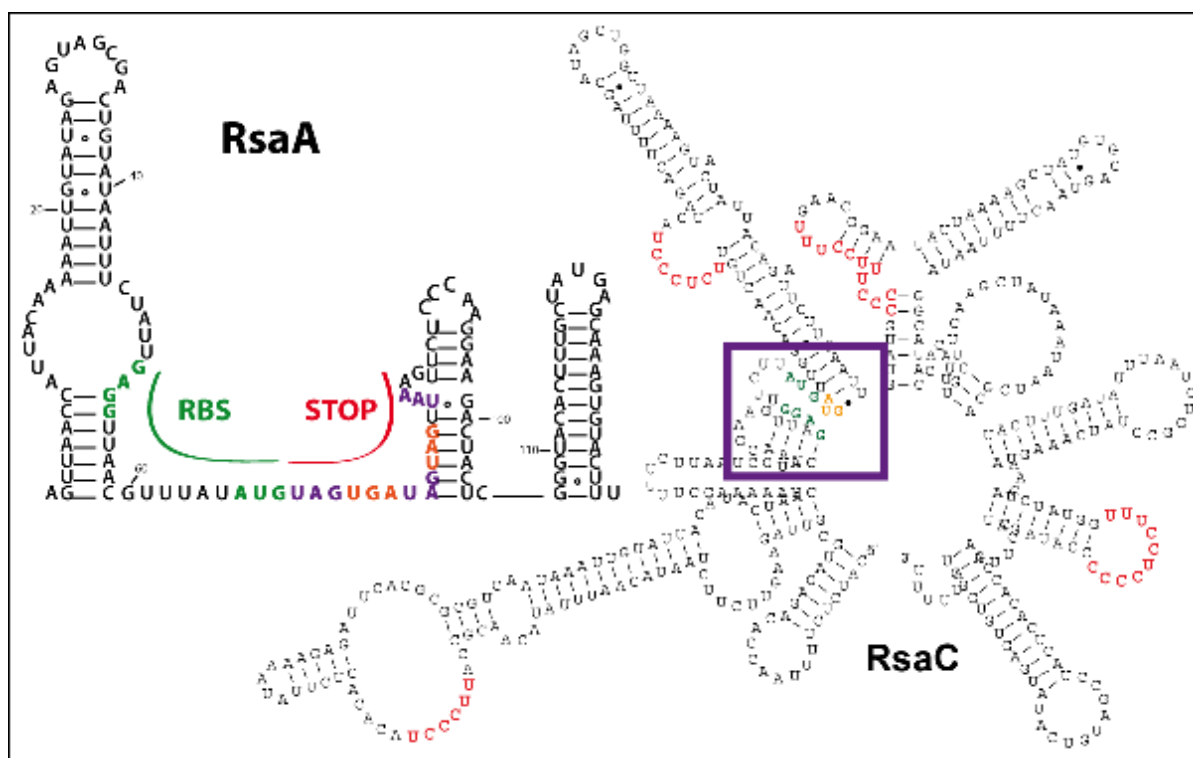


## II.D.2. ARN régulateurs et traduction

Comme décrit en III.A.3, le ribosome est un acteur essentiel dans la régulation médiée par l'ARN régulateur *SgrS*. Le ribosome traduisant l'ARNm *pstG* empêche la liaison de *SgrS*. Ce dernier interagit préférentiellement avec l'ARNm, lorsque le polypeptide naissant entraîne la localisation de l'ARNm à la membrane, là où il est moins accessible au ribosome pour d'autres cycles de traduction. En plus de la séquence complémentaire à l'ARNm *pstG*, *SgrS* possède une ORF nécessaire à la traduction de la protéine *SgrT*, qui constitue le deuxième domaine régulateur de l'ARN. *SgrT*, via un mécanisme indépendant de *SgrS*, régule par une interaction protéine-protéine le transporteur codé par l'ARNm *pstG* pour inhiber l'activité de ce dernier (Wadler *et al.*, 2007). Le site de terminaison du peptide se trouve à 15 nts du site qui se fixe aux ARNm. Cette proximité suggère que les deux activités sont mutuellement exclusives. Or, des mutations qui affectent la traduction diminuent la capacité riborégulatrice de l'ARN, ce qui veut dire que la traduction a un effet positif sur l'activité de riborégulateur (Vanderpool *et al.*, 2011). Quel est le mécanisme de régulation et les changements structuraux apportés par le ribosome ?

Un autre exemple de lien entre un ARN régulateur et le ribosome est l'ARNIII. Sa région en 5', complémentaire à la région 5'UTR de l'ARNm *hla*, active la traduction de l'ARNm en empêchant la formation d'une structure inhibitrice. De manière surprenante, ce domaine se situe dans une région relativement structurée de l'ARNIII et l'interaction ARNIII-ARNm *hla* ne se forme pas dans les conditions où les deux ARN ne sont pas dénaturés au préalable. De plus, les séquences complémentaires à l'ARNm *hla* et l'ORF codant pour *hld* se chevauchent partiellement. **La traduction de Hld est-elle reliée à la fonction activatrice de l'ARNIII, ou ces fonctions sont-elles exclusives ?** Il sera intéressant de réaliser des mutants au niveau de la séquence Shine et Dalgarno sur l'ARNIII, puis de mesurer les conséquences sur la traduction de l'ARNm *hla*. Il est intéressant de noter que de manière similaire à *SgrS*, toutes les fonctions de l'ARNIII sont reliées à une fonction biologique commune. L'ARNIII permet la synthèse des exotoxines que ce soit en codant directement la protéine Hld, en activant l'expression de l'ARNm *hla* ou en inhibant la synthèse de Rot, un répresseur de toxines.

**RsaA peut-il être reconnu par le ribosome ?** Récemment, une recherche des protéines pouvant se fixer à RsaA a été entreprise au laboratoire par Arnaud Tomasini et Isabelle Caldelari. Une séparation des complexes protéiques sur un gradient de sucrose montre que RsaA est présent dans les fractions légères, mais également dans des fractions plus lourdes. L'identification par spectroscopie de masse des protéines du gradient lourd a permis de mettre en évidence des protéines ribosomiques. Aucune ORF n'est détectée pour RsaA. Néanmoins, il possède ce qui semble être un RBS suivi de manière étonnante par 5 codons STOP soulevant l'hypothèse que RsaA puisse piéger des ribosomes (Figure 50). Il sera donc essentiel de vérifier que RsaA se fixe au ribosome. Des caractéristiques semblables ont été identifiées pour RsaC.



**Figure 51 :** Structure secondaire de RsaA et RsaC avec un RBS en vert et une succession de codons STOP en violet et orange.

Si RsaA piège le ribosome, le fait-il en même temps qu'il régule les ARNm cibles ? Ce mécanisme pourrait-il participer à l'activation sélective de la traduction de certains ARNm par l'intermédiaire de RsaA ? Existe-t-il d'autres ARN régulateurs aux caractéristiques similaires ?

### II.D.3. Comment identifier les protéines cibles ?

Existe-t-il une protéine capable de se lier à *RsaA* pour l'assister dans la régulation de l'expression des gènes ? Pour répondre à cette question, deux approches sont en cours de développement au laboratoire par Arnaud Tomasini et Isabelle Caldelari. La première est basée sur l'utilisation d'un gradient de glycérol pour séparer les protéines qui copurifient avec nos ARN d'intérêt. L'avantage de cette technique est qu'elle est globale et permet donc d'étudier plusieurs ARN dans le même gradient de glycérol. La seconde approche consiste à purifier l'ARN *RsaA* portant un aptamère MS2 avec ses partenaires protéiques. Les complexes seront ainsi purifiés sur des billes de sépharose présentant des fusions MS2::MBP (*maltose binding protein*), qui seront spécifiquement reconnues par l'ARN étiqueté et les protéines identifiées par spectrométrie de masse.

Si une ou plusieurs protéines sont identifiées par ces approches, l'importance qu'elles occupent dans la régulation par les sARN pourraient ensuite être estimée en utilisant les techniques de HITS-CLIP (*Hight-throughput sequencing of RNA isolated by Crosslinking and Immunoprecipitation*) et de PAR-CLIP (*Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation*). La base de ces approches consiste à lier de manière covalente l'ensemble des ARNs interagissant avec une protéine d'intérêt *via* un pontage aux ultraviolets, à 254 nm pour la technique HITS-CLIP ou à 365 nm pour le PAR-CLIP. Cette méthode alternative utilise des NTP analogues portant un groupement thiolé qui sont incorporés aux transcrits naissants au sein d'une cellule (Hafner *et al.*, 2010). Après purification de la protéine d'intérêt par chromatographie d'affinité, le séquençage à haut débit des ARN entraînés avec elle fournit une liste des ARN liés par la protéine, ainsi que le site de fixation de cette dernière.

### III. Dynamique des réseaux de régulation au sein de la population de *S. aureus*

Des études récentes, rendues possibles grâce aux développements des outils de fluorescence à l'échelle d'une cellule unique, ont mis en évidence un nouveau principe dans la régulation de l'expression des gènes, la bistabilité. Ce phénomène décrit au sein d'une population donnée, l'existence de plusieurs états différents (revue Siebring *et al.*, 2011). Un exemple est donné par la synthèse du flagelle chez *B. subtilis*. En milieu de phase exponentielle, une sous-partie d'une population exhibe le flagelle. Les réseaux impliqués dans la synthèse de ce dernier sont dits de "*all or nothing*", c'est à dire du « tout ou rien », et impliquent une activation complète de la voie de synthèse du flagelle ou à l'inverse une inhibition totale. Ce système bistable évite ainsi la synthèse incomplète d'un flagelle, ce qui constituerait une perte d'énergie inutile pour la bactérie (Kearns *et al.*, 2005). Chez *E. coli*, la synthèse de *curli* est aussi contrôlée par un système bistable. En effet, les bactéries constituant la population expriment ou non CsgD, un facteur de transcription impliqué dans la synthèse des curli. De manière intéressante, ce système est contrôlé par deux ARN régulateurs, *OmrA* et *OmrB*, qui inhibent la traduction de l'ARNm *csgD* et donc la synthèse de curli (Holmqvist *et al.*, 2010). Il est intéressant de noter que plusieurs sARN régulent de manière opposée la synthèse de CsgD et de la flagelline. Ainsi en fonction de signaux extérieurs ou des conditions de stress, l'une ou l'autre voie sera privilégiée. C'est le cas de l'ARN régulateur *MacS* chez *E. coli*. Ce dernier, transcrit en phase exponentielle ou en réponse à une limitation en glucose, inhibe *csgD* et active l'expression de *flhD* (régulateur de la synthèse du flagelle) ayant pour conséquence l'augmentation de la mobilité de la bactérie. En parallèle, *MacS* est également un activateur de la synthèse de biofilm puisqu'il active l'expression de *PgdA*, un transporteur impliqué dans l'export de N-acetyl-glucosamine nécessaire à la formation de biofilm. Dans ce système, les niveaux d'expression de *MacS* influencent la mise en place de l'une ou l'autre des voies, et sont nécessaires à la mise en place de biofilm en réponse au stress (Thomason *et al.*, 2012).

Les réseaux de régulation des bactéries sont dynamiques, ils intègrent une multitude de signaux aboutissant à la régulation fine de nombreux processus cellulaires. Chez *S. aureus*, le réseau de régulation de l'ARNIII peut être régulé indépendamment du système de densité cellulaire *via* le facteur de transcription MgrA, lui-même cible de RsaA (voir [Discussions et Perspectives I](#)). Le réseau ainsi formé est capable de moduler l'expression des facteurs de virulence en fonction de la densité cellulaire et du stress.

**RsaA pourrait être impliqué dans la synthèse des capsules et la régulation de l'autolyse.** Comme nous l'avons vu en II.A, l'ARNm *yabJ-spoVG* est une cible directe potentielle réprimée par RsaA. Cet ARNm, sous le contrôle de  $\sigma^B$ , est responsable de l'activation des gènes de capsules (Meier *et al.*, 2007). Les analyses du transcriptome montrent également que ces gènes sont activés par MgrA (Luong *et al.*, 2006). Ainsi, RsaA, en régulant l'ARNm *mgrA* et l'ARNm *yabJ-spoVG*, formerait une *feedforward loop* négative aboutissant à une diminution de l'expression des capsules.

Une seconde *feed forward loop* négative impliquerait cette fois l'ARNm *sa1956* et l'ARNm *mgrA* dans la régulation de l'autolyse. Pour rappel, MgrA a été identifiée dans une souche de *S. aureus* possédant un transposon au niveau du gène *mgrA*, caractérisé phénotypiquement par une autolyse plus importante des cellules et une sensibilité accrue au détergent (Ingavale *et al.*, 2003). L'ARNm *sa1956* a été prédit par homologie de séquence comme étant une protéine régulatrice de l'autolyse. Cette dernière fait également partie du régulon de MgrA puisqu'elle semble activée par ce dernier (Luong *et al.*, 2006). Ainsi, en régulant à la fois SA1956 et MgrA, RsaA pourrait moduler l'autolyse de la bactérie et réduire la viabilité de la cellule en réponse au stress.

Cette régulation de l'autolyse peut également être reliée à la régulation de l'ARNm *glyS*. La GlyRS exerce un rôle important dans la synthèse de la paroi cellulaire. En effet, le peptidoglycane de *S. aureus* contient des ponts penta-glycine, spécifiques aux staphylocoques. Les enzymes de la paroi FemA, FemB et FemX utilisent des ARNt<sup>Gly</sup> aminoacylés comme donneurs de glycine. De manière remarquable, *S. aureus* possède 5 ARNt-Gly dont seulement deux sont protéogéniques. Néanmoins la spécificité de recrutement d'un ARNt-gly en particulier par l'une des enzymes impliquées dans la synthèse de peptidoglycane reste à démontrer (Schneider *et al.*, 2004). En réprimant la synthèse de la

GlyRS, *RsaA* pourrait indirectement perturber la paroi cellulaire et la rendre plus sensible aux muréines hydrolases dont l'expression augmente en l'absence de MgrA (Ingavale *et al.*, 2003). Pour rappel, les biofilms synthétisés en présence de *RsaA* sont d'origines mixtes, c'est-à-dire qu'ils sont formés à partir d'ADN, de polysaccharides et de protéines. Cette mixité est conservée chez des mutants de *RsaA*, néanmoins, la part apportée par les protéines et l'ADN diminue (voir **Résultats** II.C.1).

**En régulant le métabolisme de la paroi *via* l'expression de l'ARNm *mgrA*, *glyS* et *sa1956*, *RsaA* pourrait avoir des effets négatifs sur la viabilité de la cellule. Pourquoi *S. aureus* exprimerait-il un tel réseau de régulation ?**

En réponse au stress, l'expression de *RsaA* pourrait modifier et perturber les propriétés de la paroi cellulaire. Si certaines des cellules se lysent plus facilement, l'ADN exogène ainsi généré pourrait servir de matrice à la formation de biofilms. Le potentiel rôle de *RsaA* dans la régulation de l'autolyse et l'activation des biofilms semblent aller dans ce sens.

Pour vérifier ces hypothèses, la viabilité des cellules avant et après stress pourrait être évaluée *via* le kit *LIVE/DEAD BacLight bacterial viability*. Le principe est de marquer simultanément les membranes des bactéries avec une sonde fluorescente et l'ADN génomique *via* l'iodure de propidium. Si la viabilité des cellules est mise en péril lors d'un stress, le marquage à l'iodure de propidium sera davantage visible que celui des membranes.

De manière générale, il serait intéressant d'étudier le réseau de régulation dépendant de *RsaA* et de l'ARNIII dans les cellules uniques en utilisant les outils de microfluidique qui sont en cours de développement au laboratoire (M. Ryckelynck et E. Westhof). En effet, il serait alors possible d'évaluer l'expression des ARN à l'échelle de la cellule unique au sein d'une population de *S. aureus*. L'idée est que les cellules en périphérie de la population, plus exposées à un stress environnemental, seraient plus aptes à activer le régulon de  $\sigma^B$  et par conséquent de *RsaA* en réponse à un stimuli. La formation de biofilm résultant de cette activation aiderait la population à se protéger. De plus, la régulation de l'autolyse pouvant être perturbée par *RsaA*, les bactéries incapables de s'adapter au stress pourraient s'engager sur une voie de mort programmée et libérer nutriments et ADN ; ce

dernier servirait alors de matrice à la formation de couche supplémentaires de biofilms. Au centre de la population, la densité cellulaire élevée activerait le système *agr* et l'expression de l'ARNIII de manière plus importante qu'à la périphérie.

En résumé, en ciblant le facteur général de la transcription MgrA, le réseau de régulation de RsaA pourrait non seulement réguler le système *agr* et l'ARNIII indépendamment de la densité cellulaire, mais également être relié à la synthèse de biofilm, capsules, via la régulation de l'ARNm *yabJ-spoVG*, et à la régulation de l'autolyse via la régulation des cibles potentielles *sa1956* et *glyS*.

RsaE semble induit par un stress anaérobie chez *B. subtilis*, via le système *resDE*. Si cette voie d'induction est conservée chez *S. aureus* via le TCS *srrAB*, le réseau de régulation de RsaE pourrait alors rejoindre celui de l'ARNIII en plus de la régulation du cycle de Krebs. Une telle interconnectivité (Figure 51) permet à la bactérie d'intégrer de multiples signaux provenant du milieu extracellulaire pour adapter efficacement l'expression de ses gènes. Ainsi, les ARN régulateurs, et en particulier ceux qui ciblent des facteurs de transcription globaux seront des cibles de choix dans la recherche de nouvelles molécules thérapeutiques contre le pathogène opportuniste *Staphylococcus aureus*.

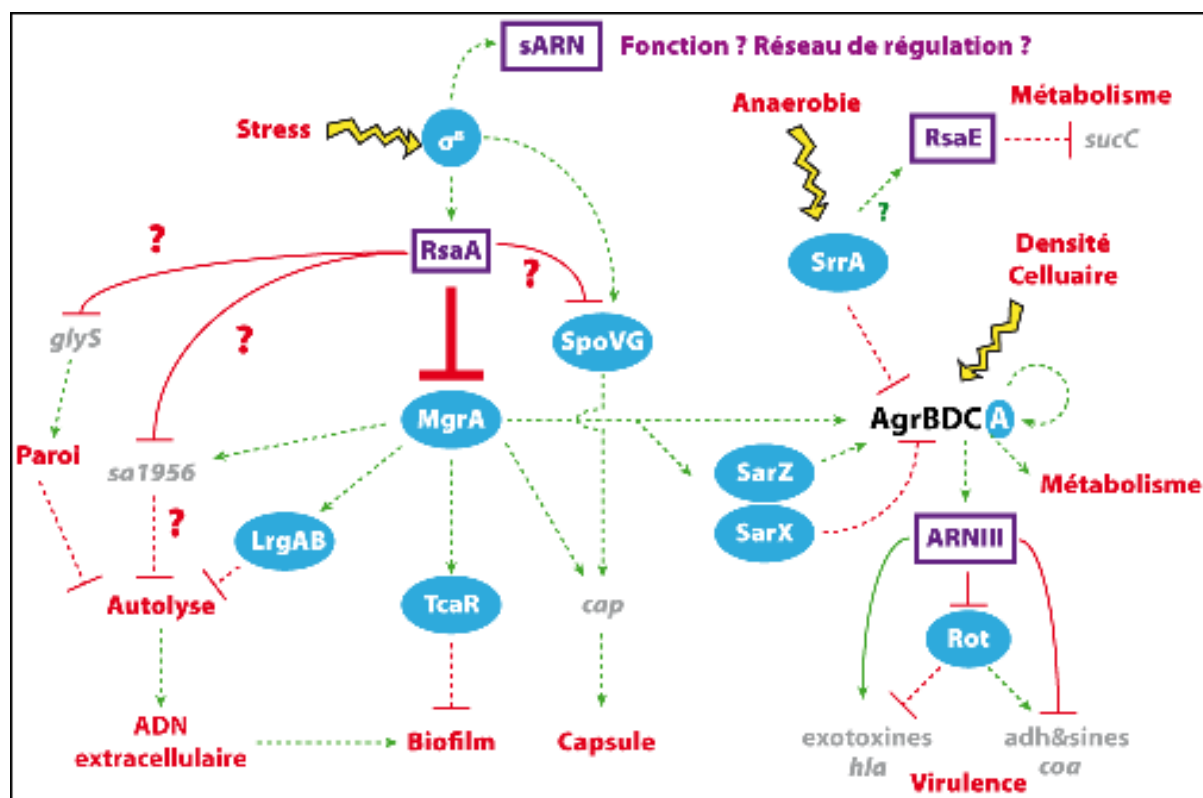


Figure 52 : Interconnectivité entre les réseaux de régulation de RsaA, RsaE et ARNIII.

De manière intéressante, nos collaborateurs de Lyon ont récemment montré que les ARN régulateurs sur lesquels nous travaillons (*RsaA*, *RsaE*, *RsaG* et *RsaH*) sont exprimés dans des prélèvements de patients hospitalisés (Song *et al.*, 2012). Néanmoins, plusieurs questions restent à élucider. *In vivo*, les ARN régulateurs sont-ils exprimés à des niveaux semblables quelle que soit la niche du corps humain occupée ? Les réseaux de régulation exprimés sont-ils les mêmes pour tous les processus d'infection ?



## IV. La recherche de nouvelles molécules thérapeutiques

Comme expliqué en [introduction](#) (voir IV.A), *S. aureus* est un problème majeur en termes de santé publique en raison de l'émergence de souches multi-résistantes aux antibiotiques, aussi bien en milieu hospitalier qu'en dehors. L'apparition de souches de *S. aureus* résistantes à la vancomycine rend urgent la mise au point de nouvelles molécules thérapeutiques. En effet, les antibiotiques à large spectre, longtemps privilégiés par les compagnies pharmaceutiques car plus rentables, sont également ceux pour lesquels les résistances sont le plus vite apparues. Ainsi, plusieurs axes de recherche ont été entrepris dans la mise au point de molécules dirigées spécifiquement contre certaines bactéries. En plus de limiter les risques d'émergence de bactéries résistantes, cette stratégie a également un impact modéré sur les flores microbiennes présentes au sein du corps humain (voir [introduction](#) I.B). Ces avancées ont été rendues possibles grâce aux progrès effectués dans l'étude des relations hôte-pathogène (Hunter 2012).

Deux grandes stratégies sont actuellement développées dans la mise au point de molécules thérapeutiques.

La première consiste à cibler des gènes essentiels à la survie de la bactérie. Les ARN constituent une cible de choix dans cet optique. En effet, la moitié des antibiotiques à notre disposition ciblent l'ARN ribosomique bactérien, comme par exemple les aminoglycosides ou les macrolides. En parallèle, la découverte des *riboswitch*, éléments régulateurs en *cis* (voir [Introduction](#) et revue Lioliou *et al.*, 2011), constituent un nouveau réservoir de cibles. Ces derniers sont localisés dans la région 5'UTR de certains opérons, sont capables de lier spécifiquement un métabolite donné. La fixation de ce dernier aboutit à un changement conformationnel de la plateforme de régulation entraînant une régulation de la transcription, de la traduction ou de la stabilité de l'ARNm (voir [introduction](#) III.B.1.a). De manière remarquable, les gènes codés par l'opéron appartiennent à la voie de synthèse et de transport du métabolite reconnu (voir [introduction](#) III.B.1.a). La recherche de nouvelles molécules thérapeutiques se base sur la synthèse d'analogues d'affinité équivalente ou supérieure au métabolite normalement reconnu par le *riboswitch*. Ainsi, des analogues de la

guanine capables de se lier au *riboswitch* reconnaissant la purine et inhibant la croissance de *B. subtilis* ont été synthétisés (Kim *et al.*, 2009). De manière similaire, un dérivé des pyrimidines a été mis au point pour se lier au *riboswitch* senseur des purines chez *S. aureus*. Cet analogue, nommé PC1, a une activité antimicrobienne contre un spectre restreint de bactérie, dont *S. aureus*, diminuant ainsi les risques d'émergence de souches résistantes (Mulhbachet *et al.*, 2010).

Un autre axe de recherche dans la lutte contre *S. aureus* consiste à perturber les fonctions reliées à l'infection par la bactérie, comme les facteurs de virulence, les toxines, les systèmes de sécrétion ou les protéines d'adhésion. Cette stratégie cible donc des mécanismes propres à chaque pathogène, diminuant ainsi la pression sélective et les effets au sein des microbiomes endogènes (revue Clatworthy *et al.*, 2007). Pour rappel, un inhibiteur de RpnA a un effet global sur la stabilité des ARNm et sur la croissance de *S. aureus* (Olson *et al.*, 2010). Récemment une molécule inhibant le facteur de transcription  $\sigma^B$  a été mise en évidence chez *L. monocytogenes* et *B. subtilis*. Chez *L. monocytogenes*, l'inhibition du facteur de transcription  $\sigma^B$  par cette molécule empêche l'invasion des cellules épithéliales de l'intestin (Palmer *et al.*, 2011). Chez *S. aureus*, le facteur de transcription MgrA est la cible d'une molécule, MDSA, qui empêche la fixation de ce dernier à l'ADN. Cette molécule atténue la virulence de la bactérie dans un modèle murin, mais augmente également la résistance de *S. aureus* contre certains antibiotiques tels la vancomycine (Sun *et al.*, 2011).

**L'étude des ARN régulateurs est le chaînon manquant dans la compréhension des réseaux de régulation et ces derniers reflètent les capacités d'adaptation de la bactérie à son environnement. Comprendre leurs fonctions dans l'établissement de relations hôte pathogènes durables ou au contraire dans l'établissement de l'infection ouvrira de nouvelles stratégies anti-microbiennes et des nouvelles perspectives en microbiologie (revue Lioliou *et al.*, 2010).**



# Annexes



## I. Matériel & Méthodes

### Liste des souches utilisées

Souche <i>S. aureus</i>	
RN4220	dérivée de 8325, restriction -
RN1HG	RsbU réparé
RN1HG $\Delta$ <i>rsaA</i>	RN1HG : $\Delta$ <i>rsaA</i> région:: <i>cat86</i>
RN1HG $\Delta$ <i>rsaA</i> complémentée	RN1HG $\Delta$ <i>rsaA</i> avec plasmide pLUG274::P3 <i>rsaA</i>
Newman	
Newman $\Delta$ <i>rsaA</i>	Newman: $\Delta$ <i>rsaA</i> région:: <i>cat86</i>
Newman $\Delta$ <i>rsaA</i> complémentée	Newman $\Delta$ <i>rsaA</i> avec plasmide pLUG274::P3 <i>rsaA</i>
Becker	
Becker $\Delta$ <i>rsaA</i>	Becker: $\Delta$ <i>rsaA</i> région:: <i>cat86</i>
Becker $\Delta$ <i>rsaA</i> complémentée	RN1HG $\Delta$ <i>rsaA</i> avec plasmide pLUG274::P3 <i>rsaA</i>
RN1HG MgrA:: $\beta$ -gal	RN1HG avec plasmide pLUG200 $P_{rpoB}$ P1 <i>mgrA</i> (5'UTR et +1/+258):: <i>lacZ</i>
RN1HG $\Delta$ <i>rsaA</i> MgrA:: $\beta$ -gal	RN1HG avec plasmide pLUG200 $P_{rpoB}$ P1 <i>mgrA</i> (5'UTR et +1/+258):: <i>lacZ</i>
Plasmide de <i>S. aureus</i>	
pE194	plasmide de <i>S. aureus</i> , erm résistance
pLUG 274	pE194::EcoRV site dans MCS
pLUG 959	pLUG274::P3 <i>rsaA</i>
Plasmides navettes	
pTCV- <i>lac</i>	vecteur navette avec promoteur et gène <i>lacZ</i> , résistant à la kanamycine et à l'erythromycine
pLUG 220	pTCV- <i>lac</i> $\Delta$ RBS et AUG
pLUG mgrA:: $\beta$ -gal	pLUG200 $P_{rpoB}$ P1 <i>mgrA</i> (5'UTR et +1/+258):: <i>lacZ</i>

### Culture cellulaire

Une colonie de *S. aureus* ou *B. subtilis* est mise en croissance à 37°C pendant une nuit dans 2 ou 3 ml de milieu BHI (Brain heart infusion, AES laboratories) ou YT (Difco™). De cette préculture, une aliquote est diluée (1:100) dans le milieu correspondant puis incubé à 37°C.

Les prélèvements sont faits à 2 h (DO<sub>600nm</sub> 0,5 dans BHI), 4 h (DO<sub>600nm</sub> 2 dans BHI) et 6 h (DO<sub>600nm</sub> 2 dans BHI).

### **Northern**

Un extrait d'ARN total est préparé à partir de 20 ml de bactérie cultivées dans du milieu BHI ou YT comme décrit ci-dessus. Après centrifugation (3200 g, 4°C, 10 min), les ARN totaux sont extraits en utilisant le kit **fastRNA Pro Blue kit** (Qbiogen, France) et conservé à -20°C dans de l'eau stérile. Les ARN (20 µg / pistes) sont séparés sur des gels de polyacrylamide dénaturant (8M urée) ou d'agarose (1 ou 2%) contenant 20mM de guanidine thiocyanate. Les ARN sont révélés soit par l'utilisation d'ARNs complémentaires aux ARN testés, transcrits *in vitro* à l'aide du kit DIG (ROCHE), soit par des oligonucléotides antisens marqué au  $\gamma$ -P<sup>32</sup>-ATP à leurs extrémités 5'.

### **Préparation des ARN et expériences de retard sur gel.**

Les ARN non-codant, ARNm et ARNs (sondes northern) sont transcrits *in vitro* en utilisant l'ARN polymérase T7 à partir de fragments de PCR ou de plasmides linéarisés, comme décrit dans Benito *et al.*, 2000. Avant utilisation, les ARN sont dénaturés dans de l'eau stérile pendant 2 min à 90°C puis 1 min à 4°C, suivi d'une étape de renaturation de 20 min à température ambiante dans du tampon TMN (20mM Tris-HCl pH 7,5, 10mM MgCl<sub>2</sub> et 150mM NaCl).

La formation des complexes ARN-ARN par retard sur gel nécessite le marquage d'un des deux partenaires à son extrémité 5' au  $\gamma$ -P<sup>32</sup>-ATP. Après une étape de dénaturation-renaturation, l'ARN marqué est incubé seul et avec des concentrations croissantes d'ARN complémentaire non marqué pendant 15 min à 37°C dans du tampon TMN (volume final de 10 µl). Les échantillons sont ensuite séparés sur un gel natif de polyacrylamide 6% dans du tampon Tris-borate 1x (0,09 M).

### **Analyse du protéome.**

Une préculture sur la nuit de *S. aureus* et *B. subtilis* sert à ensemercer (dilution 1:100) 50 ml de milieu BHI ou YT. Les protéines sont préparées à partir d'extrait réalisés à partir de cellules récoltées en phase exponentielle ou stationnaire de croissance. Après centrifugation

des cellules, les protéines sont obtenues par une extraction au Trizol (Sigma) suivie d'une précipitation à l'acétone. Les protéines sont dissoutes dans le tampon C (7M urée, 2M thiourée, 4% CHAPS, 20mM Tris-HCl pH 8) and 50 µg sont marquées avec 400 pmol de Cy2, Cy3 or Cy5 (Amersham) (réactives via des groupements amines) dans l'obscurité, pendant 30 min à 4°C. La réaction est stoppée par l'ajout de 20 µl de lysine à 10mM pendant 10 min, suivi d'une étape d'alkylation à 37°C pendant 15 min en présence de DTT (50 mM) et 10 µl d'iodoacétamide 1M. Les protéines sont ensuite désalées par une chromatographie d'exclusion G-25. Après ajout de bromophénol et d'ampholytes (Biorad), les échantillons sont utilisés pour ré-hydrater des strips IPG (gradient de pH 4 à 7) de Biorad. L'isofocalisation électrique est effectuée en utilisant un "IEF-cell" (Biorad) à 75kV/h. Les strips IPG sont ensuite incutés dans du tampon d'échange (1 M Tris-HCl pH 8,45, 6 M urée, 2% SDS et 30% de glycérol). Pour la seconde dimension, les strips IPG sont chargés sur le haut du gel polyacrylamide 12% / SDS 2%. Après électrophorèse, les gels sont scannés pour mesurer les intensités de fluorescence puis coloré au bleu colloïdal. Les images obtenues sont analysées par le logiciel PQuest (Biorad). Les protéines montrant une expression différentielle sont identifiées par spectroscopie de masse.

