

**UNIVERSITE DE STRASBOURG**

**ECOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE**

**THESE**

Présentée  
pour l'obtention du grade de

**DOCTEUR DE L'UNIVERSITE DE STRASBOURG**  
**Domaine: BIOLOGIE MOLECULAIRE ET STRUCTURALE**

Par

**Clément CHEVALIER**

**FONCTIONS ET MECANISMES D'ACTION DE L'ARNIII ET DE  
NOUVEAUX ARN NON CODANTS DE *STAPHYLOCOCCUS AUREUS***

Soutenue le 20 octobre 2009 devant le jury

Docteur Hilde De Reuse	Rapporteur externe
Docteur Marc Dreyfus	Rapporteur externe
Docteur Anne-Catherine Dock-Bregeon	Rapporteur interne
Docteur Gérard Lina	Examineur
Professeur Eric Westhof	Examineur
Docteur Pascale Romby	Directeur de thèse
Professeur François Vandenesch	Membre invité

UPR 9002 du CNRS, Institut de Biologie Moléculaire et Cellulaire, Strasbourg

---

# Sommaire

<b>SOMMAIRE</b> .....	<b>I</b>
<b>ABREVIATIONS</b> .....	<b>V</b>
<b>INTRODUCTION</b> .....	<b>1</b>
<b>I. Généralités sur l'ARN</b> .....	<b>3</b>
<b>II. Les ARN régulateurs eucaryotiques</b> .....	<b>8</b>
A. Les longs ARN non codants .....	9
B. ARN interférence et micro ARN .....	10
<b>III. Les ARN régulateurs bactériens</b> .....	<b>14</b>
A. Les "riboswitch" .....	17
B. ARN régulateurs piégeant les protéines .....	21
B.1. Les ARN CsrB et CsrC et le métabolisme du glucose.....	22
B.2. L'ARN 6S et la régulation du facteur $\sigma^{70}$ .....	24
C. Les ARN régulateurs antisens .....	25
C.1. Généralités .....	25
C.2. Les ARN régulateurs strictement complémentaires .....	28
C.2.1. Cas des ARN plasmidiques CopA et CopT .....	28
C.2.2. L'ARN IstR1 et le mécanisme de "stand by" .....	30
C.3. Les ARN antisens à complémentarité partielle .....	32
C.3.1. La protéine Hfq, un cofacteur essentiel chez <i>E. coli</i> .....	32
C.3.2. L'ARN RhyB régule simultanément de nombreux gènes impliqués dans le métabolisme du fer..	34
C.3.3. L'ARN GcvB, un motif conservé pour réguler de multiples ARNm .....	36
D. Les éléments CRISPR .....	37
E. Nouveaux ARN .....	40
F. ARN régulateurs et virulence.....	41
<b>IV. Régulation des gènes de virulence chez <i>Staphylococcus aureus</i></b> .....	<b>43</b>
A. <i>Staphylococcus aureus</i> .....	43
B. Les facteurs de virulence.....	45
B.1. Les protéines de surface .....	45
B.1.1. Les MSCRAMM .....	45
B.1.2. Les SERAM.....	46
B.1.3. La protéine A .....	47
B.2. Les protéines sécrétées .....	49
B.2.1. Les toxines.....	49
Les toxines à tropisme membranaire.....	49

Les toxines superantigéniques.....	50
B.2.2. Les protéases.....	50
B.2.3. Les autres enzymes.....	51
C. La régulation des facteurs de virulence.....	51
C.1. Les systèmes à deux composants.....	53
C.2. Les régulateurs de la transcription, la famille Sar.....	54
C.2.1. Le facteur SarA.....	54
C.2.2. Les facteurs homologues à SarA.....	54
C.3. Le facteur de stress $\sigma^B$ .....	55
D. Le système agr et la synthèse de l'ARNIII.....	58
D.1. Le système <i>agr</i> .....	58
D.2. Conclusions préliminaires.....	62
<b>OBJECTIFS DE LA THESE.....</b>	<b>63</b>
<b>I. Validation de cibles de l'ARNIII.....</b>	<b>65</b>
<b>II. Identification de nouveaux ARN régulateurs chez <i>S. aureus</i>.....</b>	<b>66</b>
<b>RESULTATS &amp; DISCUSSIONS.....</b>	<b>69</b>
<b>I. L'ARNIII et la régulation des gènes de virulence de <i>S. aureus</i>.....</b>	<b>71</b>
A. Recherche des cibles et validation expérimentale.....	71
B. L'ARNIII réprime la synthèse des facteurs SA1000 et Rot.....	74
C. L'ARNIII réprime la synthèse du facteur SA2353.....	76
C.1. L'ARNIII s'hybride au niveau de la séquence SD de l'ARNm <i>sa2353 in vitro</i> .....	77
C.2. L'ARNIII inhibe la formation du complexe d'initiation de la traduction.....	79
C.3. Conclusion.....	80
D. L'ARNIII réprime la synthèse de la staphylocoagulase.....	81
E. Discussion.....	82
L'ARNIII utilise un mécanisme commun pour réguler de multiples gènes.....	82
L'ARNIII régule la transcription de gènes cibles de manière indirecte.....	84
La relation répression de la traduction-dégradation des ARNm, un rôle pour la RNase III.....	84
<b>II. Les propriétés de la RNase III de <i>Staphylococcus aureus</i>.....</b>	<b>147</b>
A. La RNase III: état des lieux.....	147
B. Est-ce que la RNase III pourrait stabiliser les interactions ARNIII-ARNm ?.....	171
C. La RNase III reconnaît les interactions boucles-boucles.....	172
C.1. Les complexes ARNIII-ARNm <i>rot</i> et ARNIII-ARNm <i>coa</i> .....	172
C.2. La séquence ne semble pas être un déterminant de reconnaissance.....	173
C.3. La topologie de l'interaction boucle-boucle est l'élément de reconnaissance.....	174
D. Discussion.....	176
La topologie des interactions boucles-boucles guide la coupure par la RNase III.....	176
Rôles possibles de la RNase III dans la régulation par l'ARNIII.....	177

---

<b>III. Identification et caractérisation de nouveaux ARN régulateurs chez <i>S. aureus</i>..</b>	<b>181</b>
A. Stratégie expérimentale.....	181
A.1. Recherche de nouveaux ARN non codants .....	181
A.2. Identification des cibles .....	183
B. Etudes des régions intergéniques .....	184
C. L'expression des ARN Rsa est fortement régulée.....	185
D. Caractéristiques structurales des ARN Rsa: découverte d'une signature conservée .....	187
E. Vers la fonction des ARNnc .....	188
E.1. RsaE et le métabolisme.....	188
E.2. Cibles potentielles des autres ARNnc.....	190
F. Conclusions.....	191
Vers la complication des réseaux de régulation .....	191
Le génome de <i>S. aureus</i> cache-t-il d'autres riborégulateurs.....	193
<b>CONCLUSION GENERALE &amp; PERSPECTIVES.....</b>	<b>251</b>
L'évolution de l'ARNIII et ses multiples fonctions.....	253
Une signature possible pour les ARNnc agissant au niveau des ARNm .....	254
Fonctions et mode d'action des ARN Rsa.....	255
Conclusion générale .....	257
<b>BIBLIOGRAPHIE.....</b>	<b>259</b>
<b>ANNEXES .....</b>	<b>285</b>
ANNEXE 1: ROLE DE L'ARN DANS LA REGULATION DE L'INITIATION DE LA TRADUCTION CHEZ LES BACTERIES.....	287
ANNEXE 2: TECHNIQUES DE CARTOGRAPHIE EN SOLUTION DES ARN.....	303
ANNEXE 3: TECHNIQUE DE "TOEPRINT" .....	323
ANNEXE 4: POSTERS .....	343
ANNEXE 5: COMMUNICATIONS SCIENTIFIQUES.....	347

---

# Abréviations

A	Adénine
aaRS	Aminoacyl-ARNt synthétase
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
AIP	Peptide autoinducteur
ARN	Acide ribonucléique
ARNcr	ARN CRISPR
ARNdb	ARN double brin
ARNi	ARN interférence
ARNlnc	Long ARN non-codant
ARNnc	ARN non-codant
ARNr	ARN ribosomique
ARNt	ARN de transfert
ARNt <sup>Tyr</sup>	ARNt tyrosine
C	Cytosine
°C	Degré Celsius
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
<i>E. coli</i>	<i>Escherichia coli</i>
G	Guanine
IGR	Régions intergéniques
K <sub>D</sub>	Constante de dissociation
kDa	Kilo Daldon
<i>L. monocytogenes</i>	<i>Lysteria monocytogenes</i>
miARN	microARN
nM	Nanomolaire
NTP	Nucléosides tri-phosphates
ORF	Phase ouverte de lecture
pb	Paire de bases

---

RBS	Site de fixation du ribosome
RNase	Ribonucléase
RNP	Ribonucléoprotéine
RT	Reverse transcriptase
<i>S. typhimurium</i>	<i>Salmonella typhimurium</i>
SD	Séquence Shine Dalgarno
siARN	Small interfering ARN
snARN	Small nuclear ARN
snoARN	Small nucleolar ARN
<i>S. aureus</i>	<i>Staphylococcus aureus</i>
<i>S. pyogènes</i>	<i>Streptococcus pyogenes</i>
T	Thymine
tyrRS	tyrosyl-ARNt synthétase
U	Uracile
<i>V. cholerae</i>	<i>Vibrio cholerae</i>
$\Delta G^\circ$	Variation d'énergie libre
3'UTR	Région 3' non traduite
5'UTR	Région 5' non traduite

---

# Introduction

## I. Généralités sur l'ARN

De nos jours, de nombreuses observations suggèrent qu'un "monde à ARN" aurait précédé l'apparition des cellules "modernes" dont le fonctionnement est principalement basé sur l'ADN (Forterre et Gribaldo, 2007). Ce sont les virus qui auraient assuré le passage vers l'ADN. Ce dernier offre plus de possibilités en terme d'évolution car il est plus stable que l'ARN, sa réplication est plus fidèle et il permet l'augmentation de la taille du génome en fonction des besoins du monde vivant (Forterre, 2006). Il est probable que durant de longues années, l'ARN était suffisant pour assurer l'ensemble des fonctions nécessaires à la vie. Même si les protéines existaient déjà dans ce "monde à ARN", elles n'exerçaient qu'un rôle de chaperonnes alors que les ARN étaient responsables de la plupart des processus biologiques fondamentaux comme la réplication ou la traduction (Forterre et Gribaldo, 2007). Si l'ARN est capable de maintenir par lui-même des fonctions biologiques majeures, c'est essentiellement grâce à ses propriétés structurales qui lui confèrent une grande dynamique (Cruz et Westhof, 2009; Leontis et Westhof, 2003) ainsi qu'une grande liberté d'action (Sharp, 2009). En effet, l'ARN est capable d'adopter un nombre varié de structures tertiaires correspondant à une fonction particulière. Selon les cas, l'ARN peut masquer une séquence pour la protéger du milieu extérieur, ou bien exposer un motif structural spécifique pour favoriser une interaction directe avec un partenaire. L'ARN peut également adopter des conformations qui lui permettent, entre autre, de séquestrer des ions métalliques et d'acquérir ainsi une activité catalytique.

Ces observations soulignent l'impact que peut avoir l'ARN sur la régulation des gènes en général. Néanmoins, pendant longtemps il a été négligé. C'est en partie grâce aux séquençages des génomes et à l'analyse comparative de ces génomes à travers l'évolution que notre notion de la génétique a été révolutionnée par la découverte d'une myriade de nouveaux ARN.

Dans le monde du vivant, la régulation de l'expression de gènes est un processus complexe qui nécessite la production et la coordination de nombreux facteurs. De manière générale, ces facteurs font partie ou sont issus du dogme général de la biologie selon lequel une séquence d'acide désoxyribonucléique (ADN) est transcrite en acide ribonucléique messager (ARNm) puis traduite en protéine. Dans ce schéma, les ARN de transfert (ARNt) jouent le rôle d'adaptateur entre le code génétique et la synthèse protéique, et les ARN ribosomiques (ARNr) constituent le squelette des deux sous-unités du ribosome nécessaire au mécanisme de traduction (Jacob et Monod, 1961). Durant de longues années l'ARN était donc considéré comme un simple intermédiaire entre l'information génétique et la synthèse protéique. Ce dogme a finalement été rompu au début des années 80 avec la découverte chez *Tetrahymena thermophila* du premier intron possédant une activité catalytique (Kruger et al., 1982). La découverte de ce premier ribozyme a ainsi ouvert de nouveaux horizons et a rapidement conduit à l'identification d'autres ARN catalytiques comme la Ribonucléase P, qui telle une enzyme agit en *trans* sur la maturation en 5' des ARNt (Guerrier-Takada et al., 1983; Marvin et Engelke, 2009), ou les familles d'introns de groupe I et II impliqués dans l'épissage des ARNr (Doudna et Cech, 2002). Plus récemment, la fonction des ARNr et des ARNt a également été élargie. Chez *Thermus thermophilus*, lors de la réaction de transpeptidation, l'ARNr 23S semble directement responsable de l'activité peptidyl-transférase (Noller et al., 2001) et l'ARNt présent dans le site P serait indispensable pour un bon positionnement des deux substrats (Korostelev et al., 2008). Cette dernière observation renforce l'idée que les ribosomes seraient issus du "monde à ARN" dans lequel les ARNr gouvernaient les étapes essentielles du mécanisme de traduction.

Aujourd'hui, il est clairement établi que l'ARN remplit de multiples fonctions biologiques en usant de ses propriétés catalytiques et structurales (tableau 1). Ainsi, plusieurs classes d'ARN non codants (ARNnc) coexistent. Ces ARN ne portent pas l'information nécessaire à la synthèse protéique et plusieurs d'entre eux participent aux processus essentiels "ménagers" de la cellule comme la maturation des ARN, l'épissage ou encore la traduction et le transport des protéines vers la membrane (Mattick et Makunin, 2005).

ARN	Fonction	Facteurs associées	Mode d'action	Références
RNase P	Maturation des ARNt		Catalyse le clivage des ARNt en 5'	(Marvin et Engelke, 2009)
snARN U1	Régulateur nucléaire	Protéines TAF15	?	(Jobert et al., 2009)
snARN U1, U2, U4, U5, U6, U11, U15	Epissage des pré-ARNm	Protéines du spliceosome	Interactions antisens	(Wahl et al., 2009)
snoARN à Boîte C/D	Maturation des ARN des ARNr, ARNt et snARN	Nop56p, Nop58p, Nop1p	Catalyse des réactions de méthylation en 2'O du ribose	(Bachelierie et al., 2002; Reichow et al., 2007)
snoARN à Boîte H/ACA	Maturation des ARN des ARNr, ARNt et snARN	Cbf5p, Gar1p, Nhp2p, Nop10p	Catalyse des réactions de pseudouridylation	(Bachelierie et al., 2002; Reichow et al., 2007)
H/ACA "like"	Ségrégation des chromosomes, maintenance des télomères	Cbf5p, Nap57, Gar1p, Nop10p	Catalyse des réactions de pseudouridylation	(Meier, 2005)
7SK	Régulation de la transcription	CDK9/cyclin T	Inhibe la phosphorylation de l'ARN polymérase II	(Yang et al., 2001)
SL	Adressage des protéines	Protéines de la "signal recognition particule"	Noyau de la "signal recognition particule"	(Neher et al., 2008)
Evf2	Activateur de la transcription	DLX2	Recrute le facteur de transcription DLX2	(Feng et al., 2006)
Antisens Zeb2	Inhibiteur de l'épissage et activateur de la traduction	ARNm Zeb2	Interaction antisens	(Beltran et al., 2008)
Xist	Inactivation du chromosome X	PCR1, PRC2, ADN	Méthylation des histones	(Wutz et Gribnau, 2007; Zhao et al., 2008)
siARN	Défense antivirale, régulation des ARNm, régulation épigénétique	Complexe RISC, DICER	Interaction antisens	(Brodersen et Voinnet, 2009; Filipowicz et al., 2008; Umbach et Cullen, 2009)
miARN	Régulation traductionnelle et dégradations de ARNm	Drosha, Complexe RISC, DICER	Interaction antisens	(Brodersen et Voinnet, 2009; Filipowicz et al., 2008; Umbach et Cullen, 2009)

**Tableau 1:** Quelques exemples d'ARN eucaryotiques impliqués dans la maturation des ARN et dans la régulation de l'expression des gènes.

Une classe d'ARNnc bien connue est représentée par les "small nuclear" ARN (snARN). Chez les eucaryotes, ces ARN participent aux mécanismes d'épissage des pré-ARNm. Les snARN permettent la reconnaissance des sites d'épissage des pré-ARNm et avec l'aide des protéines Sm conduisent à la formation d'une particule ribonucléoprotéique complexe, le spliceosome (Wahl et al., 2009). Cette particule induit une réorganisation structurale du pré-ARNm et apporte l'activité catalytique nécessaire à l'excision de l'intron. Même si le mécanisme d'excision du spliceosome reste encore mal connu, de récents résultats suggèrent néanmoins que le snARN U6 pourrait être responsable de l'activité catalytique en coordonnant l'action de deux ions métalliques (Toor et al., 2008). Les "small nucleolar" ARN (snoARN) constituent une deuxième classe d'ARNnc qui agit notamment sur la maturation des ARNr, ARNt, snARN et même des pré-ARNm. Il existe deux familles distinctes selon leur mode de conservation, les snoARN à boîte C/D et les snoARN à boîte H/ACA. Ces deux familles induisent respectivement une méthylation en 2'-O du ribose et une pseudouridylation sur leurs ARN cibles. Comme les snARN, les snoARN s'associent à diverses protéines pour former les complexes ribonucléiques appelés snoRNP. Dans ces complexes, les snoARN servent de guide pour interagir avec les ARN cibles et les protéines sont responsables de la modification des ARN (Bachellerie et al., 2002; Reichow et al., 2007). Enfin, un dernier exemple est donné par l'ARN 7SL qui constitue le noyau du complexe ribonucléoprotéique de la "signal recognition particle" (SRP). Ce complexe est essentiel pour l'adressage des protéines en cours de synthèse vers la membrane plasmique de la cellule et son mécanisme est conservé aussi bien chez les eucaryotes que chez les bactéries (Halic et Beckmann, 2005; Neher et al., 2008).

Aujourd'hui, il est estimé que plus de 95% des génomes seraient en réalité transcrits (Frith et al., 2005; Mattick et Makunin, 2006). Même si actuellement il est encore difficile de distinguer les ARNnc des ARN non fonctionnels résultant d'une transcription abortive, ces dernières années ont pourtant révélé l'existence de nouvelles classes d'ARN régulateurs capables de moduler directement l'expression de gènes cibles (Storz et al., 2005). De manière générale, ces ARN agissent à différents niveaux de l'expression des gènes par des contrôles épigénétiques, transcriptionnels, post-transcriptionnels ou post-translationnels. Par l'intermédiaire de ces mécanismes, ces ARN influencent non seulement les étapes du développement d'un organisme mais également sa physiologie globale (Mattick et Makunin, 2006). Pour ces différentes raisons, il devient donc indispensable d'identifier les fonctions, de définir leurs modes d'action et leurs réseaux de régulation. Dans la suite de ce manuscrit, je

décrierai brièvement certains des ARN régulateurs eucaryotiques puis je focaliserai essentiellement mon attention sur les ARN régulateurs bactériens. Actuellement, il est suspecté que près de 10 à 20 % des génomes bactériens pourraient générer des ARN régulateurs. Avec l'aide de plusieurs exemples bien précis, je montrerai la grande diversité de ces ARN. Je décrirai leurs mécanismes d'action, leurs propriétés et les facteurs protéiques auxquels ils sont associés.

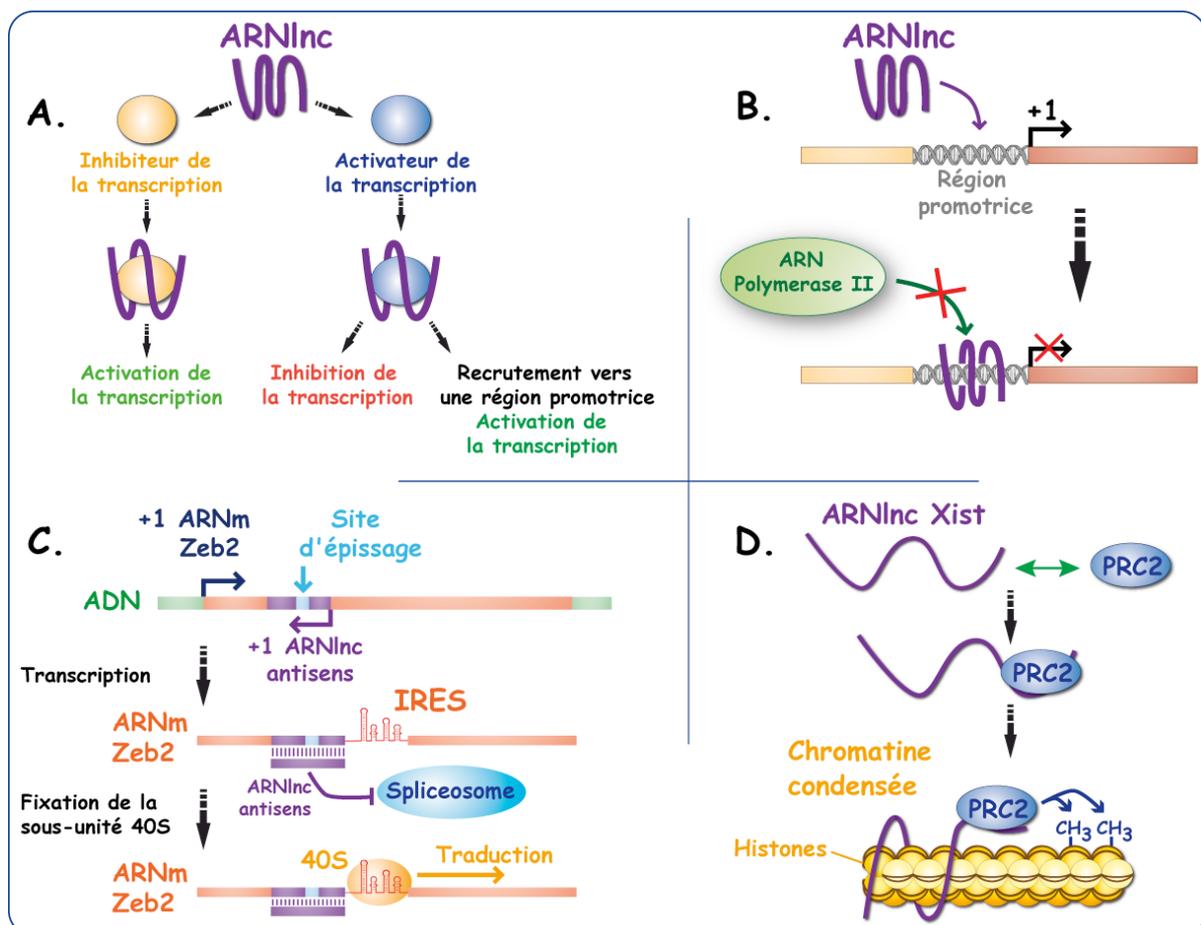
## II. Les ARN régulateurs eucaryotiques

Chez les mammifères, seulement 2% de l'euchromatine est nécessaire à l'expression d'ARNm fonctionnels capable d'initier la synthèse protéique (Consortium, 2004). De plus, l'évolution des génomes eucaryotes montre que la proportion des régions non codantes est en constante augmentation ce qui favorise l'apparition d'ARN régulateurs, notamment au niveau des régions intergéniques (IGR) (Frith et al., 2005; Mattick, 2001; Mattick et Makunin, 2006). Le séquençage des génomes et le développement de nouvelles techniques ont mis en évidence la grande diversité de nouveaux ARNnc. Ces ARNnc, avec les protéines, représentent désormais les acteurs essentiels de la régulation des gènes. Chez les eucaryotes, de nombreux ARNnc influencent directement de nombreux processus cellulaires comme le contrôle qualité ou le transport des ARNm. D'autres sont les effecteurs de régulations au niveau transcriptionnel. Ainsi, l'ARN 7SK régule la transcription des gènes en inhibant l'activité kinase du complexe CDK9/cyclin T qui est nécessaire à la phosphorylation de l'ARN polymérase II (Yang et al., 2001). Enfin, certains de ces ARN participent à la maintenance ou à la ségrégation des chromosomes lors des différentes étapes de cycle cellulaire. Un exemple est donné par un ARN analogue aux snoARN à boîte H/ACA qui est associé au complexe de la télomérase (Meier, 2005). Une mutation dans cet ARN serait responsable d'une maladie génétique liée à un défaut de la télomérase, la dyskératose congénitale (Vulliamy et al., 2001).

Si les ARN régulateurs amènent la spécificité de régulation, ils sont cependant dépendants et/ou associés à des machineries protéiques complexes. En effet, les eucaryotes supérieurs ont l'avantage de posséder un large répertoire de protéines capables de reconnaître non seulement les ARN mais également les molécules d'ADN ainsi que la chromatine (Mattick, 2009; Mattick et Makunin, 2005; Mattick et Makunin, 2006). Les propriétés de ces protéines ont permis le développement de systèmes complexes de maturation et d'adressage des ARN qui offrent de nombreuses possibilités en terme de régulation. Finalement, l'association entre machinerie protéique et ARNnc a donné naissance à des systèmes sophistiqués dans lesquels des interactions ARN-ARN et même ARN-ADN permettent une régulation directe de l'expression de nombreux gènes (Mattick et Makunin, 2005; Mattick et Makunin, 2006). Généralement, ces systèmes font intervenir les longs ARNnc (ARNInc), les "Small interfering"ARN (siARN) et les microARN (miARN)

## A. Les longs ARN non codants

Les ARNlnc sont des ARN faiblement conservés dont la taille est généralement supérieure à 200 nucléotides (Pang et al., 2006). Pendant longtemps ils étaient considérés comme des ARN non fonctionnels résultant d'un bruit de fond transcriptionnel causé par la faible fidélité de l'ARN polymérase II. Ces dernières années ont pourtant révélé que nombre d'entre eux exercent en réalité des rôles importants dans la régulation génique (Mercer et al., 2009; Ponting et al., 2009; Wilusz et al., 2009). D'une manière intéressante, la plupart de ces ARN sont exprimés spécifiquement durant le développement embryonnaire, notamment lors des étapes de différenciation des cellules souches (Amaral et Mattick, 2008; Dinger et al., 2008). Les mécanismes d'action de ces ARN sont diversifiés, ils peuvent agir au niveau transcriptionnel, post-transcriptionnel et même induire des effets épigénétiques.



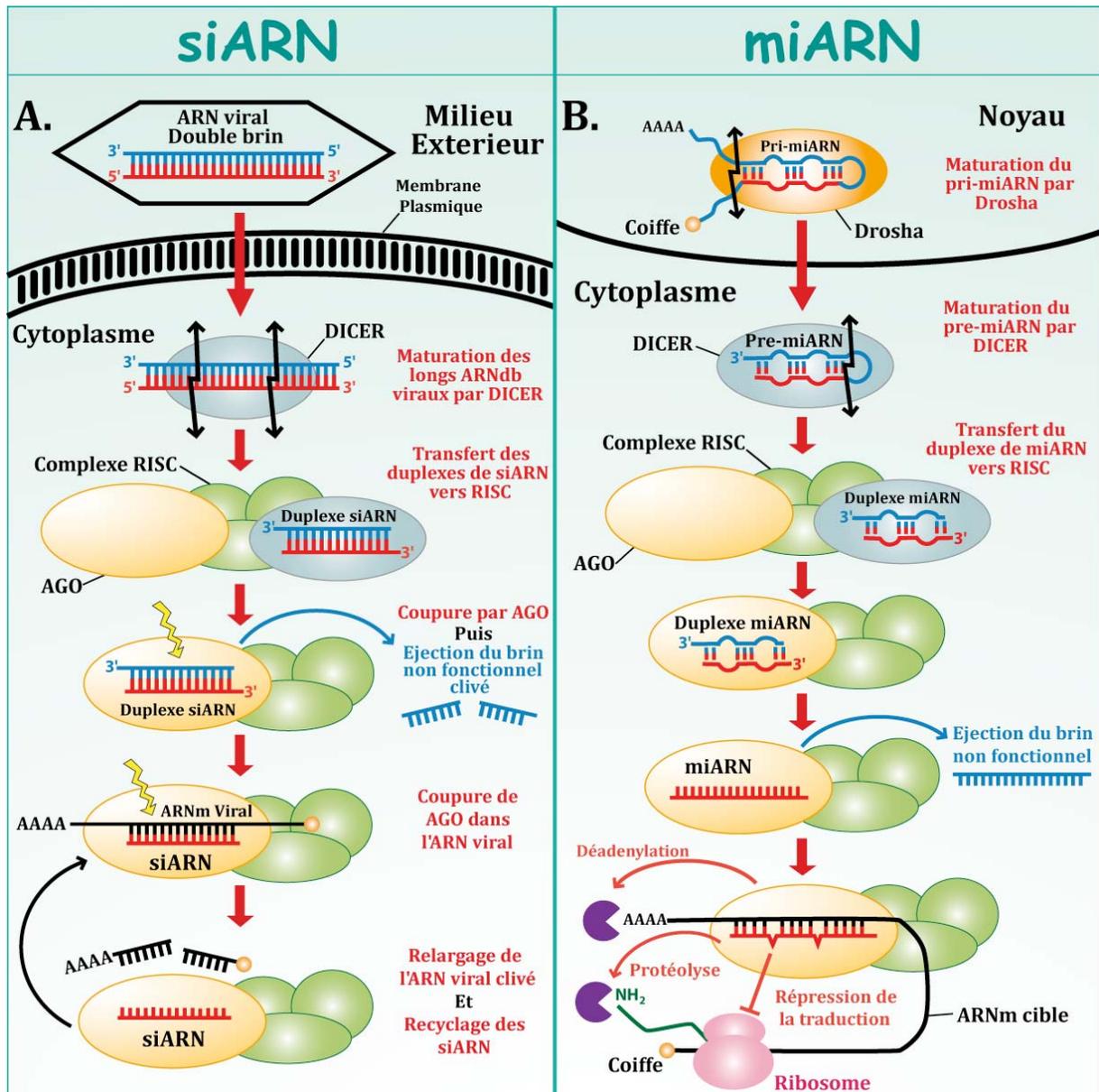
**Figure 1: Mécanismes d'action des ARNlnc.** Les ARNlnc régulent l'expression des gènes à différents niveaux. Ils peuvent induire des effets transcriptionnels en séquestrant des facteurs de transcription ou en interférant avec l'ARN polymérase II (A, B), favoriser la traduction en empêchant l'épissage d'une séquence IRES (C), ou encore induire des effets épigénétiques en recrutant des facteurs qui vont permettre la méthylation des histones (D).

Au niveau transcriptionnel, les ARNlnc peuvent moduler l'expression d'un promoteur en recrutant des facteurs de régulation (Wang et al., 2008), séquestrer directement des facteurs de régulation de la transcription (Feng et al., 2006) ou encore interférer avec l'ARN polymérase II en formant des interactions de type triple hélice au niveau des certaines régions promotrices (Martianov et al., 2007) (figure 1, A et B). Au niveau post-transcriptionnel, les ARNlnc peuvent réguler l'épissage alternatif de certains ARNm. C'est le cas de l'ARNm *Zeb2* qui code pour un répresseur de la transcription et dont l'épissage est lié à un ARNlnc antisens strictement complémentaire. Dans cet exemple, l'ARNlnc antisens interagit directement au niveau des sites d'épissage et bloque totalement l'excision de l'intron (figure 1, C). D'une manière intéressante, l'intron concerné a la particularité de contenir une région importante pour la fixation du ribosome (IRES); l'ARNlnc agit donc indirectement comme un activateur de la traduction de l'ARNm *Zeb2* (Beltran et al., 2008).

Enfin, les ARNlnc sont essentiellement connus pour leurs rôles dans la régulation épigénétique de plusieurs gènes. En effet, certains ont la capacité de recruter des complexes de remodelage de la chromatine afin de réprimer des loci bien précis. Le plus connu de ces ARN est certainement l'ARNlnc *Xist* qui est responsable de l'inactivation d'un des deux chromosomes X chez les femelles. Cet ARN échappe à l'export cytoplasmique puis s'associe au chromosome X à inactiver (Clemson et al., 1996). L'ARN *Xist* recrute ensuite les complexes protéiques PRC1 et PRC2 qui sont responsables de la méthylation des histones (Wutz et Gribnau, 2007; Zhao et al., 2008) (figure 1, D). Ces modifications ont pour effet de compacter la chromatine du chromosome X et mènent à l'extinction de l'expression des gènes. A coté de *Xist*, il existe de nombreux autres ARNlnc présents dans le noyau dont la fonction reste toujours inconnue à ce jour (Mercer et al., 2009). Ainsi, il est tentant de penser que ces ARN puissent posséder des caractéristiques de structure ou de séquence communes qui seraient importantes pour influencer le remodelage de la chromatine.

## **B. ARN interférence et micro ARN**

Au cours de l'évolution, les plantes et les invertébrés ont développé un puissant système immunitaire inné qui répond aux infections virales. Si pendant longtemps ce système de défense était considéré comme unique chez ces organismes, des études réalisées dans les années 90 ont mis en évidence un autre mécanisme de défense insoupçonné, l'ARN interférence (ARNi) (figure 2) (Fire et al., 1998; Napoli et al., 1990).



**Figure 2: Biogenèse et mode d'action des siARN et miARN.** Alors que les siARN sont pour la plupart exogènes, les miARN sont endogènes. Malgré cette grande différence, ces deux types d'ARN utilisent des machineries protéiques similaires pour leur biogenèse ainsi que pour cibler les ARN cibles. Parmi ces protéines se retrouvent notamment DICER, AGO et le complexe RISC.

Lors d'une infection, la plupart des virus à ADN et à ARN génèrent de longs ARN double brin (ARNdb) indispensables au bon déroulement de leur cycle répliatif. Par l'intermédiaire d'un mécanisme impliquant plusieurs facteurs protéiques, l'ARNi utilise cette étape clé du cycle répliatif viral pour affaiblir et détruire les virus (Brodersen et Voinnet, 2009; Umbach et Cullen, 2009). Brièvement, en période d'infection la cellule hôte est envahie par de nombreux ARNdb viraux qui sont rapidement reconnus par DICER, une endoribonucléase (RNase) spécifique des ARNdb. DICER clive les ARNdb et induit la formation de petits duplexes d'ARN de 20 à 22 paires de bases, les siARN. Les siARN sont

ensuite dirigés vers le complexe multiprotéique RISC au sein duquel l'un des deux brins d'ARN est éliminé par l'action d'une protéine Argonaute (AGO). Le brin d'ARN restant est finalement éjecté et guide la protéine AGO vers les ARN viraux avec lesquels il partage une complémentarité stricte. La protéine AGO induit finalement la dégradation des ARN viraux et libère le siARN qui est recyclé vers un nouveau cycle d'ARNi (figure 2). Ainsi, ce mécanisme fournit un moyen de défense antiviral efficace mais il est important de noter que la voie de l'ARNi peut également agir sur d'autres processus. Par exemple, l'ARNi peut participer à la dégradation d'ARNdb endogènes (Hamilton et al., 2002) ou même engendrer des effets épigénétiques en influençant le remodelage de la chromatine (Cam et al., 2009; Zhang et al., 2008).

A côté des siARN, les miARN forment une autre classe importante d'ARN régulateurs eucaryotiques. Les miARN sont des petits ARN endogènes (~21 nucléotides) impliqués dans la plupart des processus biologiques et qui réguleraient au niveau post-transcriptionnel plus de 30% des gènes chez l'homme. Bien que ces ARN soient directement codés par le génome, ils sont maturés par une machinerie similaire à celle des siARN. Leur biogenèse est néanmoins différente (Brodersen et Voinnet, 2009; Umbach et Cullen, 2009) et peut être soumise à une régulation fine par des facteurs protéiques comme la protéine KSRP (Trabucchi et al., 2009). Ces ARN sont tout d'abord transcrits par l'ARN polymérase II sous forme de précurseurs de 80 nucléotides, les pri-miARN. Dans le noyau, ces précurseurs sont ensuite maturés en pré-miARN par la protéine Drosha (RNase spécifique des ARNdb) avant d'être exportés vers le cytoplasme. Une fois dans le cytoplasme, les pré-miARN sont à leur tour maturés en miARN par DICER et rejoignent finalement un système d'adressage proche de celui des siARN (figure 2). Via ce système, les miARN sont guidés vers leurs ARN cibles. Cependant, à la grande différence des siARN, les miARN interagissent le plus souvent de manière imparfaite avec leurs cibles et usent de mécanismes variés, souvent mal connus, pour réguler de nombreux gènes (Brodersen et Voinnet, 2009; Filipowicz et al., 2008). En effet, en plus d'induire la dégradation de leurs cibles via la protéine AGO, les miARN peuvent également bloquer certaines étapes de la traduction (Chendrimada et al., 2007; Pillai et al., 2005). Dans certains cas, ce mode de répression peut être réversible sous l'effet d'un régulateur comme la protéine Hur (Bhattacharyya et al., 2006). Par ailleurs les miARN peuvent également induire la déadenylation des ARNm (Eulalio et al., 2007) ou encore induire la protéolyse du peptide en cours de synthèse (Nottrott et al., 2006). Si on ne peut exclure que les miARN puissent se fixer dans les régions 5' non traduites des ARNm (5'UTR), la plupart d'entre eux semble

interagir au niveau de séquences conservées situées dans les régions 3' non traduites (3'UTR) (Brodersen et Voinnet, 2009). L'impact des miARN dans la cellule pourrait donc enfin expliquer pourquoi chez les eucaryotes, les régions 3'UTR sont souvent longues. En effet, l'action des miARN suggère que les régions 3'UTR seraient tout simplement aussi importantes pour les régulations post-transcriptionnelles que les promoteurs pour les régulations transcriptionnelles, (Bhattacharyya et Filipowicz, 2007; Vasudevan et Steitz, 2007).

### III. Les ARN régulateurs bactériens

Les premiers ARN régulateurs bactériens ont été découverts au début des années 80, bien avant les premiers siARN et miARN eucaryotiques. Ces ARN agissent de manière antisens et régulent notamment le nombre de copies des plasmides et les mécanismes de conjugaison ou de transposition (Simons et Kleckner, 1983; Stougaard et al., 1981; Wagner et al., 2002). C'est seulement un peu plus tard que de tels ARN ont été identifiés dans le chromosome bactérien (Mizuno et al., 1984; Wassarman et al., 1999). Aujourd'hui, près de 10% d'un génome bactérien coderait pour des ARN régulateurs et nombre d'entre eux ont déjà été identifiés, aussi bien chez les bactéries Gram-négatif (*Escherichia coli*, *Salmonella typhimurium*, *Vibrio cholerae*...) que Gram-positif (*Staphylococcus aureus*, *Streptococcus pyogenes*, *Listeria monocytogenes*,...) (Waters et Storz, 2009) (tableau 2). Ces ARN sont très hétérogènes, leur taille varie entre 50 et 500 nucléotides, et leur structure est généralement constituée de plusieurs motifs en tige-boucle nécessaires à leur stabilité et bien souvent à leur fonction.

A ce jour, il existe trois classes distinctes d'ARN régulateurs bactérien. La première classe est représentée par les "riboswitch". Ces ARN situés généralement dans les régions 5'UTR des ARNm sont capables de sonder les variations du milieu extérieur (i.e. température) ou la concentration intracellulaire d'un métabolite ou d'un ion et d'induire en réponse soit l'activation soit la répression de l'ARNm en aval. La deuxième classe d'ARN régulateurs regroupe les ARN qui vont se lier spécifiquement à des protéines régulatrices. Ainsi, ces ARN agissent d'une manière indirecte sur la régulation de nombreux gènes. Enfin, la classe la plus répandue est constituée par des ARN qui interagissent via un mécanisme antisens avec les ARNm.

Les ARN régulateurs confèrent différents avantages à la cellule. D'un point de vue énergétique, il est beaucoup moins coûteux à la bactérie d'utiliser une molécule d'ARN comme effecteur intracellulaire que de passer par la voie classique de la synthèse protéique. Par ailleurs, l'absence de synthèse protéique confère à l'ARN une plus grande rapidité d'action. La stabilité des ARN est également plus facile à moduler que celle des protéines. Ainsi, les ARN régulateurs activés dans des conditions particulières sont généralement très stables afin de ne pas être limitant alors que les ARN transcrits de manière constitutive sont souvent peu stables.

ARN	Bactérie	Cibles	Fonctions	Mécanisme d'action	Références
MicA, MiC, MicF	<i>E. coli</i>	<i>ompA, ompC, ompF</i>	Régulation des protéines de la membrane externe	Inhibition de la traduction et dégradation des ARNm	(Johansen et al., 2006; Rasmussen et al., 2005; Udekwi et al., 2005)
OmrA, OmrB	<i>E. coli</i>	<i>ompt, cirA, fecA, fepA</i>	Régulation des protéines de la membrane externe	Inhibition de la traduction et dégradation des ARNm	(Guillier et Gottesman, 2006)
GcvB	<i>Salmonella Typhimurium</i>	<i>oppA, dppA, gtl, livK, livJ, argT, stm4351</i>	Régulation des protéines de la membrane externe	Inhibition de la traduction et dégradation des ARNm	(Sharma et al., 2007)
RyhB	<i>E. coli</i>	<i>sodB, sdhD, fumA</i>	Métabolisme du fer	Inhibition de la traduction et dégradation des ARNm	(Masse et al., 2007; Masse et al., 2005)
OxyS	<i>E. coli</i>	<i>Rpos, fhla</i>	Stress oxydatif	Inhibition de la traduction	(Repoila et al., 2003)
IstR1	<i>E. coli</i>	<i>TisAB</i>	Système toxine-antitoxine	Inhibition du site de "stand by"	(Darfeuille et al., 2007)
GadY	<i>E. coli</i>	<i>gadX</i>	Réponse au stress acide	Stabilisation de la région 3' de l'ARNm	(Opdyke et al., 2004)
RatA	<i>B. subtilis</i>	<i>tpxA</i>	Système toxine-antitoxine	Dégradation de la région 3' de l'ARNm	(Silvaggi et al., 2005)
6S	<i>E. coli</i>	Facteur $\sigma^{70}$	Réarrangement de l'expression des gènes en phase stationnaire	Séquestration du facteur $\sigma^{70}$	(Wassarman, 2007)
CsrB, CsrC	<i>E. coli</i>	Protéine CsrA	Métabolisme du glucose	Séquestration de la protéine CsrA	(Babitzke et Romeo, 2007)
CRISPR	<i>E. coli</i>	ARN virale, ADN plasmidique	Défense antivirale, conjugaison plasmidique	Appariement antisens	(Barrangou et al., 2007; Brouns et al., 2008; Marraffini et Sontheimer, 2008)

**Tableau 2:** Quelques exemples d'ARN régulateurs bactériens.

Les ARN régulateurs procèdent par des mécanismes diversifiés (tableau 2). Si dans la plupart des cas ils régulent la traduction et/ou la stabilité des ARNm (Marzi et al., 2008; Romby et al., 2006), ils peuvent également agir au niveau transcriptionnel, par exemple en bloquant les effets d'un régulateur de la transcription (Argaman et Altuvia, 2000; Boisset et al., 2007; Repoila et al., 2003). Ces différentes possibilités font que l'ARN est particulièrement approprié pour adapter la croissance bactérienne à divers stress ou à des changements environnementaux. Par exemple, chez les bactéries Gram-négatif, un réseau d'ARNnc régule la synthèse de nombreuses protéines membranaires (Guillier et Gottesman, 2006; Papenfort et al., 2008; Rasmussen et al., 2009) alors que d'autres jouent un rôle

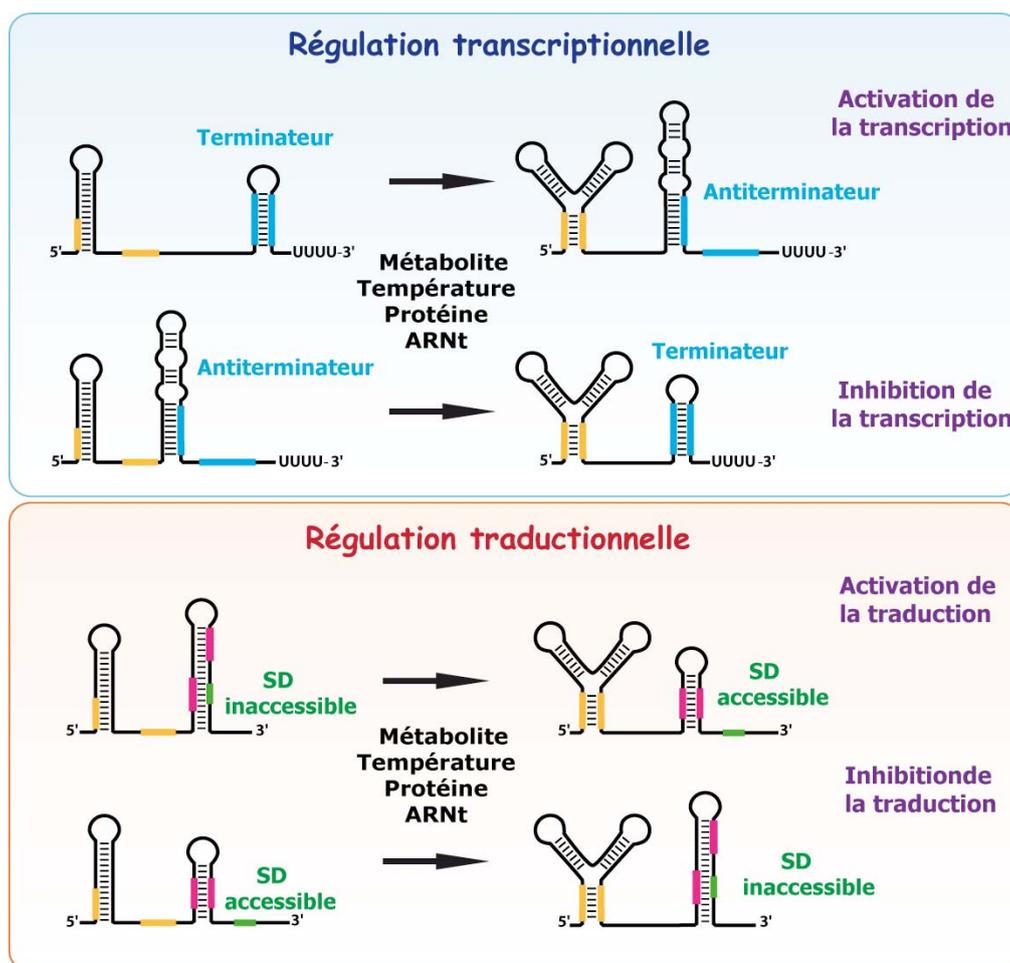
important dans le métabolisme du fer ou dans l'utilisation des sucres (Gorke et Vogel, 2008; Masse et al., 2007). Plus récemment, il a également été montré que certains de ces ARN exercent le rôle d'effecteur intracellulaire de la régulation des gènes de virulence dans des bactéries pathogènes comme *Staphylococcus aureus* ou *Listeria monocytogenes* (Geissmann et al., 2006; Toledo-Arana et al., 2007).

Actuellement, l'un des défis majeurs est d'identifier l'ensemble des ARN régulateurs et leurs mécanismes d'action. Cependant, puisque ces ARN sont très hétérogènes, il n'existe pas d'approche unique pour les détecter. Ainsi, le cumul d'approches expérimentales (clonage "shotgun", puces à ADN, purification par affinité...) mais également bioinformatiques ont notamment conduit à la découverte de plus de 80 ARN chez *E. coli* (Altuvia, 2007; Livny et Waldor, 2007; Mandin et Gottesman, 2009; Vogel et Sharma, 2005). Si chacune de ces techniques possède des avantages et des inconvénients, elles n'ont conduit qu'à l'identification d'un nombre limité d'ARNnc. C'est pourquoi, de nouvelles approches ont été développées pour identifier ces ARN à une large échelle. Par exemple, une étude bioinformatique a permis de prédire plus de 500 ARNnc conservés dans plusieurs génomes simultanément (Livny et al., 2008). Parallèlement, les derniers progrès technologiques comme le "deep sequencing" ou les "tiling arrays" offrent désormais la possibilité d'analyser rapidement l'expression d'un groupe d'ARN en fonction de conditions données. Ces deux dernières techniques ont notamment mené à la découverte de nouveaux ARN régulateurs chez *S. typhimurium* et *L. monocytogenes* (Sittka et al., 2008; Toledo-Arana et al., 2009). Cependant, trouver leurs cibles et donc leurs fonctions est un autre défi. A l'instar des ARNnc, il n'existe pas d'approche de choix pour identifier les cibles. Encore une fois, l'utilisation de stratégies complémentaires est souvent obligatoire (Vogel et Wagner, 2007).

Dans la suite de cette partie, je décrirai plusieurs exemples qui me semblent représentatifs de chaque classe d'ARN régulateurs. Je décrirai leurs modes d'action et leur impact sur la physiologie de la cellule.

## A. Les "riboswitch"

Les "riboswitch" constituent certainement la forme la plus simple des ARN régulateurs et découlent probablement d'un mécanisme primitif de régulation des gènes. Ils ont été décrits pour la première fois au début des années 80, notamment pour leur rôle dans le mécanisme d'atténuation de la transcription de l'opéron tryptophane (Merino et al., 2008). Ces éléments situés dans les régions 5'UTR des ARNm adoptent des conformations alternatives en réponse à la concentration intracellulaire d'un petit ligand (ions, métabolites, ARNt...) ou à un changement de température (Dambach et Winkler, 2009; Grundy et Henkin, 2006). Ces régions 5'UTR sont généralement divisées en deux parties: une région aptamère qui reconnaît l'effecteur et une région appelée plateforme d'expression qui module l'expression de l'ARNm via différentes structures alternatives (figure 3).



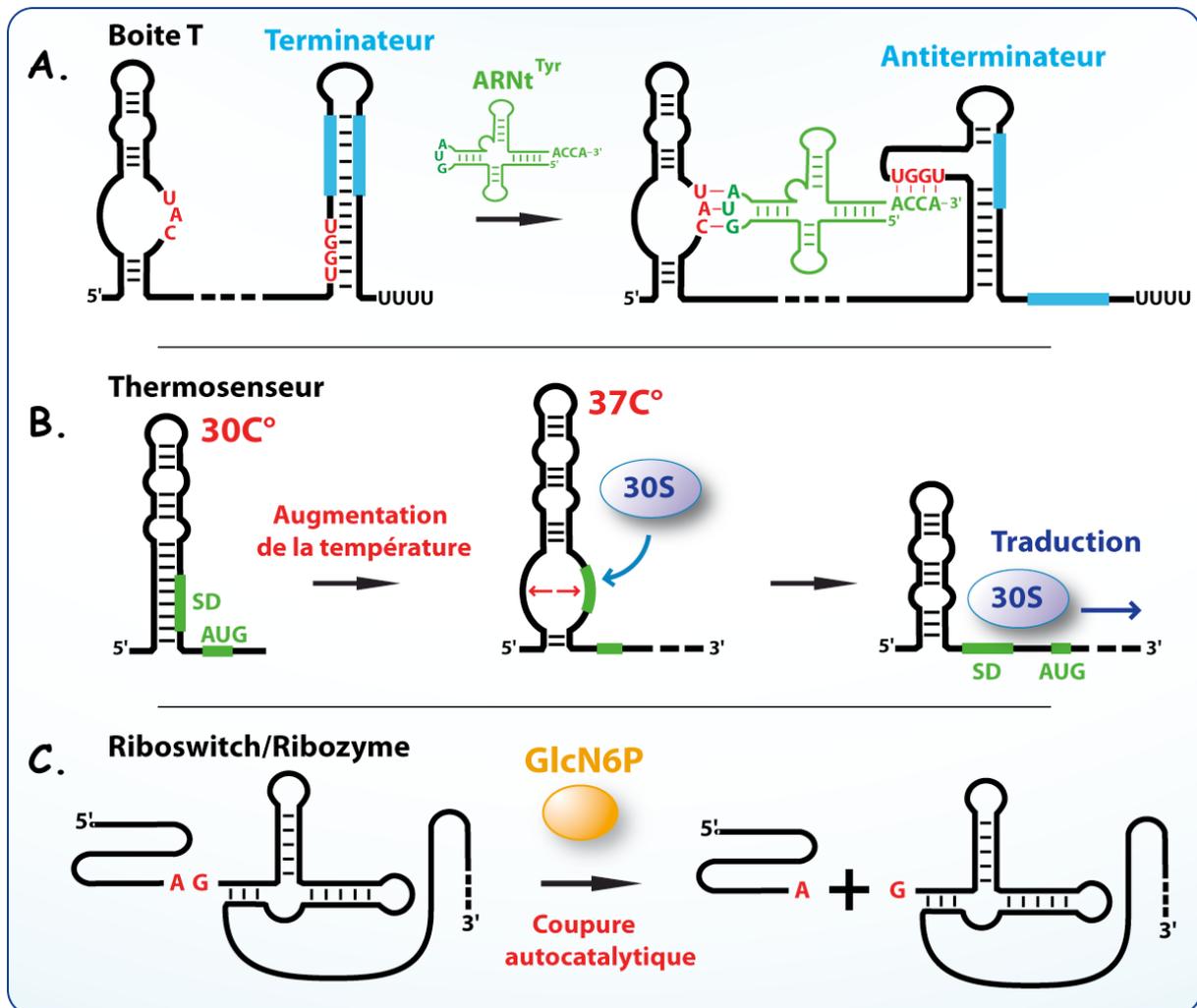
**Figure 3: Régulation de l'expression des gènes par les "riboswitch".** L'action d'un métabolite ou de la température peut réguler les gènes dont l'expression est dépendante d'un "riboswitch". Ce contrôle s'exerce au niveau transcriptionnel en modulant la formation de structures terminatrices, ou au niveau traductionnel en modulant l'accessibilité des sites de fixation du ribosome. Selon les cas, les riboswitch active ou réprime les gènes.

Selon les cas, ces éléments régulateurs agissent positivement ou négativement sur l'expression de l'ARNm (Henkin, 2008; Mandal et Breaker, 2004; Montange et Batey, 2008; Nudler et Mironov, 2004; Winkler et Breaker, 2005). Bien souvent, les "riboswitch" se retrouvent en amont d'opérons qui codent pour les enzymes impliquées dans le métabolisme de leur ligand. Ainsi, ces éléments induisent un rétrocontrôle qui permet d'ajuster à tout moment la concentration des enzymes en fonction de la concentration du ligand intracellulaire donné. Ces systèmes sont largement impliqués dans la régulation des gènes du métabolisme ou de la machinerie traductionnelle (Mandal et Breaker, 2004; Serganov et Patel, 2007; Winkler et Breaker, 2005). Récemment, il a également été montré qu'un même "riboswitch" pouvait se situer en amont de plusieurs opérons différents. Ceci est un moyen efficace pour réguler différentes voies cellulaires de manière coordonnée et à partir d'un seul et même ligand (Sudarsan et al., 2008). Par ailleurs, certaines régions 5'UTR peuvent contenir deux aptamères différents qui répondent à deux effecteurs différents. Ceci permet de répondre à de multiples signaux extérieurs (Dambach et Winkler, 2009; Sudarsan et al., 2006).

Chez les bactéries Gram-positif, de nombreux "riboswitch" appelés boîte T régulent l'expression des aminoacyl-ARNt synthétases (aaRS) en fonction de la concentration intracellulaire d'un ARNt non aminoacylé (Grundy et Henkin, 2003). Chez *Bacillus subtilis* (*B. subtilis*) un exemple est donné par l'ARNt Tyrosine (ARNt<sup>Tyr</sup>) qui reconnaît spécifiquement une boîte T située en amont des gènes codant pour la Tyrosyl-ARNt synthétase (TyrRS) (Mandal et Breaker, 2004). La reconnaissance entre les deux partenaires est guidée par des interactions de type Watson-Crick entre une région située en amont de la boîte T et la boucle de l'anticodon de l'ARNt<sup>Tyr</sup>. Une deuxième interaction intervient ensuite entre l'extrémité 3' de l'ARNt<sup>Tyr</sup> et la boîte T. Ceci a pour conséquence de stabiliser une structure antiterminatrice de la transcription, induisant la transcription du gène *tyrRS* (figure 4, A). Par ce mécanisme, les ARNt<sup>Tyr</sup> non aminoacylés activent la synthèse de la TyrRS et ajustent la concentration de l'enzyme à celle de l'ARNt aminoacylé.

D'autres "riboswitch" qui fonctionnent comme des ARN senseurs sont capables de mesurer des variations de température (Narberhaus et al., 2006; Waldminghaus et al., 2007). C'est le cas de l'ARNm *prfa* qui code pour un régulateur de la transcription impliqué dans l'activation des gènes de virulence chez *L. monocytogenes* (Johansson et al., 2002). A une température de 30°C, la région 5'UTR de cet ARNm se replie sous la forme d'une longue structure en tige-boucle qui séquestre la séquence Shine Dalgarno (SD) et bloque l'initiation de la traduction. Lorsque la bactérie infecte l'hôte, la température du milieu extérieur change

brutalement pour atteindre 37°C. La structure de l'ARNm *pfa* est sensible à ce changement (figure 4, B). En effet, à 37°C des appariements non canoniques présents dans la longue tige-boucle sont déstabilisés. Ceci libère finalement la séquence SD et le ribosome peut initier la traduction du facteur Prfa (Johansson et al., 2002). Grâce à ce mécanisme, les facteurs de virulence de la bactérie ne sont exprimés que lorsqu'ils sont nécessaires, c'est-à-dire au moment de l'infection de l'hôte.



**Figure 4: Exemples de "riboswitch" connus.** (A) Les "riboswitch" à boîte T reconnaissent spécifiquement un ARNt non aminoacylé et activent la transcription de l' aminoacyl-ARNt synthétase correspondante. (B) L'ARNm *Prfa* réagit en fonction de la température. A 37°C, la structure de la région 5' de cet ARN se réorganise et permet la traduction du facteur Prfa. SD indique la séquence Shine Dalgarno (C) Le riboswitch *glmS* reconnaît la glucosamine-6-phosphate et agit à la manière d'un ribozyme pour induire un clivage dans l'ARNm.

De nombreux "riboswitch" sont également connus pour influencer le métabolisme de la cellule (Henkin, 2008; Mandal et Breaker, 2004; Montange et Batey, 2008; Nudler et Mironov, 2004; Winkler et Breaker, 2005). Généralement, au contact d'un ligand spécifique (i.e. vitamines, purines, acides aminés) ces "riboswitch" induisent des réarrangements structuraux qui modulent la transcription ou la traduction des ARNm auxquels ils sont associés. Si cette règle s'applique dans la plupart des cas, il existe cependant quelques "riboswitch" qui possèdent des propriétés différentes. Un exemple connu concerne le "riboswitch" *glmS* que l'on retrouve chez plusieurs bactéries Gram-positif comme *B. subtilis* ou *S. aureus*. Ce "riboswitch" a la particularité de fonctionner comme un ribozyme (Collins et al., 2007). Ainsi, au contact de son ligand, la glucosamine-6-phosphate (GlcN6P), il n'induit pas un réarrangement de structure mais une coupure autocatalytique qui inactive l'ARN *glmS* (figure 4, C). Dans ce mécanisme, la GlcN6P est elle-même un cofacteur nécessaire à l'action du ribozyme (Tinsley et al., 2007). Puisque l'ARN *glmS* code pour l'enzyme nécessaire à la synthèse de la GlcN6P, le riboswitch est directement responsable d'un rétrocontrôle négatif dépendant de la concentration intracellulaire de GlcN6P.

Dans le monde bactérien, si certains "riboswitch" sont bien conservés, il n'est pas rare que leurs mécanismes d'action varient d'une espèce à l'autre (Nudler et Mironov, 2004). Même s'il existe des exceptions, les "riboswitch" agissent généralement au niveau traductionnel chez les bactéries Gram-négatif alors qu'ils régulent préférentiellement la transcription chez les bactéries Gram-positif. Cette différence est probablement liée aux habitats variés des bactéries. Chez *B. subtilis*, près de 4% des gènes sont régulés par des riboswitch "transcriptionnels". La préférence des bactéries Gram-positif à réguler leurs gènes par cette voie est sans doute liée à l'organisation génétique des opérons. En effet, chez ces organismes, les opérons sont souvent très longs. Le "riboswitch" est donc un moyen simple pour la cellule de réguler une multitude de gènes de manière coordonnée. Une autre explication concernant la forte présence des "riboswitch" chez ces bactéries Gram-positif serait le faible nombre d'ARN régulateurs agissant en *trans* chez ces organismes. Cependant, cette observation est certainement biaisée par l'orientation des recherches actuelles qui est majoritairement tournée sur l'étude des bactéries Gram-négatif.

"Riboswitch"	Inducteur	Gènes cible	Type de régulation	Distribution	Références
FMN	Flavin mononucléotide	Métabolisme et transport de la vitamine B2	Transcriptionnelle et traductionnelle	Bactéries gram (-) et gram (+)	(Mironov et al., 2002; Winkler et al., 2002)
TPP	Thiamine pyrophosphate	Métabolisme et transport de la vitamine B1	Transcriptionnelle et traductionnelle	Bactéries, archae, champignons, plantes	(Rentmeister et al., 2007)
B12-box	Adenosylcobalamin	Métabolisme et transport de la vitamine B12	Transcriptionnelle et traductionnelle	Bactéries gram (-), gram (+) et archae	(Nahvi et al., 2002)
SAM	S-Adénosyl méthionine	Métabolisme des dérivés soufrés	Transcriptionnelle	Principalement bactéries gram(+)	(Gilbert et al., 2008; Montange et Batey, 2006)
SAM-B12	Adenosylcobalamin/ S-Adénosyl méthionine	<i>metE</i> , métabolisme de la méthionine	Régulation en tandem de la transcription	Principalement bactéries gram(+)	(Sudarsan et al., 2006)
Glycine	Glycine	Métabolisme de la glycine	Transcriptionnelle	Bactéries gram(+)	(Mandal et al., 2004)
Purine	Adénine, guanine	Métabolisme et transport des purines	Transcriptionnelle	Bactéries gram(+)	(Serganov et al., 2004)
glmS	Glucosamine-6-phosphate	Synthèse de la glucosamine-6-phosphate	Ribozyme	Bactéries gram(+)	(Collins et al., 2007; Tinsley et al., 2007)
di-GMP cyclique	di-GMP cyclique	Virulence, synthèse du flagèle	Transcriptionnelle et traductionnelle	Bactéries gram(-) et gram (+)	(Sudarsan et al., 2008)
T-box	ARNt	Aminoacyl ARNt synthétase	Transcriptionnelle	Bactéries gram(+)	(Grundy et Henkin, 2003)
L-box	Lysine	<i>lysC</i>	Transcriptionnelle	Bactéries gram (-) et gram (+)	(Garst et al., 2008; Serganov et al., 2008)
Prfa	Température	Facteur de régulation de la transcription	Traductionnelle	<i>Lystéria monocytogenes</i>	(Johansson et al., 2002)
FourU	Température	Protéine de choc thermique	Traductionnelle	<i>Salmonella</i>	(Waldminghaus et al., 2007)

Tableau 3: Quelques exemples de "riboswitch" connus.

## B. ARN régulateurs piégeant les protéines

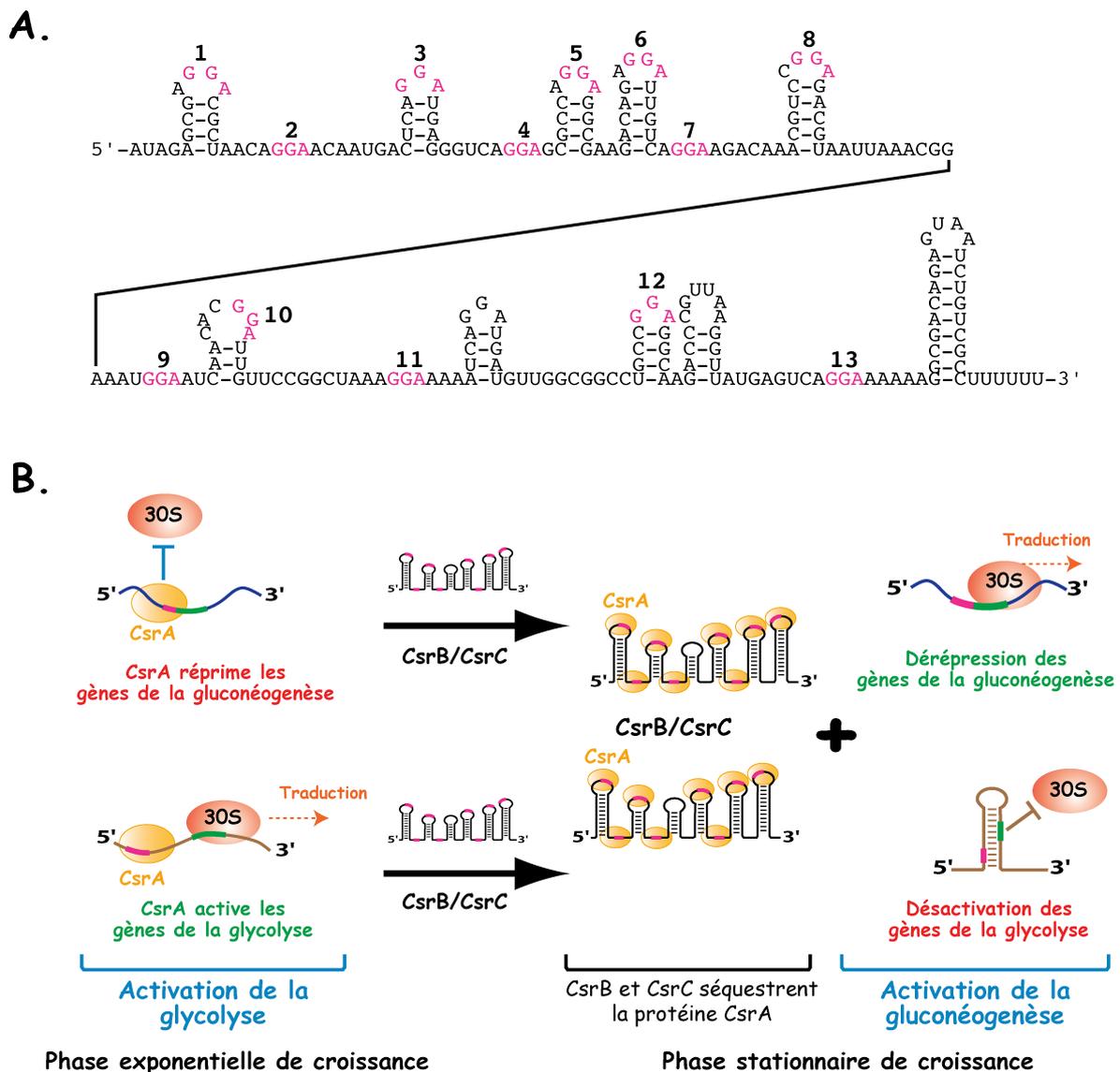
Chez les bactéries, de nombreux ARNm sont régulés au niveau post-transcriptionnel par l'action de protéines. Certaines protéines régulent directement leur propre ARNm en les piégeant dans des complexes ribonucléiques stables. Cette rétroinhibition est souvent observée pour les gènes impliqués dans la traduction comme la thréonyl-ARNt synthétase (Caillet et al., 2003; Torres-Larios et al., 2002) ou les protéines ribosomiques S15, S8 et L20

(Babitzke et al., 2009; Marzi et al., 2007). Dans chacun de ces cas, la structure de l'ARNm mime les sites de reconnaissance des substrats naturels des protéines régulatrices. Dans d'autres circonstances, ce sont les ARN qui vont piéger des protéines. Ces protéines peuvent être des régulateurs post-transcriptionnels (CsrA) ou transcriptionnels (facteur sigma). Piéger une protéine régulatrice confère l'avantage de réguler de manière coordonnée et indirecte de nombreux gènes, et ces effets peuvent être facilement inversés, notamment lorsque la transcription de l'ARN régulateur diminue.