### **Cartographie en solution**

**Les expériences de cartographie en solution des duplexes *RsaA* / ARNm *mgrA* sont décrites dans la publication présentée ci dessous à laquelle j'ai participé.**

Helfer A.C, Romilly C., Lioliou E., Chevalier C. & Romby P. (2011) **Probing the RNA structure and the effect of ligand binding or of the environmental cues using enzymes and chemicals. In Handbook of RNA Biochemistry** (Hartmann RK B.A., Schon A., Westhof E., ed.) Wiley-VCH Verlag, Weinheim.



**Probing RNA structure *in vitro* with enzymes and chemicals**

Anne-Catherine Helfer, Cédric Romilly, Clément Chevalier, Efthimia Lioliou, Stefano Marzi  
& Pascale Romby\*

Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, 15 rue René  
Descartes, Strasbourg, France.

\*Corresponding author:

UPR 9002-ARN CNRS, 15 Rue R. Descartes 67084 Strasbourg; Tél: 33 [0]3 88 41 70 51; e-  
mail: p.romby@ibmc-cnrs.unistra.fr

Running title: RNA structure probing

## 1. Introduction

In all organisms, RNAs play a diverse range of essential functions, including acting as major regulators of gene expression [1-3]. These functions result from specific RNA tertiary structures, and RNA structural transitions or conformational rearrangements often occur during regulation, leading to different genetic responses [4, 5]. Indeed, the regulatory regions of bacterial mRNAs can sense physical and metabolic changes. These *cis*-acting regions respond to temperature [6, 7], pH [8], divalent ions [9], and the intracellular concentration of a wide variety of trans-acting factors including metabolites [5], uncharged tRNAs [10], proteins [11], and non coding RNAs [3, 12, 13]. Hence, the biological functions of RNA are linked to the dynamics and versatility of its RNA structure, which promotes the formation of stable or transient intra- or intermolecular interactions. Consequently, there is increasing interest in studies on the structure of both free RNAs and RNAs interacting with ligands.

Although the structure of the whole ribosome has been now solved [14], analysis of the architecture of long RNA molecules is still a challenging experimental problem. Among the wide array of methods applied to this problem, chemical and enzymatic probing has the obvious advantage of probing RNA molecules of any size under a wide range of experimental conditions (i.e., by varying temperature, pH, and the concentration of monovalent and divalent ions). These methods map the reactivity and/or the accessibility of each nucleotide toward enzymes and chemicals and unambiguously define unpaired regions of the RNA. These data are often used as constraints in computer folding programs, facilitating the prediction of RNA secondary structure inferred from the sequence [15, 16]. The resulting secondary structure model can be further validated by site-directed mutagenesis studies coupled with a probing approach to analyze the effects of each mutation on the RNA structure. Chemical probing also provides information regarding the tertiary folding of RNA. Large RNAs are usually composed of stable secondary structural elements that are brought together by long-distance interactions to form compact domains [17]. Crystallographic studies performed on the ribosome and various regulatory elements of mRNAs have revealed the presence of numerous RNA motifs that either mediate specific interactions to induce compact folding of rRNAs or constitute specific ligand binding sites [18, 19]. These motifs are composed of stacked arrays of non Watson-Crick base pairs and are characterized by a specific pattern of chemical reactivity of the bases. Hence, the correlation between the X-ray structure and chemical modifications of the bases can be used to determine the existence of structural motifs in RNAs. Other chemicals, such as hydroxyl radicals generated by Fe-

EDTA, which cleave ribose [20], and N-ethyl-N-nitrosourea, which alkylates phosphates [21], are useful to probe the solvent accessible surface and divalent ion binding pocket of large, structured RNAs. Other methods that probe RNA in a sequence-independent manner have been shown to be powerful in predicting RNA secondary structure and aid in visualizing conformational changes [22-25]. These methods are based on the relative flexibility of specific RNA linkages (in-line probing, [22]) or of the 2'OH group of the ribose units toward acylation (SHAPE; [26, 27]).

To define the rules that govern the RNA folding process, time-resolved probing techniques have been developed using either base-specific reagents [28] or hydroxyl radicals [29]. With these tools, it is now possible to monitor conformational rearrangements and transient interactions that occur during the RNA folding process or during the assembly of RNA-ligand complexes. As a new development, using time-resolved chemical probing in single-RNA molecule experiments would be of interest to analyze the intermediates generated during the formation of the native and functional RNA state. Indeed, recent single-molecule fluorescence resonance energy transfer experiments have revealed that RNA can fold into multiple distinct states that interconvert [30]. This ability of RNA to adopt different conformations may have an impact on cell biology, perhaps by providing different responses to environmental changes [31].

Structure-specific chemical probes are also unique tools to map the structure of RNA *in vivo* under different cell growth conditions. The use of probes, however, is limited by their inability to penetrate the cell wall of bacteria or membranes. To date the reagents that have been used for *in vivo* RNA probing are dimethylsulfate (*e.g.*, [7]), kethoxal [32], and lead(II)-induced cleavages [33]. The novel SHAPE approach based on selective 2'-hydroxyl acylation of the ribose units might also be used *in vivo* [34]. With the development of high throughput sequencing analysis combined with lead(II)-induced cleavages *in vivo*, it is now feasible to obtain structural information on a pool of RNA molecules expressed under specific conditions. Such an approach might also allowing mapping of the binding sites of regulators that target multiple mRNAs. Other technical improvements using capillary electrophoresis with fluorescent DNA probes for primer extension allow quantitative analysis of the chemical and enzymatic probing data [35, 36]. This technique based on DNA sequencing technology allows structure probing of very long RNA molecules in one step, such as the entire HIV-1 RNA genome, in a single step [37].

Other methods such as chemical interference approaches have been derived from footprinting experiments. Chemical interference defines a set of nucleotides that have lost the

capability to interact with a ligand when they are modified by a chemical probe (e.g. [38, 39]). In this respect, nucleotide analog interference mapping (NAIM) is a powerful method to investigate the effects of substituting specific nucleotide functional groups on the RNA structure or of ligand binding [40, 41]. Chemical probes tethered to proteins or RNAs can provide topographical information on ligand-RNA complexes by inducing site-specific cleavage of a proximal RNA after binding (e.g., [42, 43]).

In this review, an experimental guide is given for several probes commonly used to map the structure of RNAs and RNA-ligand interactions.

## 2. Enzymatic and chemical probes

These probes are used under limited conditions in which less than one cleavage or modification occurs per RNA molecule with a statistical distribution. The identification of the cleavages or modifications can be achieved by various methodologies depending on the length of the RNA molecule and the nature of the modification. The first approach uses end-labeled RNA, but it only detects scissions and is limited to RNAs of less than 300 nucleotides. The second approach uses primer extension to detect stops of reverse transcriptase at modified nucleotides or cleavages and can be applied to RNAs of any size. Primer extension with reverse transcriptase can be done with a 5'-<sup>32</sup>P-labeled DNA oligonucleotide followed by separation of the labeled cDNA fragments with conventional denaturing polyacrylamide gel electrophoresis, allowing the identification of modifications/cleavages for 150-200 nt per experiment. A more recent approach uses primer extension with a 5'-fluorescently tagged DNA primer; the cDNA products are then resolved by capillary electrophoresis. Although the classical approach using a radiolabeled primer is sensitive, better quality of the quantification of the nucleotide reactivities/cleavages has been obtained using capillary electrophoresis [26, 35, 36]. In addition, modifications or cleavages can be identified for 650-1000 nt per experiment.

The proposed experimental guide was adapted from the conditions used to probe the structure of the regulatory RNAlII from *Staphylococcus aureus* [44, 45] and is limited to several commonly used probes (Table 1). In addition, experimental procedures are given to gain insight into RNA structure in bacterial cells. Other protocols have been used for *in vivo* RNA probing in eukaryotic cells (this issue, [46]).

### 2.1. Enzymes

RNase T1 (specific for unpaired guanines), RNase T2 (with preference for unpaired adenines), and RNase V1 (specific for double-stranded regions) are easy to use and are able to identify secondary structure RNA elements such as hairpins. RNase V1 is the only probe that provides positive evidence for the existence of helical regions. Due to their size, however, the RNases are sensitive to steric hindrance. The cleavages caused by RNases may potentially induce conformational rearrangements in the RNA that provide new targets (secondary cuts) to the RNase. These secondary effects are especially observed if the reaction is too strong due to the addition of a high concentration of the RNase.

### 2.2. Base-specific chemical probes

Base-specific reagents modify specific positions on bases. The combination of dimethylsulfate (DMS) and 1-cyclohexyl-3-(2-morpholinoethyl)carbodiimide metho-*p*-toluene sulfonate (CMCT) allow probing of the four bases at one of their Watson-Crick positions (Table 1). DMS methylates position N1 of adenines and, to a lesser extent, N3 of cytosines. CMCT modifies position N3 of uridines and, to a lesser extent, N1 of guanines ([47] and ref therein).

Position N7 of purines can be probed by diethyl pyrocarbonate (DEPC) for adenines, and by DMS or Nickel complex for guanines. Nickel complex [48] and DEPC [49] are very sensitive to the stacking of the base rings; and therefore, position N7 of purines within a helix is never reactive unless if the deep groove of the helix is widened.

### 2.3. Backbone-specific chemical probes

A new generation of chemicals has been designed that modify 2'-OH groups of ribose units. These reagents include N-methylisatoic anhydride (NMIA) [26], 1-methyl-7-nitroisatoic anhydride (1M7) [37, 50], and benzoyl cyanide (BzCN) [36, 51]. All of them react in a very similar way, but their half-lives differ; thus, their reaction times extend from 20 min (NMIA) and 1 min (1M7) to the second range (BzCN). All of these reagents are highly sensitive to the local flexibility of the nucleotide, which is required for the formation of a 2'-O-adduct between the 2'-OH group of ribose units and the hydroxyl-selective electrophile [51]. Hence, these reagents unambiguously identify the unpaired nucleotides. Although purine residues are slightly more reactive than pyrimidines, it has been shown that these differences are subtle in

contrast to the reactivity differences between paired and unpaired nucleotides [50]. These ribose modifications can only be detected as stops in reverse transcription.

In-line probing exploits the fact that RNA phosphodiester linkages are spontaneously cleaved as the result of a nucleophilic attack of the 2' OH group resulting in the formation of a transient 2'3' cyclic phosphate. It has been proposed that the maximal rates of this spontaneous cleavage are achieved when the 2' oxygen, the phosphorus center and the 5' oxygen leaving group form a perfect 180° angle. Thus, the ribose-phosphate linkages that participate in stable Watson-Crick base pairs exhibit lower rates of spontaneous cleavage than unpaired nucleotides [24, 52]. The in-line probing consists of a long incubation (> 40 h) of the RNA in a basic buffer containing monovalent and divalent ions. This simple approach has been used to monitor the conformational states of a tRNA in solution and in crystals [25] and to visualize the conformational changes and structures of riboswitches, the *cis*-acting regulatory regions of mRNAs that respond to the intracellular concentration of metabolite and magnesium (*e.g.*, [5, 9, 53]). The combination of in-line probing and sequence alignment led to the establishment of secondary structure models for several riboswitches that have been well correlated with the high resolution crystal structures and NMR structures [17, 54].

### 3. Experimental Protocols

Probing the conformation of RNAs with different enzymes and chemicals requires the use of defined buffer conditions (*e.g.*, pH, ionic strength, magnesium concentration, temperature). Indeed, the optimal conditions vary slightly with different probes, and the possibility exists that subtle conformational changes may occur under different incubation conditions (Table 1). The probe:RNA ratio must also be adapted so that the experiments are conducted under limited and statistical conditions. For the first experiment, different concentrations of the probes and time scales should be used, which is also necessary when the commercial source of the probes has been changed.

#### 3.1. Equipment and reagents

##### Equipment

Equipment for denaturing polyacrylamide gel electrophoresis (Model S2, Gibco BRL) is used to size end-labeled RNA products and labeled cDNA fragments. If the primer extension

reactions are sized on capillary gels, instruments from Applied Biosystems (ABI) or Beckman can be used. Eppendorf tubes, tips, and buffers should all be sterilized before use.

#### **Chemicals and enzymes**

CMCT can be purchased from Fluka (ref. 29469); DMS from Acros Organics (ref. 11682-0100); DEPC, from Sigma-Aldrich (ref. D578); benzoyl cyanide, from Sigma-Aldrich (ref. 115959), and should be stored in a dessicator at room temperature; calf intestinal phosphatase from Roche (ref. 713023; 1 U/ $\mu$ l) and T4 RNA ligase from Biolabs (ref. M0204L; 20 U/ $\mu$ l); avian myeloblastosis virus (AMV) reverse transcriptase from MP Biochemicals (ref. EMAMV200; 20 U/ $\mu$ l); T4 polynucleotide kinase from Fermentas (ref. EK0031; 10 U/ $\mu$ l), [ $\gamma$ - $^{32}$ P]ATP (3200 Ci/mmol) and [5'- $^{32}$ P]pCp (3000 Ci/mmol), from Amersham. RNase T1 can be purchased from Fermentas (ref. EN0541; 1000 U/ $\mu$ l); RNase T2, from Invitrogen (ref. 18031-013; 20 U/ $\mu$ l); and RNase V1, from Ambion (ref. 2275; 0,1 U/ $\mu$ l).

#### **Safety rules using chemicals**

Most of these chemical reagents are potential carcinogens; therefore, chemical modifications until the removal of the first ethanol supernatant (see below) are carried out under a fume hood while wearing protective gloves. DMS and solutions are discarded in 1 M sodium hydroxide waste, and CMCT, in 10% acetic acid waste.

#### **Buffers**

The buffer conditions given here are indicative and can be modulated according to the system used and the nature of the ligand: *Buffer N1*: 20 mM Tris-HCl pH 7.5, 10 mM MgCl<sub>2</sub>, 150 mM KCl. *Buffer N2*: 50 mM sodium cacodylate pH 7.5, 10 mM MgCl<sub>2</sub>, 100 mM KCl. *Buffer N3*: 50 mM sodium borate pH 8, 10 mM MgCl<sub>2</sub>, 100 mM KCl. *Buffer N4*: 40 mM MOPS pH 8, 80 mM acetate potassium, 10 mM MgCl<sub>2</sub>. *Buffer in-line*: 50 mM Tris-HCl pH 8.3, 100 mM KCl, 20 mM MgCl<sub>2</sub>. *Buffer  $\Delta$ T1*: 20 mM sodium citrate pH 4.5, 1 mM EDTA, 7 M urea, 0.025% xylene cyanol, 0.025% bromophenol blue. *Ladder buffer*: Na<sub>2</sub>CO<sub>3</sub> 0.1 M/NaHCO<sub>3</sub> 0.1 M pH 9. *RNA loading buffer*: 0.025% xylene cyanol, 0.025% bromophenol blue in 8 M urea. *DNA loading buffer*: 0.025% xylene cyanol, 0.025% bromophenol blue in formamide. *RTB buffer*: 50 mM Tris-HCl pH 7.5, 20 mM MgCl<sub>2</sub> and 50 mM KCl. *TBE buffer*: 0,09M Tris-borate pH 8.3, 1 mM EDTA. All buffers are given as 1x concentrations, but they are usually prepared as 5x solutions.

### 3.2. RNA preparation

The RNA is usually transcribed *in vitro* from a DNA template using T7 RNA polymerase. It is necessary to produce a sufficient amount of RNA for all of the experiments (50-100 µg). The RNAs produced are then purified from the DNA template and excess of nucleoside triphosphate (NTP) by using gel filtration column [55], monoQ [56], or ion-pairing reversed-phase high-performance liquid chromatography (IP-RPLC) [57]. For large RNAs, DNase RNase free treatment should be performed to remove the DNA template. Denaturing polyacrylamide-urea gel electrophoresis is the best strategy to obtain homogenous RNA [58] but the efficiency of the RNA elution from the gel decreases with the length of the RNA.

For 5' end labeling, the RNA is first dephosphorylated at its 5' end and then labeled using [ $\gamma$ - $^{32}$ P]ATP and T4 polynucleotide kinase [22, 59]. The dephosphorylation step can be avoided if transcription is carried out in the presence of GMP, or with UpG (Sigma). The 3' end labeling is performed using [5'- $^{32}$ P]pCp and T4 RNA ligase [60] with 5 µg of RNA. The labeled RNAs are purified by electrophoresis on 8% polyacrylamide (0.5% bis)- 8 M urea slab gels. Before each experiment, the RNA is eluted from the gel slice in 100 µl of 500 mM ammonium acetate/1 mM EDTA in the presence of 1/6 (vol/vol) phenol. The RNA is then phenol extracted and precipitated with 2.5 volumes of cold ethanol in the presence of 1 µg of glycogen. After two washing steps with 200 µl of cold 80% ethanol, the pellet is then dissolved in sterile H<sub>2</sub>O (to obtain about 50000 cpm/µl). Given that the RNA is often in contact with denaturing reagents during its purification, it is worth making an effort to carry out a renaturation process before the probing experiments.

### 3.3. Enzymatic cleavages and in-line probing of end-labeled RNA

This direct method is restricted to the detection of cleavages (see Table 1). Enzymatic probing has been adapted for structure probing of the *Staphylococcus aureus* RNAlII (514 nts) and for the determination of its footprint on its target mRNAs and *vice versa*. Experiments are illustrated in Figure 1. All reactions are conducted in a total volume of 10 µl. Appropriate dilutions of enzymes are made in sterile water just before use. For RNA-ligand footprinting experiments, the complex is pre-formed for 15 min at 20°C by adding increasing concentrations of the ligand before the enzymatic or chemical reactions in the appropriate buffer for optimal binding. Incubation controls in the absence of the probes are performed on the RNA either free or bound to its ligand (e.g., RNA, protein) to detect spontaneous



cleavages of the RNA. In these controls, the RNase and the RNA ligand are replaced by sterile H<sub>2</sub>O.

Labeled RNA (50,000 cpm/ $\mu$ l) sufficient for the planned experiments was first denatured in sterile H<sub>2</sub>O at 90°C for 1 min then cooled on ice for 1 min.

#### **RNase hydrolysis**

- RNase T1: Labeled mRNA (1  $\mu$ l, 50,000 cpm) is renatured in the presence of 5  $\mu$ l of H<sub>2</sub>O and 2  $\mu$ l of buffer N1 (5x concentrated) at 20°C or 37°C for 20 min. 1  $\mu$ l of total tRNA (2  $\mu$ g/ $\mu$ l) is then added, and the reaction is performed with 1  $\mu$ l of RNase T1 (0.2 U) for 5 min at 20°C or 2 min at 37°C.
- RNase T2: The same protocol is used for RNase T1 except that the reaction is performed with 1  $\mu$ l of RNase T2 (0.05 U) for 5 min at 20°C or 2 min at 37°C.
- RNase V1: The same protocol is used for RNase T1 except that reaction is performed with 1  $\mu$ l of RNase V1 (0.05 U) for 5 min at 20°C or 2 min at 37°C.
- For the first experiment, at least three different concentrations of the enzymes should be tried: RNase T1 (0.1-0.2-0.5 U), RNase T2 (0.01-0.05-0.1 U), and RNase V1 (0.01-0.05-0.1 U).

#### **Stopping the reactions**

Enzymatic assays are stopped by phenol extraction.

- 40  $\mu$ l of 0.3 M sodium acetate pH 6 and 50  $\mu$ l of phenol saturated with chloroform (vol/vol) are added to all samples. The samples are mixed for 1 min, and then centrifuged 1 min at high speed.
- The aqueous phase is carefully removed, and transferred into a new sterile 1.5 ml micro tube, and 2.5 volumes of cold ethanol (150  $\mu$ l) is added to precipitate the RNA. After mixing, the samples are left in a dry ice-ethanol bath for 10 min, and centrifuged (13000 rpm at 4°C for 15 min).
- The supernatant is discarded and the pellet is washed twice with 200  $\mu$ l of cold 80% ethanol. After a short centrifugation (13,000 rpm for 5 min at 4°C), the supernatant is removed, and the pellets are vacuum-dried (no more than 5 min), and dissolved in 6  $\mu$ l of RNA loading buffer.

#### **In-line probing**

- Labeled mRNA (75,000-100,000 cpm) is incubated with 2  $\mu$ l of 5x buffer in-line in a total volume of 10  $\mu$ l at room temperature for 40 h.
- In-line probing can be conducted under various conditions. The pH can range from 7.5 to 8.5, and the time of incubation should decrease if the temperature is increased. In-line probing can also be used to study magnesium-dependent effects on RNA folding. For these experiments, a high concentration of monovalent cations (1-3 M) should be used to eliminate or vary the  $Mg^{2+}$  concentration [61].
- The reaction is stopped by adding 10  $\mu$ l of RNA loading buffer and frozen at  $-20^{\circ}C$ .
- A control corresponding to the RNA, which is not incubated in the buffer, is also loaded in parallel to the in-line probing reactions on the polyacrylamide-urea gel electrophoresis.

#### **Fractionation of end-labeled RNA fragments**

RNase T1 and alkaline ladders are required to identify the cleavage positions.

- RNase T1 ladder: Labeled mRNA (1  $\mu$ l, 50,000 cpm) is pre-incubated at  $50^{\circ}C$  for 5 min in 5  $\mu$ l of buffer  $\Delta$ T1 containing 1  $\mu$ g of total tRNA. The reaction is then carried out at  $50^{\circ}C$  for 10 min in the presence of 1  $\mu$ l of RNase T1 (0.5 U).
- Alkaline ladder: Labeled mRNA (2  $\mu$ l, 100,000 cpm) is incubated at  $90^{\circ}C$  for 3 min in the presence of total tRNA (2  $\mu$ g) in 5  $\mu$ l of ladder buffer. The time of the hydrolysis must be adapted according to the length of the RNA.

The end-labeled RNA fragments are sized by electrophoresis on 12% or 15% polyacrylamide (0.5% bis)- 8 M urea slab gels (0.5 mm x 30 cm x 40 cm) in 1X TBE. Gels should be pre-run (30 min at 75 W) and run warm (75 W) to avoid band compression. The migration conditions must be adapted to the length of the RNA, given that on a 15% gel, xylene cyanol migrates as a 39-nucleotides long RNA and bromophenol blue as 9 nucleotides. A 15% polyacrylamide gel is convenient to collect data on small fragments (1- to 50-nucleotide RNA fragments). For a 250-nucleotide RNA, two migrations are necessary to interpret correctly the reactivity of nucleotides of the entire RNA molecule. At the end of the run, the 12% gel is fixed for 5 min in a 10% ethanol/6% acetic acid solution, transferred to Whatman 3 MM paper and dried. The 15% gel is transferred without drying on a plastic support and wrapped with a plastic film. Exposure is carried out at  $-80^{\circ}C$  using an intensifying screen.

### 3.4. Chemical modifications

Examples of chemical modifications performed on a *Staphylococcus aureus* mRNA target of the regulatory RNAIII are shown in Figure 2. For enzymatic reactions, the same experimental conditions can be used as described above except that the reactions are carried out with 1 pmol of unlabeled RNA. Reactions are carried out in a total volume of 20  $\mu$ l except for the modification of riboses by 1M7 which is carried out in 50  $\mu$ l. An unmodified control is performed in parallel to detect pauses of reverse transcriptase due to stable secondary structures and/or spontaneous cleavage. Unlabeled mRNA is first heated in sterile H<sub>2</sub>O at 90°C for 1 min, then cooled on ice for 1 min and treated as described below.

#### DMS modification (N3C, N1A)

- 1 pmol of mRNA (per assay) is renatured by incubation at 37°C for 15 min in buffer N2 in a total volume of 15  $\mu$ l.
- When probing mRNA-RNAIII, increasing concentrations of RNAIII (1 to 250 nM) are added, and complex formation is carried out at 37°C for 15 min.
- 1  $\mu$ l of tRNA (2  $\mu$ g/ $\mu$ l) and 1  $\mu$ l of pure DMS or DMS freshly diluted in ethanol at 1:2, 1:5 or 1:10 is added, and the tubes are gently mixed. The reaction is performed in a total volume of 20  $\mu$ l at 37°C for 5 min. The optimal chemical modification is usually obtained with DMS diluted at 1:10 (Figure 2).

#### DMS modification (N7G)

- The same conditions are used as for modifications of N3C and N1A except that the reaction is carried out for 10 min at 37°C
- After ethanol precipitation (see first step of "stopping the reactions"), the pellet is dissolved in 10  $\mu$ l of 1 M Tris-HCl pH 8.3 and 10  $\mu$ l of 8 mg/ml sodium borohydrate (dissolved in H<sub>2</sub>O extemporaneously). The reaction is done in the dark and on ice for 10 min, followed by ethanol precipitation of the modified RNA.
- The RNA pellet is dissolved in 10  $\mu$ l of aniline (100  $\mu$ l of bidistilled aniline (Fluka), 60  $\mu$ l acetic acid, 930  $\mu$ l H<sub>2</sub>O) and incubated at 60°C for 10 min in the dark. For further steps, follow the "stopping the reactions" as described below.

**CMCT modification (N3U, N1G)**

- 1 pmol of mRNA (per assay) is first incubated at 37°C for 10 min in buffer N3 in a total volume of 15  $\mu$ l.
- When probing mRNA-RNAIII, increasing concentrations of RNAIII (1 to 250 nM) are added, and complex formation is carried out at 37°C for 15 min.
- 1  $\mu$ l of tRNA (2  $\mu$ g/ $\mu$ l) and 4  $\mu$ l of CMCT (40 or 60 mg/ml dissolved in water just before use) are added, and the samples are gently mixed. The modification is carried out in a total volume of 20  $\mu$ l at 37°C for 10 min at 20°C for 20 min. The optimal modification of mRNA is seen with CMCT at 40 mg/ml (Figure 2).

**DEPC carbethoxylation (N7A)**

- 1 pmol of mRNA (per assay) is first renatured by incubation at 37°C for 15 min in buffer N2 in a total volume of 15  $\mu$ l.
- When probing mRNA-RNAIII pairings, increasing concentrations of RNAIII (1 to 250 nM) are added, and complex formation is carried out at 37°C for 15 min.
- 1  $\mu$ l of tRNA (2  $\mu$ g/ $\mu$ l) and 4  $\mu$ l of pure DEPC are added, and the tubes are mixed gently. The reaction is carried out at 37°C for 15 min (Figure 2).

**Ribose acetylation with 1M7**

- 2 pmol of RNA (per assay) are renatured in buffer N4 at 20°C for 15 min in a total volume of 18  $\mu$ l.
- The reagent 1-methyl-7-nitroisatoic (1M7) is not commercially available but has been chemically synthesized by Mortimer & Weeks [23]. It undergoes rapid hydrolysis in buffer such that the reaction is complete after 70 s. Therefore, this reagent should be stored in a desiccator at 4°C. 1M7 must be diluted with anhydrous DMSO to obtain an 80 mM solution so that the reagent remains stable and can be stored at 20°C in a desiccator.
- The sample is splitted into two tubes (9  $\mu$ l each). 1  $\mu$ l of 80 mM 1M7 is added to one tube, and 1  $\mu$ l of DMSO is added to the other tube (incubation control). The tubes are mixed gently, and the reaction is performed at 37°C for 70 sec.
- The optimal concentration of 1M7 can vary with RNA length. Thus, it is recommended to determine the concentration dependence ranging from 2 to 20 mM, with the typical optimum concentration being 8 mM. No quenching step is required.

#### **Ribose acetylation with benzoyl cyanide**

- 2 pmol of RNA (per assay) are renatured in buffer N4 at 20°C for 15 min in a total volume of 20 µl.
- Benzoyl cyanide solution should be freshly prepared for each experiment. BzCN is dissolved in DMSO to obtain a 10x (400 mM) stock solution.
- 2 µl of 400 mM BzCN, and the tubes is mixed gently. The reaction is performed at 37°C for 1 sec or at 20°C for 3 sec. 5 µl of DMSO is added to the control sample instead of BzCN. As for 1M7, it is also recommended to determine the concentration dependence ranging from 10 to 80 mM final concentration of BzCN, with the typical optimum concentration being 40 mM. No quenching reaction is required [36].

#### **Stopping the reactions**

All of the reactions are stopped by ethanol precipitation of the RNA.

- 50 µl of 0.3 M sodium acetate pH 6, and 250 µl of cold ethanol are added to each sample. For RNA-protein footprinting experiments, the protein is removed by phenol extraction. The samples are then mixed, placed in a dry-ice/ethanol bath for 15 min, and centrifuged (13000 rpm (10000 g)) at 4°C for 15 min.
- The supernatants are removed carefully (without touching the pellet), and 200 µl of cold 80% ethanol is added to each pellet. The samples are centrifuged (13,000 rpm at 4°C for 5 min), and the supernatants are removed with the same caution. This step is repeated once.
- The pellets are vacuum-dried (not more than 5 min) and dissolved in 4 µl of sterile H<sub>2</sub>O.

### **3.5. Primer extension analysis**

Primer extension with reverse transcriptase was originally developed to probe the structure of ribosomal RNAs [62]. Reverse transcriptase stops its incorporation of dNTPs at the residue preceding a cleavage or a modification at a Watson-Crick position. While carbethoxylation of A(N7) by DEPC is sufficient to stop reverse transcriptase, subsequent treatment is necessary to induce cleavage at G(N7) after DMS modification (Table 1).

The length of the primer varies, but 12 to 18 nucleotides typically provide sufficient specificity, even if the primers are used on a mixture of RNAs. For long RNAs, primers are selected every 200 nucleotides due to gel resolution limits. To obtain efficient hybridization, pilot assays should be performed previously to define the ideal concentration of the RNA, the

choice of the primer sequence, and the hybridization conditions to have an efficient primer extension.

In most of the cases, <sup>32</sup>P-labeled DNA probes are used for primer extension followed by separation of the labeled cDNA fragments on a denaturing polyacrylamide gel. The experimental conditions are given below. As an alternative approach, 5'-fluorescently labeled DNA primers have also been used for primer extension, and the generated cDNA fragments have been fractionated using capillary electrophoresis and a classical DNA sequencing apparatus. For these experiments, at least three DNA primers labeled with a different fluorophore are required. Hence, the control, experimental and sequencing reactions can be loaded on the same capillary gel. Detailed protocols have been published previously [35, 36].

#### **Hybridization**

- 1 µl of 5' end <sup>32</sup>P-labeled DNA primer (around 100000 cpm) is added to the 4 µl of modified mRNA (1 to 5 pmol). The samples are then heated at 90°C for 1 min and quickly cooled on ice. Annealing of the primer can also be done by incubating the reaction at 65°C for 5 min and then quickly placing it on ice for 1 min.
- 1 µl of 5x RTB buffer is added and the samples are incubated for 20 min at 37°C.

#### **Primer extension**

- The reaction is carried out in a total volume of 15 µl. 2 µl of RTB 5x, 2 µl of dNTP mix (2.5 mM of each dNTP), 4 µl sterile H<sub>2</sub>O, 1 µl of AMV RT (2 U/µl freshly diluted in sterile H<sub>2</sub>O) are added to the hybridization mix. Primer extension is done at 37°C for 30 min.
- To improve the resolution of the gels, the RNA template is destroyed by alkaline treatment. Just after primer extension, 20 µl of the buffer containing 50 mM Tris-HCl pH 7.5, 7.5 mM EDTA, 0.5% SDS, and 3.5 µl of 3 M KOH is added. The samples are incubated at 90°C for 3 min and then at 37°C for at least 1 h.
- 6 µl of 3 M acetic acid, 100 µl 0.3 M sodium acetate pH 6 and 300 µl of cold ethanol are added to all samples. The samples are then mixed, placed in a dry-ice/ethanol bath for 15 min, and centrifuged (13,000 rpm (10,000 g)) at 4°C for 15 min.
- The supernatants are carefully removed, and the pellets are washed twice with 200 µl of cold 80% ethanol. After centrifugation (13,000 rpm at 4°C for 5 min), the supernatants are removed, and the end-labeled DNA fragments are dissolved in 6 µl of DNA loading buffer.

#### 4. Commentary

##### 4.1. Critical parameters

###### RNA preparation

The RNA is usually synthesized by *in vitro* transcription using T7 RNA polymerase, therefore, the 3' or 5' end of the RNA might be heterogeneous. The incorporation of ribozyme at the 5' and 3' end of the transcripts is one solution to generate homogenous ends [68]. Furthermore, several abortive transcription products might accumulate. To obtain homogenous RNA molecules, the method of choice remains fractionation of the transcript by electrophoresis on a denaturing polyacrylamide-urea gel. For long RNA molecules, electro-elution might help to increase the elution efficiency.

###### Homogeneous RNA conformation

During purification, the RNA can be partially denatured, it is thus essential to design renaturation protocols to have a "conformationally" homogeneous RNA population and to test whether this conformation is biologically relevant (enzymatic activity for ribozyme, efficient ligand binding). Alternative RNA conformations may co-exist and can be revealed by the simultaneous presence of specific single-stranded and double-stranded cleavages or modifications at the same position. If the conformers have different electrophoretic mobilities on a native polyacrylamide gel, chemical probing can be used to distinguish them. After chemical modification, the co-existing structures are separated on a native polyacrylamide gel, and the modification sites for each conformer are identified by primer extension [69].

###### Chemical and enzymatic probing *in vitro*

- The chemical reactions and RNase T1 hydrolysis can be conducted under a variety of experimental conditions. For instance, the influence of divalent ion (such as magnesium) can be tested on the folding of the RNA, and by varying the temperature (between 4°C and 90°C) one can follow thermal transitions of the RNA [70]. It is essential to adapt modifications and the enzymatic hydrolysis for each condition to have less than one cut or modification per molecule (i.e., more than 80% of the RNA should not be modified or cleaved). For example, for DMS modification, the reaction is carried out at 4°C for 20 min in the presence of 1  $\mu$ l of DMS whereas at 50°C, the reaction is carried out for 5 min with 1  $\mu$ l of DMS diluted 1:16.

- RNase cleavages can induce conformational rearrangements that are potentially able to provide new targets for secondary cleavages. These secondary cleavages occur when the reaction is too strong, but they are also of weak intensity and are not reproducibly found in all experiments. These cleavages can be distinguished from primary cuts by comparing the hydrolysis patterns obtained from 5' or 3' end-labeled RNA.
- RNase V1 hydrolysis generates RNA fragments with a 3' OH group in contrast to alkaline treatment. Therefore, 5' end-labeled fragments generated by alkaline treatment will migrate faster than RNase V1-treated fragments. Meanwhile, 3' end-labeled fragments generated by RNase V1 will migrate faster than those induced by alkaline treatment. This difference is only observed for the shortest RNA fragments.
- Appropriate incubation controls are required to identify cleavages that are induced during the incubation treatment, or during pauses of reverse transcriptase due to stable secondary structures or cleavages. Nucleotides for which strong bands are visible in the control lanes are not considered in the interpretation. For primer extension of structured RNAs, SuperScript III reverse transcriptase (Invitrogen, ref. 18080-093) can be used for primer extension as the enzyme is still active at 55°C [36].
- Each experiment should be repeated several times, and only the reproducible cleavages and modifications should be considered in the interpretation.
- Footprinting assays. These experiments should be done in the presence of increasing concentrations of the ligand (Figure 1). Lead(II)-induced cleavages, hydroxyl radicals, and IM7 or BzCN are appropriate probes to map ligand binding sites due to their small size and the fact that they are less sensitive to the RNA structure. In-line probing is perhaps less appropriate for protein footprinting as the time of incubation is quite long. Chemicals such as DEPC react strongly with proteins; therefore, it is important to avoid such compounds for footprinting assays. The results should be interpreted with care as decreased reactivity does not necessarily result from a direct shielding effect but could be due to steric hindrance effect (particularly observed with the bulky RNases) or a conformational change in the RNA.
- For the acylation of riboses, only NMIA and BzCN are commercially available (Sigma). Due to its short reaction time, BzCN is particularly appropriate to study the RNA folding process. For NMIA, detailed protocols have already been described [22, 26]. 1 µl of NMIA (30, 65 and 140 mM dissolved in DMSO) is added to 9 µl of buffer N4 containing 1 pmol of renatured RNA. An incubation control is performed in the presence of DMSO. The samples are then mixed and incubated at 37°C for 45 min.



#### 4.2. Troubleshooting

- Compression of bands due to stable secondary structure (in general rich in G-C base pairs): The end-labeled RNA samples can be heated at 90°C for 3 min before gel loading. The alkaline ladder and the RNase T1 ladder, however, should never be heated.
- Too many bands in the incubation controls of end-labeled RNA: This problem is likely due to RNase contamination. The RNA should be purified again on a denaturing polyacrylamide-urea gel, and new sterile buffers should be prepared.
- Radiolysis of end-labeled RNAs can also be a source of noise, particularly for in-line probing. The end-labeled RNA must be kept in water with an activity below 200,000 cpm/ $\mu$ l. It is sometimes better to make a rapid elution of the end-labeled RNA from the gel (several hours at 4°C) and to store the gel at -20°C for further elution.
- Aggregation of cDNA fragments in the gel pockets: Only fragments of small sizes can be visualized, and the data cannot be taken into account. Degradation of the RNA template after primer extension prevents these artifacts of gel electrophoresis migration.
- Samples do not migrate correctly during electrophoresis: This complication is probably due to the presence of salt. Several washing steps with 80% ethanol should be added at the end of the procedure.
- No full-length RNA after modification or enzymatic hydrolysis: The conditions should be adapted by reducing the amount of probe to obtain more than 80% unmodified RNA molecules.
- Very weak signals for modifications or cleavages although an intense full-length product is observed: This problem reflects insufficient modification or hydrolysis. It is always useful to determine the concentration dependence of the probes to use optimal conditions for probing. The preparation of the probe (enzymes and chemicals) must be done just prior to the experiments. For the enzymes, they must be diluted in a buffer, and for the chemicals (DMS, DEPC, IM7 and BzCN), particular storage conditions must be taken into consideration.
- Absence of a signal after primer extension: It is likely that the modified RNA did not precipitate efficiently. As the RNA is not labeled, particular caution should be taken to prevent loss of the pellet after precipitation. The concentration of the RNA can also be increased (2 to 5 pmol per assay). Another explanation could be the quality of the reverse transcriptase, which is sensitive to both MgCl<sub>2</sub> concentration (3 mM is the optimal concentration) and freezing. Optimal conditions for primer hybridization should be

established in a series of pilot experiments. The optimal temperature for annealing varies from RNA to RNA, depending on the G+C content, the propensity of the RNA to form secondary structure, and the length of the primer.

- No proper extension or weak bands in the sequencing lanes: The concentration of the ddNTP and the ratio of ddNTP/dNTP should be adjusted.
- To maintain high resolution of the gels, acrylamide, urea solutions, and particularly ammonium persulfate should be freshly prepared.

### Acknowledgements

We thank P. Fechter and I Caldelari for helpful discussions. We are grateful to E. Westhof for his constant support and helpful advice. This work was supported by the Centre National de Recherche (CNRS), the Agence Nationale de la Recherche (ANR-07-BLAN-0351-01, and ANR-09-BLAN-0024-01). EL receives the support from FEBS, CR from the Conseil Régional Alsace and the Délégation Générale de l'Armement, and CC, from the Région Alsace.

### References

1. CARON M.P., D.A. LAFONTAINE, E. MASSE. *RNA Biol.* 2010, 7, in press.
2. MATTICK J.S., R.J. TAFT, G.J. FAULKNER. *Trends Genet.* 2010, 26, 21–28.
3. WATERS L.S., G. STORZ. *Cell*, 2009, 136, 615–628.
4. Dambach M.D., W.C. Winkler. *Curr Opin Microbiol.* 2009, 12, 161–169.
5. Roth A., R.R. Breaker. *Annu Rev Biochem.* 2009, 78, 305–334.
6. NARBERHAUS F. *RNA Biol.* 2010, 7, in press.
7. GIULIODORI A.M., F. DI PIETRO, S. MARZI, B. MASQUIDA, R. WAGNER, P. ROMBY, C.O. GUALERZI, C. PON. *Mol Cell*, 2010, 37, 21–33.
8. NECHOOSHTAN G., M. ELGRABLY-WEISS, A. SHEAFFER, E. WESTHOF, S. ALTUVIA. *Genes Dev.* 2009, 23, 2650–2662.
9. RAMESH A., W.C. WINKLER. *RNA Biol.* 2010, 7, in press.
10. GREEN N.J., F. J. GRUNDY, T.M. HENKIN. *FEBS Lett.* 2010, 584, 318–324.
11. BABITZKE P., C. S. BAKER, T. ROMEO. *Annu Rev Microbiol.* 2009, 63, 27–44.
12. VOGEL J. *Mol Microbiol.* 2009, 71, 1–11.
13. FIGUEROA-BOSSI N., M. VALENTINI, L. MALLERET, L. BOSSI. *Genes Dev.* 2009, 23, 2004–2015.
14. FRANK J., R. L. GONZALEZ JR. *Annu Rev Biochem.* 2010, in press.
15. REUTER J.S., D.H. MATHEWS. *BMC Bioinformatics.* 2010, 11, 129.
16. VASA S. M., N. GUEX, K.A. WILKINSON, K.M. WEEKS, M.C. GIDDINGS. *RNA*, 2009, 14, 1979–1990.
17. SERGANOV A. *RNA Biol.* 2010, 7, in press.
18. LESCOUTE A., N.B. LEONTIS, C. MASSIRE, E. WESTHOF. *Nucleic Acids Res.* 2005, 33, 2395–2409.

19. LEONTIS N.B., A. LESCOUTE, E. WESTHOF. *Curr Opin Struct Biol*, 2006, 16, 279–287.
20. LATHAM J.A., T.R. CECH. *Science*, 1996, 245, 276–282.
21. ROMBY P., D. MORAS, M. BERGDOLL, P. DUMAS, V.V. VLASSOV, E. WESTHOF, J.P. EBEL, R. GIEGE. *J Mol Biol*, 1985, 184, 455–471.
22. WAKEMAN C.A., W.C. WINKLER. *Methods Mol Biol*, 2009, 540, 173–191.
23. MORTIMER S.A., K.W. WEEKS. *J Am Chem Soc*, 2007, 129, 4144–4145.
24. SOUKUP G.A., R.R. BREAKER. *RNA*, 1999, 5, 1308–1325.
25. MORAS D., A.C. DOCK, P. DUMAS, E. WESTHOF, P. ROMBY, J.P. EBEL, R. GIEGE. *Proc Natl Acad Sci U S A*, 1986, 83, 932–936.
26. WILKINSON K.A., E.J. MERINO, K.M. WEEKS. *Nat Protoc*, 2006, 1, 1610–1616.
27. MERINO E.J., K.A. WILKINSON, J.L. COUGHLAN, K.M. WEEKS. *J Am Chem Soc*, 2005, 127, 4223–4231.
28. FABBRETTI A., P. MILON, A.M. GIULIODORI, C.O. GUALERZI, C.L. PON. *Methods Enzymol*, 2007, 430, 45–58.
29. RALSTON C.Y., B. SCLAVI, M. SULLIVAN, M.L. DERAS, S.A. WOODSON, M.R. CHANCE, M. BRENOWITZ. *Methods Enzymol*, 2000, 317, 353–368.
30. SOLOMATIN S.V., M. GREENFELD, S. CHU, D. HERSCHLAG. *Nature*, 2010, 463, 681–684.
31. CRUZ J.A., E. WESTHOF. *Cell*, 2009, 136, 604–609.
32. BALZER M., R. WAGNER. *Anal Biochem*, 1998, 256, 240–242.
33. LINDELL M., P. ROMBY, E.G. WAGNER. *RNA*, 2002, 8, 534–541.
34. GHERGHE C., C.W. LEONARD, R.J. GORELICK, K.M. WEEKS. *J Virol*, 2010, 84, 898–906.
35. MITRA S., I.V. SHCHERBAKOVA, R.B. ALTMAN, M. BRENOWITZ, A. LAEDERACH. *Nucleic Acids Res*, 2008, 36, e63.
36. MORTIMER S.A., K.M. WEEKS. *Nat Protoc*, 2009, 4, 1413–1421.
37. WATTS J.M., K.K. DANG, R.J. GORELICK, C.W. LEONARD, J.W.J. BESS, R. SWANSTROM, C.L. BURCH, K.M. WEEKS. *Nature*, 2009, 460, 711–716.
38. SKRIPKIN E., J.C. PAILLART, R. MARQUET, B. EHRESMANN, C. EHRESMANN. *Proc Natl Acad Sci U S A*, 1994, 91, 4945–4949.
39. SCHLEGL J., V. GEGOUT, B. SCHLAGER, M.W. HENTZE, E. WESTHOF, C. EHRESMANN, B. EHRESMANN, P. ROMBY. *RNA*, 1997, 3, 1159–1172.
40. RYDER S.P., L. ORTOLEVA-DONNELLY, A.B. KOSEK, S.A. STROBEL. *Methods Enzymol*, 2000, 317, 92–109.
41. WALDSICH C. *Nat Protoc*, 2008, 3, 811–823.
42. JOSEPH S., H.F. NOLLER. *Methods Enzymol*, 2000, 318, 175–190.
43. CULVER G.M., H.F. NOLLER. *Methods Enzymol*, 2000, 318, 461–475.
44. BENITO Y., F.A. KOLB, P. ROMBY, G. LINA, J. ETIENNE, F. VANDENESCH. *RNA*, 2000, 6, 668–679.
45. BOISSET S., T. GEISSMANN, E. HUNTZINGER, P. FECHTER, N. BENDRIDJ, M. POSSEDKO, C. CHEVALIER, A.C. HELFER, Y. BENITO, A. JACQUIER, C. GASPIN, F. VANDENESCH, P. ROMBY. *Genes Dev*, 2007, 21, 1353–1366.
46. WELLS S.F., J.M. HUGHES, A.H. IGEL, M.J. ARES. *Methods Enzymol*, 2007, 318, 479–493.
47. EHRESMANN C., F. BAUDIN, M. MOUGEL, P. ROMBY, J.P. EBEL, B. EHRESMANN. *Nucleic Acids Res*, 1987, 15, 9109–9128.
48. CHEN X., S.A. WOODSON, C.J. BURROWS, S.E. ROKITA. *Biochemistry*, 1993, 32, 7610–7616.
49. WEEKS K.M., D.M. CROTHERS. *Science*, 1993, 261, 1574–1577.

50. WILKINSON K.A., S.M. VASA, K.E. DEIGAN, S.A. MORTIMER, M.C. GIDDINGS, K.M. WEEKS. *RNA*, 2009, 15, 1314–1321.
51. MORTIMER S.A., K.M. WEEKS. *J Am Chem Soc*, 2008, 130, 16178–16180.
52. DOCK-BREGEON A.C., D. MORAS. *Cold Spring Harb Symp Quant Biol*, 1987, 52, 113–121.
53. WINKLER W.C., R.R. BREAKER. *Annu Rev Microbiol*, 2005, 59, 487–517.
54. BUCK J., B. FURTIG, J. NOESKE, J. WOHNERT, H. SCHWALBE. *Proc Natl Acad Sci U S A*, 2007, 104, 15699–15704.
55. ROMANIUK P.J., I.L. DE STEVENSON, H.H. WONG. *Nucleic Acids Res*, 1987, 15, 2737–2755.
56. JAHN M.J., D. JAHN, A.M. KUMAR, D. SOLL. *Nucleic Acids Res*, 1991, 19, 2786.
57. GELHAUS S.L., W.R. LACOURSE, N.A. HAGAN, G.K. AMARASINGHE, D. FABRIS. *Nucleic Acids Res*, 2003, 31, e135.
58. MILLIGAN J.F., O.C. UHLENBECK. *Methods Enzymol*, 1989, 180, 51–62.
59. SILBERKLANG M., A.M. GILLUM, U.L. RAJBHANDARY. *Methods Enzymol*, 1979, 59, 58–109.
60. ENGLAND T.E., O.C. UHLENBECK. *Nature*, 1978, 275, 560–561.
61. DANN C.E., C.A. WAKEMAN, C.L. SIELING, S.C. BAKER, I. IRNOV, W.C. WINKLER. *Cell*, 2007, 130, 878–892.
62. QU H.L., B. MICHOT, J.P. BACHELLERIE. *Nucleic Acids Res*, 1983, 11, 5903–5920.
63. LAEDERACH A., R. DAS, Q. VICENS, S.M. PEARLMAN, M. BRENOWITZ, D. HERSCHLAG, R.B. ALTMAN. *Nat Protoc*, 2008, 3, 1395–1401.
64. MORENO R., S. MARZI, P. ROMBY, F. ROJO. *Nucleic Acids Res*, 2009, 37, 7678–7690.
65. WILKINSON K.A., R.J. GORELICK, S.M. VASA, N. GUEX, A. REIN, D.H. MATHEWS, M.C. GIDDINGS, K.M. WEEKS. *PLoS Biol*, 2008, 6, e96.
66. DEIGAN K.E., T.W. LI, D.H. MATHEWS, K.M. WEEKS. *Proc Natl Acad Sci U S A*, 2009, 106, 97–102.
67. MATHEWS D.H., M.D. DISNEY, J.L. CHILDS, S.J. SCHROEDER, M. ZUKER, D.H. TURNER. *Proc Natl Acad Sci U S A*, 2004, 101, 7287–7292.
68. WALKER S.C., J.M. AVIS, G.L. CONN. *Nucleic Acids Res*, 2003, 31, e82.
69. SCHRODER A.R., T. BAUMSTARK, D. RIESNER. *Nucleic Acids Res*, 1998, 26, 3449–3450.
70. JAEGER L., E. WESTHOF, F. MICHEL. *J Mol Biol*, 1993, 234, 331–346.
71. CHEVALIER C., S. BOISSET, C. ROMILLY, B. MASQUIDA, P. FECHTER, T. GEISSMANN, F. VANDENESCH, P. ROMBY. *PLoS Pathog*, 2010, 6, e1000809.

### Figures and table

**Figure 1: Enzymatic probing of *Staphylococcus aureus rot* mRNA (A) and the 3' end domain of RNAIII (B).** (A) RNase T1 hydrolysis performed on 5' end-labeled *rot* mRNA either free (-) or in the presence of increasing concentrations of RNAIII (1, 10, 50, 100, 250 nM). The concentration of RNase T1 was 0.2 U per assay. Lanes C1, C2: incubation controls of *rot* mRNA alone (C1) and in the presence of RNAIII (C2). Lanes T1, L: RNase T1 under denaturing conditions (T1) and alkaline ladder (L). Changes induced by *rot* mRNA binding are shown by bars on the autoradiograph. (B) RNase T1 cuts are shown on the secondary structure model of the regulatory region of *rot* mRNA: medium cut, black arrow; weak cut, dashed arrow. Protections induced by RNAIII binding are shown as black circles. SD is for Shine-Dalgarno sequence. (C) RNase T2 and RNase V1 hydrolysis performed on the 3' end domain of RNAIII labeled at its 5' end free (-) or bound to *rot* mRNA at a concentration of 1, 10, 50, 100, or 250 nM. RNase T2 (0.04 U); RNase V1 (0.01 U). Same legend is used as above. Bars on the autoradiograph represent changes induced by RNAIII binding. (D) Secondary structure of the 3' end domain of RNAIII: RNase T2, black and dashed arrows for medium and weak cleavages, respectively. RNase V1, black and empty arrowheads for strong and medium cleavages, respectively. The results were derived from Boisset et al. [45]. Nucleotides shown in red in the *rot* mRNA and RNAIII are complementary and form loop-loop interactions between the hairpin loop 14 of RNAIII and hairpin loop II of *rot* and between the hairpin loop 13 of RNAIII and the apical loop I of *rot*.

**Figure 2: Chemical probing of *Staphylococcus aureus coa* mRNA free or bound to RNAIII.** (A) Primer extension performed on *coa* mRNA modified by DMS and CMCT. The mRNA was either free (T) or bound to wild type RNAIII (WT) or various RNAIII derivatives carrying deletions of hairpin 7 ( $\Delta 7$ ), hairpin 13 ( $\Delta 13$ ) or hairpin 14 ( $\Delta 14$ ). Lanes C1, C2: incubation controls of the *coa* mRNA either free (C1) or bound to RNAIII (C2). Lanes G, A, U, C: DNA sequencing ladders. Experimental details are given in the text. (B) A summary of the data are shown on the secondary structure of the *coa* mRNA: circled nucleotides are reactive toward DMS (N1A>>N3C) and CMCT (N3U>>N1G) modifications, while black and dashed lines show moderate and low reactivity, respectively. The effects of RNAIII binding: protections (black arrows) and enhanced reactivity (star) are shown. A schematic view of the secondary structure of RNAIII is given with hairpins 7, and 13 highlighted. The

position of the primer use for reverse transcription is complementary to nucleotides A99 through A116. The enzymatic cleavages are specified. For RNase T2, strong cuts are shown with black arrows, and medium cuts are shown with dashed arrows. For RNase V1, medium and weak cuts are shown with black and empty arrowheads, respectively. The concomitant presence of RNase V1 cuts in hairpin I and the reactivity of several bases at one Watson-Crick position suggests that alternative structure might occur. Nucleotides in red in the *coa* mRNA are complementary to two domains of RNAIII: hairpin 13 of RNAIII binds to the 5' untranslated region of the mRNA, including the Shine-Dalgarno (SD) sequence, while hairpin 7 forms a loop-loop interaction with hairpin loop III of the *coa* mRNA. These data were derived from Chevalier et al. [71]. It is worth noting that on the autoradiograph showing the DMS modifications, there are fewer labeled cDNA products in the lane corresponding to the free mRNA (T) compared to the other lane. The modification levels are expected to be identical to those found in lane  $\Delta 13$ .

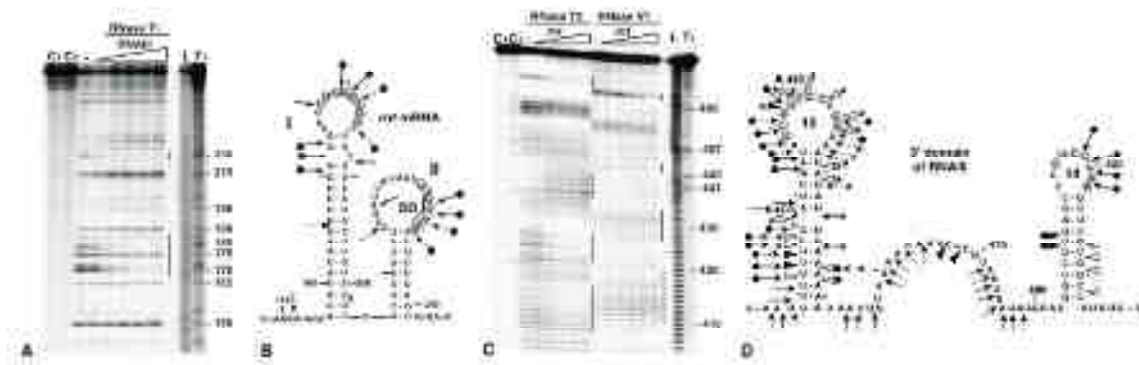


Figure 1- Hefter et al.

IM7	222	As for NMIA	-	+	II, III F	Some comment as for NMIA
BzCN	117	As for NMIA	-	+	II, III F	Some comment as for NMIA
<b>Divalent ions and hydrolytic cleavages</b>						
Fe-EDTA (radical hydroxyl)	Cleavage at ribose (C1', C4')		+	+	II, III F	? Reactivity relatively insensitive to buffers, pH, T°C, Mg and cation. Only sodium phosphate should be avoided as well as glyceral (lower than 0.5%)
Pb(II)-acetate	Unpaired nt and unstable helices Phosphate linkage cleavage		+	+	II, III F	+
In-line	Unpaired nt Phosphate linkage cleavage		+	+	II, III F*	? Work under a wide range of conditions salt and Mg, T°C, optimal pH 7.5-8

- **Table 1: Structure-specific probes for RNA.** DMS, dimethylsulfoxide; DEPC, diethylpyrocarbonate; CMCT, 1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide metho-p-toluene sulfonate; NMIA, N-methylisatoic anhydride; IM7, 1-methyl-7-nitroisatoic anhydride; BzCN, benzoyl cyanide. Detection method: (direct) detection of cleavages on end-labeled RNA molecule; (indirect) detection by primer extension with reverse transcriptase using either <sup>32</sup>P-labeled primer or primer labeled at the 5' end with a fluorophore. (+) the corresponding detection method can be used; (s) a chemical treatment is necessary to cleave the ribose-phosphate chain prior to the detection; *In vivo* mapping: probes which diffuse efficiently across membranes and walls (+), the other probes can be used only after permeabilization of the cell (-). Molecular weight, specificity, and products generated by the probe action are indicated. Most of the probes provided information useful to build a secondary structure model (II) and elements of the tertiary folding (III), and to map the binding sites of proteins, RNA and metabolites (F). F\*: due to the long time of modification, the reagent might be not appropriate for mapping the sites of interactions. However, in-line probing was a useful tool to visualize the conformational changes of the RNA upon metabolite binding (see text).

## II. Publications scientifiques

**Romilly C.**, Lays C., Lioliou E., Boisset S., Vandenesch F. Geissmann T. & Romby P. *Staphylococcus aureus* RsaA represses translation of the global regulator MgrA, in preparation.

**Romilly C.**, Chevalier C., Lioliou E., Masquida B., Marzi S., Westhof E. and Romby P. *Staphylococcus aureus* endoribonuclease III recognizes a specific topology promoted by a loop-loop interaction, submitted.

**Romilly C.**, Caldelari I., Parmentier D., Lioliou E., Romby P. and Fechter P. (2012) **Current knowledge on regulatory RNAs and their machineries in *Staphylococcus aureus***. RNA biology 9:4; 1-12

**Abstract** : *Staphylococcus aureus* is one of the major human pathogens, which causes numerous community-associated and hospital-acquired infections. The regulation of the expression of numerous virulence factors is coordinated by complex interplays between two component systems, transcriptional regulatory proteins, and regulatory RNAs. Recent studies have identified numerous novel RNAs comprising cis-acting regulatory RNAs, antisense RNAs, small non coding RNAs and small mRNAs encoding peptides. We present here several examples of RNAs regulating *S. aureus* pathogenicity and describe various aspects of antisense regulation.

Lioliou E., **Romilly C.**, Geissmann T., Vandenesch F., Romby P. (2011) **RNA-mediated regulation of virulence gene expression : another layer of complexity**. in " Bacterial regulation network , Horizon Scientific Press, Norwich UK, (ed. A. Filloux, UK).

**Abstract** : Many pathogenic bacteria cause serious diseases in humans, animals, and plants. Due to the appearance of resistance to multiple antibiotics, it has become important to fully understand the regulatory networks that lead to the production of virulence factors that help the bacteria combat the host defense machinery, acquire nutrients, and survive and/or proliferate within the host. In recent years, complex interplays between transcriptional regulatory proteins, two-component systems, and regulatory RNAs have been described, establishing the gene expression patterns in pathogenic bacteria. In this review, several examples will illustrate the diversity of regulatory RNAs and how they are integrated into the regulatory circuits required for virulence gene expression, with special emphasis on the mechanisms of regulation at the molecular level.



Caldelari I., Fechter P., Lioliou E., **Romilly C.**, Chevalier C., Gaspin C. & Romby P. (2011) **A current overview of regulatory RNAs in *Staphylococcus aureus***. In “Regulatory RNAs in prokaryotes” (eds, Edited by A. Marchfelder and W. Hess), Wiley Verlag, Wien and New York.

This review provides a summary of the newly identified ncRNAs and gives an overview of the functions they regulate. A particular focus is made on RNAs that are involved in the regulation of metabolism, stress response, DNA maintenance, defense mechanisms against foreign DNA, and virulence. This review also gives a view of regulatory elements acting in *cis* that are predicted in *S. aureus*, such like riboswitch and T-boxes

Lioliou E., **Romilly C.**, Romby P. & Fechter P. (2010). **RNA-mediated regulation in bacteria: from natural to artificial systems**. N Biotechnol. 31, 222-235.

**Abstract** : Bacteria use various means of RNA-mediated gene regulation. Regulatory RNAs include mRNA leaders that affect expression in *cis* or in *trans*, non-coding RNAs that trap regulatory proteins or interact with one or multiple target mRNAs, and RNAs that protect the bacteria against foreign and invasive DNA. The aim of this review is to outline the basic principles of bacterial RNA-mediated regulation, with a special focus on both *cis*-acting regulatory regions of mRNAs and antisense RNAs (asRNAs), and to give a brief overview of selected examples of RNA-based technology that have paved the way for biotechnological applications.

Chevalier C., Boisset S. **Romilly C.**, Masquida B., Fechter P., Geissmann T., Vandenesch F. & Romby P. (2010). ***Staphylococcus aureus* RNAIII binds to two distant regions of *coa* mRNA to arrest translation and promote mRNA degradation**. PLoS Pathog 6, e1000809.

**Abstract** : *Staphylococcus aureus* RNAIII is the intracellular effector of the quorum sensing system that temporally controls a large number of virulence factors including exoproteins and cell-wall-associated proteins. Staphylocoagulase is one major virulence factor, which promotes clotting of human plasma. Like the major cell surface protein A, the expression of staphylocoagulase is strongly repressed by the quorum sensing system at the post-exponential growth phase. Here we used a combination of approaches in vivo and in vitro to analyze the mechanism used by RNAIII to regulate the expression of staphylocoagulase. Our data show that RNAIII represses the synthesis of the protein through a direct binding with the mRNA. Structure mapping shows that two distant regions of RNAIII interact with *coa* mRNA and that the mRNA harbors a conserved signature as found in other RNAIII-target mRNAs. The resulting complex is composed of an imperfect duplex masking the Shine-Dalgarno sequence of *coa* mRNA and of a loop-loop interaction occurring downstream in the coding region. The imperfect duplex is sufficient to prevent the formation of the ribosomal initiation complex and to repress the expression of a reporter gene in vivo. In addition, the double-strand-specific endoribonuclease III cleaves the two regions of the mRNA bound to RNAIII that may contribute to the degradation of the repressed mRNA. This study validates another direct target of RNAIII that plays a role in virulence. It also illustrates the diversity of RNAIII-mRNA topologies and how these multiple RNAIII-mRNA interactions would mediate virulence regulation.

# *Staphylococcus aureus* RNAIII Binds to Two Distant Regions of *coa* mRNA to Arrest Translation and Promote mRNA Degradation

Clément Chevalier<sup>1</sup>, Sandrine Boisset<sup>2</sup>, Cédric Romilly<sup>1</sup>, Benoit Masquida<sup>1</sup>, Pierre Fechter<sup>1</sup>, Thomas Geissmann<sup>2</sup>, François Vandenesch<sup>2\*</sup>, Pascale Romby<sup>1\*</sup>

<sup>1</sup>Architecture et Fonction de l'ARN, Université de Strasbourg, CNRS, IMC, Strasbourg, France, <sup>2</sup>INSERM U911, Centre National de Référence des Staphylocoques, Lysis, France, <sup>3</sup>Université de Lyon, Lyon, France

## Abstract

*Staphylococcus aureus* RNAIII is the intracellular effector of the quorum sensing system that temporally controls a large number of virulence factors including exoproteins and cell-wall-associated proteins. Staphylocoagulase is one major virulence factor, which promotes clotting of human plasma. Like the major cell surface protein A, the expression of staphylocoagulase is strongly repressed by the quorum sensing system at the post-exponential growth phase. Here we used a combination of approaches *in vivo* and *in vitro* to analyze the mechanism used by RNAIII to regulate the expression of staphylocoagulase. Our data show that RNAIII represses the synthesis of the protein through a direct binding with the mRNA. Structure mapping shows that two distant regions of RNAIII interact with *coa* mRNA and that the mRNA harbors a conserved signature as found in other RNAIII-target mRNAs. The resulting complex is composed of an imperfect duplex marking the Shine-Dalgarno sequence of *coa* mRNA and of a loop-loop interaction occurring downstream in the coding region. The imperfect duplex is sufficient to prevent the formation of the ribosomal initiation complex and to repress the expression of a reporter gene *in vivo*. In addition, the double-strand-specific endonuclease II cleaves the two regions of the mRNA bound to RNAIII that may contribute to the degradation of the repressed mRNA. This study validates another direct target of RNAIII that plays a role in virulence. It also illustrates the diversity of RNAIII-mRNA topologies and how their multiple RNAIII-mRNA interactions would mediate virulence regulation.