### **B.1. Les ARN CsrB et CsrC et le métabolisme du glucose**

Au cours de leur développement, les bactéries ont besoin de différentes sources d'énergies qu'elles trouvent généralement sous forme de glucose. Si les besoins énergétiques sont variables selon les étapes de croissance, c'est en phase exponentielle qu'ils vont être maximum. A ce stade, les bactéries dirigent toutes leurs ressources de glucose vers la voie de la glycolyse afin de dégager l'énergie nécessaire à leur survie. Une fois la phase stationnaire atteinte, les besoins énergétiques diminuent et les bactéries procèdent à un réajustement de l'expression des gènes pour détourner le glucose vers une voie de stockage, la voie de biogenèse du glycogène encore appelée gluconéogenèse. L'utilisation du glucose est donc finement régulée en fonction des phases de croissance. Chez *E. coli*, c'est la protéine CsrA qui est en grande partie responsable de cette régulation (Romeo et al., 1993). Cette protéine est synthétisée de manière constitutive durant toute la phase de croissance et exerce un double rôle au niveau traductionnel: elle active les gènes impliqués dans la glycolyse et réprime les gènes nécessaires à la gluconéogenèse. A l'entrée en phase stationnaire, la protéine CsrA est régulée par deux ARN régulateurs, les ARN CsrB et CsrC (Babitzke et Romeo, 2007; Weilbacher et al., 2003). Ces ARN possèdent des motifs répétés riches en purine (GGA) localisés à la fois dans des régions en simple brin et dans des motifs en tige-boucle (figure 5, A). Ces motifs (18 sur CsrB et 13 sur CsrC) imitent les sites de reconnaissance situés en 5' UTR des ARNm régulés par la protéine CsrA. Ainsi, CsrB et CsrC inhibent l'action de la protéine CsrA en la séquestrant dans des complexes ribonucléiques compacts (Liu et al., 1997). Ces ARN entrent en compétition directe avec les ARNm régulés en phase exponentielle par CsrA et contrôlent ainsi indirectement la répression de la glycolyse et l'activation de la gluconéogenèse (figure 5, B). De manière intéressante, plusieurs homologues à CsrB et CsrC ont été identifiés chez *Salmonella*, *Pseudomonas* ou *Vibrio*. Ces

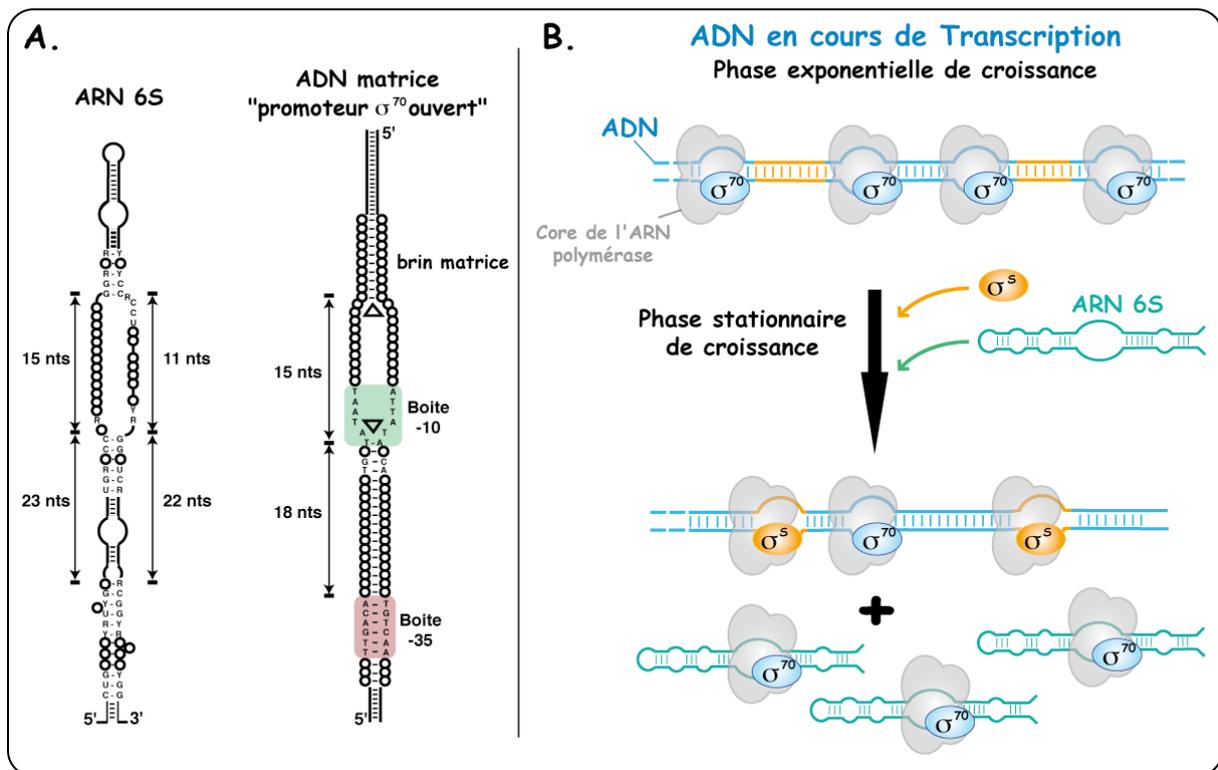
ARN participent à la régulation de protéines homologues à CsrA mais influencent également des processus comme le métabolisme, le *quorum sensing* et la virulence (Lapouge et al., 2008; Lucchetti-Miganeh et al., 2008). Globalement, ces systèmes impliquent à chaque fois plusieurs ARN régulateurs qui agissent par un mécanisme commun. Cette redondance fonctionnelle a certainement pour but de réguler finement la séquestration des protéines mais pourrait également contribuer à maintenir le mécanisme en cas de perte de l'un ou l'autre des ARN régulateurs. Par ailleurs, il n'est pas exclu que certains de ces ARN puissent répondre à des signaux différents.



**Figure 5: Régulation de la protéine CsrA chez *E. coli*.** (A) Structure secondaire de l'ARN CsrC. Les séquences marquées en rose représentent les régions reconnues par la protéine CsrA. (B) Mécanisme d'action des ARN CsrB et CsrC. En phase stationnaire de croissance, les ARN CsrB et CsrC séquestrent la protéine CsrA en mimant des sites de reconnaissance trouvés sur les ARNm cibles. Les régions roses correspondent aux séquences reconnues par CsrA et les régions en vert représentent les sites de fixation du ribosome.

## B.2. L'ARN 6S et la régulation du facteur $\sigma^{70}$

L'ARN 6S a été découvert dans les années 60 (Hindley, 1967) mais sa fonction n'a été identifiée que récemment. Cet ARN est exprimé majoritairement en phase stationnaire de croissance et contrôle l'activité de l'ARN polymérase bactérienne (ARNP) en modérant l'utilisation du facteur  $\sigma^{70}$  (Wassarman, 2007). Cet ARN piège l'ARNP dépendante du facteur  $\sigma^{70}$  (ARNP- $\sigma^{70}$ ) dans un complexe ribonucléoprotéique stable (Gildehaus et al., 2007; Trotochaud et Wassarman, 2005; Wassarman et Storz, 2000). L'ARN 6S agirait par des similarités structurales en imitant le repliement de l'ADN en cours de transcription (figure 6, A) (Gildehaus et al., 2007; Wassarman et Saecker, 2006). Par ce mécanisme, cet ARN fait chuter la concentration cellulaire du facteur  $\sigma^{70}$  et indirectement favorise la transcription des gènes dépendant du facteur  $\sigma^s$  qui sont plus exprimés en phase stationnaire car généralement impliqués dans la réponse au stress ou dans la virulence (figure 6, B) (Trotochaud et Wassarman, 2004). Récemment, il a également été montré que les sites de liaison à l'ARN et à l'ADN du facteur  $\sigma^{70}$  sont chevauchants mais néanmoins bien distincts (Klocko et Wassarman, 2009). Ceci pourrait expliquer pourquoi certains promoteurs  $\sigma^{70}$  sont sensibles à la régulation de l'ARN 6S. Lorsque la concentration en NTP augmente, le positionnement de l'ARN 6S sur l'ARN polymérase change, et l'ARN sert alors de matrice pour la synthèse de petits fragments d'ARN de 14 à 22 nucléotides (Gildehaus et al., 2007; Wassarman et Saecker, 2006). Si la fonction de ces ARN reste encore énigmatique, ils seraient requis pour libérer les facteurs  $\sigma^{70}$  en cas de reprise de la croissance bactérienne (Wassarman et Saecker, 2006). L'ARN 6S est donc un bon moyen pour la cellule de cibler l'expression des gènes en fonction de la croissance bactérienne. D'une manière intéressante, des cellules mutées de l'ARN 6S ne développent aucun défaut de croissance mais leur survie en phase stationnaire est considérablement affectée (Trotochaud et Wassarman, 2005; Wassarman et Storz, 2000). L'ARN 6S joue donc un rôle important dans la physiologie cellulaire et il n'est donc pas surprenant de le retrouver conservé chez d'autres bactéries. Par exemple, chez *B. subtilis*, deux homologues de l'ARN 6S ont été identifiés mais le rôle de ces copies est une question qui reste encore sans réponse (Barrick et al., 2005).



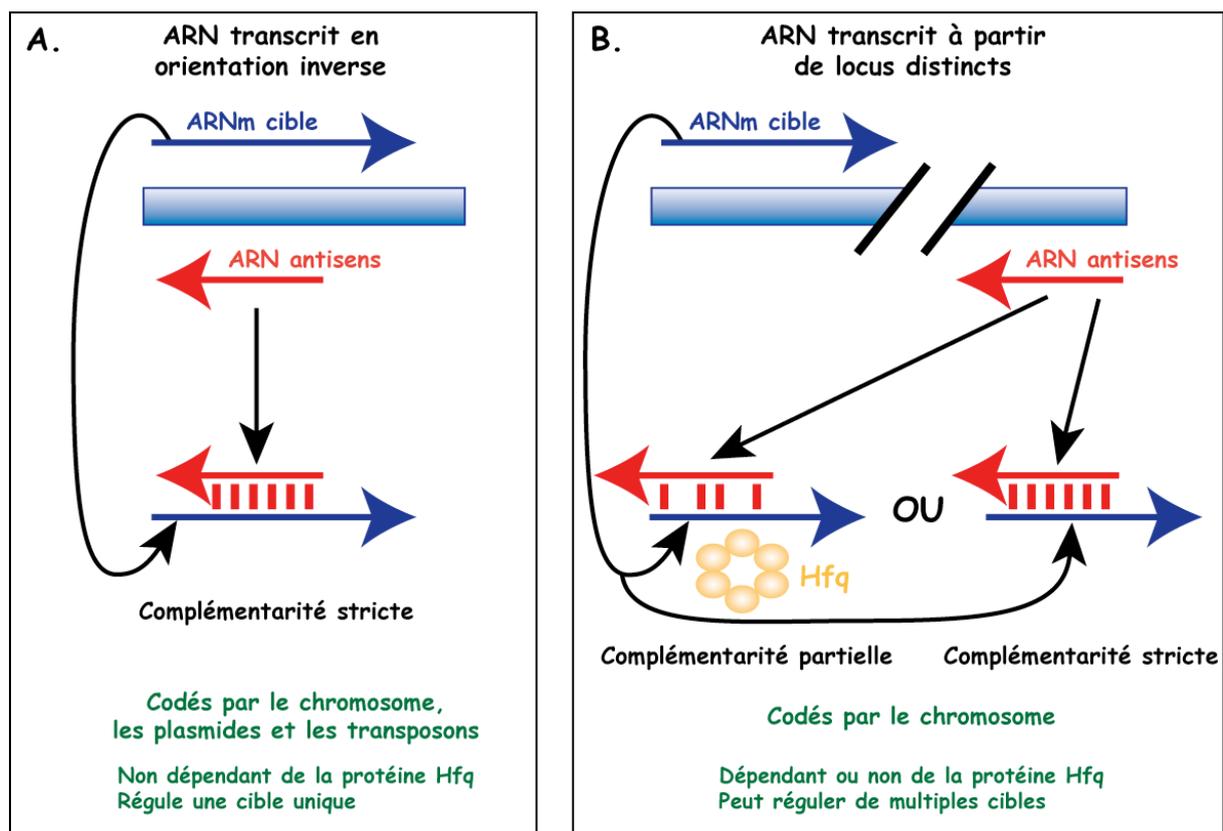
**Figure 6: L'ARN 6S de *E. coli* régule l'utilisation des facteurs sigma en phase stationnaire. (A)** Comparatif entre la structure secondaire de l'ARN 6S et la structure d'un promoteur "ouvert" en cours de transcription. **(B)** Mécanisme d'action de l'ARN 6S. En phase stationnaire, l'ARN 6S séquestre le facteur sigma 70 et favorise l'utilisation du facteur sigma S. Les régions jaunes représentent les promoteurs sigma S.

## C. Les ARN régulateurs antisens

### C.1. Généralités

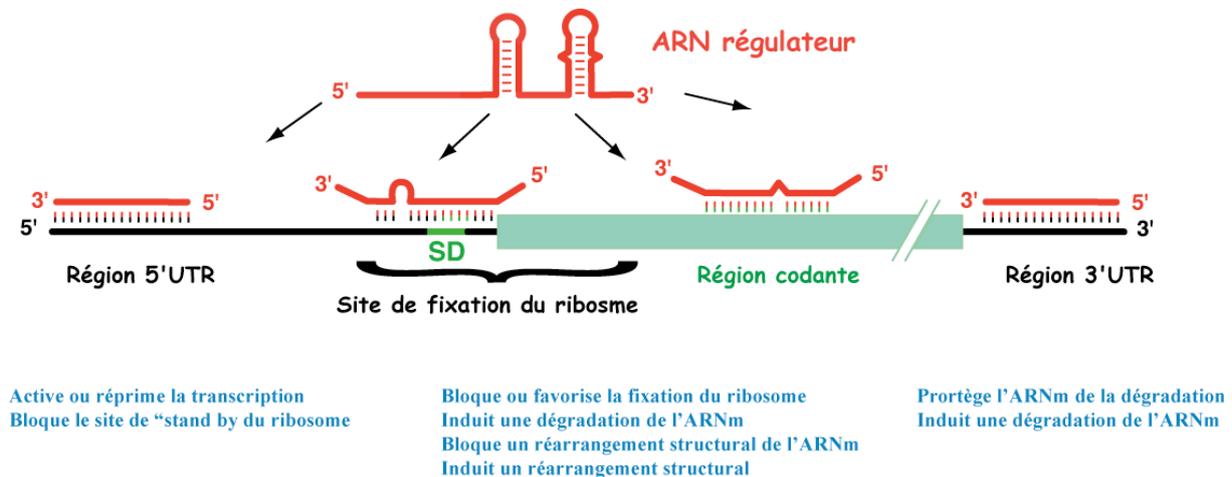
Les ARNnc antisens constituent la classe majeure des ARN régulateurs bactériens. Certains d'entre eux sont directement codés en orientation inverse d'un gène cible et interagissent via des appariements de bases strictement complémentaires avec leur ARNm (figure 7, A). Néanmoins, la plupart de ces ARN est transcrite à partir de loci différents de leurs cibles. S'il existe quelques exemples où ces derniers forment également des appariements parfaits, le plus souvent ils interagissent avec leurs cibles via des interactions partiellement complémentaires (figure 7, B). Ce type d'interaction permet de cibler des séquences variées, il n'est donc pas rare qu'un ARNnc contrôle l'expression de plusieurs gènes simultanément ou qu'un ARNm soit régulé par plusieurs ARN régulateurs différents (Guillier

et Gottesman, 2008; Papenfort et Vogel, 2009). Ainsi, tout comme les facteurs protéiques régulent indifféremment plusieurs promoteurs, les ARN antisens agiraient de manière croisée sur l'expression de nombreux ARNm. D'une certaine manière, ces ARN agissent au niveau post-transcriptionnel comme les facteurs de transcription agissent au niveau transcriptionnel (Bejerano-Sagie et Xavier, 2007; Masse et al., 2007). Par ailleurs, certains ARN peuvent également induire des effets multiples grâce à certaines caractéristiques inhabituelles. Par exemple, l'ARN SgrS de *E. coli* code lui-même pour un petit peptide (Wadler et Vanderpool, 2007). SgrS et son petit peptide sont tous deux impliqués dans l'assimilation du glucose. Ainsi, comme les régions codantes et antisens de SgrS sont distinctes, il est probable que cet ARN ait acquis au cours de l'évolution deux fonctions indépendantes qui sont néanmoins impliquées dans la même voie cellulaire.



**Figure 7: Modes d'action des ARN antisens bactériens.** (A) Les ARN antisens transcrits en orientation inverse de leur gène cible forment des interactions strictement complémentaires. (B) Les ARN antisens transcrits indépendamment de leurs cibles peuvent former des interactions partiellement ou strictement complémentaires.

Les ARN antisens peuvent agir sur la transcription, la traduction et/ou la stabilité d'ARNm cibles, soit comme activateur (Morfeldt et al., 1995; Prevost et al., 2007; Urban et Vogel, 2007), soit comme répresseur (Gottesman, 2005; Vogel, 2009). Les ARN antisens sont très variables en structure et en séquence. Ils possèdent aussi de multiples modes d'action et interagissent de différentes manières avec les ARNm (figure 8).



**Figure 8: Mécanismes d'action des ARN régulateurs antisens bactériens.** Les ARN antisens bactériens peuvent interagir dans les régions 5' ou 3' et dans la phase codante des ARNm. Selon les cas, ces interactions influencent la traduction et/ou la stabilité des ARNm

Le plus souvent, ils ciblent les sites de fixation du ribosome (RBS). Ces régions englobent la séquence SD et le codon initiateur AUG, et sont donc préférentiellement ciblées par les ARN antisens pour réguler la traduction des ARNm (Guillier et Gottesman, 2006; Storz et al., 2004; Vogel et Papenfort, 2006). Certains ARN peuvent également s'hybrider loin en amont du site RBS et réguler la traduction d'un ARNm par compétition directe avec un site de "stand by" du ribosome. C'est le cas de l'ARN IstR1 chez *E. coli* qui régule une toxine impliquée dans la réponse SOS (Darfeuille et al., 2007). Plus récemment, il a également été montré que les ARN antisens ciblent directement la séquence codante d'un ARNm pour induire une dégradation rapide par la RNase E (Pfeiffer et al., 2009; Wagner, 2009). Chez *Salmonella*, l'ARN RyhB interagit aussi dans la région codante proche du codon AUG et réprime la traduction de l'ARNm *ompR* qui code pour une protéine de la membrane externe (Bouvier et al., 2008). Enfin, les ARN antisens peuvent interagir avec la région 3'UTR des ARNm. Ce mécanisme pourrait stabiliser des ARNm possédant une extrémité 3' hétérogène et/ou induire une maturation. C'est le cas de l'ARNm *gadX* impliqué dans la réponse au stress acide chez *E. coli* et dont l'extrémité 3' est stabilisée par l'action coordonnée de l'ARN antisens GadY et de la protéine Hfq (Chen et al., 2002; Opdyke et al., 2004). Ce type de

mécanisme a également été observé dans la bactérie Gram-positif *B. subtilis*. Dans cet exemple, l'ARN RatA régule un ARNm codant pour une toxine et induit la dégradation de son ARNm cible (Silvaggi et al., 2005).

Récemment il a été montré que le phénomène inverse était possible et qu'un ARNnc pouvait être lui-même être réprimé par un ARNm (Figueroa-Bossi et al., 2009). Chez *Salmonella*, l'ARNm *ybfM* code pour une porine qui permet l'entrée de la chitine. En absence de son inducteur, la chitine, cet ARNm est réprimé au niveau post-transcriptionnel par l'ARN antisens MicM. En présence de chitine, le mécanisme de répression est levé dû à l'accumulation d'un ARNm qui s'hybride dans la région 3' de MicM et qui induit sa dégradation en déstabilisant le terminateur de transcription. L'ARNm anti-MicM provient d'un opéron impliqué dans le métabolisme de la chitine dont l'induction est dépendante de ce type de sucre. D'une manière intéressante, à deux nucléotides près, l'ARNm anti-MicM mime la séquence de l'ARNm *ybfM*. Ce mécanisme est particulièrement approprié pour les ARNnc constitutivement exprimés. En effet, le taux d'un ARNnc peut être régulé soit au niveau transcriptionnel (induction par un stress) soit au niveau de sa stabilité. Ainsi, dans cet exemple un ARNm peut être soit le prédateur, soit la proie (Vogel, 2009).

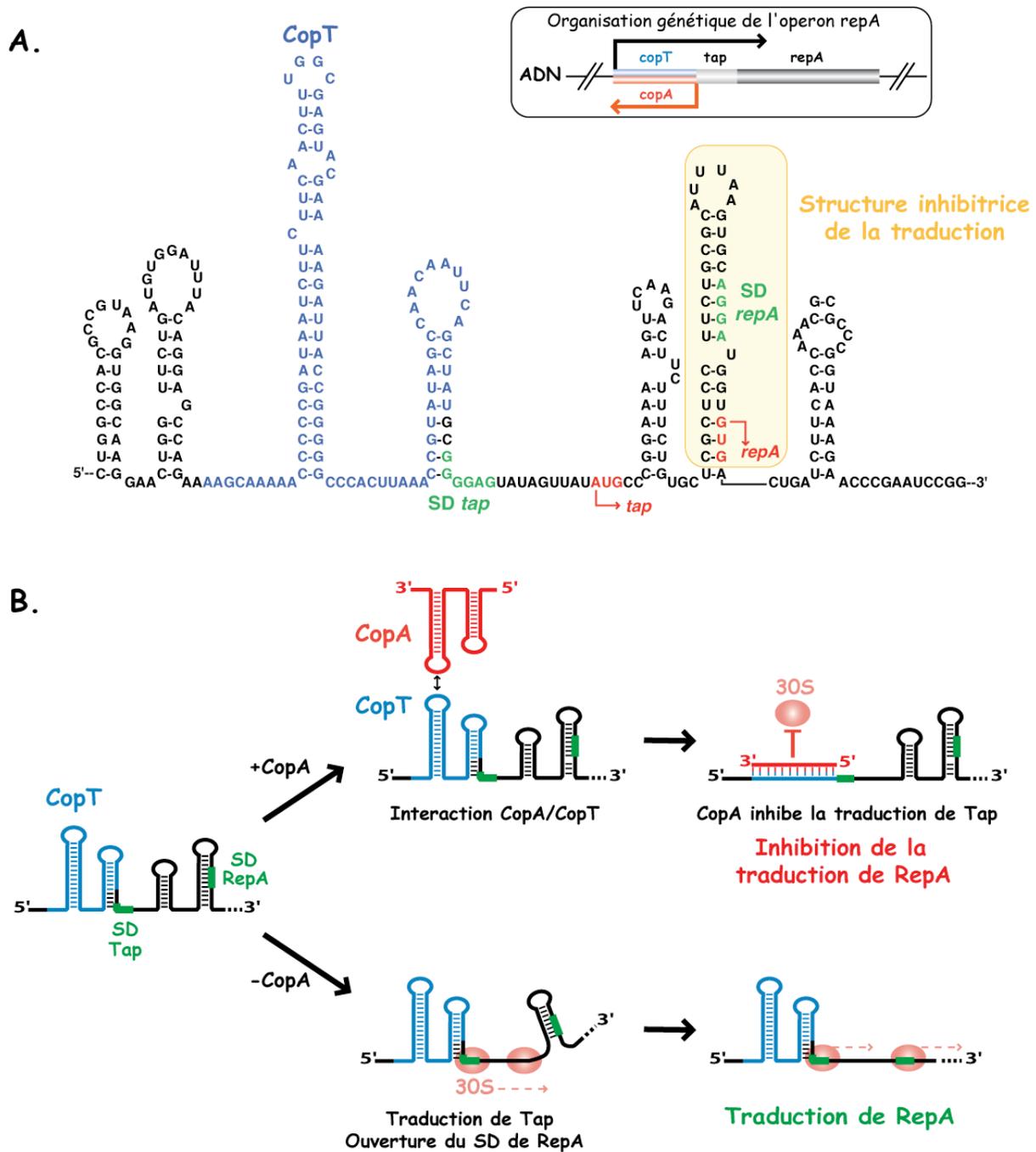
## C.2. Les ARN régulateurs strictement complémentaires

### C.2.1. Cas des ARN plasmidiques CopA et CopT

La plupart des ARN antisens strictement complémentaires à leur cible ont été identifiés dans des éléments mobiles comme les bactériophages, les transposons et les plasmides (Wagner et al., 2002). Ces ARN présentent plusieurs avantages, ils ne nécessitent pas de cofacteurs pour reconnaître leurs cibles et bien souvent ils forment des interactions suffisamment longues pour recruter la RNase III et ainsi induire la dégradation de l'ARNm.

Les plasmides sont des éléments génétiques extra-chromosomiques qui apportent à la bactérie de nouveaux gènes. Ils sont notamment responsables de la résistance croissante aux antibiotiques de nombreuses bactéries pathogènes. La régulation du nombre de copies des plasmides est un processus finement régulé car un taux trop haut peut entraîner des dommages irréversibles pouvant aller jusqu'à la mort cellulaire. A l'inverse, un trop faible nombre de copies peut conduire à la perte du plasmide (Wagner et Brantl, 1998). Ainsi, pour maintenir un nombre constant de copies, les plasmides utilisent des ARN antisens comme effecteurs

principaux du contrôle de la réplication, de la conjugaison ou de la ségrégation plasmidique (Brantl, 2002). Généralement ces ARN sont constitutivement synthétisés mais rapidement dégradés, ainsi leur concentration est à tout moment représentative de celle du plasmide.



**Figure 9: CopA, un ARN strictement complémentaire à CopT. (A) Structure secondaire et organisation génétique de l'ARN *repA*. (B) Mécanisme d'action de l'ARN CopA. En absence de CopA, la traduction du peptide Tap libère la séquence SD de RepA et permet au ribosome d'initier la synthèse protéique. En présence de CopA, le peptide Tap n'est plus traduit ce qui bloque également la traduction de RepA.**

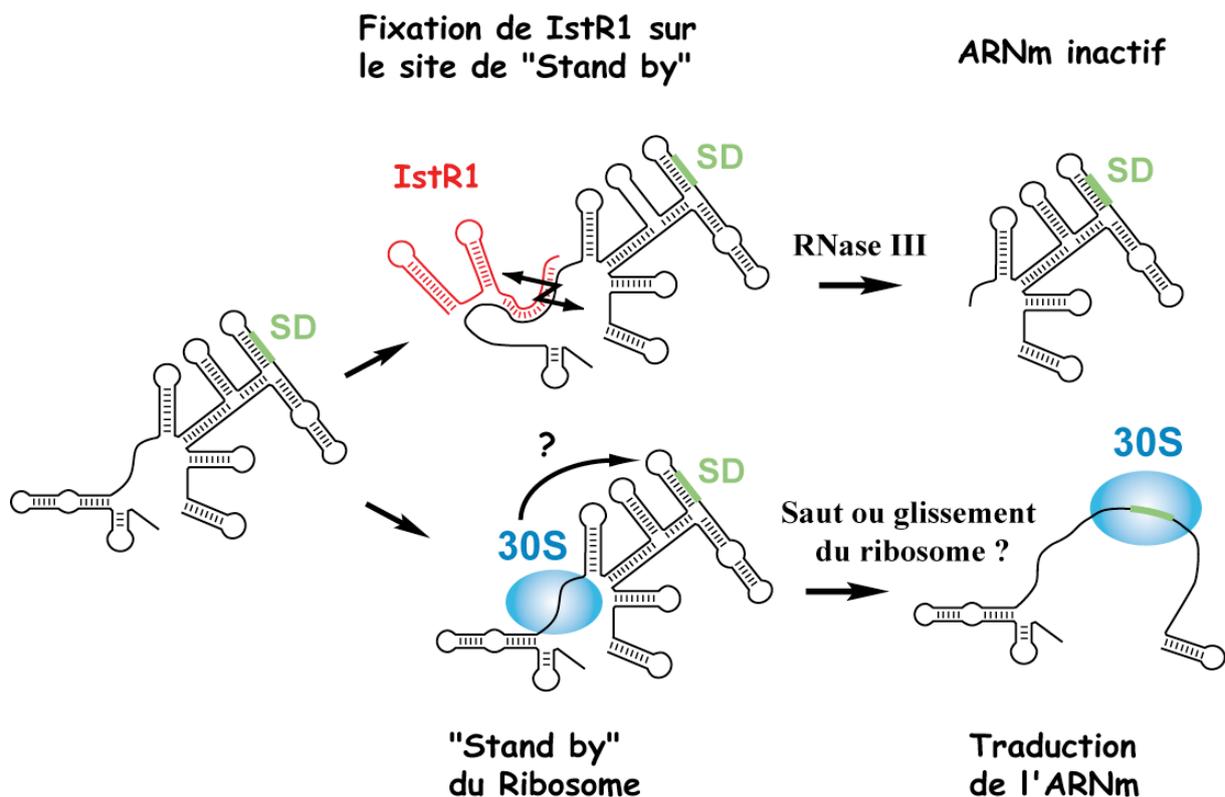
Un système bien connu concerne le mécanisme de réplication du plasmide R1 chez *E. coli*. C'est un plasmide à faible nombre de copies (1 à 2 copies par cellule) dont la réplication est initiée par la protéine RepA. La synthèse de RepA est régulée au niveau traductionnel par l'ARN antisens CopA. CopA est transcrit à partir du même locus que la protéine RepA mais en orientation inverse, il est donc strictement complémentaire à une séquence située en amont de RepA, la séquence CopT. La synthèse de la protéine RepA nécessite la traduction d'un petit peptide Tap codé entre CopT et *repA* (Figure 9, A) (Blomberg et al., 1992). En effet, lors de la synthèse du peptide Tap, le ribosome détruit une structure secondaire qui bloque l'accessibilité de la séquence SD du gène *repA*. Indirectement, c'est donc la traduction du petit peptide Tap qui permet au ribosome d'initier la traduction de RepA (Blomberg et al., 1994). L'interaction entre CopA et CopT inhibe la traduction de Tap. En conséquence, puisque les traductions de *tap* et de *repA* sont couplées, CopA est indirectement responsable de l'inhibition traductionnelle de *repA* (Figure 9, B) (Malmgren et al., 1996).

La structure des ARN CopA et CopT est déterminante pour leur activité *in vitro* et *in vivo* (Kolb et al., 2000a). L'appariement entre les deux ARN est initié par une interaction boucle-boucle riche en GC puis converti sous la forme d'une jonction à quatre hélices avant d'aboutir à la formation d'un long duplexe étendu. L'efficacité du mécanisme d'inhibition est reliée à la vitesse d'association des deux ARN (Wagner et Simons, 1994). Plusieurs résultats ont révélé que la structure des tiges-boucles (séquences riche en GC, nucléotides non appariés dans les hélices) est essentielle à la reconnaissance initiale et à la formation rapide de la jonction à quatre hélices (Hjalt et Wagner, 1995; Kolb et al., 2000b; Kolb et al., 2001). Cette topologie conduit à la formation d'une hélice intermoléculaire de 30 paires de bases (pb), stabilisant le complexe de manière irréversible. Ces règles de base s'avèrent être vérifiées pour de nombreux ARNnc.

### **C.2.2. L'ARN IstR1 et le mécanisme de "stand by"**

Chez les bactéries, certains ARN régulateurs répriment spécifiquement des ARNm codant pour des toxines. Ces systèmes toxine-antitoxine sont utilisés pour éviter à la bactérie de perdre des gènes importants codés par des plasmides, pour réguler la réponse SOS ou différents stress (Fozo et al., 2008; Gerdes et Wagner, 2007). Un exemple connu chez *E. coli* concerne l'opéron *tisAB* qui est sous le contrôle de la réponse SOS en cas de dommages au

niveau de l'ADN. Cet opéron code pour l'ARN régulateur IstR1 mais également pour une toxine qui inhibe la croissance cellulaire. Alors que IstR1 est exprimé de manière constitutive, la toxine est fortement inhibée au niveau transcriptionnel par la protéine LexA. Cependant, cette répression transcriptionnelle n'est pas totale et quelques molécules d'ARNm parviennent à être synthétisées. Ces dernières sont finalement réprimées par l'ARN IstR1 qui forme un long hybride strictement complémentaire à l'ARNm *tisAB*. Cette interaction a pour effet de bloquer la traduction de l'ARNm et induit finalement une dégradation par la RNase III. En cas de réponse SOS, la répression médiée par LexA est suspendue ce qui entraîne une synthèse importante de l'ARNm *tisAB*. Une partie des ARNm va alors échapper à IstR1, induisant la synthèse de la toxine. Ainsi, l'action de la toxine induit un ralentissement de la croissance ce qui donnerait le temps nécessaire à la cellule pour récupérer des dommages subits au niveau de l'ADN (Vogel et al., 2004).



**Figure 10:** L'ARN IstR1 réprime la traduction de l'ARNm *tisAB* par compétition directe avec le site de "stand by" du ribosome (d'après Darfeuille et al., 2007). En rouge est représenté l'ARN IstR1, 30S correspond à la petite sous unité du ribosome, et SD à la séquence Shine Dalgarno.

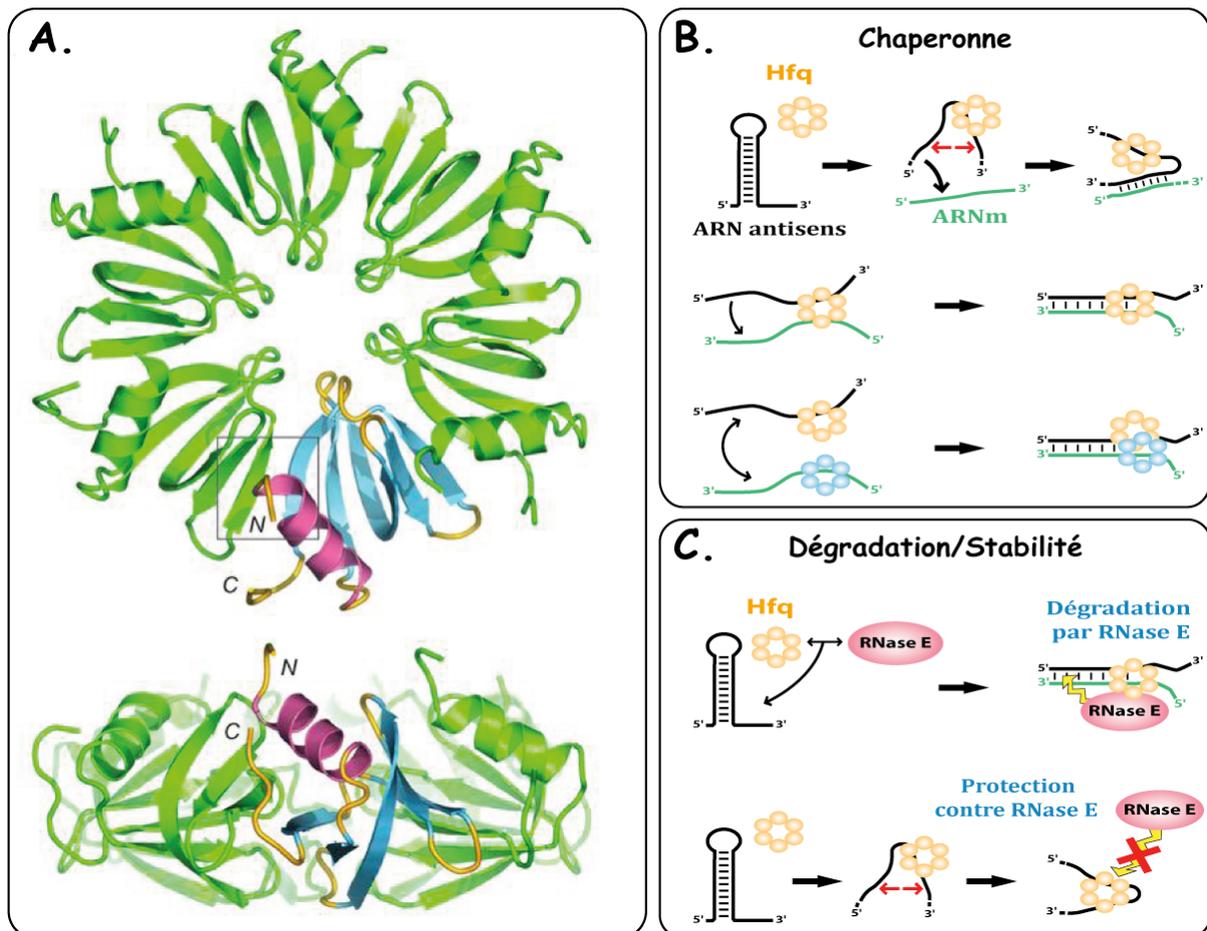
La particularité de cet exemple concerne la régulation traductionnelle de l'ARNm *tisAB*. Dans ce mécanisme, le rôle de l'ARN IstR1 n'était pas évident car des données *in vitro* montraient que le RBS de l'ARNm était séquestré dans une structure en tige-boucle qui devait inhiber l'initiation de la traduction (Darfeuille et al., 2007). En effet, en comparant le  $\Delta G^\circ$  de la structure RBS de l'ARNm *tisAB* avec le taux de diffusion et la concentration des ribosomes libres dans la cellule, il est virtuellement impossible que le ribosome puisse traduire l'ARNm *tisAB* (de Smit et van Duin, 2003). Pourtant, la traduction est efficace. Ce paradoxe a finalement été expliqué par le modèle de "stand by" du ribosome (figure 10) (de Smit et van Duin, 2003). Dans ce modèle, le ribosome interagit avec l'ARNm au niveau d'une région en simple brin, appelée site de "stand by", situé loin en amont du site RBS. Grâce à cette interaction, la diffusion du ribosome n'est plus limitante et le ribosome peut alors glisser ou sauter sur le site RBS afin d'initier la traduction. C'est ce mécanisme qui a été proposé pour l'ARNm *tisAB* (Darfeuille et al., 2007). L'ARN IstR1 réprime la traduction de l'ARNm *tisAB* en s'hybridant au niveau du site de "stand by". Cette interaction bloque ainsi complètement l'accès du ribosome à la région RBS et induit la dégradation rapide de l'ARNm par la RNase III (Darfeuille et al., 2007). Cette étude a ainsi révélé un nouveau mécanisme de recrutement du ribosome par l'ARNm et a montré qu'une interaction éloignée du site RBS peut servir de point d'ancrage à la sous unité 30S pour initier la traduction.

### **C.3. Les ARN antisens à complémentarité partielle**

#### **C.3.1. La protéine Hfq, un cofacteur essentiel chez *E. coli***

Dans bien des cas, les ARN antisens bactériens à complémentarité partielle reconnaissent leurs cibles avec l'aide de la protéine Hfq (Soper et Woodson, 2008; Udekwu et al., 2005). Cette protéine homologue aux protéines Sm eucaryotiques est extrêmement conservée dans de nombreuses bactéries (Axmann et al., 2005; Sun et al., 2002). C'est une protéine hexamérique de 11 kDa dont la structure tridimensionnelle se replie d'une manière caractéristique en forme d'anneau (figure 11, A) (Nikulin et al., 2005; Sauter et al., 2003). Chaque face de l'hexamère possède des domaines indépendants de liaison à l'ARN et la région C terminale semble particulièrement importante pour fixer les ARNm (Brennan et Link, 2007; Vecerek et al., 2008). Grâce à ces différentes propriétés, la protéine Hfq est capable de reconnaître deux molécules d'ARN simultanément. Plusieurs modes d'actions ont été proposés

selon les systèmes étudiés (figure 11, B et C). Tout d'abord, la protéine Hfq favoriserait les interactions de type ARNnc-ARNm en agissant à la manière d'une chaperonne (Aiba, 2007; Valentin-Hansen et al., 2004). Ainsi, elle peut déplier un ARN fortement structuré afin de favoriser des interactions avec un ARN antisens ou alors servir de guide aux ARN pour faciliter une interaction entre un ARNnc et sa cible (Hopkins et al., 2009). Selon les cas, une ou deux molécules d'Hfq sont nécessaires pour associer deux ARN. En outre, la protéine Hfq est également impliquée dans la stabilité et la dégradation des ARNnc. D'une part, elle peut protéger les ARNnc de la dégradation en masquant les sites de coupures de la RNase E. Elle favorise ainsi le maintien des ARNnc à un niveau élevé dans la cellule. D'autre part, chez *E. coli*, la protéine Hfq permet la formation d'un complexe ARNnc-Hfq-RNase E ce qui entraîne la dégradation de l'ARNm cible (Moll et al., 2003; Morita et al., 2005).

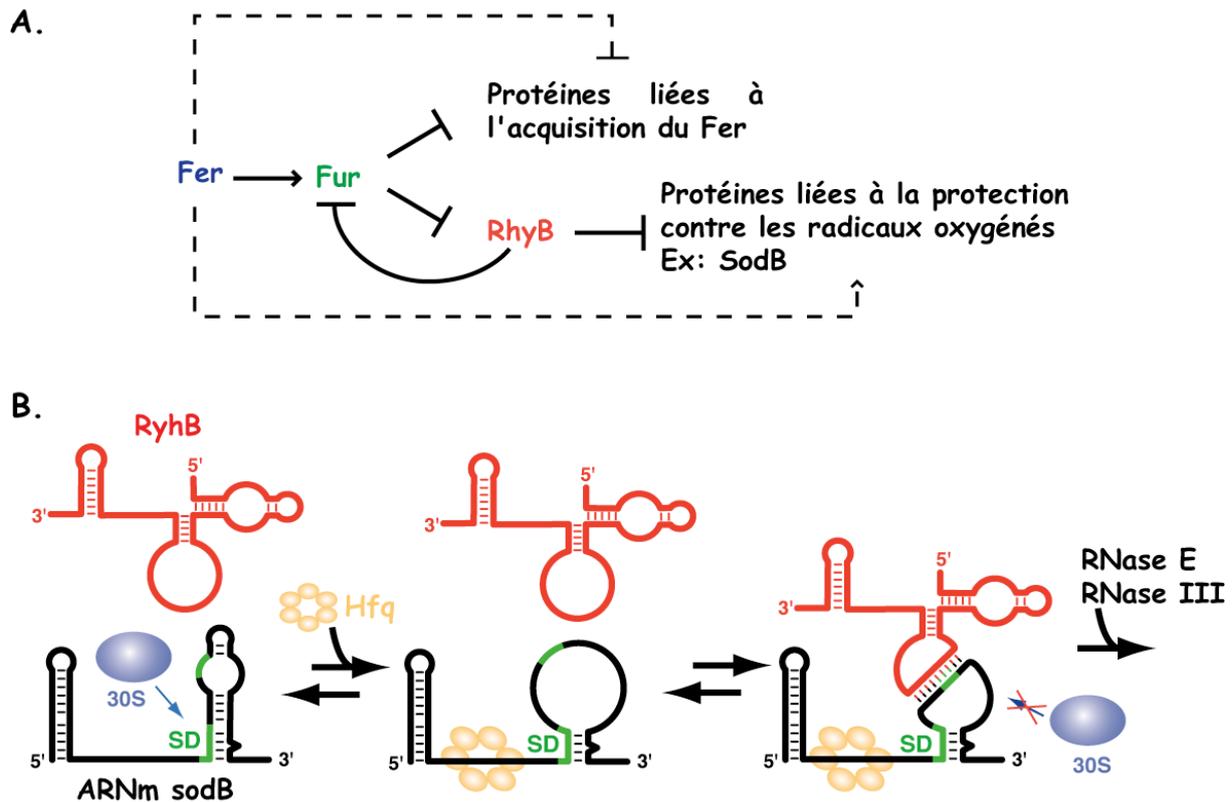


**Figure 11: Structure et fonction de la protéine Hfq.** (A) Structure cristallographique de la protéine Hfq d'après Sauter et al., 2003. N et C représentent respectivement les régions N terminale et C terminale de la protéine. (B et C) Mécanismes d'action de la protéine Hfq. Soit la protéine Hfq agit comme chaperonne en facilitant les interactions ARN-ARN, soit elle joue sur la stabilité des ARNnc en modulant l'action de la RNase E.

Ainsi, la protéine Hfq participe activement à la régulation post-transcriptionnelle des gènes en modulant l'activité et la stabilité des ARNnc antisens. Il est cependant important de noter que sa présence n'est pas toujours indispensable. En fait, Hfq joue un rôle essentiel lorsque l'ARN régulateur et sa cible forment des appariements de bases limités de 8 à 10 nucléotides et donc peu stables. Si les ARN forment de longs hybrides, si la concentration intracellulaire de l'ARN régulateur est forte et si les structures des ARN sont appropriées, les interactions antisens peuvent se former spontanément. C'est par exemple le cas de l'ARN VrrA chez *V. Cholerae* (Song et al., 2008).

### **C.3.2. L'ARN RhyB régule simultanément de nombreux gènes impliqués dans le métabolisme du fer**

Chez les bactéries, l'acquisition du fer est un processus indispensable, que ce soit pour la croissance ou pour la survie au sein d'un hôte. Cependant, quand il est en excès, le fer s'oxyde et génère des radicaux hydroxyles OH<sup>•</sup> qui dégradent les molécules d'ARN et d'ADN. Au dessus d'une certaine concentration seuil, ce phénomène peut même entraîner la mort cellulaire. Pour éviter cela, les bactéries doivent continuellement mesurer et doser la présence du fer afin de maintenir un niveau suffisant pour les besoins vitaux de la cellule. Chez *E. coli*, l'entrée et l'utilisation du fer sont régulées par l'action coordonnée de la protéine Fur ("Ferric uptake regulator") et de l'ARN RyhB (figure 12, A) (Masse et al., 2007). La protéine Fur sonde la concentration intracellulaire en fer et ajuste l'expression des gènes impliqués dans le métabolisme du fer. Cette protéine agit sur l'activation et la répression de nombreux gènes, soit directement en stimulant des régions promotrices (Escolar et al., 1999; Hantke, 2001), soit indirectement en modulant l'expression de l'ARN non codant RhyB. L'ARN RhyB est un régulateur post-transcriptionnel essentiel. Non seulement il réprime la protéine Fur mais également plusieurs opérons (~50 gènes) dont certains sont activés au niveau transcriptionnel par Fur (Desnoyers et al., 2009; Masse et Gottesman, 2002; Mitarai et al., 2009; Vecerek et al., 2007). Plusieurs orthologues de RhyB ont été détectés dans d'autres bactéries Gram-négatif comme *S. typhimurium* et *Pseudomonas*. Ainsi, la protéine Fur et l'ARN RhyB forment le cœur d'un circuit de régulation dynamique qui est indispensable en cas de stress lié à la concentration intracellulaire du fer.



**Figure 12: L'ARN RyhB et le métabolisme du Fer.** (A) Chez *E. coli*, le métabolisme du Fer est finement contrôlé par un réseau de régulation impliquant comme régulateur intracellulaire la protéine Fur et l'ARN RyhB. (B) Avec l'aide de la protéine Hfq, RyhB réprime la traduction de l'ARNm *sodB* et induit une dégradation rapide par l'action coordonnée de la RNase E et de la RNase III. 30S représente la petite sous unité du ribosome.

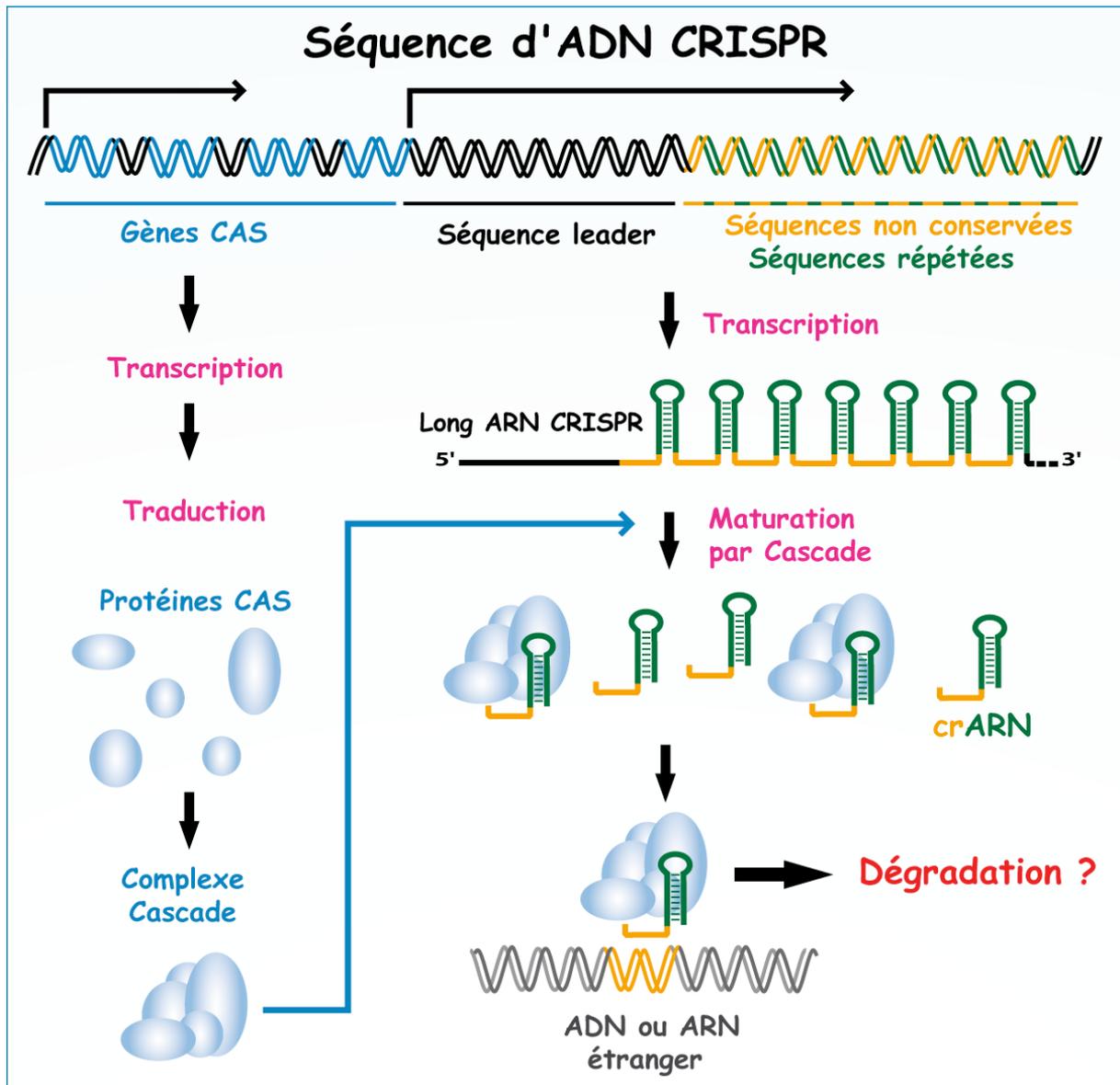
Au contact de l'eau, les métaux génèrent des radicaux hydroxyles  $\text{OH}^\bullet$ . Pour empêcher l'oxydation de ces différents métaux, la cellule régule finement la synthèse d'une famille d'enzymes, les superoxydes dismutases. Chacune de ces enzymes reconnaît un métal différent et bloque ainsi la formation des radicaux hydroxyles. Dans le cas du fer, la superoxyde dismutase impliquée est codée par le gène *sodB*. Lorsque la concentration cellulaire du fer est faible, la synthèse de la superoxyde dismutase n'est pas nécessaire et l'ARN RyhB réprime alors la traduction de l'ARNm *sodB* (Geissmann et Touati, 2004; Masse et al., 2005). Avec l'aide de la protéine Hfq qui déroule le site RBS de l'ARNm, l'ARN RyhB interagit au niveau du codon initiateur et bloque ainsi l'accessibilité au ribosome qui ne peut plus initier la synthèse protéique. Cette interaction mène ensuite à la dégradation de l'ARNm qui est clivé par l'action coordonnée de la RNase E et de la RNase III (figure 12, B) (Afonyushkin et al., 2005; Geissmann et Touati, 2004; Morita et al., 2005). Inversement, lorsque la concentration cellulaire du fer est haute, la protéine Fur inhibe la transcription de RyhB en se fixant au niveau des régions promotrices de l'ARN. Ainsi, l'ARNm *sodB* et tous les autres gènes

réprimés par RhyB sont traduits. Ceci a pour conséquence d'induire l'expression des gènes qui sont nécessaires à l'assimilation du fer.

### C.3.3. L'ARN GcvB, un motif conservé pour réguler de multiples ARNm

Chez *S. typhimurium* de nombreux ARN non codants ont été identifiés ces dernières années (Vogel, 2009). Généralement, ces ARN ciblent des ARNm qui sont synthétisés à partir d'unités de transcription différentes mais dont les fonctions sont reliées (Papenfort et al., 2008; Pfeiffer et al., 2007). Ce mode de régulation donne un moyen simple à la cellule de coordonner la synthèse de protéines distantes sur le génome en fonction d'un stress particulier ou de la disponibilité des nutriments. Parmi ces ARN, l'ARN GcvB réprime à lui seul l'expression de sept ARNm simultanément. Tous ces ARNm codent pour des protéines périplasmiques impliquées dans l'importation d'acides aminés ou de petits peptides dans la cellule (Sharma et al., 2007). GcvB agit de manière coordonnée avec la protéine Hfq et mène à une inhibition traductionnelle qui est rapidement suivie par une dégradation des ARNm cibles par la RNase E (Sharma et al., 2007). D'une manière intéressante, l'ARN GcvB utilise un motif en simple brin conservé et riche en G/U pour s'hybrider à une séquence riche en C/A présente sur la plupart des ARNm cibles (figure 13, A). Cette séquence C/A est importante pour l'initiation de la traduction mais le mécanisme moléculaire dans lequel elle est impliquée reste encore inconnu. Sur plusieurs ARNm cibles comme *oppA* et *dppA*, cette séquence est localisée proche du RBS ce qui explique pourquoi l'interaction avec GcvB inhibe l'initiation de la traduction. Cependant, sur d'autres ARNm cette séquence C/A peut se trouver jusqu'à 35 nucléotides en amont du RBS sans annuler l'effet de répression traductionnelle induit par GcvB (figure 13, B). Un tel phénomène pourrait s'expliquer par le modèle de "stand by" du ribosome (Darfeuille et al., 2007), mais dans ces exemples, la région RBS reste en permanence accessible. Une autre explication serait alors la forte affinité de la protéine ribosomique S1 pour les régions riches en C/A (Komarova et al., 2005). Dans ce cas là, les régions C/A pourraient tout simplement servir de point d'ancrage pour la fixation du ribosome. Globalement, cette étude a montré pour la première fois que des séquences conservées sur un ARN régulateur ainsi que sur ses ARNm cibles étaient requises pour réprimer efficacement l'initiation de la traduction. Le futur révélera quels sont les mécanismes moléculaires réellement engagés entre GcvB et les motifs C/A situés en amont de la région RBS des ARNm.





**Figure 14: Biogenèse et mode d'action des éléments CRISPR.** Les régions CRISPR permettent la synthèse d'un long ARN qui est constitué d'une alternance de séquences répétées (vert) et de séquences non-conservées (orange). Cet ARN est ensuite mûri par le complexe Cascade (bleu) qui induit la formation des ARNcr. Ces petits ARN représentent les effecteurs du système CRISPR et induisent la répression des gènes en ciblant, soit l'ADN, soit l'ARN. Le mécanisme de répression des ARNcr est encore mal connu mais il est probable qu'ils agissent sur la dégradation des molécules cibles.

La plupart des régions non conservées présentes dans les séquences CRISPR correspondent à des fragments d'ADN phagiques provenant d'infections antérieures. La synthèse des ARNcr aurait pour fonction de contrer une nouvelle attaque virale. Ces régions constitueraient donc une sorte de mémoire et les éléments CRISPR représenteraient ainsi la découverte du premier système immunitaire bactérien (Barrangou et al., 2007). Plus récemment, une autre étude a montré que le rôle des ARNcr n'était pas limité à la seule défense antivirale. En effet, chez *Staphylococcus epidermidis* certaines régions CRISPR possèdent des homologies de séquence avec des gènes plasmidiques. Ces régions

permettraient la synthèse de ARNcr impliqués dans la répression du mécanisme de conjugaison des plasmides (Marraffini et Sontheimer, 2008). Comment la bactérie incorpore des fragments d'ADN phagique ou plasmidique dans son propre génome est un processus encore mystérieux. Chez *E coli*, ce mécanisme pourrait impliquer la protéine CAS 1 qui possède un domaine endonucléase spécifique pour les ADN en double brin (Marraffini et Sontheimer, 2009; Wiedenheft et al., 2009).

Les protéines CAS ont la particularité de posséder de fortes homologies avec les enzymes impliquées dans l'ARNi eucaryotique (Makarova et al., 2006). Grâce à ces observations, plusieurs données expérimentales ont conduit au modèle selon lequel les ARNcr seraient maturés par l'action d'une machinerie complexe (figure 14). Dans ce modèle, les régions d'ADN CRISPR seraient tout d'abord transcrites sous la forme d'un long ARN précurseur de plusieurs centaines de nucléotides. Cet ARN serait alors reconnu par le complexe Cascade, probablement via des séquences spécifiques PAM situées au niveau des régions non conservées (Mojica et al., 2009), puis coupé à plusieurs reprises afin de générer des ARNcr de 50 à 60 nucléotides de long (Brouns et al., 2008; Carte et al., 2008). Ces petits ARN, constitués d'une région répétée et d'une région non conservée homologue à de l'ADN phagique ou plasmidique, resteraient ensuite associés au complexe Cascade avant d'être guidés vers leurs cibles pour induire la répression des gènes (Brouns et al., 2008). D'une manière surprenante les ARNcr cibleraient directement les molécules d'ADN (Marraffini et Sontheimer, 2008). Si ce mécanisme de répression est encore mal compris, il est probable que les éléments CRISPR agissent sur la dégradation de l'ADN ou sur la modification des nucléotides.

Les éléments CRISPR constituent un domaine de recherche qui suscite actuellement beaucoup d'intérêt. Si ce système commence tout juste à être compris, plusieurs questions se posent encore quant aux rôles exacts des protéines CAS ou de la région "leader" des ARNcr. Par ailleurs, puisque le nombre des protéines CAS varie d'une espèce à l'autre, plusieurs questions restent également ouvertes concernant les véritables fonctions de l'ensemble du système CRISPR. Les mois à venir apporteront certainement quelques réponses à ces questions. Chez *L. monocytogenes*, un ARN de type CRISPR mais non associé aux protéines CAS a été identifié (Mandin et al., 2007). Il se peut que cet ARN ait évolué vers d'autres fonctions ou que la machinerie protéique soit différente.

## E. Nouveaux ARN

L'identification des ARN régulateurs reste complexe, aussi bien chez les eucaryotes que chez les bactéries. Néanmoins, de récentes avancées technologiques comme le développement des "tiling arrays" ont permis de faire de grands progrès dans ce domaine. A l'inverse des "microarrays" qui testent uniquement l'expression de gènes annotés, les "tiling arrays" sont des puces à ADN (oligonucléotides de petites tailles et chevauchants) qui recouvrent l'intégralité d'un génome. Le "tiling arrays" est donc une technique sensible qui peut être associée à de nombreuses applications (Gregory et al., 2008; Mockler et al., 2005; Yazaki et al., 2007) et notamment à l'identification d'ARN régulateurs.

Récemment, cette technique a conduit à l'identification de 29 nouveaux ARNnc chez *L. monocytogenes* (Toledo-Arana et al., 2009). Ces ARN possèdent des caractéristiques distinctes. Leurs tailles varient entre 77 et 534 nucléotides, certains sont dépendants du facteur de stress  $\sigma^B$  alors que d'autres semblent requis pour la virulence. Cette étude a également identifié 12 nouveaux "riboswitch" mais elle a surtout soulevé l'existence probable de nouvelles classes d'ARN. Une première classe regrouperait cinq petits ARN possédant une phase ouverte de lecture (ORF). Ces ARN permettraient la synthèse de petits peptides dont la taille varie entre 24 et 64 acides aminés. Chez *S. aureus*, l'ARNIII exerce un double rôle d'ARN régulateur et d'ARNm (Novick et al., 1993), il n'est donc pas impossible que ces cinq nouveaux ARN agissent également comme des antisens sur la régulation des gènes. Une autre classe regroupe des ARN inhabituellement longs (>200 nucléotides), capables de recouvrir plusieurs ORF simultanément. Chez les eucaryotes, les ARNlnc ont désormais de nombreuses fonctions connues (Wilusz et al., 2009). Il est donc tentant de penser que les longs ARN bactériens pourraient avoir des fonctions comparables. Entre autre, ces ARN pourraient servir de matrice pour la maturation de petits ARN régulateurs, séquestrer et orienter l'utilisation de protéines régulatrices, ou encore induire des effets épigénétiques sur l'enroulement/superenroulement du chromosome bactérien. Enfin, cette étude révèle des complémentarités de bases entre plusieurs petits ARN qui seraient donc capables d'interagir ensemble. Cette dernière observation souligne pour la première fois l'existence d'un circuit de régulation dans lequel une cascade d'interaction entre ARN régulateurs pourrait prendre place.

## F. ARN régulateurs et virulence

Au cours d'une infection, les bactéries pathogènes doivent pouvoir s'adapter rapidement afin d'échapper aux défenses immunitaires de l'hôte. Ainsi, pour assurer leur survie, elles utilisent des systèmes variés pour sonder à tout moment le milieu extérieur et synthétisent en réponse de nombreux facteurs de virulence. Ces facteurs jouent de nombreux rôles en fonction de l'état plus ou moins avancé de l'infection. Ils sont notamment impliqués dans l'adhésion, la destruction des cellules hôtes, le transfert de l'information ou encore dans le transport transmembranaire des protéines. Du fait de cette grande diversité, ces facteurs de virulence sont donc finement régulés par des protéines mais aussi par des ARNnc (tableau 4). Plusieurs de ces ARN ont déjà été identifiés chez des bactéries Gram-négatif comme *E. coli* ou *S. typhimurium* (Geisinger et al., 2006; Vogel, 2009) et agissent sous la dépendance de la protéine Hfq dont l'implication dans la virulence a été démontré à plusieurs reprises (Hansen et Kaper, 2009; Sittka et al., 2008). Chez les bactéries Gram-positif, si de nombreux "riboswitch" sont déjà bien décrits, seul un petit nombre d'ARNnc sont connus pour être associés à la virulence (Geissmann et al., 2006; Johansson et Cossart, 2003; Toledo-Arana et al., 2007). De ce fait, il existe un réel manque de connaissance concernant les mécanismes d'action et les éventuels cofacteurs de ces ARN. Certes, la protéine Hfq est conservée chez les Gram-positif mais jusqu'à présent son importance est restée mineure. Ainsi, l'un des défis actuels est de déterminer la contribution des ARNnc dans la virulence de bactéries pathogènes comme par exemple *S. aureus*.

ARN Régulateurs	Bactéries	Cibles connues	Fonction de la cible	Mécanisme de régulation	Références
<b>Bactéries Gram négatives</b>					
SgrS	<i>E. coli</i>	ARNm <i>ptsG</i>	Transport du glucose	Inhibition de la traduction et dégradation de l'ARNm	(Kawamoto et al., 2005; Morita et al., 2006)
Qrr (1-4)	<i>V. cholerae</i>	ARNm <i>hapR</i>	Virulence	Inhibition de la traduction et dégradation de l'ARNm	(Lenz et al., 2005; Svenningsen et al., 2009; Svenningsen et al., 2008)
RsmY/RsmZ	<i>P. aeruginosa</i>	Protéine RsmA	Formation des biofilms, synthèse d'exotoxines	Séquestration de protéines	(Kay et al., 2006; Mulcahy et al., 2006)
PrrF1/PrrF2	<i>P. aeruginosa</i>	ARNm <i>sodB</i> , <i>sdhD</i>	Métabolisme du fer	Inhibition de la traduction et dégradation de l'ARNm	(Wilderman et al., 2004)
tmARN	<i>S. typhimurium</i>	?	Virulence	Dégradation des protéines?	(Julio et al., 2000)
<b>Bactéries Gram positives</b>					
VirX	<i>C. perfringens</i>	?	Régulation des toxines sécrétées	?	(Ohtani et al., 2002)
ARNIII	<i>S. aureus</i>	ARNm <i>spa</i> , <i>rot</i> , <i>sa1000</i> , <i>sa2353</i>	Régulation des gènes de virulence	Inhibition de la traduction et dégradation des ARNm	(Boisset et al., 2007; Huntzinger et al., 2005)
SprA	<i>S. aureus</i>	?	Virulence ?	?	(Pichon et Felden, 2005)
Ril32, Ril40	<i>L.monocytogenes</i>	?	Virulence	?	(Toledo-Arana et al., 2009)
RilB	<i>L.monocytogenes</i>	ARNm <i>Lmo2104</i> , <i>2105</i>	Transport du fer	Appariement antisens	(Mandin et al., 2007)
RilA, RilC, RilD, RilF, RilG, RilH	<i>L.monocytogenes</i>	?	?	Appariement antisens?	(Mandin et al., 2007)
csRNA1, 2, 3, 4 et 5	<i>S. pneumoniae</i>	?	?	?	(Halfmann et al., 2007)
FasX	<i>S. pyogenes</i>	ARNm <i>fbp54</i> , <i>mrp</i>	Adhésion au fibrinogène	?	(Klenk et al., 2005; Kreikemeyer et al., 2001)

**Tableau 4:** Quelques exemples d'ARN régulateurs impliqués dans la virulence des bactéries.

## IV. Régulation des gènes de virulence chez *Staphylococcus aureus*

### A. *Staphylococcus aureus*

Le genre *Staphylococcus* appartient à la famille des Micrococcaceae. Ce genre se compose de *cocci* à Gram positif groupés en amas, asporulés et aéro-anaérobie facultatifs. Il comprend de nombreuses espèces mais la plus fréquente en pathologie humaine est représentée par *Staphylococcus aureus*. Cette bactérie commensale se retrouve dans diverses niches écologiques comme la peau ou la plupart des muqueuses chez les mammifères à sang chaud. Elle est principalement extracellulaire mais peut dans certaines circonstances subsister dans le milieu intracellulaire (Sendi et Proctor, 2009). Chez l'homme, *S. aureus* se loge principalement au niveau des fausses nasales et près de 30% de la population comprendrait des porteurs sains (Rich, 2005; Wertheim et al., 2005). La transmission entre deux individus se produit le plus souvent au contact de la peau car les infections surviennent généralement lors d'une rupture de la barrière cutané-muqueuse. Cependant, un affaiblissement du système immunitaire peut également conduire à des infections très diverses en brisant le fragile équilibre qui existe entre la bactérie et son hôte. En effet, une fois dans le sang, la bactérie peut facilement coloniser l'ensemble des organes vitaux et induire des pathologies mortelles comme des endocardites, des chocs toxiques, des septicémies ou encore des pneumonies nécrotiques (Lina et al., 1997). De plus, *S. aureus* est une bactérie difficile à soigner car l'utilisation démesurée des antibiotiques dans les années 90 a induit une sélection de souches multirésistantes à la plupart des dérivés méticilline ainsi qu'à certains glycopeptides (Hiramatsu, 2001; Hiramatsu et al., 2001). *S. aureus* est également capable de former des biofilms compacts qui la protège contre la plupart des agressions extérieures comme l'action de détergents ou des processus de stérilisation. Ainsi, dans les hôpitaux il devient difficile d'éliminer cette bactérie qui cause dorénavant plus de 30% des infections nosocomiales chaque année. Finalement, la fréquence et la gravité des infections à *S. aureus* font que cette bactérie pose aujourd'hui un problème de santé publique.