**Citation:** Chevalier C, Boisset S, Romilly C, Masquida B, Fechter P, et al. (2010) *Staphylococcus aureus* RNAIII Binds to Two Distant Regions of *coa* mRNA to Arrest Translation and Promote mRNA Degradation. *PLoS Pathogens* 6(3): e1000809. doi:10.1371/journal.ppat.1000809

**Received:** November 17, 2009; **Accepted:** February 5, 2010; **Published:** March 12, 2010

**Copyright:** © 2010 Chevalier et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the Centre National de la Recherche Scientifique (CNRS, PR), the Institut National pour la Recherche Médicale (INSERM, PV), the Agence Nationale pour la Recherche (ANR09-BLAN-0024-01, PR & PV), and the European Community (FP6-018618, PR, FP7-EC22718, TG), Regas Abacc ICC, CR and DGA ICR. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

\* Email: densesch@univ-lyon1.fr (FV); p.romby@imc.u-strasbg.fr (PR)

## Introduction

*Staphylococcus aureus* is ubiquitous in the environment and is a commensal organism found on human skin. This major human pathogen is the most common cause of hospital- and community-acquired infections. Therefore, *S. aureus* has developed a plethora of strategies to survive in various environmental niches. The broad range of human infections caused by *S. aureus* is in part due to the production of a large number of virulence factors. These factors mediate cell and tissue adhesion, contribute to tissue damage and spreading, and protect the bacteria against the host immune defense system. Coordinated virulence gene expression is thought to be critical for infection and is orchestrated by multiple factors involving two-component systems, global regulatory proteins, and the quorum-sensing system [1,2]. Quorum-sensing regulation in staphylococci is mainly driven by the *agr* system, which was shown to exert a variety of functions in bacterial physiology and pathogenesis [2,3]. The *agr* system is composed of two divergent transcription units, RNAII and RNAIII. RNAII contain a density-sensing cassette (*agrD* and *B*) and a two-component sensory transduction system (*agrA* and *C*). Upon a threshold level

of cell density, the response regulatory protein, AgrA, activates the transcription of its own operon and of the regulatory RNAIII [4]. Recent data show the existence of two distinct *agr* regulatory circuits; one is RNAIII-independent and the other is RNAIII-dependent [3]. Although RNAIII controls the expression of many virulence factors, the expression of several enzymes involved in carbohydrate and amino acid metabolisms are downregulated by an unknown mechanism that is independent of RNAIII. Furthermore, AgrA directly activates the synthesis of several phenol-soluble modulins (PSM) peptides at the transcriptional level [3]. Hence, AgrA and RNAIII act in concert to regulate the synthesis of many proteins in response to cell density, interconnecting metabolism, and virulence gene expression [3,5,6].

RNAIII has a dual function because it acts as a mRNA that encodes a PSM peptide,  $\Delta$ -hemolysin, and temporally controls the switch between early expression of surface proteins and late expression of several exotoxins [1]. RNAIII belongs to the class of *trans*-acting RNAs, which regulate several mRNAs at the post-transcriptional level [7,8]. The 5' domain of RNAIII activates translation of *hla* mRNA (encoding  $\alpha$ -hemolysin) by preventing the formation of an intramolecular mRNA structure that sequesters

### Author Summary

*Staphylococcus aureus* causes a wide spectrum of diseases in humans and is one of the main causes of community as well as hospital-acquired infections. *S. aureus* produces a large number of virulence factors that are expressed in a coordinated manner and at appropriate time and place. To this end, a set of multiple interacting regulators, including regulatory proteins and RNA, is brought into play. The ability of organisms to use RNA to modulate gene expression is a relatively new concept. This is the case for the largest regulatory RNA, *S. aureus* RNAIII, which controls the switch between the expression of surface proteins and excreted toxins. Here we used a combination of approaches *in vivo* and *in vitro* to analyze the mechanism used by RNAIII to regulate the expression of the major virulence factor, staphylocoagulase, which promotes clotting of human plasma. We found that RNAIII regulates the expression of staphylocoagulase through direct interactions with its mRNA. RNAIII binds to two distant regions of *coa* mRNA to arrest translation and in a coordinated manner, the endoribonuclease III recognizes the formed duplex to initiate degradation of the repressed mRNA. Staphylocoagulase belongs to the early expressed virulence factors, such as protein A, that are repressed by RNAIII using a similar mechanism. This study illustrates the diversity of RNAIII-mRNA topologies and how these multiple RNAIII-mRNA interactions would mediate virulence regulation.

the *dy* ribosome binding site [13]. The 3' end and the central domain of RNAIII (Fig. S1) repress the synthesis of early expressed cell surface virulence factors (protein A, fibrinogen-binding protein) as well as the transcriptional regulator, Rot, the repressor of toxins [1,10–12]. We have previously shown that the 3' domain of RNAIII, which is the most highly conserved domain, could also form base pairings with *coa* mRNA encoding staphylocoagulase [12]. Staphylocoagulase is an extracellular protein produced by almost all clinical isolates of *S. aureus*, which specifically forms a complex with prothrombin, the so-called staphylothrombin, to promote fibrin formation in human plasma. Like the major cell surface protein A, the synthesis of staphylocoagulase is growth-phase dependent, and the protein is produced during exponential growth and rapidly repressed by the *agr* system [13].

We show here that RNAIII is responsible for the *in vivo* repression of staphylocoagulase at the post-transcriptional level. This results from a direct interaction of two distant domains of RNAIII with *coa* mRNA. The complex is formed rapidly and is stable enough to prevent the binding of the ribosomal 30S subunit and, in addition, provides binding sites for the endoribonuclease III. Thus, *coa* mRNA belongs to the RNAIII-dependent repressed mRNAs that are regulated by a similar mechanism. This work and previous data also illustrate the variety of RNAIII-mRNA topologies that are sufficient to block the access of the ribosome at the initiation step.

## Results

### RNAIII regulates the synthesis of staphylocoagulase at the post-transcriptional level

Sequence complementarity between RNAIII (nucleotides (nt) 391 to 437) and *coa* mRNA (nucleotides 15 to 52) suggested that the 3' domain of RNAIII can repress *coa* expression at the post-transcriptional level through the formation of RNAIII-mRNA interactions [12]. To validate the *in vivo* relevance of such a

mechanism, we analyzed the expression of gene reporter constructs in various *S. aureus* strains that expressed the wild-type RNAIII or mutated versions of RNAIII. The 3' start of *coa* mRNA was determined by 5' rapid amplification of cDNA ends (RACE) showing that the 5' untranslated region contains 25 nucleotides upstream the AUG initiation codon. The entire leader regulatory region of the *coa* gene, including 50 nucleotides of the coding sequence, was cloned *in-frame* with the *lacZ* gene into the pTCY-*lacZ* shuttle vector [14]. This construct is under control of an *agr*-independent promoter (P<sub>III</sub>). The  $\beta$ -galactosidase activity was determined in the *S. aureus* strain LUG1407 (wt, *mslIII*), which express RNAIII and in LUG1457 (*ΔmslIII*), which carries a deletion of the *mslIII* gene (Fig. 1A). We also measured the synthesis of the  $\beta$ -galactosidase from the *mslIII* locus in the strains lacking the *coa* gene, carrying the endoribonuclease III (RNase III) LUG1405,  $\Delta$ rot, or of *dy* gene, encoding the 50-kDa Hly protein (LUG114),  $\Delta$ hly. The  $\beta$ -galactosidase activity was reduced six-fold in the LUG1407 strain (wt, *mslIII*) compared to the LUG1457 (*ΔmslIII*) strain. Furthermore, Hly had no significant effect on the RNAIII-dependent repression, while the deletion of *coa* abolished the repression of the *mslIII* locus (Fig. 1A).

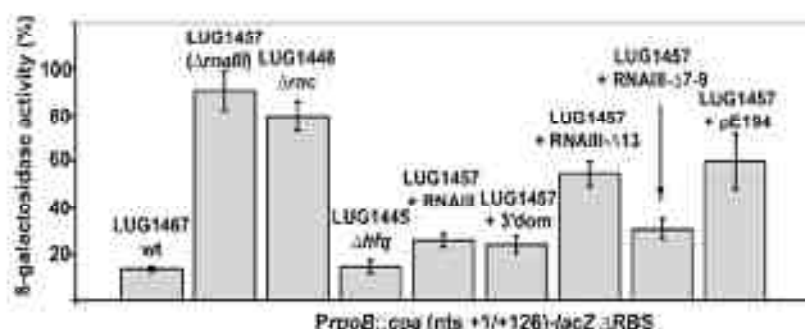
Experiments were also carried out on the LUG1457 strain (*ΔmslIII*, complemented) with different sets of plasmid pLEG274 expressing either the wild-type RNAIII, the 3' end domain comprising nt 331 to 516, RNAIII- $\Delta$ 13 deletion of hairpin 13 or RNAIII- $\Delta$ 7-9 deletion of hairpins 7 to 9 (Fig. 1A). A control experiment was carried out with the plasmid pE194 with *psl* insert. Unsurprisingly, this plasmid slightly decreased the  $\beta$ -galactosidase levels compared to the LUG1457 (*ΔmslIII*) strain. Our explanation would be that the copy number of the pTCY-*lacZ* derivatives was affected by the presence of the multiple plasmid, pE194, even though both plasmids are compatible. However, derivatives of plasmid pE194 producing high levels of wild-type RNAIII, the 3' domain, or RNAIII- $\Delta$ 7-9 reproducibly decreased the synthesis of  $\beta$ -galactosidase (Fig. 1A). Conversely, the expression of RNAIII- $\Delta$ 13, which lacks the base-pairing complementarity with *coa* mRNA, did not alter  $\beta$ -galactosidase synthesis (Fig. 1A).

We also analyzed the steady-state level of *coa* mRNA in different *S. aureus* strains in late-exponential phase (Fig. S2). The mRNA was not detected in RN6390 (wt, *mslIII*) while its level was significantly enhanced in the isogenic strain lacking *mslIII* gene (*ΔmslIII*). Of interest, in the  $\Delta$ rot strain, the level of *coa* mRNA was reproducibly found to be slightly higher than in the parental wt strain (Fig. S2). This result suggests that the RNase III-dependent degradation of the mRNA contributes in part to the disappearance of the mRNA pool. Complementation assays were also done with the mutant *ΔmslIII* strain transformed with plasmids expressing several variants of RNAIII. The expression of the 3' domain of RNAIII strongly reduced the level of *coa* mRNA while significant levels of the mRNA were still detected in the strain expressing RNAIII- $\Delta$ 13 (Fig. S2).

Taken together, these results strongly suggest that RNAIII and RNase III coordinately repress *coa* expression at the post-transcriptional level and that the hairpin 13 of RNAIII is essential for the repression.

### RNAIII binds to two distant regions of *coa* mRNA

The predicted base-pairing between RNAIII and *coa* mRNA and the *in vivo* experiments suggested that the RNAIII-dependent repression of *coa* mRNA was governed by direct RNAIII-mRNA pairing. We thus mapped the regions of interactions using enzymatic and chemical probing. The conformation of *coa* mRNA was probed using RNase T1 [specific for unpaired guanines], RNase



**Figure 1. RNAIII-dependant regulation of *coa* mRNA in vivo.**  $\beta$ -galactosidase activity detected from different gene fusions.  $\beta$ -galactosidase activity measured from PprpB-*coa* (+1/+126)-*lacZ* fusions in various *S. aureus* strains: LUG1467 (rnaIII<sup>+</sup>, wt), LUG1446- $\Delta$ rnc (deletion of *rnc* gene encoding RNase III), LUG1445- $\Delta$ hfg (deletion of *hfg* gene), LUG1457 ( $\Delta$ rnaIII), and LUG1457 transformed with the plasmid expressing the wild type RNAIII, the 3' domain, RNAIII- $\Delta$ 13 (RNAIII deleted of hairpin 13), or RNAIII- $\Delta$ 7-9 (RNAIII deleted of hairpins 7 to 9), or with the plasmid containing no insert (pE194). The  $\beta$ -galactosidase activity was normalized for total cell density and is represented as a percentage of the uninhibited control (LUG1457). The results represented a mean of three independent experiments. doi:10.1371/journal.ppat.1000809.g001

VI (specific for helical regions), and several base-specific chemicals such as dimethylsulphate (methylates N1A>N3C), a carbodiimide derivative (modifies N3U>N1G), and diethylpyrocarbonate (carboxyethylates N7A). Several experiments of RNA structure probing are shown in Fig. 2. The secondary structure model of *coa* mRNA, which explains most of the probing data, is comprised of three stem-loop structures connected by unpaired residues (Fig. 3A). The AU-rich hairpin I is of weak stability but is proposed to occur based on the enzymatic cleavage pattern. However, the coexistence of alternative structures in the region encompassing nucleotides 10 to 70 may explain the concomitant presence of RNase VI cleavages and the reactivity of many nucleotides at one of their Watson-Crick positions. In contrast, the long hairpin structure III located in the coding region of *coa* mRNA is well supported by the enzymatic cleavage patterns and the non reactivity of the Watson-Crick position of A77 to U85 towards chemicals (Fig. 2B-E). Binding of RNAIII induced changes in the region encompassing the ribosome binding site (RBS, nucleotides U10 to A88). RNAIII protected the guanines of the Shine-Dalgarno (SD) sequence against RNase T1 as well as the nucleotides U10 to U18 and A40 to A88 against chemical modifications (Fig. 2D, E). Concomitantly, RNAIII binding induced new RNase VI cleavages at positions 39-41 and enhanced reactivity of A21, A24, A25, A29 to A31 at position N1, of A21 at position N7, and of U26 and U27 at position N3 in *coa* mRNA (Fig. 2D-E, Fig. 3B). These reactivity changes in the RBS of *coa* mRNA most likely resulted from the binding of the hairpin 13 of RNAIII because its deletion in RNAIII conferred no additional effect on the accessibility of the RBS of *coa* mRNA (Fig. 2C). Binding of *coa* mRNA to the 3' domain or to RNAIII induced correlated changes in hairpin 13. Strong protections were observed at G441 against RNase T1 and at positions 411-415 and 448-449 against RNase VI (Fig. 2A). Concomitantly, increased RNase VI cleavages were observed at positions 433-434 and 444. All these data are thus consistent with the formation of a RNAIII-mRNA duplex that sequestered the RBS of *coa* mRNA (Fig. 3). This imperfect duplex involves two consecutive regions of 13 base-pairings interrupted by an internal loop and a bulged adenine 21 (Fig. 3B).

Unexpectedly, we also found a second RNAIII binding site restricted to the apical loop III of *coa* mRNA (Fig. 3B). Binding of RNAIII reduced considerably the RNase T1 cuts at G94-97 and

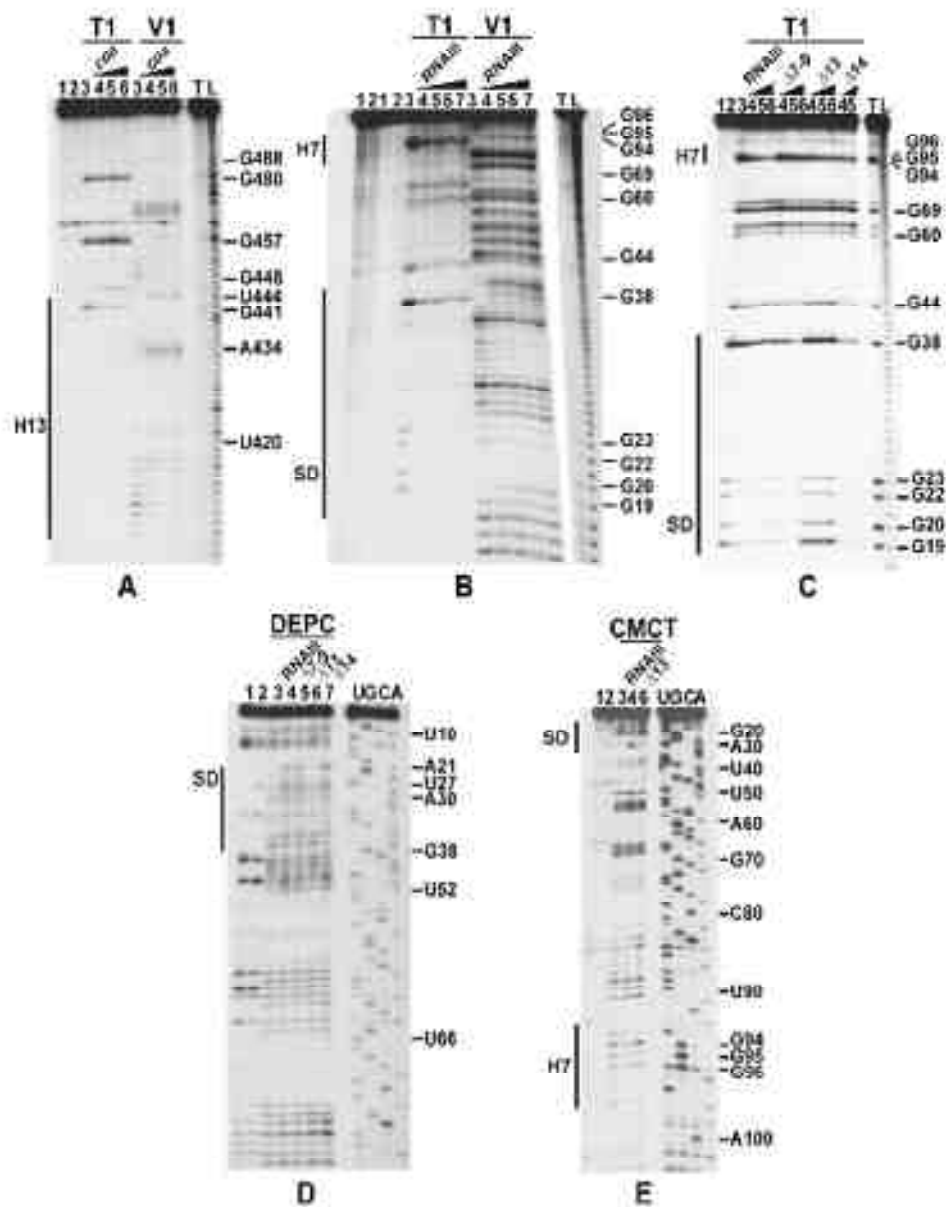
the modifications of the nucleotides UGGGAU98 mediated by the chemicals (Fig. 2B, E). Concomitantly, RNAIII binding induced several RNase VI cuts at positions 96 to 98 (Fig. 2B). These changes were abolished if the complex was formed between *coa* mRNA and the RNAIII deleted of hairpins 7 to 9 (Fig. 2C). Furthermore, *coa* mRNA binding to RNAIII reduced significantly the reactivity of the nucleotides CCGA243 towards IDMS in the apical loop 7 of RNAIII. These data are strengthened by the sequence complementarities between the apical loop III of *coa* mRNA and the hairpin loop 7 of RNAIII and support the existence of a loop-loop interaction (Fig. 3A, B). This interaction is, however, strongly dependent on the formation of the imperfect duplex because the reactivity changes in the hairpin loop III of *coa* mRNA were significantly decreased if complex formation was performed with RNAIII- $\Delta$ 13 (Fig. 2C). Molecular modeling of the RNA interaction between the two loops shows an almost continuous stacking from the 3' side of the helix III of *coa* mRNA, through the loop-loop intermolecular helix to the helix of the hairpin 7 of RNAIII. The two connecting loops of three and two nucleotides bridge the grooves of the newly formed helix (Fig. 3C).

Altogether, the data show that the mRNA-RNAIII complex is composed of a bipartite site, which implies the formation of an imperfect duplex and a loop-loop interaction.

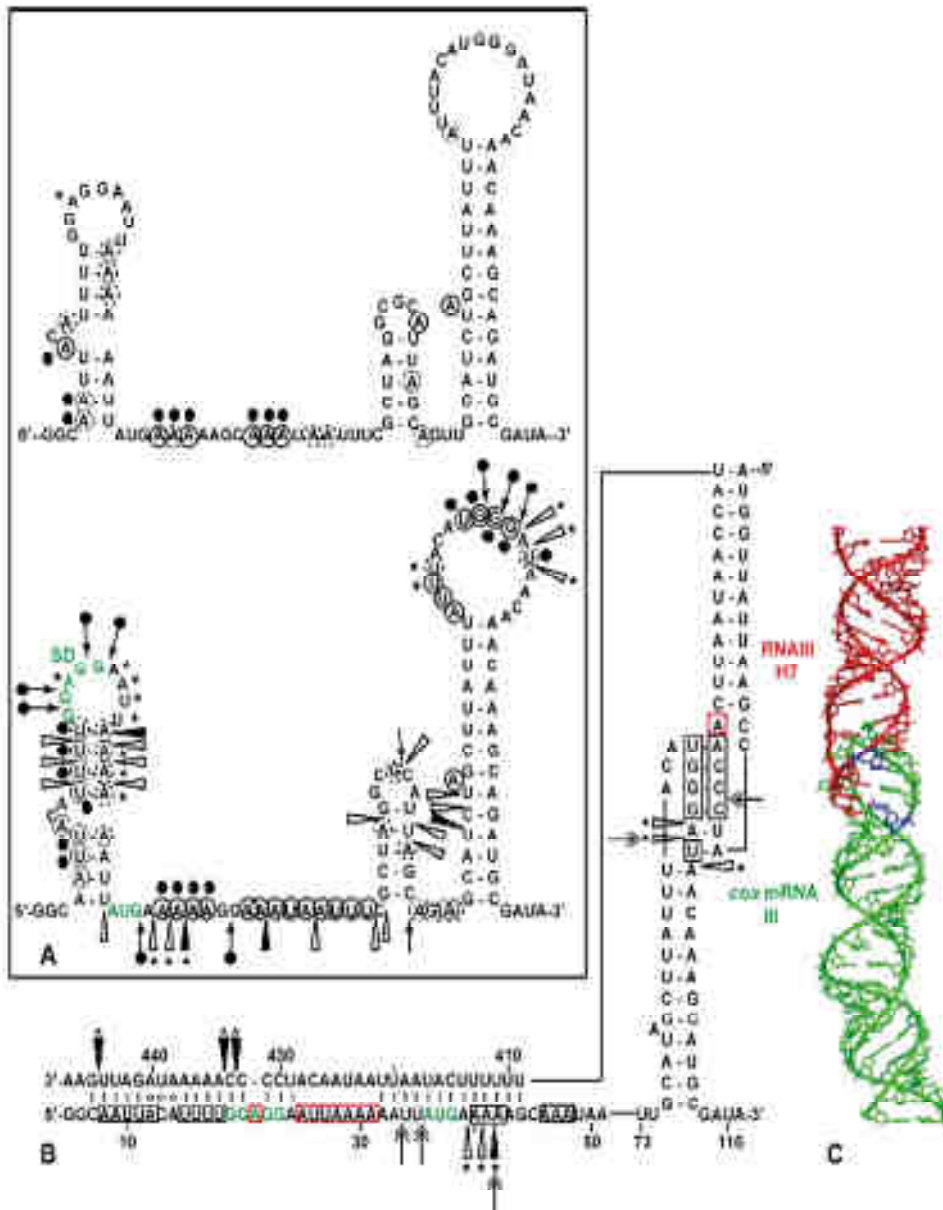
#### The ribosome binding site of *coa* mRNA is the major recognition site for RNAIII

The contribution of the two binding sites toward complex formation was further evaluated by gel shift assays. Each experiment has been reproduced four times. *In situ* labeled *coa* mRNA was first incubated with increasing concentrations of RNAIII or its variants (RNAIII- $\Delta$ 13, RNAIII- $\Delta$ 7-9, and the 3' domain) at 37°C for 15 min (Fig. 4A). This experiment shows that *coa* mRNA binds to RNAIII or its 3' domain with a  $K_d$  value of around 10 nM. The deletion of hairpins 7 to 9 in RNAIII had only a two-fold effect on the dissociation constant (around 25 nM), while the deletion of hairpin 13 in RNAIII increased significantly the  $K_d$  value by one order of magnitude (around 150 nM).

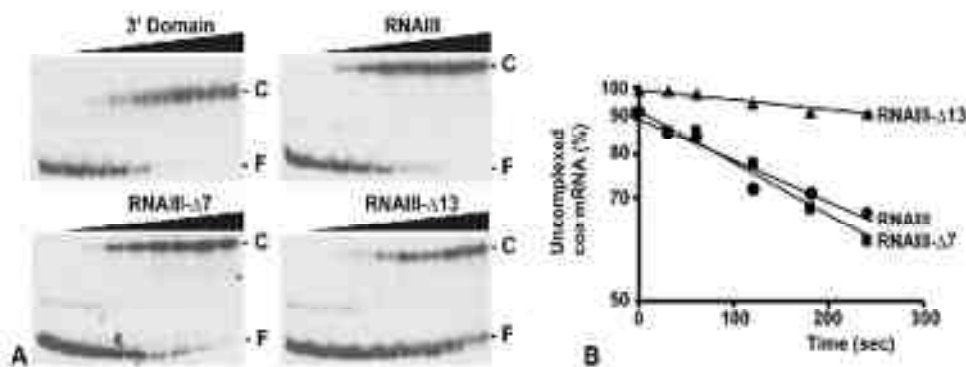
The initial rate of wild type RNAIII binding to 5' end-labeled *coa* mRNA was estimated from a time-course analysis and resulted in an association rate constant of  $1.1 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$  (Fig. 4B).



**Figure 2. Enzymatic and chemical probing of the structure of the inhibitory RNAIII-*coa* mRNA complex.** (A) Enzymatic hydrolysis of 5'-end-labeled 3' domain of RNAIII, alone (lane 3) or in the presence of an excess of *coa* mRNA (lane 4, 20 nM; lane 5, 50 nM; lane 6, 100 nM; lane 7, 250 nM). Lanes 1, 2: incubation controls on free RNA or bound to *coa* mRNA, respectively. Lanes T, L: RNase T1 under denaturing conditions and alkaline ladders, respectively. T1, V1: RNase T1 and RNase V1 hydrolysis, respectively. (B, C) Enzymatic hydrolysis of 5'-end-labeled *coa* mRNA, alone (lane 3) or bound to the wild type RNAIII, or to the mutant RNAIII deleted of hairpins 7 to 9 ( $\Delta 7-9$ ), of hairpin 13 ( $\Delta 13$ ) or of hairpin 14 ( $\Delta 14$ ). Concentrations of wild type or mutant RNAIII: lane 4, 20 nM; lane 5, 50 nM; lane 6, 100 nM; lane 7, 250 nM. Same legend as in A. (D) DEPC (N7A) modification of unlabeled *coa* mRNA, free (lane 3) or bound to the wild type RNAIII (lane 4), or the mutant RNAIII deleted of hairpins 7 to 9 ( $\Delta 7-9$ , lane 5), of hairpin 13 ( $\Delta 13$ , lane 6) or of hairpin 14 ( $\Delta 14$ , lane 7) at 200 nM. Lanes U, G, C, A: dideoxy-sequencing reactions performed on *coa* mRNA. (E) CMCT modification of unlabeled *coa* mRNA. Same legend as in D. Reactivity changes are indicated by bars on one side of each autoradiography. SD is for Shine-Dalgarno sequence and H7 is for hairpin 7 of RNAIII.  
doi:10.1371/journal.ppat.1000809.g002



**Figure 3. Structure of the RNase III-coo mRNA complex.** (A) Summary of the enzymatic cleavages and chemical reactivities of nucleotides of *coo* mRNA. Enzymatic cleavages are given as follows: RNase V1 (black arrow), and RNase III (white arrowhead) nucleotides, (black arrowhead) strong cleavage. Chemical modifications of cytosines at N3, and adenines at N1 by DMS, of uridines at N3 and guanines at N1 by CMCT, and of adenines at N7 by DEFC. Full and dashed circled nucleotides are for strong and moderate reactivity, respectively. No symbol is for non reactive, nd is for not determined due to non-specific cleavages or pauses of RT in the incubation control. The reactivity of A at N7 is reported on the secondary structure shown in the insert. Reactivity changes induced by the binding of RNase III are indicated as follows: black circles denote strong protection, enhancements and new RNase V1 cleavages are represented by asterisks. (B) Secondary structure model of the RNase III-coo mRNA complex showing the reactivity changes induced by complex formation: RNase V1 (white and black arrowheads), RNase III cleavages (double arrow). Effect of RNase III binding: Protected nucleotides are squared in black and the nucleotides, which become accessible, are squared in red. (C) The topology of the loop-loop interaction built by graphic modeling based on the probing data. The hairpin III of *coo* mRNA is in green and the hairpin 7 of RNase III in red. The RNase III cleavages are shown in blue.  
 doi:10.1371/journal.ppat.1000809.g003



**Figure 4. RNAIII binds efficiently to *coa* mRNA in vitro.** (A) Determination of the apparent dissociation constant for RNAIII-*coa* mRNA complex. 5'-end-labeled *coa* mRNA was incubated alone (–) or with various concentrations of unlabeled wild type RNAIII, the 3' domain, RNAIII-Δ7-9, and RNAIII-Δ13 (1, 5, 10, 20, 50, 100, 200, and 250 nM). The fraction of labeled *coa* mRNA associated with RNAIII or its derivatives was calculated from the counts in the corresponding band relative to the total counts in the lane. The  $K_d$  value was estimated as the concentration of RNAIII allowing 50% of *coa* mRNA binding. (B) Binding rate constant for various RNAIII-*coa* mRNA complexes as determined from three independent experiments. 5'-end-labeled *coa* mRNA (0.1 nM) was incubated with unlabeled RNAIII (20 nM), RNAIII-Δ7-9 (20 nM), and RNAIII-Δ13 (20 nM) at 37°C. Aliquots were withdrawn at various times (from 0 to 350 sec). The percentage of free *coa* mRNA was plotted as a function of time to estimate the association rate constant according to [40]. The values for the binding rate constants are the means of three independent experiments:  $1.1 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$  (RNAIII),  $9.5 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$  (RNAIII-Δ7-9), and  $1.1 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$  (RNAIII-Δ13). doi:10.1371/journal.ppat.1000809.g004

Similar values were observed for three other RNAIII-mRNA (*pta*, SA1900, *ntf*) target complexes [11,12]. These data indicate that the complexes are rapidly formed as observed for several fully complementary antisense-target RNA systems [15,16]. We then investigated whether deletion of hairpin 7 or hairpin 13 of RNAIII involved in the binding would affect binding rates (Fig. 4B). The binding rate constant for the mutant RNAIII-Δ7-9-*coa* mRNA pair was identical to the wild type complex ( $9.35 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ ). However, for the mutant RNAIII-Δ13-*coa* mRNA pair, the value was significantly decreased by one order of magnitude lower ( $1.1 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ ). These experiments strongly suggest that initial pairings involved the hairpin 13 of RNAIII and that this motif confers stable binding to *coa* mRNA.

#### RNAIII binding to *coa* mRNA interferes with the formation of the initiation complex

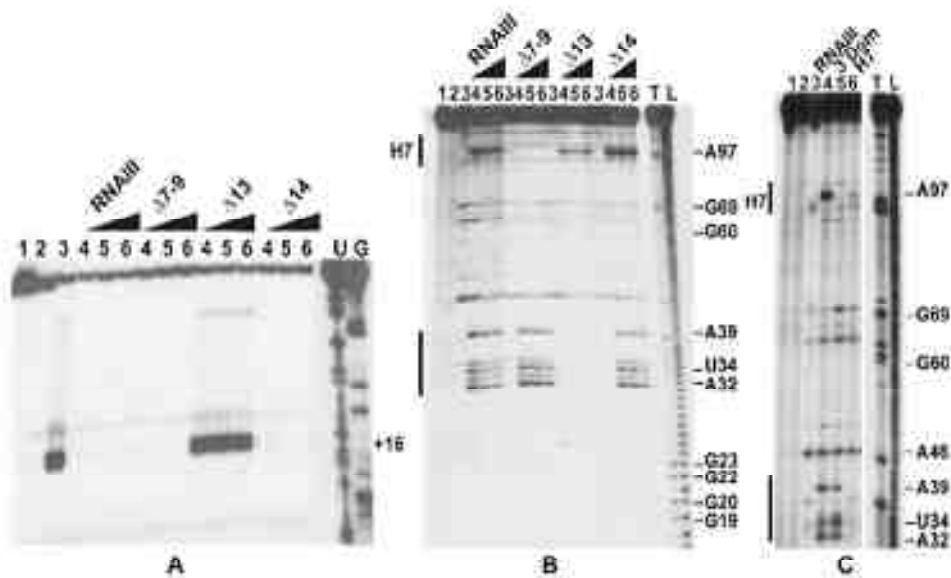
Since RNAIII binds to the SD sequence of *coa* mRNA, we analyzed whether RNAIII binding is sufficient to prevent the formation of the ternary initiation complex formed with the *S. aureus* 30S subunit, initiator tRNA<sup>Met</sup>, and *coa* mRNA. Formation of the ternary complex, which blocked the elongation of a cDNA primer by reverse transcriptase, produced a toeprint at U49/A50, 15 nucleotides downstream of the initiation codon ([17]; Fig. 5A). Intriguingly, a second toeprint resulting from ribosome binding was also observed at A14; this weak toeprint was not detected with the *E. coli* ribosomal 30S subunit (result not shown). Binding of RNAIII, RNAIII-Δ7-9 (deleted of hairpins 7 to 9), or RNAIII-Δ14 (deleted of hairpin 14) strongly decreased the two toeprint signals. This indicates that the regulatory RNAIII totally blocks access of the ribosome at the RBS site of *coa* mRNA. The inhibition was observed whether the RNAIII-mRNA complex was pre-formed or RNAIII was added together with the 30S subunit (Fig. 5A). This shows that the resulting inhibitory complex is rapidly formed and sufficiently stable to prevent the formation of the ribosomal initiation complex. Using this assay, we were not able to analyze the contribution of the loop-loop interaction in the inhibition of ribosome binding because the primer used for

elongation hybridized in the long hairpin loop III. However, the RNAIII-ΔH13 exerted no inhibitory effect on ribosome binding, showing that the specific RNAIII-mediated inhibition of ribosome binding to *coa* mRNA resulted mainly from the sequestration of the RBS by the hairpin 13 of RNAIII (Fig. 5A).