Le fort pouvoir pathogène de *S. aureus* est étroitement lié à l'expression de nombreux facteurs de virulence au niveau du site d'infection (Foster, 2004; Holden et al., 2004; Novick, 2003b). Ces facteurs codent pour des protéines de surfaces ou des exoprotéines, et permettent à la bactérie de combattre le système immunitaire, d'adhérer aux cellules, de se disséminer dans l'hôte mais aussi d'utiliser les nutriments et l'énergie disponible (Novick, 2003a). Chaque type d'infection nécessite une combinaison particulière de facteurs de virulence. Ainsi, les protéines de surface sont synthétisées en début de phase de croissance puis réprimées en phase exponentielle alors que la plupart des exoprotéines sont synthétisées en phase post exponentielle (cf § **IV.C.**) (Dunman et al., 2001; Vandenesch et al., 1991). Ce mode de régulation alterné mime les étapes de l'infection *in vivo* et suggère le modèle suivant: en début d'infection, les bactéries synthétisent prioritairement des adhésines pour adhérer et coloniser l'hôte, puis lorsque les cellules s'accumulent, la synthèse des adhésines est inhibée au profit de celle des exoprotéines. Cette réorganisation de l'expression des facteurs de virulence permet à la bactérie de récupérer des nutriments mais surtout de se disperser dans l'hôte et de coloniser de nouveaux sites d'infection.

Chez *S. aureus* de nombreux facteurs de virulence ainsi que plusieurs gènes de résistances aux antibiotiques sont codés au niveau d'îlots de pathogénie (Hochhut et al., 2005; Novick, 2003b). Ces régions chromosomiques mobiles rendent le génome très dynamique en favorisant les transferts horizontaux entre différentes souches bactériennes. Ainsi, ces dernières années, le séquençage de différents génomes a montré que des souches considérées jusqu'à présent comme peu pathogènes avaient acquis des gènes de virulence par transfert d'îlot de pathogénie. C'est par exemple le cas d'une souche de *Staphylococcus epidermidis* qui aurait acquis des gènes provenant de *S. aureus* (Gill et al., 2005). Ces régions fournissent donc un système complémentaire à la conjugaison des plasmides et permettent un échange de gènes spécifiques à la virulence. Les îlots de pathogénie expliquent en partie pourquoi *S. aureus* s'adapte à des conditions environnementales multiples.

## B. Les facteurs de virulence

### B.1. Les protéines de surface

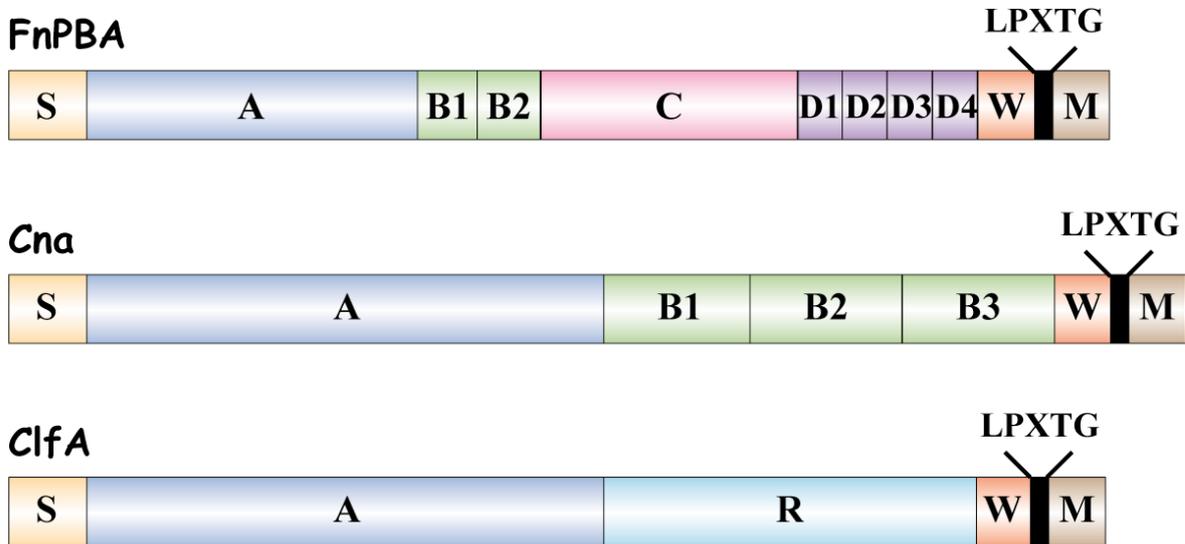
L'adhésion de *S. aureus* permet de commencer son cycle de division cellulaire et donc de démarrer le processus d'invasion de l'hôte. Pour se fixer efficacement, *S. aureus* procède via deux types de mécanismes: un mécanisme aspécifique basé sur des interactions de type hydrophobe (Carruthers et Kabat, 1983) et un mécanisme faisant appel à des protéines de surface qui reconnaissent spécifiquement les tissus de l'hôte. Il existe deux grandes classes de protéines d'adhésion, les adhésines liées à la membrane bactérienne ou MSCRAMM ("Microbial Surface Components Recognizing Adhesive Matrix Molecules") et les adhésines sécrétées ou SERAM ("Secretable Expanded Repertoire Adhesive Molecules") (Chavakis et al., 2005; Menzies, 2003).

#### B.1.1. Les MSCRAMM

Ces protéines sont associées à la membrane par des liaisons covalentes et partagent des caractéristiques structurales communes (figure 15) (Foster et Hook, 1998). Chacune possède un peptide signal en région N terminale pour l'adressage vers la membrane plasmique et un motif consensus LPXTG pour l'ancrage dans la paroi bactérienne par l'intermédiaire d'une réaction de transpeptidation (Mazmanian et al., 1999).

Certaines de ces adhésines se lient spécifiquement à la fibronectine. Cette glycoprotéine présente sous forme soluble dans le sang et sous forme insoluble dans la matrice extracellulaire est notamment impliquée dans l'adhésion des bactéries entre elles. Par exemple, les adhésines FnBPA et FnBPB utilisent les propriétés de la fibronectine pour permettre à la bactérie de se fixer efficacement sur les tissus de l'hôte (Grundmeier et al., 2004; Heilmann et al., 2004; Jonsson et al., 1991). D'autres adhésines comme ClfA et ClfB ("clumping factors A et B") ciblent plus spécifiquement le facteur fibrinogène qui est lui impliqué dans la coagulation des cellules (Ni Eidhin et al., 1998). Récemment, il a été montré que ClfB était particulièrement important pour l'adhésion des bactéries au niveau des fausses nasales (Corrigan et al., 2009) et la protéine ClfA a été utilisée dans le but de créer un potentiel vaccin contre *S. aureus* (Yin et al., 2009). Toutes ces protéines qui favorisent l'adhésion de la bactérie aux tissus de l'hôte sont également impliquées dans

l'adhésion sur des cathéters médicaux. Par conséquent, elles sont également à l'origine d'importants problèmes d'infections dans le milieu hospitalier (Cheung et Fischetti, 1990; Loughman et al., 2005).



**Figure 15: Organisation structurale des MSCRAMM.** "S" représente la séquence "peptide signal", "A" est le domaine de liaison au fibrinogène, les Domaines "R" et "D" sont impliqués dans la liaison à la fibronectine, "B" et "C" sont de fonctions inconnues, le domaine "W" correspond au site d'insertion dans la paroi, et "M" est la séquence d'adressage membranaire. La position du motif LPXTG est indiquée en noir. FnPBA: "fibrinogen bonding protein A", Cna: "collagen binding protein", Clfa: Clumping factor A".

### B.1.2. Les SERAM

Ces adhésines n'ont été identifiées que récemment grâce à leur capacité à fixer diverses protéines de la matrice extracellulaire comme le fibrinogène, la fibronectine, le collagène ou encore la prothrombine (Chavakis et al., 2005). Cependant, à l'inverse des MSCRAMM, les SERAM ne forment pas d'interactions covalentes au niveau de la membrane bactérienne. A ce jour, il n'est pas encore défini si ces adhésines forment des liaisons faibles avec les protéines membranaires, ou si elles sont sécrétées dans le milieu extracellulaire (Chavakis et al., 2005).

La coagulase est une des protéines majeures de la famille SERAM. Cette protéine se lie à la prothrombine pour former un complexe appelé staphylothrombine. Ce complexe entraîne ensuite la polymérisation du fibrinogène en fibrine et abouti finalement à l'agglutination des cellules (Kawabata et al., 1985). Cette propriété est notamment utilisée dans les laboratoires médicaux afin de diagnostiquer la présence de *S. aureus*. En plus de leur fonction d'adhésines, certaines protéines de la famille SERAM possèdent également

des propriétés immunorégulatrices qui peuvent provoquer des maladies vasculaires chez l'hôte. C'est par exemple le cas des protéines Eap et Efb qui modulent la réponse immunitaire (Palma et al., 2001). En outre, la protéine Eap favorise l'internalisation des bactéries par les cellules phagocytaires et contrecarre la réponse inflammatoire en bloquant l'accès aux leucocytes au niveau des sites lésés (Chavakis et al., 2002; Hagggar et al., 2003; Hansen et al., 2006).

### **B.1.3. La protéine A**

La protéine A est la protéine de surface majeure chez *S. aureus*, elle est présente dans plus de 95% des souches et représente 7% des protéines totales exprimées par la cellule (Patel et al., 1992). Cette protéine joue un rôle clé dans la virulence car elle possède plusieurs propriétés importantes pour contourner les défenses immunitaires de l'hôte. Tout d'abord, la protéine A possède une grande affinité pour la partie constante Fc des immunoglobulines G (IgG) (Foster, 2005; King et Wilkinson, 1981). Grâce à cette propriété, elle contrecarre la reconnaissance spécifique des IgG et bloque le phénomène de phagocytose par les macrophages (Foster, 2005). La protéine A joue également le rôle de superantigène contre les lymphocytes B et induit leur prolifération de manière totalement anarchique (Kristiansen et al., 1994). Selon la concentration de la protéine A, cette prolifération anarchique mène, soit à l'anergie, soit à la lyse des lymphocytes B (Goodyear et Silverman, 2004; Goodyear et al., 2006). La protéine A peut également induire l'activation du complément ou la sécrétion d'histamine par les basophiles (Espersen, 1985; Gomez et al., 2004; Nguyen et al., 2000). Ces effets affaiblissent les défenses de l'hôte en désorganisant la coordination du système immunitaire. Du fait de ses propriétés d'adhésine, la protéine A est directement responsable de plusieurs pathologies graves chez l'homme. Par exemple, en interagissant avec le facteur von Willebrand, une glycoprotéine importante pour l'adhésion des plaquettes au niveau des endothéliums vasculaires lésés, la protéine A serait impliquée dans certains cas d'endocardites infectieuses (Hartleib et al., 2000). La protéine A est également à l'origine de diverses pathologies respiratoires comme des pneumonies. En effet, cette dernière peut activer les récepteurs TNFR1 situés au niveau des poumons. L'activation de TNFR1 induit l'expression des chemokines IL-8 et entraîne une réaction pro-inflammatoire qui perturbe les fonctions de l'épithélium respiratoire. Parallèlement, la protéine A initie, via l'enzyme TACE, la coupure du TNFR1 qui est alors relargué sous forme soluble vers le milieu extracellulaire. Grâce à cette dernière propriété,

la protéine A abolie la réponse inflammatoire mais aussi la transmission des signaux entre les cellules. Il existe donc un équilibre entre réponse pro et anti-inflammatoire (Gomez et al., 2004; Gomez et al., 2006).

Facteurs de Virulence	Gènes	Rôle dans laPathogénie	<i>agr</i>	Références
<b>Superantigènes</b>				
Entérotoxine A	<i>sea</i>	intoxication alimentaire	O	(Tremaine et al., 1993)
Entérotoxine B	<i>seb</i>	intoxication alimentaire	+	(Gaskill et Khan, 1988)
Entérotoxine C	<i>sec</i>	intoxication alimentaire	+	(Regassa et al., 1991)
Entérotoxine D	<i>sed</i>	intoxication alimentaire	+	(Zhang et Stewart, 2000)
Exfoliatine A, B	<i>eta</i>	syndrome de la peau ébouillantée	+	(Sheehan et al., 1992)
TSST-0, TSST-1	<i>etb</i>	syndrome du choc toxique	+	(Recsei et al., 1986)
<b>Cytotoxines</b>				
Hémolysine $\alpha$	<i>hla</i>	hémolyse, nécrose	+	(Schmidt et al., 2001)
Hémolysine $\beta$	<i>hlb</i>	hémolyse, nécrose	+	(Giraud et al., 1997)
Hémolysine $\delta$	<i>hld</i>	hémolyse légère	+	(Schmidt et al., 2001)
Hémolysine $\gamma$	<i>hlg</i>	hémolyse, nécrose	+	(Bronner et al., 2000)
Leucolysine	<i>lukS/F</i>	leucolyse, pneumopathies,...	+	(Said-Salim et al., 2003)
<b>Enzymes</b>				
Protéases à sérine SplA-F	<i>splA-F</i>	diffusion, nutrition	+	(Said-Salim et al., 2003)
Protéase V8 (SspA)	<i>sspA</i>	diffusion, nutrition	+	(Arvidson et Tegmark, 2001)
Auréolysine	<i>aur</i>	diffusion, nutrition	+	(Arvidson et Tegmark, 2001)
Protéase à cystéine	<i>sspB</i>	diffusion, nutrition	+	(Said-Salim et al., 2003)
Glycérol ester hydrolase	<i>geh</i>	diffusion, nutrition	+	(Said-Salim et al., 2003)
Lipase	<i>lip</i>	diffusion, nutrition	+	(Chamberlain et Imanoel, 1996)
Nucléase	<i>nuc</i>	nutrition	+	(Smeltzer et al., 1993)
Coagulase	<i>coa</i>	adhésion	-	(Giraud et al., 1997)
Staphylokinase	<i>sak</i>	activateur du plasminogène	+	(Recsei et al., 1986)
<b>Protéines de surface</b>				
Protéine A	<i>spa</i>	adhésion, anti-immunitaire	-	(Said-Salim et al., 2003)
Collagène BP	<i>cna</i>	fixation au collagène	O	(Blevins et al., 1999)
Fibronectin BPA	<i>fnbA</i>	fixation à la fibronectine	O	(Saravia-Otten et al., 1997)
Fibronectin BPB	<i>fnbB</i>	fixation à la fibronectine	O	(Saravia-Otten et al., 1997)
Clumping Factor A	<i>clfA</i>	fixation au fibrinogène	O	(Said-Salim et al., 2003)
Clumping Factor B	<i>clfB</i>	fixation au fibrinogène	O	(Said-Salim et al., 2003)

**Tableau 5:** Principaux facteurs de virulence de *S. aureus*. (+) activation par *agr*, (-) répression par *agr*. (O), pas d'effet d'*agr*.

## B.2. Les protéines sécrétées

Les protéines sécrétées regroupent toute une batterie de toxines et de protéases. Ces protéines participent d'une part à la lutte contre le système immunitaire et à la destruction des tissus cellulaires de l'hôte, et d'autre part, à la dégradation des protéines de surface bactériennes.

### B.2.1. Les toxines

#### Les toxines à tropisme membranaire

Ces toxines ciblent préférentiellement les membranes plasmiques afin d'induire la lyse cellulaire. Elles peuvent procéder de deux manières différentes, soit elles provoquent la formation de canaux dans la membrane, soit elles déstabilisent les lipides membranaires par l'intermédiaire d'une réaction enzymatique. Ces toxines sont divisées en quatre classes distinctes: les hémolysines  $\alpha$ ,  $\beta$ ,  $\delta$  et les toxines à deux composants.

L'hémolysine  $\alpha$  est la toxine majeure chez *S. aureus*, elle serait présente dans plus de 90% des souches (Palmer, 1998). Cette toxine est sécrétée sous forme de monomères qui se regroupent ensuite en un heptamère lytique au niveau de la membrane plasmique des cellules cibles (Song et al., 1996). Si les érythrocytes sont les plus vulnérables, la plupart des cellules immunitaires (lymphocytes, plaquettes, kératinocytes...) sont aussi attaquées par l'hémolysine  $\alpha$  (Walev et al., 1993; Walev et al., 1994). L'hémolysine  $\beta$  provoque une hémolyse partielle à 37°C. Ce phénomène est considérablement augmenté à 4°C d'où la dénomination d'hémolyse "chaud-froid". C'est une phospholipase C synthétisée en phase stationnaire de croissance et qui possède une activité sphingomyelase. Ainsi, elle est plus ou moins active sur les hématies de différentes espèces en fonction de la teneur en sphingomyeline de leur membrane. Enfin, l'hémolysine  $\delta$  est une petite protéine de 26 acides aminés qui perméabilise les cellules de l'hôte en s'insérant dans les membranes plasmiques (Dufourc et al., 1990). Dans certains cas, l'action de l'hémolysine  $\delta$  peut entraîner la lyse des érythrocytes.

Parmi les toxines à deux composants, la leucocidine de Pantone Valentine (PLV) est l'une des plus virulente. Cette toxine est notamment responsable de nombreux cas de pneumonies nécrosantes mortelles (Barcelo et al., 2009; Thomas et al., 2009a). Elle se compose de deux sous-unités protéiques sécrétées qui s'assemblent en un octamère à la surface des cellules cibles. Cet octamère favorise la formation d'un pore dans la membrane plasmique et induit finalement la lyse de plusieurs types de cellules immunitaires comme les monocytes et les lymphocytes (Meyer et al., 2009; Prevost et al., 1995). A forte concentration, la PLV induit la nécrose des cellules cibles par activation des canaux calciques et entraîne une fuite osmotique dans le cytosol. A faible concentration, la PLV génère des pores dans les membranes mitochondriales et induit l'apoptose (Genestier et al., 2005).

### **Les toxines superantigéniques**

*S. aureus* est également à l'origine d'intoxications alimentaires variées. Ce type d'infection est généralement du à l'expression de la toxine du choc toxique staphylococcique (TSST-1) et de diverses entérotoxines (SE) (Thomas et al., 2009b; Thomas et al., 2006). La TSST-1 et certaines de ces entérotoxines possèdent des propriétés superantigéniques, c'est-à-dire qu'elles sont capables d'activer de manière anarchique la multiplication des lymphocytes T et B. La prolifération aberrante des lymphocytes entraîne alors une production massive de cytokines pro-inflammatoires ce qui provoque le choc toxique (Proft et Fraser, 2003). Ces toxines sont, entre autre, responsables du syndrome de la peau ébouillantée ("Staphylococcal Scalded Skin Infection") qui survient le plus souvent chez le nourrisson (Kapoor et al., 2008; Ladhani, 2001).

### **B.2.2. Les protéases**

*S. aureus* exprime plusieurs sortes de protéases. La serine protéase SspA (ou protéase V8) joue un rôle important dans la dégradation des protéines de surface de la bactérie (Karlsson et al., 2001; McGavin et al., 1997). Ainsi, elle est responsable de la dégradation de la protéine A et de la plupart des protéines de liaison au fibrinogène en fin de phase exponentielle de croissance. L'auréolysine est une autre protéase majeure qui a une double fonction. D'une part, elle agit de manière similaire à SspA sur la dégradation des protéines de surface, et d'autre part, elle catalyse l'activation de nombreuses protéases

en clivant leurs précurseurs au niveau de sites spécifiques (McAleese et al., 2001; Rice et al., 2001). Il y a donc une réelle interdépendance dans l'activation des différentes protéases. Par exemple, l'auréolysine mature SspA sous sa forme active, puis SspA mature à son tour la cystéine protéase SspB (Rice et al., 2001). Si dans l'ensemble les protéases de *S. aureus* sont importantes pour remodeler la membrane bactérienne, elles participent aussi activement à la lutte contre le système immunitaire. Certaines sont notamment capables de dégrader les chaînes lourdes des immunoglobulines et plusieurs inhibiteurs de protéases (Prokesova et al., 1992). Enfin, les protéases fournissent une source d'énergie à la bactérie puisqu'elles dégradent l'ensemble des macromolécules présentes dans le milieu extracellulaire après l'apoptose des cellules de l'hôte (Shaw et al., 2004).

En 2001, six nouvelles protéases à serine codées sur un même opéron ont été identifiées, les protéases SplA-F (Reed et al., 2001). La fonction de ces protéines est encore mal connue mais la structure cristallographique de SplA montre de fortes similarités avec SspA (Stec-Niemczyk et al., 2009). Les protéases SplA-F sont donc probablement impliquées dans la virulence de la bactérie, d'autant plus que l'expression de SplC a déjà été confirmée dans le passé chez des patients atteints d'endocardites (Rieneck et al., 1997).

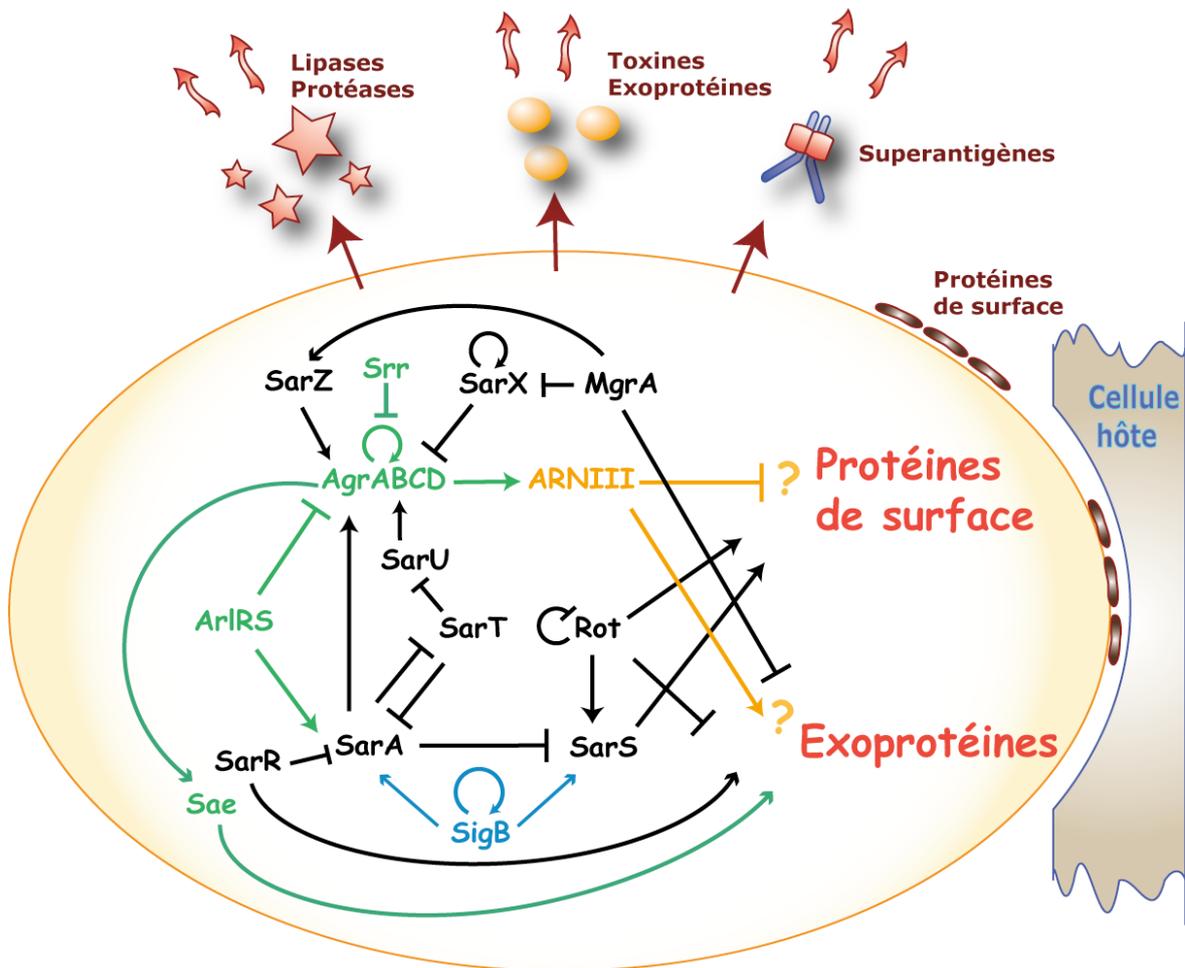
### **B.2.3. Les autres enzymes**

La plupart des souches de *S. aureus* expriment également de nombreuses lipases. Ces lipases auraient deux rôles majeurs au cours de l'infection. Le premier serait de déstabiliser l'organisation des acides gras présents au niveau de la peau afin de favoriser la pénétration des bactéries à travers la barrière cutanéomuqueuse de l'hôte. Leur deuxième rôle serait de protéger les bactéries de l'apoptose en détruisant les acides gras bactéricides produits par l'hôte (Long et al., 1992).

## **C. La régulation des facteurs de virulence**

Chez *S. aureus*, chaque facteur de virulence possède des fonctions précises pour adapter les besoins de la bactérie en fonction de la variation des conditions environnementales. Au cours d'une infection, ces facteurs sont régulés de manière

temporelle. Leur expression est finement régulée et passe en réalité par l'intermédiaire d'un réseau de régulation complexe qui implique des systèmes à deux composants, des régulateurs de la transcription et le facteur de stress  $\sigma^B$  (figure 16). Ces régulateurs agissent simultanément sur plusieurs voies de régulation et selon les cas jouent le rôle d'activateur ou de répresseur. Ainsi, il n'est pas rare que deux régulateurs contrôlent l'expression d'un même gène, l'un positivement et l'autre négativement. De cette manière, les facteurs de virulence se retrouvent enfermés dans un véritable circuit de régulation qui contrôle à chaque instant l'expression des gènes, que soit au niveau transcriptionnel ou post-transcriptionnel (Novick, 2003a).



**Figure 16: Réseaux de régulation des facteurs de virulence chez *S. aureus*.** L'expression des facteurs de virulence (protéines de surface, exoprotéines) est dépendante de réseaux de régulation complexes, souvent interconnectés, qui font intervenir des systèmes à deux composants (vert), des facteurs de régulation de la transcription (noir), et le facteur sigma B (en bleu).

### C.1. Les systèmes à deux composants

Les systèmes à deux composants (TCS) sondent le milieu extérieur et relayent l'information vers la cellule. Ces systèmes, généralement organisés en opéron, codent pour deux types de protéines: une protéine membranaire (histidine kinase) capable de mesurer un signal spécifique comme la densité cellulaire ou la présence d'un métabolite particulier, et une protéine régulatrice de la transcription. Les protéines membranaires sont activées par la fixation d'un ligand spécifique qui induit l'autophosphorylation de la kinase. Ce phosphate est ensuite transféré au facteur de régulation qui sous cette forme peut activer la synthèse des gènes cibles (Stock et al., 2000). Chez *S. aureus*, il existe quatre principaux TCS impliqués dans la virulence, les systèmes *agr*, *sae*, *srr* et *arlRS*.

Le système *agr* ("accessory response regulator) sonde la densité cellulaire et régule à lui seul plus d'une centaine de gènes (Recsei et al., 1986). Ses fonctions seront expliquées plus en détails au chapitre IV.D.

Le système *sae* ("*S. aureus* exoprotein expression") répond à divers signaux comme la forte concentration en sels, le pH acide ou la présence de glucose (Giraud et al., 1999; Giraud et al., 1994; Novick et Jiang, 2003). Ce système est activé par *agr*. *Sae* active plusieurs hémolysines mais joue également un rôle important dans l'adhésion des cellules en régulant plusieurs protéines de liaison au fibrinogène (Liang et al., 2006).

Le système *srr* ("Staphylococcal respiratory response") répond au taux d'oxygénation du milieu extérieur et régule à la fois les protéines de surface et les exoprotéines (Yarwood et al., 2001). En condition anaérobie, *srr* réprime le système *agr* et favorise l'expression des gènes impliqués dans la fermentation (Throup et al., 2001).

Le système *arlRS* ("autolysis related locus") régule l'expression de gènes impliqués dans le remodelage de la membrane bactérienne. Cet opéron est particulièrement important pour les processus de lyse et de division cellulaire. En parallèle, il participe à la virulence en régulant l'opéron *agr* et le facteur Rot (Liang et al., 2005).

## C.2. Les régulateurs de la transcription, la famille Sar

### C.2.1. Le facteur SarA

Le facteur SarA ("Staphylococcal accessory regulator A") est le premier régulateur de la famille Sar à avoir été identifié (Cheung et al., 1992). Ce facteur de 14,3 kDa est synthétisé à partir de trois promoteurs différents (P1, P2 et P3). Alors que les promoteurs P1 et P2 sont dépendants du facteur ménager  $\sigma^A$ , le promoteur P3 est lui dépendant du facteur de stress  $\sigma^B$ . Grâce à la diversité de ces promoteurs, SarA est synthétisé de manière constitutive et régule de nombreux gènes durant tout le cycle cellulaire (Manna et al., 1998).

Le facteur SarA s'assemble sous forme de dimères et reconnaît des séquences conservées, riches en AT, situées proches de la région promotrice des gènes cibles (Novick, 2003a). SarA est largement impliqué dans la régulation des gènes de virulence où il agit à la fois comme activateur et comme répresseur. Par exemple, il active l'opéron *agr* et l'hémolysine  $\alpha$  alors qu'il réprime la protéine A (Cheung et al., 1997a; Cheung et al., 1997b). Les mécanismes d'action de SarA sont encore mal connus. En tant qu'activateur, une hypothèse suggère qu'il pourrait jouer sur l'élasticité de l'ADN et favoriser la transcription en rapprochant les régions promotrices -35 et -10 (Rechtin et al., 1999). Cependant, des études réalisées *in vitro* montrent que SarA agirait uniquement en tant que répresseur (Chakrabarti et Misra, 2000). En réalité, il activerait indirectement la synthèse de certains gènes en inhibant l'action de répresseurs transcriptionnels (Arvidson et Tegmark, 2001). Par exemple, SarA activerait l'hémolysine  $\alpha$  en réprimant la synthèse d'un autre facteur Sar, le facteur SarT (Schmidt et al., 2001). Le facteur SarA pourrait également réprimer des gènes au niveau post-transcriptionnel en régulant l'expression de petits ARN régulateurs mais à ce jour aucune donnée expérimentale n'a encore confirmé cette hypothèse.

### C.2.2. Les facteurs homologues à SarA

Plusieurs autres facteurs de la famille Sar ont été identifiés aujourd'hui. Si la plupart de ces facteurs effectuent des régulations croisées, tous jouent un rôle important dans l'expression des gènes de virulence en interagissant notamment avec l'opéron *agr* (Cheung et al., 2008). Leurs principales propriétés sont données ci-dessous:

- SarR (13,4 kDa) possède 51% d'homologie avec SarA. SarR réprime le facteur SarA et régule de multiples facteurs de virulence, aussi bien en phase exponentielle qu'en phase stationnaire (Manna et Cheung, 2001; Manna et Cheung, 2006b).

- SarS (29 kDa) possède deux domaines homologues à SarA (Tegmark et al., 2000). Cette protéine active plusieurs gènes de virulence et notamment la protéine A en début de phase exponentielle de croissance. SarS est réprimé en fin de phase exponentielle par l'action coordonnée de SarA et de l'opéron *agr* (Oscarsson et al., 2006; Schmidt et al., 2003).

- SarT (16,1 kDa) possède 36% d'homologie avec SarA et 20% d'homologie avec SarR. SarT active le facteur SarS et réprime l'hémolysine  $\alpha$ . La transcription de SarT est réprimée par SarA et par l'opéron *agr* (Cheung et al., 2004; Schmidt et al., 2001).

- SarU (29,3 kDa) possède une structure comparable à SarA. Ce facteur active l'opéron *agr* mais est réprimé par SarT (Cheung et al., 2004; Manna et Cheung, 2003).

- SarX possède un domaine homologue à SarA. Il réprime l'opéron *agr* mais est activé par le facteur MgrA. Il régule également sa propre expression (Manna et Cheung, 2006a).

- SarZ active l'opéron *agr* ainsi que la protéase SspA et réprime le facteur SarS. Ce facteur est lui-même activé par MgrA et régule des gènes importants pour la formation des biofilms (Ballal et al., 2009; Tamber et Cheung, 2009).

- MgrA ("multiple gene regulator") active l'opéron *agr* et réprime le facteur SarS. Ce facteur participe également à l'expression du polysaccharide de type 8 et réprime la synthèse de plusieurs protéases sécrétées (Ingavale et al., 2005; Luong et al., 2003).

-Rot ("Repressor of toxins"; 15,6 kDa) agit comme antagoniste à l'opéron *agr*. Il active les protéines de surface et réprime les exoprotéines. La synthèse de Rot est réprimée par l'opéron *agr* en phase stationnaire de croissance (McNamara et al., 2000; Said-Salim et al., 2003).

### **C.3. Le facteur de stress $\sigma^B$**

Le facteur  $\sigma^B$  est impliqué dans certains cas de septicémie chez la souris (Jonsson et al., 2004) mais généralement il n'est pas toujours indispensable à la pathogénie de *S. aureus*. Il joue cependant un rôle essentiel dans l'adaptation au stress (van Schaik et Abee, 2005). Ce facteur est principalement exprimé en phase stationnaire de croissance et

permet à l'ARN polymérase bactérienne de reconnaître des régions promotrices spécifiques (Palma et al., 2001). Le facteur  $\sigma^B$  contrôle plus de 30 gènes impliqués dans la réponse au stress et dans le métabolisme, des facteurs de régulation (i.e. SarA, *agr*) et des gènes de virulence (Bischoff et al., 2001; Entenza et al., 2005; Gertz et al., 2000). Il est aussi impliqué dans les mécanismes de résistance aux antibiotiques, la formation des biofilms ou encore, l'internalisation dans les cellules de l'hôte (Nair et al., 2003; Rachid et al., 2000; Singh et al., 2003). Enfin, nous avons montré que  $\sigma^B$  induit la synthèse de petits ARNnc et qu'il réprime ainsi de manière indirecte l'expression de certains gènes (cf Résultats et Discussions, § **III**).

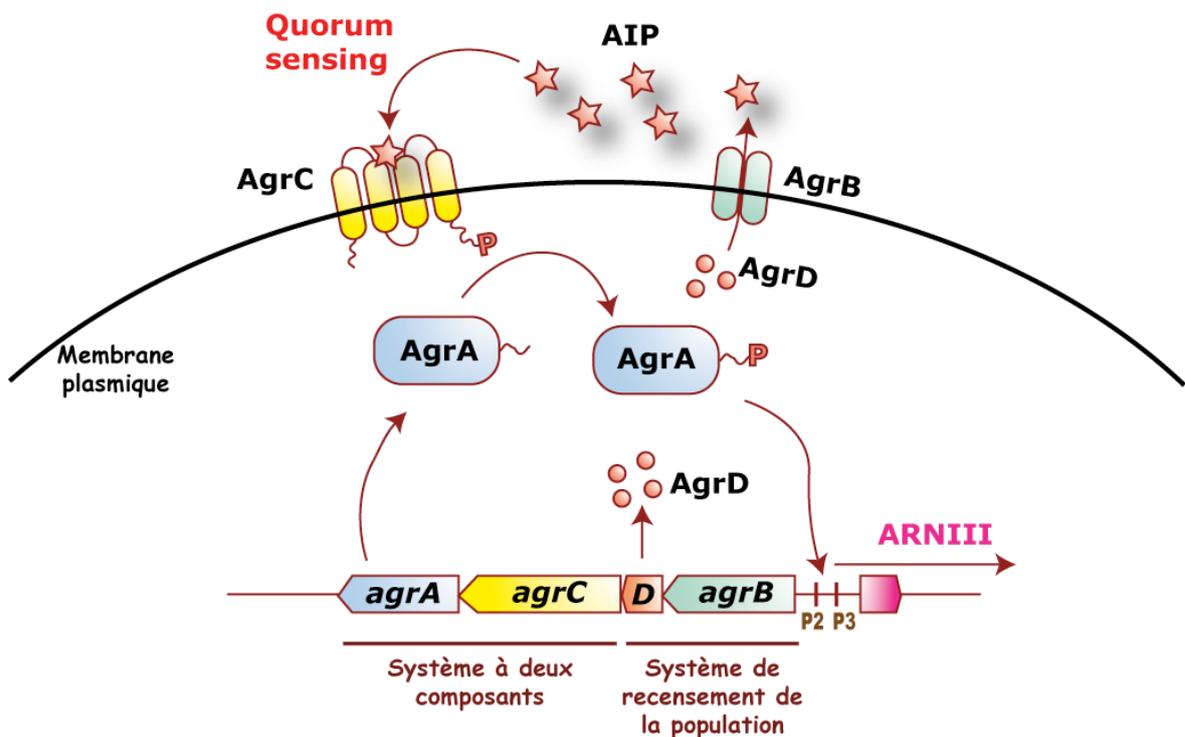
Système régulateur	Type	Rôle	Références
<i>agr/ARNIII</i>	Systèmes à deux composants	Régulateur de l'expression de nombreuses protéines cytoplasmiques et extracellulaires	(Novick et al., 1993)
<i>saeRS</i>	"	Régulation de l'expression de protéines extracellulaires	(Novick et Jiang, 2003)
<i>arlRS</i>	"	Régulation de l'autolyse et de certains facteurs de virulence	(Liang et al., 2005)
<i>srrAB</i>	"	Régulation de certains facteurs de virulence aux basses pressions en oxygène	(Yarwood et al., 2001)
<i>lytRS</i>	"	Régulation de l'autolyse	(Brunskill et Bayles, 1996)
<i>yycFG</i>	"	Régulation de la perméabilité cellulaire	(Martin et al., 1999)
SvrA	Protéine membranaire	Régulation de l'expression de <i>agr</i>	(Garvis et al., 2002)
SarA	Facteur de régulation de la transcription	Activation de <i>agr</i> , répresseur de la transcription	(Manna et al., 1998)
SarR	"	Facteur de transcription pour SarA et SarS	(Manna et Cheung, 2006b)
SarS	"	Activation de la transcription de <i>spa</i> et peut-être d'autres protéines de surface	(Oscarsson et al., 2006)
SarT	"	Répression de l'expression de <i>hla</i> et peut-être d'autres toxines extracellulaires	(Cheung et al., 2004)
SarU	"	Activation de la transcription de <i>agr</i>	(Cheung et al., 2004)
SarX	"	Répression de la transcription de <i>agr</i>	(Manna et Cheung, 2006a)
SarZ	"	Répression de <i>agr</i> et du facteur SarS. Régulation des gènes du biofilm	(Ballal et al., 2009; Tamber et Cheung, 2009)
MgrA	"	Activation de la transcription de <i>agr</i>	(Ingavale et al., 2005)
Rot	"	Répresseur de la transcription de <i>hla</i> et d'autres exoprotéines, activation <i>spa</i>	(McNamara et al., 2000; Said-Salim et al., 2003)
$\sigma^B$	facteur sigma	Régulation de certains facteurs de virulence en fin de phase exponentielle	(van Schaik et Abee, 2005)

**Tableau 6:** Liste des principaux régulateurs des gènes de virulence chez *S. aureus*.

## D. Le système agr et la synthèse de l'ARNIII

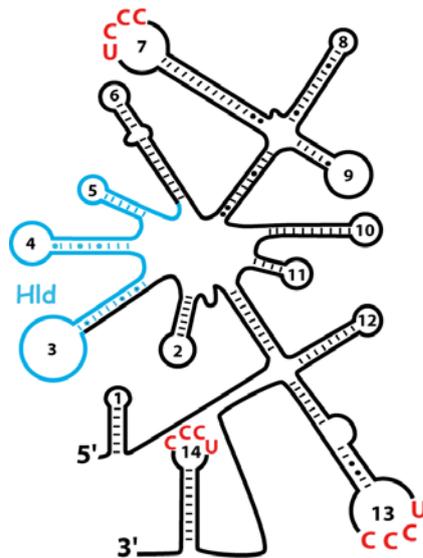
### D.1. Le système agr

L'opéron *agr* est un régulateur global exprimé en fin de phase de croissance qui contrôle l'expression de plus de 100 facteurs de virulence. En plus de coder pour un système à deux composants, l'opéron *agr* a la particularité d'exprimer un système de recensement de la population. Les deux systèmes fonctionnent de manière coordonnée et permettent à l'opéron *agr* de sonder à tout moment la densité cellulaire par un mécanisme appelé "quorum sensing" (Novick et Geisinger, 2008; Swem et al., 2009). Ce mécanisme établit une communication entre les cellules (Bassler, 2002) et, chez *S. aureus*, détermine le passage entre la phase de colonisation et la phase de dissémination.



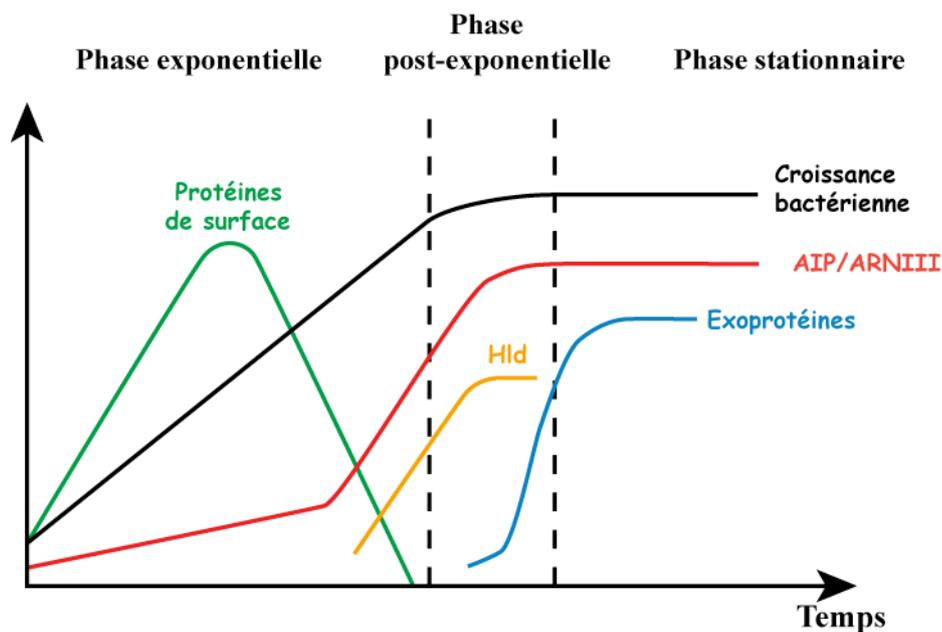
**Figure 17: L'opéron *agr*.** AgrB et AgrD mettent place le système de recensement de la population. Durant la croissance cellulaire, le peptide autoinducteur AIP s'accumule dans le milieu extracellulaire et, à une certaine concentration seuil, se fixe au niveau de la protéine membranaire AgrC. La fixation de AIP induit l'autophosphorylation de AgrC. Le groupement phosphate est ensuite transféré sur la protéine AgrA qui sous cette forme active les promoteurs P2 et P3.

L'opéron *agr* est composé de deux unités de transcription divergentes contrôlées par les promoteurs P2 et P3 (figure 17). Le promoteur P2 gouverne l'expression de quatre gènes qui permettent d'une part, l'expression du système à deux composants (*agrC* et *agrA*), et d'autre part, l'expression du système de recensement de la population (*agrB* et *agrD*). Le promoteur P3 induit la synthèse de l'ARNIII. En début de phase exponentielle de croissance, une faible transcription à lieu au niveau du promoteur P2, ce qui permet la synthèse à un niveau basal des protéines AgrA, AgrB, AgrC et AgrD. La protéine AgrB mature AgrD en un peptide autoinducteur (AIP) qui est ensuite sécrété vers le milieu extérieur (Novick et Geisinger, 2008; Qiu et al., 2005). A une certaine concentration seuil, le peptide autoinducteur se fixe sur le récepteur membranaire AgrC, une histidine kinase. Cette fixation provoque un réarrangement structural de AgrC et induit une autophosphorylation de la protéine (Lyon et al., 2002). La protéine AgrC modifiée provoque ensuite la phosphorylation du régulateur transcriptionnel AgrA, qui sous cette forme régule plusieurs gènes du métabolisme (Queck et al., 2008) mais surtout stimule l'expression des promoteurs P2 et P3 (Koenig et al., 2004). Ainsi, ce mécanisme permet à la bactérie de mesurer la densité cellulaire en fonction des peptides autoinducteurs produits par l'ensemble de la population bactérienne. Finalement, ce système permet une autoamplification de l'opéron *agr* ainsi que la synthèse massive de son effecteur principal, l'ARNIII (Novick, 2003a).



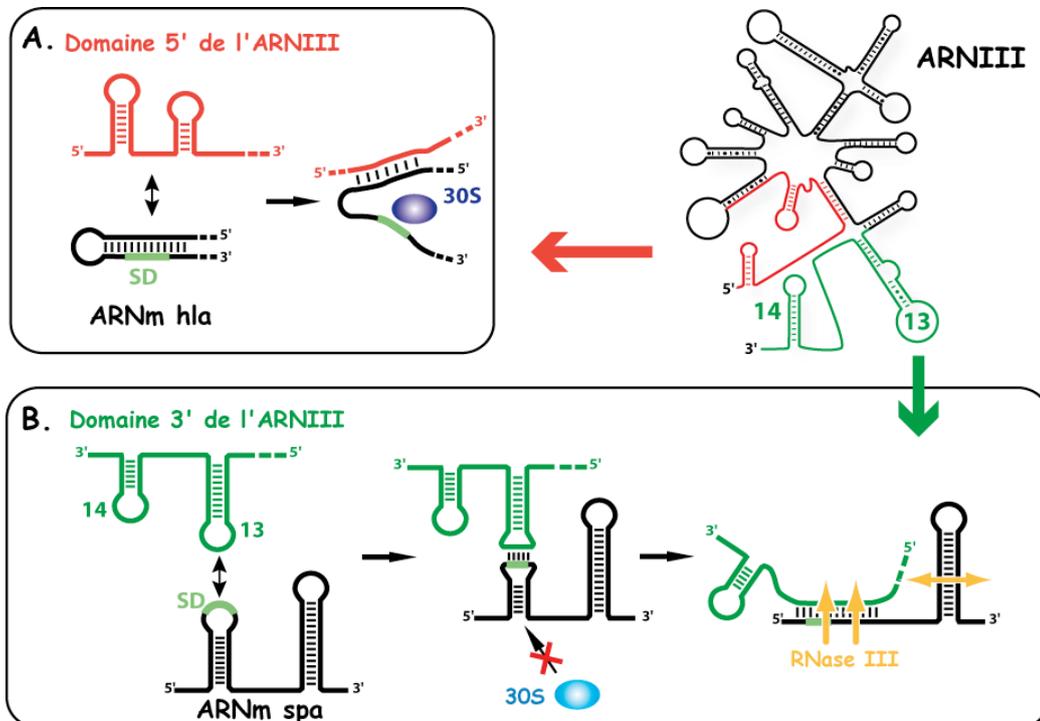
**Figure 18: Structure secondaire de l'ARNIII.** L'ARNIII est un ARN de 514 nucléotides constitué de quatorze motifs en tige-boucle. Cet ARN a également la particularité de coder lui-même pour l'hémolysine  $\delta$  (en bleu) et possède plusieurs motifs conservés riches en résidus C dans ses tige-boucles 7, 13 et 14 (en rouge).

L'ARNIII est un ARN régulateur de 514 nucléotides qui contrôle l'expression de plus de 40 gènes de virulence chez *S. aureus* (Novick et al., 1993). Cet ARN a également la particularité de coder lui-même pour un facteur de virulence, l'hémolysine  $\delta$ , mais cette propriété ne semble pas liée à sa fonction de régulateur. La structure secondaire de l'ARNIII est constituée de quatorze motifs en tiges-boucles (Benito et al., 2000). Parmi ces motifs, les tige-boucles 7, 13 et 14 sont extrêmement conservées, et possèdent toutes des séquences redondantes riches en résidus C (figure 18). L'ARNIII s'accumule en fin de phase de croissance et régule de manière antagoniste deux types de facteurs de virulence. Il active les exoprotéines (hémolysines, protéases) et réprime les protéines de surface (protéine A, protéines de liaison au fibrinogène...). Ces régulations se déroulent au niveau post-transcriptionnel ainsi qu'au niveau transcriptionnel (figure 19) (Novick et al., 1993).



**Figure 19: La régulation temporelle des gènes de virulence chez *S. aureus*.** L'ARNIII est l'un des principaux régulateurs des gènes de virulence chez *S. aureus*; il réprime les protéines de surface et active les exoprotéines. En début de phase de croissance l'ARNIII est faiblement exprimé ce qui permet une forte production des protéines de surface. Au niveau de l'infection, ces protéines favorisent les étapes d'adhésion et de colonisation de l'hôte. L'ARNIII répond à la densité cellulaire, il s'accumule donc en fin de phase exponentielle et favorise l'expression des exoprotéines. Ces dernières vont permettre à la bactérie de lutter contre les défenses immunitaires. Cette étape est aussi requise pour la dissémination de la bactérie dans les tissus de l'hôte. Hld représente l'hémolysine  $\delta$ , la toxine directement codée par l'ARNIII et AIP le peptide autoinducteur codé par le gène *agrD* de l'opéron *agr*.

Au niveau post-transcriptionnel, l'ARNIII agit sur la traduction et la dégradation d'ARNm cibles par des interactions de type antisens. Par exemple, le domaine 5' de l'ARNIII joue le rôle d'activateur en empêchant la formation d'une structure inhibitrice de la traduction au niveau de l'ARNm *hla* qui code pour l'hémolysine  $\alpha$  (figure 20, A) (Morfeldt et al., 1995). Cependant, l'ARNIII agit le plus souvent en tant que répresseur traductionnel. Le premier cas a été montré avec l'ARNm *spa* qui code pour la protéine A (Huntzinger et al., 2005). En absence d'ARNIII, la séquence SD de l'ARNm *spa* est accessible au ribosome qui peut facilement initier la synthèse protéique. Dès que l'ARNIII s'accumule, ce dernier interagit au niveau de la région RBS de l'ARNm *spa* (figure 20, B). Cette interaction est initiée entre le motif riche en C de la tige-boucle 13 de l'ARNIII et la région SD de l'ARNm, puis est rapidement étendue sous la forme d'un long duplexe d'ARN. La formation de ce duplexe inhibe la traduction de l'ARNm et fournit un site de reconnaissance pour la RNase III spécifique des ARN en double brin. Par ailleurs, l'ARNm contient une structure en tige-boucle stable en aval du site de fixation de l'ARNIII qui est fortement coupée par la RNase III. Ainsi, l'absence du ribosome permet à la RNase III d'initier la dégradation rapide de l'ARNm *spa*.



**Figure 20: L'ARNIII régule les facteurs de virulence par des mécanismes de type antisens. (A)** Le domaine 5' de l'ARNIII (en rouge) active la traduction de l'hémolysine  $\alpha$  en empêchant la formation d'une structure inhibitrice pour la fixation du ribosome. **(B)** Le domaine 3' de l'ARNIII, et notamment sa tige-boucle 13 (en vert), réprime la synthèse de la protéine A en interagissant au niveau de la séquence Shine Dalgarno de l'ARNm *spa*. Cette interaction bloque l'accès au ribosome et induit une dégradation par la RNase III.

Au niveau transcriptionnel, l'ARNIII agirait principalement de manière indirecte en inhibant la traduction de certains facteurs de régulation. De récents résultats ont montré que l'ARNIII réprime fortement la traduction du facteur de régulation Rot en phase stationnaire de croissance (Boisset et al., 2007; Geisinger et al., 2006). De manière intéressante, en phase exponentielle de croissance, le facteur Rot induit des effets directement opposés à l'ARNIII (Said-Salim et al., 2003). En réprimant la traduction de Rot, l'ARNIII renforce donc un peu plus son action sur la régulation des facteurs de virulence. L'ARNIII pourrait également induire des effets transcriptionnels en séquestrant directement des facteurs de régulation. Cependant, à ce jour, il n'existe aucune preuve expérimentale qui confirme cette hypothèse.

## D.2. Conclusions préliminaires

En réponse à la densité cellulaire, AgrA active l'ARNIII qui en fin de phase exponentielle de croissance agit ensuite sur l'expression des gènes de virulence par des mécanismes antisens. Par ailleurs, une récente étude a montré qu'en plus d'activer l'ARNIII, le facteur AgrA agissait indépendamment sur la régulation de nombreux gènes du métabolisme ainsi que sur l'activation des peptides PSM (i.e. l'hémolysine  $\delta$ ) (Queck et al., 2008). Comme l'ARNIII est particulièrement long et code lui même pour l'hémolysine  $\delta$ , il est probable que sa première fonction était uniquement celle d'un ARNm classique (Queck et al., 2008). Il aurait ensuite évolué vers sa fonction d'ARN régulateur afin d'affiner la régulation des gènes de virulence en agissant au niveau post-transcriptionnel. Les premiers résultats ont montré que l'ARNm *spa* est directement réprimé par le domaine 3' de l'ARNIII (Huntzinger et al., 2005). Puisque ce domaine est très conservé chez *S. aureus*, ce résultat suggérait fortement que l'ARNIII puisse réguler d'autres ARNm par des mécanismes analogues.

---

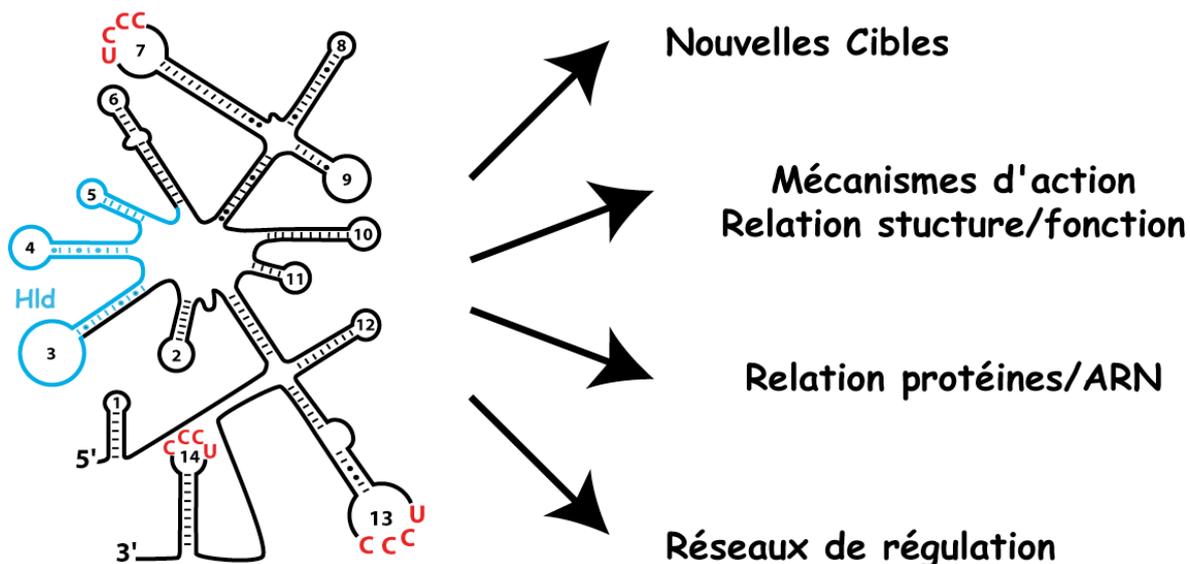
## **Objectifs de la thèse**

Comme décrit précédemment, je me suis donc intéressé au rôle des ARN régulateurs dans l'expression des gènes chez *S. aureus*. Dans un premier temps, j'ai contribué à mieux comprendre le rôle et les mécanismes d'action de l'ARNIII, puis je me suis focalisé sur l'identification et la caractérisation de nouveaux ARN non codants de *S. aureus*, basé sur une analyse bioinformatique couplée à l'étude de leur expression.

## I. Validation de cibles de l'ARNIII

L'ARNIII est constitué de plusieurs domaines fonctionnels distincts qui influencent indépendamment la régulation des gènes (Huntzinger et al., 2005; Morfeldt et al., 1995). De plus, plusieurs de ses structures en tiges-boucles (1, 7, 13 et 14) sont très conservées entre différentes espèces de Staphylocoques et trois d'entre elles possèdent un motif redondant potentiellement complémentaire à la région SD des ARNm (Benito et al., 2000). Ainsi, l'ARNIII pourrait constituer une sorte d'opéron de petits ARN régulateurs. Cependant, même s'il régule à lui seul plus de 40 gènes de virulence (Novick et al., 1993), au début ma thèse seules deux cibles directes avaient été identifiées (Huntzinger et al., 2005; Morfeldt et al., 1995). Ce travail avait donc pour objectif de caractériser de nouvelles cibles directes de l'ARNIII, et de déterminer les différents partenaires impliqués dans les réseaux de régulation de la virulence afin de replacer le rôle de l'ARNIII dans la physiologie globale de la bactérie (figure 21).

Plusieurs cibles directes de l'ARNIII ont été identifiées grâce à l'utilisation combinée d'approches expérimentales *in vivo* et *in vitro*. Les études du mécanisme d'action ont révélé un mécanisme commun de régulation qui implique de longs hybrides d'ARN ou des interactions boucle-boucle. Nos résultats montrent que la topologie des interactions boucle-boucle induit une reconnaissance hautement spécifique par la RNase III. Plusieurs approches ont aussi été réalisées *in vitro* pour tenter de définir la place que tient la RNase III dans les mécanismes de répression des ARNm dépendants de l'ARNIII.



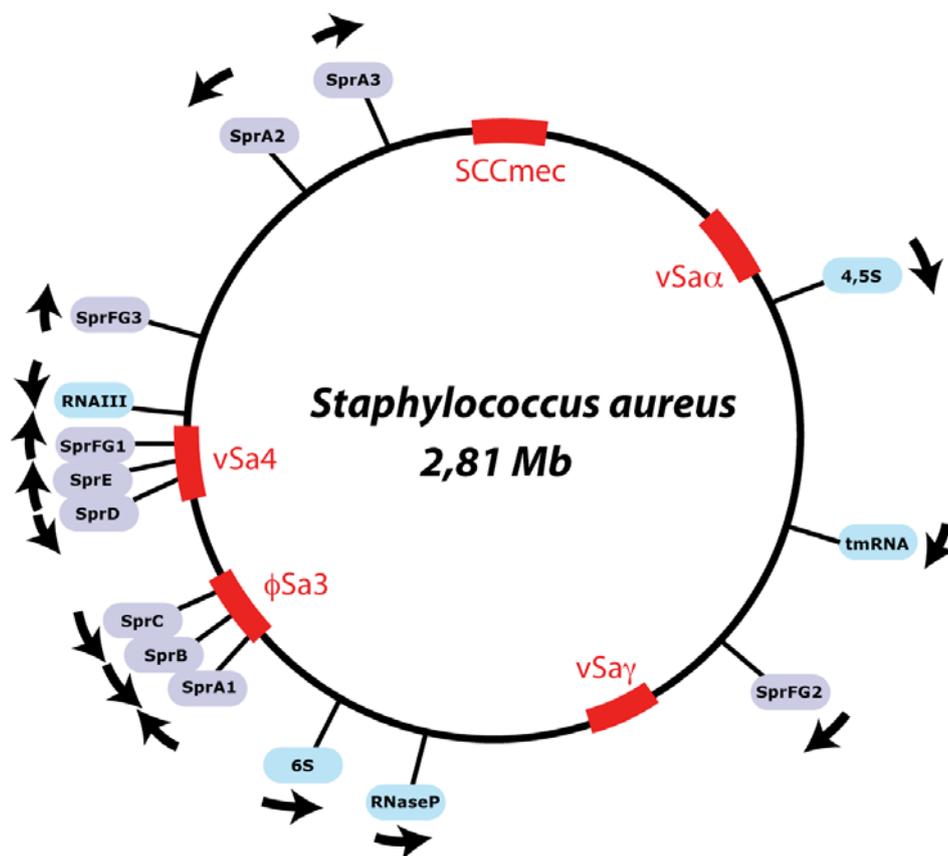
**Figure 21: Objectifs de l'étude réalisée sur l'ARNIII et ses cibles.** Le projet démarré sur l'ARNIII avait pour but d'identifier des cibles directes, de déterminer les mécanismes moléculaires mis en jeu, d'identifier les différents interactants et enfin de resituer ces régulations dans la physiologie de la bactérie.

## II. Identification de nouveaux ARN régulateurs chez *S. aureus*

Les ARN non codants exercent de multiples rôles dans le monde bactérien (cf introduction). Cependant, la plupart des ARNnc de fonction connue ont été identifiés chez les bactéries Gram-négatif comme *E. coli* ou *S. typhimurium* (Gottesman, 2005; Vogel, 2009). Même si récemment plus de 50 nouveaux ARNnc ont été découverts chez *L. monocytogenes* (Mandin et al., 2007; Toledo-Arana et al., 2009), de manière générale, beaucoup moins de données existent chez les bactéries Gram-positif. Pendant longtemps l'ARNIII était le seul ARN régulateur connu chez *S. aureus*.

Au début de ma thèse, différentes études avaient prédit l'existence de nouveaux ARNnc chez *S. aureus* (Pichon et Felden, 2005; Roberts et al., 2006). Parmi ces prédictions, six ARN (SprA-F) ont été validés expérimentalement (Pichon et Felden, 2005). Ces six ARN sont tous codés au niveau d'îlots de pathogénie et deux d'entre eux présentent plusieurs copies sur le génome (Figure 22) (Pichon et Felden, 2005). Si certains de ces ARN sont exprimés différemment selon les souches et pourraient donc contribuer à

l'adaptation à des niches écologiques variées, aucune donnée expérimentale n'a confirmé leur implication dans la virulence. A cette époque, il existait donc un réel vide concernant le rôle de l'ARN dans la régulation des gènes chez *S. aureus*. Ainsi, afin d'approfondir les connaissances dans ce domaine, le deuxième objectif de ma thèse était d'identifier de nouveaux ARNnc chez *S. aureus* et de déterminer les fonctions auxquelles ils sont associés. Cette étude avait pour but d'étoffer les réseaux de régulation des gènes et de préciser les rôles de l'ARN dans ces réseaux. Pour cela, nous avons choisi une approche prédictive basée sur l'analyse des régions intergéniques (IGR) du génome combinée à l'analyse de l'expression des ARN.



**Figure 22: Etat des lieux des ARN régulateurs connus chez *S. aureus* en 2006.** Carte génétique de *S. aureus* qui représente la localisation de l'ARNIII ainsi que de certains ARN non codants conservés (en bleu). En violet sont représentés les ARN régulateurs SprA-F identifiés par Pichon et Felden en 2006. Les zones rouges correspondent aux différents îlots de pathogénicité et les flèches noires indiquent l'orientation des gènes. Mb correspond à Mégabase.

---

# Résultats & Discussions

# I. L'ARNIII et la régulation des gènes de virulence de *S. aureus*

## A. Recherche des cibles et validation expérimentale

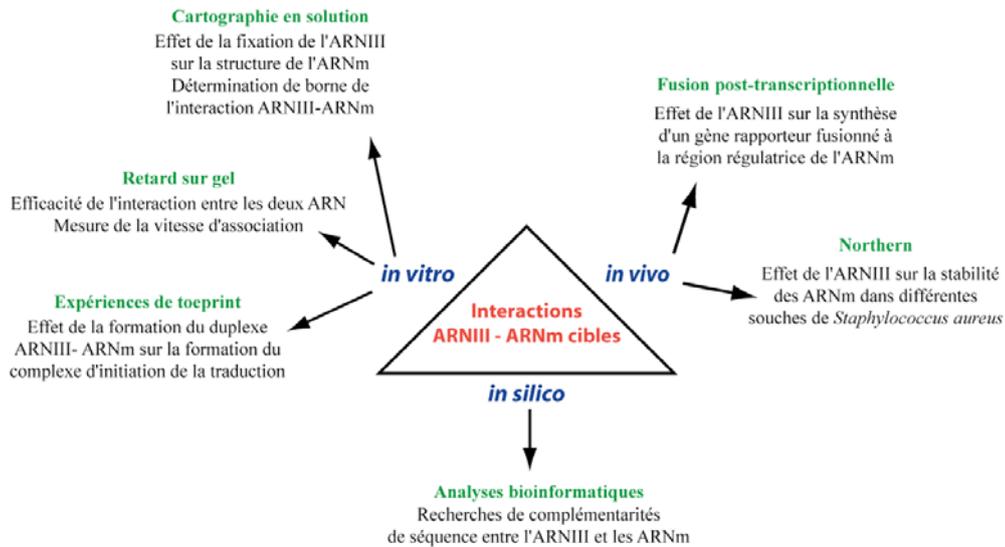
Comme décrit précédemment l'ARNIII, effecteur intracellulaire du système de densité cellulaire, possède une double fonction, celle d'un ARNm qui code pour l'hémolysine  $\delta$ , et celle d'un ARN régulateur. L'organisation modulaire de la structure secondaire de l'ARNIII indique que chacun des domaines structuraux pourrait exercer une fonction spécifique. En effet, par son domaine 5', l'ARNIII agit comme ARN antisens et active la traduction de l'ARNm *hla* et par son domaine 3' il réprime la traduction de l'ARNm *spa*. Cependant ces deux cibles directes n'expliquent toujours pas pourquoi l'ARNIII est requis pour la régulation temporelle d'un grand nombre de facteurs de virulence. Ceci suggérerait que l'une des cibles possible puisse être un facteur global de régulation. Deux possibilités étaient donc envisageables, soit l'ARNIII séquestrerait une protéine de régulation, soit il régulerait au niveau post-transcriptionnel la synthèse de cette protéine. C'est ce deuxième aspect que nous avons privilégié et qui a fait l'objet de la première partie de ma thèse.

Nous avons tout d'abord utilisé une approche *in silico* pour la recherche d'appariements stables entre l'ARNIII et les ARNm. Sachant que de nombreux ARNnc d'*E. coli* qui régule la traduction d'ARNm cibles interagissent avec les régions 5' non codantes incluant le codon d'initiation, nous avons tout d'abord restreint notre recherche sur l'ensemble de ces régions annotées dans le génome de *S. aureus*. En général, la région régulatrice de l'ARNnc est non appariée et très conservée (Marzi et al., 2008; Vogel, 2009). Ainsi, comme la tige-boucle 13 est très conservée et porte une des séquences redondantes riches en C qui interagit avec l'ARNm *spa*, nous avons restreint la recherche des appariements sur ce motif de l'ARNIII. L'approche informatique que nous avons utilisé a été mis en œuvre grâce à Christine Gaspin (INRA, Toulouse). Dans cette approche, la région 5' régulatrice (-30 à +30 par rapport au AUG) de l'ensemble des ARNm annotés dans le génome a été reliée par un "linker" à la séquence de la tige-boucle 13. Ces constructions ont ensuite été repliées *in silico* et l'énergie libre ( $\Delta G$ ) de l'interaction



La validation des appariements a été effectuée par une combinaison d'approches expérimentales. Les approches *in vivo* ont été réalisées dans l'équipe du Prof. F. Vandenesch par S. Boisset (INSERM, Lyon). Des fusions post-transcriptionnelles contenant la région régulatrice de l'ARNm cible fusionnée en phase à un gène rapporteur codant pour la  $\beta$ -galactosidase ont été réalisées. Ces fusions ont été exprimées dans différentes souches de *S. aureus*, exprimant l'ARNIII sauvage ou des ARNIII mutés. Des expériences northern ont également été effectuées dans l'équipe afin de suivre le taux des ARNm endogènes dans différentes souches exprimant ou pas l'ARNIII. Nous avons ainsi analysé la fonction de la protéine Hfq et de l'endoribonucléase III spécifique des ARN en double brin dans la régulation des gènes cibles. Toutes ces données permettent d'évaluer la contribution des interactions ARNIII-ARNm et de ces deux protéines sur la traduction et la stabilité des ARNm. Pour caractériser en détail, les interactions ARNm-ARNIII, j'ai utilisé une batterie de méthodes *in vitro*. Dans un premier temps, la stabilité des complexes et la vitesse d'association ont été évaluées par retard sur gel. Ce dernier aspect est particulièrement important puisque la vitesse d'association est en général l'une des constantes clés pour un contrôle efficace. Toutes ces méthodes ne permettent cependant pas de déterminer les régions des deux ARN impliquées dans l'interaction. C'est pourquoi, j'ai utilisé la cartographie en solution pour obtenir une image précise de la structure secondaire des ARNm cibles ainsi que de la topologie des complexes ARNm-ARNIII. Cette méthode a révélé des caractéristiques communes aux ARNm cibles et mis en évidence des topologies variées qui ne pouvaient pas être prédites par l'approche *in silico*. Enfin, des expériences de toeprint m'ont permis d'analyser l'effet des complexes formés sur l'interaction du ribosome et la formation du complexe d'initiation de la traduction. Nous avons par ailleurs montré que la RNase III est une sonde idéale pour révéler des topologies très différentes de complexes ARN-ARN (figure 24).

L'ensemble des résultats de ce travail est résumé ci-dessous. A mon arrivée dans l'équipe, la validation expérimentale de l'ARNm *SA1000* et *rot* avait déjà été commencée par E. Huntzinger et T. Geissmann. Ma contribution personnelle a été déterminante pour deux des autres cibles, l'ARNm *SA2353* et l'ARNm *coa*, et sur la reconnaissance des différents hybrides par la RNase III.

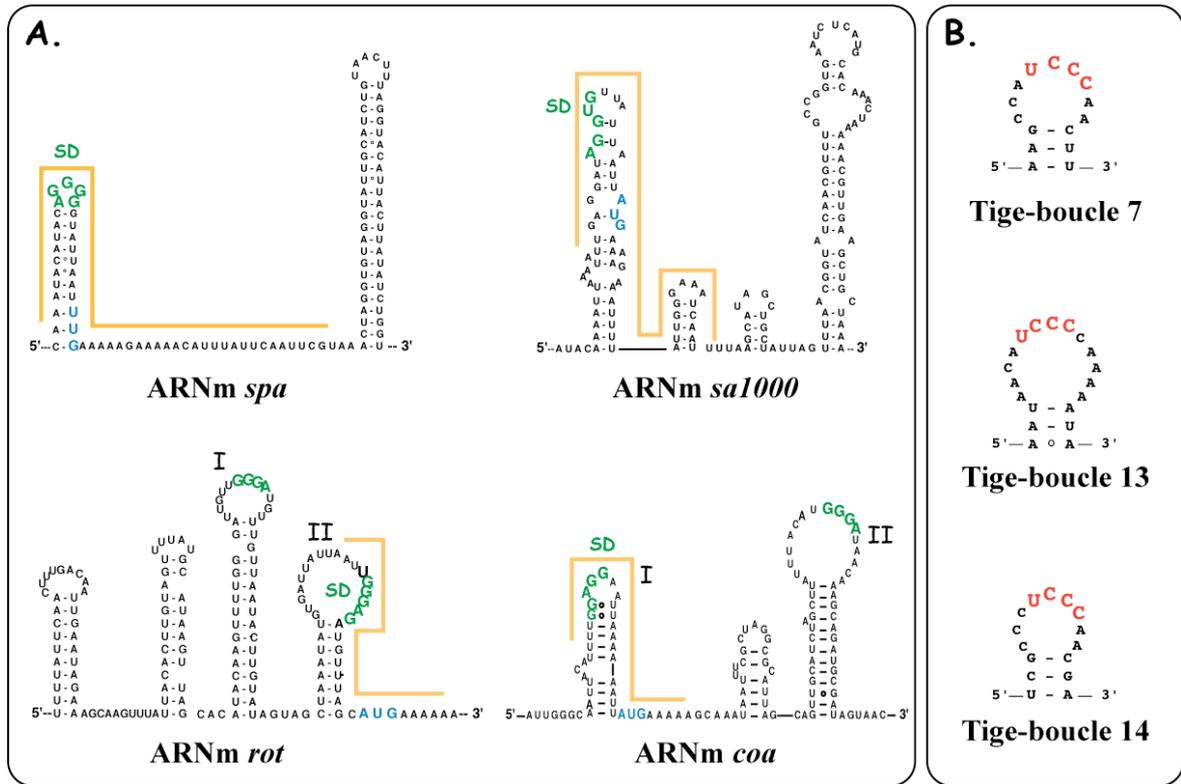


**Figure 24: Stratégie expérimentale pour l'étude des cibles de l'ARNIII.** Les expériences *in vivo* ont été réalisées dans différentes souches de *S. aureus* (i.e. sauvage, délétée de l'ARNIII, délétée de la RNase III, etc...). Les expériences *in vitro* ont été réalisées avec l'ARNIII sauvage ou délété de ses tige-boucles 7, 13 ou 14.

## B. L'ARNIII réprime la synthèse des facteurs SA1000 et Rot

(Article I: Boisset et al., 2007)

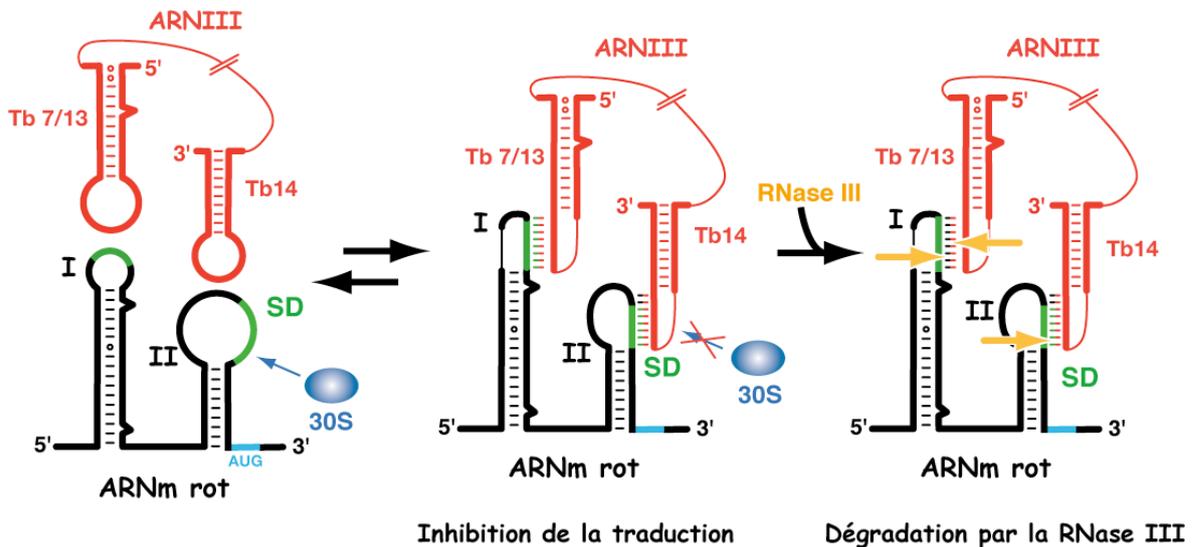
Le facteur SA1000 est conservé dans toutes les souches de *S. aureus* qui ont été séquencées à ce jour. Ce facteur possède le même motif N que certaines MSCRAMM connues (figure 15) et joue un rôle dans l'adhésion, notamment grâce à ces propriétés de liaison au fibrinogène (Liang et al., 2006). En plus de posséder une longue séquence complémentaire avec l'ARNIII, la région 5' régulatrice de l'ARNm *sa1000* à la particularité de se replier d'une manière identique à celle de l'ARNm *spa* (figure 25, A). Nos données *in vitro* et *in vivo* confirment que l'ARNm *sa1000* est également réprimé au niveau post-transcriptionnel par l'ARNIII via un mécanisme analogue, à celui de la protéine A (figure 20). Les expériences de fusion post-transcriptionnelle ont montré que la répression du facteur SA1000 est dépendante d'une interaction ARNIII-ARNm *in vivo*. *In vitro*, la tige-boucle 13 de l'ARNIII se fixe à la séquence SD de l'ARNm *sa1000* et entraîne la formation rapide d'un long duplexe d'ARN. Cet hybride bloque l'accès au ribosome qui ne peut plus initier la synthèse protéique. Le complexe formé devient alors une cible pour la RNase III qui induit non seulement des coupures dans le duplexe mais aussi dans une longue structure en tige-boucle de l'ARN située ~60 nucléotides en aval du codon initiateur.



**Figure 25: Structure secondaire des ARNm cibles de l'ARNIII et de ses tiges-boucles 7, 13 et 14.** (A) Structures secondaires des ARNm *spa*, *sa1000*, *rot* et *coa*. Les codons initiateurs (en bleu) et les séquences complémentaires aux motifs riches en C de l'ARNIII (en vert) sont présentés. SD désigne la séquence Shine Dalgarno des ARNm alors que la ligne jaune représente les séquences complémentaires à la tige-boucle 13 de l'ARNIII qui ont été identifiées par analyse bioinformatique. (B) Tige-boucles 7, 13 et 14 de l'ARNIII. Les motifs riches en C sont marqués en rouge.

Le facteur Rot est un régulateur transcriptionnel de la famille Sar. Ce facteur est particulièrement intéressant car il a la particularité d'induire des effets directement antagonistes à ceux de l'ARNIII. En effet, contrairement à l'ARNIII, le facteur Rot active les protéines de surface et réprime les exoprotéines (Said-Salim et al., 2003). La région 5'UTR de l'ARNm *rot* présente plusieurs tiges-boucles dont deux contiennent une séquence de type SD redondante et complémentaire à la séquence riche en C des tiges-boucles 7, 13 et 14 de l'ARNIII (figure 25, A et B). Bien que l'ARNm *rot* puisse potentiellement former un long duplexe avec la tige-boucle 13 de l'ARNIII, nos résultats ont révélé que l'ARNIII se fixe à l'ARNm *rot* par l'intermédiaire de deux interactions de type boucle-boucle. L'utilisation combinée de la RNase III et de mutants de l'ARNIII a permis de déterminer les nucléotides impliqués dans ces interactions boucle-boucle. Ainsi, la tige-boucle 14 de l'ARNIII interagit au niveau de la séquence SD de l'ARNm *rot* (tige boucle II) alors que la tige-boucle 7 ou 13 forme la deuxième interaction au niveau d'une région de l'ARNm (tige boucle I) située ~60 nucléotides en amont du codon AUG (figure 26). Ces deux interactions impliquent 6 à 7 nucléotides mais sont suffisamment

stables pour bloquer l'accès au ribosome. Elles génèrent également la formation d'une longue structure hélicoïdale créant un site spécifique pour la RNase III. Cependant, contrairement aux ARNm *spa* et *sa1000*, cette coupure par la RNase III ne conduit pas à une dégradation totale de l'ARNm.



**Figure 26: L'ARNIII réprime l'ARNm rot par deux interactions de type boucle-boucle.** Tb7, 13 et 14 correspondent aux tiges-boucles 7, 13 et 14 de l'ARNIII. La région bleue représente le codon initiateur, les régions vertes les sites riches en G de l'ARNm rot, et les flèches jaunes les coupures induites par la RNase III. 30S et SD représentent la petite sous-unité du ribosome et la séquence Shine Dalgarno, respectivement.

### C. L'ARNIII réprime la synthèse du facteur SA2353

Le facteur SA2353 code pour un facteur de virulence homologue à SsaA de *S. epidermidis*. Ce facteur possède des propriétés immunostimulantes qui peuvent causer la prolifération anarchique des lymphocytes T et ainsi affaiblir le système immunitaire. SsaA est, entre autre, à l'origine de certains cas d'endocardites chez l'homme (Lang et al., 2000). Plus récemment, il a été proposé que ce facteur appartiendrait à une famille d'amidases impliquée dans la synthèse du peptidoglycane (Dubrac et al., 2008; Dubrac et Msadek, 2004). LytM et SA2093, sont deux autres cibles de l'ARNIII qui appartiennent à cette famille (figure 27). Ainsi, la répression coordonnée de ces protéines en phase stationnaire serait essentielle pour éviter des dommages cellulaires causés par une trop forte expression de ces peptidoglycane hydrolases.

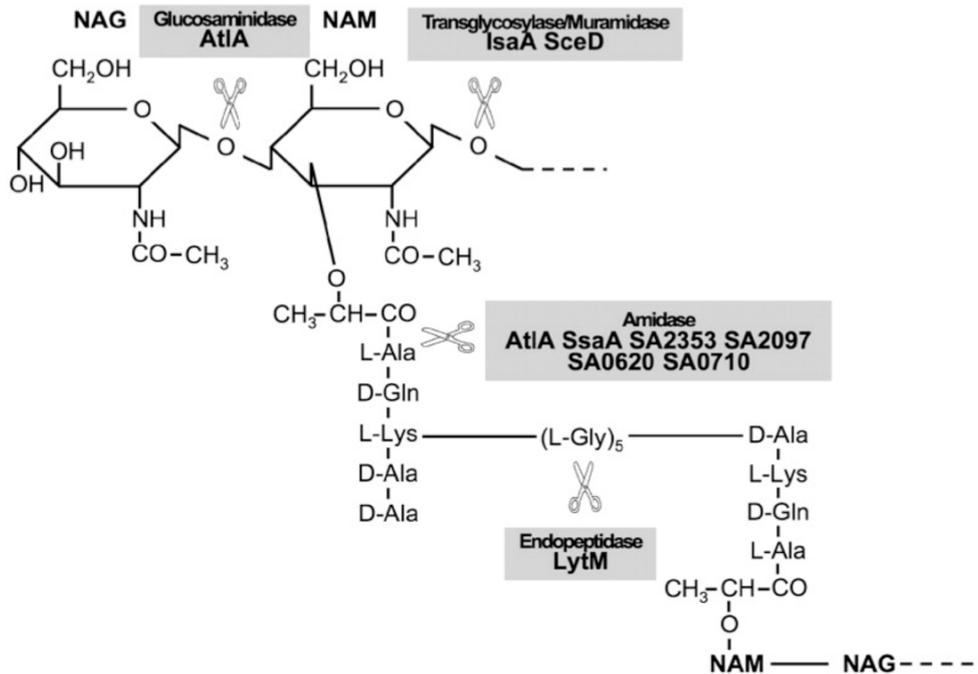
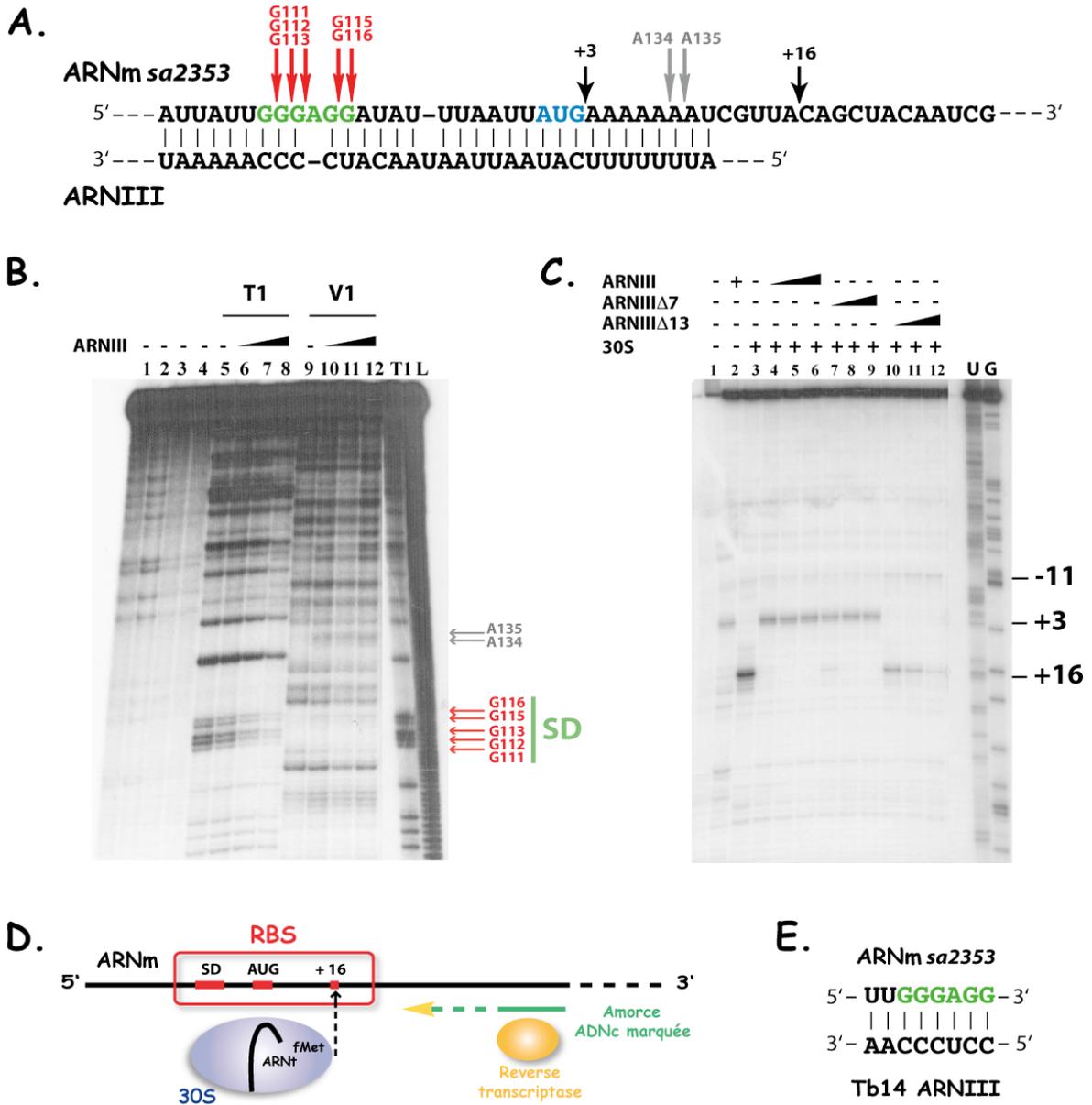


Figure 27: Les facteurs SA2353 et LytM sont responsable du remodelage du peptidoglycane.

### C.1. L'ARNIII s'hybride au niveau de la séquence SD de l'ARNm *sa2353* *in vitro*

Comme pour les cibles précédentes, le site de reconnaissance du ribosome de l'ARNm *sa2353* possède une forte complémentarité avec la tige-boucle 13 de l'ARNIII (Figure 28, A). Ainsi, après avoir déterminé le +1 de l'ARNm *sa2353* (-126/AUG) par la technique de 5'RACE ("**R**andom **A**mplification of **c**DNA **E**nd"), des expériences de cartographie de l'ARNm en solution ont été réalisées (cf annexe 2). Deux sondes enzymatiques ont été testées sur l'ARNm seul ou en présence d'ARNIII: la RNase T1 coupe spécifiquement en 3' des résidus guanines non appariés et la RNase V1 clive les régions en double brin. Cette méthode a permis de cartographier les modifications structurales induites par la fixation de l'ARNIII sur l'ARNm, et donc de déterminer précisément les sites d'interaction. Les résultats de ces expériences sont donnés (figure 28, B).



**Figure 28: La tige boucle 13 de l'ARNIII réprime la traduction de l'ARNm *sa2353*.** (A) Prédiction bioinformatique de l'hybride formé entre la tige-boucle 13 de l'ARNIII et la région RBS de l'ARNm *sa2353*. Les flèches rouges représentent les coupures protégées par l'ARNIII contre la RNase T1, les flèches grises les coupures par la RNase V1 en présence d'ARNIII et les flèches noires, les pauses observées lors des expériences de "toeprint". La séquence SD et le codon initiateur sont représentés en vert et en bleu respectivement (B) Expérience de cartographie en solution. Les lignes 1 à 4 représentent les contrôles d'incubation. Les lignes 5 à 8 et 9 à 12 correspondent aux digestions par la RNase T1 et V1, respectivement. Ces digestions ont été réalisées sur l'ARNm seul (-) ou en présence d'une concentration croissante en ARNIII (50, 100 ou 200 nM). T1 désigne l'échelle T1 et L, l'hydrolyse alcaline. Les flèches rouges et grises suivent la même nomenclature que pour (A). (C) Expérience de "toeprint". Les lignes 1, 2 et 3 représentent les contrôles d'incubations réalisés sur l'ARNm seul, en présence de la sous-unité 30S ou de l'ARNIII respectivement. La sous-unité 30S a été mise en compétition avec différentes concentrations d'ARNIII (25, 50 et 100 nM; lignes 4 à 6), d'ARNIIIΔ7 (ligne 7 à 9) ou d'ARNIIIΔ13 (lignes 10 à 12). +3 et +16 représentent les positions par rapport au codon initiateur. (D) Principe des expériences de "toeprint". La reverse transcriptase allonge une amorce d'ADN complémentaire à l'ARNm d'intérêt. Le ribosome bloque la reverse transcriptase en position +16, pause appelée "toeprint". (E) Séquence complémentaire entre la tige-boucle 14 de l'ARNIII et la région Shine Dalgarno de l'ARNm *sa2353*.

Bien que la région 5' UTR de l'ARNm *sa2353* soit relativement longue, la séquence SD est cependant accessible à la RNase T1 (figure 28, A et B). La présence de nombreuses coupures par la RNase V1 suggère que la région 5' non codante de cet ARNm est structurée. Cependant n'ayant pas effectué les sondes chimiques pour analyser la réactivité des positions Watson-Crick des bases, nous n'avons pas proposé de modèle de structure secondaire pour cette région. La formation du complexe avec l'ARNIII induit des changements de coupures par les RNases restreints à la région de fixation du ribosome. L'ARNIII protège la séquence SD (G111-113, G115-116) contre l'hydrolyse par la RNase T1. De manière concomitante, des coupures par la RNase V1 apparaissent en position 134 et 135 (figure 28, A et B). Ces résultats sont donc en faveur de la formation d'un duplexe relativement étendu prédit par l'analyse bioinformatique.

## **C.2. L'ARNIII inhibe la formation du complexe d'initiation de la traduction**

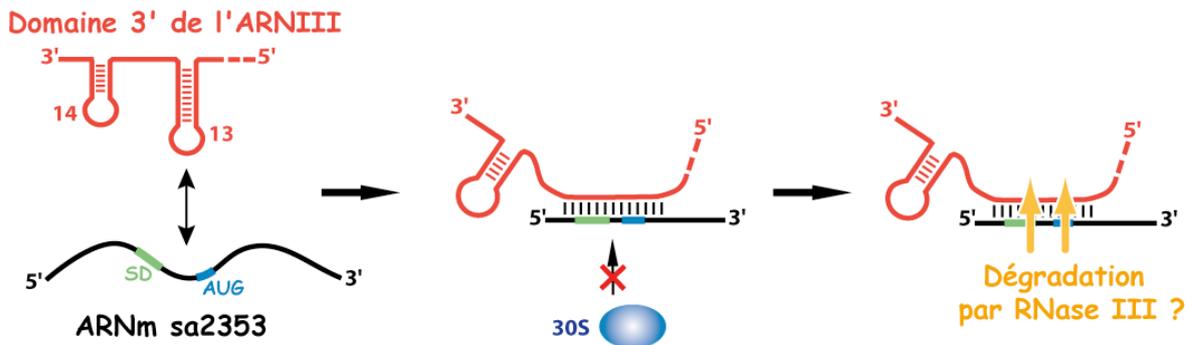
Puisque l'ARNIII s'hybride à la séquence SD de l'ARNm *sa2353*, nous avons ensuite regardé par des expériences de "toeprint" si cette interaction était suffisamment stable pour bloquer la formation du complexe d'initiation de la traduction. Au cours des expériences de "toeprint", la reverse transcriptase (RT) allonge une amorce d'ADN complémentaire à l'ARNm cible. Si le ribosome reconnaît l'ARNm, ce dernier bloque l'élongation de la RT qui induit une forte pause appelée "toeprint", située généralement 16 nucléotides en aval du codon initiateur (figure 28, D) (Hartz et al., 1988). Si l'ARNIII empêche la fixation du ribosome, la reverse transcriptase n'est plus stoppée et la pause au niveau du "toeprint" disparaît. Les résultats de ces expériences sont donnés (figure 28, C).

Lorsque l'ARNIII est ajouté à l'ARNm *sa2353*, une forte pause de la RT en position +3 apparaît correspondant à la signature de la formation du duplexe. Une pause beaucoup plus faible est aussi observée proche de la séquence SD en position -11. Cette signature s'accompagne de la disparition complète du toeprint, suggérant que l'ARNIII bloque l'accès au ribosome. La délétion de la tige-boucle 7 de l'ARNIII (ARNIII- $\Delta$ 7) ne perturbe pas l'activité inhibitrice de l'ARNIII. En revanche, la délétion de la tige-boucle 13 (ARNIII- $\Delta$ 13) s'accompagne d'une perte de la pause en position +3, et d'une intensification de la pause de la RT en position -11. En présence de l'ARNIII- $\Delta$ 13, la

formation du complexe d'initiation se forme mais celui-ci diminue aux fortes concentrations de l'ARNIII- $\Delta$ 13. Ainsi la tige-boucle 13 est l'élément régulateur principal de l'ARNIII (figure 28, C). Nous avons cependant observé que la boucle 14 (CCUCCCA) peut former une interaction de 8 nucléotides avec la séquence SD de l'ARNm (figure 28, E). Ceci pourrait expliquer la pause de la RT en position -11, et le fait que le toeprint soit diminué aux fortes concentrations de l'ARNIII- $\Delta$ 13 (figure 28, C).

### C.3. Conclusion

Les expériences de cartographie en solution et de "toeprint" permettent d'établir un modèle selon lequel l'ARNIII serait directement impliqué dans la répression post-transcriptionnelle de l'ARNm *sa2353*. Dans ce modèle, la tige-boucle 13 s'hybride au niveau de la séquence SD de l'ARNm *sa2353* et bloque ainsi l'initiation de la traduction (figure 29). Ces résultats confortent les expériences de northern qui ont montré que l'ARNm *sa2353* est détecté en phase post-exponentielle dans les souches de *S. aureus* n'exprimant pas l'ARNIII, ou exprimant l'ARNIII délété de la tige-boucle 13. Inversement, l'ARNm n'est plus détectable dans les souches exprimant l'ARNIII, le domaine 3' (contenant les tiges-boucles 13 et 14), ou le domaine 3' délété de la tige-boucle 14 (voir papier joint, Boisset et al. 2007). Cette régulation dépendante de la tige-boucle 13 de l'ARNIII reste inchangée dans la souche n'exprimant plus Hfq. En revanche, une accumulation de l'ARNm est observée dans la souche n'exprimant plus la RNase III. Ainsi, comme pour l'ARNm *spa* et *sa1000*, la répression de la synthèse de la protéine SA2353 dépend à la fois de l'ARNIII et de la RNase III.

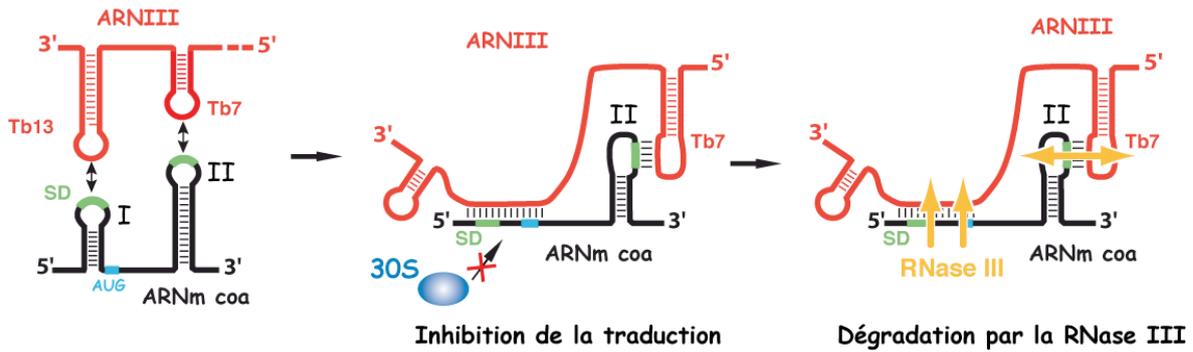


**Figure 29: Modèle de régulation de l'ARNm *sa2353*.** La tige-boucle 13 de l'ARNIII séquestre la séquence SD de l'ARNm *sa2353* et bloque ainsi sa traduction. Ce duplexe serait ensuite probablement dégradé par la RNase III afin de rendre la répression irréversible.

## D. L'ARNIII réprime la synthèse de la staphylocoagulase

(Article II: Chevalier et al., soumis)

La coagulase est une protéine de 60 kDa qui participe activement à l'adhésion et à l'agglutination des cellules pendant les premières étapes de l'infection. Elle joue donc un rôle important dans la virulence de *S. aureus*, notamment lors de la phase de colonisation de l'hôte. Comme les cibles précédentes, ce facteur est dépendant du système *agr in vivo* (Lebeau et al., 1994) et la région 5' régulatrice de son ARNm possède une longue séquence complémentaire avec la tige-boucle 13 de l'ARNIII. Cependant, cet ARN est particulièrement intéressant car il partage certaines caractéristiques communes avec les ARNm *spa* et *sa1000* d'une part, et avec l'ARNm *rot* d'autre part. En effet, la structure secondaire l'ARNm *coa* se replie d'une manière similaire aux ARNm *spa* et *sa1000*, et comme l'ARNm *rot* sa séquence porte une deuxième région riche en G (tige-boucle II) complémentaire aux motifs C des tiges-boucles 7, 13 ou 14 de l'ARNIII (figure 25, A et B). Du fait de ces différentes caractéristiques, le mécanisme par lequel l'ARNIII réprime l'ARNm *coa* implique des topologies structurales comparables à celles retrouvées dans les mécanismes de répression de la protéine A et du facteur Rot. Ainsi, la tige-boucle 13 de l'ARNIII interagit avec la séquence SD de l'ARNm *coa* (tige-boucle I) ce qui mène à l'extension d'un long duplexe étendu d'ARN. Parallèlement une deuxième interaction de type boucle-boucle se forme entre la tige-boucle 7 de l'ARNIII et la tige-boucle II de l'ARNm *coa*, située 30 nucléotides en aval du codon AUG (figure 30). Ces deux interactions deviennent ensuite une cible pour la RNase III. Alors que la RNase III induit plusieurs coupures au niveau du duplexe étendu, elle n'induit qu'une seule coupure forte et spécifique dans l'interaction boucle-boucle. Par ailleurs, plusieurs données *in vitro* et *in vivo* montrent que seule la formation du duplexe étendu est essentielle pour réprimer efficacement la traduction de l'ARNm *coa*. Le rôle de l'interaction boucle-boucle serait de stabiliser le duplexe mais également de faciliter sa dégradation par la RNase III.



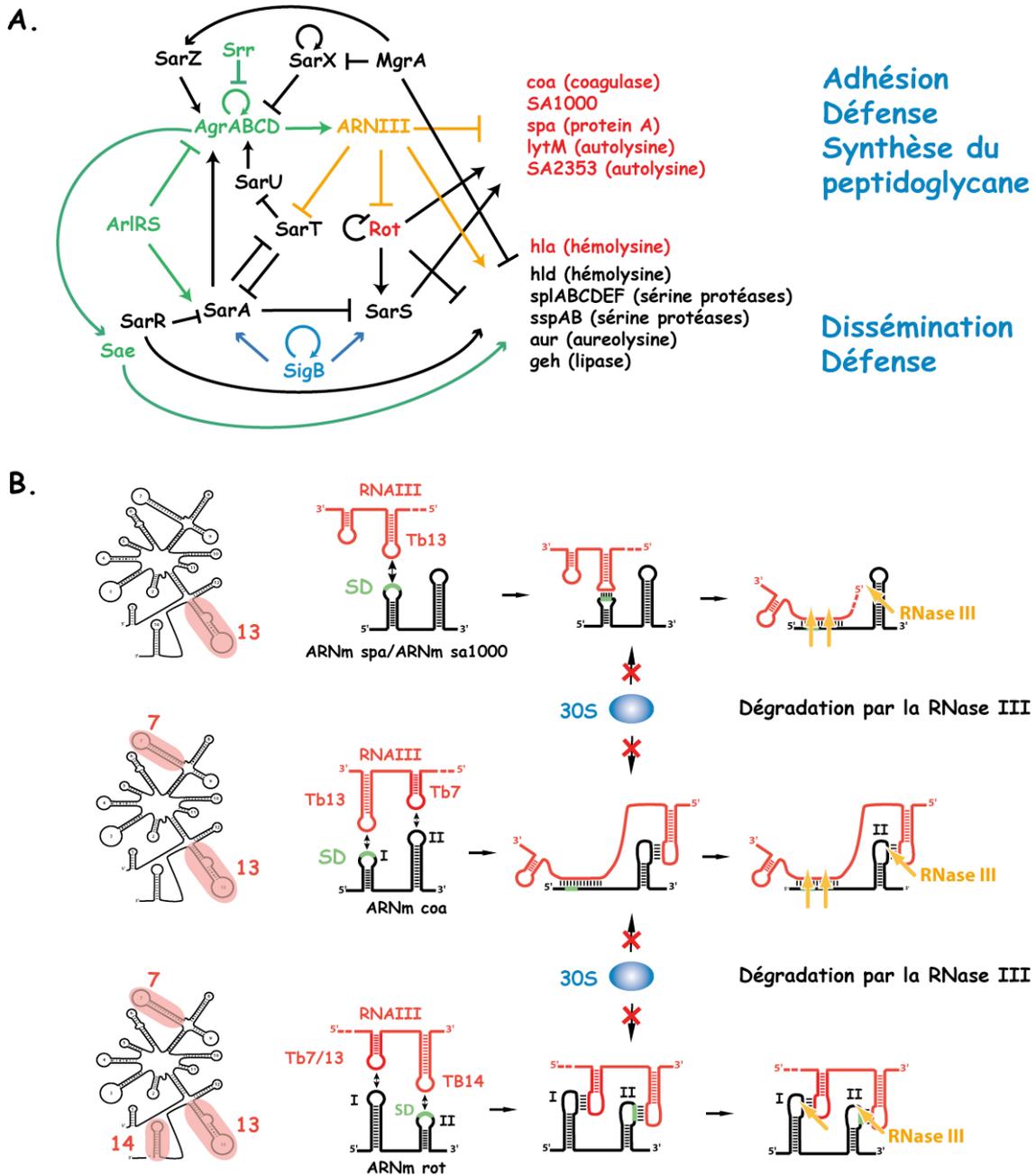
**Figure 30: L'ARNIII réprime l'ARNm *coa* par l'intermédiaire d'une longue interaction étendue et d'une interaction boucle-boucle.** Tb7 et 13 correspondent aux tiges-boucles 7 et 13 de l'ARNIII. La région en bleu représente le codon initiateur, les régions en vert les sites riches en G de l'ARNm *coa*, et les flèches jaunes les coupures induites par la RNase III. 30S et SD représentent la petite sous-unité du ribosome et la séquence Shine et Dalgarno respectivement.

## E. Discussion

### L'ARNIII utilise un mécanisme commun pour réguler de multiples gènes

Notre étude montre que l'ARNIII est un ARN multifonctionnel qui régule plusieurs ARNm cibles à travers un réseau de régulation complexe (figure 31, A). Parmi ces cibles, nous avons montré que l'ARNIII réprime directement toute une classe de facteurs de virulence (protéine A, SA1000, SA2353 et coagulase) ainsi que le régulateur transcriptionnel Rot. Dans chaque cas, l'ARNIII agit au niveau post-transcriptionnel et procèdent par des mécanismes de type antisens (figure 31, B). Dans nos différents modèles, cet ARN utilise ses motifs riches en résidus C pour séquestrer les séquences SD de ses ARNm cibles. L'ARNIII utilise donc un motif commun pour coordonner l'expression de gènes aux fonctions variées. Ce mode de régulation est un moyen efficace d'inhiber rapidement la traduction des ARNm le moment venu. Bien que théoriquement l'ARNIII pourrait interagir avec tous les ARNm dont la séquence SD est accessible, il ne régule en réalité que des ARNm possédant des propriétés bien spécifiques. En effet, nos résultats montrent qu'un seul appariement limité de bases G/C ne suffit pas pour maintenir une interaction stable qui bloque la fixation du ribosome. Pour interagir efficacement avec l'ARNIII, un ARNm doit posséder des caractéristiques de séquence et de structure favorables qui permettent, soit la formation d'au moins deux interactions de type boucle-boucle, soit la formation d'un long duplexe étendu d'ARN. Ces interactions médiées par des appariements G/C rappellent le système CopA/CopT (cf introduction, § **III.C.2.1**). Ainsi, elles seraient essentielles pour initier le contact entre les deux partenaires et

permettrait une conversion rapide vers un complexe plus stable. Ce type d'interactions pourrait notamment expliquer pourquoi la protéine Hfq n'est pas requise dans les mécanismes d'action de l'ARNIII puisque dans chaque exemple ce dernier interagit rapidement avec ces cibles.



**Figure 31: L'ARNIII régule de multiples gènes à travers un réseau de régulation complexe.** (A) Réseau de régulation de l'ARNIII, en rouge sont présentées les cibles dont la répression post-transcriptionnelle dépendante de l'ARNIII a été validée expérimentalement. (B) Bilan des mécanismes d'action de l'ARNIII. Selon les cibles l'ARNIII induit la formation de long duplexe étendu (ARNm *spa*, *sa1000*, *sa2353* et *coa*) ou des interactions de type boucle-boucle (ARNm *rot* et *coa*)

### **L'ARNIII régule la transcription de gènes cibles de manière indirecte**

Il avait été postulé que l'ARNIII utiliserait ses motifs redondants pour séquestrer un facteur de transcription et ainsi réguler indirectement de nombreux gènes (Novick, 2003a). Or notre analyse suggère que l'ARNIII agit principalement au niveau post-transcriptionnel, et réprime la synthèse de la protéine Rot. Au moment où nous étions en train de finaliser nos données expérimentales, Geisinger & Novick ont apporté les évidences *in vivo* de cette régulation. Notre travail montre que cette régulation est gouvernée par la formation de deux interactions boucles-boucles identiques. Rot réprime la transcription de nombreuses exoprotéines en phase exponentielle et active la synthèse de plusieurs facteurs de virulence dont la protéine A et la coagulase (Said-Salim et al., 2003). Ainsi en régulant la traduction de l'ARNm *rot*, l'ARNIII induit des effets indirects multiples, comme par exemple l'activation des exoprotéines. Par ailleurs, nos données prédictives suggèrent que l'ARNm *sarT* pourrait être une autre cible. Les facteurs de transcription SarT et Rot sont tout deux requis pour lever l'activité de répresseur transcriptionnel de SarA au promoteur du gène *sarS*. Ainsi, ces deux facteurs conduisent à une répression de *hla* et à une activation de *spa* (Oscarsson et al., 2005; Oscarsson et al., 2006).

Ces résultats montrent la complexité des réseaux de régulation des facteurs de virulence. En effet, un même gène peut être régulé de différentes façons par l'ARNIII, de manière indirecte au niveau transcriptionnel (en réprimant Rot), et directement au niveau post-transcriptionnel. Par ailleurs, la protéine A est aussi dégradée par diverses exoprotéines en phase stationnaire. Ainsi, réguler les facteurs de virulence de manière fine aurait un impact biologique important lors de l'infection bactérienne pour favoriser la transition d'adhésion des bactéries aux tissus de l'hôte vers l'étape de dissémination.

### **La relation répression de la traduction-dégradation des ARNm, un rôle pour la RNase III**

Les mécanismes d'action de l'ARNIII sont tous liés à l'action de la RNase III. Dans nos trois systèmes nous avons observé au moins deux sites forts de reconnaissance pour la RNase III (figure 31, B). Cette dernière interviendrait dans la dégradation des hybrides ARNIII-ARNm et permettrait notamment de rendre les mécanismes de répression complètement irréversible. *In vivo*, les expériences de fusions post-transcriptionnelles ont montré que la RNase III est absolument nécessaire pour obtenir une répression efficace des

ARNm *spa*, *sa1000* et *coa*. Par contre, la répression de l'ARNm *rot* semble plus dépendante de l'inhibition traductionnelle causée par l'ARNIII que de la dégradation par la RNase III. Cette observation amène à la question: pourquoi les gènes réprimés par l'ARNIII sont-ils plus ou moins dépendant de l'action de la RNase III *in vivo*? Une des raisons possibles de la dégradation rapide des ARNm *spa* et *sa1000* serait que ces ARN sont stables. Ainsi, pour obtenir une répression irréversible, ces ARN doivent être rapidement dégradés. Il se peut aussi que la régulation ait co-évolué avec la fonction des gènes cibles. Le facteur Rot régule simultanément de nombreux gènes, il joue donc un rôle important dans la physiologie globale de la bactérie (Said-Salim et al., 2003). Une dégradation partielle de son ARNm serait un moyen pour la cellule d'anticiper le changement des conditions environnementales. En effet, en cas de stress, le maintien de l'ARNm *rot* dans le cytoplasme permettrait le démarrage immédiat de sa traduction, et ainsi de répondre rapidement aux nouveaux besoins de la cellule.

Je discuterai plus en détail les rôles possibles de la RNase III de *S. aureus* dans le chapitre suivant.

Downloaded from [www.genesdev.org](http://www.genesdev.org) on June 2, 2007



## Staphylococcus aureus RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism

Sandrine Boisset, Thomas Geissmann, Eric Huntzinger, Pierre Fechter, Nadia Bendridi, Maria Possedko, Clément Chevalier, Anne Catherine Helfer, Yvonne Benito, Alain Jacquier, Christine Gaspin, François Vandenesch and Pascale Romby

*Genes & Dev.* 2007 21: 1353-1366

Access the most recent version at doi:[10.1101/gad.423507](https://doi.org/10.1101/gad.423507)

---

**Supplementary data**

"Supplemental Research Data"

<http://www.genesdev.org/cgi/content/full/21/11/1353/DC1>

**References**

This article cites 46 articles, 25 of which can be accessed free at:

<http://www.genesdev.org/cgi/content/full/21/11/1353#References>

**Email alerting service**

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#)

---

**Notes**

---

To subscribe to *Genes and Development* go to:  
<http://www.genesdev.org/subscriptions/>

---

© 2007 Cold Spring Harbor Laboratory Press



*Signalement bibliographique ajouté par :*

**L'UNIVERSITÉ DE STRASBOURG**  
**Service Commun de la Documentation**

**Staphylococcus aureus RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism**

Sandrine BOISSET, Thomas GEISSMANN, Eric HUNTZINGER, et al.

**Genes & Development, 2007, n° 21, pages 1353-1366**

**Résultats et discussions : p. 88-101**

La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.

Les utilisateurs de l'UdS peuvent consulter cette publication sur le site de l'éditeur :

<http://dx.doi.org/10.1101/gad423507>

La version imprimée de cette thèse peut être consultée à la bibliothèque ou dans un autre établissement via une demande de prêt entre bibliothèques (PEB) auprès de nos services :

<http://scd.unistra.fr/services/peb/>

## Supplementary Material

### ***SA1000 is a novel fibrinogen-binding protein.***

SA1000 was initially identified as a putative fibrinogen-binding protein by its primary sequence similarity with Efb (also designated Fib) (Boden and Flock, 1994), an extracellular adhesin of *S. aureus* known to bind fibrinogen and fibronectin (Signas et al., 1989). Here we show that SA1000 carries the properties of a fibrinogen-binding protein. Indeed, fibrinogen and fibronectin adhered to SA1000-coated plates in a dose-dependent manner using enzyme-linked immunosorbent binding assays (Fig. S2A). Furthermore, the addition of increasing concentrations of soluble recombinant SA1000 inhibited the binding of fibrinogen (5 µg/ml) and fibronectin (at 1 µg/ml) to coated SA1000. The estimated IC50 value (inhibitory concentration, *i.e.* concentration required to inhibit 50% of binding) was 4 nM and 1,5 nM for fibrinogen and fibronectin, respectively (Fig. S2B). We then determined whether SA1000 is a cell-surface protein. The protein was cloned with a C-terminus His-Tag and expressed in *S. aureus* RN6390. Western blot analysis was performed using commercial anti-His antibodies on the supernatant and cell extracts treated by lysostaphin (Fig. S2C). The protein was mostly present in the total cell extract fraction and barely detectable in the supernatant. A 2% SDS cell-surface extraction method also suggested that SA1000 was non-covalently associated to the bacterial cell-surface. Indirect immunofluorescent assay confirmed that SA1000-His is present on the cell surface of RN6390 expressing SA1000- His (Fig. S2D).

### ***Purification of recombinant SA1000 protein.***

To express His tagged recombinant SA1000 in *S. aureus*, a vector similar to pLUG389 but containing a start codon and two cohesive restriction sites was constructed in pLUG274

forming pLUG345. The coding sequence of SA1000 was amplified using RN6390 as template with primers SA1000-5' and SA1000-3' (carrying six additional His codons; Table S1), restricted with *StuI* and *BamHI*, and ligated to pLUG345 cleaved with the same enzymes. The resulting plasmid (pLUG550) was introduced in RN6390. The coding sequence of SA1000 was also amplified using RN6390 as template with primers SA1000-F and SA1000-R (Table S1), restricted with *EcoRI* and *XhoI*, and ligated into the pET43 vector (Qiagen) (modified by the adjunction of a StrepTag (WSHPQFEK) on the 5' side of the multiple cloning site) digested with the same enzymes. The resulting plasmid (pLUG693) was co-transformed with pRARE (a plasmid designed to enhance the expression of proteins that contain codons rarely used in *E. coli*, Novagen) into *E. coli* strain BL21 DE3 (Novagen) (strain LUG1105). The resulting protein carries a N-terminal StrepTag and a C-terminal hexa-Histidine tagging.

A crude extract was prepared from 2 l of culture of *E. coli* LUG1105 in LB medium containing ampicilline at 50 µg/ml and chloramphenicol at 17 µg/ml. The cells were lysed for 1 h at 4°C in a buffer (Qiagen) containing 1 mM phenylmethanesulfonyl fluoride, 0.5 M NaCl, phosphate, 10 mM imidazole and 1 mg/ml final of lysozyme. After centrifugation, the supernatant was incubated with 10 ml of a 50% Ni-nitriloacetic acid affinity resin and the protein was eluted with increasing concentrations of imidazole. The protein was further purified using a StrepTactin affinity chromatography (IBA-GmbH), and stored in PBS buffer.

#### ***Direct binding of fibrinogen and fibronectine immobilized sa1000***

Microtiter wells (Costar) were coated with SA1000 at 5 µg/ml in PBS at 20°C overnight. The wells were blocked by incubation with 5% non-fat dry milk in PBS-Tween 0.05% (PBS-T) for 1 h at 37°C. After washing 3 times with PBS-T, Fibrinogen (Fg) or fibronectin

(Fn) were added at concentration ranging from 0.005-25 µg/ml, and the plates were incubated for 1 h at 37°C. Bound Fg was detected with rabbit anti-human Fg-horseradish peroxidase-conjugated antibodies (diluted 5000-fold), and bound Fn was detected using rabbit antibodies against human fibronectin (diluted 2000-fold), followed by secondary antibodies conjugated with HRP (diluted 5000-fold). After washing the wells with PBS-T, bound Fg or Fn were detected with 75 µl of chromogenic substrate (0.09% tetramethylbenzidine, 0.0001% H<sub>2</sub>O<sub>2</sub> in 0.1 M sodium acetate buffer pH 5), for 15 min in dark. The reactions were stopped by the addition of 75 µl of 1 M of H<sub>2</sub>SO<sub>4</sub> and the absorbance at 450 nm was measured in an ELISA plate reader.

Competitive inhibition of Fg or Fn binding to SA1000-coated microtiter was performed with plates coated with SA1000 (5 µg/ml) and blocked with 5% non fat dry milk in PBS and 0.05% Tween-20 (PBS-T) for 1 h at 37°C. Increasing concentrations of SA1000 (0-10 nM) were incubated for 1 h at 37°C in PBS-T in wells in the presence of Fg (5 µg/ml) or Fn (1 µg/ml). After washing the wells with PBS-T, bound Fg or Fn were detected as described above.

#### ***SA1000 western immunoblotting***

LUG1141 (RN6390/pLUG550) was grown in 10 ml of Brain Heart broth at 37°C for 18 h and then centrifuged at 5,000 x g for 10 min. Total proteins were extracted from the pellet by lysostaphin treatment followed by centrifugation at 13,000 x g for 20 min. The supernatant was then loaded on 15% SDS-PAGE followed by electroblotting to nitrocellulose membrane (Biorad). The membrane was blocked with 5% non fat dry milk in Tween 0.01% Tris-Base Saline buffer (TTBS), 1 h at 20°C. Anti-histidine-horseradish peroxidase monoclonal antibodies (Dako 1/2000, overnight 4°C) were used to detect recombinant SA1000. To analyse whether SA1000 is a cell-surface associated protein, the

LUG1141 bacterial pellet was resuspended in 1 ml extraction buffer (10 mM Tris-HCl pH 7.0, 2% SDS), heated at 95°C for 2 min, and then centrifuged at 13,000 x g for 10 min. To 20 µl of liquid supernatant, 4 µl of 6X sample buffer was added, and the mixture was then heated at 95°C for 3 min and loaded on 15% SDS-PAGE followed by western blotting.

#### ***Detection of the surface location of SA1000 by indirect immunofluorescence***

LUG1141 (RN6390/pLUG550) cells were grown overnight, centrifuged and suspended in PBS to an optical density at 600 nm of 1. The bacterial suspension (5 µl) was dried on a immunofluorescence slide (SuperFrost® Plus, Menzel-Glaser) and then fixed with a solution of paraformaldehyde 4% diluted in PBS (15 min 20°C). After gentle washing in PBS (5 min), fixed cells were blocked with 4% bovine serum albumine and 10% normal goat serum during 1 h at 20°C, and then exposed at 1:50 dilution of mouse monoclonal antibodies against His-tagging (Dako) in blocking buffer (4°C overnight). After 3 washes with PBS, a 1:100 dilution of chicken FITC-conjugated anti-mouse Ig (Alexa Fluor® 488, Invitrogen) was added (1 h at 20°C). The slides were again washed 3 times with PBS, dried and, after the addition of mounting medium (FluoroPrep, bioMérieux), examined by fluorescence microscopy. Control slides without monoclonal antibodies against His-tagging but with chicken FITC-conjugated anti-mouse Ig gave no signal.

#### ***Extracellular proteome analysis***

Strains were grown to post-exponential phase by inoculating 50 ml of BHI medium with an overnight culture (1:100), at 37 °C for 7 h. The supernatants (1.2 ml) were precipitated by cold acetone, and the pellets washed with cold ethanol. The proteins were dissolved in 100 µl of sample buffer (7 M urea, 2 M thiourea, 4 % CHAPS, 50 mM DTT) with 4 µl of Tris-HCl 1 M pH 8 and 10 µl of iodoacetamide 1 M. The samples were incubated 15 min

at 37 °C for alkylation, desalted by a G-25 Sephadex chromatography, and used to re-swell pre-cast pH 3-10 IPG strips (Biorad). Isoelectric focusing was performed using a IEF Cell (Biorad) for a total of 75 kVh. IPG strips were buffer-exchanged in 1 M Tris-HCl pH 8.45, 6 M urea, 2 % SDS, 30 % glycerol. The strips were loaded onto the top of a 12 % gel and the second dimension was carried out at 20 mA for 16 h (Schagger and von Jagow, 1987). After electrophoresis, gels were stained with colloidal blue and analyzed using the PQuest software (Biorad). Proteins were identified by mass spectroscopy.

### ***Plasmid construction***

Total DNA and plasmid DNA were prepared with standard methods (Sambrook & Maniatis, 1989). Transformation of *Escherichia coli* DH5 $\alpha$  was performed by treatment with CaCl<sub>2</sub>, and *S. aureus* strains were transformed by electroporation (Bio-Rad gene pulser).

RNAIII was expressed in *Staphylococcus aureus* WA400 by using plasmid pE194 (Horinouchi and Weisblum, 1982), modified by adding an *EcoRV* restriction site next to the *XbaI* site forming pLUG274 (Benito et al., 2000). The P3 region was amplified with primers agr-sa1819/agr-sa1569 and the transcriptional terminator with primers agr-sa1095/agr-sa751 using PWO DNA polymerase (Roche). Blunt-ended PCR products generated were ligated together and re-amplified with the external primers to be cloned onto pLUG274 (forming pLUG 389). This plasmid pLUG389 containing only the P3 promoter and the transcriptional terminator was used as a negative control. To express the 3' domain of RNAIII with a mutation in the transcriptional terminator, nucleotides 494 to 502 (CCCUCCTAA) of the loop 14 were replaced by GAGA. The region from the P3 promoter to the 5' stem of the hairpin 14 was amplified from pLUG324 with primers agr-sa1819/agr-sa1074 (carrying the additional nucleotides GAGA) and the region

corresponding to the 3' stem of hairpin 14 was amplified using primers agr-sa1068/agr-sa751, using PWO DNA polymerase. Blunt-ended PCR products were then cloned onto pLUG274 (forming pLUG661) as described above. To express the 3' domain of RNAIII deleted in loop 13 and in loop 14, a similar construct was made using the same oligonucleotides, but PCR were performed from pLUG302 which express RNAIII deleted in loop 13 (Huntzinger et al., 2005), forming pLUG667.

### ***Deletion of the hfq gene by allelic replacement***

The deletion/replacement  $\Delta hfq/cat$  mutant of *S. aureus* RN6390 (LUG911) was obtained by using pMAD, a thermosensitive plasmid, which contains a constitutively expressed  $\beta$ -galactosidase gene. This allows positive selection of double crossing over by following the  $\beta$ -galactosidase activity on X-gal agar plates (Arnaud et al 2004). The chloramphenicol transacetylase gene *cat* was cloned in pMAD between two DNA fragments corresponding respectively to the chromosomal regions upstream and downstream the *hfq* coding sequence using primers hfq211/hfq995 and hfq1201/hfq2031 (Table S1). The resulting plasmid, pLUG533, was electroporated successively into RN4220, and RN6390. Transformants were grown at the non-permissive temperature (37°C) to select cells, which have integrated the plasmid into the chromosome by homologous recombination. To favor the second recombination event, a single colony was grown at 30°C for 10 generations and plated at 37°C overnight. Cells, which have lost the plasmid vector through a double crossing over event, were detected on Xgal agar plates. PCR amplifications were used to confirm the loss of *hfq* gene, and antibodies against Hfq (Huntzinger et al., 2005) were used to visualize the protein.

### ***Construction of translation fusions***

Translation fusions were constructed by using plasmid pTCV-*lac*, a low-copy-number promoter-less *lacZ* vector (Poyart & Trieu-Cuot, 1997), which had been modified by deletion of a region encompassing the Shine & Dalgarno sequence and the AUG sequence of *lacZ* forming pLUG220. Leader regions of SA1000 and *rot* mRNAs were cloned downstream the *rpoB* (RNA polymerase) gene promoter in frame with *lacZ* (Huntzinger et al., 2005). SA1000-mRNA (nts +1/+90) was amplified using primers SA1000-1009/SA1000-1089, and *rot*-mRNA (nts +1/+290) was amplified using primers rot1310/rot1600 (Table S1). Selected plasmids were controlled by PCR and sequencing before being electroporated into *S. aureus* RN4220 and then into other staphylococcal backgrounds.

### ***β-galactosidase measurements***

*S. aureus* containing *lacZ* fusions were grown in BHI broth to exponential growth phase. Cells (1 ml) were lysed using a FastPrep Instrument (QBiogen). Protein concentrations were determined in the lysates by using the Bradford method and specific β-galactosidase activity was measured with the β-galactosidase Enzyme Assay System with Reporter Lysis Buffer Kit (Promega) as recommended by the provider. All assays were carried out three times on duplicate cultures.

### ***RNA preparation***

RNAIII, RNAIII derivatives (RNAIII-Δ7-9 deletion of nts G207 to U319, RNAIII-Δ13 deletion of nts U409 to A451, and RNAIII-Δ14 deletion of nts G483 to C511), the isolated hairpins 7, 13 and 14 and mRNA fragments were transcribed *in vitro* using T7 RNA polymerase as previously described (Benito et al., 2000). All RNAs carry at their 5' end

two additional guanines, which have no effect on the RNA folding. *In vitro* 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, RNAs were renatured by incubation at 90°C for 2 min in RNase-free water, 1 min at 4°C 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).

### ***RNase III preparation and RNA hydrolysis***

Recombinant RNase III, containing six histidines at its N-terminus, was purified by affinity chromatography using Ni<sup>2+</sup>-beads. The purified RNase III was stored in 25 mM Tris-HCl pH 8.0, 150 mM KCl, 1 mM DTE, 0.1 mM EDTA, 50% glycerol at -20°C.

5' end-labeled SA1000 or *rot* mRNA or RNAIII ( $3 \times 10^{-8}$  M) were incubated with a five-fold excess of the unlabeled RNAIII or mRNAs respectively, in 10 µl TMN buffer containing 20 mM Tris-acetate, pH 7.5, 10 mM magnesium-acetate, 150 mM Na-acetate, 1 mM DTT at 37°C for 15 min. Full RNA duplexes between RNAIII and mRNA were formed by incubation at 90°C for 2 min followed by slow cooling to 37°C in TMN buffer. Cleavages of end-labeled RNA either free or in complex, were performed at 37°C from 30 sec to 15 min with 8 pmol of RNase III in the presence of 1 µg of carrier tRNA. Reactions were stopped by phenol extraction followed by RNA precipitation.

### ***Bandshift analysis of different RNAIII/mRNA complexes***

Binding of <sup>32</sup>P-labelled RNAIII to a ten fold excess of unlabelled SA1000 mRNA was performed at 37°C in TMN buffer as previously described (Huntzinger et al., 2005). Samples were withdrawn at various time points (0-30 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 hours and subsequently dried. Bands corresponding to the

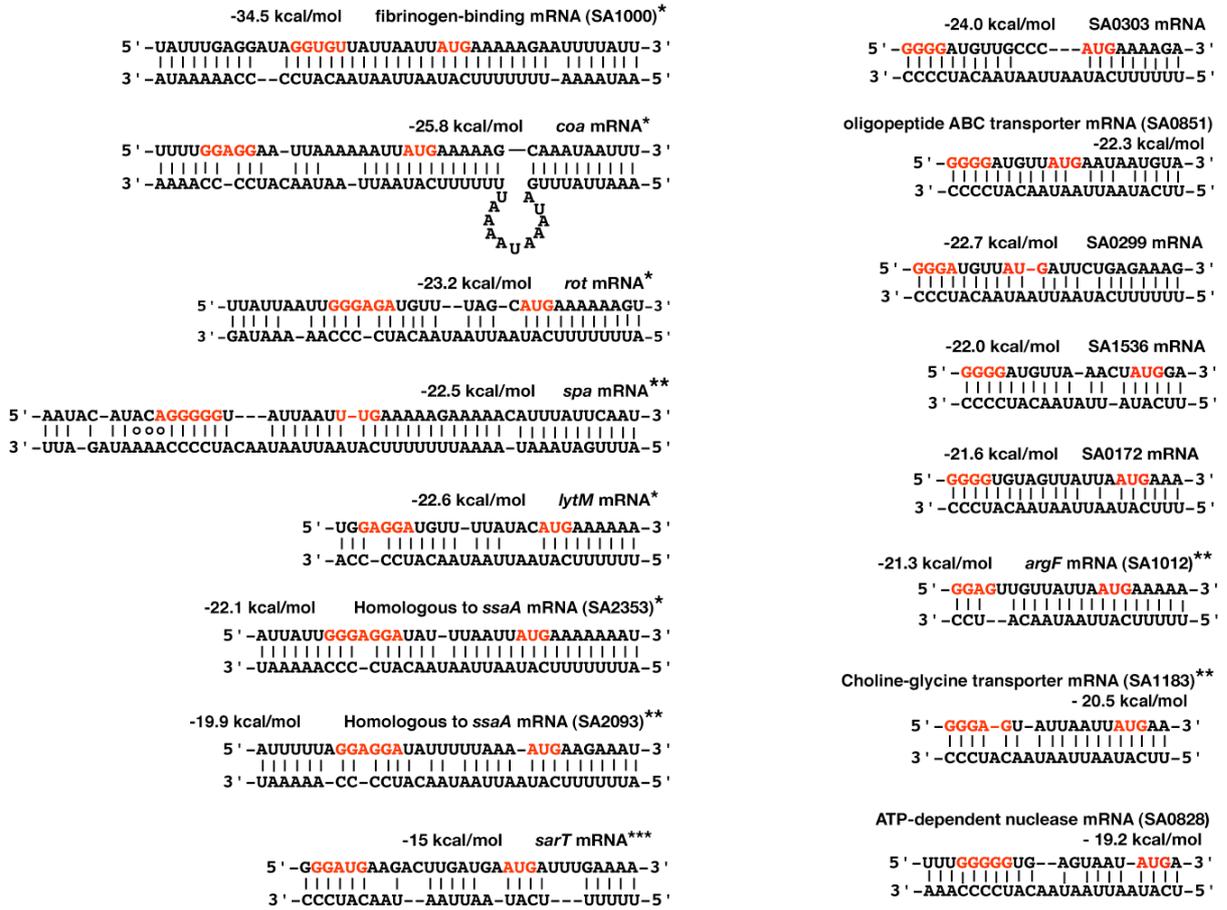
RNAIII/SA1000 complex and free RNAIII, respectively, were quantified using a Bio-imager Analyser BAS 2000 (Fuji).

For determination of the apparent dissociation rate constant of RNAIII/SA1000 mRNA or RNAIII/*rot* mRNA complexes, <sup>32</sup>P-labelled mRNA was incubated with increased molar amount of wild-type or mutant RNAIII for 15 min at 37°C in TMN buffer. Samples were then treated as described above.

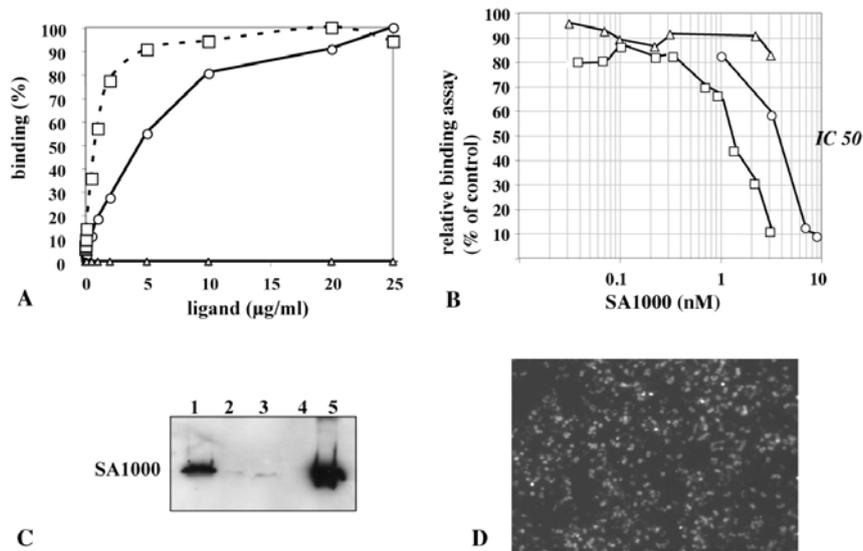
## References

- Arnaud M, A. M. Chastanet and Debarbouillé. 2004. New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl Environ Microbiol* **70**: 6887-91
- Benito, Y., F.A. Kolb, P. Romby, G. Lina, J. Etienne and F. Vandenesch. 2000. Probing the structure of RNAIII, the *Staphylococcus aureus agr* regulatory RNA, and identification of the RNA domain involved in repression of protein A expression. *RNA* **6**: 668-679.
- Boden, M.K. and J.I. Flock. 1994. Cloning and characterization of a gene for a 19 kDa fibrinogen-binding protein from *Staphylococcus aureus*. *Mol Microbiol* **12**: 599-606.
- Dunman, P.M., E. Murphy, S. Haney, D. Palacios, G. Tucker-Kellogg, S. Wu, E.L. Brown, R.J. Zagursky, D. Shlaes and S.J. Projan. 2001. Transcription profiling-based identification of *Staphylococcus aureus* genes regulated by the *agr* and/or *sarA* loci. *J Bacteriol* **183**: 7341-7353.
- Huntzinger, E., S. Boisset, C. Saveanu, Y. Benito, T. Geissmann, A. Namane, G. Lina, J. Etienne, B. Ehresmann, C. Ehresmann, A. Jacquier, F. Vandenesch and P. Romby. 2005. *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate *spa* gene expression. *EMBO J* **24**: 824-835.

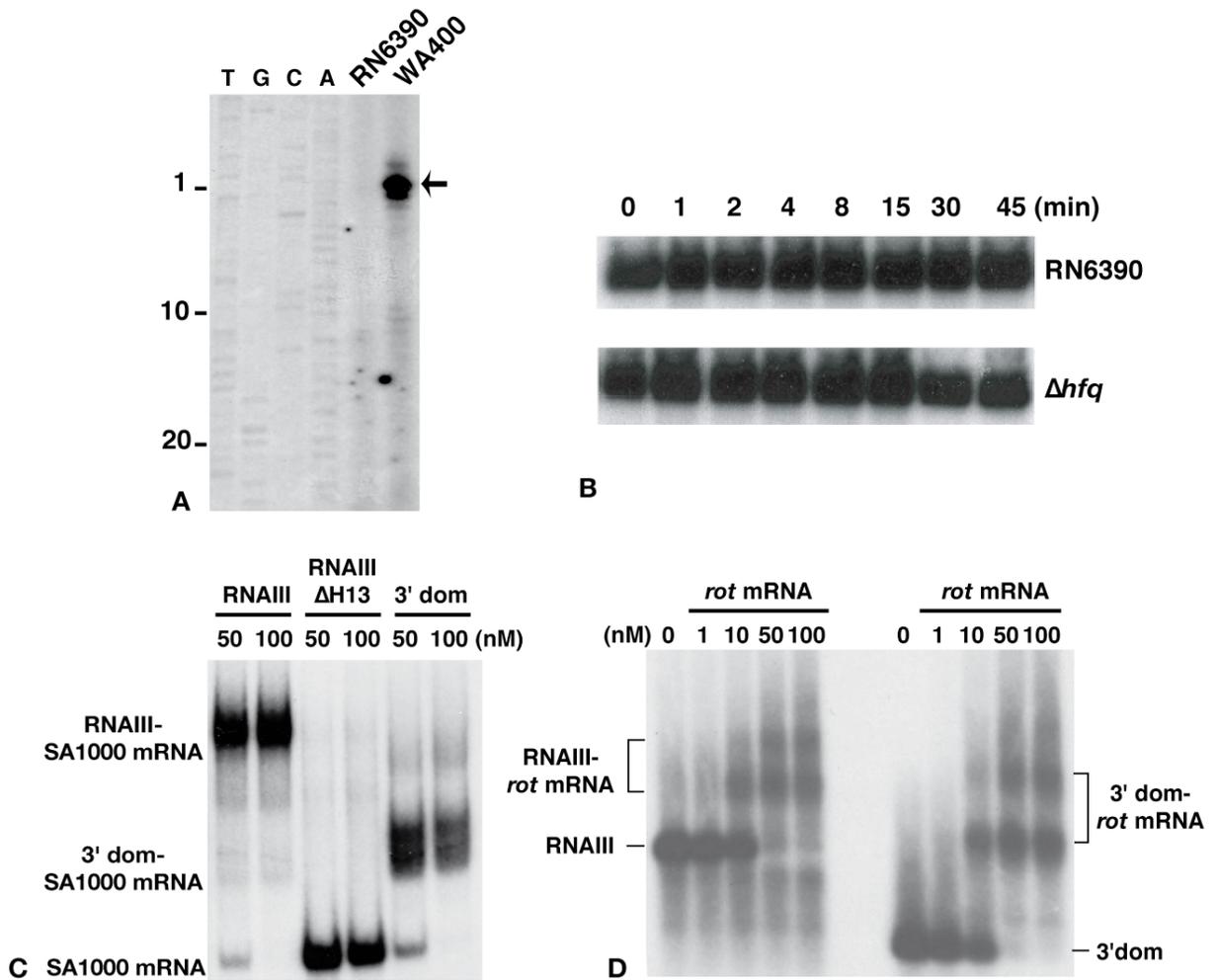
- Horinouchi, S. and B. Weisblum. 1982. Nucleotide sequence and functional map of pE194, a plasmid that specifies inducible resistance to macrolide, lincosamide, and streptogramin type B antibiotics. *J Bacteriol* **150**: 804-814.
- Hussain, M., K. Becker, C. von Eiff, J. Schrenzel, G. Peters and M. Herrmann. 2001. Identification and characterization of a novel 38.5-kilodalton cell surface protein of *Staphylococcus aureus* with extended-spectrum binding activity for extracellular matrix and plasma proteins. *J Bacteriol* **183**: 6778-6786.
- Poyart, C. and P. Trieu-Cuot. 1997. A broad-host-range mobilizable shuttle vector for the construction of transcriptional fusions to beta-galactosidase in gram-positive bacteria. *FEMS Microbiol Lett* **156**: 193-198.
- Sambrook, J., E. Fritsch and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual*.
- Schagger, H. and G. von Jagow. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166**: 368-379.
- Schmidt, K.A., A.C. Manna and A.L. Cheung. 2003. SarT influences sarS expression in *Staphylococcus aureus*. *Infect Immun* **71**: 5139-5148.
- Signas, C., G. Raucci, K. Jonsson, P.E. Lindgren, G.M. Anantharamaiah, M. Hook and M. Lindberg. 1989. Nucleotide sequence of the gene for a fibronectin-binding protein from *Staphylococcus aureus*: use of this peptide sequence in the synthesis of biologically active peptides. *Proc Natl Acad Sci U S A* **86**: 699-703.