#### RNAIII cleaves the two distant mRNA-RNAIII interactions in vitro

As RNAIII is required for efficient repression *in vivo* (Fig. 1), we analyzed whether this enzyme can cleave the complex *in vitro*. We have shown previously that cleavage assays by RNAIII can be a useful tool for probing *in vitro* RNA-RNA complexes [18]. The RNAIII-dependent cleavages were probed on the 5'-end labeled RNAs as well as on the native RNAIII-*coa* mRNA complex using a purified His-tagged RNAIII from *S. aureus* (Fig. 5B-C). Only weak RNAIII cleavages were observed in the free *coa* mRNA. When the 5'-end-labeled mRNA was incubated with RNAIII, four major cleavages occurred at positions 32, 34, 39, and 97 in the mRNA (Fig. 5B-C). Binding of *coa* mRNA induced a RNAIII-dependent cleavage at C241 of the labeled RNAIII (results not shown). Thus, the two regions of hybridization were susceptible to RNAIII cleavages.

Using truncated versions of RNAIII and the isolated hairpin 7 or the 3' domain, we were able to assign the partners involved in the RNAIII-mRNA complex. Indeed, the hairpin 7 only induced a specific RNAIII-cleavage at position 97 of the mRNA, while the hairpin 13 binding promoted major cleavages at positions 32, 34, and 39 of the mRNA (Fig. 5C). Furthermore, the complex formed between RNAIII-Δ7-9 and *coa* mRNA was cleaved efficiently by RNAIII at positions 32, 34, and 39 of the mRNA (Fig. 5B). Conversely, only one RNAIII-mediated cleavage was detected at position 97 of *coa* mRNA bound to RNAIII-Δ13 or to hairpin 7 (Fig. 5B-C). This cleavage was, however, weaker than the cleavage found in the wild type complex. These experiments correlate well with the probing data showing that the loop-loop interaction is stabilized by the duplex formed between the RBS of *coa* mRNA and the hairpin 13 of RNAIII. In the irregular duplex, RNAIII



**Figure 3. The RNAIII-*coa* mRNA complex prevents ribosome binding and promotes RNase III cleavages.** (A) Formation of the ternary complex between *coa* mRNA (15 nM), 2.5-fold 30S ribosomal subunits (250 nM), and initiator tRNA (1  $\mu$ M) was monitored in the absence (lane 1) or in the presence of increasing concentrations of wild-type RNAIII, RNAIII- $\Delta$ 7-9 ( $\Delta$ 7-9), RNAIII- $\Delta$ 13 ( $\Delta$ 13), and RNAIII- $\Delta$ 14 ( $\Delta$ 14): lane 4, 25 nM; lane 5, 50 nM; lane 6, 100 nM. The toeprint at position +16 is indicated. Lanes 1, 2: incubation controls on free RNA or RNA bound to RNAIII, respectively. Lanes U, G: dideoxy-sequencing reactions performed on *coa* mRNA. (B) RNase III hydrolysis of 5'-end-labeled *coa* mRNA, alone (lane 1) or in the presence of an excess of wild type RNAIII, RNAIII- $\Delta$ 7-9 ( $\Delta$ 7-9), RNAIII- $\Delta$ 13 ( $\Delta$ 13), RNAIII- $\Delta$ 14 ( $\Delta$ 14): lane 4, 25 nM; lane 5, 50 nM; lane 6, 100 nM. Lanes 1, 2: incubation controls on free *coa* mRNA or bound to RNAIII, respectively. Lanes T1, L: RNase T1 and alkaline ladders, respectively. (C) RNase III hydrolysis of 5'-end-labeled *coa* mRNA, alone (lane 1) or in the presence of an excess of wild type RNAIII (lane 4), the 3' domain of RNAIII (3' Dom), and the hairpin 7 of RNAIII (H7). Lanes T1, L: RNase T1 and alkaline ladders, respectively.  
doi:10.1371/journal.ppat.1000809.g005

cleaves only from the mRNA side, whereas the enzyme induces cleavages on both strands of the loop-loop interaction leading to the classical two nucleotides 3' overhang. Taken together, these data fully support the chemical and enzymatic probing showing that the hairpin 13 of RNAIII binds to the RBS of *coa* mRNA, while the hairpin loop 7 forms limited base pairings with the coding sequence. The data further indicate that the loop-loop interaction adopts a topology that is appropriate for efficient RNase III binding and catalysis [18].

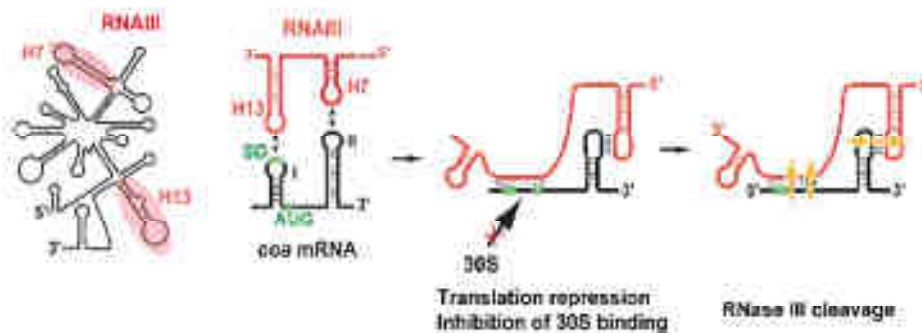
## Discussion

*S. aureus* produces a large variety of virulence factors that are required for the successful colonization of the host and that confer to the bacteria the ability to counteract the immune defense system of the host [2]. Among these virulence factors, staphylocoagulase primarily activates prothrombin, inducing the formation of a fibrin clot around the bacterial cell [19]. Coating the bacteria with host proteins contributes to hiding the bacteria from the immune system and from phagocytosis. The expression of coagulase was shown to follow a temporal regulation, as do several adhesins and surface proteins that are expressed earlier than the secreted enzymes, immunotoxins, and cytotoxins [2]. Furthermore, coagulase belongs to the early expressed virulence factors such as protein A, the fibrinogen-binding protein SA1000, and the ScaA-like protein SA2353, which were found to be repressed by the quorum sensing-controlled RNAIII. During the growth cycle, the

level of RNAIII varies inversely with that of *coa* mRNA [13]. In addition, it was shown that the coagulase expression was both positively and negatively controlled by an *agr*-dependent mechanism. A functional *agr* element resulted in a relative elevation of the *coa* mRNA level at the early exponential phase of growth followed by a strong decrease of the mRNA level at the post-exponential phase of growth [13]. We demonstrate here that the *agr*-dependent repression effect on *coa* mRNA is most probably the result of a direct binding of RNAIII to *coa* mRNA.

We show that RNAIII in conjunction with RNase III are required to fully repress the synthesis of staphylocoagulase at the stationary phase of growth (Fig. 6). The primary effect of RNAIII would be to prevent translation initiation subsequently followed by the RNase III-dependent cleavage of the repressed mRNA. Since we have previously shown that RNase III binds efficiently to RNAIII, we propose that RNAIII-dependent translation repression and RNase III cleavage are coupled. Hence, these data, together with previous works, show that RNAIII represses the synthesis of coagulase, protein A, SA1000, SA2353, and Rot by a similar mechanism [10–12]. In addition, probing the mRNA structure also shows that *coa* mRNA adopts a very similar structural organization to *yle* and SA1000 mRNAs (Fig. S3). The three mRNAs have short 5' untranslated regions, which carry a 5' hairpin structure with a strong SD sequence located in the apical loop (Fig. S3). In the absence of RNAIII, these elements may confer to the mRNAs a high stability [11]. Indeed, in *B. subtilis*, stabilization of mRNAs was





**Figure 6. Schematic view of RNase III-mediated repression of *coa* mRNA.** RNase III binds to its target mRNA masking the HT1 and part of the coding sequence. Binding of RNase III hinders ribosome binding and promotes access to RNase III. SD (Shine-Dalgarno) and AUG are in green. RNase III is in red, the mRNA target is in black, 30S is for small ribosomal subunit. doi:10.1371/journal.ppat.1000809.g006

shown to be a consequence of the blocking of the 5' end by a stalled initiating ribosome at a SD-like sequence [20,21] or by a stable 5' hairpin structure and a strong RBS [22]. Therefore, the coordinated action of RNase III and RNase III would be needed to irreversibly repress the synthesis of these virulence factors at an appropriate time.

*In vitro* binding assays show that RNase III binds to *coa* mRNA and its other mRNA targets with a rather high association rate constant. Efficient repression by non-coding RNAs (ncRNA), which act at the translational level, requires that the ncRNA binds to target mRNAs within a short time frame, i.e. before the formation of the stable ribosomal initiation complex [15,16]. Our data also indicate that *S. aureus* Ssa-like Hfq protein is not required for the RNase III-dependent repression of *coa* mRNA *in vivo* (Fig. 1), in contrast to *Escherichia coli* and *Salmonella typhimurium* ncRNAs which act in concert with Hfq to bind mRNA targets [23,24]. Despite the fact that Hfq binds to RNase III [11], the observation that the deletion of *hfq* does not exhibit severe phenotypic defects rules out the direct involvement of Hfq in regard to RNase III-mediated regulation in *S. aureus* [10,25]. Instead, we propose that the structures of RNase III and its mRNA targets may compensate for the need of a helper protein as shown for antisense RNAs fully complementary to their target mRNAs [15,16]. We, however, do not rule out that another protein or RNase III could contribute to stabilize and/or facilitate the formation of the hybrid [11].

The regions of interaction in RNase III and *coa* mRNA contained stem-loop structures that are indeed well appropriate for initial loop-loop interactions. The two conserved C-rich loops, 7 and 13, of RNase III bind to the RBS and to the hairpin loop III in the coding sequence of *coa* mRNA, respectively (Fig. 3, 6). These C-rich hairpin loops of RNase III are also used to repress the other mRNA targets, although the topologies of the resulting inhibitory complexes are different (Fig. S3). RNase III forms long duplexes with the RBS of *gta* and SA1000 mRNAs, while it forms two loop-loop interactions with the 5'UTR and the RBS of *rot* mRNA, respectively [11,12] (Fig. S3). Here we show that the RNase III-*coa* mRNA complex involves an imperfect duplex of two stretches of 13 base pairs separated by a bulged loop that sequestered the RBS, and a loop-loop interaction that took place in the coding region. In contrast to *rot* mRNA, in which the two loop-loop interactions were essential for *in vivo* repression, the sequestration of the RBS of *coa* mRNA is sufficient by itself to promote efficient repression *in vivo* and to prevent the formation of the ribosomal initiation complex (Fig. 5A, 6). Indeed, the loop-loop interaction is not

essential for efficient *in vivo* repression and contributes only moderately to the stability of the inhibitory complex. Hence, the various topologies of the repressed RNase III-mRNA complexes depend largely on the mRNA context.

The inhibitory RNase III-*coa* mRNA complex also provided specific binding sites for the double strand-specific RNase III, which induced strong cleavages in the two regions of *coa* mRNA bound to RNase III. Notably, the cleavage sites in the loop-loop interaction also occurred at a similar position in the two kissing interactions that took place in the *rot* mRNA-RNase III complex (Fig. S3; [12]). The sequences of *coa* mRNA involved in the loop-loop interactions are very similar to *rot* mRNA, showing that similar signatures exist in various RNase III-repressed mRNAs (Fig. S3). Molecular modeling of the kissing interaction, which took into account the chemical and enzymatic probing data, revealed that the loop-loop interaction induces a coaxial stacking of the two intramolecular helices (Fig. 3C). The overall topology is very similar to the RNA loop-loop structure obtained by NMR, which mimics the interaction between sense and antisense RNAs involved in the regulation of the ColE1 plasmid [26]. Such a long helical structure might well be appropriate for the binding of the homodimeric enzyme, although the sequence of the kissing interactions might also be a specific binding determinant. Of interest, *coa* mRNA was shown to be completely depleted as soon as RNase III was produced, and the deletion of *nc* caused the accumulation of *coa* mRNA [13] (Fig. S2). Therefore, as we postulated previously, RNase III might initiate rapid degradation of *coa* mRNA, and the cleavage in the loop-loop interaction may also contribute to access to several other endo- or exoribonucleases for further degradation (Fig. 6). Notably, at a similar position, *gta* and SA1000 mRNAs carry a long stem-loop structure in the coding sequence that is also cleaved efficiently by RNase III [11,12] (Fig. S3). In addition, the depletion of the mRNA might also result from an indirect effect of RNase III. Indeed, Rot protein was shown to activate the transcription of *coa* mRNA [27], while RNase III represses the synthesis of Rot at the post-transcriptional level [10]. Thus, the RNase III-mediated repression of coagulase would occur at both transcriptional and post-transcriptional levels as it was shown for *gta* mRNA [10,12,27].

It is not an exception that RNase III utilizes conserved C-rich loops to target similar regions of various mRNAs that are functionally related. In *S. typhimurium*, GcvB RNA represses translation initiation of multiple target mRNAs by binding to a C/A-rich motif present in all these mRNAs, which encode

periplasmic substrate-binding proteins of ABC uptake systems for amino acids and peptides [28,29]. *E. coli* CyaR contains a hairpin loop with a conserved anti-SD sequence that is used to target the SD sequence of a subset of mRNAs [30,31]. Similarly to RNAIII, we recently found that other *S. aureus* ncRNAs carry a similar UGGC signature always present in an unpaired region, and through its unpaired C-rich motif, one of these RNAs binds to the RBS and represses the expression of several mRNAs [32]. *S. aureus* *coo* mRNA and the other mRNA targets of RNAIII carry a strong SD sequence located in an unpaired region that is quite appropriate for the docking of the 30S subunit, but also for the formation of initial contacts with the C-rich loop of RNAIII (Fig. S1). Specificity for *coo* regulation is mainly given by the propagation of the intermolecular contacts to form a long imperfect duplex further stabilized by a loop-loop interaction in the coding sequence.

In conclusion, this study validates another direct target of RNAIII that plays a role in virulence. Our study further stresses that the RNAIII harbors highly conserved regions that provide a specific signature to generate interactions with the RBS of multiple mRNAs and that the mRNA context directs the topology of the inhibitory complexes. Recent works focusing on *E. coli* and *S. typhimurium* show that regulatory RNAs that target mRNAs regulate gene expression through a variety of unusual mechanisms and bind to mRNA regions located far away from the ribosome binding site in the 5'UTR [33], in the coding sequence [34,35], and in the 3' end [36]. Whether *S. aureus* has also evolved such a diversity of RNA-dependent regulatory mechanisms remains to be addressed.

## Materials and Methods

### Strains and plasmids

*S. aureus* RN6390 or LUG1467 derives from 8325-4. In WA400 and LUG1457 (*AraDIII*), the *PS* operon is deleted and replaced by the chloramphenicol transacetylase gene (*cat86*) [37]. LUG774 and LUG911 strains derive from RN6390, in which *we* and *hly* genes, respectively, have been replaced by *cat86* gene [11]. *Staphylococci* were grown either on BM agar plates (1% peptone, 0.5% yeast extract, 0.1% glucose, 0.5% NaCl, 0.1% K<sub>2</sub>HPO<sub>4</sub>) or in brain-heart infusion (BHI) with erythromycin (5 µg/ml) when appropriate. RNAIII and its variants were expressed in *Staphylococcus aureus* WA400 with plasmid pE194 (see Table 1).

### Construction of translation fusions and β-galactosidase measurements

Translation fusions were constructed with plasmid pLUG220, a derivative of pTCV-*lacZ*, a low-copy-number promoter-less *lacZ* vector (Table 1). The 5' end of the *coo* mRNA was first determined by rapid amplification of cDNA ends (RACE) using the First Choice RLM-RACE kit following the company's protocol (Ambion). The whole leader region of *coo* mRNA including 126 nt of the coding sequence, was cloned downstream the *spalB* promoter in frame with *lacZ* [11]. β-galactosidase activity was measured three times on duplicate cultures with the Enzyme Assay System (Promega).

### Northern blots and measure of mRNA half-life

Electrophoresis of total RNA (20 µg) was done on a 1% agarose gel containing 2.2 M formaldehyde and vacuum transfer to nylon membrane. Hybridizations with specific digoxigenin-labeled RNA probes complementary to *coo* mRNA and luminescent detection were carried out as described previously [13].

### RNA preparation

RNAIII, RNAIII derivatives (RNAIII-Δ7-9: deletion of nt G287 to U319, RNAIII-Δ13: deletion of nt U409 to A451, and RNAIII-Δ14: deletion of nt G483 to C511, the 3' domain comprises nt 391 to 516), the isolated hairpin 7, and the *coo* mRNA fragment were transcribed *in vitro* using T7 RNA polymerase as described previously [38]. The transcribed RNAs were purified by 8% polyacrylamide-8 M urea gel electrophoresis. After elution in 0.5 M ammonium acetate/1 mM EDTA buffer, the RNAs were precipitated twice with ethanol. Before use, the pellet was dissolved in sterile bi-distilled water and the concentration was measured accurately.

The 5' end-labeling of dephosphorylated RNA or DNA oligonucleotides was performed with T4 polynucleotide kinase and [ $\gamma$ -<sup>32</sup>P]ATP [39]. Before use, RNAs were renatured by incubation at 90°C for 2 min in the absence of magnesium and salt, 1 min on ice, followed by an incubation step at 20°C for 15 min in TMN buffer (20 mM Tris-acetate pH 7.5, 10 mM magnesium-acetate, 150 mM Na-acetate).

### Determination of constants of RNAIII-*coo* mRNA complex formation

Binding rate constant of RNAIII-*coo* mRNA complex was measured as described previously [40]. Binding of end-labeled *coo* mRNA to a ten-fold excess of unlabeled RNA (RNAIII, RNAIII-Δ13, RNAIII-Δ7-9) was performed at 37°C in TMN buffer. Samples were withdrawn at various time points (0-10 min), added to gel application buffer and loaded onto a native 5% polyacrylamide gel. The gel was run at 4°C and constant voltage (300 V) for 3 h and subsequently dried. Bands corresponding to the RNAIII-*coo* mRNA complex and free RNAIII, respectively, were quantified using the SAFA algorithm [41].

For determination of the dissociation rate constant of RNAIII-*coo* mRNA complex, end-labeled *coo* mRNA was incubated with an increased molar amount of wild-type RNAIII or RNAIII variants (RNAIII-Δ13, RNAIII-Δ7-9, 3' domain, hairpin 7) for 15 min at 37°C in TMN buffer. Samples were then treated as described above. All experiments were done four times giving reproducible data.

### RNA structure probing

RNAIII-*coo* mRNA formation was carried out at 37°C for 15 min in TMN buffer. Enzymatic hydrolysis was performed in 10 µl of TMN, in the presence of 1 µg carrier tRNA at 37°C for 5 min: RNase T1 (0.0025 units), RNase V1 (0.5 units). Chemical modifications were performed on 2 pmol of *coo* mRNA or RNAIII at 20°C in 20 µl of reaction buffer containing 2 µg of carrier tRNA. Alkylation of C(N3) and A(N1) positions was done with 1 µl DMS (diluted 1/8 and 1/16 in ethanol) for 2 min, and modification of A(N7) was done with 4 µl of DEPC for 20 min at 20°C in TMN buffer. Modifications of U(N3) and G(N1) were done with 5 µl of CMCT (50 mg/ml) for 10 and 20 min in a buffer containing 50 mM Na-borate pH 8, 5 mM MgAc, and 150 mM KOAc. RNase III purification and the enzymatic cleavage assays on *coo* mRNA and on RNAIII were performed as described previously [14].

End-labeled RNA fragments were sized on 12% polyacrylamide-8 M urea slab gels. Cleavage positions were identified using RNase T1 and alkaline ladders of the probed RNA. The cleavage or modification sites of unlabeled RNAs were detected by primer extension. Details for hybridization conditions, primer extension, and analysis of the data have been described previously [38].



strains. The blot was hybridized with an RNA probe antisense to *coa* mRNA. RN6390 (WT, *coaIII+*); *S. aureus* strain lacking *coaIII* gene, WA400 (*coaIII-*); *S. aureus* strain lacking *rcs* gene, *Δrcs* strain (JUG774); WA400 transformed with a plasmid expressing the 3' domain of RNAMIII (*ΔcoaIII* + 3' Dom) or the RNAMIII deleted of hairpin 13 (*ΔcoaIII* + RNAMIII-Δ13). Ribosomal RNAs were visualized on the same membrane by ethidium bromide staining, as an internal control (data not shown). Three independent experiments provided reproducible results.

Found at: doi:10.1371/journal.ppat.1000809.s002 (3.85 MB TIF)

**Figure S3** Comparison of the secondary structures of the RNAMIII targets. (A) The secondary structures of *gpa* mRNA [11], SA1000 and *rcs* mRNA [12], and *coa* mRNA. The nucleotides base-paired to RNAMIII are in red. The conserved G-rich sequence in the mRNAs, which interact with the C-rich motif of hairpin loops 7, 13, 14 of RNAMIII are encircled in pink. (B)

## References

1. Nevick RP, Ross HF, Prejan SJ, Korshak J, Kristowich B, et al. (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J* 12: 3967–3975.
2. Nevick RP, Geisinger E (2008) Quorum sensing in staphylococci. *Annu Rev Genet* 42: 541–564.
3. Quack SV, Janssen-Lee M, Villano AE, Bach TH, Khan BA, et al. (2008) RNAMIII-independent target gene control by the *agr* quorum-sensing system: insight into the evolution of virulence regulation in *Staphylococcus aureus*. *Mol Cell* 32: 150–158.
4. Nevick RP (2003) Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Mol Microbiol* 48: 1429–1449.
5. Dunham PM, Murphy E, Haney S, Palacios D, Tucker-Kellogg G, et al. (2004) Transcription profiling-based identification of *Staphylococcus aureus* genes regulated by the *agr* and/or *ica* loci. *J Bacteriol* 186: 7541–7555.
6. Semerville GA, France BA (2005) At the crossroads of bacterial metabolism and virulence synthesis in *Staphylococci*. *Microbiol Mol Biol Rev* 73: 233–248.
7. Talyko-Arono A, Repeda F, Cosset P (2007) Small noncoding RNAs controlling pathogenesis. *Curr Opin Microbiol* 10: 182–188.
8. Watson LS, Starr G (2005) Regulatory RNAs in bacteria. *Cell* 116: 615–628.
9. Morfitt E, Taylor D, von Gabain A, Arvidson S (1995) Activation of alpha-toxin translation in *Staphylococcus aureus* by the trans-encoded antisense RNA, RNAMIII. *EMBO J* 14: 4569–4577.
10. Geisinger E, Adhikari RP, Ju R, Ross HF, Nevick RP (2008) Inhibition of *rcs* translation by RNAMIII, a key feature of *agr* function. *Mol Microbiol* 61: 1030–1048.
11. Hantlinger E, Boisset S, Savarino C, Bruto Y, Geismann T, et al. (2005) *Staphylococcus aureus* RNAMIII and the endonuclease III coordinately regulate *gpa* gene expression. *EMBO J* 24: 824–835.
12. Boisset S, Geismann T, Hantlinger E, Focher P, Bruchli N, et al. (2007) *Staphylococcus aureus* RNAMIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev* 21: 1533–1544.
13. Lebrun G, Vandersch F, Greenhalgh T, Nevick RP, Etienne J (1994) Coagulase expression in *Staphylococcus aureus* is positively and negatively modulated by an *agr*-dependent mechanism. *J Bacteriol* 176: 5538–5539.
14. Poyart C, Trias-Cast P (1997) A broad-host-range mobilizable shuttle vector for the construction of nonreplicating fusions to  $\beta$ -galactosidase in gram-positive bacteria. *FEMS Microbiol Lett* 156: 195–198.
15. Brand S (2007) Regulatory mechanisms employed by *cis*-encoded antisense RNAs. *Curr Opin Microbiol* 10: 302–309.
16. Wagner EG, Alvaris S, Roudy P (2002) Antisense RNAs in bacteria and their genetic elements. *Adv Genet* 46: 361–398.
17. Hara D, McPherson DS, Yoon R, Gold L (1988) Extension inhibition analysis of translation initiation complexes. *Methods Enzymol* 164: 419–425.
18. Chevillon C, Hantlinger E, Focher P, Boisset S, Vandersch F, et al. (2008) *Staphylococcus aureus* endonuclease III purification and properties. *Methods Enzymol* 447: 305–327.
19. Kawahara S, Morita T, Iwanaga S, Iguchi H (1981) Enzymatic properties of staphylocoagulase, an active molecular complex formed between staphylocoagulase and human prothrombin. *J Biochem* 90: 1603–1614.
20. Mahe N, Beaud L, Pellegrini O, Daux R, Wu T, et al. (2007) 5'-to-3' endonuclease activity in bacteria: role of RNase J1 in mRNA maturation and 5' stability of mRNA. *Cell* 129: 481–492.
21. Daux-Chabot E, Mahe N, Beaud L, Coudon C (2008) Ribosomes initiating translation of the *rcs* mRNA protect it from 5'-to-3' endonucleolytic degradation by RNase J1. *Mol Microbiol* 71: 1538–1550.
22. Hansbraten G, Kallstrom K, Rothberg B (2002) A 5' stem-loop and ribosome-binding but not translation are important for the stability of *Bacillus subtilis* *opd* leader mRNA. *Microbiology* 148: 1795–1805.
23. Abe H (2007) Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr Opin Microbiol* 10: 134–139.
24. Brennan RG, Link TM (2007) Hfq structure, function and ligand binding. *Curr Opin Microbiol* 10: 125–133.
25. Bohn C, Rognvald C, Boudier P (2007) No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. *BMC Microbiol* 7: 10.
26. Lee AJ, Crothers DM (1998) The solution structure of an RNA loop-loop complex: the CAGE inverted loop sequence. *Structure* 6: 995–1005.
27. Nairi-Saleh B, Dunham PM, McAlreese PM, Macgregor D, Murphy E, et al. (2005) Global regulation of *Staphylococcus aureus* genes by Rot. *J Bacteriol* 185: 610–619.
28. Sharma CM, Duriville F, Hastings TH, Vogel J (2007) A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev* 21: 2804–2817.
29. Papenfuss K, Vogel J (2009) Multiple target regulation by small noncoding RNAs represses gene expression at the post-transcriptional level. *Res Microbiol* 160: 270–287.
30. Papenfuss K, Pfeiffer V, Lucchini S, Senowitz A, Hinton JC, et al. (2008) Systematic deletion of *Stenotrophomonas maltophilia* small RNA genes identifies CyaR, a conserved CRP-dependent riboregulator of OmpX synthesis. *Mol Microbiol* 68: 890–908.
31. De Lay N, Guterman S (2009) The CyaR-activated small noncoding regulatory RNA CyaR (RyfE) links nutritional status to group behavior. *J Bacteriol* 191: 461–476.
32. Geismann T, Chevillon C, Gros MJ, Boisset S, Focher P, et al. (2005) A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation. *Nucleic Acids Res* 33: 7239–7257.
33. Duriville F, Unson C, Vogel J, Wagner EG (2007) An antisense RNA inhibits translation by competing with standard ribosomes. *Mol Cell* 26: 381–392.
34. Boisset M, Sharma CM, Mills F, Nerlsson KH, Vogel J (2008) Small RNA binding to 5' mRNA coding region inhibits translational initiation. *Mol Cell* 32: 827–837.
35. Pfeiffer V, Papenfuss K, Lucchini S, Hinton JC, Vogel J (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* 16: 840–846.
36. Opdyke JA, Kang JG, Starr G (2004) GafY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J Bacteriol* 186: 6008–6015.
37. Janssen L, Arvidson S (1996) The role of the *Alj* gene (*AljI*) in the regulation of virulence genes by the accessory gene regulator (*agr*) in *Staphylococcus aureus*. *EMBO J* 15: 1391–1399.
38. Boisset Y, Kelle FA, Roudy P, Lass G, Etienne J, et al. (2008) Probing the structure of RNAMIII, the *Staphylococcus aureus* *agr* regulatory RNA, and identification of the RNA domain involved in repression of protein A expression. *RNA* 14: 668–679.
39. Sambrook J, Fritsch E, Maniatis T (1989) *Molecular Cloning: a laboratory manual*. Cold Spring Harbor, New York, USA: Cold Spring Harbor Press.
40. Prouss G, Wagner EG, Nordstrom K (1988) Control of replication of plasmid R1: kinetics of *in situ* interaction between the antisense RNA, CopA, and its target, CopY. *EMBO J* 7: 3279–3288.
41. Dai R, Luederach A, Pothman SM, Hirschling D, Altmann RB (2003) SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. *RNA* 11: 344–354.
42. Focher P, Chevillon C, Yungues G, Yungues M, Roudy P, et al. (2009) Ribosomal initiation complexes probed by transcripting and effect of transacting translational regulators in bacteria. *Methods Mol Biol* 540: 247–263.
43. Manic C, Westhof E (1998) MANIP: an interactive tool for modeling RNA. *J Mol Graph Model* 16: 197–205, 253–7.
44. Manic C, Westhof E (2005) Modeling the architecture of structured RNAs within a molecular and hierarchical framework. *Handbook of RNA Biochemistry*, pp 336–345.

Post-Transcriptional Regulation of *coa* mRNA

45. Kreiswirth BN, Lofthol S, Beley MJ, O'Reilly M, Schlevert PM, et al. (1993) The toxic shock syndrome exotoxin structural gene is not detectably transcribed by a prophage. *Nature* 365: 709–712.
46. Pong III, Nostik KP, Kreiswirth B, Korolik J, Schlevert P (1998) Cloning, characterization, and sequencing of an accessory gene regulator (*ag*) in *Aglycyrococcus* sp. *J Bacteriol* 170: 4365–4372.
47. Horinouchi S, Weidman B (1982) Nucleotide sequence and functional map of pE198, a plasmid that specifies inducible resistance to tetracycline, lincomycin, and streptogramin type II antibiotics. *J Bacteriol* 150: 804–815.

### III. Communications scientifiques

**RNA meeting 2012**, Ann Arbor, USA

**Communication orale** : A  $\sigma^B$  dependent regulatory RNA modulates biofilm formation and capsule formation in *Staphylococcus aureus*

**C. Romilly\***, C. Lays\*, E. Lioliou, F. Vincent, F. Vandensech, P. Romby, S. Boisset<sup>§</sup> and T. Geissmann<sup>§</sup> \*equal first authorship; <sup>§</sup>equal senior authorship

**ERA-NET PathoGenoMics 2012**, Tenerife (Espagne)

**Poster** : A  $\sigma^B$  dependent regulatory RNA modulates biofilm formation and capsule formation in *Staphylococcus aureus*

**C. Romilly\***, C. Lays\*, E. Lioliou, F. Vincent, F. Vandensech, P. Romby, S. Boisset<sup>§</sup> and T. Geissmann<sup>§</sup> \*equal first authorship; <sup>§</sup>equal senior authorship

**RNA meeting 2011**, Kyoto, (Japon)

**Poster** : *Staphylococcus aureus* Endoribonuclease III and its diverse RNA-mRNA substrates involved in gene regulation.