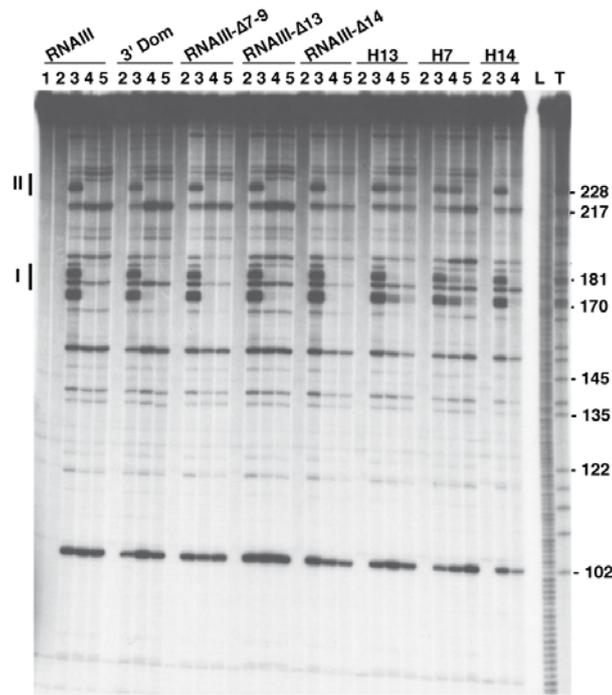
**Figure S1: Potential base-pairings between the hairpin 13 of RNAIII and target mRNAs.** (left panel) mRNA targets encoding virulence factors or transcriptional regulators (Rot et SarT). (Right panel) mRNA targets encoding for cell-surface proteins. The Shine and Dalgarno sequence and the AUG codon are indicated in red characters. \*The expression of several mRNAs was shown regulated by RNAIII (see references in legend of Fig. 1). \*\*These mRNAs were not detected in a mutant *S. aureus* strain deleted of the *agr* operon which does not express RNAIII (Dunman et al., 2001). \*\*\*The expression of SarT is negatively regulated by the *agr* operon (Schmidt et al., 2003). No star: regulation unknown. Nomenclature of genes as given for the *S. aureus* N315: *ssaA*, secretory antigen protein ; *rot*, repressor of toxins ; *spa*, protein A ; *coa*, coagulase precursor ; *lytM*, peptidoglycane hydrolase ; *sarT*, transcriptional regulator ; *argF*, ornithine carbamoyl-transferase ; SA0172, similar to integral membrane protein LMRP ; SA0851, oligopeptide ABC transporter; SA1536, predicted permease ; SA0303, sodium coupled permease ; SA0299, similar to carbohydrate kinase ; SA0828, similar to ATP-dependent nuclease subunit A.



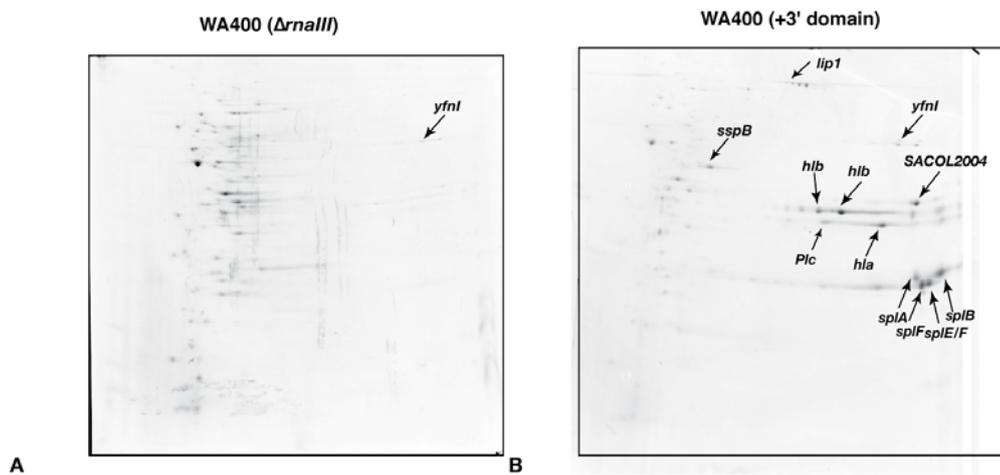
**Figure S2: SA1000 is a fibrinogen-binding protein.** (A) Binding assays of fibrinogen or fibronectin to immobilized recombinant SA1000. Increasing concentrations of Fg (circles) or Fn (squares) were incubated with recombinant SA1000 (5 μg/ml). Data represents the mean of three experiments. (B) Inhibition assays using increasing concentrations of SA1000 (0-100 μg/ml) to Fg (5 μg/ml) (circles), Fn (1 μg/ml) (squares) or Bovine serum albumine (triangles) incubated with SA1000-coated microtiter wells (5 μg/ml). Results are expressed in % of binding as compared to the control incubated with Fg or Fn alone in the absence of the competitor. Data represents the means of three experiments. (C) Detection of recombinant His tagged SA1000 protein by western blot of cell surface and secreted proteins from RN6390 using HRP-labeled anti-His monoclonal antibodies. (lane 1) : recombinant SA1000 (10 μg); (lane 2) : cell-secreted protein in the supernatant; (lane 3) : cell-secreted protein in 10 X concentrated supernatant; (lane 4) : untreated cells; (lane 5) : cell-surface proteins extracted with 2% SDS. (D) Detection of recombinant SA1000 on the surface of *S. aureus* by indirect immunofluorescence microscopy. *S. aureus* RN6390 (pLUG550) strains were incubated with mouse anti-His monoclonal antibodies (1/50) overnight at +4°C. Bound antibodies were detected using chicken FITC-labelled anti-mouse antibodies (1/100) and visualized by immunofluorescence microscopy (magnification x1000).



**Figure S3: Primer extension analysis, half-life of RNAIII and analysis of the formation of different RNAIII-mRNA complexes.** (A) Primer extension analysis of SA1000 mRNA prepared from RN6390 (*rnaIII*<sup>+</sup>, *agr*<sup>+</sup>) and WA400 ( $\Delta rnaIII$ ). The arrow indicates the transcriptional start. (B) Comparison of RNAIII levels as a function of time, prepared from RN6390 (*rnaIII*<sup>+</sup>) and RN6390 deleted of the *hfq* gene ( $\Delta hfq$ ) after treatment with rifampycin. (C) Analysis of the RNAIII/SA1000 mRNA complex. 5'-end-labeled SA1000 mRNA was incubated alone or in the presence of two concentrations of RNAIII, RNAIII- $\Delta 13$ , or the 3' domain (50 and 100 nM). (B) Visualization of the RNAIII/*rot* mRNA complexes. 5'-end-labeled RNAIII or its 3' domain was incubated alone or in the presence of increasing concentrations of *rot* mRNA (1, 10, 50, 100 nM). The fraction of labeled RNA associated with mRNA was calculated from the counts in the corresponding band relative to the total counts in the lane. The apparent  $K_d$  value was determined as the concentration of RNAIII allowing 50% of mRNA binding.



**Figure S4: Enzymatic probing of several RNAIII/rot mRNA complexes.** (A) RNase T1 hydrolysis of 5'-end-labeled *rot* mRNA either free (lane 3) or in the presence of two concentrations (lane 4, 100 nM; lane 5, 300 nM) of wild type or mutated RNAIII. Lanes 1, 2: Incubation controls in the absence (1) or in the presence of wild type or mutated RNAIII (2). Lanes T, L: RNase T1 under denaturing conditions and alkaline ladders, respectively. Bars denote the main reactivity changes induced by complex formation.



**Figure S5: Extracellular proteomes of *S. aureus* strains.** A comparative analysis of the exoproteins was done using WA400 ( $\Delta$ *rnaIII*) and the same strain in which the 3' domain was expressed. The corresponding protein spots are labelled with protein names according to the *S. aureus* N315: lip1, triacylglycerol lipase; yfnI, a putative sulfatase; sspB, cysteine protease; hlb, hla,  $\beta$  and  $\alpha$ -hemolysins, respectively; Plc, 1-phosphatidylinositol phosphodiesterase precursor; SplA-F, serine proteases.

***Staphylococcus aureus* RNAIII binds to two distant regions of *coa* mRNA to arrest translation initiation and to promote mRNA degradation**

Chevalier Clément<sup>1</sup>, Sandrine Boisset<sup>2</sup>, Benoit Masquida<sup>1</sup>, Pierre Fechter<sup>1</sup>, Thomas Geissmann<sup>2</sup>, François Vandenesch<sup>2§</sup> and 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>INSERM U851, Centre National de Référence des Staphylocoques, F-69008 Lyon, France; Université de Lyon, F-69008 Lyon, France.

Running title: Post-transcriptional regulation of *coa* mRNA

§ To whom correspondence should be sent: François Vandenesch ([denesch@univ-lyon1.fr](mailto:denesch@univ-lyon1.fr)) & Pascale Romby ([p.romby@ibmc.u-strasbg.fr](mailto:p.romby@ibmc.u-strasbg.fr))

**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. 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. We show here that RNAIII represses the synthesis of the staphylocoagulase through a direct binding with the mRNA. Structure mapping indicates that two distant regions of RNAIII interact with the ribosome binding site and the coding region of *coa* mRNA. This 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 (RNase III) cleaves the two regions of the mRNA bound to RNAIII that may contribute to the degradation of the repressed mRNA. Thus, *coa* mRNA belongs to the RNAIII-dependent repressed mRNAs that are regulated by a shared mechanism although the topologies of the inhibitory complexes are significantly different due to the mRNA context.

**Author summary**

*Staphylococcus aureus* causes a large number of diseases in humans, and is one the main cause of hospital-acquired infections. *S. aureus* produces a high number of virulence factors which are expressed in a coordinated manner and at appropriate time and space. This regulation involved the coordinated action of multiple trans-acting regulators including regulatory proteins and RNA. 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 of action of RNAIII on the synthesis of one major virulence factor, staphylocoagulase. We found that RNAIII regulates the expression of staphylocoagulase through direct interactions with the 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 shared mechanism. Our study illustrates the diversity of RNAIII-mRNA topologies and how these multiple RNAIII-mRNA interactions would be decisive for virulence regulation.

## INTRODUCTION

*Staphylococcus aureus* is ubiquitous in the environment and occurs as a commensal on human skin. *S. aureus* is also a major human pathogen, which is the most common causes of the hospital-acquired infections and in the community. Therefore, *S. aureus* has developed a plethora of strategies to experience and 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 adhesion and tissue, 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 components system, global regulatory proteins, and the quorum sensing system [1,2]. The 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 contains a density-sensing cassette (*agrD* and B) and a two-component sensory transduction system (*agrA* and C). Upon a threshold level of the 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 RNAIII-independent and the other one 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].

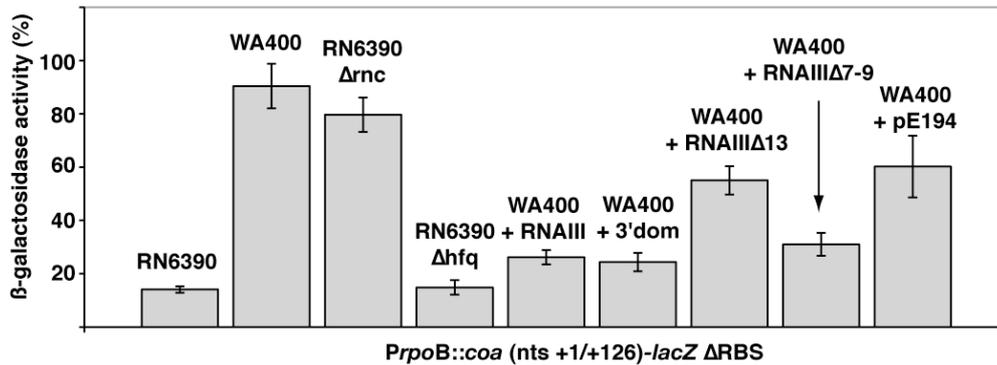
RNIII has a dual function since it acts as a mRNA which encodes a PSM peptide,  $\alpha$ -hemolysin, and temporally controls the switch between early expression of surface proteins and late expression of several exotoxins [1]. RNIII belongs to the class of *trans*-acting RNAs which regulate several mRNAs at the post-transcriptional level [7,8]. The 5' domain of RNIII activates translation of *hla* mRNA (encoding  $\alpha$ -hemolysin) by preventing the formation of an intramolecular mRNA structure that sequesters the *hla* ribosome binding site [1,9]. The 3' end and the central domain of RNIII 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 RNIII 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 all *S. aureus* strains, 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 is produced during the exponential growth and rapidly repressed by the *agr* system [13].

We show here that RNIII 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 RNIII with *coa* mRNA. The complex is rapidly formed and 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 RNIII-dependent repressed mRNAs that are regulated by a similar mechanism. This work and previous data also illustrate the variety of RNIII-mRNA topologies that are sufficient to block the access of the ribosome at the initiation step.

## RESULTS

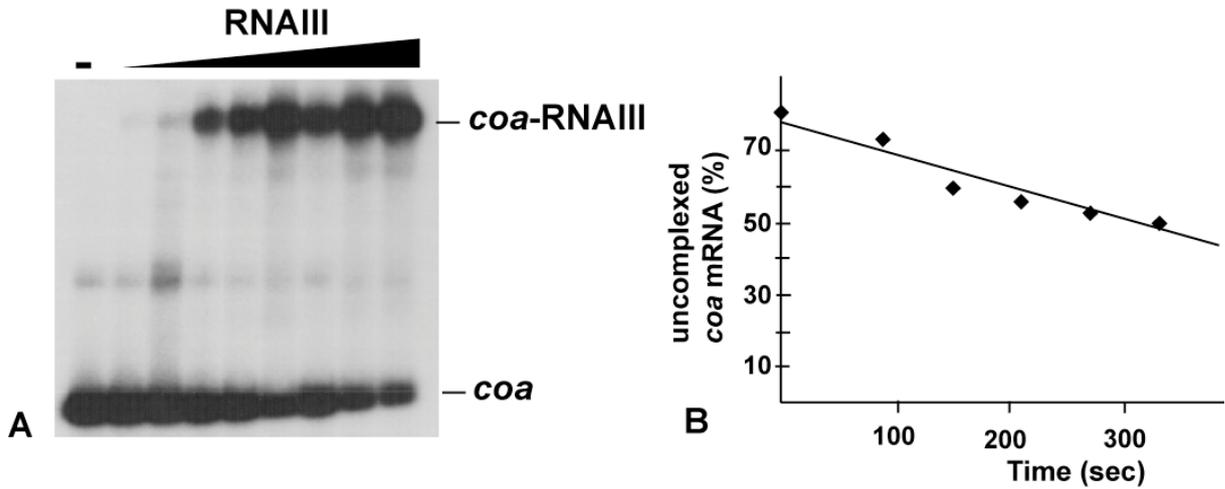
### **RNAIII regulates the synthesis of staphylocoagulase at the post-transcriptional level**

Sequence complementarity between RNAIII (nucleotides 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]. In order 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 truncated versions of RNAIII. The 5' start of *coa* mRNA was determined by 5'RACE showing that the 5' untranslated region contains 35 nucleotides upstream the AUG initiation codon. The whole leader regulatory region of *coa* gene including 88 nucleotides of the coding sequence was cloned in frame with *lacZ* gene into the pTCV-*lac* shuttle vector [14]. This construct is under control of an *agr*-independent promoter (*PrpoB*). The  $\beta$ -galactosidase activity was determined in the *S. aureus* strain LUG1467 (*rnaIII*<sup>+</sup>, *agr*<sup>+</sup>), which express RNAIII and in LUG1457 ( $\Delta$ *rnaIII*), which carries a deletion of the *rnaIII* gene (Fig. 1). We also measured the synthesis of the  $\beta$ -galactosidase from the *coa-lacZ* fusion in the strains deleted of *rnc* gene encoding RNase III (LUG1446), or of *hfq* (LUG1445). The  $\beta$ -galactosidase activity was reduced 6-fold in LUG1467 as compared to LUG1457. Furthermore, Hfq had no significant effect on the RNAIII-dependent repression while the deletion of *rnc* alleviates the repression of *coa-lacZ* fusion (Fig. 1).



**Figure 1:  $\beta$ -galactosidase activity detected from different gene fusions.**  $\beta$ -galactosidase activity measured from *PrpoB-coa (+1/+126)::lacZ* fusions in various *S. aureus* strains: RN6390 (*rnaIII*<sup>+</sup>, *agr*<sup>+</sup>, LUG1467), RN6390- $\Delta rnc$  (deletion of *rnc* gene encoding RNase III, LUG1446), RN6390- $\Delta hfq$  (deletion of *hfq* gene, LUG1445), WA400 ( $\Delta rnaIII$ , LUG1457), and WA400 transformed with the plasmid expressing the wild type RNIII (LUG1474), the 3' domain (LUG1447), RNIII- $\Delta$ 13 (RNIII deleted of hairpin 13, LUG1454), or RNIII- $\Delta$ 7-9 (RNIII deleted of hairpins 7 to 9, LUG1456), or no RNA (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.

Experiments were also carried out on LUG1457 strain ( $\Delta rnaIII$ ), complemented with different sets of plasmid pLUG274 expressing either the wild-type RNIII (LUG1474), the 3' end domain (LUG1447), RNIII- $\Delta$ 13 (deletion of hairpin 13, LUG1454) or RNA- $\Delta$ 7-8-9 (deletion of hairpins 7 to 9, LUG1456) (Fig. 1). Control experiment was carried out with the plasmid pE194, which did not contain insert. Unexpectedly, this plasmid slightly decreased the  $\beta$ -galactosidase levels as compared to LUG1457 strain. One explanation would be that the copy number of the pTCV-*lac* derivative was affected by the presence of the multicopy plasmid pE194 even though both plasmids are compatible. However, derivatives of plasmid pE194 producing high levels of wild-type RNIII, of the 3' domain, or of RNIII $\Delta$ 7-8-9 reproducibly decreased the synthesis of the  $\beta$ -galactosidase (Fig. 1). Conversely, the expression of RNIII- $\Delta$ 13 which lacks the base pairing complementarities with *coa* mRNA did not alter  $\beta$ -galactosidase synthesis (Fig. 1). Taken together, these results indicate that RNIII-mRNA pairings regulate *coa* expression at the post-transcriptional level and that the hairpin 13 of RNIII is essential for the repression.

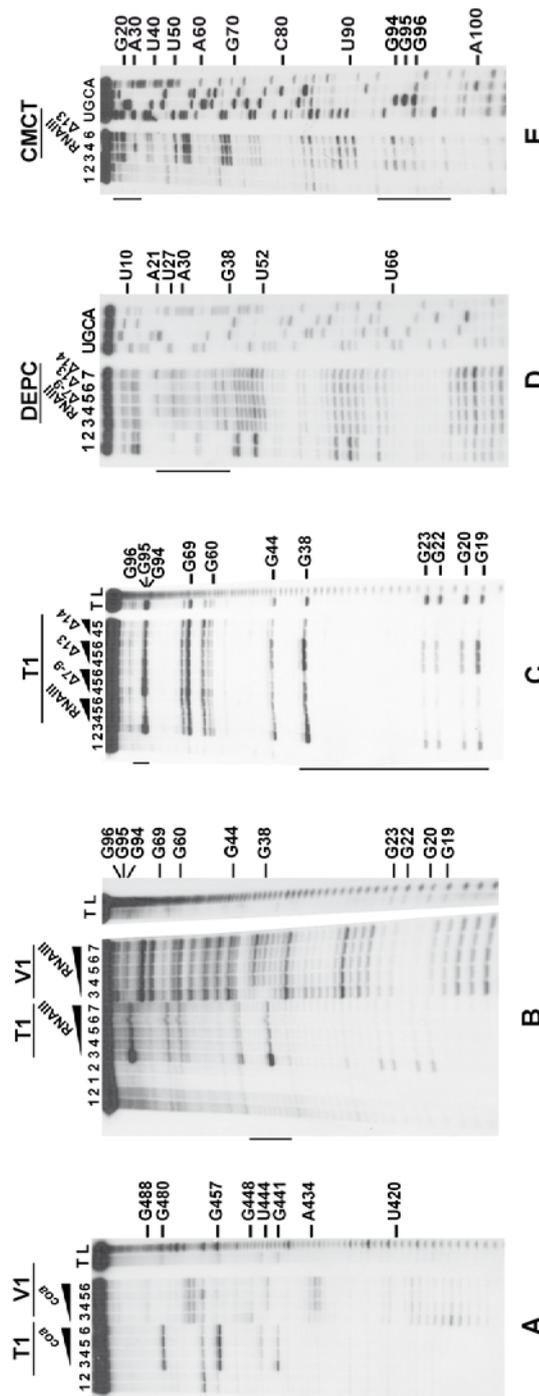


**Figure 2: RNAIII binds efficiently to *coa* mRNA *in vitro*.** (A) Determination of the apparent dissociation constant for RNAIII-*coa* mRNA complex. 5'-end-labeled *coa* mRNA was incubated alone (-) or with various concentrations of unlabeled wild type RNAIII (1, 5, 10, 20, 50, 100, 200 and 250 nM). The fraction of labeled *coa* mRNA associated with RNAIII 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 the RNAIII-*coa* mRNA complex as determined from two independent experiments. 5'-end-labeled *coa* mRNA (0.1 nM) was incubated with unlabeled RNAIII (30 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 {Persson et al., 1988, EMBO J, 7, 3279-88}.

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

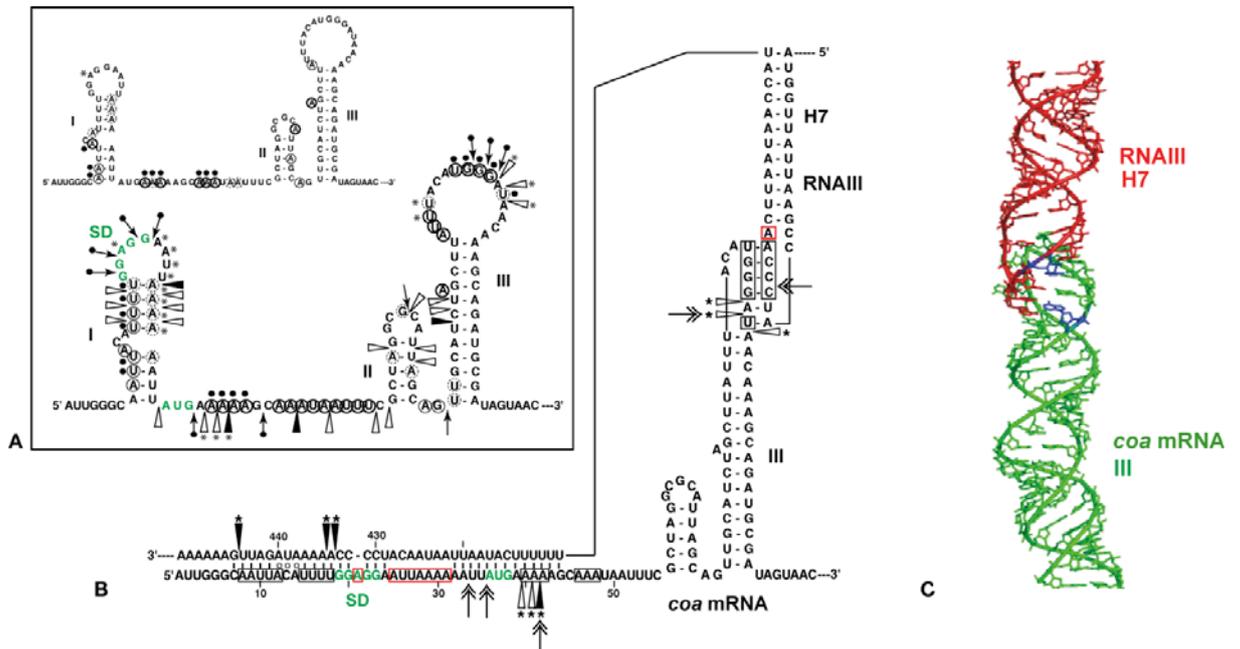
The *in vivo* experiments indicated that the RNAIII-dependent repression of *coa* mRNA is governed by direct RNAIII-mRNA pairing. Therefore, the formation of a complex between RNAIII and a fragment of *coa* mRNA (200 nt) including the whole 5' UTR and part of the coding region was investigated by gel shift assays. *In vitro* labeled *coa* mRNA was incubated with increasing concentrations of RNAIII at 37°C for 15 min (Fig. 2). This experiment shows that *coa* mRNA binds to RNAIII with a  $K_d$  value of 50 nM. Furthermore, the initial rate of RNAIII binding to *coa* mRNA was estimated from a time-course analysis to be around  $2-4 \cdot 10^6 \text{ M}^{-1}\text{s}^{-1}$ . Similar characteristics were observed for three other RNAIII-mRNA (*spa*, SA1000, *rot*) target complexes [11,12]. These data indicate that the complexes are efficiently formed as observed for several fully complementary antisense-target RNA systems [15,16].

The conformation of *coa* mRNA was then probed using RNase T1 (specific for unpaired guanines), RNase V1 (specific for helical regions), and several base-specific chemicals such as dimethylsulfate (methylates N1A>>N3C), carbodiimide derivative (modifies N3U>>N1G), and diethylpyrocarbonate (carboxyethylates N7A). Several experiments are shown in Fig. 3.



**Figure 3:** Enzymatic and chemical probing of the structure of the inhibitory RNAIII-coa mRNA complex. (A) Enzymatic hydrolysis of 5'-end-labeled 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 3, 4: 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 in the presence of RNAIII, or of the mutant RNAIII deleted of hairpins 7 to 9 (RNAIIIΔ7-9), of hairpin 13 (RNAIIIΔ13) or of hairpin 14 (RNAIIIΔ14) at 200 nM. Same legend as in A. (D) DEPC (N7A) modification of unlabeled coa mRNA, free (lane 3) or in complex with the wild type RNAIII (lane 4), or the mutant RNAIII deleted of hairpins 7 to 9 (RNAIIIΔ7-9, lane 5), of hairpin 13 (RNAIIIΔ13, lane 6) or of hairpin 14 (RNAIIIΔ14, lane 7) at 200 nM. Lanes 8, 9, C, A: dideoxy-sequencing reactions performed on coa mRNA. (E) CMCT and DMS modification of unlabeled coa mRNA. Same legend as in D.

The secondary structure model of *coa* mRNA, that explains most of the probing data, involves three stem-loop structures connected by unpaired residues (Fig. 4A). The first AU-rich hairpin is of weak stability but, based on the enzymatic cleavage pattern, is proposed to occur. However, the coexistence of alternative structures in the region encompassing nucleotides 10 to 70 may explain the concomitant presence of RNase V1 cleavages and of the reactivity of many nucleotides at one of their Watson-Crick position. 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. 3B, C). Binding of RNIII induced changes in the region encompassing the ribosome binding site (RBS, nucleotides A9 to A58). RNIII protected the guanines of the SD sequence against RNase T1 as well as the nucleotides A8 to U18 and A41 to A48 against chemical modifications (Fig. 3D, E). Concomitantly, RNIII binding induced new RNase V1 cleavages at positions 39-41, and enhanced reactivity of A21, A25, A29 to A31 at position N1, of A21 at position N7, and of U26 and U27 at position N3 in *coa* mRNA (Fig. 3A, Fig. 4). These reactivity changes in the RBS of *coa* mRNA most likely resulted from the binding of the hairpin 13 of RNIII since its deletion in RNIII had no effect on the structure of the RBS of *coa* mRNA (Fig. 3C). Binding of *coa* mRNA to the 3' domain or to RNIII induced correlated changes in hairpin 13. Strong protections were observed at G441 against RNase T1, at positions 411-415 and 448-449 against RNase V1 (Fig. 3A). Concomitantly, increased RNase V1 cleavages were observed at positions 433-434 and 444. All these data are thus consistent with the formation of RNIII-mRNA duplex which sequestered the RBS of *coa* mRNA (Fig. 4). This imperfect duplex involves two consecutive regions of 13 base pairings interrupted by an internal loop, and a bulged adenine 21 (Fig. 4B).



**Figure 4: Structure of the RNAIII-*coa* mRNA complex.** (A) Summary of the enzymatic cleavages and chemical reactivities of nucleotides of *coa* mRNA. Enzymatic cleavages are given as follows: RNase T1 (  $\longrightarrow$  ), and RNase V1 (  $\triangleright$  ) moderate, (  $\blacktriangleright$  ) strong cleavage. Chemical modifications of cytosines at N3, and adenines at N1 towards DMS, of uridines at N3 and guanines at N1 towards CMCT, and of adenines at N7 towards DEPC: full and dashed circled nucleotides are for strong and moderate reactivity, respectively. No symbol is for not reactive, nd is for not determined due to unspecific 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 RNAIII 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 RNAIII-*coa* mRNA complex showing the reactivity changes induced by complex formation: RNase V1 (  $\triangleright\blacktriangleright$  ), RNase III cleavages (  $\longrightarrow$  ). Nucleotides which become protected upon RNAIII binding 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 *coa* mRNA is in green and the hairpin 7 of RNAIII in red. The RNase III cleavages are shown in blue.

Unexpectedly, we also found a second RNAIII binding site restricted to the apical loop III of *coa* mRNA (Fig. 4B). Binding of RNAIII reduced considerably the RNase T1 cuts at G94-97 and the modifications of the nucleotides UGGGAU98 towards chemicals (Fig. 3B). Concomitantly, RNAIII binding induced several RNase V1 cuts at positions 96 to 98 (Fig. 3B). These changes were abolished if the complex was formed between *coa* mRNA and the RNAIII deleted of hairpins 7 to 9 (Fig. 3C). Furthermore, *coa* mRNA binding to RNAIII reduced significantly the reactivity of the nucleotides CCA243 towards DMS in

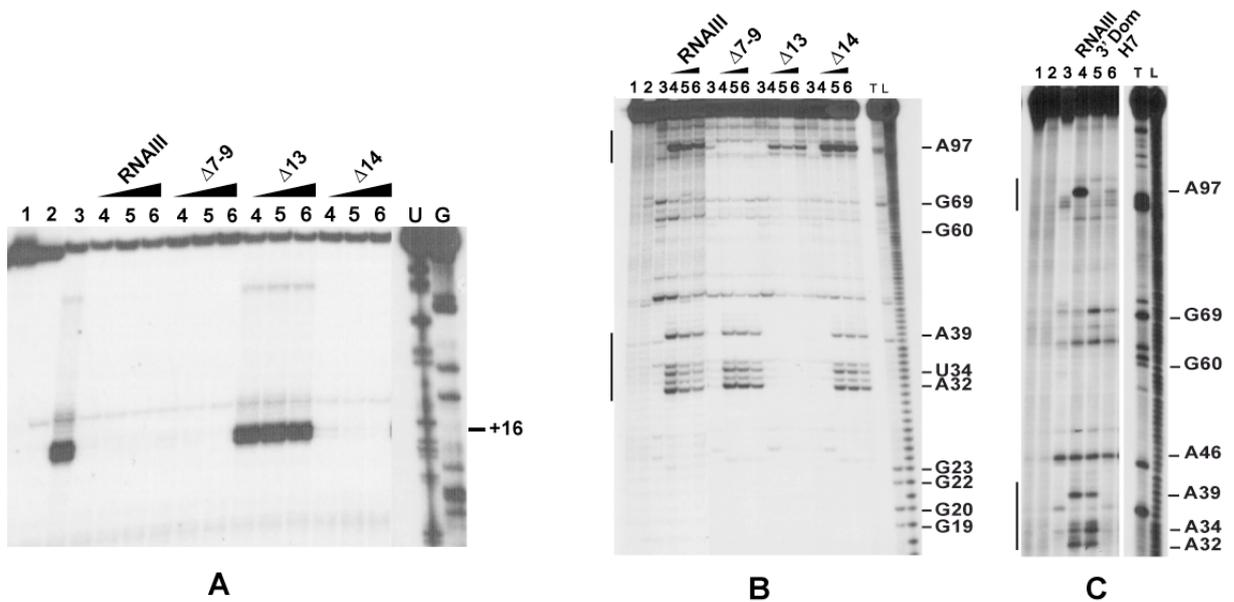
the apical loop 7 of RNAIII. These data are strengthened by the sequence complementarities between the apical loop III of *coa* mRNA and the hairpin loop 7 of RNAIII, and support the existence of a loop-loop interaction (Fig. 4A, B). This interaction is however strongly dependent on the formation of the imperfect duplex since the reactivity changes in the hairpin loop III of *coa* mRNA were significantly decreased if complex formation was performed with RNAIII- $\Delta$ 13. 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. 4C).

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

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

Since RNAIII binds to the Shine-Dalgarno sequence of *coa* mRNA, we analyzed whether RNAIII binding is sufficient to prevent the formation of the ternary initiation complex formed with *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 16 nucleotides downstream of the initiation codon (Hartz *et al*, 1988; 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- $\Delta$ 7-9 (deleted of hairpins 7 to 9), or RNAIII- $\Delta$ H14 (deleted of hairpin 14) strongly decreased the two toeprint signals. This indicates that the regulatory RNAIII totally blocks the 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 the ribosome binding since the primer used for elongation hybridized in the long hairpin loop III. However, the RNAIII- $\Delta$ H13 exerted no inhibitory effect on ribosome binding, showing that the specific RNAIII-inhibition of ribosome binding to *coa* mRNA mainly resulted from the sequestration of the RBS by the hairpin 13 of RNAIII (Fig. 5A).



**Figure 5: The RNAIII-*coa* mRNA complex prevents ribosome binding and promotes RNase III cleavages.** (A) Formation of the ternary complex between *coa* mRNA (15 nM), *S. aureus* 30S ribosomal subunits (250 nM) and initiator tRNA (1  $\mu$ M) was monitored in the absence (lane 3) 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 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 3) 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 3) 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.

**RNase III cleaves the two distant mRNA-RNAIII interactions *in vitro***

As RNase III is required for efficient repression *in vivo* (Fig. 1), we analyzed whether this enzyme can cleave the complex *in vitro*. We have previously shown that cleavage assays by RNase III can be considered as a useful tool for probing *in vitro* RNA-RNA complexes [17]. The RNase III-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 RNase III from *S. aureus* (Fig. 5B-C). Only two weak RNase III cleavages were observed in the long hairpin III of *coa* mRNA. When the 5' end-labeled mRNA was incubated with RNAIII, three major cleavages occurred at positions 32, 39, and 97 in the mRNA, and a weaker cleavage was detected at position 34 (Fig. 5B-C). Binding of *coa* mRNA induces a RNase III-dependent cleavage at C241 of the labeled RNAIII (results not shown). Thus, the two regions of hybridization were susceptible to RNase III cleavages.

Using truncated versions of RNAIII and the isolated hairpins 7 or 13, we were able to assign the partners involved in the RNAIII-mRNA complex. Indeed, the hairpin 7 only induced a specific RNase III-cleavage at position 97 of the mRNA while the hairpin 13 binding promoted a major cleavage at position 39 of the mRNA (Fig. 5C). Furthermore, the complex formed between RNAIII- $\Delta$ 7-9 and *coa* mRNA was efficiently cleaved by RNase III at positions 32, 34, and 39 of the mRNA (Fig. 5B). Conversely, only one RNase III mediated cleavage was detected at position 97 of *coa* mRNA bound to RNAIII- $\Delta$ 13 (Fig. 5B-C). In the irregular duplex, RNase III 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 [17].

## DISCUSSION

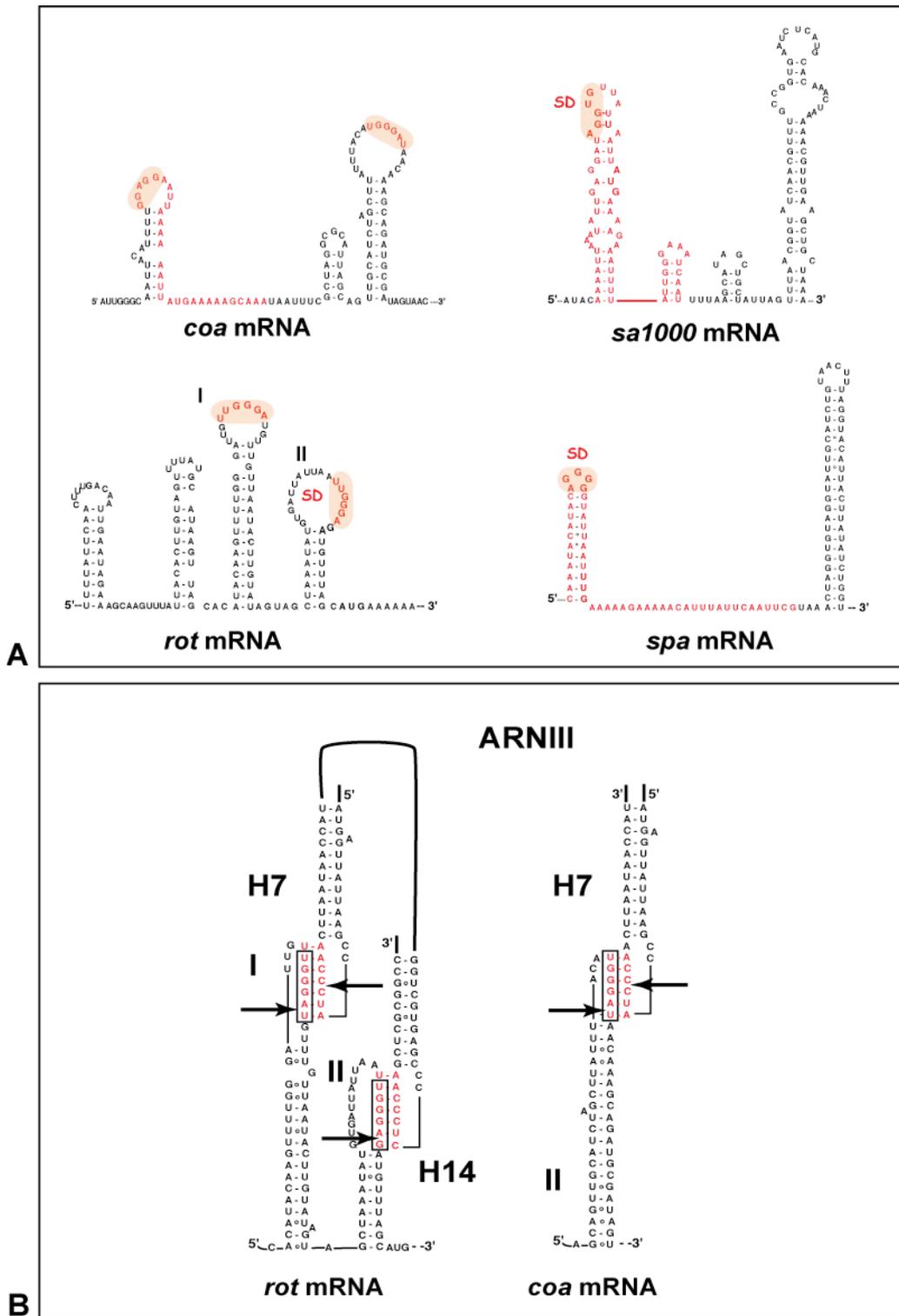
*S. aureus* produces a large variety of virulence factors which are required for the successful colonization of the host and which confer to the bacteria the ability to counteract the immune defense system of the host [2]. Among these virulence factors, staphylocoagulase has the primary function to activate the prothrombin, inducing the formation of a fibrin clot around the bacterial cell [18]. Coating the bacteria with host proteins contribute to hide the bacteria from the immune system and from phagocytosis. The expression of coagulase was shown to follow a temporal regulation as for several adhesins and surface proteins, which 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, the SsaA-like protein SA2353 that were found to be repressed by the quorum sensing regulatory RNIII. During the growth cycle, the level of RNIII varied 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 as a functional *agr* element resulted in a relative elevation of the *coa* mRNA level at the early time point followed by a diminution of the transcript at the later time point [13]. We demonstrated here that the *agr*-dependent repression effect of *coa* mRNA is the result of a direct binding of RNIII to *coa* mRNA.

We show that *in vivo*, RNIII in conjunction with the endoribonuclease III (RNase III) are required to fully repress the expression of a reporter gene fused in frame with the whole regulatory region of *coa* mRNA (Fig. 3). *In vitro*, the formation of the *coa* mRNA-

RNAIII duplex prevents the binding of the small ribosomal 30S subunit at its loading site, and RNase III specifically cleaves the mRNA bound to RNAIII. Based on this data, we propose that the primary effect of RNAIII would be to prevent translation initiation. This step would be followed by the RNase III-dependent cleavage of the repressed mRNA. Hence, the data together with previous works show that RNAIII represses the synthesis of coagulase, protein A, SA1000, SA2353, and Rot by a similar post-transcriptional mechanism [10-12]. In addition, probing the mRNA structure also shows that *coa* mRNA adopts a very similar structural organization to *spa* and SA1000 mRNAs. The three mRNAs have rather short 5' untranslated regions, which carry a 5' hairpin structure with a strong SD sequence located in the apical loop (Fig. S1). In the absence of RNAIII, these elements may confer to the mRNAs a high stability. Indeed, in *B. subtilis*, stabilization of mRNAs was shown to be a consequence of the blocking of the 5' end by a stalled initiating ribosome at a SD-like sequence [19,20] or by a stable 5' hairpin structure and a strong RBS [21]. Blocking the 5' end of mRNAs probably prevents the action of the 5'-3' exoribonuclease activity of RNase J1 which is highly conserved in *B. subtilis* and *S. aureus* [19]. Therefore the coordinated action of RNAIII and RNase III would be to irreversibly repress the synthesis of these virulence factors at an appropriate time.

*In vitro* binding assays show that RNAIII binds to *coa* mRNA with a rather fast association rate constant in the same range of magnitude as for the other target mRNAs. Efficient repression by ncRNAs, 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]. This could be achieved by various strategies exemplified by functional studies performed on various ncRNAs from *E. coli* and *S. typhimurium* [22,23]. A large class of small ncRNA requires the Sm-like protein Hfq which facilitates ncRNA-mRNA base pairings by various mechanisms and which protects

ncRNAs against RNase E-dependent degradation [24,25]. In contrast, many antisense RNAs fully complementary to their target RNAs do not require a helper protein to facilitate base pairings. Instead, the structures of RNAs have evolved so that fast pairing is initiated through a limited number of intermolecular base pairs often GC-rich to form a loop-loop or a loop-unpaired region, which subsequently propagates to form irreversible RNA-mRNA duplexes [15,16]. Our data indicate that *S. aureus* Hfq is not required for the RNIII-dependent repression of *coa* mRNA *in vivo* (Fig. 1) and does not facilitate the formation of the RNIII-mRNA complex *in vitro* (results not shown). The fact that the deletion of *hfq* does not exhibit severe phenotypic defects doubts the direct involvement of Hfq in regard to RNIII-mediated regulation in *S. aureus* [10,26] although we have previously shown that Hfq binds to RNIII [11]. Since RNIII binds rapidly to its target mRNAs, it might also be possible that the RNA structures may compensate for the need of a helper protein in a way reminiscent to the fully complementary antisense RNAs. We also do not rule out that the RNase III could contribute to stabilize and/or facilitate the formation of the hybrid since we have previously shown that the enzyme binds to RNIII [11].



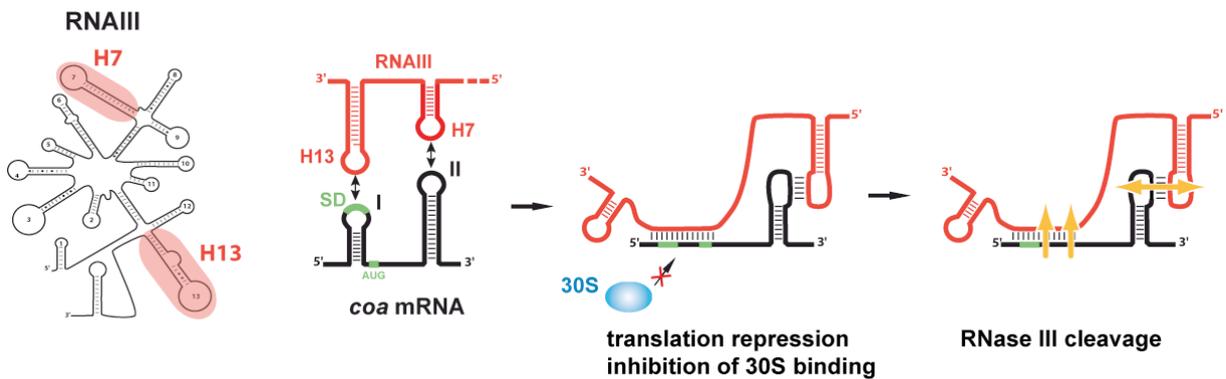
**Figure S1: Comparison of the secondary structures of the RNAIII targets.** (A) The secondary structures of *spa* mRNA [11], SA1000 and *rot* mRNA [12], and *coa* mRNA. The nucleotides base paired to RNAIII are in red. The conserved G-rich sequence in the mRNAs which interact with the C-rich motif of hairpin loops 7, 13, 14 of RNAIII are encircled. (B) Sequences of the loop-loop interactions as found in *rot* mRNA-RNAIII and *coa* mRNA-RNAIII complexes. The RNase III cleavages are shown by arrows.

The regions of interaction in RNAIII and *coa* mRNA contained stem-loop structures that are well appropriate for rapid initial loop-loop interactions. Probing the structure of the RNAIII-*coa* mRNA structure revealed that the two conserved C-rich loops 7 and 13 of RNAIII bind to the RBS and to the hairpin loop III in the coding sequence of *coa* mRNA, respectively (Fig. 4, 6). These C-rich hairpin loops of RNAIII are also used to repress other mRNAs encoding virulence factors and a transcriptional regulatory protein although the topologies of the resulting inhibitory complexes are different. RNAIII forms long duplexes with the RBS of *spa* and SA1000 mRNAs, while it forms two loop-loop interactions with respectively the 5'UTR and the RBS of *rot* mRNA [11,12]. Here we show that the RNAIII-*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. 6). Thus the various topologies of the repressed complexes depend largely on the mRNA context. The inhibitory RNAIII-*coa* mRNA complex also provided specific binding sites for the double-stranded specific RNase III which induced strong cleavages in the two regions of *coa* mRNA bound to RNAIII. Notably, the cleavage sites in the loop-loop interaction also occurred at a similar position in the two kissing interactions that took place in *rot* mRNA-RNAIII complex. The sequences of *coa* mRNA involved in the loop-loop interactions are very similar to *rot* mRNA, showing that similar signatures exist in various RNAIII-repressed mRNAs (Fig. S1; [12]). 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. 4C). The overall topology is very

similar to the RNA loop-loop structure obtained by NMR which mimics the interaction between a sense and antisense RNAs involved in the regulation of ColE1 plasmid [27]. Such a long helical structure might be well 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 RNIII is produced [13]. Therefore, as we previously postulated, RNase III might initiate rapid degradation of *coa* mRNA, and the cleavage in the loop-loop interaction within the coding sequence would provide the access to several other endo- or exoribonucleases such as RNase J1 for further degradation (Fig. 6). Notably, at a similar position *spa* and SA1000 mRNAs carry a long stem-loop structure in the coding sequence that are also well cleaved by RNase III [11,12]. The rapid depletion of the mRNA might also result from an indirect effect of RNIII. Indeed, Rot protein was shown to activate the transcription of *coa* mRNA [28] while RNIII represses the synthesis of the Rot at the post-transcriptional level [10]. Thus, the RNIII-mediated repression of coagulase would occur at both transcriptional and post-transcriptional levels as it was shown for *spa* mRNA [10,12,28].

It is not an exception that RNIII 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 [22,29]. *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 RNIII, we recently found that other *S. aureus* ncRNAs carry a similar UCCC signature always present in an unpaired region, and through its unpaired C-rich motif, one of these RNA binds to the RBS and represses the expression of several mRNAs [32]. *S. aureus coa* mRNA and the other mRNA targets of RNIII carry a

strong SD sequence located in a unpaired region well appropriate for the docking of the 30S subunit, but also for the formation of initial contacts with the C-rich loop of RNAIII. Specificity for *coa* regulation is however 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.



**Figure 6: Schematic view of RNAIII mediated repression of *coa* mRNA.** RNAIII binds to its target mRNA masking the RBS and part of the coding sequence. Binding of RNAIII hinders ribosome binding and promotes the access to RNase III. SD (Shine and Dalgarno) and AUG are in green. RNAIII is in red, the mRNA target in black, and the 30S subunit in blue.

In conclusion, this study validates another direct target of RNAIII playing a role in virulence and completes the RNAIII-dependent networks in *S. aureus*. 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 in *E. coli* and *S. typhimurium* show that regulatory RNAs which target mRNAs regulate gene expression through a variety of unusual mechanisms and bind to the mRNA regions located far away the ribosome binding site in the 5'UTR [33], in the coding sequence [34,35], and in the 3' end [36]. Whether *S. aureus* have also evolved such a diversity of sRNA-dependent regulatory mechanisms remains to be addressed.

## MATERIAL AND METHODS

### *Strains and plasmids*

*S. aureus* RN6390 derives from 8325-4. In WA400 ( $\Delta$ *rnaIII*), the P3 operon is deleted and replaced by the chloramphenicol transacetylase gene (*cat86*) [37]. LUG774 and LUG911 strains derive from RN6390 in which *rnc* and *hfq* genes, respectively, have been replaced by *cat86* gene [11]. Staphylococci were grown either on BM agar plates or in brain-heart infusion (BHI) with erythromycin (5  $\mu$ g/ml) when appropriate. RNAIII and its variants were expressed in *Staphylococcus aureus* WA400 by using plasmid pE194 (see Table 1).

### *Construction of translation fusions and b-galactosidase measurements*

Translation fusions were constructed by using plasmid pLUG220, a derivative of pTCV-*lac*, a low-copy-number promoter-less *lacZ* vector (Table 1). The 5' end of the *coa* mRNA was first determined by rapid amplification of cDNA ends (RACE) using the First Choice RLM-RACE kit following the company protocol (Ambion). The whole leader region of *coa* mRNA including 200 nts of the coding sequence was cloned downstream the *rpoB* promoter in frame with *lacZ* [11]. b-galactosidase activity was measured three times on duplicate cultures with the Enzyme Assay System (Promega).

**Table 1. Strains and plasmids**

<b><i>S. aureus</i> strains</b>	<b>Relevant characteristics</b>	<b>Reference or source</b>
RN4220	restriction- mutant of 8325-4	[44]
RN6390	derivative of 8325-4, <i>agr</i> positive	[45]
WA400	8325-4 : $\Delta rnaIII$ region:: <i>cat86</i>	[37]
LUG 404	W400/pLUG 274	[12]
LUG 450	WA400/pLUG 300	[11]
LUG 453	WA400/pLUG 304	[11]
LUG 580	WA400/pLUG 394	this study
LUG 581	W400/pLUG 298	[11]
LUG774	RN6390 : $\Delta rnc$ region:: <i>cat86</i>	[11]
LUG911	RN6390 : $\Delta hfg$ region:: <i>cat86</i>	[12]
LUG 1445	LUG 911/pLUG 745	this study
LUG 1446	LUG 774/pLUG 745	this study
LUG 1447	LUG 450/pLUG 745	this study
LUG 1454	LUG 453/pLUG 745	this study
LUG 1456	LUG 580/pLUG 745	this study
LUG 1457	WA400/pLUG 745	this study
LUG 1467	RN6390/pLUG 745	this study
LUG 1474	LUG 581/pLUG 745	this study
LUG 1478	LUG 404/pLUG 745	this study
<b><i>E. coli</i>-</b>		
<b>staphylococcal shuttle plasmids</b>		
pTCV- <i>lac</i>	Promoter- <i>lac</i> fusion shuttle vector:	[14]
	<i>spoVG-lacZ</i> , <i>ermB</i> , <i>aphA-3</i>	
pLUG 220	pTCV- <i>lac</i> D RBS and start codon	[11]
pLUG 745	pLUG220:: <i>PrpoB</i> (nts -480 to +1):: <i>coa</i> (+1 [transcriptional start] to + 126)	this study
<b>Staphylococcal plasmids</b>		
pE194	3.728 kb <i>S. aureus</i> plasmid, inducible MLS resistance ( <i>erm</i> )	[46]
pLUG 274	pE194:: <i>EcoRV</i> site in MCS	[38]
pLUG 298	pLUG274:: <i>P3</i> operon (nts 1819-751)	[11]
pLUG 300	pLUG274:: <i>P3</i> promoter link to 3'end RNAIII (nts 391-514)	[38]
pLUG 304	pLUG274:: <i>rnaIII</i> D nts 408-451 (RNAIII D13)	[11]
pLUG 394	pLUG274:: <i>rnaIII</i> D nts 208-323 (RNAIII D 7-9)	this study

***RNA preparation***

RNAIII, RNAIII derivatives (RNAIII- $\Delta$ 7-9 deletion of nts G207 to U319, RNAIII- $\Delta$ 13 deletion of nts U409 to A451, and RNAIII- $\Delta$ 14 deletion of nts G483 to C511), the isolated hairpins 7, 13 and 14 and *coa* mRNA fragment were transcribed *in vitro* using T7 RNA polymerase as previously described [38]. The 5' end-labeling of dephosphorylated RNA or DNA oligonucleotides was performed with T4 polynucleotide kinase and [ $\gamma$ - $^{32}$ 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-*coa* mRNA complex formation***

Binding rate constant of RNAIII-*coa* mRNA complex was measured as previously described [40]. Binding of end-labelled *coa* mRNA to a ten fold excess of unlabelled RNA 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-*coa* mRNA complex and free RNAIII, respectively, were quantified using SAFA algorithm [41].

For determination of the apparent dissociation rate constant of RNAIII-*coa* mRNA complex, end-labelled *coa* mRNA was incubated with increased molar amount of wild-type RNAIII for 15 min at 37°C in TMN buffer. Samples were then treated as described above.

**RNA structure probing**

RNAIII-*coa* 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 *coa* 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. Modification 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, 150 mM KOAc.

RNase III purification and the enzymatic cleavage assays on *coa* mRNA and on RNAIII were performed as previously described [17].

End-labeled RNA fragments were sized on 12% polyacrylamide/8 M urea slab gels. Cleavage positions were identified by running 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 previously described [38].

**Toeprinting assays**

*S. aureus* 30S subunits were prepared according to [42]. The formation of a simplified translational initiation complex with mRNA and the extension inhibition conditions were strictly identical to those described by [38].

***Molecular modeling***

Modelling of the regions encompassing residues U73 to A114 of the *coa* mRNA and of residues A223 to U256 of RNAlII were carried out as described [43]. Following the interactive assembly step, several cycles of geometrical least-square refinements were performed until a satisfactory solution was reached. Figure 3D was prepared using the PYMOL program (DeLano WL., The PyMOL Molecular Graphics System (2002) <http://www.pymol.org>).

**ACKNOWLEDGEMENTS**

We are thankful to Efthimia Lioliou, Stefano Marzi, and Gérard Lina for helpful comments and critical reading of the manuscript. This work was supported by the Centre National de la Recherche Scientifique (CNRS; PR), the Institut National pour la Recherche Médicale (INSERM; FV), the Agence Nationale pour la Recherche (ANR05-MIME, ANR09-BLAN-; PR & FV), and the European Community (BacRNA, FP6-018618; PR; EC222718; TG).

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

**REFERENCES**

1. Novick RP, Ross HF, Projan SJ, Kornblum J, Kreiswirth B, et al. (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J* 12: 3967-3975.
2. Novick RP, Geisinger E (2008) Quorum sensing in staphylococci. *Annu Rev Genet* 42: 541-564.

3. Queck SY, Jameson-Lee M, Villaruz AE, Bach TH, Khan BA, et al. (2008) RNAIII-independent target gene control by the agr quorum-sensing system: insight into the evolution of virulence regulation in *Staphylococcus aureus*. *Mol Cell* 32: 150-158.
4. Novick RP (2003) Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Mol Microbiol* 48: 1429-1449.
5. Dunman PM, Murphy E, Haney S, Palacios D, Tucker-Kellogg G, et al. (2001) Transcription profiling-based identification of *Staphylococcus aureus* genes regulated by the agr and/or sarA loci. *J Bacteriol* 183: 7341-7353.
6. Somerville GA, Proctor RA (2009) At the crossroads of bacterial metabolism and virulence synthesis in Staphylococci. *Microbiol Mol Biol Rev* 73: 233-248.
7. Toledo-Arana A, Repoila F, Cossart P (2007) Small noncoding RNAs controlling pathogenesis. *Curr Opin Microbiol* 10: 182-188.
8. Waters LS, Storz G (2009) Regulatory RNAs in bacteria. *Cell* 136: 615-628.
9. Morfeldt E, Taylor D, von Gabain A, Arvidson S (1995) Activation of alpha-toxin translation in *Staphylococcus aureus* by the trans-encoded antisense RNA, RNAIII. *EMBO J* 14: 4569-4577.
10. Geisinger E, Adhikari RP, Jin R, Ross HF, Novick RP (2006) Inhibition of *rot* translation by RNAIII, a key feature of *agr* function. *Mol Microbiol* 61: 1038-1048.
11. Huntzinger E, Boisset S, Saveanu C, Benito Y, Geissmann T, et al. (2005) *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate *spa* gene expression. *EMBO J* 24: 824-835.
12. Boisset S, Geissmann T, Huntzinger E, Fechter P, Bendridi N, et al. (2007) *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev* 21: 1353-1366.
13. Lebeau C, Vandenesch F, Greenland T, Novick RP, Etienne J (1994) Coagulase expression in *Staphylococcus aureus* is positively and negatively modulated by an agr-dependent mechanism. *J Bacteriol* 176: 5534-5536.
14. Poyart C, Trieu-Cuot P (1997) A broad-host-range mobilizable shuttle vector for the construction of transcriptional fusions to beta-galactosidase in gram-positive bacteria. *FEMS Microbiol Lett* 156: 193-198.
15. Brantl S (2007) Regulatory mechanisms employed by cis-encoded antisense RNAs. *Curr Opin Microbiol* 10: 102-109.

16. Wagner EG, Altuvia S, Romby P (2002) Antisense RNAs in bacteria and their genetic elements. *Adv Genet* 46: 361-398.
17. Chevalier C, Huntzinger E, Fechter P, Boisset S, Vandenesch F, et al. (2008) *Staphylococcus aureus* endoribonuclease III purification and properties. *Methods Enzymol* 447: 309-327.
18. Kawabata S, Morita T, Iwanaga S, Igarashi H (1985) Enzymatic properties of staphylothrombin, an active molecular complex formed between staphylocoagulase and human prothrombin. *J Biochem* 98: 1603-1614.
19. Mathy N, Benard L, Pellegrini O, Daou R, Wen T, et al. (2007) 5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell* 129: 681-692.
20. Daou-Chabo R, Mathy N, Benard L, Condon C (2009) Ribosomes initiating translation of the hbs mRNA protect it from 5'-to-3' exoribonucleolytic degradation by RNase J1. *Mol Microbiol* 71: 1538-1550.
21. Hambræus G, Karhumaa K, Rutberg B (2002) A 5' stem-loop and ribosome binding but not translation are important for the stability of *Bacillus subtilis* aprE leader mRNA. *Microbiology* 148: 1795-1803.
22. Papenfort K, Vogel J (2009) Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level. *Res Microbiol* 160: 278-287.
23. Repoila F, Darfeuille F (2009) Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects. *Biol Cell* 101: 117-131.
24. Aiba H (2007) Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr Opin Microbiol* 10: 134-139.
25. Brennan RG, Link TM (2007) Hfq structure, function and ligand binding. *Curr Opin Microbiol* 10: 125-133.
26. Bohn C, Rigoulay C, Bouloc P (2007) No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. *BMC Microbiol* 7: 10.
27. Lee AJ, Crothers DM (1998) The solution structure of an RNA loop-loop complex: the Cole1 inverted loop sequence. *Structure* 6: 993-1005.
28. Said-Salim B, Dunman PM, McAleese FM, Macapagal D, Murphy E, et al. (2003) Global regulation of *Staphylococcus aureus* genes by Rot. *J Bacteriol* 185: 610-619.

29. Sharma CM, Darfeuille F, Plantinga 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.
30. Papenfort K, Pfeiffer V, Lucchini S, Sonawane A, Hinton JC, et al. (2008) Systematic deletion of *Salmonella* small RNA genes identifies CyaR, a conserved CRP-dependent riboregulator of OmpX synthesis. *Mol Microbiol* 68: 890-906.
31. De Lay N, Gottesman S (2009) The Crp-activated small noncoding regulatory RNA CyaR (RyeE) links nutritional status to group behavior. *J Bacteriol* 191: 461-476.
32. Geissmann T, Chevalier C, Cros MJ, Boisset S, Fechter P, et al. (2009) A search for small non coding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation. *Nucleic Acids Res* in press:
33. Darfeuille F, Unoson C, Vogel J, Wagner EG (2007) An antisense RNA inhibits translation by competing with standby ribosomes. *Mol Cell* 26: 381-392.
34. Bouvier M, Sharma CM, Mika F, Nierhaus KH, Vogel J (2008) Small RNA binding to 5' mRNA coding region inhibits translational initiation. *Mol Cell* 32: 827-837.
35. Pfeiffer V, Papenfort 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, Storz G (2004) GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J Bacteriol* 186: 6698-6705.
37. Janzon L, Arvidson S (1990) The role of the  $\delta$ -lysin gene (*hld*) in the regulation of virulence genes by the accessory gene regulator (*agr*) in *Staphylococcus aureus*. *EMBO J* 9: 1391-1399.
38. Benito Y, Kolb FA, Romby P, Lina G, Etienne J, et al. (2000) Probing the structure of RNAIII, the *Staphylococcus aureus agr* regulatory RNA, and identification of the RNA domain involved in repression of protein A expression. *RNA* 6: 668-679.
39. Sambrook J, Fritsch E, Maniatis T (1989) *Molecular Cloning: a laboratory manual*. Cold Spring Harbor, New York, USA: Cold Spring Harbor Press.
40. Persson C, Wagner EG, Nordstrom K (1988) Control of replication of plasmid R1: kinetics of in vitro interaction between the antisense RNA, CopA, and its target, CopT. *EMBO J* 7: 3279-3288.

41. Das R, Laederach A, Pearlman SM, Herschlag D, Altman RB (2005) SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. *RNA* 11: 344-354.
42. Fechter P, Chevalier C, Yusupova G, Yusupov M, Romby P, et al. (2009) Ribosomal initiation complexes probed by toeprinting and effect of trans-acting translational regulators in bacteria. *Methods Mol Biol* 540: 247-263.
43. Massire C, Westhof E (1998) MANIP: an interactive tool for modelling RNA. *J Mol Graph Model* 16: 197-205, 255-7.
44. Kreiswirth BN, Lofdahl S, Betley MJ, O'Reilly M, Schlievert PM, et al. (1983) The toxic shock syndrome exotoxin structural gene is not detectably transmitted by a prophage. *Nature* 305: 709-712.
45. Peng HL, Novick RP, Kreiswirth B, Kornblum J, Schlievert P (1988) Cloning, characterization, and sequencing of an accessory gene regulator (*agr*) in *Staphylococcus aureus*. *J Bacteriol* 170: 4365-4372.
46. Horinouchi S, Weisblum B (1982) Nucleotide sequence and functional map of pE194, a plasmid that specifies inducible resistance to macrolide, lincosamide, and streptogramin type B antibiotics. *J Bacteriol* 150: 804-814.

## II. Les propriétés de la RNase III de *Staphylococcus aureus*

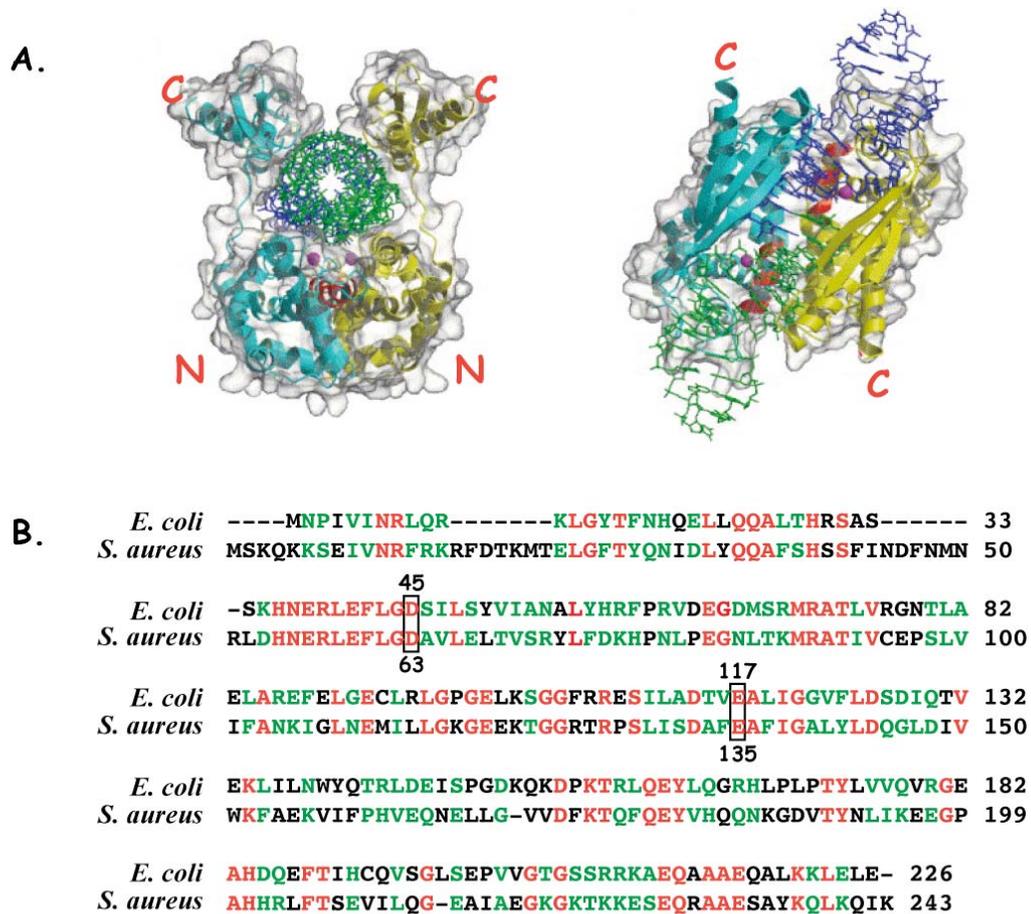
Au cours de ce travail, la RNase III s'est avérée être une sonde déterminante pour analyser *in vitro* les différents complexes impliquant l'ARNIII et ses ARNm cibles. Pour la première fois, nous avons révélé que les interactions boucles-boucles sont reconnues de manière spécifique par cette enzyme. Une étude plus approfondie sur le complexe ARNIII-ARNm *coa* suggère que la RNase III pourrait en plus de son activité catalytique, exercer une action plus directe sur la formation des hybrides.

### A. La RNase III: état des lieux

(Article III: Chevalier et al., 2009)

La RNase III fait partie du groupe I des endoribonucléases qui coupent spécifiquement les ARN en double brin (Drider et Condon, 2004). Elle est représentée sur l'ensemble du règne vivant (Bactéries, champignons, plantes et mammifères). Chez *E. coli*, cette enzyme est codée par le gène *mnc*; elle est constituée de 226 acides aminés et s'organise sous la forme d'un homodimère. Chaque monomère est divisé en deux domaines majeurs. Dans la région N terminale se trouve le domaine catalytique qui porte l'activité endoribonucléase alors que la région C terminale constitue le domaine de liaison à l'ARN (figure 32, A) (Gan et al., 2006). A ce jour, il n'existe pas de séquence consensus d'ARN associée à l'activité de l'enzyme. Cependant, certaines séquences exercent la fonction d'anti-déterminant et protègent l'ARN contre l'action de la RNase III (Nicholson, 1999). En général, l'enzyme requiert comme substrat minimal, une hélice d'ARN de 22 paires de bases (Nicholson, 1999). Chez *E. coli*, sa fonction majeure est la maturation des ARN ribosomique 16S et 23S (Drider et Condon, 2004; Nicholson, 1999) mais elle a également été associée à la régulation de l'expression des gènes. En effet, elle régule sa propre expression en se fixant dans la région 5'UTR de son ARNm (Matsunaga et al., 1996). Cette enzyme est également souvent associée aux ARN antisens qui régulent l'expression des gènes de plasmides, phages, transposons (Santos et al., 1997; Wagner et al., 2002) et chromosomiques (Darfeuille et al., 2007). Par ces différents exemples, il apparaît que la RNase III reconnaît des structures d'ARN variées comme des hélices interrompues par des

boucles internes ou des hélices empilées. Enfin, l'enzyme régule l'expression d'un ARNm de bactériophage par son activité de fixation et non pas par son activité catalytique (Altuvia et al., 1987). Toutes ces indications suggèrent fortement que cette RNase exerce une activité importante dans la régulation de l'expression des gènes chromosomiques et que celle-ci est souvent détournée au profit des éléments invasifs comme les bactériophages, ou les plasmides. D'une manière intéressante, un certain parallèle peut être effectué avec l'enzyme DICER et son implication dans l'ARNi et les mécanismes de défense contre les virus de plantes. Ainsi, il ne serait pas étonnant que la RNase III puisse avoir une action dans la maturation des ARN CRISPR dans les bactérie Gram-positif (cf introduction, § III.D).



**Figure 32: Caractéristiques de la RNase III:** (A) Structure tridimensionnelle de la RNase III de *E. coli* d'après Gan et al., 2006. Les deux monomères de la RNase III sont représentés en jaune et en turquoise. N et C représentent les extrémités N et C terminales qui portent les sites catalytiques et de liaison à l'ARN respectivement. Les structures vertes et bleues représentent les ARN double brin. (B) Alignement de séquence entre les RNase III de *E. coli* et de *S. aureus*. En rouge sont représentés les acides aminés parfaitement conservés et en vert ceux dont les propriétés biochimiques sont conservées. Les résidus encadrés sont essentiels à l'activité catalytique de la protéine.

La RNase III de *B. subtilis* est essentielle contrairement à celle de *E. coli* où les souches inactivées de cette RNase sont thermosensibles (Li et al., 1993). Ceci n'est pas relié à la maturation des ARNr mais serait probablement issu d'une maturation requise pour un ARNm codant une protéine essentielle de *B. subtilis*. En revanche, dans *S. aureus*, dans les conditions de croissance en milieu riche, l'enzyme n'est pas essentielle (Huntzinger et al., 2005). La RNase III de *S. aureus* est très conservée et par analogie à *E. coli* et *B. subtilis*, il est donc probable qu'elle exerce une fonction dans la maturation de l'ARNr 23S. Par des expériences de northern et de coloration au bromure d'ethidium, nous n'avons cependant pas détecté d'accumulation de précurseur d'ARNr dans la souche mutante. Chez *S. aureus*, des voies de substitution existent ce qui pourrait palier à l'absence de la RNase III (Drider et Condon, 2004). Cependant, des expériences northern plus sensibles devront être réalisées en utilisant des oligonucléotides spécifiques. En utilisant des ARN substrats d'*E. coli*, nous avons montré que la RNase III de *S. aureus* induit les mêmes coupures que l'enzyme d'*E. coli*. Par ailleurs, une mutation du résidu E117 conservé dans *S. aureus* conduit à une perte de l'activité catalytique de l'enzyme alors que la reconnaissance à l'ARN n'est pas perturbée (figure 32, B). Enfin, comme pour l'enzyme de *E. coli*, l'activité catalytique est inhibée par le remplacement du magnésium par le calcium. Ceci indique que les deux enzymes présentent des propriétés proches voir identiques.

Dans *S. aureus*, nous avons montré que les ARNm *spa* et *sa1000* sont des substrats de la RNase III. Ces deux ARNm possèdent dans leur séquence codante une longue structure en tige-boucle qui est fortement coupée *in vitro* (figure 31). De plus, tous les hybrides formés (ARNIII-ARNm *coa*, ARNIII-ARNm *spa*, ARNIII-ARNm *rot*, ARNIII-ARNm *sa1000*) sont coupés efficacement. Ces complexes adoptent des topologies variées incluant des duplexes irréguliers et des interactions boucles-boucles. Dans le cas de l'ARNm *rot* qui contient deux interactions boucle-boucle très similaires, l'utilisation de mutants de l'ARNIII combinée à la RNase III a permis d'identifier les partenaires impliqués dans ces interactions. Ainsi la RNase III est une sonde appropriée pour identifier de tels motifs souvent observés dans les régulations impliquant les ARN antisens.

*Signalement bibliographique ajouté par :*

**L'UNIVERSITÉ DE STRASBOURG**  
**Service Commun de la Documentation**

**Staphylococcus aureus Endoribonuclease III : Purification and Properties**

Clément CHEVALIER, Eric HUNTZINGER, Pierre FECHTER, Sandrine BOISSET,  
François VANDENESCH, Pascale ROMBY and Thomas GEISSMANN

**Methods in enzymology, 2008, vol. 447, chapter 16, pages 309-327**

**Résultats et discussions : p. 151-169**

La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.

Les utilisateurs de l'UdS peuvent consulter cette publication sur le site de l'éditeur :

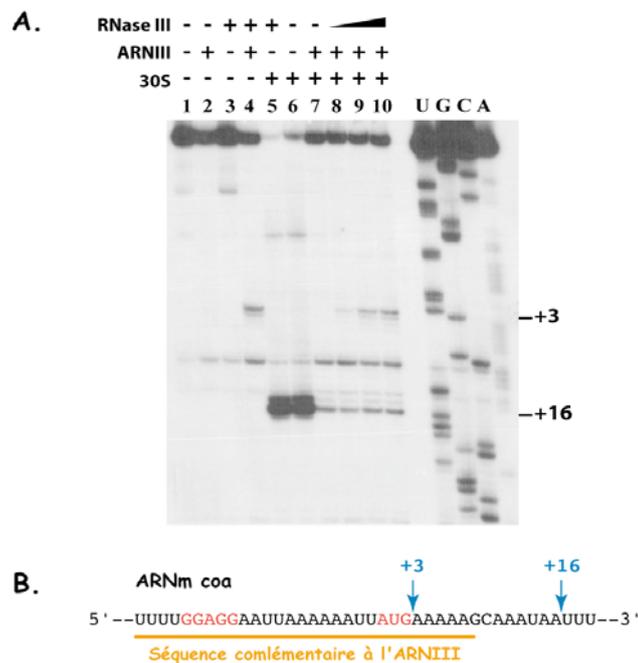
[http://dx.doi.org/10.1016/S0076-6879\(08\)02216-7](http://dx.doi.org/10.1016/S0076-6879(08)02216-7)

La version imprimée de cette thèse peut être consultée à la bibliothèque ou dans un autre établissement via une demande de prêt entre bibliothèques (PEB) auprès de nos services :

<http://scd.unistra.fr/services/peb/>

## B. Est-ce que la RNase III pourrait stabiliser les interactions ARNIII-ARNm ?

Sachant que la RNase III est essentielle *in vivo* pour la répression du gène *coa*, nous avons utilisé l'enzyme mutante E117 afin de découpler l'activité catalytique de la fixation à l'ARN (figure 32, B). Dans un premier temps, nous avons analysé par des expériences de "toeprint" si cette enzyme mutante était capable de stimuler l'action inhibitrice de l'ARNIII sur la fixation du ribosome au niveau des ARNm *spa*, *coa* et *rot*. Ces expériences ont été réalisées sur des complexes ARNIII-ARNm préformés à 37°C. Comme nous avons utilisé un excès d'ARNIII, nous n'avons pas observé d'effet additionnel de la RNase III sur la fixation de la sous-unité 30S. Cependant, dans le cas de l'ARNm *coa*, nous avons observé l'apparition d'une pause de la RT en position +3 (figure 33). Cette pause dépend à la fois de la présence de l'ARNIII et de la RNase III. Cet effet n'a pas été observé pour les ARNm *spa* et *rot*. Ce résultat suggère que la RNase III pourrait stabiliser l'interaction ARNIII-ARNm *coa*.



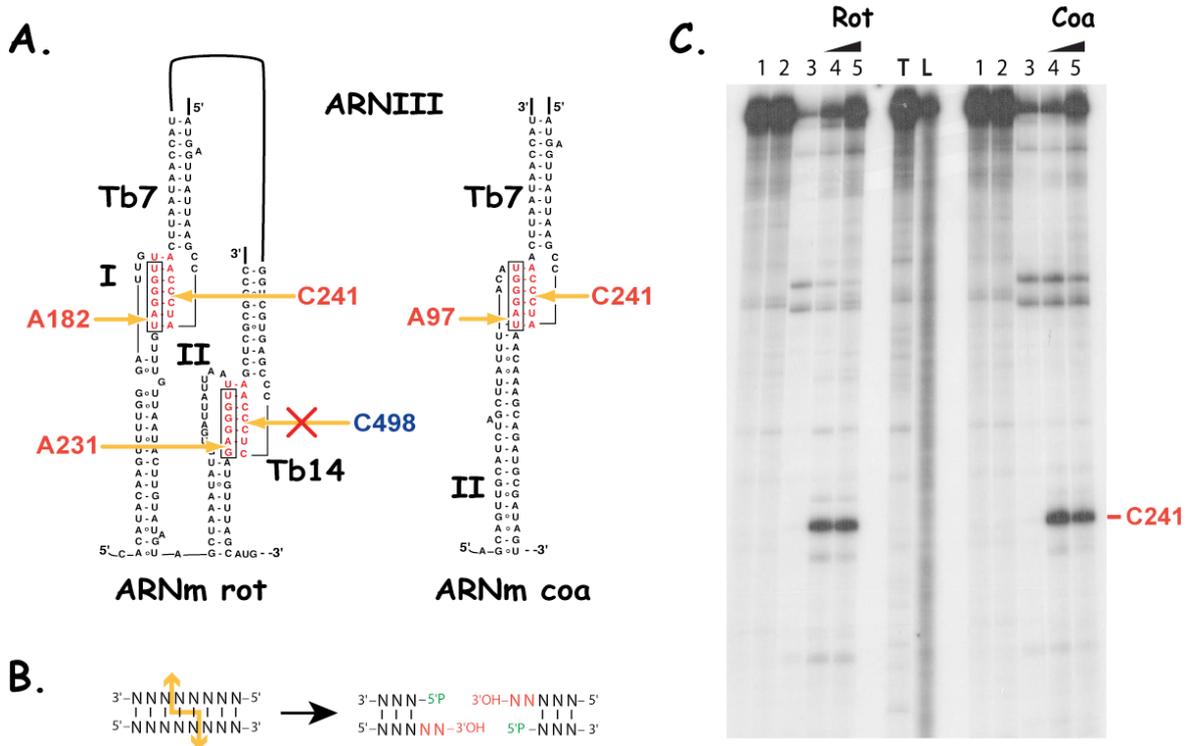
**Figure 33: La RNase III mutante stabilise le duplexe ARNIII-ARNm *coa*.** (A) Toeprint réalisé sur le duplexe ARNIII-ARNm, en présence ou en absence de la RNase III mutante. (1 à 4) contrôles d'incubation. (5, 6) formation du complexe ternaire ARNm-30S-ARNt. (7, 8, 9 et 10) effet de la RNase III sur ce complexe seul ou en présence de concentrations croissantes en ARNIII (200 nM, 400nM et 600 nM). (B) Séquence de la région RBS de l'ARNm *coa*. Les nucléotides en rouge représentent la séquence Shine Dalgarno, et le codon initiateur. +3 indique la pause induite par la RNase III et +16 indique la position attendue du "toeprint".

## C. La RNase III reconnaît les interactions boucles-boucles

### C.1. Les complexes ARNIII-ARNm *rot* et ARNIII-ARNm *coa*

Au cours de notre étude, nous avons constaté que la RNase III était particulièrement adaptée pour reconnaître et couper les interactions boucles-boucles formées entre l'ARNIII d'une part, et les ARNm *rot* ou *coa* d'autre part (Boisset et al., 2007; Chevalier et al., soumis). D'une manière intéressante, les séquences engagées dans la formation des hybrides ARN-ARN sont strictement identiques et la RNase III induit des coupures exactement au même endroit dans la séquence des ARNm (5'-UGGGA↑U-3') alors que la taille des hélices intramoléculaires varie (figure 34, A). Nous avons regardé si cette spécificité se retrouvait du côté de l'ARNIII. En effet, chacun des homodimères de la RNase III coupe un brin de l'hélice d'ARN. Cette double coupure libère des extrémités 5' phosphates ainsi que deux bases chevauchantes du côté 3'OH (figure 34, B). Ainsi, connaissant les propriétés catalytiques de l'enzyme, il était attendu de retrouver des coupures du côté ARNIII, au niveau des résidus C241 de la tige-boucle 7 et C498 de la tige-boucle 14 (figure 34, A). Pour vérifier cette hypothèse, l'ARNIII marqué en 5' a été digéré par la RNase III, en absence ou en présence de différentes concentrations des ARNm *rot* et *coa*. Ces expériences montrent que la formation des différents duplexes induit l'apparition d'une forte coupure au niveau du résidu C241 mais pas au niveau du résidu C498 (figure 34, C). Ces résultats confirment que la RNase III coupe spécifiquement les interactions impliquant la tige-boucle 7 de l'ARNIII. Cependant ils montrent aussi que l'interaction entre la tige-boucle 14 de l'ARNIII et la tige-boucle II de l'ARNm n'induit pas de coupure côté ARNIII. La boucle II de l'ARNm *rot* étant particulièrement grande (17 nucléotides), la taille des boucles serait donc un élément important pour induire une reconnaissance spécifique par la RNase III.

Comme les interactions boucles-boucles engagées entre l'ARNIII et ses cibles sont très similaires, il était difficile de dire à ce stade qui de la séquence ou de la structure était réellement nécessaire pour induire une coupure spécifique. La deuxième partie de ce travail avait donc pour but de répondre à cette question.

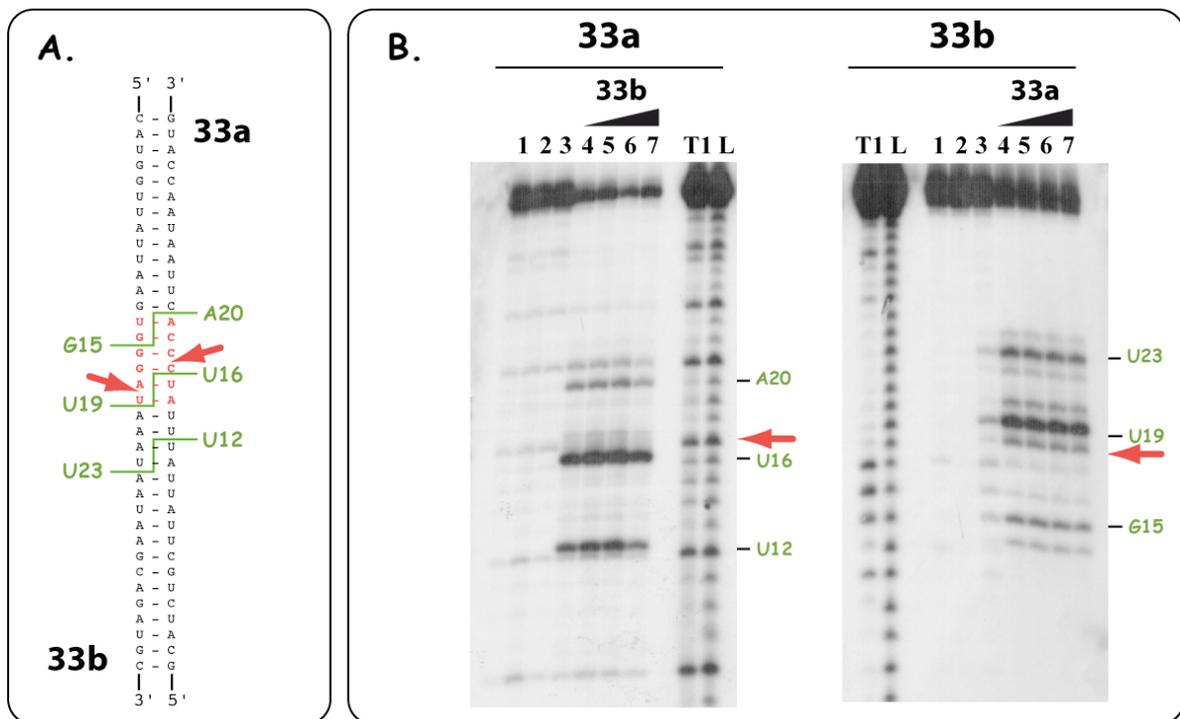


**Figure 34: la RNase III coupe spécifiquement l'ARNIII.** (A) Structure secondaire des interactions boucle-boucle entre l'ARNIII et les ARNm *rot* et *coa*. (B) Mode de coupure de la RNase III. Les extrémités 5' phosphates et 3'OH sont représentés en vert et rouge, respectivement (C) Digestion de l'ARNIII par la RNase III. (1 et 2) contrôles d'incubation. L'ARNIII a été digéré seul (3) ou en présence d'une concentration croissante en ARNm *rot* ou *coa* (100 nM (4) et 200 nM (5)). T et L, RNase T1 en condition dénaturante et hydrolyse alcaline respectivement. Les flèches jaunes représentent les coupures RNase III.

## C.2. La séquence ne semble pas être un déterminant de reconnaissance

Dans le but d'identifier qui de la structure ou de la séquence est déterminante pour induire une coupure spécifique de la RNase III, deux oligonucléotides (33a et 33b) ont été synthétisés chimiquement. Ces oligonucléotides s'hybrident sous la forme d'un long duplexe étendu d'ARN (33 nucléotides) et portent les séquences impliquées dans les interactions entre la tige-boucle 7 de l'ARNIII et les ARNm *rot* ou *coa* (figures 34 et 35, A). Après les étapes classiques de dénaturation/renaturation puis d'hybridation, les oligonucléotides marqués ont été digérés par la RNase III afin de déterminer les sites préférentiels de coupure. Les résultats montrent que le duplexe d'ARN est fortement clivé à trois reprises par la RNase III (figure 35, B). Ces coupures sont observées sur les deux oligonucléotides mais aucune ne survient au niveau des sites attendus de l'ARNm (5'-UGGGA↑U-3') ou de l'ARNIII (5-AUC↑CCA-3') (figure 35). Puisque 33a et 33b se replient sous la forme d'une hélice d'ARN parfaite, la seule contrainte imposée à la

RNase III était la séquence de l'hélice. Or, les trois coupures interviennent au centre de l'hélice, certainement les positions les plus favorables à la fixation de la RNase III, et semblent totalement indépendantes de la séquence. Ceci suggère que la séquence n'est pas déterminante pour induire la coupure par la RNase III. Il est donc probable que la topologie structurale soit le facteur clé pour le recrutement de la RNase III.

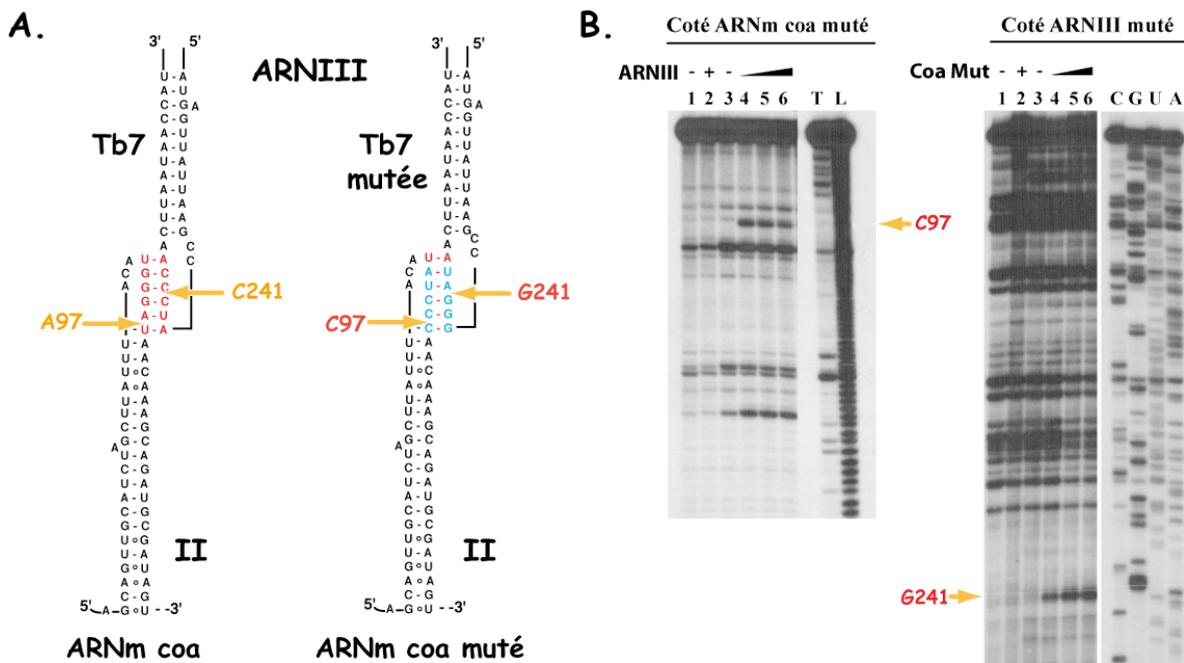


**Figure 35: Effet de la RNase III sur des oligos d'ARN synthétisés chimiquement.** (A) Les oligonucléotides 33a et 33b s'assemblent sous la forme d'un long duplexe d'ARN. 33a porte la séquence d'interaction de la tige-boucle 7 de l'ARNIII et 33b celle des ARNm (en rouge). Les séquences en noir sont totalement aléatoires. (B) Digestion par la RNase III du duplexe 33a/33b. Le gel de gauche montre les coupures du côté 33a et le gel de droite du côté 33b. (1 et 2) contrôles d'incubation. Les oligonucléotides ont été digérés seul (3) ou en présence d'une concentration croissante d'oligonucléotide antisens ((4) 50 nM, (5) 100 nM, (6) 200 nM, (7) 300 nM ). Les flèches rouges représentent les coupures attendues telles qu'elles sont observées dans les ARN naturels. Les lignes vertes sont les coupures obtenues dans ces expériences.

### C.3. La topologie de l'interaction boucle-boucle est l'élément de reconnaissance

Afin de confirmer les résultats obtenus avec les oligonucléotides synthétisés chimiquement, des mutations compensatoires ont ensuite été réalisées au niveau des séquences complémentaires de la tige-boucle 7 de l'ARNIII et de la tige-boucle II de l'ARNm *coa*. Ces mutations ont été réalisées de manière à conserver du mieux possible la structure secondaire des deux tiges-boucles. Les paires de bases complémentaires ont simplement été échangées entre l'ARNIII et l'ARNm et l'ordre de la séquence d'interaction

a été modifié. Ainsi, les sites de coupure RNase III C241 de l'ARNIII et A97 de l'ARNm deviennent respectivement G241 et C97 dans les ARN mutés (figure 36, A). Les ARN mutés ont ensuite été hybridés *in vitro* et le duplexe résultant a été digéré par la RNase III (figure 36, B). Du côté ARNm *coa*, les résultats montrent que la RNase III induit une forte coupure exactement à la position attendue C97. Du côté ARNIII, la RNase III induit une coupure en position G241. L'ensemble de ces résultats confirment que les ARN sauvages et mutants sont clivés par la RNase III exactement aux mêmes positions C241/A97 et G241/C97, respectivement. Comme séquences d'interactions ont été modifiées sans altérer les structures secondaires des ARN, la topologie des interactions boucles-boucles est donc bien le facteur décisif pour orienter une coupure spécifique par la RNase III.

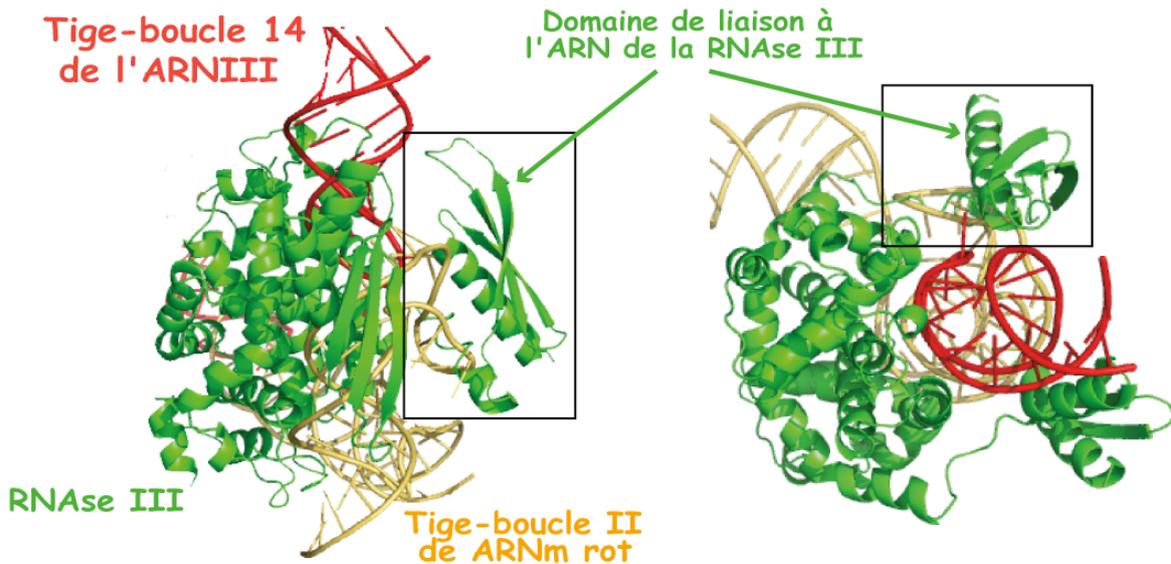


**Figure 36: La RNase III coupe à l'identique les interactions boucle-boucle ARNIII-ARNm sauvage ou mutées.** (A) Structure secondaire de l'interaction entre la tige-boucle 7 de l'ARNIII et la tige-boucle II de l'ARNm *coa*. La séquence en rouge représente le site d'interaction sauvage alors que la séquence en bleu montre les mutations compensatoires introduites dans l'ARNIII et dans l'ARNm. Les flèches jaunes montrent les coupures induites par la RNase III. (B) Digestion par la RNase III des duplexes ARNIII-ARNm mutés, côté ARNm (gel de gauche) ou côté ARNIII (gel de droite). (1 et 2) contrôles d'incubation. L'ARNm *coa* ou l'ARNIII marqués ont été digérés seuls (3) ou en présence d'une concentration croissante d'ARNIII ou d'ARNm *coa* non marqués, respectivement (100 nM (4), 200 nM (5) et 300 nM (6)). Dans le cas de l'ARNIII, la digestion par la RNase III a été suivie par une élévation par la reverse transcriptase. T et L, digestion en condition dénaturante par la RNase T1 et hydrolyse alcaline respectivement. C, G, U et A séquence réalisée par la reverse transcriptase.

## D. Discussion

### La topologie des interactions boucles-boucles guide la coupure par la RNase III

Notre étude montre que la RNase III est particulièrement utile pour reconnaître les interactions de type boucle-boucle. En effet, ces interactions induisent un empilement coaxial des hélices intermoléculaires formant une longue structure propre à la fixation de la RNase III (figure 34, A). D'une manière surprenante, la RNase III induit des coupures uniques et spécifiques qui interviennent au niveau des ARNm *rot* ou *coa* mais également au niveau de l'ARNIII (figure 34, B). A priori, trois facteurs pourraient influencer la reconnaissance: les séquences d'interaction boucle-boucle, la taille des hélices intermoléculaires et la topologie des duplexes d'ARN. Même si les séquences engagées dans les interactions ARNIII-ARNm cibles sont identiques, nos résultats montrent que ces dernières n'influencent pas directement le positionnement de la RNase III sur le duplexe d'ARN (figure 35); elles ne sont donc pas responsables de la spécificité de coupure. En outre, la taille des hélices des tige-boucles impliquées dans les interactions boucle-boucle varie, par conséquent ce facteur ne serait pas non plus déterminant. Il ne reste plus que la topologie structurale. Les mutations compensatoires introduites dans l'ARNIII et dans l'ARNm *coa* confirment que ce facteur est important pour orienter les coupures RNase III. Il est à noter que la RNase III ne coupe pas l'ARNIII lorsque l'interaction a lieu entre la tige-boucle 14 de l'ARNIII et la tige-boucle II de l'ARNm *rot* (figure 34, A). Basé sur nos résultats de cartographie en solution, un modèle moléculaire de l'interaction entre l'ARNIII et l'ARNm *rot* a été réalisé avec la collaboration de Benoît Masquida (figure 37). Ce modèle montre que la boucle II de l'ARNm *rot* qui interagit avec la boucle 14 de l'ARNIII interfère également avec l'un des domaines de liaison à l'ARN de la RNase III. Ainsi, la fixation d'un des homodimères de l'enzyme serait affectée, ce qui empêcherait la coupure au niveau de l'ARNIII.



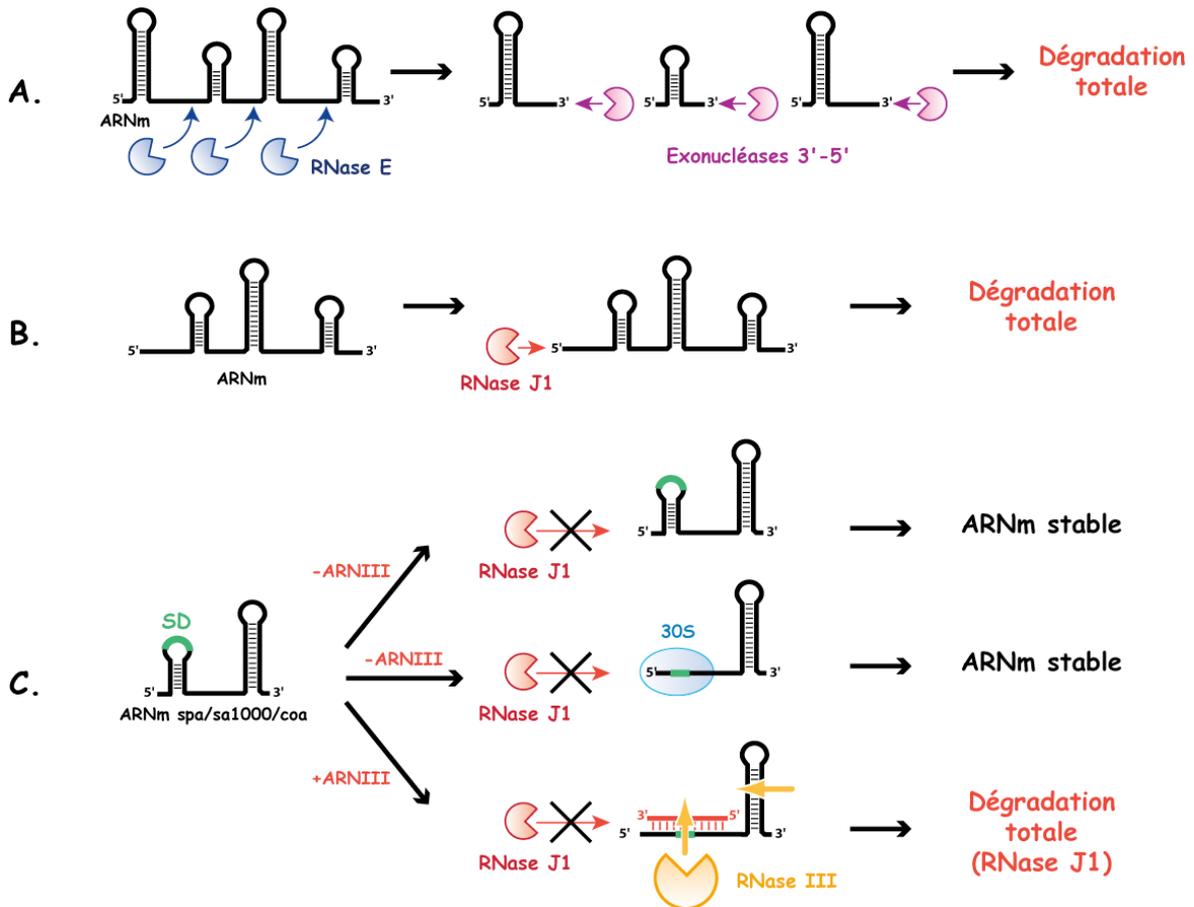
**Figure 37: Modélisation de l'interaction boucle-boucle formée entre la tige-boucle II de l'ARNm *rot* et la tige-boucle 14 de l'ARNIII, en présence de la RNase III.** Les cadres noirs montrent que la longue boucle II de l'ARNm interfère avec le bon positionnement du domaine de liaison à l'ARN de la RNase III. Les coordonnées de la structure cristallographique de la RNase III sont de Gan et al., (1996)

### Rôles possibles de la RNase III dans la régulation par l'ARNIII

#### *La RNase III et la régulation: l'ARNIII un guide pour la RNase?*

Nous avons montré que la RNase III mutante dépourvue de son activité catalytique stabilise le duplexe ARNIII-ARNm *coa*. Cependant, la coupure de l'enzyme pourrait être l'événement nécessaire à son relargage. On ne peut donc pas exclure que la mutation dans le site catalytique stabilise la RNase III sur l'hybride ce qui bloque la reverse transcriptase en position +3 (figure 33). Cependant, le duplexe ARNIII-ARNm *coa* est plus court que les duplexes ARNIII-ARN *spa* et ARNIII-ARNm *sa1000*. Ainsi, il est probable que le duplexe ARNIII-ARNm *coa* ait besoin d'une aide extérieure pour être stabilisé et que la RNase III soit le facteur requis. Pour tester cette hypothèse, des expériences complémentaires seront nécessaires. Nous envisageons de préformer le complexe ARNIII-RNase III mutante et d'analyser si celui-ci stabilise l'interaction ARNIII-ARNm *coa* par des expériences de retard sur gel. Des études cinétiques analyseront la contribution éventuelle de la RNase III sur la vitesse d'association des deux ARN. Enfin des expériences de traduction *in vitro* seront réalisées en testant l'ordre d'addition des deux différents partenaires (ARNIII, RNase III mutante). Récemment, Efthimia Lioliou, chercheur post-doctoral dans notre équipe, a co-immunoprécipité l'ARNIII et la RNase III mutante portant un "flag-tag". Par ailleurs, de précédents résultats avaient déjà montré que

la RNase III sauvage pouvait être copurifiée avec un ARNIII biotinylé sans induire de coupure (Huntzinger et al., 2005). Même si les sites d'interaction restent encore à identifier, ces observations confirment que les deux molécules interagissent ensemble *in vivo*. Ainsi, l'ARNIII pourrait très bien servir de guide à la RNase III et la diriger vers des ARNm cibles. Dans ce cas de figure, la traduction et la dégradation des ARNm seraient alors couplées.



**Figure 38: Mécanismes de dégradation des ARNm.** (A) Dans les bactéries Gram-négatif, la dégradation des ARNm est initiée par la RNase E et implique ensuite des exonucléases 3'-5'. (B) Dans les bactéries Gram-positif comme *B. subtilis*, les ARNm sont dégradés en partie par leur extrémité 5' grâce à l'activité 5'-3' de la RNase J1. (C) Plusieurs ARNm de *S. aureus* possèdent de courtes régions 5'UTR qui les protègent de la RNase J1 en formant une structure secondaire stable ou en recrutant la sous unité 30S du ribosome. L'ARNIII interagit avec ces régions 5'UTR et recrute la RNase III qui peut alors dégrader les ARNm. SD représente la séquence Shine Dalgarno.

### *La RNase III et la dégradation des ARNm réprimés par l'ARNIII*

Que la répression des gènes soit fortement ou partiellement dépendante de la RNase III, cette enzyme semble jouer un rôle clé dans la régulation des gènes de virulence chez *S. aureus*. Chez *E. coli*, de récentes études montrent que l'inhibition traductionnelle causée par certains ARN régulateurs est suffisante pour induire une répression efficace des

gènes, et que la dégradation ne serait qu'une conséquence de l'interaction ARNnc-ARNm (Morita et al., 2006). Ces derniers résultats posent donc la question de l'importance de la RNase III chez *S. aureus*. Si la dégradation n'est qu'une conséquence d'un mécanisme de répression antisens, pourquoi cette enzyme est-elle si importante pour réprimer les gènes de virulence chez *S. aureus*? Une réponse pourrait venir de la différence qu'il existe entre les bactérie Gram-négatif et Gram-positif au niveau de leur machinerie de dégradation des ARNm (Condon, 2007). En effet, dans les bactéries Gram-négatif, la dégradation des ARNm est principalement initiée par la RNase E. Cette endoribonucléase coupe les régions en simple brin et libère des extrémités 3'OH afin que les exoribonucléases 3'-5' dégradent les ARNm. (figure 38, A) Dans les bactéries Gram-positif, un récent modèle propose que la dégradation des ARNm serait initiée par l'activité exoribonucléases 5'-3' de la RNase J1 (figure 38, B) (Mathy et al., 2007). Chez *S. aureus*, la plupart des ARNm cibles de l'ARNIII possède des régions 5' UTR courtes caractérisées par une séquence SD forte et une tige-boucle à l'extrémité 5'. Ces trois caractéristiques protègent l'ARNm contre la dégradation par la RNase J1 (Daou-Chabo et al., 2009). En effet, le recrutement de la sous-unité 30S pourrait se faire en tout début de transcription de l'ARNm ce qui bloquerait la progression de l'exoribonucléase J1 (Daou-Chabo et al., 2009; Mathy et al., 2007). Ainsi, en ciblant les hybrides ARNIII-ARNm, la RNase III fournirait un mécanisme de dégradation efficace qui permettrait de palier à l'inefficacité de la RNase J1 envers certains ARNm particulièrement stables, i.e. ARNm *spa*, *sa1000* (figure 38, C).

Comme décrit ci-dessus, E. Lioliou a exprimé la RNase III mutante portant un "flag-tag" dans *S. aureus* afin de co-immunoprécipiter l'ensemble des ARN reconnus par l'enzyme. Après séquençage par la méthode de "deep sequencing" des ADNc, cette étude devrait conduire à l'identification de nouveaux substrats ARN de la RNase III et peut être révéler de nouveaux hybrides ARNnc-ARNm. D'ores et déjà, E. Lioliou a retrouvé l'ARNm *spa* et l'ARNm *rnc* (codant pour la RNase III) dans les fractions immunoprécipitées. Ces résultats sont encourageants puisque l'ARNm *spa* est une cible de la RNase III (Huntzinger et al., 2005) et ils pourraient indiquer que comme dans *E. coli*, la RNase III autorégulerait sa propre expression (Matsunaga et al., 1996). Ainsi, l'ensemble de ces résultats devrait conduire à définir la contribution de cette enzyme dans la régulation de l'expression des gènes de *S. aureus*.

### **III. Identification et caractérisation de nouveaux ARN régulateurs chez *S. aureus***

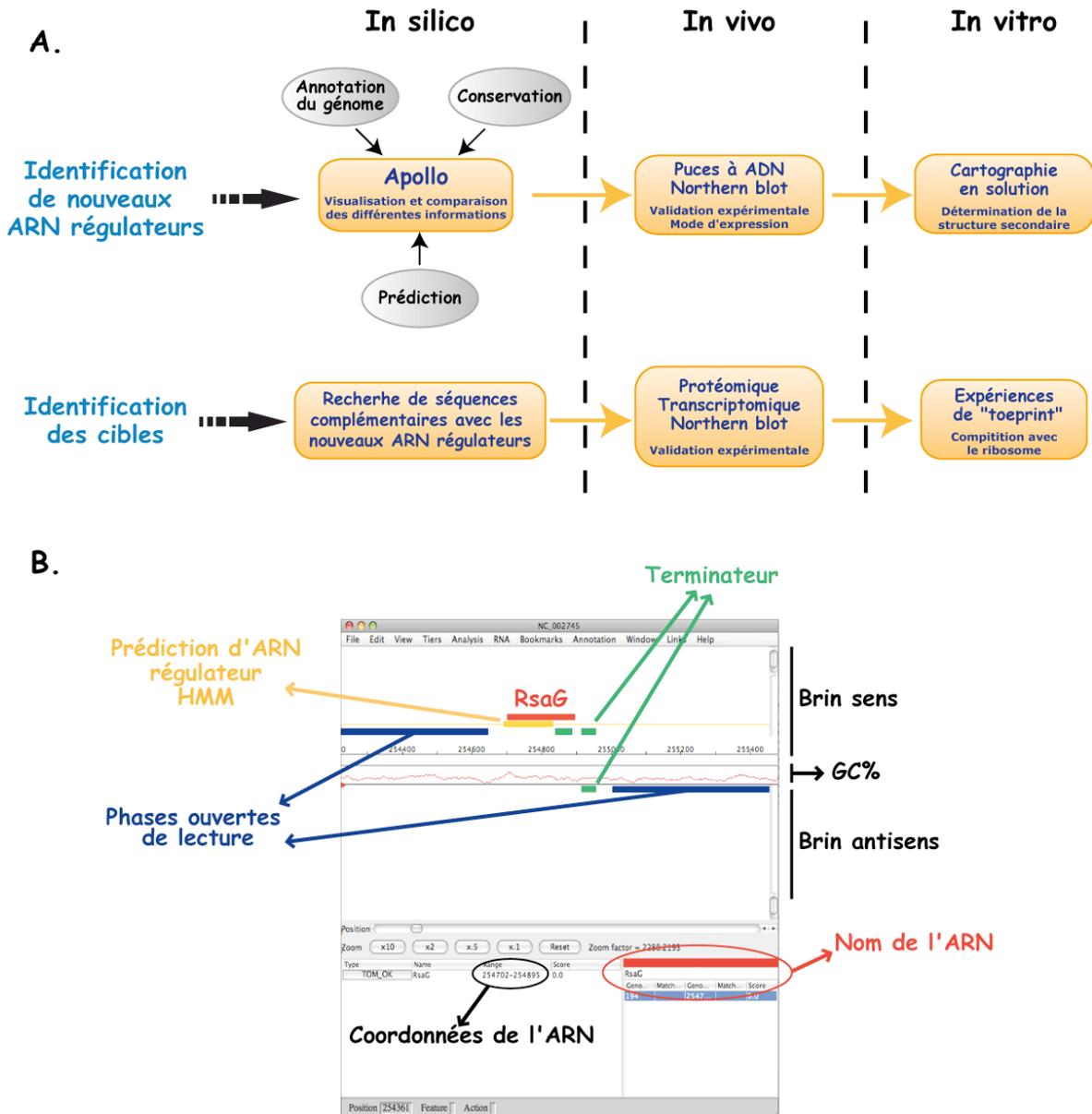
(Article IV: Geissmann et al., 2009)

Au début de ma thèse, l'ARNIII était le seul ARN régulateur à posséder une fonction bien établie chez *S. aureus*. Nous avons choisi une analyse qui couple une recherche génomique des régions intergéniques et l'analyse de leur expression. Une telle méthode a été réalisée avec succès dans de nombreuses bactéries (Argaman et al., 2001; Livny et al., 2008; Marchais et al., 2009; Pichon et Felden, 2005; Rivas et Eddy, 2001; Wassarman et al., 2001). Elle a l'inconvénient de générer de nombreux faux-positifs mais à l'inverse permet de caractériser les régions potentiellement intéressantes en terme de séquence ou de motifs structuraux. Il est clair que dans un proche avenir, les méthodologies ("deep sequencing", "tiling arrays") qui ont été développées récemment aborderont de manière plus systématique l'ensemble des ARN qui sont induits en réponse à un signal donné.

#### **A. Stratégie expérimentale**

##### **A.1. Recherche de nouveaux ARN non codants**

Les ARN régulateurs sont très hétérogènes. Il est donc difficile de définir des critères de choix. Pour la première fois, nous avons utilisé un ensemble d'outils expérimentaux qui incluent la recherche de motifs particuliers (promoteurs de transcription, terminateur de transcription Rho-indépendant, structure en tige-boucle, séquence répétée), les régions conservées, et une analyse statistique (figure 39, A). L'ensemble de ces outils a été utilisé et mis en place grâce à la collaboration avec C. Gaspin (INRA, Toulouse).



**Figure 39: Stratégie utilisée pour identifier les ARN Rsa ainsi que leurs cibles.** (A) Une combinaison d'approches *in silico*, *in vivo* et *in vitro* ont été nécessaires aussi bien pour identifier les ARN Rsa que pour identifier des cibles. (B) Représentation d'une fenêtre ouverte du programme de visualisation Apollo. Grâce à cette approche, de multiples données (annotation, conservations, prédictions) peuvent être comparées simultanément. Cet exemple montre les différents critères qui ont permis d'identifier RsaG (en rouge).

La recherche des conservation de séquence a été réalisée de manière simultanée sur plusieurs génomes, soit dans diverses souches de *S. aureus*, soit dans au moins quatre souches de Staphylocoques (*S. epidermidis*, *S. haemolyticus*, *S. saprophyticus*, *S. aureus*), soit dans les firmicutes. Cette recherche de conservation de séquence a été couplée à la recherche de motifs de structure secondaire conservés (Rivas et Eddy, 2001). L'approche statistique inclue le pourcentage de séquence riche en G/C le long du génome, et l'énergie libre. L'ensemble de ces données a été visualisé par le programme APOLLORNA qui

donne en plus les informations d'annotation des génomes. Ce programme inclue la prédiction des structures secondaires d'ARN ainsi que la recherche de complémentarités ARN-ARN (figure 39, B). A partir de ces résultats nous avons sélectionné plus de 500 régions intergéniques possédant au moins deux critères, et ayant une taille supérieure à 150 nucléotides. Nous avons tenu compte des régions spécifiques aux *S. aureus*, conservées dans plusieurs staphylocoques. Dans un premier temps, nous avons sélectionné 35 régions dont l'expression a été ensuite analysée par des expériences de puces « ADN » mises au point dans l'équipe par T. Geissmann. Cette étude grossière a ainsi mis en évidence 11 nouveaux ARN non codants (RsaA-K) dont l'expression a ensuite été analysée individuellement dans différentes souches de *S. aureus*. Après avoir déterminé les bornes des ARN par la technique de RACE sur des ARN circularisés par l'ARN ligase du phage T4, nous avons produit les ARN *in vitro* par l'ARN polymérase du phage T7. La structure secondaire de plusieurs d'entre eux a été réalisée par cartographie en solution.

## A.2. Identification des cibles

Pour chacun des ARN, nous avons tout d'abord recherché les complémentarités de séquences avec les régions 5' non codantes des ARNm annotés par le biais de différents outils. Nous avons utilisé la méthode mise au point pour l'ARNIII ainsi que le programme développé sur les ARNnc de *I. monocytogenes* (Mandin et al., 2007). Cette dernière méthode tient compte de la force d'interaction des appariements Watson-Crick et pénalise les résidus non appariés et les boucles. Ce programme a été entraîné sur des hybrides déterminés expérimentalement comme l'ARNIII-ARNm *spa*. Il nous est apparu important d'utiliser plusieurs approches qui malheureusement ne génèrent que très peu de recoupement. Un autre critère que nous avons retenu est la conservation des hybrides.

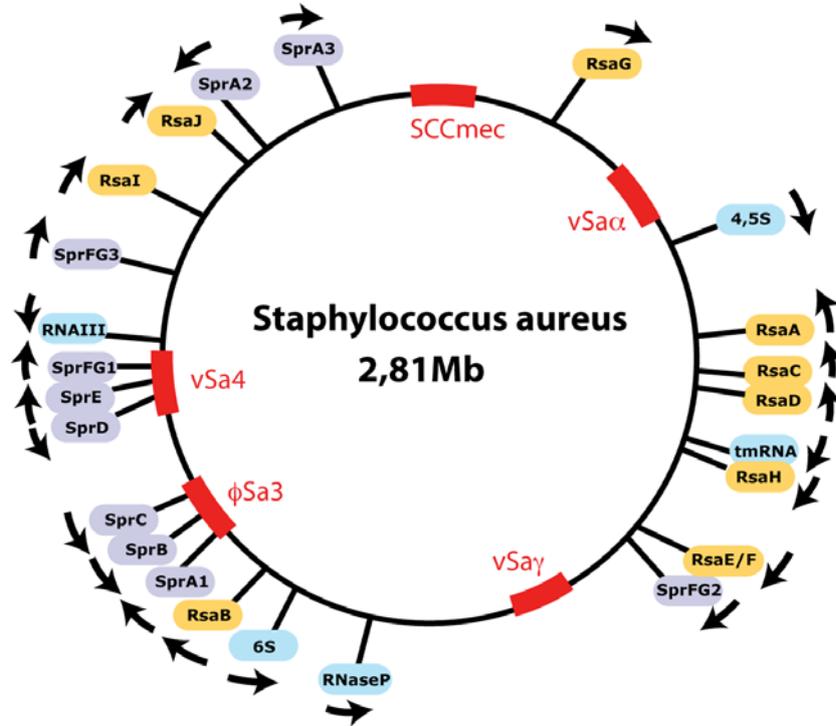
L'équipe du Prof. F. Vandenesch a réalisé plusieurs mutants des ARNnc par remplacement allélique (RsaA, RsaG, RsaE, RsaH). Pour deux des mutants (RsaE, RsaH), des expériences de protéome comparatif (DIGE) ont été réalisées avec P. Fechter dans l'équipe pour analyser l'effet des ARN sur la synthèse des protéines cytosoliques et sécrétées. En collaboration avec le Prof. J. Schrenzel (Genève, Suisse), l'effet des ARNnc sur le taux des ARNm a été analysé par des expériences de transcriptome. L'ensemble de ces méthodes prédictives et comparatives nous ont permis de sélectionner plusieurs cibles

potentielles. Nous avons par la suite validé ces interactions par retard sur gel, et cartographie en solution. L'effet de la fixation de ces ARN sur la formation des complexes de la traduction a été mesuré par des expériences de toeprint.

## B. Etudes des régions intergéniques

En étudiant les régions intergéniques nous avons retrouvé et identifié trois types de régions régulatrices. (1) Nous avons notamment observé de nombreuses régions répétées STAR ("*Staphylococcus aureus* repeat") réparties sur tout l'ensemble du génome. De manière intéressante, ces régions se retrouvent dans certains ARNnc (i.e. SprA3) et sont souvent proches des sites d'intégration. La fonction des éléments STAR reste encore mystérieuse à ce jour. Cependant, la conservation unique chez *S. aureus*, et leur localisation génomique suggèrent qu'ils pourraient contribuer à l'évolution des génomes, et faciliter l'intégration de nouveaux gènes. La ressemblance de certains motifs STAR avec les éléments CRISPR laisse penser que ces régions pourraient avoir évolué vers de nouvelles fonctions régulatrices. (2) Notre approche a également permis de retrouver plusieurs régions régulatrices agissant en *cis* comme les ARN à boîte T qui régulent l'expression des aminoacyl-ARNt synthétases. Par ailleurs, la plupart des "riboswitch" qui répondent à la concentration intracellulaire d'un métabolite (thiamine pyrophosphate, S- adénosyl-méthionine, lysine, guanine, glucosamine-6 phosphate ...) ont été identifiés. Toutes ces régions en *cis* sont très bien conservées dans la plupart des firmicutes et sont généralement impliquées dans la régulation du métabolisme ou des composants de la machinerie de traduction. (3) Cette analyse combinée a permis de retrouver de nombreux ARNnc stables et conservés dans toutes les bactéries (RNase P, SRP, 6S, tmARN...) ainsi que tous les ARN Spr issus des îlots de pathogénie et précédemment identifiés par Pichon & Felden en 2005. Enfin, nous avons identifié 10 nouveaux ARNnc (RsaA-J) répartis de manière aléatoire sur l'ensemble du génome (figure 40). Ces ARN sont différemment conservés. Certains sont uniquement présents chez *S. aureus* (RsaB, C, G, H et J, K) et ont sans doute évolué récemment avec l'adaptation de la bactérie à ses nouveaux habitats et l'apparition de la pathogénie. Les autres ARNnc sont conservés chez les staphylocoques (RsaA, D, F et I) et un seul a été retrouvé dans l'ensemble des *Staphylococcaeae* et *Bacillaceae* (RsaE). Ces derniers ont une origine plus ancienne et doivent probablement contrôler des gènes impliqués dans le métabolisme ou dans divers stress. Ainsi, la

conservation peut apporter des indices importants pour aider à identifier les cibles et la fonction de ces nouveaux ARNnc.

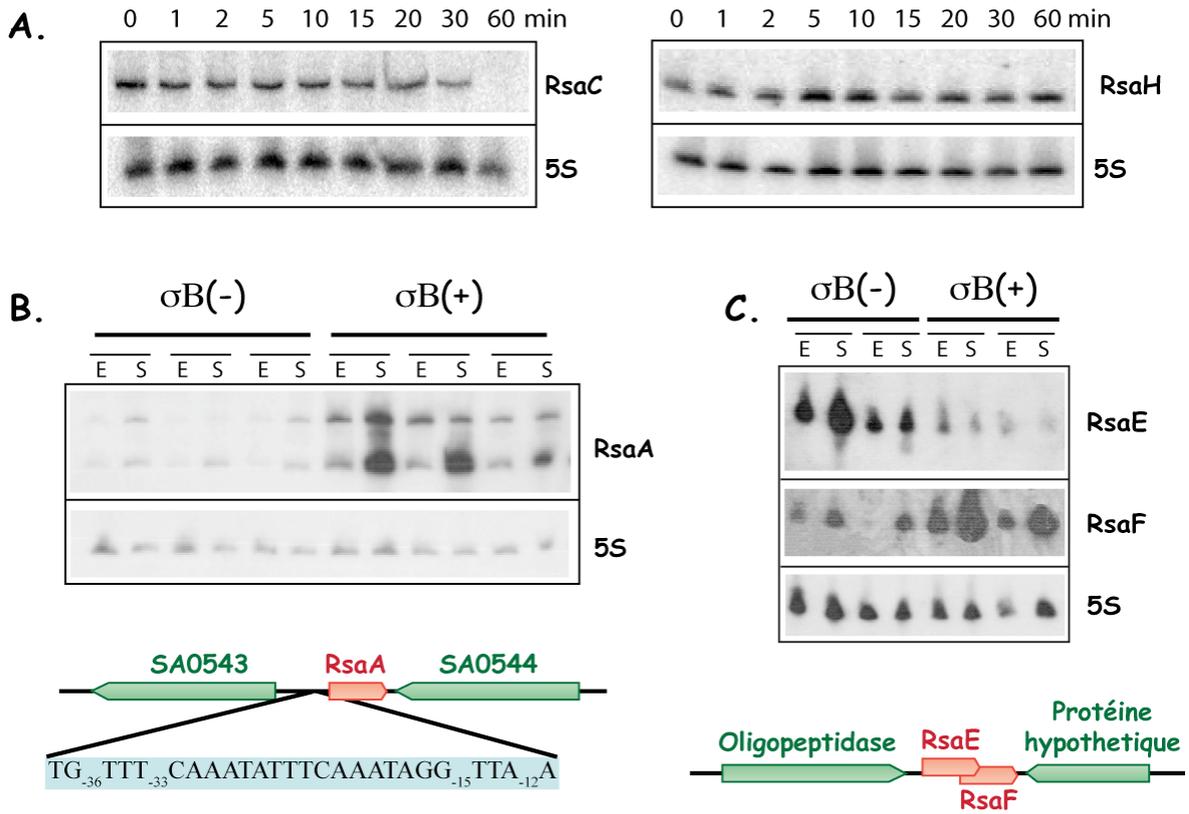


**Figure 40: Localisation et orientation des ARN Rsa.** Carte génétique du génome de *S. aureus* qui représente la localisation et l'orientation des ARN Rsa (en orange). L'ARNIII ainsi que des ARN non codants conservés sont indiqués en bleu. En violet sont représentés les ARN régulateurs SprA-F identifiés par Pichon et Felden en 2005. Les zones rouges correspondent aux différents îlots de pathogénies et les flèches noires indiquent l'orientation des gènes. Mb correspond à Mégabase.

### C. L'expression des ARN Rsa est fortement régulée

L'expression des ARN Rsa a été testée en fonction de différents stress (pH, chocs thermiques, pression osmotique...) et dans différentes souches de *S. aureus* cliniques (Newman, COL) et mutantes (délétées du facteur  $\sigma^B$ , de l'opéron *agr*, de la protéine Hfq ou encore de la RNase III). Nos résultats montrent que la stabilité des ARN Rsa n'est pas dépendante de la protéine Hfq. Leur temps de demie-vie varie entre 15 minutes (RsaC) et plus d'une heure (RsaH), ils sont donc très stables (figure 41, A). Comme l'ARNIII, ils sont exprimés principalement en phase stationnaire de croissance. Néanmoins, certains sont transcrits en fonction de conditions environnementales particulières. Par exemple, RsaA, E et I sont exprimés en réponse à divers stress (choc osmotique, stress oxydatif, pH acide)

alors que RsaC et D sont plutôt induits lors d'un choc thermique froid. Par ailleurs, RsaA, D et F sont dépendants du facteur  $\sigma^B$ . En fin de phase de croissance, ce facteur adapte rapidement le métabolisme de la bactérie en réponse à différents stress induits par le milieu extérieur, comme par exemple lors de l'infection d'un hôte (voir introduction, § **IV.C.3**). Ainsi, RsaA, D et F pourraient être reliés au régulon  $\sigma^B$  et participer à l'adaptation de la bactérie à ses multiples environnements (figure 41, B et C).



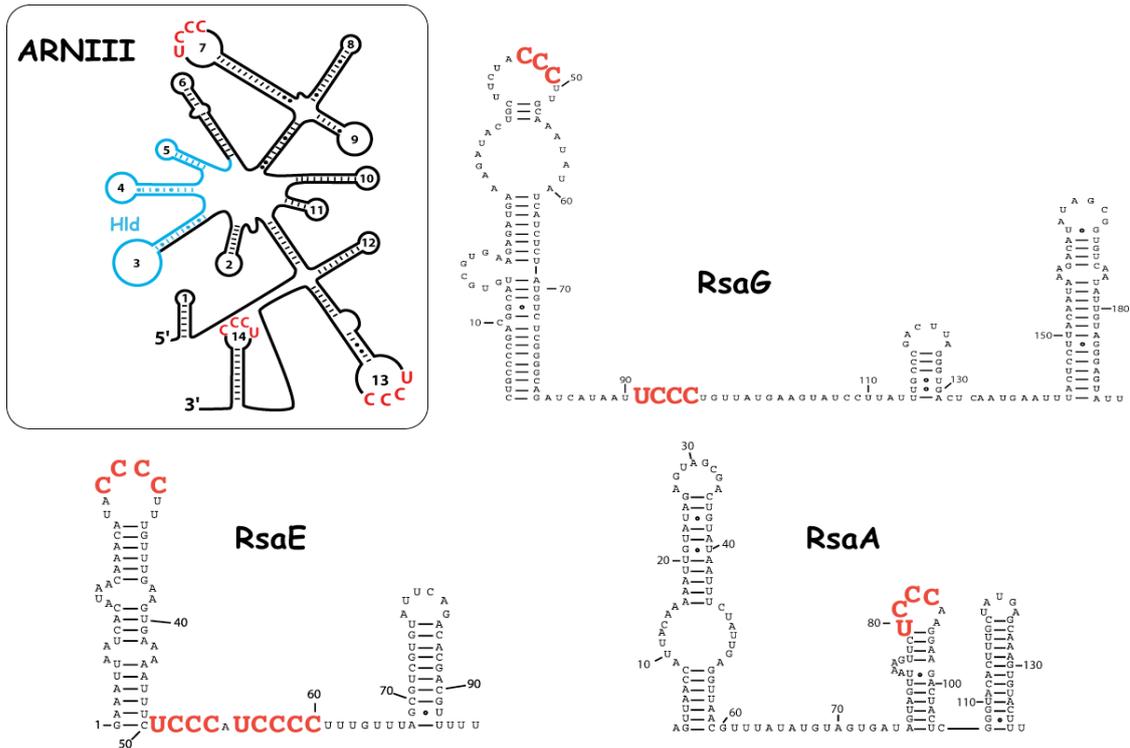
**Figure 41: L'expression des ARN Rsa est analysée par des expériences de northern.** (A) L'expression des ARN RsaC et RsaH est visualisée à différents temps (en minutes) après traitement à la rifampicine. Les deux ARN sont stables: le temps de demi-vie de RsaC est ~15 minutes alors que celui de RsaH est supérieur à 1 heure. L'expression de RsaA (B) et de RsaE et RsaF (C) est visualisée dans différentes souches de *S. aureus* sigma B négatives ou positives. L'organisation génétique des ARN est également donnée: les ARN Rsa sont représentés en rouge, les gènes flanquant en vert et le promoteur sigma B en bleu. E et S représentent les phases de croissance exponentielle et stationnaire respectivement. 5S indique le contrôle interne des expériences de northern.

L'organisation génétique de RsaE et RsaF est complexe. L'extrémité 3' de RsaE correspond à l'extrémité 5' de RsaF et pourtant ces deux ARN sont exprimés de manière différentielle (figure 41, C). En effet, alors que RsaE est fortement exprimé dans les souches  $\sigma^B$  négatives et faiblement dans les souches  $\sigma^B$  positives, le phénomène est inversé pour RsaF (figure 41, C). Comme RsaF est directement dépendant de  $\sigma^B$ , la synthèse de ce

dernier pourrait donc inhiber la transcription de RsaE par un phénomène d'interférence transcriptionnelle ce qui expliquerait pourquoi RsaE n'est plus exprimé dans les souches  $\sigma^B$  positives (Shearwin et al., 2005). Ainsi, le facteur  $\sigma^B$  régulerait de manière indirecte l'expression de RsaE. Cependant RsaF pourrait aussi être issu d'un transcrit plus long incluant RsaE et qui aurait subi une dégradation ultérieure. Nous avons effectivement détecté un long transcrit par des expériences de northern, mais malheureusement je n'ai pas réussi à déterminer les bornes de cet ARN. Des études récentes dans *I. monocytogenes* ont montré que le facteur  $\sigma^B$  induit la synthèse de longs ARN antisens dont l'un d'entre eux pourrait réguler la motilité de la bactérie (Toledo-Arana et al., 2009). Ainsi on ne peut pas exclure que l'induction de ce long ARN incluant RsaE et RsaF puisse réguler les gènes en aval codés en orientation inverse. Il sera donc essentiel de définir avec plus de précision les acteurs régulant la transcription de cet ARN et le rôle fonctionnel de RsaF.

## **D. Caractéristiques structurales des ARN Rsa: découverte d'une signature conservée**

La taille des ARN Rsa varie entre 60 et 550 nucléotides mais tous possèdent des caractéristiques communes avec l'ARNIII. L'analyse de leur structure secondaire montre que tous ces ARN possèdent des structures en tiges-boucles à leurs extrémités 5' et 3'. Ceci confère une protection contre la dégradation par les exoribonucléases (i.e. RNase J1) et explique en partie pourquoi la plupart de ces ARN sont stables dans la cellule. Par ailleurs, à l'exception de RsaF et I, tous les ARN Rsa possèdent le même motif conservé riche en C (UCCC) localisé dans des régions en simple brin (figure 42). Ceci souligne pour la première fois l'existence d'une signature commune à plusieurs ARNnc. Comme pour l'ARNIII, ces motifs pourraient s'avérer décisif pour réprimer au niveau post-transcriptionnel de nombreux gènes. Ainsi, cette signature spécifique pourrait résulter d'une évolution convergente ayant abouti à l'émergence d'une classe d'ARN régulateurs qui agiraient via un mode d'action commun.



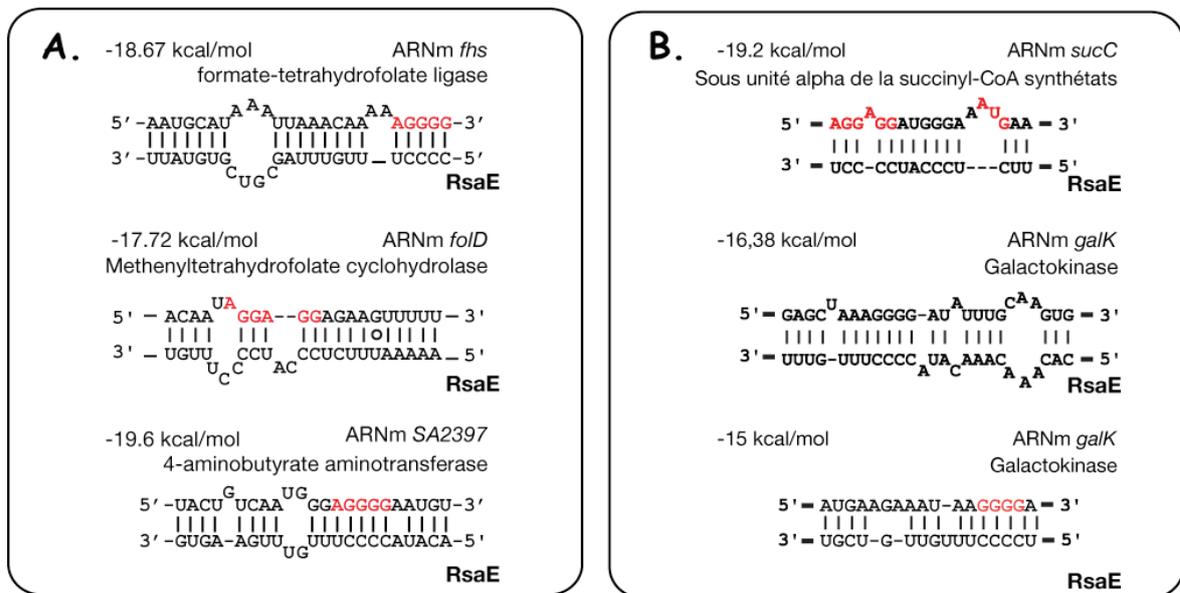
**Figure 42:** Les ARN RsaA, RsaE et RsaG portent le même motif riche en C. La structure secondaire des ARN RsaA, RsaE et RsaG a été obtenue par des expériences de cartographie en solution. Le motif conservé riche en C (en rouge) est présent dans des régions en simple brin des différents ARN. La séquence codante de l'ARNIII (hémolysine  $\delta$ ) est indiquée en bleu.