**C. Romilly<sup>1</sup>**, C. Chevalier<sup>2</sup>, E. Lioliou, B. masquida, T. Geissmann<sup>3</sup>, F. Vandenesch<sup>3</sup> and P. Romby<sup>1</sup>

**Journée DGA 2011**, Paris

**Poster** : L'endoribonuclease III de *Staphylococcus aureus* et ses divers substrats : un nouvel outil pour détecter les interactions de type boucle-boucle entre deux ARN.

**C. Romilly**, C. Chevalier, E. Lioliou, B. Masquida, T. Geissmann, F. Vandenesch\* et P. Romby

**All roads lead to RNA : news and perspectives**, Ecole doctorale de Strasbourg

**Communication orale** : An example of a multifunctional regulatory RNA in *Staphylococcus aureus*

**C. Romilly**

**sifrARN 2010**, Dourdan (France)

**Poster** : L'endoribonuclease III de *Staphylococcus aureus* et ses divers substrats : un nouvel outil pour détecter les interactions de type boucle-boucle entre deux ARN.

**C. Romilly**, C. Chevalier, E. Lioliou, B. Masquida, T. Geissmann, F. Vandenesch\* et P. Romby

**GCOE Summer school 2010**, Tokyo (Japon).

**Communication orale** : « The role of RNA in gene regulation of bacterial pathogen : adaptation of virulence and metabolism ».

**C. Romilly**

**FJC 2009, SFBBM** Nancy (France)

**Poster** : L'endoribonuclease III de *Staphylococcus aureus* et ses divers substrats : un nouvel outil pour détecter les interactions de type boucle-boucle entre deux ARN.

**C. Romilly**, C. Chevalier, E. Lioliou, B. Masquida, T. Geissmann, F. Vandenesch\* et P. Romby

## IV. Bibliographie

Ades SE. Regulation by destruction: design of the sigmaE envelope stress response. *Curr Opin Microbiol.* 2008 Dec; 11 (6) :535-40. PubMed PMID:18983936.

Aiba H. Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr Opin Microbiol.* 2007 Apr; 10 (2) :134-9. PubMed PMID:17383928.

Alix E, Blanc-Potard AB. MgtC: a key player in intramacrophage survival. *Trends Microbiol.* 2007 Jun; 15 (6) :252-6. PubMed PMID:17416526.

Anderson I, Göker M, Nolan M, Lucas S, Hammon N, Deshpande S, Cheng JF, Tapia R, Han C, Goodwin L, Pitluck S, Huntemann M, Liolios K, Ivanova N, Pagani I, Mavromatis K, Ovchinnikova G, Pati A, Chen A, Palaniappan K, Land M, Hauser L, Brambilla EM, Huber H, Yasawong M, Rohde M, Spring S, Abt B, Sikorski J, Wirth R, Detter JC, Woyke T, Bristow J, Eisen JA, Markowitz V, Hugenholtz P, Kyrpides NC, Klenk HP, Lapidus A. Complete genome sequence of the hyperthermophilic chemolithoautotroph *Pyrolobus fumarii* type strain (1A). *Stand Genomic Sci.* 2011 Jul 1; 4 (3) :381-92. PubMed PMID:21886865; PubMed Central PMCID: PMC3156397.

André G, Even S, Putzer H, Burguière P, Croux C, Danchin A, Martin-Verstraete I, Soutourina O. S-box and T-box riboswitches and antisense RNA control a sulfur metabolic operon of *Clostridium acetobutylicum*. *Nucleic Acids Res.* 2008 Oct; 36 (18) :5955-69. PubMed PMID:18812398; PubMed Central PMCID: PMC2566862.

Archer NK, Mazaitis MJ, Costerton JW, Leid JG, Powers ME, Shirliff ME. *Staphylococcus aureus* biofilms: properties, regulation, and roles in human disease. *Virulence.* 2011 Sep-Oct; 2 (5) :445-59. PubMed PMID:21921685; PubMed Central PMCID: PMC3322633.

Argaman L, Hershberg R, Vogel J, Bejerano G, Wagner EG, Margalit H, Altuvia S. Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr Biol.* 2001 Jun 26; 11 (12) :941-50. PubMed PMID:11448770.

Argaman L, Elgrably-Weiss M, Hershko T, Vogel J, Altuvia S. RelA protein stimulates the activity of RyhB small RNA by acting on RNA-binding protein Hfq. *Proc Natl Acad Sci U S A.* 2012 Mar 20; 109 (12) :4621-6. PubMed PMID:22393021; PubMed Central PMCID: PMC3311362.

Arraiano CM, Andrade JM, Domingues S, Guinote IB, Malecki M, Matos RG, Moreira RN, Pobre V, Reis FP, Saramago M, Silva IJ, Viegas SC. The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol Rev.* 2010 Sep; 34 (5) :883-923. PubMed PMID:20659169.

Awano N, Rajagopal V, Arbing M, Patel S, Hunt J, Inouye M, Phadtare S. *Escherichia coli* RNase R has dual activities, helicase and RNase. *J Bacteriol.* 2010 Mar; 192 (5) :1344-52. PubMed PMID:20023028; PubMed Central PMCID: PMC2820863.

Babitzke P, Romeo T. CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Curr Opin Microbiol.* 2007 Apr; 10 (2) :156-63. PubMed PMID:17383221.

Babitzke P, Granger L, Olszewski J, Kushner SR. Analysis of mRNA decay and rRNA processing in *Escherichia coli* multiple mutants carrying a deletion in RNase III. *J Bacteriol.* 1993 Jan; 175 (1) :229-39. PubMed PMID:8416898; PubMed Central PMCID: PMC196118.

Ballal A, Ray B, Manna AC. sarZ, a sarA family gene, is transcriptionally activated by MgrA and is involved in the regulation of genes encoding exoproteins in *Staphylococcus aureus*. *J Bacteriol.* 2009 Mar; 191 (5) :1656-65. PubMed PMID:19103928; PubMed Central PMCID: PMC2648185.



Bardwell JC, Régnier P, Chen SM, Nakamura Y, Grunberg-Manago M, Court DL. Autoregulation of RNase III operon by mRNA processing. *EMBO J.* 1989 Nov; 8 (11) :3401-7. PubMed PMID:2583104; PubMed Central PMCID: PMC401487. Barrick JE, Breaker RR. The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol.* 2007; 8 (11) :R239. PubMed PMID:17997835; PubMed Central PMCID: PMC2258182.

Behammer W, Shao Z, Mages W, Rachel R, Stetter KO, Schmitt R. Flagellar structure and hyperthermophily: analysis of a single flagellin gene and its product in *Aquifex pyrophilus*. *J Bacteriol.* 1995 Nov; 177 (22) :6630-7. PubMed PMID:7592443; PubMed Central PMCID: PMC177518.

Beich-Frandsen M, Vecerek B, Konarev PV, Sjöblom B, Kloiber K, Hämmerle H, Rajkowitsch L, Miles AJ, Kontaxis G, Wallace BA, Svergun DI, Konrat R, Bläsi U, Djinovic-Carugo K. Structural insights into the dynamics and function of the C-terminus of the *E coli* RNA chaperone Hfq. *Nucleic Acids Res.* 2011 Jun; 39 (11) :4900-15. PubMed PMID:21330354; PubMed Central PMCID: PMC3113564.

Beisel CL, Storz G. Base pairing small RNAs and their roles in global regulatory networks. *FEMS Microbiol Rev.* 2010 Sep; 34 (5) :866-82. PubMed PMID:20662934; PubMed Central PMCID: PMC2920360.

Beisel CL, Updegrove TB, Janson BJ, Storz G. Multiple factors dictate target selection by Hfq-binding small RNAs. *EMBO J.* 2012 Mar 2; 31 (8) :1961-74. PubMed PMID:22388518; PubMed Central PMCID: PMC3343335.

Bekeredjian-Ding I, Inamura S, Giese T, Moll H, Endres S, Sing A, Zähringer U, Hartmann G. *Staphylococcus aureus* protein A triggers T cell-independent B cell proliferation by sensitizing B cells for TLR2 ligands. *J Immunol.* 2007 Mar 1; 178 (5) :2803-12. PubMed PMID:17312124.

Benito Y, Kolb FA, Romby P, Lina G, Etienne J, Vandenesch F. Probing the structure of RNAIII, the *Staphylococcus aureus* agr regulatory RNA, and identification of the RNA domain involved in repression of protein A expression. *RNA.* 2000 May; 6 (5) :668-79. PubMed PMID:10836788; PubMed Central PMCID: PMC1369947.

Bhakdi S, Tranum-Jensen J. Alpha-toxin of *Staphylococcus aureus*. *Microbiol Rev.* 1991 Dec; 55 (4) :733-51. PubMed PMID:1779933; PubMed Central PMCID: PMC372845.

Bischoff M, Dunman P, Kormanec J, Macapagal D, Murphy E, Mounts W, Berger-Bächi B, Projan S. Microarray-based analysis of the *Staphylococcus aureus* sigmaB regulon. *J Bacteriol.* 2004 Jul; 186 (13) :4085-99. PubMed PMID:15205410; PubMed Central PMCID: PMC421609.

Blaszczak J, Gan J, Tropea JE, Court DL, Waugh DS, Ji X. Noncatalytic assembly of ribonuclease III with doublestranded RNA. *Structure.* 2004 Mar; 12 (3) :457-66. PubMed PMID:15016361.

Bohn C, Rigoulay C, Boulloc P. No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. *BMC Microbiol.* 2007 Feb 9; 7:10. PubMed PMID:17291347; PubMed Central PMCID: PMC1800855.

Boisset S, Geissmann T, Huntzinger E, Fechter P, Bendridi N, Possedko M, Chevalier C, Helfer AC, Benito Y, Jacquier A, Gaspin C, Vandenesch F, Romby P. *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev.* 2007 Jun 1; 21 (11) :1353-66. PubMed PMID:17545468; PubMed Central PMCID: PMC1877748.

Bokarewa MI, Jin T, Tarkowski A. *Staphylococcus aureus*: Staphylokinase. *Int J Biochem Cell Biol.* 2006; 38 (4) :504-9. PubMed PMID:16111912.

Bordi C, Lamy MC, Ventre I, Termine E, Hachani A, Fillet S, Roche B, Bleves S, Méjean V, Lazdunski A, Filloux A. Regulatory RNAs and the HptB/RetS signalling pathways fine-tune *Pseudomonas aeruginosa* pathogenesis. *Mol Microbiol.* 2010 Jun; 76 (6) :1427-43. PubMed PMID:20398205; PubMed Central PMCID: PMC2904497.

Brantl S. Regulatory mechanisms employed by cis-encoded antisense RNAs. *Curr Opin Microbiol.* 2007 Apr; 10 (2) :102-9. PubMed PMID:17387036.

Breaker RR. Riboswitches and the RNA world. *Cold Spring Harb Perspect Biol.* 2012 Feb 1; 4 (2) PubMed PMID:21106649.

Brenchley JM, Douek DC. Microbial translocation across the GI tract. *Annu Rev Immunol.* 2012; 30:149-73. PubMed PMID:22224779.

Brennan RG, Link TM. Hfq structure, function and ligand binding. *Curr Opin Microbiol.* 2007 Apr; 10 (2) :125-33. PubMed PMID:17395525.

Brunskill EW, Bayles KW. Identification and molecular characterization of a putative regulatory locus that affects autolysis in *Staphylococcus aureus*. *J Bacteriol.* 1996 Feb; 178 (3) :611-8. PubMed PMID:8550490; PubMed Central PMCID: PMC177702.

Bukowski M, Rojowska A, Wladyka B. Prokaryotic toxin-antitoxin systems--the role in bacterial physiology and application in molecular biology. *Acta Biochim Pol.* 2011; 58 (1) :1-9. PubMed PMID:21394325.

Busch A, Richter AS, Backofen R. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics.* 2008 Dec 15; 24 (24) :2849-56. PubMed PMID:18940824; PubMed Central PMCID: PMC2639303.

Carpousis AJ. The RNA degradosome of *Escherichia coli*: an mRNA-degrading machine assembled on RNase E. *Annu Rev Microbiol.* 2007; 61:71-87. PubMed PMID:17447862.

Cavanagh AT, Sperger JM, Wassarman KM. Regulation of 6S RNA by pRNA synthesis is required for efficient recovery from stationary phase in *E coli* and *B subtilis*. *Nucleic Acids Res.* 2012 Mar; 40 (5) :2234-46. PubMed PMID:22102588; PubMed Central PMCID: PMC3299989.

Centers for Disease Control and Prevention (CDC). *Staphylococcus aureus* resistant to vancomycin--United States, 2002. *MMWR Morb Mortal Wkly Rep.* 2002 Jul 5; 51 (26) :565-7. PubMed PMID:12139181.

Chabelskaya S, Gaillot O, Felden B. A *Staphylococcus aureus* small RNA is required for bacterial virulence and regulates the expression of an immune-evasion molecule. *PLoS Pathog.* 2010 Jun 3; 6 (6) :e1000927. PubMed PMID:20532214; PubMed Central PMCID: PMC2880579.

Chakrabarti SK, Misra TK. SarA represses agr operon expression in a purified in vitro *Staphylococcus aureus* transcription system. *J Bacteriol.* 2000 Oct; 182 (20) :5893-7. PubMed PMID:11004191; PubMed Central PMCID: PMC94714.

Chao Y, Vogel J. The role of Hfq in bacterial pathogens. *Curr Opin Microbiol.* 2010 Feb; 13 (1) :24-33. PubMed PMID:20080057.

Chavakis T, Hussain M, Kanse SM, Peters G, Bretzel RG, Flock JI, Herrmann M, Preissner KT. *Staphylococcus aureus* extracellular adherence protein serves as anti-inflammatory factor by inhibiting the recruitment of host leukocytes. *Nat Med.* 2002 Jul; 8 (7) :687-93. PubMed PMID:12091905.

Chen C, Deutscher MP. Elevation of RNase R in response to multiple stress conditions. *J Biol Chem.* 2005 Oct 14; 280 (41) :34393-6. PubMed PMID:16135521.

Chen S, Zhang A, Blyn LB, Storz G. MicC, a second small-RNA regulator of Omp protein expression in *Escherichia coli*. *J Bacteriol.* 2004 Oct; 186 (20) :6689-97. PubMed PMID:15466019; PubMed Central PMCID: PMC522180.

Cheung AL, Bayer MG, Heinrichs JH. Genetic determinants necessary for transcription of RNAII and RNAIII in the *agr* locus of *Staphylococcus aureus*. *J Bacteriol.* 1997 Jun; 179 (12) :3963-71. PubMed PMID:9190813; PubMed Central PMCID: PMC179206.

Cheung AL, Bayer AS, Zhang G, Gresham H, Xiong YQ. Regulation of virulence determinants in vitro and in vivo in *Staphylococcus aureus*. *FEMS Immunol Med Microbiol.* 2004 Jan 15; 40 (1) :1-9. PubMed PMID:14734180.

Cheung AL, Nishina K, Manna AC. SarA of *Staphylococcus aureus* binds to the *sarA* promoter to regulate gene expression. *J Bacteriol.* 2008 Mar; 190 (6) :2239-43. PubMed PMID:18178734; PubMed Central PMCID: PMC2258893.

Cheung AL, Nishina KA, Trottonda MP, Tamber S. The SarA protein family of *Staphylococcus aureus*. *Int J Biochem Cell Biol.* 2008; 40 (3) :355-61. PubMed PMID:18083623; PubMed Central PMCID: PMC2274939.

Chevalier C, Boisset S, Romilly C, Masquida B, Fechter P, Geissmann T, Vandenesch F, Romby P. *Staphylococcus aureus* RNAIII binds to two distant regions of *coa* mRNA to arrest translation and promote mRNA degradation. *PLoS Pathog.* 2010 Mar 12; 6 (3) :e1000809. PubMed PMID:20300607; PubMed Central PMCID: PMC2837412.

Chien Y, Manna AC, Cheung AL. SarA level is a determinant of *agr* activation in *Staphylococcus aureus*. *Mol Microbiol.* 1998 Dec; 30 (5) :991-1001. PubMed PMID:9988476.

Chow J, Lee SM, Shen Y, Khosravi A, Mazmanian SK. Host-bacterial symbiosis in health and disease. *Adv Immunol.* 2010; 107:243-74. PubMed PMID:21034976; PubMed Central PMCID: PMC3152488.

Chowdhury S, Maris C, Allain FH, Narberhaus F. Molecular basis for temperature sensing by an RNA thermometer. *EMBO J.* 2006 Jun 7; 25 (11) :2487-97. PubMed PMID:16710302; PubMed Central PMCID: PMC1478195.

Chung H, Kasper DL. Microbiota-stimulated immune mechanisms to maintain gut homeostasis. *Curr Opin Immunol.* 2010 Aug; 22 (4) :455-60. PubMed PMID:20656465.

Clatworthy AE, Pierson E, Hung DT. Targeting virulence: a new paradigm for antimicrobial therapy. *Nat Chem Biol.* 2007 Sep; 3 (9) :541-8. PubMed PMID:17710100.

Cogen AL, Nizet V, Gallo RL. Skin microbiota: a source of disease or defence?. *Br J Dermatol.* 2008 Mar; 158 (3) :442-55. PubMed PMID:18275522; PubMed Central PMCID: PMC2746716.

Collins MD, Lawson PA, Labrenz M, Tindall BJ, Weiss N, Hirsch P. *Nesterenkonia lacusekhoensis* sp nov, isolated from hypersaline Ekho Lake, East Antarctica, and emended description of the genus *Nesterenkonia*. *Int J Syst Evol Microbiol.* 2002 Jul; 52 (Pt 4) :1145-50. PubMed PMID:12148619.

Commichau FM, Rothe FM, Herzberg C, Wagner E, Hellwig D, Lehnik-Habrink M, Hammer E, Völker U, Stülke J. Novel activities of glycolytic enzymes in *Bacillus subtilis*: interactions with essential proteins involved in mRNA processing. *Mol Cell Proteomics.* 2009 Jun; 8 (6) :1350-60. PubMed PMID:19193632; PubMed Central PMCID: PMC2690492.

Condon C, Putzer H. The phylogenetic distribution of bacterial ribonucleases. *Nucleic Acids Res.* 2002 Dec 15; 30 (24) :5339-46. PubMed PMID:12490701; PubMed Central PMCID: PMC140075.

Condon C. Maturation and degradation of RNA in bacteria. *Curr Opin Microbiol.* 2007 Jun; 10 (3) :271-8. PubMed PMID:17560162.

Coppins RL, Hall KB, Groisman EA. The intricate world of riboswitches. *Curr Opin Microbiol.* 2007 Apr; 10 (2) :176-81. PubMed PMID:17383225; PubMed Central PMCID: PMC1894890.

Cunnion KM, Hair PS, Buescher ES. Cleavage of complement C3b to iC3b on the surface of *Staphylococcus aureus* is mediated by serum complement factor I. *Infect Immun.* 2004 May; 72 (5) :2858-63. PubMed PMID:15102797; PubMed Central PMCID: PMC387901.

Cunnion KM, Lee JC, Frank MM. Capsule production and growth phase influence binding of complement to *Staphylococcus aureus*. *Infect Immun.* 2001 Nov; 69 (11) :6796-803. PubMed PMID:11598052; PubMed Central PMCID: PMC100057.

Dam P, Olman V, Harris K, Su Z, Xu Y. Operon prediction using both genome-specific and general genomic information. *Nucleic Acids Res.* 2007; 35 (1) :288-98. PubMed PMID:17170009; PubMed Central PMCID: PMC1802555.

Darfeuille F, Unoson C, Vogel J, Wagner EG. An antisense RNA inhibits translation by competing with standby ribosomes. *Mol Cell.* 2007 May 11; 26 (3) :381-92. PubMed PMID:17499044.

de Hoon MJ, Makita Y, Nakai K, Miyano S. Prediction of transcriptional terminators in *Bacillus subtilis* and related species. *PLoS Comput Biol.* 2005 Aug; 1 (3) :e25. PubMed PMID:16110342; PubMed Central PMCID: PMC1187862.

Deana A, Celesnik H, Belasco JG. The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature.* 2008 Jan 17; 451 (7176) :355-8. PubMed PMID:18202662.

Deguchi K, Kawamura S, Sugita R, Fujimaki Y, Ohsawa H. [Sensitivity to second-generation cephem agents of clinically isolated strains of bacteria in otorhinolaryngological field (author's transl)]. *Jpn J Antibiot.* 1982 Mar; 35 (3) :812-20. PubMed PMID:6285017.

Deikus G, Babitzke P, Bechhofer DH. Recycling of a regulatory protein by degradation of the RNA to which it binds. *Proc Natl Acad Sci U S A.* 2004 Mar 2; 101 (9) :2747-51. PubMed PMID:14976255; PubMed Central PMCID: PMC365692.

Delaune A, Poupel O, Mallet A, Coic YM, Msadek T, Dubrac S. Peptidoglycan crosslinking relaxation plays an important role in *Staphylococcus aureus* WalkR-dependent cell viability. *PLoS One.* 2011 Feb 28; 6 (2) :e17054. PubMed PMID:21386961; PubMed Central PMCID: PMC3046168.

Delumeau O, Chen CC, Murray JW, Yudkin MD, Lewis RJ. High-molecular-weight complexes of RsbR and paralogues in the environmental signaling pathway of *Bacillus subtilis*. *J Bacteriol.* 2006 Nov; 188 (22) :7885-92. PubMed PMID:16963570; PubMed Central PMCID: PMC1636303.

Didier JP, Cozzone AJ, Duclos B. Phosphorylation of the virulence regulator SarA modulates its ability to bind DNA in *Staphylococcus aureus*. *FEMS Microbiol Lett.* 2010 May; 306 (1) :30-6. PubMed PMID:20337713. Do CB, Woods DA, Batzoglou S. CONTRAfold: RNA secondary structure prediction without physics-based models. *Bioinformatics.* 2006 Jul 15; 22 (14) :e90-8. PubMed PMID:16873527.

Dubrac S, Boneca IG, Poupel O, Msadek T. New insights into the Walk/WalR (YycG/YycF) essential signal transduction pathway reveal a major role in controlling cell wall metabolism and biofilm formation in *Staphylococcus aureus*. *J Bacteriol.* 2007 Nov; 189 (22) :8257-69. PubMed PMID:17827301; PubMed Central PMCID: PMC2168699.

Dumitrescu O, Dauwalder O, Boisset S, Reverdy M<sup>É</sup>, Tristan A, Vandenesch F. [Staphylococcus aureus resistance to antibiotics: key points in 2010]. *Med Sci (Paris).* 2010 Nov; 26 (11) :943-9. PubMed PMID:21106176.

Durand S, Gilet L, Bessièrès P, Nicolas P, Condon C. Three essential ribonucleases-RNase Y, J1, and III-control the abundance of a majority of *Bacillus subtilis* mRNAs. *PLoS Genet.* 2012; 8 (3) :e1002520. PubMed PMID:22412379; PubMed Central PMCID: PMC3297567.

Eggenhofer F, Tafer H, Stadler PF, Hofacker IL. RNApredator: fast accessibility-based prediction of sRNA targets. *Nucleic Acids Res.* 2011 Jul; 39 (Web Server issue) :W149-54. PubMed PMID:21672960; PubMed Central PMCID: PMC3125805.

Even S, Pellegrini O, Zig L, Labas V, Vinh J, Bréchemmier-Baey D, Putzer H. Ribonucleases J1 and J2: two novel endoribonucleases in *Bsubtilis* with functional homology to *Ecoli* RNase E. *Nucleic Acids Res.* 2005; 33 (7) :2141-52. PubMed PMID:15831787; PubMed Central PMCID: PMC1079966.

Falord M, Karimova G, Hiron A, Msadek T. GraXSR proteins interact with the *VraFG* ABC transporter to form a five-component system required for cationic antimicrobial peptide sensing and resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother.* 2012 Feb; 56 (2) :1047-58. PubMed PMID:22123691; PubMed Central PMCID: PMC3264281.

Felden B, Vandenesch F, Bouloc P, Romby P. The *Staphylococcus aureus* RNome and its commitment to virulence. *PLoS Pathog.* 2011 Mar; 7 (3) :e1002006. PubMed PMID:21423670; PubMed Central PMCID: PMC3053349.

Fender A, Elf J, Hampel K, Zimmermann B, Wagner EG. RNAs actively cycle on the Sm-like protein Hfq. *Genes Dev.* 2010 Dec 1; 24 (23) :2621-6. PubMed PMID:21123649; PubMed Central PMCID: PMC2994036.

Fierer N, Hamady M, Lauber CL, Knight R. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proc Natl Acad Sci U S A.* 2008 Nov 18; 105 (46) :17994-9. PubMed PMID:19004758; PubMed Central PMCID: PMC2584711.

Flint HJ, Bayer EA, Rincon MT, Lamed R, White BA. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat Rev Microbiol.* 2008 Feb; 6 (2) :121-31. PubMed PMID:18180751.

Foster TJ. Immune evasion by staphylococci. *Nat Rev Microbiol.* 2005 Dec; 3 (12) :948-58. PubMed PMID:16322743. Fournier B, Hooper DC. A new two-component regulatory system involved in adhesion, autolysis, and extracellular proteolytic activity of *Staphylococcus aureus*. *J Bacteriol.* 2000 Jul; 182 (14) :3955-64. PubMed PMID:10869073; PubMed Central PMCID: PMC94580.

Fozo EM, Makarova KS, Shabalina SA, Yutin N, Koonin EV, Storz G. Abundance of type I toxin-antitoxin systems in bacteria: searches for new candidates and discovery of novel families. *Nucleic Acids Res.* 2010 Jun; 38 (11) :3743-59. PubMed PMID:20156992; PubMed Central PMCID: PMC2887945.

Fozo EM, Kawano M, Fontaine F, Kaya Y, Mendieta KS, Jones KL, Ocampo A, Rudd KE, Storz G. Repression of small toxic protein synthesis by the *Sib* and *OhcC* small RNAs. *Mol Microbiol.* 2008 Dec; 70 (5) :1076-93. PubMed PMID:18710431; PubMed Central PMCID: PMC2597788.

Fredricks DN. Microbial ecology of human skin in health and disease. *J Investig Dermatol Symp Proc.* 2001 Dec; 6 (3) :167-9. PubMed PMID:11924822.

Fukuda S, Toh H, Hase K, Oshima K, Nakanishi Y, Yoshimura K, Tobe T, Clarke JM, Topping DL, Suzuki T, Taylor TD, Itoh K, Kikuchi J, Morita H, Hattori M, Ohno H. Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature.* 2011 Jan 27; 469 (7331) :543-7. PubMed PMID:21270894.

Gan J, Tropea JE, Austin BP, Court DL, Waugh DS, Ji X. Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. *Cell.* 2006 Jan 27; 124 (2) :355-66. PubMed PMID:16439209.

Gan J, Shaw G, Tropea JE, Waugh DS, Court DL, Ji X. A stepwise model for double-stranded RNA processing by ribonuclease III. *Mol Microbiol.* 2008 Jan; 67 (1) :143-54. PubMed PMID:18047582.

Geiger T, Goerke C, Mainiero M, Kraus D, Wolz C. The virulence regulator Sae of *Staphylococcus aureus*: promoter activities and response to phagocytosis-related signals. *J Bacteriol.* 2008 May; 190 (10) :3419-28. PubMed PMID:18344360; PubMed Central PMCID: PMC2395011.

Geisinger E, Adhikari RP, Jin R, Ross HF, Novick RP. Inhibition of rot translation by RNAll, a key feature of agr function. *Mol Microbiol.* 2006 Aug; 61 (4) :1038-48. PubMed PMID:16879652.

Genestier AL, Michallet MC, Prévost G, Bellot G, Chalabreysse L, Peyrol S, Thivolet F, Etienne J, Lina G, Vallette FM, Vandenesch F, Genestier L. *Staphylococcus aureus* Panton-Valentine leukocidin directly targets mitochondria and induces Bax-independent apoptosis of human neutrophils. *J Clin Invest.* 2005 Nov; 115 (11) :3117-27. PubMed PMID:16276417; PubMed Central PMCID: PMC1265849.

Geng H, Zuber P, Nakano MM. Regulation of respiratory genes by ResD-ResE signal transduction system in *Bacillus subtilis*. *Methods Enzymol.* 2007; 422:448-64. PubMed PMID:17628154.

Georg J, Hess WR. cis-antisense RNA, another level of gene regulation in bacteria. *Microbiol Mol Biol Rev.* 2011 Jun; 75 (2) :286-300. PubMed PMID:21646430; PubMed Central PMCID: PMC3122628.

Georg J, Voss B, Scholz I, Mitschke J, Wilde A, Hess WR. Evidence for a major role of antisense RNAs in cyanobacterial gene regulation. *Mol Syst Biol.* 2009; 5:305. PubMed PMID:19756044; PubMed Central PMCID: PMC2758717.

Gertz S, Engelmann S, Schmid R, Ziebandt AK, Tischler K, Scharf C, Hacker J, Hecker M. Characterization of the sigma(B) regulon in *Staphylococcus aureus*. *J Bacteriol.* 2000 Dec; 182 (24) :6983-91. PubMed PMID:11092859; PubMed Central PMCID: PMC94824.

Ghosh S, Deutscher MP. Oligoribonuclease is an essential component of the mRNA decay pathway. *Proc Natl Acad Sci U S A.* 1999 Apr 13; 96 (8) :4372-7. PubMed PMID:10200269; PubMed Central PMCID: PMC16339.

Giachino P, Engelmann S, Bischoff M. Sigma(B) activity depends on RsbU in *Staphylococcus aureus*. *J Bacteriol.* 2001 Mar; 183 (6) :1843-52. PubMed PMID:11222581; PubMed Central PMCID: PMC95078.

Giraud AT, Cheung AL, Nagel R. The sae locus of *Staphylococcus aureus* controls exoprotein synthesis at the transcriptional level. *Arch Microbiol.* 1997 Jul; 168 (1) :53-8. PubMed PMID:9211714.

Gogol EB, Rhodius VA, Papenfort K, Vogel J, Gross CA. Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon. *Proc Natl Acad Sci U S A.* 2011 Aug 2; 108 (31) :12875-80. PubMed PMID:21768388; PubMed Central PMCID: PMC3150882.

Goldman SR, Sharp JS, Vvedenskaya IO, Livny J, Dove SL, Nickels BE. NanoRNAs prime transcription initiation in vivo. *Mol Cell.* 2011 Jun 24; 42 (6) :817-25. PubMed PMID:21700226; PubMed Central PMCID: PMC3130991.

Gómez MI, Lee A, Reddy B, Muir A, Soong G, Pitt A, Cheung A, Prince A. *Staphylococcus aureus* protein A induces airway epithelial inflammatory responses by activating TNFR1. *Nat Med.* 2004 Aug; 10 (8) :842-8. PubMed PMID:15247912.

Gordon RJ, Lowy FD. Pathogenesis of methicillin-resistant *Staphylococcus aureus* infection. *Clin Infect Dis.* 2008 Jun 1; 46 Suppl 5:S350-9. PubMed PMID:18462090; PubMed Central PMCID: PMC2474459. Görke B, Vogel J. Noncoding RNA control of the making and breaking of sugars. *Genes Dev.* 2008 Nov 1; 22 (21) :2914-25. PubMed PMID:18981470.

Gorwitz RJ, Kruszon-Moran D, McAllister SK, McQuillan G, McDougal LK, Fosheim GE, Jensen BJ, Killgore G, Tenover FC, Kuehnert MJ. Changes in the prevalence of nasal colonization with

Staphylococcus aureus in the United States, 2001-2004. *J Infect Dis.* 2008 May 1; 197 (9) :1226-34. PubMed PMID:18422434.

Gottesman S, Storz G. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harb Perspect Biol.* 2011 Dec 1; 3 (12) PubMed PMID:20980440.

Green NJ, Grundy FJ, Henkin TM. The T box mechanism: tRNA as a regulatory molecule. *FEBS Lett.* 2010 Jan 21; 584 (2) :318-24. PubMed PMID:19932103; PubMed Central PMCID: PMC2794906.

Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. *Cell.* 2010 Mar 19; 140 (6) :883-99. PubMed PMID:20303878; PubMed Central PMCID: PMC2866629.