## E. Vers la fonction des ARNc

### E.1. RsaE et le métabolisme

La régulation de la transcription de *rsaE* est complexe puisqu'elle est régulée à la fois par  $\sigma^B$  et *agr* (cf § C précédent). Il a été récemment montré que le système *agr* de densité cellulaire est avant tout requis pour adapter la bactérie à l'entrée dans la phase stationnaire (Queck et al., 2008). Il se pourrait que RsaE soit relié au métabolisme de la bactérie et indirectement qu'il contribue à la virulence. La délétion du gène *rsaE* a permis de réaliser une approche comparative du protéome et du transcriptome. Ces données comparatives montrent que RsaE régule effectivement des gènes impliqués dans de nombreuses voies métaboliques. En combinant ces expériences avec la recherche des appariements stables RsaE-ARNm, nos résultats suggèrent fortement que RsaE régule directement la traduction des ARNm *sucC* et *sucD* qui codent pour des protéines impliquées dans le métabolisme des carbohydrates (cycle de Krebs) et de l'ARNm *fold*

requis pour la synthèse du folate. Par ailleurs, l'ARNm *oppB* peut aussi former un duplexe stable avec RsaE. L'opéron *opp* code pour un système d'importation de petits peptides. *In vitro*, des expériences de "toeprint" confirme que RsaE forme un complexe stable avec la région 5' régulatrice des ARNm *oppB* et *sucC* qui empêche la formation du complexe d'initiation de la traduction. Ainsi, RsaE pourrait agir de la même manière que l'ARNIII en ciblant la séquence SD des ARNm par son motif UCCC. Pour l'instant, seules quelques cibles directes ont été validées expérimentalement mais les effets multiples de RsaE révélés par les analyses comparatives du transcriptome suggèrent que cet ARN pourrait induire de nombreux effets indirects en régulant un facteur de transcription comme nous l'avons démontré pour l'ARNIII. Par ailleurs d'autres hybrides stables ont été prédits par la recherche *in silico* (figure 43, A).



**Figure 43:** (A) Hybrides formés entre RsaE et certains ARNm de *S. aureus* ou (B) de *B. subtilis*. Les bases en rouges indiquent les séquences Shine Dalgarno.

RsaE est d'autant plus intéressant qu'il est le seul à être conservé sur l'ensemble des *Bacillaceae*, hormis *L. monocytogenes*. Il est aussi exprimé chez *B. subtilis* ou la prédiction de cibles indique que les appariements avec l'ARNm *sucD* seraient conservés. On retrouve aussi l'ARNm codant pour *galK* qui code pour une protéine impliquée dans le choix des sources en carbone (figure 43, B). Il est intéressant de signaler que dans *E. coli*, l'ARN Spot 42 régule la traduction de ces mêmes cibles (Gottesman, 2004; Moller et al., 2002). Ainsi, même si ces deux ARN sont très différents, le même type de mécanisme pourrait

être conservé entre bactéries Gram-négatif et Gram-positif pour adapter le choix des sources en carbone lors de la croissance cellulaire (Gorke et Vogel, 2008). L'étude fonctionnelle plus précise de l'ARN RsaE dans des bactéries aussi éloignées que *S. aureus* et *B. subtilis* devraient aider à identifier les cibles communes mais aussi les cibles spécifiques de chaque bactérie et ainsi obtenir une image évolutive des mécanismes de régulation. Pour cela, nous envisageons de réaliser la délétion de *rsaE* dans *B. subtilis*.

## E.2. Cibles potentielles des autres ARNnc

La fonction de la plupart des ARN Rsa reste encore à déterminer. Cependant, partant de l'idée que le motif riche en C serait approprié pour initier la fixation à divers ARNm, nous avons recherché *in silico* les complémentarités de séquences entre les ARNnc et les ARNm. Cette analyse peut donner quelques idées sur leur fonction (tableau 7).

ARNnc	cibles	Fonction du gène cible
<b>RsaA</b>	<i>glpF</i> Protéine membranaire	Acquisition du glycérol Adaptation
	<i>sa2093</i> (Homologue à SsaA) Antigène sécrété Peptidoglycane hydrolase	Virulence Remodelage de la paroi bactérienne
	<i>ampS</i> aminopeptidase	Virulence?
	<i>citM</i> Protéine de transport de diverses toxines	Virulence
	<i>sa0022</i> 5' nucléotidase	Virulence?
	<i>sa2226</i> Système de transport APC pour acides aminés	Adaptation Virulence?
<b>RsaG</b>	<i>sigB</i> Facteur de transcription $\sigma^B$	Adaptation, virulence
	<i>saeR</i> Système à deux composants	Virulence
	<i>sa0215</i> système à deux composants	Adaptation? Virulence?
<b>RsaH</b>	<i>geh</i> glycérol ester hydrolase	Virulence?
	<i>sa2093</i> (Homologue à SsaA) Antigène sécrété Peptidoglycane hydrolase	Virulence Remodelage de la paroi bactérienne

**Tableau 7:** Quelques exemples d'ARNm cibles prédits pour les ARN RsaA, RsaG et RsaH.

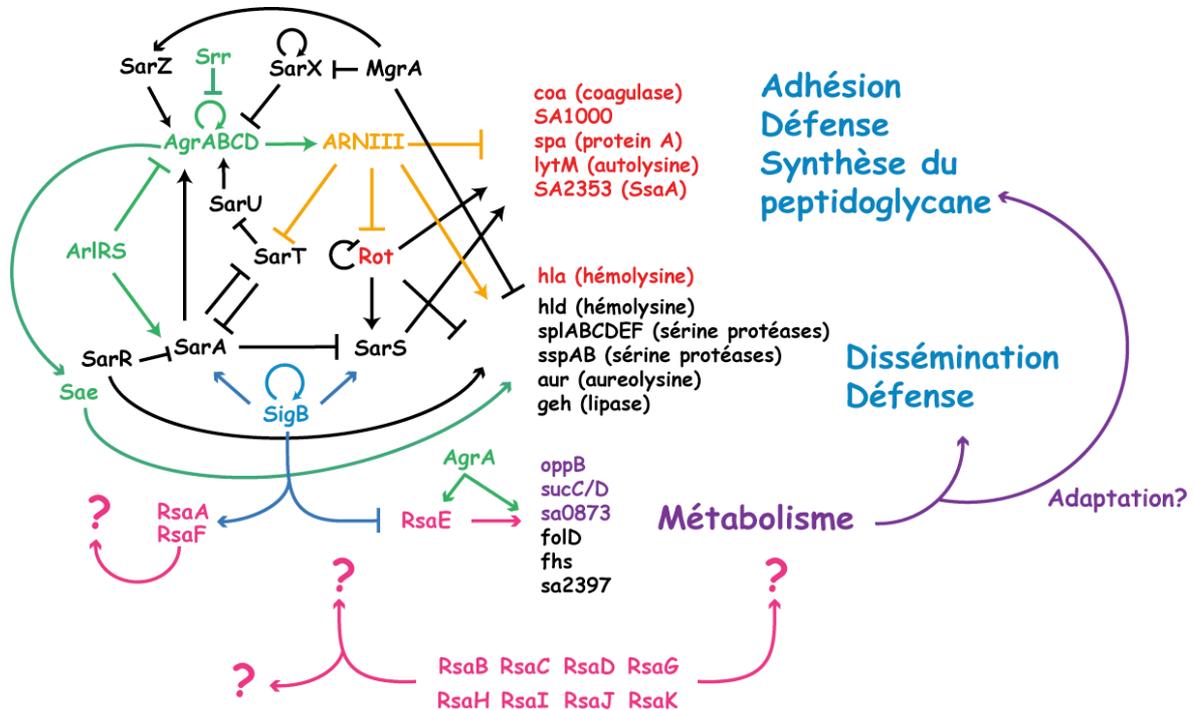
L'ARN RsaA est induit par le facteur de  $\sigma^B$ . L'analyse du transcriptome de souches exprimant ou pas le facteur  $\sigma^B$  actif, a révélé de nombreux gènes activés et réprimés par  $\sigma^B$ . La répression de ces gènes doit être indirecte, et on peut imaginer qu'elles résulteraient de l'expression de RsaA. Ainsi nous avons regardé si les ARNm réprimés par  $\sigma^B$  pouvaient former des complexes stables avec RsaA. C'est le cas de l'ARNm *citM* qui code pour une protéine de transport du citrate et du magnésium, et de plusieurs ARNm codant pour diverses toxines. Trois autres cibles potentielles codent pour deux aminopeptidases (SA2093 et AmpS) et une protéine membranaire. L'ARN RsaG possède plusieurs copies du motif UCCC et peut potentiellement former des complexes stables avec les ARNm *sigB* et *saeR* qui codent respectivement pour le facteur  $\sigma^B$  et pour le système à deux composants SaeR. Ces deux gènes sont largement impliqués dans la virulence de la bactérie et font partie des réseaux de régulation de l'ARNIII (cf introduction **IV.C**). Ainsi en régulant des facteurs de régulation, RsaG devrait induire des effets pléiotropiques. Enfin, RsaH peut former un hybride stable avec l'ARNm *geh* qui code pour la glycérol ester hydrolase, une enzyme impliquée dans le remodelage de la paroi bactérienne. Une première analyse comparative du protéome a montré que dans la souche délétée du gène *rsaH*, la protéine Geh est présente en plus forte quantité. Par ailleurs, tout comme RsaA et l'ARNIII, RsaH peut également former un duplexe stable avec l'ARNm *sa2093*. Comme la plupart des ARN Rsa possède le motif UCCC, on ne peut pas exclure que certains d'entre eux agissent de manière concertée sur la régulation de gènes identiques. En effet, la redondance fonctionnelle est souvent utilisée par de nombreuses bactéries pathogènes (Svenningsen et al., 2009; Svenningsen et al., 2008). Ces données prédictives pourront être validées puisque nous disposons depuis peu des souches mutantes pour RsaH, RsaA, et RsaH.

## F. Conclusions

### Vers la complication des réseaux de régulation

Notre étude confirme que l'ARN doit remplir des rôles variés dans *S. aureus*. Nos résultats montrent que RsaE est un régulateur post-transcriptionnel global du métabolisme dont l'expression est contrôlée par les facteurs  $\sigma^B$  et AgrA. Ainsi, RsaE pourrait faire partie des réseaux de régulation de l'ARNIII. Il pourrait non seulement participer à l'adaptation aux stress causés par le milieu extérieur mais également à la virulence en adaptant le

métabolisme de la bactérie lors de l'infection d'un hôte (figure 44). En effet, il est clairement établi que le métabolisme et l'acquisition des nutriments ont un impact sur la synthèse des facteurs de virulence de *S. aureus* (Somerville et Proctor, 2009). La fonction des autres ARN Rsa reste encore à déterminer. Cependant, leur mode d'expression nous donne quelques indications quant à leur rôle dans la cellule. Ainsi plusieurs des ARN Rsa (RsaA, RsaF, RsaD) font partie intégrante du régulon  $\sigma^B$  et certains (RsaB, C, G, H, J et K) sont uniquement présents dans *S. aureus*. Ces ARN résultent probablement d'une évolution récente et pourraient être liés à la pathogénie de la bactérie. La recherche de cibles potentielles des ARNnc indique que RsaA, RsaG et RsaH réguleraient la synthèse de facteurs de virulence ou de protéines régulatrices (tableau 7). Enfin, les ARN Rsa possèdent le motif UCCC conservé et des caractéristiques structurales communes, il est donc probable que certains d'entre eux agissent au niveau post-transcriptionnel d'une manière similaire à l'ARNIII ou à RsaE. En plus des systèmes à deux composants et des facteurs de régulation de la transcription, l'apport de ces nouveaux ARNnc favoriserait ainsi une régulation très fine de l'expression des gènes.



**Figure 44: Réseaux de régulation dépendant des ARN.** En modulant l'expression de plusieurs gènes du métabolisme RsaE pourrait influencer l'expression des gènes de virulence. La fonction et la place des autres ARN Rsa dans ce réseau reste encore à déterminer. Les cibles directes de l'ARNIII sont indiquées en rouge et celles de RsaE en violet.

### **Le génome de *S. aureus* cache-t-il d'autres riborégulateurs**

Une étude parallèle à la notre mais basée sur une nouvelle approche bioinformatique a été récemment utilisée sur tous les génomes bactériens (Marchais et al., 2009). Cette méthode est basée sur une étude phylogénétique (présence des IGR au sein des espèces) et sur la coévolution des régions IGR par rapport aux ARNm. Ainsi, l'étude de l'expression de 24 IGR a révélé l'existence de 7 ARN chez *S. aureus* dont six correspondent à des régions 5' UTR et un s'est avéré être RsaI (Marchais et al., 2009). Une telle approche phylogénétique est puissante pour révéler les régions régulatrices agissant *in cis* comme les " riboswitch " qui sont conservées dans des bactéries très éloignées dans la phylogénie. Ce qui n'est pas le cas des ARNnc agissant *in trans* qui sont en général spécifiques à une espèce. Ceci pourrait indiquer que les éléments régulateurs des ARNm seraient à l'origine d'un mécanisme de régulation primitif alors que les ARNnc seraient apparus plus tardivement lors de l'adaptation de la bactérie à une nouvelle niche écologique. Ceci souligne également que la détection des IGR fonctionnelles est réellement dépendante des approches utilisées et suggère que de nombreux ARN restent à découvrir dans le génome de *S. aureus*. Par exemple, de longs ARNnc antisens qui recouvrent des opérons entiers ont récemment été identifiés chez *L. monocytogenes* (Toledo-Arana et al., 2009). Même si la fonction de ces ARN est encore inconnue, on ne peut pas exclure qu'ils soient également exprimés chez *S. aureus*. Les multiples régions répétées du génome posent aussi de nombreuses questions comme les éléments STAR qui sont spécifiques de *S. aureus*. A quoi ces éléments correspondent-ils et sont-ils transcrits? Les éléments CRISPR semblent eux fonctionner différemment selon les bactéries. Chez *S. epidermis* ces éléments sont associés à 3 protéines CAS contre 5 en moyenne chez *E. coli*. Ceci sous entend que la maturation et le fonctionnement des ARNcr varient selon les espèces.

Nucleic Acids Research Advance Access published September 28, 2009

*Nucleic Acids Research*, 2009, 1–19  
doi:10.1093/nar/gkp668

## A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation

Thomas Geissmann<sup>1</sup>, Clément Chevalier<sup>1</sup>, Marie-Josée Cros<sup>2</sup>, Sandrine Boisset<sup>3</sup>, Pierre Fechter<sup>1</sup>, Céline Noirot<sup>2</sup>, Jacques Schrenzel<sup>4</sup>, Patrice François<sup>4</sup>, François Vandenesch<sup>3</sup>, Christine Gaspin<sup>2</sup> and 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, <sup>2</sup>INRA, UR 875, F-31320 Castanet-Tolosan, <sup>3</sup>INSERM U851, Centre National de Référence des Staphylocoques, Université de Lyon I, F-69008 Lyon, France and <sup>4</sup>Genomic Research Laboratory, Service of Infectious Diseases, University of Geneva Hospitals, Rue Micheli-du-Crest, 24, CH-1211 Geneva 14, Switzerland

Received May 22, 2009; Revised July 24, 2009; Accepted July 29, 2009

### ABSTRACT

Bioinformatic analysis of the intergenic regions of *Staphylococcus aureus* predicted multiple regulatory regions. From this analysis, we characterized 11 novel noncoding RNAs (RsaA–K) that are expressed in several *S. aureus* strains under different experimental conditions. Many of them accumulate in the late-exponential phase of growth. All ncRNAs are stable and their expression is Hfq-independent. The transcription of several of them is regulated by the alternative sigma B factor (RsaA, D and F) while the expression of RsaE is *agrA*-dependent. Six of these ncRNAs are specific to *S. aureus*, four are conserved in other *Staphylococci*, and RsaE is also present in *Bacillaceae*. Transcriptomic and proteomic analysis indicated that RsaE regulates the synthesis of proteins involved in various metabolic pathways. Phylogenetic analysis combined with RNA structure probing, searches for RsaE–mRNA base pairing, and toeprinting assays indicate that a conserved and unpaired UCCC sequence motif of RsaE binds to target mRNAs and prevents the formation of the ribosomal initiation complex. This study unexpectedly shows that most of the novel ncRNAs carry the conserved C-rich motif, suggesting that they are

members of a class of ncRNAs that target mRNAs by a shared mechanism.

### INTRODUCTION

RNAs are recognized as major regulators of gene expression, and many of them affect the fate of mRNAs at different levels. In bacteria, systematic searches for functional intergenic regions using a combination of methods ranging from *in silico* prediction to genome-wide expression studies have led to the discovery of more than 150 small *trans*-acting non-coding RNAs [generally referred to as small RNAs (sRNAs)]. Most of them have been identified in *Escherichia coli* and closely related bacteria (1,2). These sRNAs participate in regulatory pathways that allow the bacteria to sense the population density, to modulate their cell surface composition in response to various stresses, and to adjust their metabolism during cell growth. It was thus not surprising to discover that RNA-dependent regulation also enables several pathogenic bacteria to express virulence genes and to adapt their metabolic needs during the infection process (3,4).

*Staphylococcus aureus* is a Gram-positive pathogen responsible for a wide-range of human diseases, from minor skin infections to life-threatening diseases such as endocarditis, pneumonia, or septic shock. It is also one of the leading causes of hospital-acquired infections, often causing post-surgical wound infections. Moreover, the emergence of multiple antibiotic resistances renders the

\*To whom correspondence should be addressed. Tel: 33 3 88417051; Fax: 33 3 88602218; Email: p.romby@ibmc.u-strasbg.fr

Correspondence may also be addressed to Christine Gaspin. Email: gaspin@toulouse.inra.fr

Correspondence may also be addressed to François Vandenesch. Email: denesch@univ-lyon1.fr

Present address:

Thomas Geissmann, INSERM U851, Centre National de Référence des Staphylocoques, Université de Lyon I, F-69008 Lyon, France.

© The Author(s) 2009. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/2.5/uk/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

2 *Nucleic Acids Research*, 2009

treatment of the infections difficult (5). The high diversity of disease manifestations caused by *S. aureus* depends on the expression of numerous virulence factors and stress-response pathways. This requires the coordinated action of multiple regulators including several two-component systems, transcriptional regulatory proteins, the alternative sigma B factor and the regulatory RNAIII (6,7). Remarkably, regulation of the genes encoding virulence factors occurs at multiple levels and in a temporal manner, suggesting that the amount and timing of production of the virulence factors need to be precisely controlled during the course of infection. The complexity of these regulatory networks likely benefits the pathogen by allowing it to integrate multiple external signals. Although many studies have focused on regulatory proteins, few studies have addressed the importance of RNA-dependent regulation in *S. aureus* (8–12). Only RNAIII, which is so far the largest known regulatory RNA, has been shown to control the switch between the expression of many surface proteins and that of excreted toxins (13). RNAIII plays a key role in the quorum sensing-dependent regulatory circuit and coordinately regulates the synthesis of virulence-associated genes and of a transcriptional regulator at the post-transcriptional level through multiple RNAIII-mRNA interactions (6,9,14,15). *Staphylococcus aureus* RNAIII belongs to a major class of regulatory RNAs that acts on mRNAs by an antisense mechanism (1,16–18). In *Enterobacteriaceae*, many sRNAs require the Sm-like protein Hfq to facilitate mRNA-sRNA interaction and to enhance sRNA stability (19). However, *S. aureus* Hfq is required neither for virulence nor for RNAIII function (10,20) and it is characterized by different RNA-binding properties from that of *E. coli* Hfq (21). Several stable RNAs have been identified in infectious *S. aureus* strains (8,12,22). Based on their sequence conservation and GC-rich content, seven sRNAs were found to be encoded by pathogenicity islands, although their functions remain unknown (8). Their localization on mobile genetic elements, which contain clusters of virulence factors and genes encoding antibiotic resistance, suggests that they may contribute to niche adaptation and pathogenicity. Numerous stable RNAs, identified by microarrays, were differently expressed under stringent, cold and heat shock responses, but many of them may be co-transcribed as part of an operon (11,12). More recently, a large-scale computational work predicted numerous ncRNA genes in 932 different bacteria, revealing the potential presence of 47 new ncRNAs in the N315 strain of *S. aureus* (23).

In the present work, we use computational tools and expression studies to identify sRNA genes in intergenic regions (IGR) of *S. aureus* and related Gram-positive bacteria. Eleven new sRNAs are described and characterized with respect to transcriptional regulation, Hfq-dependence, stability, their target mRNAs and their likely biological functions. Most notably, one ncRNA regulates metabolism and uses a conserved and unpaired C-rich sequence that binds to the Shine-Dalgarno sequence of mRNA targets. This specific signature was found in many of the identified ncRNAs in *S. aureus*,

suggesting that they regulate gene expression by a shared mechanism.

**MATERIALS AND METHODS****Genomes, databases and known sRNA annotation**

Genome sequences and ORF files (.gbk, .fna and .ptt extensions) were downloaded from the NCBI ftp database for *S. saprophyticus* ATCC15305, *S. haemolyticus* JCSC1435, *S. aureus* N315 and *S. epidermidis* RP62A. Currently, there are 110 known annotated RNAs in the database for the *S. aureus* N315 genome. In this analysis, the coordinates of 78 tRNAs and rRNAs were downloaded from the NCBI. The coordinates of other 32 putative sRNAs and riboswitches were obtained from the RFAM (version 8.1) database. The coordinates of other known sRNAs including RNAIII, sprA, sprB, sprC, sprD, sprE, sprFG and other stable RNAs of *S. aureus* N315 were obtained from data available in the literature (8,11,12,22,24).

Variable regions resulting from a comparative analysis of available *S. aureus* strains were obtained from MOSAIC, a relational database developed to compare closely related bacterial genomes (25,26). Tandem repetitive elements were assigned using the Microorganisms Tandem Repeats Database [(27); <http://minisatellites.u-psud.fr/GPMS/>] for *S. aureus* N315. Additional sequence comparisons were performed with Blastn (default values) in order to complete the list of repeats. PatScan (28) allowed us to search for the STAR element (29) and we also used it to find  $\sigma^A$ - and  $\sigma^B$ -dependent promoters, Fur-boxes (30) and hairpins with lengths greater than 6 bp. Rho-independent transcription terminators were found with TransTermHP with default parameters [(31); <http://transterm.cbcb.umd.edu/>].

**In silico prediction of sRNA genes**

The possibility of a biased composition analysis in *S. aureus* N315 was considered since the G+C content was significantly higher in the set of the 110 annotated ncRNAs than in the rest of the genome (Gaspin *et al.*, unpublished data). This result lead us to use a two state Hidden Markov Model (HMM) to segment the genomic sequence into RNA and other regions. Parameters of the HMM were adapted from the analysis performed on the AT-rich hyperthermophilic Archaea *Methanococcus jannaschii* and *Pyrococcus furiosus* (32). From HMM predictions, only putative RNAs of lengths greater than 70 nt located in intergenic regions (not overlapping long ORFs) were kept for further analysis. A comparative analysis of the chromosomes of *S. saprophyticus* ATCC15305 (NC\_007350.1), *S. haemolyticus* JCSC1435 (NC\_0071168.1), *S. aureus* N315 (NC\_002745.2) and *S. epidermidis* RP62A (NC\_002976.3) was performed. A file containing fasta nucleotide sequences of intergenic regions (IGR) greater than 150 nt and ORFs with a size lower than 100 aa was built. For each chromosome, all the RNAs (tRNAs, tmRNA and rRNAs) annotated in Genbank were removed. All the IGR sequences were compared pairwise with the RNAsim software

(Supplementary Figures S1 and S2). RNAsim searches for conserved sequences and structural regions between different genomes. It is based on Wu-blast 2.0 pairwise comparisons (this work,  $W = 11$ ,  $E < 0.001$ ) of sequences for searching similarities, and it uses QRNA (33) to identify base substitution patterns in pairwise alignments that could correspond to a conserved RNA secondary structure. Finally, in the context of a multi-genome comparison, it combines overlapping predicted loci that are conserved in multiple replicons into a prediction. Only putative RNAs with lengths greater than 150 nt were selected to search for similarities against the NCBI complete bacterial genome database ( $W = 7$  and default parameters).

#### Integration of data and results of predictions

ApolloRNA (<http://www.carlit.toulouse.inra.fr/ApolloRNA>) is an extension of the Apollo environment [<http://apollo.berkeleybop.org/current/index.html>; (34)], which provides information on the orientation of the flanking genes and on all annotated genes. ApolloRNA was used to integrate the HMM and comparative analysis predictions (Supplementary Figure S1). We also assigned riboswitch elements, known ncRNAs (8,11), STAR (30), other repeats (29) and mobile elements (transposons, pathogenicity islands). All these data are reported in Supplementary Tables S1 and S2. ApolloRNA also contains some extensions to Apollo that are devoted to RNA analysis. It includes the integration of RNAfold of the Vienna package to predict the RNA secondary structure of a selected sequence along the genome (35). In addition, ApolloRNA integrates tools to interactively search for possible RNA/RNA interactions, and to visualize the graphs of nucleotide composition and minimum free energy associated with a sliding window all along a genome. In ApolloRNA, all predicted RNAs were analyzed with regard to available data and orientation of the flanking genes. Several were selected for further expression studies (Table 1, Supplementary Figure S1).

#### Strains, vector and plasmids

*Staphylococcus aureus* RN6390 derives from the NCTC8325-4 strain (Table 2). In WA400 ( $\Delta rnaIII$ ), the P3 operon is deleted and replaced by the chloramphenicol transacetylase gene (*cat86*) (36). In the LUG774 and LUG911 strains, the RNase III (*rnc*) and *hfq* genes, respectively, have been replaced by the *cat86* gene (9,14) (Table 2). In the LUG1430 strain, the *rsaE* gene has been replaced by *aphA-3*, which confers kanamycin resistance. These strains are derived from RN6390. For the deletion of the *rsaE* gene, the deletion/replacement  $\Delta rsaE/aphA-3$  mutant of *S. aureus* RN6390 was obtained by using pMAD, a thermosensitive plasmid (37). Clones were transferred by electroporation into RN4220, a nitrosoguanidine-induced mutant capable of accepting *E. coli* DNA, before transfer to other strains. Staphylococci were grown on BM agar plates, in brain-heart infusion (BHI) with kanamycin (5  $\mu$ g/ml) when appropriate, or in NZM medium (10 g/l Casein Tryptone digested, Fluka 95039;

5 g/l NaCl; 2 g/l  $MgSO_4 \cdot 7H_2O$ ; pH 7.0). The allelic replacement of *rsaE* was verified by PCR.

*Staphylococcus aureus* COL strains were submitted to various stress conditions. The cells were grown to mid-exponential phase (OD 600 nm 0.5) at 37°C and stress was induced by the addition of 15% NaCl (osmotic stress), 5 mM hydrogen peroxide (oxidative stress), 0.1 mM paraquat (oxidative stress), 0.25 mM dipyrindyl (an iron chelating agent), or HCl to decrease the pH of the medium to 4.5. The cells were submitted to either cold shock (25°C) or heat shock (42°C). After 30 min, the cells were centrifuged and the total RNA was extracted as described previously (9). Total RNAs were prepared from different *S. aureus* or *B. subtilis* strains at early (2 h) and late (6 h) exponential phases of growth, or from *S. aureus* COL strain submitted to stress conditions.

#### Northern analysis, RACE and primer extension analysis

Electrophoresis of total RNAs was done on 5% polyacrylamide gels containing 8 M urea or 1% agarose gels containing 20 mM guanidine thiocyanate (Sigma, G9277). Hybridizations were carried out with end-labeled specific oligodeoxynucleotides to detect mRNA, Rsa RNAs, RNAlII or 5S rRNA (14).

Primer extension was carried out on 2  $\mu$ g of total RNA prepared from *S. aureus* or *B. subtilis* strains. Total RNA was first denatured for 1 min at 90°C followed by 1 min on ice in 15  $\mu$ l of RNase-free water in the presence of 2  $\mu$ l of 5'-end labeled oligodeoxynucleotide complementary to a specific Rsa RNA (Supplementary Table S4). The hybridization was done at 20°C for 15 min in the presence of 3  $\mu$ l of the commercial AMV 10 $\times$  buffer (Qbiogen) and 6  $\mu$ l of dNTP 2.5 mM. The extension was done with 1  $\mu$ l of AMV reverse transcriptase (20 U/ $\mu$ l, Qbiogen) for 30 min at 45°C. The RNA was then hydrolyzed by RNase A (1 U) for 10 min at 37°C. After the addition of 4  $\mu$ l of formamide-bromophenol blue, the samples were loaded onto an 8% polyacrylamide gel containing 8 M urea and 1 $\times$  TBE. The 5' start and the size of the RNA were evaluated by running sequencing ladders in parallel.

The 5'-end of the Rsa RNAs was also determined by rapid amplification of cDNA ends (RACE) using the First Choice RLM-RACE kit (Ambion). To assign the 5'- and 3'-ends of the RNAs, the RACE approach was done on RNAs previously circularized using T4 RNA ligase according to Redko *et al.* (38).

#### Microarray manufacturing and design

The microarray was manufactured by *in situ* synthesis of 10 807 long oligonucleotide probes (Agilent, Palo Alto, CA, USA), selected as previously described (39). It covers >98% of all ORFs annotated in strains N315 and Mu50 (40), MW2 (41), COL NCTC8325 (42), USA300 (43), MRSA252 and MSSA476 (44) including their respective plasmids.

#### Preparation of labeled nucleic acids for expression microarrays

Total RNAs were purified from late-exponential (6 h at 37°C) phase cultures grown in BHI. After additional

4 *Nucleic Acids Research*, 2009

DNase treatment, the absence of remaining DNA traces was evaluated by quantitative PCR (SDS 7700; Applied Biosystems, Framingham, MA) with assays specific for 16S rRNA (45,46). Batches of 8 µg total *S. aureus* RNA were labeled by Cy-5dCTP (RN6390 strain) or Cy-3dCTP (LUG1430 strain) using SuperScript II (Invitrogen, Basel, Switzerland) following the manufacturer's instructions. Labeled products were then purified onto Qiaquick columns (Qiagen).

The Cy5-labeled DNA and the Cy3-labeled cDNA mixture (1 µg) were diluted in Agilent hybridization buffer (final volume of 50 µl) and hybridized at 60°C for 17 h in a dedicated hybridization oven (Robbins Scientific, Sunnyvale, CA, USA). Slides were washed, dried under nitrogen flow and scanned with an Agilent scanner (Agilent, Palo Alto, CA, USA) using 100% Photon Multiplier Tube (PMT) power for both wavelengths.

Fluorescence intensities were extracted using the feature extraction software (Agilent, version 9.3). Local background-subtracted signals were corrected for unequal dye incorporation or unequal loads of the labeled product and normalization per gene and per slide was performed using GeneSpring. Data consisting of two independent biological experiments were expressed as log<sub>10</sub> ratios and analyzed using GeneSpring (Agilent). The statistical significance of differentially expressed genes was identified by variance analysis (ANOVA,  $P < 0.05$ ) (45,47), performed using GeneSpring, including the Benjamini and Hochberg false discovery rate correction (5%).

**Microarray data accession number**

The complete microarray data set has been posted on the Gene Expression Omnibus database (<http://www.ncbi.nlm.nih.gov/geo/>) under accession number GPL7137 for the platform design and GSE17135 for the original data set.

**Proteomic analysis**

Strains were grown at 37°C to exponential (2 h) and post-exponential (6 h) phases by inoculating 50 ml of BHI medium with an overnight culture (1:100). The supernatants (1.2 ml) were precipitated by cold acetone, and the pellets were washed with cold ethanol. Crude extracts from the pellets were obtained by Trizol (Sigma) extraction, followed by acetone precipitation. The proteins were dissolved in sample buffer (7 M urea, 2 M thiourea, 4% CHAPS, 20 mM Tris-HCl pH 8) and 50 µg were labeled with 400 pmol of amine-reactive-Cy2, Cy3 or Cy5 (Amersham) for 30 min in the dark at 4°C. The reaction was stopped by the addition of 20 µl of 10 mM lysine for 10 min. Alkylation then proceeded at 37°C for 15 min in the presence of DTT (up to 50 mM) and 10 µl of 1 M iodoacetamide. The protein extracts were desalted by a G-25 Sephadex chromatography. Ampholytes (Bio-Rad) and bromophenol blue were added, and the samples were used to re-swell pre-cast pH 3–10 or 4–7 IPG strips (Bio-Rad). Isoelectric focusing was performed using an IEF Cell (Bio-Rad) at 75 kV/h. The IPG strips were buffer-exchanged in 1 M Tris-HCl pH 8.45, 6 M urea, 2%

SDS, 30% glycerol. For the second dimension, the strips were loaded onto the top of a 12% polyacrylamide–2% SDS gel (48). After electrophoresis, the gels were scanned for fluorescence detection, stained with colloidal blue and analyzed using the PQuest software (Bio-Rad). The differentially expressed proteins were characterized by mass spectroscopy.

**RNA preparation and band shift analysis**

Rsa RNAs (RsaA, E, G and H) and mRNA (*oppB*, SA0873) fragments were transcribed *in vitro* using T7 RNA polymerase as previously described (24). Before use, the RNAs were renatured by incubation at 90°C for 2 min and 1 min at 4°C in RNase-free water, followed by an incubation step at 20°C for 15 min in TMN buffer (20 mM Tris-HCl pH 7.5, 10 mM MgCl<sub>2</sub>, 150 mM NaCl).

Complex formation was done using 5'-end-labeled RsaE (5 nM) and an increasing amount of unlabeled target mRNA (5–400 nM) for 15 min at 37°C in TMN buffer in a total volume of 10 µl. The samples were then separated on a native 6% polyacrylamide gel in 1× Tris-borate (0.09 M) buffer. Quantification of the complex formation was done using a Bioimager analyzer.

**Structure probing and toeprinting**

Enzymatic hydrolysis of free or bound RNAs was performed as previously described (14). The formation of a simplified translational initiation complex with mRNA and the extension inhibition conditions were strictly identical to those described by Benito *et al.* (24).

**RESULTS****Computational selection of IGRs using multiple approaches**

Based on functional studies performed on *S. aureus* RNAIII (9,14), we made the assumption that most sRNAs may carry specific structural motifs required for efficient target recognition. A combination of different tools was then used to retain IGRs that present particular sequence/structure features (Supplementary Figure S1). First, the degree of conservation of the IGRs in different microbial genomes was analyzed. Regions conserved either in various firmicutes (*Staphylococcus*, *Bacillus*, *Listeria* and *Clostridium*), or in four different staphylococcal species (*S. aureus* N315, *S. epidermidis*, *S. saprophyticus* and *S. haemolyticus*) were analyzed using RNAsim (Supplementary Figure S2 and Table S1). Because unique genes could also contribute to pathogenesis, we searched *S. aureus*-specific regions (Supplementary Table S1) using data downloaded from MOSAIC, a database developed for comparison of closely related bacterial genomes (25). MOSAIC identified identical sequences constituting the backbone interrupted by strain-specific sequences (loop regions). Finally, in all the conserved and *S. aureus*-specific IGRs, we searched for orphan Rho-independent terminators located far downstream of the flanking coding genes (49).

**Table 1.** List of selected intergenic regions of *S. aureus* for expression studies

Rsa <sup>a</sup>	Start-end (size)	Start <sup>b</sup>	Flanking genes	Strand	Selection criteria <sup>c</sup>	Comments <sup>d</sup>
RsaA	637111–637250 (139)	RT, 5'RACE	SA0543/SA0544	<<<	RNAsim, ΔG, Term	E<<<S (BHI, NZM), Staphylococci
RsaB	1778009–1778068 (60)	RT, 5'3'RACE	SA1552/ <i>fhs</i>	<<<	RNAsim	E<<<S (BHI), <i>S. aureus</i>
RsaC	679909–680452 (544)	5'3'RACE	SA0586/SA0587	<<>	ΔG, RNAsim, Term	E<<<S (BHI, NZM), <i>S. aureus</i>
RsaD	695867–696043 (177)	RT	SA0600/SA0601	>>>	RNAsim, Term	E<<<S (BHI, NZM), Staphylococci
RsaE	975383–975482 (100)	RT, 5'RACE	SA0859/SA0860	>>>	RNAsim, Term	E<<<S (BHI, NZM), <i>Bacillaceae</i>
RsaF	975461–975564 (104)	5'RACE	SA0859/SA0860	>>>	ΔG, RNAsim	E<<<S (BHI, NZM), <i>S. aureus</i>
RsaG	254702–254895 (194)	RT, 5'RACE	<i>uhpT</i> /SA0215	>><	HMM, ΔG, Term	S (BHD), <i>S. aureus</i>
RsaH	829511–829634 (124)	RT, 5'3'RACE	SA0724/SA0725	<<>	ΔG, RNAsim, Term	E<<<S (BHI), Staphylococci
RsaI	2367918–2368061 (144)	RT, 5'RACE	SA2104/SA2105	<<>	RNAsim, Term	E<<<S (BHI), Staphylococci
RsaJ	2486806–2487092 (287)	nd	<i>bioD</i> /SA2216	<<<	RNAsim, ΔG, Term	E<<<S (BHI), <i>S. aureus</i>
RsaK	216920–217128 (209)	RT	<i>glcA</i> /SA0184	>><	RNAsim, Term	<i>Cis</i> -acting leader of <i>glcA</i>
RsaX01	43000–43600		SA0035/SA0036		HMM, RNAsim	–
RsaX02	81706–81855		<i>kdpC</i> /SA0072		RNAsim, Term	(23)
RsaX03	95295–95689		SA0084–SA0085		RNAsim, Term	(22)
RsaX04	149756–149850		SA0129–SA0130		RNAsim, Term	–
RsaX05	470600–470900		SA0410/ <i>ndhF</i>		Prom, Term	<i>Cis</i> -acting leader of <i>ndhF</i>
RsaX06	659823–659968		SA0565/SA0566		RNAsim	–
RsaX07	666122–666257		SA0575/ <i>sarA</i>		RNAsim, Term	–
RsaX08	761087–761230		SA0667/SA0668		RNAsim, Term	PreQ1 riboswitch, (22)
RsaX09	905650–905835		SA0801/SA0802		Term	<i>Cis</i> -acting leader of SA0802?
RsaX10	1138640–1138848		SA1005/SA1006		RNAsim, Term	–
RsaX11	1175489–1175753		SA1037/ <i>lsp</i>		HMM, RNAsim, Term	–
RsaX12	1399044–1399209		SA1224/ <i>lysC</i>		RNAsim, HMM	Upstream of lysine riboswitch
RsaX13	1421983–1422128		<i>odhA</i> / <i>arlS</i>		Term	UCCC motif
RsaX14	1801062–1801380		SA1570/SA1571		RNAsim, Term	–
RsaX15	1839300–1839600		SA1602/ <i>trunc.SA</i>		HMM, RNAsim, Term	–
RsaX16	1914236–1914464		SA1676/ <i>tnp</i>		RNAsim, Term	–
RsaX17	1992090–1992414		SA1738/SA056		HMM, RNAsim, Term	(11)
RsaX18	2436200–2436800		SA2168–SA2169		RNAsim, ΔG	(11)
RsaX19	2440668–2440850		<i>scrA</i> / <i>gltT</i>		HMM, RNAsim, Term	Repeat motifs, STAR
RsaX20	2473450–2473566		SA2203/SA2204		RNAsim, ΔG, Term	UCCC motif, sORF upstream
RsaX21	2544796–2545113		SA2268/SA2269		Term	–
RsaX22	2586551–2586821		SA2301/SA2302		RNAsim	–
RsaX23	2590928–2591046		<i>fbp</i> /SA2305		RNAsim, Term	sprA3 (8)
RsaX24	2649194–2649400		<i>isaA</i> /SA2357		RNAsim, Term	–
RsaX25	2769241–2769725		SA2457/ <i>icaR</i>		HMM, RNAsim, Term	UCCCC motif

<sup>a</sup>The Rsa sRNAs have been referred in GenBank under the accession numbers: GQ403615 (RsaA), GQ403616 (RsaB), GQ403617 (RsaC), GQ403618 (RsaD), GQ403619 (RsaE), GQ403620 (RsaF), GQ403621 (RsaG), GQ403622 (RsaH), GQ403623 (RsaI), GQ403624 (RsaJ) and GQ403625 (RsaK). RsaD, E, H and I were predicted by Livny *et al.* (23) and RsaI was also found expressed in the *S. aureus* N315 strain (22). In RN6390, no significant expression was detected for RsaX01 to RsaX26. sORF is for small open reading frame.

<sup>b</sup>The 5' start of Rsa RNAs was determined by primer extension (RT), and Rapid Amplification of cDNA Ends (RACE); 5'3'RACE was done on circularized RNA; nd is for not determined.

<sup>c</sup>Term is for Rho-independent terminator.

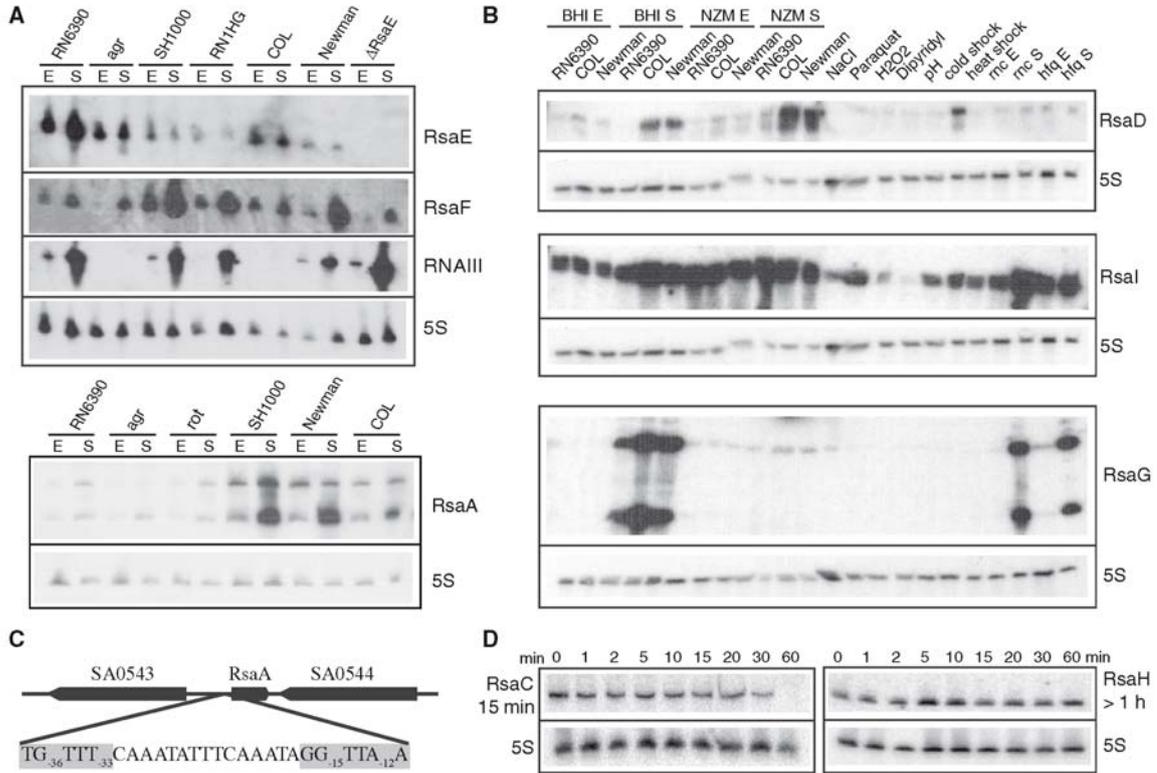
<sup>d</sup>The expression of RsaA to RsaK was analyzed at the exponential (E) and the stationary (S) phase of growth in various *S. aureus* strains.

As a second approach, a Hidden Markov Model (HMM) derived from Klein *et al.* (32) was used to exploit the compositional bias between known RNAs (RFAM, GenBank) and the rest of the genome (Supplementary Figure S1). The HMM model was able to distinguish ncRNAs including tRNAs and rRNAs, from mRNAs on the basis of local variations in genomic base composition of AT rich genomes (8,32,50). Although this approach failed to find RNAIII, it allowed the detection of known ncRNAs such as 6S RNA, SRP, tmRNA and sprA RNAs as well as the identification of 28 candidate sRNAs. In addition, this method identified many repeated sequences and structural motifs known as *Staphylococcus aureus* repeats (STAR) elements (29) (Supplementary Table S2).

The obtained information was assigned to the annotated sequence of the *S. aureus* N315 genome using ApolloRNA (<http://carlit.toulouse.inra.fr/ApolloRNA/>)

(51). Using this environment, we have selected 36 IGRs based on at least two different criteria (conserved and *S. aureus*-specific IGR, orphan terminators, promoters, structural motifs and the orientation and nature of the flanking genes), and the expression patterns of these putative transcripts were analyzed (Table 1, Supplementary Figure S1). Home-made arrays were designed using the 36 PCR-amplified IGR-specific dsDNAs which were spotted on nitrocellulose membranes (for experimental details, see Supplementary Data). As positive control, we used dsDNA corresponding to the regulatory RNAIII. Total RNAs were prepared from strain RN6390 grown in BHI (rich) and NZM (stringent) medium and in exponential and stationary phases. Total RNAs were fractionated by a denaturing polyacrylamide gel electrophoresis to remove ribosomal RNAs. The RNAs with a size of 50–600 nt were eluted, labeled at their 3'-end and hybridized to the membrane under stringent conditions

6 *Nucleic Acids Research*, 2009



**Figure 1.** The expression of Rsa RNAs as monitored under various stresses, culture conditions, and in various *S. aureus* strains. (A) Northern blot analysis of RsaE, RsaF and RsaA in different *S. aureus* strains: RN6390 ( $\Delta$ rbsU,  $\sigma^B$ ), RN6911 ( $\Delta$ agr,  $\sigma^B$ ), SH1000, RN1HG, COL and Newman ( $\sigma^B$ ). The strains LUG1430 ( $\Delta$ rnaE) and LUG1160 ( $\Delta$ rot) are isogenic to RN6390. Total RNAs were prepared from cells grown in BHI medium and stopped at the exponential (E) or stationary (S) phases of growth. *Staphylococcus aureus* RNAIII was used as positive control, and 5S rRNA was probed as a quality RNA control. (B) Northern blot analysis of RsaD, RsaI and RsaG in various *S. aureus* strains (RN6390, COL, Newman, LUG774 (RN6390- $\Delta$ rnc) and LUG911 (RN6390- $\Delta$ hfg)) grown in BHI or NZM at exponential (E) and stationary (S) phases of growth. The COL strain was grown under various stress conditions: osmotic stress (NaCl), oxidative stress (H<sub>2</sub>O<sub>2</sub>, paraquat), iron chelating agent (dipyridyl), acidic pH, cold shock (25°C) and heat shock (42°C). (C) Genomic organization of the *rsaE* gene with the  $\sigma^B$ -binding site; the consensus sequence is highlighted. (D) Northern blot analysis of the half-lives of RsaC and RsaH. The cells were treated with rifampicin and total RNAs were extracted after 1, 2, 5, 10, 20, 30 and 60 min at 37°C in BHI. 5S rRNA was probed to quantify the yield of RNA in each lane.

(for experimental details, see Supplementary Data). Strong expression signals were obtained for 11 RNA candidates prepared from cultures grown in the stationary phase. The expression of these 11 RNAs was then analyzed by northern blots in various strains and under stress conditions (Figure 1, Supplementary Figure S3). In the rest of the article, these 11 RNAs are denoted Rsa (A–K), for RNA from *S. aureus* (Table 1). We used the nomenclature of the *S. aureus* N315 strain for convenience but all of these IGRs are conserved in all other *S. aureus* strains sequenced to date.

**Expression studies reveal tight transcriptional regulation of Rsa sRNAs**

The expression of the 11 Rsa (A–K) transcripts was further analyzed in several *S. aureus* reference strains (RN6390, COL and Newman). RN6390, a derivative of NCTC8325-4, has been frequently used for molecular

analysis (Table 2). However, this strain carries a mutation in *rsbU* which results in defective  $\sigma^B$  activity (52,53). *Staphylococcus aureus* COL was one of the first methicillin-resistant isolates identified (42). The *S. aureus* Newman strain was originally isolated from a human infection and has been extensively studied in animal models (54). Based on phylogenetic analysis, these three strains are closely related, while their pathogenicity islands show polymorphisms. The expression of the 11 Rsa RNAs was analyzed in cells grown to exponential and stationary phases. In addition, the COL strain was submitted to stress conditions that might be encountered by the bacteria during infection, such as osmotic and oxidative stress, cold and heat shock, iron limitation and acidity.

Most of the Rsa RNAs (RsaA, C, D, E, F, G, H and I) were highly expressed and showed distinct bands consistent with the length of the predicted IGR (Figure 1, Supplementary Figure S3, Table 1). No significant expression was found for the antisense strands of RsaA, C, H, E,

**Table 2.** Strains and plasmids

Name	Relevant characteristics	Reference or source
<i>Escherichia coli</i>	DH5 $\alpha$ , cloning strain	Laboratory stock
<i>Bacillus subtilis</i>	Wild type strain	ATCC6051
<i>S. aureus</i> strains		
RN4220	restriction- mutant of 8325-4	(105)
RN6390	derivative of 8325-4, <i>agr</i> positive	(106)
RN6911	RN6390: $\Delta$ <i>agr</i> region:: <i>tetM</i>	(107)
WA400	8325-4: $\Delta$ <i>rnaIII</i> region:: <i>cat86</i>	(36)
COL	Methicillin-resistant laboratory strain	(108)
Newman	Wild type strain	(109)
SH1000	<i>rsbU</i> + strain derivative of 8325-4	(56)
RN1HG	<i>rsbU</i> + strain derivative of RN1	(57)
LUG774	RN6390: $\Delta$ <i>rnc</i> region:: <i>cat86</i>	(14)
LUG911	RN6390: $\Delta$ <i>hfq</i> region:: <i>cat86</i>	(9)
LUG1160	RN6390: $\Delta$ <i>rot</i> region:: <i>cat86</i>	This work
LUG1397	RN4220/pLUG766	This work
LUG1408	RN6390/pLUG766	This work
LUG1430	RN6390: $\Delta$ <i>rsaE</i> :: <i>aphA-3</i>	This work
<i>E. coli</i> plasmids		
pUT7_RsaG	T7 promoter/RsaG	This work
pUC18_T7_RsaA	T7 promoter/RsaA	This work
pUC18_T7_RsaH	T7 promoter/RsaH	This work
pUC18_T7_RsaE	T7 promoter/RsaE	This work
pUC18_T7_oppB	T7 promoter/5'-region of <i>oppB</i> (nt -182 to +182, +1 being the A of AUG codon)	This work
pUC18_T7_SA0873	T7 promoter/5'-region of SA0873 (nt -22 to +128)	This work
<i>E. coli-staphylococcal</i> shuttle plasmids		
pMAD	Thermosensitive origin of replication, constitutively expressed $\beta$ <i>galB</i> gene	(37)
pLUG766	pMAD derivative for deletion/replacement of <i>S. aureus</i> <i>rsaE</i> gene	This work

or G (results not shown). The expression of many Rsa RNAs appeared to be tightly regulated, as observed for *E. coli* and *Salmonella typhimurium* sRNAs (1,4,55). In rich culture media, RsaB, D, G and H were predominantly expressed in the stationary phase while the expression of RsaA, C, E, F, I and J was observed in the exponential phase, but was strongly enhanced in the stationary phase (Figure 1, Supplementary Figure S3). When cultures were grown in stringent NZM medium, RsaE, F, G and H were not expressed, whereas RsaI expression was significantly enhanced in the exponential phase. Moreover, several Rsa RNAs responded to stress conditions but RsaF and RsaG did not. Indeed, RsaA, E and I were expressed under various stress conditions (osmotic and oxidative stress, acidic pH and temperature variations) while the expression of RsaC and D were significantly induced during cold shock (Figure 1B, Supplementary Figure S3).

RsaA, D, E, F and J were differently expressed in the three *S. aureus* strains. While RsaA, D and F bands were faint in RN6390, they showed strong expression in the COL and Newman strains (Figure 1A and B, Supplementary Figure S3). An opposite pattern of expression was also observed for RsaE and RsaF (Figure 1A).

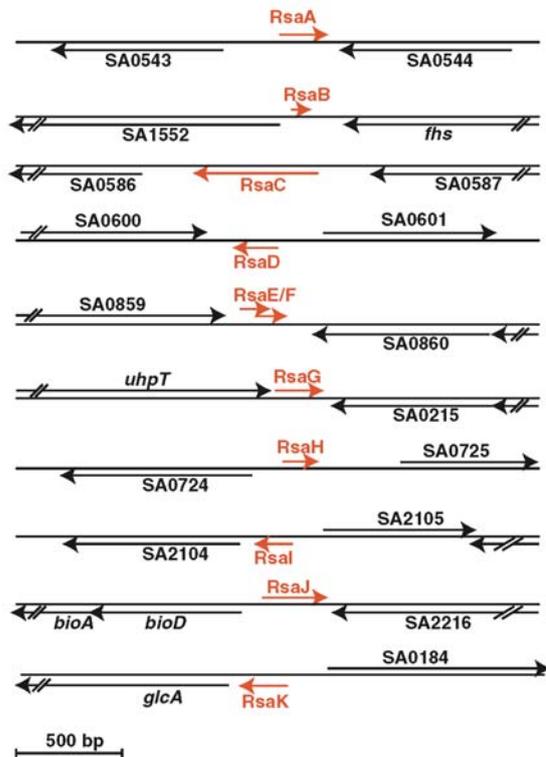
Since RN6390 has a defect in the alternative sigma B factor due to a mutation in the *rsbU* gene, we analyzed *rsaA*, *rsaE* and *rsaF* expression in the isogenic *rsbU*-proficient strain SH1000 (56) and in strain RN1HG (57). The results show that RsaA and RsaF are strongly enhanced in SH1000 and RN1HG, approaching the level observed in the COL and Newman strains (Figure 1A, Supplementary Figure S3). This suggests that transcription of these sRNAs is dependent on *rsbU* and/or  $\sigma^B$ . Since the transcription start sites of all sRNAs have been determined either by primer extension and/or RACE (Table 1), we looked for putative promoters directly upstream. We identified a conserved  $\sigma^B$  promoter sequence [TG<sub>-36</sub>TTT(N15)GG<sub>-15</sub>TTA<sub>-12</sub>A] upstream of the *rsaA* gene, suggesting a direct role of  $\sigma^B$  in *rsaA* transcription (Figure 1C). Moreover, the synthesis of RsaE was significantly repressed in  $\sigma^{B+}$  strains by an unknown mechanism. Interestingly, the partially overlapping RsaE and RsaF transcripts had opposite expression patterns (Figure 1A). RsaF was strongly enhanced in  $\sigma^{B+}$  strains, whereas RsaE was mainly detected in the  $\sigma^{B-}$  deficient strain RN6390. The overlap between a  $\sigma^A$  recognition motif [TTGAAA(N16)TATAT T] and a potential  $\sigma^B$  [AG<sub>-36</sub>GTT<sub>-33</sub>GAA(N15)TG<sub>-15</sub>TTA<sub>-12</sub>T]-binding site may account for transcriptional interference resulting in the differential expression of *rsaE* and *rsaF* (58). RsaE levels were also strongly decreased in an *agr* mutant strain in the late-exponential phase, suggesting that the synthesis of this RNA is regulated by the quorum sensing system (Figure 1A).

#### Primary and secondary features of the ncRNAs

The 5'- and 3'-ends of RsaB, C and H were determined by 5'/3'RACE after circularization of total RNAs by T4 RNA ligase (59). The 5'-ends of the other Rsa sRNAs were identified by 5' RACE (RsaA, E, F, G and I) and primer extension (RsaA, B, D, E, G, H, I and K; Table 1). The transcriptional start sites for most of the sRNAs were consistent with the size of the RNAs as evaluated by northern blots. The 5'-end-mapping of RsaJ failed, probably due to its rather poor expression. Determination of the 5' starts of the Rsa RNAs helped to predict consensus  $\sigma^A$ -dependent promoters for RsaB-E/I and a  $\sigma^B$ -dependent promoter for RsaA (Figure 1C). Most of the Rsa RNAs have at their 3'-ends a typical Rho-independent transcription terminator, except for RsaB and RsaF. The presence of a Rho-independent terminator 136 nt downstream of the 3'-end of RsaF suggests that this RNA might be processed from a longer RNA precursor, which was detectable by Northern blots (Supplementary Figure S3). Noteworthy, the sequences of the terminator structures are identical for RsaJ and SprA, an sRNA expressed from the pathogenicity island (8). The sizes of the sRNAs range from 59 (RsaB) to 544 nt (RsaC) (Table 1).

Their genomic locations indicate that most of the sRNAs are encoded by autonomous genes (Figure 2) and that none of the Rsa RNAs, including the longest one (RsaC, 544 nt), carry a small ORF as has been shown for RNIII. Four of them (RsaC, H, I and K)

8 *Nucleic Acids Research*, 2009



**Figure 2.** Genomic organization of *rsa* genes in the *S. aureus* N315 strain. Red arrows denote the *rsa* transcripts and their orientations, and the black arrows are for the flanking open reading frames (ORFs). The genes are represented to scale. The sizes of the *rsa* genes were determined based on the experimental determination of the 5' start by RACE mapping and primer extension, estimates from denaturing gel electrophoresis, and the presence of transcription terminator. For RsaJ, the transcription start site was not determined and the length is estimated. RsaB and RsaF do not contain transcription terminator at their 3'-ends. In Table 1, the exact genomic location of *rsa* genes is given.

are transcribed in the same orientation as their downstream genes, suggesting that they could arise from mRNA leaders, whereas the Rho-independent terminators of RsaC, H and I are located 300, 431, 50 nt upstream of their downstream genes, respectively (Figure 2). The distance and the absence of signals corresponding to longer transcripts suggest that these RNAs are *bona fide* sRNAs (Figure 1D, Supplementary Figure S3). RsaK, on the other hand, appears to be part of the leader sequence of the *glcA* mRNA, which encodes an enzyme of the glucose-specific phosphotransferase system (Supplementary Figure S4). RsaK contains a conserved ribonucleic antiterminator sequence (RAT) motif recognized by the transcriptional antiterminator protein GlcT [e.g. (60,61)]. RsaK was detected in several *S. aureus* strains together with a longer transcript corresponding to the full-length *glcA* mRNA (Supplementary Figure S4).

Rsa RNA sequences were compared across phylogeny using the RNAsim software (Supplementary Figure S2).

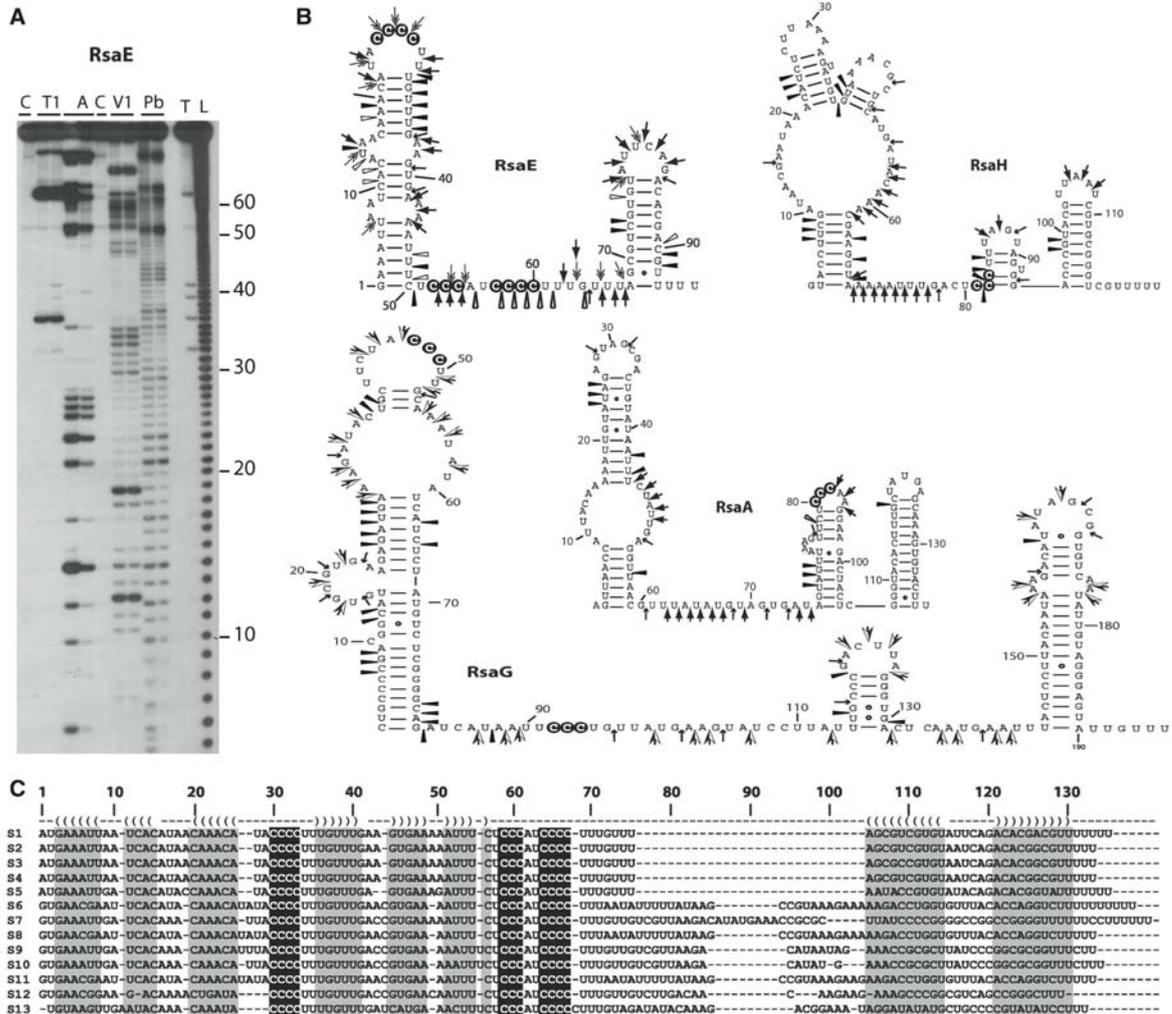
Their genes and genomic localizations are well-conserved in all sequenced *S. aureus* strains although expression of some of Rsa RNAs varies in COL and RN6390. Five *rsa* genes (*rsaB*, C, F, G and J) are specific to *S. aureus* whereas all the others (*rsaA*, D, E, H, I and K) are conserved in four different staphylococcal species (*S. aureus*, *S. epidermidis*, *S. saprophyticus* and *S. haemolyticus*). RsaE is unique since its conservation extends to *Macrococcus caseolyticus* and *Bacillus* while the partially overlapping RsaF and its upstream gene are only conserved in *S. aureus* species (Table 1, Supplementary Figure S5). Interestingly, the *rsaE* gene is always located downstream of the *pepF* gene, which encodes the conserved oligoendopeptidase F. These two genes, *pepF* and *rsaE*, are not always located in the same genomic environment.

The structures of four RNAs (RsaA, E, G and H) were probed by using single-strand-specific (RNases T1, T2 and A) or double-strand-specific (RNase V1) ribonucleases (Figure 3A). The enzymatic cleavage patterns obtained on these RNAs are well correlated with the existence of several hairpin structures and a well-defined terminator of transcription interconnected by long unpaired regions. Multiple sequence alignment was performed with PARADISE (62) which took into account the secondary structure of *S. aureus* Rsa RNAs in several staphylococcal species (RsaA, D, E, H and I; Figure 3C and Supplementary Figure S6). This analysis unravels conserved regions that are expected to be functional regulatory elements (4). Although the sequences of the helices vary, the overall secondary structures of RsaA, D, E, H and I remain conserved in all staphylococci (Figure 3C, Supplementary Figure S6). Strikingly, all of the Rsa RNAs, except RsaI and F, contain a C-rich sequence motif (UCCC) located in unpaired regions (Figure 3B, Supplementary Figure S6). RsaG, RsaE and RsaC each carry several copies of this conserved sequence. RsaI is characterized by several conserved stretches of nucleotides: two regions are G-rich, while the third is a long, unpaired region that contains stretches of uridines interrupted by adenines and cytosines (Supplementary Figure S6).

**The ncRNAs are stable**

In many enterobacteria, the activity of sRNA acting as antisense RNA depends on the Hfq protein (19,63). Among the multiple functions of this pleiotropic regulator, Hfq stabilizes many sRNAs to provide regulatory activity. Therefore, we compared steady-state levels of the RNAs in RN6390 and its isogenic  $\Delta hfq$  derivative (representative data is shown in Figure 1D and Supplementary Figure S3). Although these data do not exclude that Hfq binds to the sRNAs, deletion of *hfq* had no major effect on sRNA levels, consistent with previous results for RNAIII (9).

Half-lives of all sRNAs, except for RsaB and RsaF, were determined in BHI medium at 37°C after rifampicin treatment (Figure 1D and Supplementary Figure S3). Most sRNAs are characterized by a long half-life: 15 min for RsaC and RsaI, 20 min and RsaD, 24 min for RsaE,

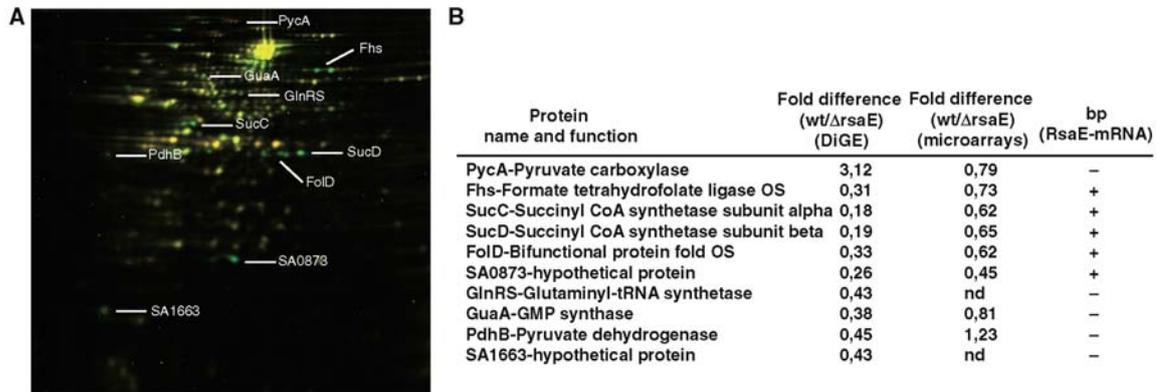


**Figure 3.** Secondary structures of various Rsa RNAs. (A) Gel fractionation of enzymatic cleavages and lead-induced cleavages of 5'-end-labeled RsaE: lanes T1: RNase T1, 0.01 and 0.02 U for 10 min at 20°C; lanes A: RNase A, 0.01 µg/ml and 0.001 µg/ml for 10 min at 20°C; lanes V1: RNase V1, 0.05 and 0.1 U for 5 min at 20°C; lanes Pb: lead-induced cleavage, 12.5 and 25 mM for 5 min at 20°C; lane T: RNase T1 under denaturing conditions; lane L: alkaline ladder; lane C: Incubation control. Bars denote the main reactivity changes induced by complex formation. (B) Probing data represented on the secondary structures of RsaE, RsaH, RsaA and RsaG RNAs. Enzymatic cleavages are as follows: RNase T1 (→), RNase A (→), RNase T2 (→), Lead-cleavages (→) moderate and (→) strong cleavages. The cytosines of the C-rich motif are encircled in black. (C) Sequence and structure alignment of RsaE coming from various Staphylococci and Bacillaceae: S1, *S. aureus*; S2, *S. epidermidis*; S3, *S. saprophyticus*; S4, *S. haemolyticus*; S5, *Micrococcus caseolyticus*; S6, *B. anthracis*; S7, *B. amyloliquefaciens*; S8, *B. cereus*; S9, *B. licheniformis*; S10, *B. subtilis*; S11, *B. thuringiensis*; S12, *Geobacillus thermodenitrificans*; S13, *Oceanobacillus iheyensis*. The helices are shown in grey. The C-rich conserved residues are highlighted in black. The alignment was done with the PARADISE platform (<https://simtk.org/home/paradise>).