Grundmeier M, Hussain M, Becker P, Heilmann C, Peters G, Sinha B. Truncation of fibronectin-binding proteins in *Staphylococcus aureus* strain Newman leads to deficient adherence and host cell invasion due to loss of the cell wall anchor function. *Infect Immun.* 2004 Dec; 72 (12) :7155-63. PubMed PMID:15557640; PubMed Central PMCID: PMC529102.

Grundy FJ, Yousef MR, Henkin TM. Monitoring uncharged tRNA during transcription of the *Bacillus subtilis* glyQS gene. *J Mol Biol.* 2005 Feb 11; 346 (1) :73-81. PubMed PMID:15663928.

Güell M, van Noort V, Yus E, Chen WH, Leigh-Bell J, Michalodimitrakis K, Yamada T, Arumugam M, Doerks T, Kühner S, Rode M, Suyama M, Schmidt S, Gavin AC, Bork P, Serrano L. Transcriptome complexity in a genome-reduced bacterium. *Science.* 2009 Nov 27; 326 (5957) :1268-71. PubMed PMID:19965477.

Guillier M, Gottesman S. The 5' end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. *Nucleic Acids Res.* 2008 Dec; 36 (21) :6781-94. PubMed PMID:18953042; PubMed Central PMCID: PMC2588501.

Guillier M, Gottesman S, Storz G. Modulating the outer membrane with small RNAs. *Genes Dev.* 2006 Sep 1; 20 (17) :2338-48. PubMed PMID:16951250.

Gustafsson E, Oscarsson J. Maximal transcription of *aur* (aureolysin) and *sspA* (serine protease) in *Staphylococcus aureus* requires staphylococcal accessory regulator R (*sarR*) activity. *FEMS Microbiol Lett.* 2008 Jul; 284 (2) :158-64. PubMed PMID:18576947.

Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jungkamp AC, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T. PAR-CLIP--a method to identify transcriptome-wide the binding sites of RNA binding proteins. *J Vis Exp.* 2010 Jul 2; PubMed PMID:20644507; PubMed Central PMCID: PMC3156069.

Hall-Stoodley L, Costerton JW, Stoodley P. Bacterial biofilms: from the natural environment to infectious diseases. *Nat Rev Microbiol.* 2004 Feb; 2 (2) :95-108. PubMed PMID:15040259.

Han J, Lee Y, Yeom KH, Kim YK, Jin H, Kim VN. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev.* 2004 Dec 15; 18 (24) :3016-27. PubMed PMID:15574589; PubMed Central PMCID: PMC535913.

Harraghy N, Hussain M, Haggar A, Chavakis T, Sinha B, Herrmann M, Flock JI. The adhesive and immunomodulating properties of the multifunctional *Staphylococcus aureus* protein Eap. *Microbiology.* 2003 Oct; 149 (Pt 10) :2701-7. PubMed PMID:14523103.

Hartz D, McPheeters DS, Traut R, Gold L. Extension inhibition analysis of translation initiation complexes. *Methods Enzymol.* 1988; 164:419-25. PubMed PMID:2468068.

Hecker M, Pané-Farré J, Völker U. SigB-dependent general stress response in *Bacillus subtilis* and related gram-positive bacteria. *Annu Rev Microbiol.* 2007; 61:215-36. PubMed PMID:18035607.

Herbert S, Ziebandt AK, Ohlsen K, Schäfer T, Hecker M, Albrecht D, Novick R, Götz F. Repair of global regulators in *Staphylococcus aureus* 8325 and comparative analysis with other clinical isolates. *Infect Immun*. 2010 Jun; 78 (6) :2877-89. PubMed PMID:20212089; PubMed Central PMCID: PMC2876537.

Herskovitz MA, Bechhofer DH. Endoribonuclease RNase III is essential in *Bacillus subtilis*. *Mol Microbiol*. 2000 Dec; 38 (5) :1027-33. PubMed PMID:11123676.

Hiramatsu K, Aritaka N, Hanaki H, Kawasaki S, Hosoda Y, Hori S, Fukuchi Y, Kobayashi I. Dissemination in Japanese hospitals of strains of *Staphylococcus aureus* heterogeneously resistant to vancomycin. *Lancet*. 1997 Dec 6; 350 (9092) :1670-3. PubMed PMID:9400512.

Hiron A, Falord M, Valle J, Débarbouillé M, Msadek T. Bacitracin and nisin resistance in *Staphylococcus aureus*: a novel pathway involving the BraS/BraR two-component system (SA2417/SA2418) and both the BraD/BraE and VraD/VraE ABC transporters. *Mol Microbiol*. 2011 Aug; 81 (3) :602-22. PubMed PMID:21696458.

Holmqvist E, Reimegård J, Sterk M, Grantcharova N, Römling U, Wagner EG. Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. *EMBO J*. 2010 Jun 2; 29 (11) :1840-50. PubMed PMID:20407422; PubMed Central PMCID: PMC2885931.

Hopkins JF, Panja S, McNeil SA, Woodson SA. Effect of salt and RNA structure on annealing and strand displacement by Hfq. *Nucleic Acids Res*. 2009 Oct; 37 (18) :6205-13. PubMed PMID:19671524; PubMed Central PMCID: PMC2764445.

Horsburgh MJ, Aish JL, White IJ, Shaw L, Lithgow JK, Foster SJ. sigmaB modulates virulence determinant expression and stress resistance: characterization of a functional rsbU strain derived from *Staphylococcus aureus* 8325-4. *J Bacteriol*. 2002 Oct; 184 (19) :5457-67. PubMed PMID:12218034; PubMed Central PMCID: PMC135357.

Hunter P. Where next for antibiotics? The immune system and the nature of pathogenicity are providing vital clues in the fight against antibiotic-resistant bacteria. *EMBO Rep*. 2012 Aug 1; 13 (8) :680-3. PubMed PMID:22767249; PubMed Central PMCID: PMC3410381.

Huntzinger E, Boisset S, Saveanu C, Benito Y, Geissmann T, Namane A, Lina G, Etienne J, Ehresmann B, Ehresmann C, Jacquier A, Vandenesch F, Romby P. *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate spa gene expression. *EMBO J*. 2005 Feb 23; 24 (4) :824-35. PubMed PMID:15678100; PubMed Central PMCID: PMC549626.

Ibrahim EH, Sherman G, Ward S, Fraser VJ, Kollef MH. The influence of inadequate antimicrobial treatment of bloodstream infections on patient outcomes in the ICU setting. *Chest*. 2000 Jul; 118 (1) :146-55. PubMed PMID:10893372.

Ingavale S, van Wamel W, Luong TT, Lee CY, Cheung AL. Rat/MgrA, a regulator of autolysis, is a regulator of virulence genes in *Staphylococcus aureus*. *Infect Immun*. 2005 Mar; 73 (3) :1423-31. PubMed PMID:15731040; PubMed Central PMCID: PMC1064946.

Ingavale SS, Van Wamel W, Cheung AL. Characterization of RAT, an autolysis regulator in *Staphylococcus aureus*. *Mol Microbiol*. 2003 Jun; 48 (6) :1451-66. PubMed PMID:12791130.

Jahn N, Preis H, Wiedemann C, Brantl S. BsrG/SR4 from *Bacillus subtilis*--the first temperature-dependent type I toxinantitoxin system. *Mol Microbiol*. 2012 Feb; 83 (3) :579-98. PubMed PMID:22229825.

Jeong DW, Cho H, Lee H, Li C, Garza J, Fried M, Bae T. Identification of the P3 promoter and distinct roles of the two promoters of the SaeRS two-component system in *Staphylococcus aureus*. *J*



Bacteriol. 2011 Sep; 193 (18) :4672-84. PubMed PMID:21764914; PubMed Central PMCID: PMC3165640.

Jett BD, Gilmore MS. Internalization of *Staphylococcus aureus* by human corneal epithelial cells: role of bacterial fibronectin-binding protein and host cell factors. *Infect Immun*. 2002 Aug; 70 (8) :4697-700. PubMed PMID:12117986; PubMed Central PMCID: PMC128182.

JEVONS MP, COE AW, PARKER MT. Methicillin resistance in staphylococci. *Lancet*. 1963 Apr 27; 1 (7287) :904-7. PubMed PMID:13957735.

Jiang X, Belasco JG. Catalytic activation of multimeric RNase E and RNase G by 5'-monophosphorylated RNA. *Proc Natl Acad Sci U S A*. 2004 Jun 22; 101 (25) :9211-6. PubMed PMID:15197283; PubMed Central PMCID: PMC438955.

Johansen J, Rasmussen AA, Overgaard M, Valentin-Hansen P. Conserved small non-coding RNAs that belong to the sigmaE regulon: role in down-regulation of outer membrane proteins. *J Mol Biol*. 2006 Nov 17; 364 (1) :1-8. PubMed PMID:17007876.

Johansson J, Mandin P, Renzoni A, Chiaruttini C, Springer M, Cossart P. An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. *Cell*. 2002 Sep 6; 110 (5) :551-61. PubMed PMID:12230973. Jönsson K, Signäs C, Müller HP, Lindberg M. Two different genes encode fibronectin binding proteins in *Staphylococcus aureus* The complete nucleotide sequence and characterization of the second gene. *Eur J Biochem*. 1991 Dec 18; 202 (3) :1041-8. PubMed PMID:1837266.

Kaberdin VR, Singh D, Lin-Chao S. Composition and conservation of the mRNA-degrading machinery in bacteria. *J Biomed Sci*. 2011 Mar 22; 18:23. PubMed PMID:21418661; PubMed Central PMCID: PMC3071783. Kaito C, Morishita D, Matsumoto Y, Kurokawa K, Sekimizu K. Novel DNA binding protein SarZ contributes to virulence in *Staphylococcus aureus*. *Mol Microbiol*. 2006 Dec; 62 (6) :1601-17. PubMed PMID:17087772.

Kamada N, Kim YG, Sham HP, Vallance BA, Puente JL, Martens EC, Núñez G. Regulated virulence controls the ability of a pathogen to compete with the gut microbiota. *Science*. 2012 Jun 8; 336 (6086) :1325-9. PubMed PMID:22582016; PubMed Central PMCID: PMC3439148.

Kaneko J, Kamio Y. Bacterial two-component and hetero-heptameric pore-forming cytolytic toxins: structures, pore-forming mechanism, and organization of the genes. *Biosci Biotechnol Biochem*. 2004 May; 68 (5) :981-1003. PubMed PMID:15170101.

Kawamoto H, Koide Y, Morita T, Aiba H. Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Mol Microbiol*. 2006 Aug; 61 (4) :1013-22. PubMed PMID:16859494.

Kearns DB, Losick R. Cell population heterogeneity during growth of *Bacillus subtilis*. *Genes Dev*. 2005 Dec 15; 19 (24) :3083-94. PubMed PMID:16357223; PubMed Central PMCID: PMC1315410.

Kim JN, Blount KF, Puskarz I, Lim J, Link KH, Breaker RR. Design and antimicrobial action of purine analogues that bind Guanine riboswitches. *ACS Chem Biol*. 2009 Nov 20; 4 (11) :915-27. PubMed PMID:19739679.

Kim T, Oh PI, Simor AE. The economic impact of methicillin-resistant *Staphylococcus aureus* in Canadian hospitals. *Infect Control Hosp Epidemiol*. 2001 Feb; 22 (2) :99-104. PubMed PMID:11232886.

King BF, Wilkinson BJ. Binding of human immunoglobulin G to protein A in encapsulated *Staphylococcus aureus*. *Infect Immun*. 1981 Sep; 33 (3) :666-72. PubMed PMID:7287175; PubMed Central PMCID: PMC350759.

Ko YP, Liang X, Smith CW, Degen JL, Höök M. Binding of Efb from *Staphylococcus aureus* to fibrinogen blocks neutrophil adherence. *J Biol Chem*. 2011 Mar 18; 286 (11) :9865-74. PubMed PMID:21247890; PubMed Central PMCID: PMC3059020.

Kolar SL, Nagarajan V, Oszmiana A, Rivera FE, Miller HK, Davenport JE, Riordan JT, Potempa J, Barber DS, Koziel J, Elasri MO, Shaw LN. NsaRS is a cell-envelope-stress-sensing two-component system of *Staphylococcus aureus*. *Microbiology*. 2011 Aug; 157 (Pt 8) :2206-19. PubMed PMID:21565927; PubMed Central PMCID: PMC3167884.

Kolb FA, Malmgren C, Westhof E, Ehresmann C, Ehresmann B, Wagner EG, Romby P. An unusual structure formed by antisense-target RNA binding involves an extended kissing complex with a four-way junction and a side-by-side helical alignment. *RNA*. 2000 Mar; 6 (3) :311-24. PubMed PMID:10744017; PubMed Central PMCID: PMC1369915.

Kolb FA, Engdahl HM, Slagter-Jäger JG, Ehresmann B, Ehresmann C, Westhof E, Wagner EG, Romby P. Progression of a loop-loop complex to a four-way junction is crucial for the activity of a regulatory antisense RNA. *EMBO J*. 2000 Nov 1; 19 (21) :5905-15. PubMed PMID:11060041; PubMed Central PMCID: PMC305787.

Kullander J, Forslund O, Dillner J. *Staphylococcus aureus* and squamous cell carcinoma of the skin. *Cancer Epidemiol Biomarkers Prev*. 2009 Feb; 18 (2) :472-8. PubMed PMID:19155437.

Kuroda M, Kuroda H, Oshima T, Takeuchi F, Mori H, Hiramatsu K. Two-component system *VraSR* positively modulates the regulation of cell-wall biosynthesis pathway in *Staphylococcus aureus*. *Mol Microbiol*. 2003 Aug; 49 (3) :807-21. PubMed PMID:12864861.

Kuroda M, Ohta T, Uchiyama I, Baba T, Yuzawa H, Kobayashi I, Cui L, Oguchi A, Aoki K, Nagai Y, Lian J, Ito T, Kanamori M, Matsumaru H, Maruyama A, Murakami H, Hosoyama A, Mizutani-Ui Y, Takahashi NK, Sawano T, Inoue R, Kaito C, Sekimizu K, Hiramatsu K, Kuhara S, Goto S, Yabuzaki J, Kanehisa M, Yamashita A, Oshima K, Furuya K, Yoshino C, Shiba T, Hattori M, Ogasawara N, Hayashi H, Hiramatsu K. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. *Lancet*. 2001 Apr 21; 357 (9264) :1225-40. PubMed PMID:11418146.

Lacey RW, Mitchell AA. Gentamicin-resistant *Staphylococcus aureus*. *Lancet*. 1969 Dec 27; 2 (7635) :1425-6. PubMed PMID:4188305.

Lappin E, Ferguson AJ. Gram-positive toxic shock syndromes. *Lancet Infect Dis*. 2009 May; 9 (5) :281-90. PubMed PMID:19393958.

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. Genome-wide antisense transcription drives mRNA processing in bacteria. *Proc Natl Acad Sci U S A*. 2011 Dec 13; 108 (50) :20172-7. PubMed PMID:22123973; PubMed Central PMCID: PMC3250193.

Lee EJ, Groisman EA. An antisense RNA that governs the expression kinetics of a multifunctional virulence gene. *Mol Microbiol*. 2010 May; 76 (4) :1020-33. PubMed PMID:20398218; PubMed Central PMCID: PMC2909850.

Lee HH, Molla MN, Cantor CR, Collins JJ. Bacterial charity work leads to population-wide resistance. *Nature*. 2010 Sep 2; 467 (7311) :82-5. PubMed PMID:20811456; PubMed Central PMCID: PMC2936489.

Lee LY, Liang X, Höök M, Brown EL. Identification and characterization of the C3 binding domain of the *Staphylococcus aureus* extracellular fibrinogen-binding protein (Efb). *J Biol Chem*. 2004 Dec 3; 279 (49) :50710-6. PubMed PMID:15337748.

Lee ST, Chu K, Im WS, Yoon HJ, Im JY, Park JE, Park KH, Jung KH, Lee SK, Kim M, Roh JK. Altered microRNA regulation in Huntington's disease models. *Exp Neurol*. 2011 Jan; 227 (1) :172-9. PubMed PMID:21035445.

Lehnik-Habrink M, Schaffer M, Mäder U, Diethmaier C, Herzberg C, Stülke J. RNA processing in *Bacillus subtilis*: identification of targets of the essential RNase Y. *Mol Microbiol.* 2011 Sep; 81 (6) :1459-73. PubMed PMID:21815947.

Ley RE, Peterson DA, Gordon JI. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell.* 2006 Feb 24; 124 (4) :837-48. PubMed PMID:16497592.

Liang X, Yu C, Sun J, Liu H, Landwehr C, Holmes D, Ji Y. Inactivation of a two-component signal transduction system, SaeRS, eliminates adherence and attenuates virulence of *Staphylococcus aureus*. *Infect Immun.* 2006 Aug; 74 (8) :4655-65. PubMed PMID:16861653; PubMed Central PMCID: PMC1539584.

Lioliou E, Sharma CM, Caldelari I, Helfer AC, Fechter P, Vandenesch F, Vogel J, Romby P. Global regulatory functions of the *Staphylococcus aureus* endoribonuclease III in gene expression. *PLoS Genet.* 2012 Jun; 8 (6) :e1002782. PubMed PMID:22761586; PubMed Central PMCID: PMC3386247.

Liu GY, Essex A, Buchanan JT, Datta V, Hoffman HM, Bastian JF, Fierer J, Nizet V. *Staphylococcus aureus* golden pigment impairs neutrophil killing and promotes virulence through its antioxidant activity. *J Exp Med.* 2005 Jul 18; 202 (2) :209-15. PubMed PMID:16009720; PubMed Central PMCID: PMC2213009.

Liu MY, Romeo T. The global regulator CsrA of *Escherichia coli* is a specific mRNA-binding protein. *J Bacteriol.* 1997 Jul; 179 (14) :4639-42. PubMed PMID:9226279; PubMed Central PMCID: PMC179305.

Liu Y, Wu N, Dong J, Gao Y, Zhang X, Mu C, Shao N, Yang G. Hfq is a global regulator that controls the pathogenicity of *Staphylococcus aureus*. *PLoS One.* 2010 Sep 29; 5 (9) PubMed PMID:20927372; PubMed Central PMCID: PMC2947504.

Livny J, Teonadi H, Livny M, Waldor MK. High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. *PLoS One.* 2008 Sep 12; 3 (9) :e3197. PubMed PMID:18787707; PubMed Central PMCID: PMC2527527.

Lo WT, Wang CC. Pantone-Valentine leukocidin in the pathogenesis of community-associated methicillin-resistant *Staphylococcus aureus* infection. *Pediatr Neonatol.* 2011 Apr; 52 (2) :59-65. PubMed PMID:21524624.

Loh E, Dussurget O, Gripenland J, Vaitkevicius K, Tiensuu T, Mandin P, Repoila F, Buchrieser C, Cossart P, Johansson J. A trans-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. *Cell.* 2009 Nov 13; 139 (4) :770-9. PubMed PMID:19914169.

López D, Vlamakis H, Kolter R. Biofilms. *Cold Spring Harb Perspect Biol.* 2010 Jul; 2 (7) :a000398. PubMed PMID:20519345; PubMed Central PMCID: PMC2890205.

Luong T, Sau S, Gomez M, Lee JC, Lee CY. Regulation of *Staphylococcus aureus* capsular polysaccharide expression by agr and sarA. *Infect Immun.* 2002 Feb; 70 (2) :444-50. PubMed PMID:11796569; PubMed Central PMCID: PMC127668.

Luong TT, Newell SW, Lee CY. Mgr, a novel global regulator in *Staphylococcus aureus*. *J Bacteriol.* 2003 Jul; 185 (13) :3703-10. PubMed PMID:12813062; PubMed Central PMCID: PMC161569.

Luong TT, Lee CY. The arl locus positively regulates *Staphylococcus aureus* type 5 capsule via an mgrA-dependent pathway. *Microbiology.* 2006 Oct; 152 (Pt 10) :3123-31. PubMed PMID:17005991.

Luong TT, Dunman PM, Murphy E, Projan SJ, Lee CY. Transcription Profiling of the mgrA Regulon in *Staphylococcus aureus*. *J Bacteriol.* 2006 Mar; 188 (5) :1899-910. PubMed PMID:16484201; PubMed Central PMCID: PMC1426550.

Ma J, Campbell A, Karlin S. Correlations between Shine-Dalgarno sequences and gene features such as predicted expression levels and operon structures. *J Bacteriol.* 2002 Oct; 184 (20) :5733-45. PubMed PMID:12270832; PubMed Central PMCID: PMC139613.

Mackie GA. Stabilization of circular rpsT mRNA demonstrates the 5'-end dependence of RNase E action in vivo. *J Biol Chem.* 2000 Aug 18; 275 (33) :25069-72. PubMed PMID:10871599.

MacRae IJ, Doudna JA. Ribonuclease revisited: structural insights into ribonuclease III family enzymes. *Curr Opin Struct Biol.* 2007 Feb; 17 (1) :138-45. PubMed PMID:17194582.

Malachowa N, DeLeo FR. Mobile genetic elements of *Staphylococcus aureus*. *Cell Mol Life Sci.* 2010 Sep; 67 (18) :3057-71. PubMed PMID:20668911; PubMed Central PMCID: PMC2929429.

Malfatti S, Tindall BJ, Schneider S, Fährnich R, Lapidus A, Labuttii K, Copeland A, Glavina Del Rio T, Nolan M, Chen F, Lucas S, Tice H, Cheng JF, Bruce D, Goodwin L, Pitluck S, Anderson I, Pati A, Ivanova N, Mavromatis K, Chen A, Palaniappan K, D'haeseleer P, Göker M, Bristow J, Eisen JA, Markowitz V, Hugenholtz P, Kyrpides NC, Klenk HP, Chain P. Complete genome sequence of *Haloquadratum walsbyi* strain (PR3). *Stand Genomic Sci.* 2009 Sep 24; 1 (2) :150-9. PubMed PMID:21304651; PubMed Central PMCID: PMC3035229.

Mandal M, Lee M, Barrick JE, Weinberg Z, Emilsson GM, Ruzzo WL, Breaker RR. A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science.* 2004 Oct 8; 306 (5694) :275-9. PubMed PMID:15472076.

Manna AC, Bayer MG, Cheung AL. Transcriptional analysis of different promoters in the sar locus in *Staphylococcus aureus*. *J Bacteriol.* 1998 Aug; 180 (15) :3828-36. PubMed PMID:9683479; PubMed Central PMCID: PMC107366.

Manna AC, Ingavale SS, Maloney M, van Wamel W, Cheung AL. Identification of sarV (SA2062), a new transcriptional regulator, is repressed by SarA and MgrA (SA0641) and involved in the regulation of autolysis in *Staphylococcus aureus*. *J Bacteriol.* 2004 Aug; 186 (16) :5267-80. PubMed PMID:15292128; PubMed Central PMCID: PMC490931.

Manna AC, Cheung AL. Expression of SarX, a negative regulator of agr and exoprotein synthesis, is activated by MgrA in *Staphylococcus aureus*. *J Bacteriol.* 2006 Jun; 188 (12) :4288-99. PubMed PMID:16740935; PubMed Central PMCID: PMC1482969.

Marchais A, Naville M, Bohn C, Bouloc P, Gautheret D. Single-pass classification of all noncoding sequences in a bacterial genome using phylogenetic profiles. *Genome Res.* 2009 Jun; 19 (6) :1084-92. PubMed PMID:19237465; PubMed Central PMCID: PMC2694484.

Marincola G, Schäfer T, Behler J, Bernhardt J, Ohlsen K, Goerke C, Wolz C. RNase Y of *Staphylococcus aureus* and its role in the activation of virulence genes. *Mol Microbiol.* 2012 Sep; 85 (5) :817-32. PubMed PMID:22780584.

Marrie TJ, Nelligan J, Costerton JW. A scanning and transmission electron microscopic study of an infected endocardial pacemaker lead. *Circulation.* 1982 Dec; 66 (6) :1339-41. PubMed PMID:7139907.

Massé E, Gottesman S. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc Natl Acad Sci U S A.* 2002 Apr 2; 99 (7) :4620-5. PubMed PMID:11917098; PubMed Central PMCID: PMC123697.

Massé E, Escorcía FE, Gottesman S. Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev.* 2003 Oct 1; 17 (19) :2374-83. PubMed PMID:12975324; PubMed Central PMCID: PMC218075.

McGavin MJ, Zahradka C, Rice K, Scott JE. Modification of the *Staphylococcus aureus* fibronectin binding phenotype by V8 protease. *Infect Immun*. 1997 Jul; 65 (7) :2621-8. PubMed PMID:9199429; PubMed Central PMCID: PMC175371.

Meier S, Goerke C, Wolz C, Seidl K, Homerova D, Schulthess B, Kormanec J, Berger-Bächi B, Bischoff M. *sigmaB* and the *sigmaB*-dependent *arlRS* and *yabJ-spoVG* loci affect capsule formation in *Staphylococcus aureus*. *Infect Immun*. 2007 Sep; 75 (9) :4562-71. PubMed PMID:17635871; PubMed Central PMCID: PMC1951174.

Menestrina G, Dalla Serra M, Comai M, Coraiola M, Viero G, Werner S, Colin DA, Monteil H, Prévost G. Ion channels and bacterial infection: the case of beta-barrel pore-forming protein toxins of *Staphylococcus aureus*. *FEBS Lett*. 2003 Sep 18; 552 (1) :54-60. PubMed PMID:12972152.

Menzies BE. The role of fibronectin binding proteins in the pathogenesis of *Staphylococcus aureus* infections. *Curr Opin Infect Dis*. 2003 Jun; 16 (3) :225-9. PubMed PMID:12821812.

Mittenhuber G. An inventory of genes encoding RNA polymerase sigma factors in 31 completely sequenced eubacterial genomes. *J Mol Microbiol Biotechnol*. 2002 Jan; 4 (1) :77-91. PubMed PMID:11763973.

Mizuno T, Chou MY, Inouye M. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc Natl Acad Sci U S A*. 1984 Apr; 81 (7) :1966-70. PubMed PMID:6201848; PubMed Central PMCID: PMC345417.

Moll I, Afonyushkin T, Vytvytska O, Kaberdin VR, Bläsi U. Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA*. 2003 Nov; 9 (11) :1308-14. PubMed PMID:14561880; PubMed Central PMCID: PMC1287052.

Montero Llopis P, Jackson AF, Sliusarenko O, Surovtsev I, Heinritz J, Emonet T, Jacobs-Wagner C. Spatial organization of the flow of genetic information in bacteria. *Nature*. 2010 Jul 1; 466 (7302) :77-81. PubMed PMID:20562858; PubMed Central PMCID: PMC2896451.

Moreno R, Fonseca P, Rojo F. Two small RNAs, *CrcY* and *CrcZ*, act in concert to sequester the *Crc* global regulator in *Pseudomonas putida*, modulating catabolite repression. *Mol Microbiol*. 2012 Jan; 83 (1) :24-40. PubMed PMID:22053874.

Morfeldt E, Taylor D, von Gabain A, Arvidson S. Activation of alpha-toxin translation in *Staphylococcus aureus* by the trans-encoded antisense RNA, *RNAIII*. *EMBO J*. 1995 Sep 15; 14 (18) :4569-77. PubMed PMID:7556100; PubMed Central PMCID: PMC394549.

Morikawa K, Inose Y, Okamura H, Maruyama A, Hayashi H, Takeyasu K, Ohta T. A new staphylococcal sigma factor in the conserved gene cassette: functional significance and implication for the evolutionary processes. *Genes Cells*. 2003 Aug; 8 (8) :699-712. PubMed PMID:12875655.

Morita T, Maki K, Aiba H. RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev*. 2005 Sep 15; 19 (18) :2176-86. PubMed PMID:16166379; PubMed Central PMCID: PMC1221888.

Mulcahy H, Sibley CD, Surette MG, Lewenza S. *Drosophila melanogaster* as an animal model for the study of *Pseudomonas aeruginosa* biofilm infections in vivo. *PLoS Pathog*. 2011 Oct; 7 (10) :e1002299. PubMed PMID:21998591; PubMed Central PMCID: PMC3188550.

Mulhbacher J, Brouillette E, Allard M, Fortier LC, Malouin F, Lafontaine DA. Novel riboswitch ligand analogs as selective inhibitors of guanine-related metabolic pathways. *PLoS Pathog*. 2010 Apr 22; 6 (4) :e1000865. PubMed PMID:20421948; PubMed Central PMCID: PMC2858708.

Nakano MM, Zuber P, Glaser P, Danchin A, Hulett FM. Two-component regulatory proteins *ResD-ResE* are required for transcriptional activation of *fnr* upon oxygen limitation in *Bacillus subtilis*. *J*

Bacteriol. 1996 Jul; 178 (13) :3796-802. PubMed PMID:8682783; PubMed Central PMCID: PMC232639.

Narberhaus F. Translational control of bacterial heat shock and virulence genes by temperature-sensing mRNAs. RNA Biol. 2010 Jan-Feb; 7 (1) :84-9. PubMed PMID:20009504.

Nechooshtan G, Elgrably-Weiss M, Sheaffer A, Westhof E, Altuvia S. A pH-responsive riboregulator. Genes Dev. 2009 Nov 15; 23 (22) :2650-62. PubMed PMID:19933154; PubMed Central PMCID: PMC2779765.

Nevo-Dinur K, Nussbaum-Shochat A, Ben-Yehuda S, Amster-Choder O. Translation-independent localization of mRNA in E coli. Science. 2011 Feb 25; 331 (6020) :1081-4. PubMed PMID:21350180.

Nicholson JK, Holmes E, Wilson ID. Gut microorganisms, mammalian metabolism and personalized health care. Nat Rev Microbiol. 2005 May; 3 (5) :431-8. PubMed PMID:15821725.

Nielsen JS, Lei LK, Ebersbach T, Olsen AS, Klitgaard JK, Valentin-Hansen P, Kallipolitis BH. Defining a role for Hfq in Gram-positive bacteria: evidence for Hfq-dependent antisense regulation in *Listeria monocytogenes*. Nucleic Acids Res. 2010 Jan; 38 (3) :907-19. PubMed PMID:19942685; PubMed Central PMCID: PMC2817478.

Nordström K. Plasmid R1--replication and its control. Plasmid. 2006 Jan; 55 (1) :1-26. PubMed PMID:16199086.

Novick RP. Autoinduction and signal transduction in the regulation of staphylococcal virulence. Mol Microbiol. 2003 Jun; 48 (6) :1429-49. PubMed PMID:12791129.

Novick RP, Ross HF, Projan SJ, Kornblum J, Kreiswirth B, Moghazeh S. Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. EMBO J. 1993 Oct; 12 (10) :3967-75. PubMed PMID:7691599; PubMed Central PMCID: PMC413679.

O'Riordan K, Lee JC. Staphylococcus aureus capsular polysaccharides. Clin Microbiol Rev. 2004 Jan; 17 (1) :218-34. PubMed PMID:14726462; PubMed Central PMCID: PMC321462.

O'Rourke JL, Grehan M, Lee A. Non-pylori Helicobacter species in humans. Gut. 2001 Nov; 49 (5) :601-6. PubMed PMID:11600455; PubMed Central PMCID: PMC1728516.

Obana N, Shirahama Y, Abe K, Nakamura K. Stabilization of *Clostridium perfringens* collagenase mRNA by VR-RNA dependent cleavage in 5' leader sequence. Mol Microbiol. 2010 Sep; 77 (6) :1416-28. PubMed PMID:20572941.

Olson PD, Kuechenmeister LJ, Anderson KL, Daily S, Beenken KE, Roux CM, Reniere ML, Lewis TL, Weiss WJ, Pulse M, Nguyen P, Simecka JW, Morrison JM, Sayood K, Asojo OA, Smeltzer MS, Skaar EP, Dunman PM. Small molecule inhibitors of *Staphylococcus aureus* RnpA alter cellular mRNA turnover, exhibit antimicrobial activity, and attenuate pathogenesis. PLoS Pathog. 2011 Feb 10; 7 (2) :e1001287. PubMed PMID:21347352; PubMed Central PMCID: PMC3037362.

Opdyke JA, Fozo EM, Hemm MR, Storz G. RNase III participates in GadY-dependent cleavage of the gadX-gadW mRNA. J Mol Biol. 2011 Feb 11; 406 (1) :29-43. PubMed PMID:21147125; PubMed Central PMCID: PMC3030618.

Opdyke JA, Kang JG, Storz G. GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. J Bacteriol. 2004 Oct; 186 (20) :6698-705. PubMed PMID:15466020; PubMed Central PMCID: PMC522195.

Oscarsson J, Kanth A, Tegmark-Wisell K, Arvidson S. SarA is a repressor of hla (alpha-hemolysin) transcription in *Staphylococcus aureus*: its apparent role as an activator of hla in the prototype strain NCTC 8325 depends on reduced expression of sarS. J Bacteriol. 2006 Dec; 188 (24) :8526-33. PubMed PMID:17012389; PubMed Central PMCID: PMC1698246.

Paige JS, Wu KY, Jaffrey SR. RNA mimics of green fluorescent protein. *Science*. 2011 Jul 29; 333 (6042) :642-6. PubMed PMID:21798953; PubMed Central PMCID: PMC3314379.

Palma M, Cheung AL. sigma(B) activity in *Staphylococcus aureus* is controlled by RsbU and an additional factor(s) during bacterial growth. *Infect Immun*. 2001 Dec; 69 (12) :7858-65. PubMed PMID:11705968; PubMed Central PMCID: PMC98882.

Palmer ME, Chaturongakul S, Wiedmann M, Boor KJ. The *Listeria monocytogenes*  $\sigma$ B regulon and its virulence-associated functions are inhibited by a small molecule. *MBio*. 2011; 2 (6) PubMed PMID:22128349; PubMed Central PMCID: PMC3225968.

Palmqvist N, Patti JM, Tarkowski A, Josefsson E. Expression of staphylococcal clumping factor A impedes macrophage phagocytosis. *Microbes Infect*. 2004 Feb; 6 (2) :188-95. PubMed PMID:14998517.

Pandey SP, Minesinger BK, Kumar J, Walker GC. A highly conserved protein of unknown function in *Sinorhizobium meliloti* affects sRNA regulation similar to Hfq. *Nucleic Acids Res*. 2011 Jun; 39 (11) :4691-708. PubMed PMID:21325267; PubMed Central PMCID: PMC3113577.

Pané-Farré J, Jonas B, Förstner K, Engelmann S, Hecker M. The sigmaB regulon in *Staphylococcus aureus* and its regulation. *Int J Med Microbiol*. 2006 Aug; 296 (4-5) :237-58. PubMed PMID:16644280.  
Pantosti A, Venditti M. What is MRSA?. *Eur Respir J*. 2009 Nov; 34 (5) :1190-6. PubMed PMID:19880619.

Papenfort K, Pfeiffer V, Mika F, Lucchini S, Hinton JC, Vogel J. SigmaE-dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global omp mRNA decay. *Mol Microbiol*. 2006 Dec; 62 (6) :1674-88. PubMed PMID:17427289; PubMed Central PMCID: PMC1804206.

Pedersen K. Exploration of deep intraterrestrial microbial life: current perspectives. *FEMS Microbiol Lett*. 2000 Apr 1; 185 (1) :9-16. PubMed PMID:10731600.

Peer A, Margalit H. Accessibility and evolutionary conservation mark bacterial small-rna target-binding regions. *J Bacteriol*. 2011 Apr; 193 (7) :1690-701. PubMed PMID:21278294; PubMed Central PMCID: PMC3067639.

Perret M, Badiou C, Lina G, Burbaud S, Benito Y, Bes M, Cottin V, Couzon F, Juruj C, Dauwalder O, Goutagny N, Diep BA, Vandenesch F, Henry T. Cross-talk between *Staphylococcus aureus* leukocidins-intoxicated macrophages and lung epithelial cells triggers chemokine secretion in an inflammasome-dependent manner. *Cell Microbiol*. 2012 Jul; 14 (7) :1019-36. PubMed PMID:22329718.

Pertzev AV, Nicholson AW. Characterization of RNA sequence determinants and antideterminants of processing reactivity for a minimal substrate of *Escherichia coli* ribonuclease III. *Nucleic Acids Res*. 2006; 34 (13) :3708-21. PubMed PMID:16896014; PubMed Central PMCID: PMC1540722.

Petersohn A, Bernhardt J, Gerth U, Höper D, Koburger T, Völker U, Hecker M. Identification of sigma(B)-dependent genes in *Bacillus subtilis* using a promoter consensus-directed search and oligonucleotide hybridization. *J Bacteriol*. 1999 Sep; 181 (18) :5718-24. PubMed PMID:10482513; PubMed Central PMCID: PMC94092.

Pfeiffer V, Papenfort K, Lucchini S, Hinton JC, Vogel J. Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol*. 2009 Aug; 16 (8) :840-6. PubMed PMID:19620966.

Pichon C, Felden B. Small RNA genes expressed from *Staphylococcus aureus* genomic and pathogenicity islands with specific expression among pathogenic strains. *Proc Natl Acad Sci U S A*. 2005 Oct 4; 102 (40) :14249-54. PubMed PMID:16183745; PubMed Central PMCID: PMC1242290.

Poiata E, Meyer MM, Ames TD, Breaker RR. A variant riboswitch aptamer class for S-adenosylmethionine common in marine bacteria. *RNA*. 2009 Nov; 15 (11) :2046-56. PubMed PMID:19776155; PubMed Central PMCID: PMC2764483.

Pragman AA, Yarwood JM, Tripp TJ, Schlievert PM. Characterization of virulence factor regulation by SrrAB, a two-component system in *Staphylococcus aureus*. *J Bacteriol*. 2004 Apr; 186 (8) :2430-8. PubMed PMID:15060046; PubMed Central PMCID: PMC412142.

Preis H, Eckart RA, Gudipati RK, Heidrich N, Brantl S. CodY activates transcription of a small RNA in *Bacillus subtilis*. *J Bacteriol*. 2009 Sep; 191 (17) :5446-57. PubMed PMID:19542274; PubMed Central PMCID: PMC2725614.

Proshkin S, Rahmouni AR, Mironov A, Nudler E. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science*. 2010 Apr 23; 328 (5977) :504-8. PubMed PMID:20413502; PubMed Central PMCID: PMC2930199.

Queck SY, Jameson-Lee M, Villaruz AE, Bach TH, Khan BA, Sturdevant DE, Ricklefs SM, Li M, Otto M. RNA-independent target gene control by the agr quorum-sensing system: insight into the evolution of virulence regulation in *Staphylococcus aureus*. *Mol Cell*. 2008 Oct 10; 32 (1) :150-8. PubMed PMID:18851841; PubMed Central PMCID: PMC2575650.

Rabhi M, Espéli O, Schwartz A, Cayrol B, Rahmouni AR, Arluison V, Boudvillain M. The Sm-like RNA chaperone Hfq mediates transcription antitermination at Rho-dependent terminators. *EMBO J*. 2011 Jun 14; 30 (14) :2805-16. PubMed PMID:21673658; PubMed Central PMCID: PMC3160251.

Ramirez-Peña E, Treviño J, Liu Z, Perez N, Sumbly P. The group A *Streptococcus* small regulatory RNA FasX enhances streptokinase activity by increasing the stability of the ska mRNA transcript. *Mol Microbiol*. 2010 Dec; 78 (6) :1332-47. PubMed PMID:21143309; PubMed Central PMCID: PMC3071709.

Rammelkamp CH, Keefer CS. PENICILLIN: ITS ANTIBACTERIAL EFFECT IN WHOLE BLOOD AND SERUM FOR THE HEMOLYTIC STREPTOCOCCUS AND STAPHYLOCOCCUS AUREUS. *J Clin Invest*. 1943 Sep; 22 (5) :649-57. PubMed PMID:16695048; PubMed Central PMCID: PMC435281.

Resch A, Večerek B, Palavra K, Bläsi U. Requirement of the CsdA DEAD-box helicase for low temperature riboregulation of rpoS mRNA. *RNA Biol*. 2010 Nov-Dec; 7 (6) :796-802. PubMed PMID:21045550; PubMed Central PMCID: PMC3073337.

Rhodus VA, Suh WC, Nonaka G, West J, Gross CA. Conserved and variable functions of the sigmaE stress response in related genomes. *PLoS Biol*. 2006 Jan; 4 (1) :e2. PubMed PMID:16336047; PubMed Central PMCID: PMC1312014.

Rieger-Fackeldey E, Plano LR, Kramer A, Schulze A. Staphylococcal scalded skin syndrome related to an exfoliative toxin A- and B-producing strain in preterm infants. *Eur J Pediatr*. 2002 Dec; 161 (12) :649-52. PubMed PMID:12447663.

Rivas E, Klein RJ, Jones TA, Eddy SR. Computational identification of noncoding RNAs in *E coli* by comparative genomics. *Curr Biol*. 2001 Sep 4; 11 (17) :1369-73. PubMed PMID:11553332.

Roberts C, Anderson KL, Murphy E, Projan SJ, Mounts W, Hurlburt B, Smeltzer M, Overbeek R, Disz T, Dunman PM. Characterizing the effect of the *Staphylococcus aureus* virulence factor regulator, SarA, on log-phase mRNA half-lives.

*J Bacteriol*. 2006 Apr; 188 (7) :2593-603. PubMed PMID:16547047; PubMed Central PMCID: PMC1428411.

Rochat T, Bouloc P, Yang Q, Bossi L, Figueroa-Bossi N. Lack of interchangeability of Hfq-like proteins. *Biochimie*. 2012 Jul; 94 (7) :1554-9. PubMed PMID:22326874.



Rogasch K, Rühmling V, Pané-Farré J, Höper D, Weinberg C, Fuchs S, Schmutde M, Bröker BM, Wolz C, Hecker M, Engelmann S. Influence of the two-component system SaeRS on global gene expression in two different *Staphylococcus aureus* strains. *J Bacteriol.* 2006 Nov; 188 (22) :7742-58. PubMed PMID:17079681; PubMed Central PMCID: PMC1636327.

Romilly C, Caldelari I, Parmentier D, Lioliou E, Romby P, Fechter P. Current knowledge on regulatory RNAs and their machineries in *Staphylococcus aureus*. *RNA Biol.* 2012 Apr 1; 9 (4) PubMed PMID:22546940.

Rosenstein R, Götz F. Staphylococcal lipases: biochemical and molecular characterization. *Biochimie.* 2000 Nov; 82 (11) :1005-14. PubMed PMID:11099797.

Rosenstein R, Götz F. Genomic differences between the food-grade *Staphylococcus carnosus* and pathogenic staphylococcal species. *Int J Med Microbiol.* 2010 Feb; 300 (2-3) :104-8. PubMed PMID:19818681.

Roux CM, DeMuth JP, Dunman PM. Characterization of components of the *Staphylococcus aureus* mRNA degradosome holoenzyme-like complex. *J Bacteriol.* 2011 Oct; 193 (19) :5520-6. PubMed PMID:21764917; PubMed Central PMCID: PMC3187390.

Saïd-Salim B, Dunman PM, McAleese FM, Macapagal D, Murphy E, McNamara PJ, Arvidson S, Foster TJ, Projan SJ, Kreiswirth BN. Global regulation of *Staphylococcus aureus* genes by Rot. *J Bacteriol.* 2003 Jan; 185 (2) :610-9. PubMed PMID:12511508; PubMed Central PMCID: PMC145333.

Sauer E, Schmidt S, Weichenrieder O. Small RNA binding to the lateral surface of Hfq hexamers and structural rearrangements upon mRNA target recognition. *Proc Natl Acad Sci U S A.* 2012 Jun 12; 109 (24) :9396-401. PubMed PMID:22645344; PubMed Central PMCID: PMC3386104.

Sauer E, Weichenrieder O. Structural basis for RNA 3'-end recognition by Hfq. *Proc Natl Acad Sci U S A.* 2011 Aug 9; 108 (32) :13065-70. PubMed PMID:21737752; PubMed Central PMCID: PMC3156190.

Sauter C, Basquin J, Suck D. Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from *Escherichia coli*. *Nucleic Acids Res.* 2003 Jul 15; 31 (14) :4091-8. PubMed PMID:12853626; PubMed Central PMCID: PMC167641.

Sayed N, Jousselin A, Felden B. A cis-antisense RNA acts in trans in *Staphylococcus aureus* to control translation of a human cytolytic peptide. *Nat Struct Mol Biol.* 2011 Dec 25; 19 (1) :105-12. PubMed PMID:22198463.

Schmidt KA, Manna AC, Gill S, Cheung AL. SarT, a repressor of alpha-hemolysin in *Staphylococcus aureus*. *Infect Immun.* 2001 Aug; 69 (8) :4749-58. PubMed PMID:11447147; PubMed Central PMCID: PMC98561.

Schneider T, Senn MM, Berger-Bächli B, Tossi A, Sahl HG, Wiedemann I. In vitro assembly of a complete, pentaglycine interpeptide bridge containing cell wall precursor (lipid II-Gly5) of *Staphylococcus aureus*. *Mol Microbiol.* 2004 Jul; 53 (2) :675-85. PubMed PMID:15228543.

Schubert M, Lapouge K, Duss O, Oberstrass FC, Jelesarov I, Haas D, Allain FH. Molecular basis of messenger RNA recognition by the specific bacterial repressing clamp RsmA/CsrA. *Nat Struct Mol Biol.* 2007 Sep; 14 (9) :807-13. PubMed PMID:17704818.

Schulthess B, Meier S, Homerova D, Goerke C, Wolz C, Kormanec J, Berger-Bächli B, Bischoff M. Functional characterization of the sigmaB-dependent yabJ-spoVG operon in *Staphylococcus aureus*: role in methicillin and glycopeptide resistance. *Antimicrob Agents Chemother.* 2009 May; 53 (5) :1832-9. PubMed PMID:19223635; PubMed Central PMCID: PMC2681525.

Scotti M, Monzó HJ, Lacharme-Lora L, Lewis DA, Vázquez-Boland JA. The PrfA virulence regulon. *Microbes Infect.* 2007 Aug; 9 (10) :1196-207. PubMed PMID:17764998.

Shahbadian K, Jamalli A, Zig L, Putzer H. RNase Y, a novel endoribonuclease, initiates riboswitch turnover in *Bacillus subtilis*. *EMBO J.* 2009 Nov 18; 28 (22) :3523-33. PubMed PMID:19779461; PubMed Central PMCID: PMC2782095.

Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, Chabas S, Reiche K, Hackermüller J, Reinhardt R, Stadler PF, Vogel J. The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature.* 2010 Mar 11; 464 (7286) :250-5. PubMed PMID:20164839.

Sharp JS, Bechhofer DH. Effect of 5'-proximal elements on decay of a model mRNA in *Bacillus subtilis*. *Mol Microbiol.* 2005 Jul; 57 (2) :484-95. PubMed PMID:15978079

Shaw LN, Lindholm C, Prajsnar TK, Miller HK, Brown MC, Golonka E, Stewart GC, Tarkowski A, Potempa J. Identification and characterization of sigma, a novel component of the *Staphylococcus aureus* stress and virulence responses. *PLoS One.* 2008; 3 (12) :e3844. PubMed PMID:19050758; PubMed Central PMCID: PMC2585143.

Shimoni Y, Friedlander G, Hetzroni G, Niv G, Altuvia S, Biham O, Margalit H. Regulation of gene expression by small non-coding RNAs: a quantitative view. *Mol Syst Biol.* 2007; 3:138. PubMed PMID:17893699; PubMed Central PMCID: PMC2013925.

Silverman GJ, Goodyear CS. Confounding B-cell defences: lessons from a staphylococcal superantigen. *Nat Rev Immunol.* 2006 Jun; 6 (6) :465-75. PubMed PMID:16724100.

Singh OV, Labana S, Pandey G, Budhiraja R, Jain RK. Phytoremediation: an overview of metallic ion decontamination from soil. *Appl Microbiol Biotechnol.* 2003 Jun; 61 (5-6) :405-12. PubMed PMID:12764555.

Smith AM, Fuchs RT, Grundy FJ, Henkin TM. Riboswitch RNAs: regulation of gene expression by direct monitoring of a physiological signal. *RNA Biol.* 2010 Jan-Feb; 7 (1) :104-10. PubMed PMID:20061810.

Song J, Lays C, Vandenesch F, Benito Y, Bes M, Chu Y, Lina G, Romby P, Geissmann T, Boisset S. The expression of small regulatory RNAs in clinical samples reflects the different life styles of *Staphylococcus aureus* in colonization vs infection. *PLoS One.* 2012; 7 (5) :e37294. PubMed PMID:22629378; PubMed Central PMCID: PMC3358344.

Song L, Hobaugh MR, Shustak C, Cheley S, Bayley H, Gouaux JE. Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science.* 1996 Dec 13; 274 (5294) :1859-66. PubMed PMID:8943190.

Sonnleitner E, Gonzalez N, Sorger-Domenigg T, Heeb S, Richter AS, Backofen R, Williams P, Hüttenhofer A, Haas D, Bläsi U. The small RNA PhrS stimulates synthesis of the *Pseudomonas aeruginosa* quinolone signal. *Mol Microbiol.* 2011 May; 80 (4) :868-85. PubMed PMID:21375594.

Sonnleitner E, Abdou L, Haas D. Small RNA as global regulator of carbon catabolite repression in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A.* 2009 Dec 22; 106 (51) :21866-71. PubMed PMID:20080802; PubMed Central PMCID: PMC2799872.

Steen EJ, Kang Y, Bokinsky G, Hu Z, Schirmer A, McClure A, Del Cardayre SB, Keasling JD. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature.* 2010 Jan 28; 463 (7280) :559-62. PubMed PMID:20111002.

Stock AM, Robinson VL, Goudreau PN. Two-component signal transduction. *Annu Rev Biochem.* 2000; 69:183-215. PubMed PMID:10966457.

Storz G, Vogel J, Wassarman KM. Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell.* 2011 Sep 16; 43 (6) :880-91. PubMed PMID:21925377; PubMed Central PMCID: PMC3176440.

Sudarsan N, Hammond MC, Block KF, Welz R, Barrick JE, Roth A, Breaker RR. Tandem riboswitch architectures exhibit complex gene control functions. *Science*. 2006 Oct 13; 314 (5797) :300-4. PubMed PMID:17038623.

Sun F, Zhou L, Zhao BC, Deng X, Cho H, Yi C, Jian X, Song CX, Luan CH, Bae T, Li Z, He C. Targeting MgrA-mediated virulence regulation in *Staphylococcus aureus*. *Chem Biol*. 2011 Aug 26; 18 (8) :1032-41. PubMed PMID:21867918; PubMed Central PMCID: PMC3163066.

Sun F, Ji Q, Jones MB, Deng X, Liang H, Frank B, Telser J, Peterson SN, Bae T, He C. AirSR, a [2Fe-2S] clustercontaining two-component system, mediates global oxygen sensing and redox signaling in *Staphylococcus aureus*. *J Am Chem Soc*. 2012 Jan 11; 134 (1) :305-14. PubMed PMID:22122613; PubMed Central PMCID: PMC3257388.

Takada A, Umitsuki G, Nagai K, Wachi M. RNase E is required for induction of the glutamate-dependent acid resistance system in *Escherichia coli*. *Biosci Biotechnol Biochem*. 2007 Jan; 71 (1) :158-64. PubMed PMID:17213667.

Tamber S, Cheung AL. SarZ promotes the expression of virulence factors and represses biofilm formation by modulating SarA and agr in *Staphylococcus aureus*. *Infect Immun*. 2009 Jan; 77 (1) :419-28. PubMed PMID:18955469; PubMed Central PMCID: PMC2612251.

Tao L, Wu X, Sun B. Alternative sigma factor sigmaH modulates prophage integration and excision in *Staphylococcus aureus*. *PLoS Pathog*. 2010 May 13; 6 (5) :e1000888. PubMed PMID:20485515; PubMed Central PMCID: PMC2869324.

Tegmark K, Karlsson A, Arvidson S. Identification and characterization of SarH1, a new global regulator of virulence gene expression in *Staphylococcus aureus*. *Mol Microbiol*. 2000 Jul; 37 (2) :398-409. PubMed PMID:10931334.

Thomason MK, Storz G. Bacterial antisense RNAs: how many are there, and what are they doing?. *Annu Rev Genet*. 2010; 44:167-88. PubMed PMID:20707673; PubMed Central PMCID: PMC3030471.

Thomason MK, Fontaine F, De Lay N, Storz G. A small RNA that regulates motility and biofilm formation in response to changes in nutrient availability in *Escherichia coli*. *Mol Microbiol*. 2012 Apr; 84 (1) :17-35. PubMed PMID:22289118; PubMed Central PMCID: PMC3312966.

Tjaden B. TargetRNA: a tool for predicting targets of small RNA action in bacteria. *Nucleic Acids Res*. 2008 Jul 1; 36 (Web Server issue) :W109-13. PubMed PMID:18477632; PubMed Central PMCID: PMC2447797.

Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, Balestrino D, Loh E, Gripenland J, Tiensuu T, Vaitkevicius K, Barthelemy M, Vergassola M, Nahori MA, Soubigou G, Régnault B, Coppée JY, Lecuit M, Johansson J, Cossart P. The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature*. 2009 Jun 18; 459 (7249) :950-6. PubMed PMID:19448609.

Trotochaud AE, Wassarman KM. A highly conserved 6S RNA structure is required for regulation of transcription. *Nat Struct Mol Biol*. 2005 Apr; 12 (4) :313-9. PubMed PMID:15793584.

Trotonda MP, Tamber S, Memmi G, Cheung AL. MgrA represses biofilm formation in *Staphylococcus aureus*. *Infect Immun*. 2008 Dec; 76 (12) :5645-54. PubMed PMID:18852246; PubMed Central PMCID: PMC2583562.

Truong-Bolduc QC, Zhang X, Hooper DC. Characterization of NorR protein, a multifunctional regulator of norA expression in *Staphylococcus aureus*. *J Bacteriol*. 2003 May; 185 (10) :3127-38. PubMed PMID:12730173; PubMed Central PMCID: PMC154082.

Tung H, Guss B, Hellman U, Persson L, Rubin K, Rydén C. A bone sialoprotein-binding protein from *Staphylococcus aureus*: a member of the staphylococcal Sdr family. *Biochem J*. 2000 Feb 1; 345 Pt 3:611-9. PubMed PMID:10642520; PubMed Central PMCID: PMC1220796.

Türck M, Bierbaum G. Purification and activity testing of the full-length YycFGHI proteins of *Staphylococcus aureus*. *PLoS One*. 2012; 7 (1) :e30403. PubMed PMID:22276191; PubMed Central PMCID: PMC3262814.

Udekwi KI, Darfeuille F, Vogel J, Reimegård J, Holmqvist E, Wagner EG. Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes Dev*. 2005 Oct 1; 19 (19) :2355-66. PubMed PMID:16204185; PubMed Central PMCID: PMC1240044.

Ulrich M, Bastian M, Cramton SE, Ziegler K, Pragman AA, Bragonzi A, Memmi G, Wolz C, Schlievert PM, Cheung A, Döring G. 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*. 2007 Sep; 65 (5) :1276-87. PubMed PMID:17697253.

Vanderpool CK, Balasubramanian D, Lloyd CR. Dual-function RNA regulators in bacteria. *Biochimie*. 2011 Nov; 93 (11) :1943-9. PubMed PMID:21816203; PubMed Central PMCID: PMC3185123.

Vecerek B, Rajkowitsch L, Sonnleitner E, Schroeder R, Bläsi U. The C-terminal domain of *Escherichia coli* Hfq is Required for regulation. *Nucleic Acids Res*. 2008 Jan; 36 (1) :133-43. PubMed PMID:18000007; PubMed Central PMCID: PMC2248732.

Ventre I, Goodman AL, Vallet-Gely I, Vasseur P, Soscia C, Molin S, Bleves S, Lazdunski A, Lory S, Filloux A. Multiple sensors control reciprocal expression of *Pseudomonas aeruginosa* regulatory RNA and virulence genes. *Proc Natl Acad Sci U S A*. 2006 Jan 3; 103 (1) :171-6. PubMed PMID:16373506; PubMed Central PMCID: PMC1324988.

Vogel J, Wagner EG. Target identification of small noncoding RNAs in bacteria. *Curr Opin Microbiol*. 2007 Jun; 10 (3) :262-70. PubMed PMID:17574901.

Vogel J, Argaman L, Wagner EG, Altuvia S. The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Curr Biol*. 2004 Dec 29; 14 (24) :2271-6. PubMed PMID:15620655.

Vogel J, Luisi BF. Hfq and its constellation of RNA. *Nat Rev Microbiol*. 2011 Aug 15; 9 (8) :578-89. PubMed PMID:21760622.

Vvedenskaya IO, Sharp JS, Goldman SR, Kanabar PN, Livny J, Dove SL, Nickels BE. Growth phase-dependent control of transcription start site selection and gene expression by nanoRNAs. *Genes Dev*. 2012 Jul 1; 26 (13) :1498-507. PubMed PMID:22751503; PubMed Central PMCID: PMC3403017.

Wadler CS, Vanderpool CK. A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proc Natl Acad Sci U S A*. 2007 Dec 18; 104 (51) :20454-9. PubMed PMID:18042713; PubMed Central PMCID: PMC2154452.

Wagner EG, Altuvia S, Romby P. Antisense RNAs in bacteria and their genetic elements. *Adv Genet*. 2002; 46:361-98. PubMed PMID:11931231.

Wang Z, Hartman E, Roy K, Chanfreau G, Feigon J. Structure of a yeast RNase III dsRBD complex with a noncanonical RNA substrate provides new insights into binding specificity of dsRBDs. *Structure*. 2011 Jul 13; 19 (7) :999-1010. PubMed PMID:21742266; PubMed Central PMCID: PMC3143303.

Wassarman KM, Repoila F, Rosenow C, Storz G, Gottesman S. Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev*. 2001 Jul 1; 15 (13) :1637-51. PubMed PMID:11445539; PubMed Central PMCID: PMC312727.

Wassarman KM, Storz G. 6S RNA regulates *E coli* RNA polymerase activity. *Cell*. 2000 Jun 9; 101 (6) :613-23. PubMed PMID:10892648.

Waters LS, Storz G. Regulatory RNAs in bacteria. *Cell*. 2009 Feb 20; 136 (4) :615-28. PubMed PMID:19239884; PubMed Central PMCID: PMC3132550.

Welz R, Breaker RR. Ligand binding and gene control characteristics of tandem riboswitches in *Bacillus anthracis*. *RNA*. 2007 Apr; 13 (4) :573-82. PubMed PMID:17307816; PubMed Central PMCID: PMC1831863.

Wertheim HF, Melles DC, Vos MC, van Leeuwen W, van Belkum A, Verbrugh HA, Nouwen JL. The role of nasal carriage in *Staphylococcus aureus* infections. *Lancet Infect Dis*. 2005 Dec; 5 (12) :751-62. PubMed PMID:16310147.

Winkler WC, Breaker RR. Regulation of bacterial gene expression by riboswitches. *Annu Rev Microbiol*. 2005; 59:487-517. PubMed PMID:16153177.

Woodford N, Livermore DM. Infections caused by Gram-positive bacteria: a review of the global challenge. *J Infect*. 2009 Sep; 59 Suppl 1:S4-16. PubMed PMID:19766888.

Wright AJ, Higginbottom A, Philippe D, Upadhyay A, Bagby S, Read RC, Monk PN, Partridge LJ. Characterisation of receptor binding by the chemotaxis inhibitory protein of *Staphylococcus aureus* and the effects of the host immune response. *Mol Immunol*. 2007 Apr; 44 (10) :2507-17. PubMed PMID:17258808; PubMed Central PMCID: PMC2646901.

Wu S, de Lencastre H, Tomasz A. Sigma-B, a putative operon encoding alternate sigma factor of *Staphylococcus aureus* RNA polymerase: molecular cloning and DNA sequencing. *J Bacteriol*. 1996 Oct; 178 (20) :6036-42. PubMed PMID:8830703; PubMed Central PMCID: PMC178463.

Wurtzel O, Sesto N, Mellin JR, Karunker I, Edelheit S, Bécavin C, Archambaud C, Cossart P, Sorek R. Comparative transcriptomics of pathogenic and non-pathogenic *Listeria* species. *Mol Syst Biol*. 2012 May 22; 8:583. PubMed PMID:22617957; PubMed Central PMCID: PMC3377988.

Xue T, You Y, Hong D, Sun H, Sun B. The *Staphylococcus aureus* KdpDE two-component system couples extracellular K<sup>+</sup> sensing and Agr signaling to infection programming. *Infect Immun*. 2011 Jun; 79 (6) :2154-67. PubMed PMID:21422185; PubMed Central PMCID: PMC3125826.

Yin S, Daum RS, Boyle-Vavra S. VraSR two-component regulatory system and its role in induction of pbp2 and vraSR expression by cell wall antimicrobials in *Staphylococcus aureus*. *Antimicrob Agents Chemother*. 2006 Jan; 50 (1) :336-43. PubMed PMID:16377706; PubMed Central PMCID: PMC1346790.

Yousef MR, Grundy FJ, Henkin TM. Structural transitions induced by the interaction between tRNA(Gly) and the *Bacillus subtilis* glyQS T box leader RNA. *J Mol Biol*. 2005 Jun 3; 349 (2) :273-87. PubMed PMID:15890195.

Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W. Single processing center models for human Dicer and bacterial RNase III. *Cell*. 2004 Jul 9; 118 (1) :57-68. PubMed PMID:15242644.

Zhang K, Nicholson AW. Regulation of ribonuclease III processing by double-helical sequence antideterminants. *Proc Natl Acad Sci U S A*. 1997 Dec 9; 94 (25) :13437-41. PubMed PMID:9391043; PubMed Central PMCID: PMC28323.

Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*. 2003 Jul 1; 31 (13) :3406-15. PubMed PMID:12824337; PubMed Central PMCID: PMC1691



**Cédric ROMILLY**

**Fonction de nouveaux ARN non codant dans la régulation de l'expression des gènes de *Staphylococcus aureus* : Adaptation à l'environnement et virulence**

*Staphylococcus aureus*, pathogène opportuniste de l'homme, est responsable de 30% des infections nosocomiales. La pathogénie de la bactérie résulte de l'expression d'une pléthore de facteurs de virulence, mais quels sont les mécanismes de régulation contrôlant l'expression de ces gènes ? Aujourd'hui, il est clairement établi que les ARN non-codant sont des molécules clés dans la régulation de l'expression des gènes. Plus de 50 ARN ont été identifiés chez *S. aureus*. Néanmoins la fonction de peu d'entre eux est connue. Durant ce travail de thèse, l'étude de la fonction et du mécanisme de régulation des ARN RsaA et RsaE a été entreprise. RsaA est un ARN sous le contrôle du facteur de stress sigmaB. Les résultats obtenus montrent que ce dernier régule la traduction de l'ARNm *mgrA* qui code pour un facteur de transcription important dans l'expression des gènes de virulence et la régulation de l'autolyse. Par appariement de base, RsaA cible l'ARNm en utilisant deux sites distants et coopératifs, permettant une interaction forte qui empêche la traduction de l'ARNm. *In vivo*, la délétion du gène *rsaA* perturbe la synthèse de biofilm de capsule. En régulant la traduction de sa cible, RsaA permet de relier l'adaptation au stress à l'expression des gènes de virulence. De manière plus générale, les réseaux de régulation des ARN se connectent les uns aux autres pour permettre à la bactérie d'intégrer une multitude de signaux provenant du milieu extracellulaire afin de moduler finement l'expression des gènes.

*Staphylococcus aureus* is a versatile and opportunist human pathogen, which is responsible of 30% of nosocomial infections. The pathogenicity of the bacteria results from the expression of numerous virulence factors. An importance focus has been made to understand what triggers virulence genes expression. Regulatory RNAs are important regulators of genes expression in bacteria. In *S. aureus*, there is more than 50 RNAs identified, but there is a lack of investigations about their functions and regulatory networks. The aim of this work was to characterize the function and mechanism of action of RsaA and RsaE RNAs. RsaA is under the control of the stress factor sigmaB. Computational analysis combined with global analysis of the proteome led to the discovery of one target : *mgrA* mRNA, which is a global transcription factor involved in autolysis and biofilm regulation. *In vitro* studies show that RsaA binds efficiently *mgrA* mRNA using two distant and cooperative interaction sites. Binding of RsaA to the mRNA prevents initiation of the translation. *In vivo*, *rsaA* gene deletion shows impact on biofilm and capsule production. By regulating the expression of *mgrA* mRNA, RsaA network is linked to *agr* system and virulence gene expression. In a more general way, this work shows that regulatory RNAs networks allow bacteria to modulate virulence, stress and metabolism gene expression depending on the signals provided by the environment of the bacteria.