55 min for RsaG and >60 min for RsaA and RsaH. Among the sRNAs, two distinct transcripts for RsaA and RsaG were visualized in northern experiments. The longer RsaA transcript is most likely generated from a transcriptional readthrough, and transcription stops 140 nt downstream at a second strong terminator. This second transcript is unstable, with a half-life of 2.5 min (Supplementary Figure S3). For RsaG, the shorter fragment most probably results from an endoribonucleolytic

cleavage in the unpaired interhelical region of RsaG, generating the stable upstream 5' hairpin motif. Recent works identified an essential enzyme, RNase J1 in *B. subtilis*, with dual endonucleolytic and 5'-3' exonucleolytic activities (64–66). This enzyme contributes significantly to the degradation of mRNAs (67,68) and cleaves small ncRNAs in *B. subtilis* (69). It remains to be addressed whether RNase J1 is responsible for site-specific cleavage of RsaG.

10 *Nucleic Acids Research*, 2009



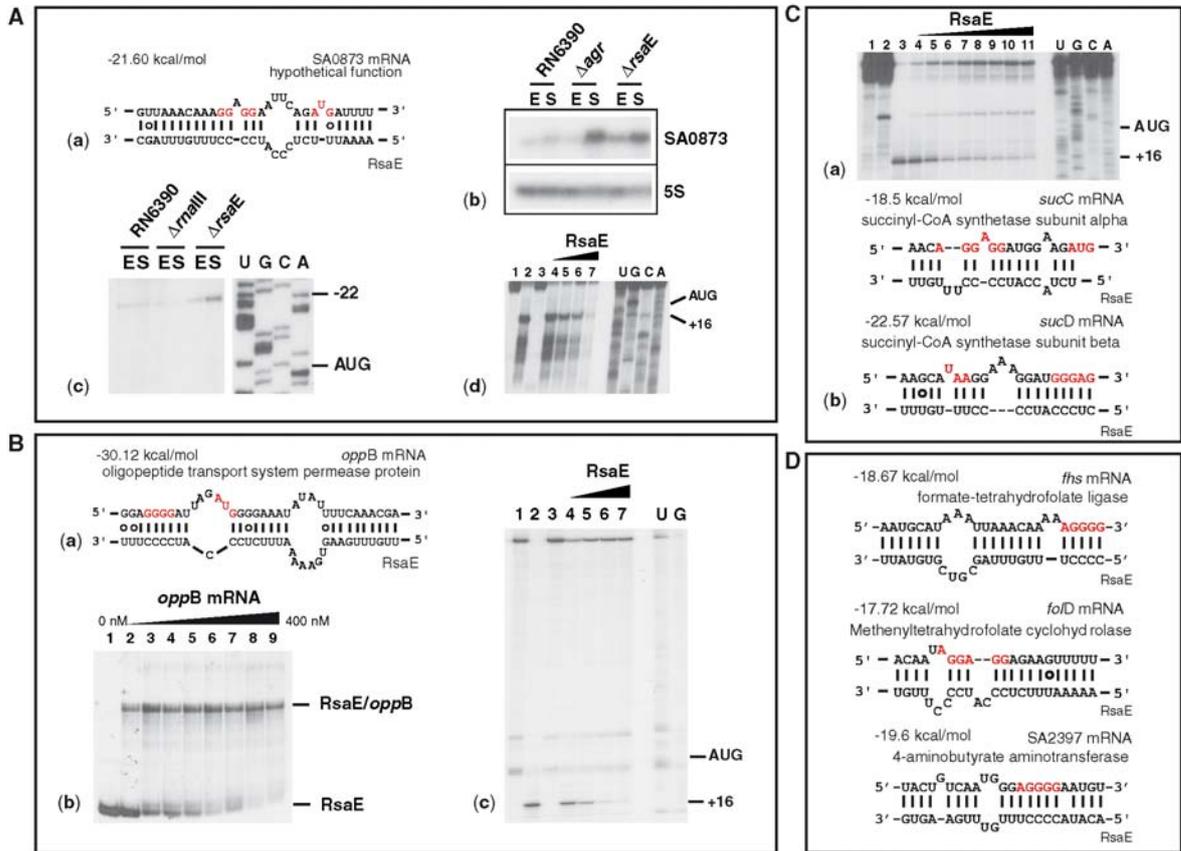
**Figure 4.** RsaE regulates the synthesis of several metabolic enzymes. (A) 2D fluorescence difference gel electrophoresis (DiGE) performed on the RN6390 and  $\Delta$ rsaE strains. Total protein extracts were prepared from cultures stopped at the stationary phase of growth and labeled with different cyanine-dyes. Yellow spots represent unchanged proteins; green spots, protein synthesis repressed in the strain expressing RsaE; red spots, protein synthesis increased in the strain expressing RsaE. The proteins were identified by mass spectrometry analysis. (B) Summary of the DiGE analysis, and comparative analysis with the microarray data. Ratios correspond to the quantification obtained for RN6390 versus the  $\Delta$ rsaE strain (data from four different experiments). (bp) for base pairing between RsaE and mRNA targets: (+) potential bp, (-) no bp.

**Functions regulated by RsaE as monitored by transcriptomic and proteomic analysis**

The function of RsaE, the only RNA found to be conserved in staphylococcal species and *Bacillaceae*, was assessed as follows. We constructed a RN6390-derived strain, in which the *rsaE* gene was deleted and replaced by *aphA-3* (aminoglycoside 3'-phosphotransferase, kanamycin resistance). In BHI medium, growth was not affected by the *rsaE* deletion. We then analyzed the effect of the mutation on gene expression and protein synthesis using microarrays and quantitative proteomics. The experiments were carried out under similar conditions on wild type (RN6390) and  $\Delta$ rsaE grown in BHI medium to early stationary phase. We have verified that under these conditions, *rsaE* expression is high in RN6390 at a level similar to *S. aureus* RNAIII while the expression of RsaE is abolished in the mutant  $\Delta$ rsaE strain (Figure 1A).

Transcriptional profiles were obtained using extensively validated oligoarrays (39). Statistically significant changes in the levels of gene expression were recorded, and only differences greater than 2-fold were considered. According to these criteria, 86 different mRNAs were differentially expressed in the wild type versus the mutant  $\Delta$ rsaE strains, representing ~4% of the transcriptome (Supplementary Table S3). The same proportions of genes were found to be up- and down-regulated by RsaE. Functional classification of these genes according to the COG database showed that many of these mRNAs encoded proteins involved in various metabolic pathways such as amino-acid transport and metabolism, lipid metabolism, inorganic ion transport, coenzyme transport and metabolism, carbohydrate metabolism and energy production. In addition, the synthesis of capsular polysaccharides (Cap5I-O) was significantly repressed (Supplementary Table S3). These cell-surface polysaccharides contribute to virulence by impeding phagocytosis and modulate *S. aureus* adherence to endothelial surfaces (70).

To assess the effect of the *rsaE* knockout on the staphylococcal proteome, cytosolic and secreted protein extracts were prepared and separated by a two-dimensional difference gel electrophoresis (DiGE), which enables quantitative differential display analysis (Figure 4A). The extracts were prepared from cells grown in BHI in the exponential and post-exponential phases of growth. Protein samples were pre-labeled with two different fluorescent dyes (Cy3 and Cy5). An internal standard, which includes the two samples present in the experiment, was pre-labeled with Cy2 and run on the same multiplexed gel. Quantitative measurements were made for each resolved protein spot relative to the cognate signal from the standard, which was used to normalize the ratio derived from one gel. Four different experiments were carried out to provide confident statistical analysis. We did not observe significant variations of secreted proteins between the two strains suggesting that RsaE does not regulate the synthesis of many of the exotoxins (serine proteases SplA-E and haemolysins  $\alpha$  and  $\beta$ ), as expected from the transcriptomic analysis (results not shown, Supplementary Table S3). For cytosolic proteins, even though we resolved only a small proportion of all proteins (around 100 proteins), we identified 10 proteins with strongly altered expression patterns (Figure 4). The synthesis of the pyruvate carboxylase PycA was shown up-regulated by RsaE. This enzyme catalyzes the ATP-dependent carboxylation of pyruvate to form oxaloacetate, which is involved in the biosynthesis of aspartate. The synthesis of several proteins (SucC/D, Fold, Fhs and SA0873) was strongly repressed even when the RsaE level was rather low (after 2 h of growth in BHI medium, Figure 1A), suggesting that their RsaE-dependent repression is most likely direct. These proteins are connected to carbohydrate metabolism: succinyl-CoA synthetase subunits  $\alpha$  and  $\beta$  (SucC/D) are part of the TCA cycle and the bi-functional protein Fold and the formate-tetrahydrofolate ligase Fhs are involved



**Figure 5.** RsaE regulates gene expression by direct interaction with target mRNAs. (A) RsaE binds to the RBS of SA0873 mRNA: (a) Base pairings predicted between SA0873 and RsaE. The Shine–Dalgarno sequence (SD) and the initiation codon AUG are shown in red. (b) Northern analysis of SA0873 mRNA prepared from various *S. aureus* strains (RN6390, RN6911- $\Delta$ agr and RN6390- $\Delta$ rsaE). The 5S rRNA was probed as an internal control. (c) The 5' start of the SA0873 mRNA was determined by primer extension with reverse transcriptase. The mRNA was only observed in the  $\Delta$ rsaE strain. Lanes U, G, C and A: sequencing ladders. (d) RsaE binding to SA0873 mRNA prevents the formation of the ribosomal initiation complex. Lanes 1 and 3: Incubation controls of mRNA alone and with 200 nM of RsaE, respectively; lane 2: formation of the ternary complex formed from the *S. aureus* 30S subunit, the initiator tRNA, and the mRNA; lanes 4–7: Formation of the ternary ribosomal mRNA–30S–tRNA complex in the presence of increasing RsaE concentrations of 50, 100, 150 and 200 nM, respectively. (B) RsaE binds to *oppB* mRNA and prevents the ribosome binding: (a) Base pairings between *oppB* and RsaE. Same legend as above. (b) Band shift experiment showing a stable RsaE-*oppB* mRNA complex. Lane 1, incubation control of RsaE alone; lanes 2–9, complex formation between 5'-end-labeled RsaE with increasing *oppB* mRNA concentrations of 10, 50, 100, 150, 200, 250, 300 and 400 nM, respectively. (c) RsaE binds to *oppB* mRNA and prevents the formation of the 30S initiation complex. Same legend as above. (C) The RsaE-*sucC/D* mRNA complex prevents the formation of the ribosomal initiation complex. (a) Toeprinting assays: lanes 1, 2, incubation controls on the mRNA alone and with 200 nM of RsaE; lane 3, ternary mRNA–30S–tRNA complex; lanes 4–11, ternary complex with increasing RsaE concentrations 50, 100, 200, 300, 400, 500, 600 and 800 nM, respectively. (b) RsaE-*sucC/D* mRNA base pairing interaction. The hybrid formed between RsaE and *sucC* mRNA includes the stop codon of the upstream *sucC* mRNA. (D) Predicted base pairing interactions between RsaE and other mRNAs repressed by RsaE.

in glyoxylate and dicarboxylate metabolism and the biosynthesis of folate. The transcriptomic analysis also shows that the levels of *sucCD* and SA0873 mRNAs are significantly decreased in the RN6390 strain (Figure 4A). We also analyzed the level of SA0873 mRNA in different *S. aureus* strains in early and late exponential phase (Figure 5A). In RN6390 and WA400 ( $\Delta$ rnaIII), the expression of the mRNA is very low whereas it significantly accumulates in the late-exponential phase in the mutant  $\Delta$ rsaE and  $\Delta$ agr strains, consistent with the AgrA-dependent expression of *rsaE* (Figure 1A).

Taken together, these data show that RsaE altered mRNA and/or protein levels for enzymes involved in several metabolic pathways that are necessary to synthesize biosynthetic intermediates. This may contribute to facilitate the transition from the exponential to the stationary phase of growth (71).

**Staphylococcus aureus RsaE regulates gene expression through mRNA pairing**

Although the transcriptomic and proteomic analyses give insights into the putative functions of RsaE, they do not

12 *Nucleic Acids Research*, 2009

permit us to distinguish between primary and secondary effects. As mentioned above, two conserved unpaired regions within RsaE contain a C-rich sequence motif. We postulated that these regions might interact with ribosome-binding sites (RBS) of target mRNAs, as does RNAIII (9,14). We searched for potential base pairing that involves the conserved unpaired regions of RsaE (nt 19–65) and the RBSs of mRNAs (30 nt downstream and upstream the AUG start codon) using the program RNAup (72). This program takes the structures of both RNAs into account and assesses the energy cost of melting local secondary structures for the formation of intermolecular base pairs. We first restricted the search to the genes whose mRNA and/or protein levels changed in an RsaE-dependent fashion. The mRNA–RsaE hybrids that were considered had a minimum free energy of  $\leq -18$  kcal/mol (Figure 5C and D). Significant base pairing was predicted between RsaE and mRNAs encoding OppB (oligopeptide transport system, *opp-3* operon), SucC and SucD, SA0873 (hypothetical protein), Fhs (formate tetrahydrofolate synthase), SA2397 (4-aminobutyrate aminotransferase) and Fold (bi-functional protein).

We used two other *ab initio* analyses to predict mRNA targets. The sequences of each RBS for all of the *S. aureus* mRNAs were concatenated to the sequence of RsaE by a linker and submitted to RNAfold for predicting base pairing (9). This approach estimates the hybridization of two RNA sequences and calculates the minimum free energy of interaction. Prediction of targets was also carried out with the algorithm developed for *Listeria monocytogenes* (73). This program quantifies the strength of RNA duplexes and takes into account positive contributions due to pairing nucleotides and negative contributions due to bulged nucleotides and internal loops. In addition, statistical significance of the pairings is assessed with respect to an ensemble of random sequences. This program was successfully trained on several regulatory *L. monocytogenes* sRNAs and *S. aureus* RNAIII (73). By comparing the two approaches, stable base pairings were predicted between RsaE and the *hutH* (encoding histidine ammonia-lyase), *oppB* and *sucD* mRNAs.

Thus, the *in silico* data predict that the *oppB*, *sucD* and SA0873 mRNAs are among the best candidates to form stable duplexes with RsaE (Figure 5). Indeed, band shift experiments show that RsaE forms a stable complex with an *oppB* mRNA fragment (370 nt) encompassing the entire 5' untranslated region and part of the coding sequence (Figure 5B). A time-course analysis suggests that this complex is rapidly formed, as was shown for *S. aureus* RNAIII and its mRNA targets (data not shown) (9). We further analyzed the effect of RsaE on the formation of initiation ribosomal complexes using toeprinting assays. The ternary initiation complex formed by *S. aureus* 30S subunits, the initiator tRNA<sup>Met</sup> and the three mRNAs (*oppB*, *sucD* and SA0873) blocked primer extension 16 nt downstream of the initiation codon. Binding of RsaE to any of the three mRNAs significantly decreased the toeprint signal (Figure 5A–C) indicating that the formation of active ribosomal initiation complexes is prevented. Therefore, we hypothesized that

RsaE inhibits the initiation of translation of the *oppB*, *sucD* and SA0873 mRNAs by masking the RBS.

#### RsaE is expressed in *Bacillus subtilis*

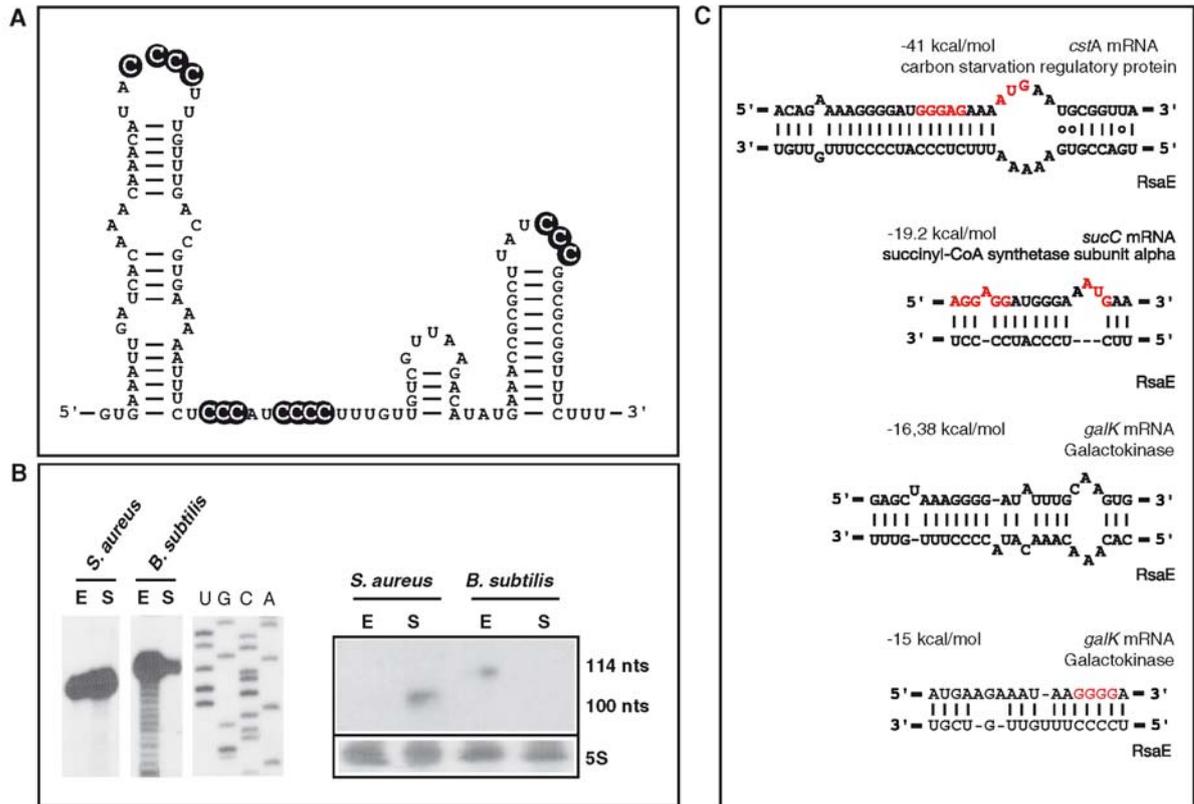
Since RsaE is conserved in *Bacillaceae*, we have analyzed the expression of *rsaE* in *B. subtilis* grown in BHI medium. Northern experiments indicated a homogenous RNA transcript that was preferentially expressed in the exponential phase (Figure 6). Determination of the 5'-end by primer extension revealed a RNA 114-nt long, consistent with the predicted length (Figure 6B). As in *S. aureus*, its secondary structure is characterized by a 5' hairpin motif and a stable terminator structure at the 3'-end, and it contains several highlighted C-rich sequences (Figure 6A). Interactions between the RBS of *B. subtilis* mRNAs and RsaE were predicted using the method applied to *S. aureus* RNAIII (9). This analysis revealed several mRNA candidates that encode proteins involved in carbohydrate and sugar utilization (Figure 6C). Indeed the most stable duplex was predicted between RsaE and the *gstA* mRNA, encoding the carbon starvation protein (CstA). Interestingly, pairings between RsaE and *sucC* mRNA appear to be conserved in *B. subtilis* and *S. aureus*. We also found that the two C-rich motifs of *B. subtilis* RsaE could interact with the RBS and part of the coding sequence of the *galK* mRNA, encoding galactokinase, respectively (Figure 6C).

## DISCUSSION

### A plethora of RNA-dependent regulatory events in *Staphylococcus aureus*

In recent years, new *in silico* and experimental strategies have been developed to identify functional regulatory intergenic regions in a variety of model bacteria (74,75). Here, we used a combination of several bioinformatic tools to predict more than 500 potential functional IGRs from *S. aureus* in addition to the highly conserved bacterial ncRNAs (tmRNA, 6S RNA, 4.5S RNA and RNase P). These IGRs were selected based on the presence of orphan terminators, on their GC content and their sequence and structure conservation among *S. aureus*, staphylococcal species and firmicutes (Supplementary Figure S1). The visualization of the data with ApolloRNA provided an easy tool for allowing the characterization of three main families of IGRs: repeated sequences, conserved leader regions of mRNAs, and *bona fide* ncRNAs. In this study, we did not search specifically for antisense RNAs (73,76).

The first class of non-coding regions includes *Staphylococcus aureus* repeats (STARs) (29). Some of them carry features reminiscent of the clustered regularly interspaced short palindromic repeats (Gaspin *et al.*, unpublished data). The expression of one of these elements failed to be significant in RN6390 grown in BHI and NZM medium (RsaX19, Table 1). However, we do not exclude that STAR elements may have evolved regulatory functions under defined stress conditions. These repeats were also present in the leader



**Figure 6.** RsaE ortholog is expressed in *B. subtilis*. (A) Secondary structure of *B. subtilis* RsaE derived from the structural alignment shown in Figure 2. The C-rich conserved boxes are shown in black. (B) Expression of *B. subtilis* RsaE: (left) Primer extension was used to map the 5' start of *S. aureus* and *B. subtilis* RsaE. The coordinates of the RNA are 1232743–1232852 corresponding to *Bacillus subtilis* subsp. *subtilis* strain 168 (NC\_000964.2). The RNA has been referred in GenBank with the accession number GQ403626. Lanes U, G, C and A: sequencing ladders. (Right) Northern experiment performed on *S. aureus* and *B. subtilis* RsaE. Both experiments were carried out on total RNA extracts prepared at the exponential (E) and stationary (S) phases. (C) Potential base pairings between *B. subtilis* RsaE and mRNAs encoding proteins involved in carbohydrate metabolism and sugar utilization. The Shine–Dalgarno (SD) sequence is in red.

region of mRNA (22) as well as in the *S. aureus* sprA3 ncRNA (Supplementary Table S1) (8).

The second class of *S. aureus* IGRs comprises numerous *cis*-acting regulatory elements located in the 5' untranslated regions of mRNAs, which regulate enzymes and proteins involved in the translational machinery and in various metabolic pathways. These regions are highly conserved among firmicutes, and in *B. subtilis* many of them provide specific binding sites for *trans*-acting regulators (metabolites, tRNAs or proteins) (77–80). RNAsim has identified all the assigned riboswitches with sizes of more than 150 nt that are regulated by small metabolites such as thiamine pyrophosphate, flavin mononucleotide, S-adenosyl methionine, glucosamine 6-phosphate, lysine, purine and preQ1 (77,78) (Supplementary Table S1). The program also revealed conserved leaders that include T-box antitermination signals. These *cis*-acting elements are sensed by uncharged tRNA to regulate the expression of genes involved in tRNA aminoacylation and the transport and biosynthesis

of amino acids (79,80). Other conserved leaders recognized by regulatory proteins have been identified in the *S. aureus* genome. A typical example is RsaK which is derived from the 5' leader of *glcA* mRNA. This region carries a long terminator structure overlapping the conserved RAT motif (Supplementary Figure S4). This regulatory region is also present upstream of the *S. aureus* *ptsG* gene, which encodes the glucose-specific phosphotransferase system. In *B. subtilis* and *S. carnosus*, the RAT motif is recognized by the transcriptional antiterminator protein GlcT, which senses the intracellular concentration of glucose (60,81). The RAT–GlcT complex prevents the formation of the terminator structure, allowing transcription of the downstream gene. Hence, we propose that *S. aureus* *glcA* mRNA is under the control of GlcT and that the detection of RsaK would be a consequence of a premature termination event. Similarly, RNAsim detected two conserved 5'UTRs with a terminator just upstream of the *pyrR* and *pyrP* genes. These two leader regions are characterized

by an AGAGAG hairpin loop motif, proposed to be the main recognition element for the regulatory protein PyrR in *B. subtilis* (82). PyrR, bound to UMP, binds to target mRNAs and induces premature termination of transcription (82) (Supplementary Figure S5). RNAsim pointed out several highly conserved leaders of operons encoding ribosomal proteins and the three initiation factors, which are regulated by a feedback mechanism (Supplementary Table S1). For instance, the *S. aureus* *infC* leader carries a typical binding site for the L20 ribosomal protein (83). In *B. subtilis*, binding of L20 to the *infC* leader region stabilizes the formation of a terminator structure to induce premature arrest of transcription (83). We have identified typical Rho-independent terminators located upstream of several genes (i.e. SA1002, SA1108, SA1841, *splA* and *uppS*) suggesting that these genes might be regulated by a premature transcription termination event (Supplementary Table S1). In agreement with this observation, *cis*-acting RNAs in low-GC Gram-positive bacteria were shown to predominantly regulate gene expression through a transcription termination/antitermination mechanism (77,80).

The third class of IGRs reported here are the *bona fide* sRNAs. The expression from 36 IGRs was analyzed in *S. aureus* RN6390 grown in two different media. From this analysis, 10 new ncRNAs (RsaA to J) of sizes ranging from 59 to 544 nt were identified (Table 1, Figure 2). However, this fraction of ncRNAs is probably an underestimate since many growth conditions and stress protocols were not investigated. None of the Rsa RNAs carries a small ORF, and the RNA genes are dispersed in the genome. The location of the *rsaC* gene is of interest since its 5' start site is located close to a repeat element and the downstream gene encodes a transposase. Many genes encoding sRNAs from various bacteria have also been found to be close to mobile elements such as transposons, prophages and pathogenicity islands (8,84,85). Based on their genomic location, it was proposed that sRNAs, like structured tRNAs and tmRNA, might play a role in genome plasticity by serving as integration sites for foreign genes (85). We have predicted most of the spr RNA expressed from the pathogenicity islands (8), except sprE (Supplementary Tables S1 and S2). The conserved region of sprE comprises only 70 nt, which is below the cut off used by RNAsim. A recent high throughput computational tool identified Rho-independent terminators downstream of conserved intergenic regions (23). This powerful analysis predicted 49 new ncRNA genes from *S. aureus*, 35 of which were also predicted by our approach (Supplementary Table S1). We show above that four of them are highly expressed (RsaH, RsaD, RsaE and RsaI), while we assigned other 14 IGRs to STAR elements (Table 1, Supplementary Tables S1 and S2). Recently, a phylogenetic profiling study has been performed on all available bacterial genomes. Based on specific signatures, this method allows clustering of intergenic and coding regions (22). From this analysis, six RNAs were shown to be expressed in the *S. aureus* N315 strain. These RNAs were also predicted by RNAsim (Supplementary Table S1), but were not further experimentally analyzed, as many of

them resemble riboswitches or leaders of mRNAs. However, our analysis confirms the high expression of *rsaI* in RN6390 and other strains. All in all, our study revealed ten new ncRNAs in addition to the pathogenicity island spr RNAs (8).

#### The sRNAs are endowed with a common signature and different expression patterns

Many of the Rsa RNAs are expressed under particular conditions. Some conditions, such as acidic pH, oxidative stress and osmotic stress, could reflect environments that the bacteria encounter within a host. The synthesis of RsaA, RsaH and RsaE is induced under osmotic stress, oxidative stress and at acidic pH, while RsaC and RsaD are cold shock-induced RNAs. Five of the RNAs (RsaA, RsaD, RsaE, RsaF and RsaJ) show significant variation in expression among *S. aureus* strains although their genes are conserved in all *S. aureus* strains. RsaA, D and F synthesis is strongly enhanced in strains expressing an active  $\sigma^B$  factor while the expression of RsaE is negatively controlled by  $\sigma^B$  via an unknown mechanism (Figure 1A and C). Interestingly, we found a typical  $\sigma^B$  consensus binding site upstream of *rsaA* (Figure 1C). Several  $\sigma^B$ -dependent ncRNAs were recently found in *L. monocytogenes* (86,87), but their functions remain to be identified. *Staphylococcus aureus* encodes three sigma factors: the housekeeping  $\sigma^A$  (a homolog of *E. coli*  $\sigma^{70}$ ), two alternative factors  $\sigma^B$  (88) and  $\sigma^H$  (89). While  $\sigma^H$  is required for transcriptional regulation of the *com* operon encoding competence factors (89),  $\sigma^B$  positively controls the synthesis of proteins involved in cell envelope biosynthesis, metabolism, virulence and signaling pathways (52,90,91). A significant number of genes that are upregulated by  $\sigma^B$  do not contain a promoter consensus, and several genes are negatively influenced by  $\sigma^B$  (90). These data suggest that  $\sigma^B$  indirectly regulates the expression of these genes most likely through the induction of transcription of regulatory proteins or ncRNAs. Of interest, we found that RsaA could potentially form stable duplexes with  $\sigma^B$ -repressed mRNAs (52,90). These mRNAs encode the magnesium-citrate transporter (CitM), an APC amino-acid permease, and the pyruvate carboxylase PycA (Supplementary Figure S7). Although further experiments will be required to validate these targets, we propose that RsaA is part of the  $\sigma^B$  regulon and regulates gene expression via an antisense mechanism.

Probing of various Rsa RNA structures combined with RNA sequence and structure alignments revealed similar common features (Figure 3, Supplementary Figure S6). Most of them carry a hairpin structure at the 5'-end and a classical terminator at the 3'-end. The 5' hairpin most probably confers the high stability of the Rsa RNAs, as has been shown for mRNAs in *B. subtilis* (92,93) and *S. aureus* (9,11). Intriguingly, most of the Rsa RNAs, except RsaF and RsaI, contain an unpaired UCCC motif which is located in the most conserved regions of the RNAs (Figure 7). *Staphylococcus aureus* RNAIII also carries UCCC motifs in three hairpin loops; all are known to interact with the Shine-Dalgarno (SD) sequences of target mRNAs (9). In RsaA, this motif is predicted to

RsaA	AAAGUUCUCCCAAGGAAG
RsaB	GAGCGCAUCCCAAUUAAA
RsaC1	CUGUCGUUCCCUUCAUCU
RsaC2	ACGCCAUUCCCUACACAC
RsaC3	ACUGUUCUCCCUACUAGA
RsaC4	AUUGUAUCUCCCUUUCU
RsaC5	GGUUUCCUCCCUCCAUAG
RsaD	AUUCAUUCUCCCAUAAAAG
RsaE1	AAAACAUUCCCUUUGUU
RsaE2	AAAUUUCUCCCAUCCCU
RsaE3	UCUCCAUUCCCUUUGUU
RsaG1	UGCUCUUCUCCCUUGCAA
RsaG2	UCAUAAUCCCUUGUUAUG
RsaG3	CUUAUUUCUCCCGACUAG
RsaH1	AUUUGACUCCCUUUAGUA
RsaH2	UUAGUGUCCCGUACGUU
RsaJ	CACUAUUCUCCCUAGUGAG
RNAIII1	UAAGCCAUCCCAACUUA
RNAIII2	AAUAACAUCCCAAAAAU
RNAIII4	AAAACAUUCCCUUAAUAA
RNAIII5	GAGCCCCUCCCAAGCUCG
Consensus	nnnnnnUCCnnnnnn

Figure 7. The C-rich box conserved motif in *S. aureus* regulatory RNAs. The alignment of the C-rich sequence motif in *S. aureus* RNAIII (9) and Rsa RNAs was done with the help of the PARADISE platform (<https://simtk.org/home/paradise>).

interact with  $\sigma^B$ -repressed mRNAs (Supplementary Figure S7). From these sequence characteristics, the ncRNAs may have resulted from a convergent evolution which selected efficient repressors. Similarly, a conserved C-rich hairpin loop of *Salmonella typhimurium* CyaR sRNA binds to the SD sequence of *ompX* mRNA to repress its translation (94). This sequence motif is also found in other ncRNAs recently identified in *L. monocytogenes* (73,86,87) and *B. subtilis* (95), suggesting that these RNAs act at the mRNA level. Even though the majority of low-GC Gram-positive bacterial mRNAs carry a strong SD sequence (96) specificity could be attained by the surrounding nucleotides and the structure of both interacting RNAs. Thus we postulate that the C-rich box might be a signature for regulatory RNAs that repress translation initiation of target mRNAs.

#### RsaE regulates several metabolic pathways

The above rule can be applied to *S. aureus* RsaE, the only sRNA which is conserved in all staphylococcal species studied and in *Bacillaceae* (Figure 3, Supplementary Figure S5). The expression pattern of RsaE is rather complex since a high level of RsaE is dependent on the quorum sensing *agr* system, whereas its expression drops considerably in strains expressing a fully active  $\sigma^B$  factor.

A stable  $\sigma^B$ -dependent RNA (RsaF) was also identified although its function remains to be determined. Both RNAs are differently expressed, suggesting that transcriptional interference regulates their synthesis. Alternatively, transcription may occur at the same start site of *rsaE* but termination readthrough could lead to the synthesis of a larger RNA precursor, which would subsequently be processed (Supplementary Figure S5). The functions of RsaE were then addressed by a combination of approaches, including a comparative analysis of the transcriptome and of the proteome performed on RN6390 and the isogenic  $\Delta$ *rsaE* mutant strain. Transcriptomic analysis revealed many genes that were up- and down-regulated by RsaE, suggesting additional indirect regulation (Supplementary Table S3). In this scenario, RsaE may regulate the expression of a regulatory protein by direct binding to the mRNA, as does RNAIII, or alternatively, it may sequester a regulatory protein. Even though further experiments are required to assign the entire set of direct RsaE targets, these data demonstrate a link between RsaE and the regulation of various metabolic pathways (including amino acid and peptide transport, cofactor synthesis, lipid metabolism, carbohydrate metabolism and the TCA cycle). Our data strongly support that RsaE directly represses the synthesis of OppB, the two subunits of the succinyl-CoA synthetase involved in the TCA cycle, two proteins (FhS and FdD) involved in the folate biosynthesis, and SA0873 of unknown function. Direct interactions are strengthened by *in vitro* experiments showing that RsaE binds efficiently to *oppB*, *sucD* and SA0873 mRNAs and inhibits the formation of ribosomal initiation complexes (Figure 5A–C). One of the main roles of bacterial Opp systems is to supply bacteria with essential amino acids and peptides. *Staphylococcus aureus* genomes carry four polycistronic *opp*-1–4 operons and one monocistronic *opp5A* gene (97). Of the four putative Opp systems, only *opp*-3 can supply *S. aureus* with peptides as nutritional source and it is the only one whose expression is modulated by amino acids (97). These proteins thus play a key role in nitrogen and carbon metabolism. Our data further suggest that RsaE directly regulates several proteins involved in carbohydrate metabolism and in the TCA cycle. TCA is an essential source of energy and biosynthetic intermediates for many organisms (71). In Gram-positive bacteria, the TCA cycle is repressed in the presence of glucose, and conversely, the derepression coincides with the depletion of the catabolized carbon source. *Staphylococcus aureus* uses the pentose phosphate and glycolytic pathways to catabolize glucose to pyruvate. Under aerobic conditions, pyruvate undergoes oxidative decarboxylation to produce acetyl-coenzyme A, which in turn leads to the accumulation of acetate in the extracellular medium during the exponential phase of growth. In stationary phase, when glucose has been consumed, acetate and amino acids are used as alternate carbon sources and are oxidized via the TCA cycle. A higher level of enzymes involved in glycolysis was indeed identified in the exponentially growing *S. aureus* COL strain while enzymes of the TCA cycle and gluconeogenesis were enhanced in the stationary phase (98). Thus, it is quite surprising that RsaE

16 *Nucleic Acids Research*, 2009

strongly represses the synthesis of SucCD enzymes in the stationary phase. RN6390 which is  $\sigma^B$ -deficient, has an impaired TCA cycle-mediated acetate catabolism (99). The isogenic strain SH1000, with restored  $\sigma^B$  activity, had significantly increased acetate catabolism (99). Since RsaE is strongly repressed by  $\sigma^B$  (Figure 1A), the ncRNA might be partly responsible for this deficiency in acetate catabolism in the stationary phase. The loss of acetate catabolism neither alters stationary-phase survival nor virulence factor expression (99). Interestingly enough, RsaE is highly conserved in most of the *Bacillaceae*. The *rsaE* gene of *B. subtilis* is mainly expressed in the exponential phase and decreases in the stationary phase (Figure 5B). The presence of RsaE in *S. aureus*, *M. caseolyticus* and *Bacillus*, but not in other firmicutes, raises evolutionary considerations (Supplementary Figure S5). The *M. caseolyticus* genome is closely related to *S. aureus* and *B. anthracis* (100). Indeed, the essential biological pathways of *M. caseolyticus* are similar to those of Staphylococci, although it lacks several sugar and amino-acid metabolism pathways and a plethora of virulence genes that are present in *S. aureus*. In addition, the genome possesses oxidative phosphorylation machineries that are closely related to those present in *Bacillaceae*. This phylogenetic analysis suggests that the *Bacillus*, *Macrococcus* and *Staphylococcus* derive from a common Gram-positive ancestor, and that the speciation of *Macrococcus* into *Staphylococcus* was accompanied by the acquisition of arrays of virulence genes for host adaptation (100). RsaE, which is present in all three genera, probably arose from the common ancestor genome before the divergence of *Bacillus* and *Macrococcus*. Therefore it is not so surprising that RsaE regulates common metabolic pathways. Consistent with this hypothesis, the *B. subtilis* RsaE ortholog is predicted to regulate the translation of *sucC* mRNA in the TCA cycle, as in *S. aureus*, as well as the translation of *galK* and *estA* mRNAs, which are both involved in carbon source utilization (Figure 6C). In *E. coli*, Spot42 sRNA inhibits the initiation of translation of *galK* mRNA in response to glucose availability (101). Spot42 is also predicted to form a stable duplex with the 5'-end of the *sucC* mRNA (101). Hence, Spot42 could be the functional equivalent of RsaE. These studies strongly suggest that, as in Gram-negative bacteria, *S. aureus* and *B. subtilis* bacteria use RsaE to control the use of carbon sources and sugar metabolism (102). Further experimental works will characterize how far the regulatory functions of RsaE have evolved in *Bacillus*, and *Staphylococcaceae*.

The present study shows that the *S. aureus* genome likely encodes a high diversity of regulatory RNAs, including mRNA leaders that affect expression in *cis*, the pleiotropic regulator RNAIII (6,103), and the small noncoding RNAs [(8,11,12,22); this work]. Deep sequencing efforts and tiling arrays would certainly help to provide a more complete view of the RNA-dependent regulatory circuits in this pathogenic bacterium (86,104). A particular challenge will now concern their roles in response to the host environment and during the infection process of this major pathogen.

## ACCESSION NUMBERS

GQ403615–GQ403626, GPL7137.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

The authors are thankful to Anne-Catherine Helfer for providing technical help for the DiGE experiments, Philippe Hammann and Philippe Wolf for the identification of proteins by mass spectroscopy analysis, and Elodie Vogel for the expression analysis of several ncRNAs. They are grateful to Gerhart Wagner, Joerg Vogel and Eric Westhof for helpful comments and critical reading of the manuscript, and to Efhimia Lioliou and Gérard Lina for discussions. This project used equipment facilities provided by Toulouse Midi-Pyrénées bioinformatic platform.

## FUNDING

Centre National de la Recherche Scientifique (CNRS; to P.R.); Institut National pour la Recherche Médicale (INSERM; to F.V.); Agence Nationale pour la Recherche (ANR05-MIME: to C.G., F.V. and P.R.; ANR09-BLAN: to F.V. and P.R.); European Community (BacRNA, FP6-018618; to P.R.); Swiss National Science Foundation PP00B-103002/1 (to J.S.) and 3100A0-116075 (to P.F.). Funding for open access charge: CNRS.

*Conflict of interest statement.* None declared.

## REFERENCES

- Waters, L.S. and Storz, G. (2009) Regulatory RNAs in bacteria. *Cell*, **136**, 615–628.
- Vogel, J. and Wagner, E.G.H. (2007) Target identification of small noncoding RNAs in bacteria. *Curr. Opin. Microbiol.*, **10**, 262–270.
- Toledo-Arana, A., Repoila, F. and Cossart, P. (2007) Small noncoding RNAs controlling pathogenesis. *Curr. Opin. Microbiol.*, **10**, 182–188.
- Vogel, J. (2009) A rough guide to the non-coding RNA world of *Salmonella*. *Mol. Microbiol.*, **71**, 1–11.
- Taubes, G. (2008) The bacteria fight back. *Science*, **321**, 356–361.
- Novick, R.P. and Geisinger, E. (2008) Quorum sensing in staphylococci. *Annu. Rev. Genet.*, **42**, 541–564.
- van Schaik, W. and Abee, T. (2005) The role of sigmaB in the stress response of Gram-positive bacteria – targets for food preservation and safety. *Curr. Opin. Biotechnol.*, **16**, 218–224.
- Pichon, C. and Felden, B. (2005) Small RNA genes expressed from *Staphylococcus aureus* genomic and pathogenicity islands with specific expression among pathogenic strains. *Proc. Natl Acad. Sci. USA*, **102**, 14249–14254.
- Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M., Chevalier, C., Helfer, A.C., Benito, Y., Jacquier, A. et al. (2007) *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev.*, **21**, 1353–1366.
- Geisinger, E., Adhikari, R.P., Jin, R., Ross, H.F. and Novick, R.P. (2006) Inhibition of rot translation by RNAIII, a key feature of *agr* function. *Mol. Microbiol.*, **61**, 1038–1048.

11. Roberts,C., Anderson,K.L., Murphy,E., Projan,S.J., Mounts,W., Hurlburt,B., Smeltzer,M., Overbeek,R., Disz,T. and Dunman,P.M. (2006) Characterizing the effect of the *Staphylococcus aureus* virulence factor regulator, SarA, on log-phase mRNA half-lives. *J. Bacteriol.*, **188**, 2593–2603.
12. Anderson,K.L., Roberts,C., Disz,T., Vonstein,V., Hwang,K., Overbeek,R., Olson,P.D., Projan,S.J. and Dunman,P.M. (2006) Characterization of the *Staphylococcus aureus* heat shock, cold shock, stringent, and SOS responses and their effects on log-phase mRNA turnover. *J. Bacteriol.*, **188**, 6739–6756.
13. Novick,R.P., Ross,H.F., Projan,S.J., Kornblum,J., Kreiswirth,B. and Moghazeh,S. (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J.*, **12**, 3967–3975.
14. Huntzinger,E., Boisset,S., Saveanu,C., Benito,Y., Geissmann,T., Namane,A., Lina,G., Etienne,J., Ehresmann,B., Ehresmann,C. et al. (2005) *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate *spa* gene expression. *EMBO J.*, **24**, 824–835.
15. Morfeldt,E., Taylor,D., von Gabain,A. and Arvidson,S. (1995) Activation of alpha-toxin translation in *Staphylococcus aureus* by the trans-encoded antisense RNA, RNAIII. *EMBO J.*, **14**, 4569–4577.
16. Altuvia,S. (2007) Identification of bacterial small non-coding RNAs: experimental approaches. *Curr. Opin. Microbiol.*, **10**, 257–261.
17. Wassarman,K.M. (2007) 6S RNA: a small RNA regulator of transcription. *Curr. Opin. Microbiol.*, **10**, 164–168.
18. Babitzke,P. and Romeo,T. (2007) CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Curr. Opin. Microbiol.*, **10**, 156–163.
19. Brennan,R.G. and Link,T.M. (2007) Hfq structure, function and ligand binding. *Curr. Opin. Microbiol.*, **10**, 125–133.
20. Bohn,C., Rigoulay,C. and Boulou,P. (2007) No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. *BMC Microbiol.*, **7**, 10.
21. Večerek,B., Rajkowitz,L., Sonnleitner,E., Schroeder,R. and Bläsi,U. (2008) The C-terminal domain of *Escherichia coli* Hfq is required for regulation. *Nucleic Acids Res.*, **36**, 133–143.
22. Marchais,A., Naville,M., Bohn,C., Boulou,P. and Gautheret,D. (2009) Single-pass classification of all noncoding sequences in a bacterial genome using phylogenetic profiles. *Genome Res.*, **19**, 1084–1092.
23. Livny,J., Teonadi,H., Livny,M. and Waldor,M.K. (2008) High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. *PLoS ONE*, **3**, e3197.
24. Benito,Y., Kolb,F.A., Romby,P., Lina,G., Etienne,J. and Vandenesch,F. (2000) Probing the structure of RNAIII, the *Staphylococcus aureus agr* regulatory RNA, and identification of the RNA domain involved in repression of protein A expression. *RNA*, **6**, 668–679.
25. Chiapello,H., Bourgain,I., Sourivong,F., Heuclin,G., Gendrault-Jacquemard,A., Petit,M.A. and El Karoui,M. (2005) Systematic determination of the mosaic structure of bacterial genomes: species backbone versus strain-specific loops. *BMC Bioinformatics*, **6**, 171.
26. Chiapello,H., Gendrault,A., Caron,C., Blum,J., Petit,M.A. and El Karoui,M. (2008) MOSAIC: an online database dedicated to the comparative genomics of bacterial strains at the intra-species level. *BMC Bioinformatics*, **9**, 498.
27. Le Flèche,P., Hauck,Y., Onteniente,L., Prier,A., Denoed,F., Ramière,V., Sylvestre,P., Benson,G., Ramière,F. and Vergnaud,G. (2001) A tandem repeats database for bacterial genomes: application to the genotyping of *Yersinia pestis* and *Bacillus anthracis*. *BMC Microbiol.*, **1**, 2.
28. Dsouza,M., Larsen,N. and Overbeek,R. (1997) Searching for patterns in genomic data. *Trends Genet.*, **13**, 497–498.
29. Cramton,S.E., Schnell,N.F., Götz,F. and Brückner,R. (2000) Identification of a new repetitive element in *Staphylococcus aureus*. *Infect. Immun.*, **68**, 2344–2348.
30. Petersohn,A., Bernhardt,J., Gerth,U., Hoper,D., Koburger,T., Volker,U. and Hecker,M. (1999) Identification of sigma(B)-dependent genes in *Bacillus subtilis* using a promoter consensus-directed search and oligonucleotide hybridization. *J. Bacteriol.*, **181**, 5718–5724.
31. Kingsford,C.L., Ayanbule,K. and Salzberg,S.L. (2007) Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol.*, **8**, R22.
32. Klein,R.J., Misulovin,Z. and Eddy,S.R. (2002) Noncoding RNA genes identified in AT-rich hyperthermophiles. *Proc. Natl Acad. Sci. USA*, **99**, 7542–7547.
33. Rivas,E. and Eddy,S.R. (2001) Noncoding RNA gene detection using comparative sequence analysis. *BMC Bioinformatics*, **2**, 8.
34. Lewis,S.E., Searle,S.M., Harris,N., Gibson,M., Lyer,V., Richter,J., Wiel,C., Bayraktaroglu,L., Birney,E., Crosby,M.A. et al. (2002) Apollo: a sequence annotation editor. *Genome Biol.*, **3**, RESEARCH0082.
35. Hofacker,I.L. (2003) Vienna RNA secondary structure server. *Nucleic Acids Res.*, **31**, 3429–3431.
36. Janson,L. and Arvidson,S. (1990) The role of the <sup>TM</sup>-lysin gene (*hid*) in the regulation of virulence genes by the accessory gene regulator (*agr*) in *Staphylococcus aureus*. *EMBO J.*, **9**, 1391–1399.
37. Arnaud,M., Chastanet,A. and Debarbouille,M. (2004) New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl. Environ. Microbiol.*, **70**, 6887–6891.
38. Redko,Y., Bechhofer,D.H. and Condon,C. (2008) Mini-III, an unusual member of the RNase III family of enzymes, catalyses 23S ribosomal RNA maturation in *B. subtilis*. *Mol. Microbiol.*, **68**, 1096–1106.
39. Charbonnier,Y., Gettler,B., François,P., Bento,M., Renzoni,A., Vaudaux,P., Schlegel,W. and Schrenzel,J. (2005) A generic approach for the design of whole-genome oligoarrays, validated for genotyping, deletion mapping and gene expression analysis on *Staphylococcus aureus*. *BMC Genomics*, **6**, 95.
40. Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I., Cui,L., Oguchi,A., Aoki,K., Nagai,Y. et al. (2001) Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. *Lancet*, **357**, 1225–1240.
41. Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A., Nagai,Y., Iwama,N., Asano,K., Naimi,T. et al. (2002) Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet*, **359**, 1819–1827.
42. Gill,S.R., Fouts,D.E., Archer,G.L., Mongodin,E.F., Deboy,R.T., Ravel,J., Paulsen,I.T., Kolonay,J.F., Brinkac,L., Beanan,M. et al. (2005) Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. *J. Bacteriol.*, **187**, 2426–2438.
43. Diep,B.A., Gill,S.R., Chang,R.F., Phan,T.H., Chen,J.H., Davidson,M.G., Lin,F., Lin,J., Carleton,H.A., Mongodin,E.F. et al. (2006) Complete genome sequence of USA300, an epidemic clone of community-acquired methicillin-resistant *Staphylococcus aureus*. *Lancet*, **367**, 731–739.
44. Holden,M.T., Feil,E.J., Lindsay,J.A., Peacock,S.J., Day,N.P., Enright,M.C., Foster,T.J., Moore,C.E., Hurst,L., Atkin,R. et al. (2004) Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. *Proc. Natl Acad. Sci. USA*, **101**, 9786–9791.
45. Scherl,A., Francois,P., Charbonnier,Y., Deshusses,J.M., Koessler,T., Huyghe,A., Bento,M., Stahl-Zeng,J., Fischer,A., Masselot,A. et al. (2006) Exploring glycopeptide-resistance in *Staphylococcus aureus*: a combined proteomics and transcriptomics approach for the identification of resistance-related markers. *BMC Genomics*, **7**, 296.
46. Renzoni,A., Barras,C., Francois,P., Charbonnier,Y., Huggler,E., Garzoni,C., Kelley,W.L., Majcherzyk,P., Schrenzel,J., Lew,D.P. et al. (2006) Transcriptomic and functional analysis of an autolysis-deficient, teicoplanin-resistant derivative of methicillin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.*, **50**, 3048–3061.
47. Churchill,G.A. (2004) Using ANOVA to analyze microarray data. *BioTechniques*, **37**, 173–175, 177.
48. Schagger,H. and von Jagow,G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of

18 *Nucleic Acids Research*, 2009

- proteins in the range from 1 to 100 kDa. *Anal. Biochem.*, **166**, 368–379.
49. de Hoon,M.J., Makita,Y., Nakai,K. and Miyano,S. (2005) Prediction of transcriptional terminators in *Bacillus subtilis* and related species. *PLoS Comput. Biol.*, **1**, e25.
50. Schattner,P. (2002) Searching for RNA genes using base-composition statistics. *Nucleic Acids Res.*, **30**, 2076–2082.
51. Cros,M.J., Moisan,A., Cierco-Ayrolles,C. and Gaspin,C. Visualizing and exploring genomic information for non-protein-coding RNA identification using ApolloRNA. *Nature Protocols*, doi: 10.1036/nprot.20.07.285.
52. Pané-Farré,J., Jonas,B., Forstner,K., Engelmann,S. and Hecker,M. (2006) The  $\sigma^B$  regulon in *Staphylococcus aureus* and its regulation. *Int. J. Med. Microbiol.*, **296**, 237–258.
53. Pané-Farré,J., Jonas,B., Hardwick,S.W., Gronau,K., Lewis,R.J., Hecker,M. and Engelmann,S. (2009) Role of RsbU in controlling SigB activity in *Staphylococcus aureus* following alkaline stress. *J. Bacteriol.*, **191**, 2561–2573.
54. Baba,T., Bae,T., Schneewind,O., Takeuchi,F. and Hiramatsu,K. (2008) Genome sequence of *Staphylococcus aureus* strain Newman and comparative analysis of staphylococcal genomes: polymorphism and evolution of two major pathogenicity islands. *J. Bacteriol.*, **190**, 300–310.
55. Padalon-Brauch,G., Hershberg,R., Elgrably-Weiss,M., Baruch,K., Rosenshine,I., Margalit,H. and Altuvia,S. (2008) Small RNAs encoded within genetic islands of *Salmonella typhimurium* show host-induced expression and role in virulence. *Nucleic Acids Res.*, **36**, 1913–1927.
56. Horsburgh,M.J., Aish,J.L., White,I.J., Shaw,L., Lithgow,J.K. and Foster,S.J. (2002)  $\sigma^B$  modulates virulence determinant expression and stress resistance: characterization of a functional *rsbU* strain derived from *Staphylococcus aureus* 8325-4. *J. Bacteriol.*, **184**, 5457–5467.
57. Pohl,K., François,P., Stenz,L., Schlink,F., Geiger,T., Herbert,S., Goerke,C., Schrenzel,J. and Wolz,C. (2009) CodY in *Staphylococcus aureus*: a regulatory link between metabolism and virulence gene expression. *J. Bacteriol.*, **191**, 2953–2963.
58. Shearwin,K.E., Callen,B.P. and Egan,J.B. (2005) Transcriptional interference—a crash course. *Trends Genet.*, **21**, 339–345.
59. Britton,R.A., Wen,T., Schaefer,L., Pellegrini,O., Uicker,W.C., Mathy,N., Tobin,C., Daou,R., Szyk,J. and Condon,C. (2007) Maturation of the 5' end of *Bacillus subtilis* 16S rRNA by the essential ribonuclease YkqC/RNase J1. *Mol. Microbiol.*, **63**, 127–138.
60. Langbein,I., Bachem,S. and Stulke,J. (1999) Specific interaction of the RNA-binding domain of the *Bacillus subtilis* transcriptional antiterminator GltT with its RNA target, RAT. *J. Mol. Biol.*, **293**, 795–805.
61. Aymerich,S. and Steinmetz,M. (1992) Specificity determinants and structural features in the RNA target of the bacterial antiterminator proteins of the BglG/SacY family. *Proc. Natl Acad. Sci. USA*, **89**, 10410–10414.
62. Jossinet,F. and Westhof,E. (2005) Sequence to Structure (S2S): display, manipulate and interconnect RNA data from sequence to structure. *Bioinformatics*, **21**, 3320–3321.
63. Aiba,H. (2007) Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr. Opin. Microbiol.*, **10**, 134–139.
64. Mathy,N., Benard,L., Pellegrini,O., Daou,R., Wen,T. and Condon,C. (2007) 5'-to-3' exonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell*, **129**, 681–692.
65. Even,S., Pellegrini,O., Zig,L., Labas,V., Vinh,J., Brechemmier-Baey,D. and Putzer,H. (2005) Ribonucleases J1 and J2: two novel endoribonucleases in *B. subtilis* with functional homology to *E. coli* RNase E. *Nucleic Acids Res.*, **33**, 2141–2152.
66. de la Sierra-Gallay,I.L., Zig,L., Jamalli,A. and Putzer,H. (2008) Structural insights into the dual activity of RNase J. *Nat. Struct. Mol. Biol.*, **15**, 206–212.
67. Mader,U., Zig,L., Kretschmer,J., Homuth,G. and Putzer,H. (2008) mRNA processing by RNases J1 and J2 affects *Bacillus subtilis* gene expression on a global scale. *Mol. Microbiol.*, **70**, 183–196.
68. Deikus,G. and Bechhofer,D.H. (2007) Initiation of decay of *Bacillus subtilis* *trp* leader RNA. *J. Biol. Chem.*, **282**, 20238–20244.
69. Yao,S., Blaustein,J.B. and Bechhofer,D.H. (2007) Processing of *Bacillus subtilis* small cytoplasmic RNA: evidence for an additional endonuclease cleavage site. *Nucleic Acids Res.*, **35**, 4464–4473.
70. O'Riordan,K. and Lee,J.C. (2004) *Staphylococcus aureus* capsular polysaccharides. *Clinical Microbiol. Rev.*, **17**, 218–234.
71. Somerville,G.A. and Proctor,R.A. (2009) At the crossroads of bacterial metabolism and virulence synthesis in Staphylococci. *Microbiol. Mol. Biol. Rev.*, **73**, 233–248.
72. Tafer,H. and Hofacker,I.L. (2008) RNAPlex: a fast tool for RNA-RNA interaction search. *Bioinformatics*, **24**, 2657–2663.
73. Mandin,P., Repoila,F., Vergassola,M., Geissmann,T. and Cossart,P. (2007) Identification of new noncoding RNAs in *Listeria monocytogenes* and prediction of mRNA targets. *Nucleic Acids Res.*, **35**, 962–974.
74. Hüttenhofer,A. and Vogel,J. (2006) Experimental approaches to identify non-coding RNAs. *Nucleic Acids Res.*, **34**, 635–646.
75. Eddy,S.R. (2002) Computational genomics of noncoding RNA genes. *Cell*, **109**, 137–140.
76. Fozo,E.M., Hemm,M.R. and Storz,G. (2008) Small toxic proteins and the antisense RNAs that repress them. *Microbiol. Mol. Biol. Rev.*, **72**, 579–589.
77. Barrick,J.E. and Breaker,R.R. (2007) The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol.*, **8**, R239.
78. Yao,Z., Barrick,J., Weinberg,Z., Nepf,S., Breaker,R., Tompa,M. and Ruzzo,W.L. (2007) A computational pipeline for high-throughput discovery of cis-regulatory noncoding RNA in prokaryotes. *PLoS Comput. Biol.*, **3**, e126.
79. Vitreschak,A.G., Mironov,A.A., Lyubetsky,V.A. and Gelfand,M.S. (2008) Comparative genomic analysis of T-box regulatory systems in bacteria. *RNA*, **14**, 717–735.
80. Gutierrez-Preciado,A., Henkin,T.M., Grundy,F.J., Yanofsky,C. and Merino,E. (2009) Biochemical features and functional implications of the RNA-based T-box regulatory mechanism. *Microbiol. Mol. Biol. Rev.*, **73**, 36–61.
81. Knezevic,I., Bachem,S., Sickmann,A., Meyer,H.E., Stulke,J. and Hengstenberg,W. (2000) Regulation of the glucose-specific phosphotransferase system (PTS) of *Staphylococcus carnosus* by the antiterminator protein GltT. *Microbiology*, **146**, 2333–2342.
82. Lu,Y., Turner,R.J. and Switzer,R.L. (1996) Function of RNA secondary structures in transcriptional attenuation of the *Bacillus subtilis* *pyr* operon. *Proc. Natl Acad. Sci. USA*, **93**, 14462–14467.
83. Choonce,N., Even,S., Zig,L. and Putzer,H. (2007) Ribosomal protein L20 controls expression of the *Bacillus subtilis* *infC* operon via a transcription attenuation mechanism. *Nucleic Acids Res.*, **35**, 1578–1588.
84. Balbontin,R., Figueroa-Bossi,N., Casades,J. and Bossi,L. (2008) Insertion hot spot for horizontally acquired DNA within a bidirectional small-RNA locus in *Salmonella enterica*. *J. Bacteriol.*, **190**, 4075–4078.
85. Sridhar,J. and Rafi,Z.A. (2007) Identification of novel genomic islands associated with small RNAs. *In Silico Biol.*, **7**, 601–611.
86. Toledo-Arana,A., Dussurget,O., Nikitas,G., Sesto,N., Guet-Revillet,H., Balestrino,D., Loh,E., Gripenland,J., Tiensuu,T., Vaitkevicius,K. et al. (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature*, **459**, 950–956.
87. Nielsen,J.S., Olsen,A.S., Bonde,M., Valentin-Hansen,P. and Kallipolitis,B.H. (2008) Identification of a sigma B-dependent small noncoding RNA in *Listeria monocytogenes*. *J. Bacteriol.*, **190**, 6264–6270.
88. Kullik,I.I. and Giachino,P. (1997) The alternative sigma factor sigmaB in *Staphylococcus aureus*: regulation of the *sigB* operon in response to growth phase and heat shock *Arch. Microbiol.*, **167**, 151–159.
89. Morikawa,K., Inose,Y., Okamura,H., Maruyama,A., Hayashi,H., Takeyasu,K. and Ohta,T. (2003) A new staphylococcal sigma factor in the conserved gene cassette: functional significance and implication for the evolutionary processes. *Genes Cells*, **8**, 699–712.
90. Bischoff,M., Dunman,P., Kormanec,J., Macapagal,D., Murphy,E., Mounts,W., Berger-Bachi,B. and Projan,S. (2004) Microarray-based analysis of the *Staphylococcus aureus*  $\sigma^B$  regulon. *J. Bacteriol.*, **186**, 4085–4099.

91. Ziebandt, A.K., Becher, D., Ohlsen, K., Hacker, J., Hecker, M. and Engelmann, S. (2004) The influence of *agr* and  $\sigma^3$  in growth phase dependent regulation of virulence factors in *Staphylococcus aureus*. *Proteomics*, **4**, 3034–3047.
92. Sharp, J.S. and Bechhofer, D.H. (2005) Effect of 5'-proximal elements on decay of a model mRNA in *Bacillus subtilis*. *Mol. Microbiol.*, **57**, 484–495.
93. Condon, C. (2007) Maturation and degradation of RNA in bacteria. *Curr. Opin. Microbiol.*, **10**, 271–278.
94. Papenfort, K., Pfeiffer, V., Lucchini, S., Sonawane, A., Hinton, J.C. and Vogel, J. (2008) Systematic deletion of *Salmonella* small RNA genes identifies CyaR, a conserved CRP-dependent riboregulator of OmpX synthesis. *Mol. Microbiol.*, **68**, 890–906.
95. Gaballa, A., Antelmann, H., Aguilar, C., Khakh, S.K., Song, K.B., Smaldone, G.T. and Helmann, J.D. (2008) The *Bacillus subtilis* iron-sparing response is mediated by a Fur-regulated small RNA and three small, basic proteins. *Proc. Natl Acad. Sci. USA*, **105**, 11927–11932.
96. Chang, B., Halgamuge, S. and Tang, S.L. (2006) Analysis of SD sequences in completed microbial genomes: non-SD-led genes are as common as SD-led genes. *Gene*, **373**, 90–99.
97. Hiron, A., Borezee-Durant, E., Piard, J.C. and Juillard, V. (2007) Only one of four oligopeptide transport systems mediates nitrogen nutrition in *Staphylococcus aureus*. *J. Bacteriol.*, **189**, 5119–5129.
98. Kohler, C., Wolff, S., Albrecht, D., Fuchs, S., Becher, D., Buttner, K., Engelmann, S. and Hecker, M. (2005) Proteome analyses of *Staphylococcus aureus* in growing and non-growing cells: a physiological approach. *Int. J. Med. Microbiol.*, **295**, 547–565.
99. Somerville, G.A., Said-Salim, B., Wickman, J.M., Raffel, S.J., Kreiswirth, B.N. and Musser, J.M. (2003) Correlation of acetate catabolism and growth yield in *Staphylococcus aureus*: implications for host-pathogen interactions. *Infect. Immun.*, **71**, 4724–4732.
100. Baba, T., Kuwahara-Arai, K., Uchiyama, I., Takeuchi, F., Ito, T. and Hiramatsu, K. (2009) Complete genome sequence of *Micrococcus caseolyticus* strain JSCS5402, reflecting the ancestral genome of the human-pathogenic staphylococci. *J. Bacteriol.*, **191**, 1180–1190.
101. Møller, T., Franch, T., Hojrup, P., Keene, D.R., Bachinger, H.P., Brennan, R.G. and Valentin-Hansen, P. (2002) Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol. Cell*, **9**, 23–30.
102. Goerke, C., Fluckiger, U., Steinhuber, A., Bisanzio, V., Ulrich, M., Bischoff, M., Patti, J.M. and Wolz, C. (2005) Role of *Staphylococcus aureus* global regulators *sae* and *sigmaB* in virulence gene expression during device-related infection. *Infect. Immun.*, **73**, 3415–3421.
103. Marzi, S., Fechter, P., Chevalier, C., Romby, P. and Geissmann, T. (2008) RNA switches regulate initiation of translation in bacteria. *Biol. Chem.*, **389**, 585–598.
104. Sittka, A., Sharma, C.M., Rolle, K. and Vogel, J. (2009) Deep sequencing of *Salmonella* RNA associated with heterologous Hfq proteins in vivo reveals small RNAs as a major target class and identifies RNA processing phenotypes. *RNA Biol.*, **6**, 266–275.
105. Kreiswirth, B.N., Lofdahl, S., Betley, M.J., O'Reilly, M., Schlievert, P.M., Bergdoll, M.S. and Novick, R.P. (1983) The toxic shock syndrome exotoxin structural gene is not detectably transmitted by a prophage. *Nature*, **305**, 709–712.
106. Peng, H.L., Novick, R.P., Kreiswirth, B., Kornblum, J. and Schlievert, P. (1988) Cloning, characterization, and sequencing of an accessory gene regulator (*agr*) in *Staphylococcus aureus*. *J. Bacteriol.*, **170**, 4365–4372.
107. Kornblum, J., Kreiswirth, B., Projan, S.J., Ross, H. and Novick, R. (1990) *Agr*: a polycistronic locus regulating exoprotein synthesis in *Staphylococcus aureus*. In Novick, R. (ed.), *Molecular Biology of the staphylococci*. Wiley-VCH, New York, pp. 373–401.
108. Dyke, K.G., Jevons, M.P. and Parker, M.T. (1966) Penicillinase production and intrinsic resistance to penicillins in *Staphylococcus aureus*. *Lancet*, **1**, 835–838.
109. Duthie, E.S. and Lorenz, L.L. (1952) Staphylococcal coagulase: mode of action and antigenicity. *J. Gen. Microbiol.*, **6**, 95–107.

## SUPPLEMENTARY MATERIAL

### MATERIALS AND METHODS

#### Cell culture

One colony of *S. aureus* (RN6390, COL) was grown overnight at 37°C in 2 to 3 ml of medium, either in brain-heart infusion (BHI), or in NZM medium (10 g/l Casein Tryptone digested; 5 g/l NaCl; 2 g/l MgSO<sub>4</sub>-7H<sub>2</sub>O; pH 7.0). From this pre-culture, 200 µl was placed in 10 to 20 ml of medium (BHI or NZM) and incubated at 37°C for 2 h (OD 600 nm 0.5 in BHI, and 0.2 in NZM, exponential phase) and for 6 h (OD 600 nm 5 in BHI, and 2.5 in NZM, late-exponential phase).

#### Preparation of the genomic DNA

DNA was prepared from 2 ml of RN6390 or COL strains grown in BHI medium for 4 h at 37°C. The cells were centrifuged (3273 g, 4°C, 10 min) and the pellet was dissolved gently in 200 µl of lysis buffer (20 mM Tris-HCl pH 8; 2 mM EDTA; 1.2% Triton X-100). Cell lysis was performed in the presence of 2 µl of lysostaphin 2.5 mg/ml at 37°C for 30 min. The genomic DNA was purified using the DNeasy Tissue kit following the protocol of the company. The genomic DNA was kept at -20°C in a commercial buffer.

#### Preparation of total RNAs and 3' end-labeling

Total RNAs were prepared from 20 ml of *S. aureus* grown in BHI or NZM medium as described above. After centrifugation of the cells (3273 g, 4°C, 10 min), total RNAs were extracted using the fastRNA Pro Blue kit (Qbiogen, France), and kept at -20°C at a concentration of 1 µg/µl in sterile water. The RNAs (20 µg) were purified by electrophoresis on a 6% polyacrylamide- 8 M urea gel to remove the highly abundant ribosomal RNAs. Bands corresponding to 50-600 nt long RNAs were cut into pieces and the RNAs were eluted at 4°C overnight in 1.5 ml of elution buffer (0.5 M ammonium acetate; 1 mM EDTA; 0.1% SDS). The RNAs were precipitated with 2.5 vol of ethanol and washed in 80% ethanol. The purified RNAs were kept in sterile water at -20°C at a concentration of 5 µg/µl.

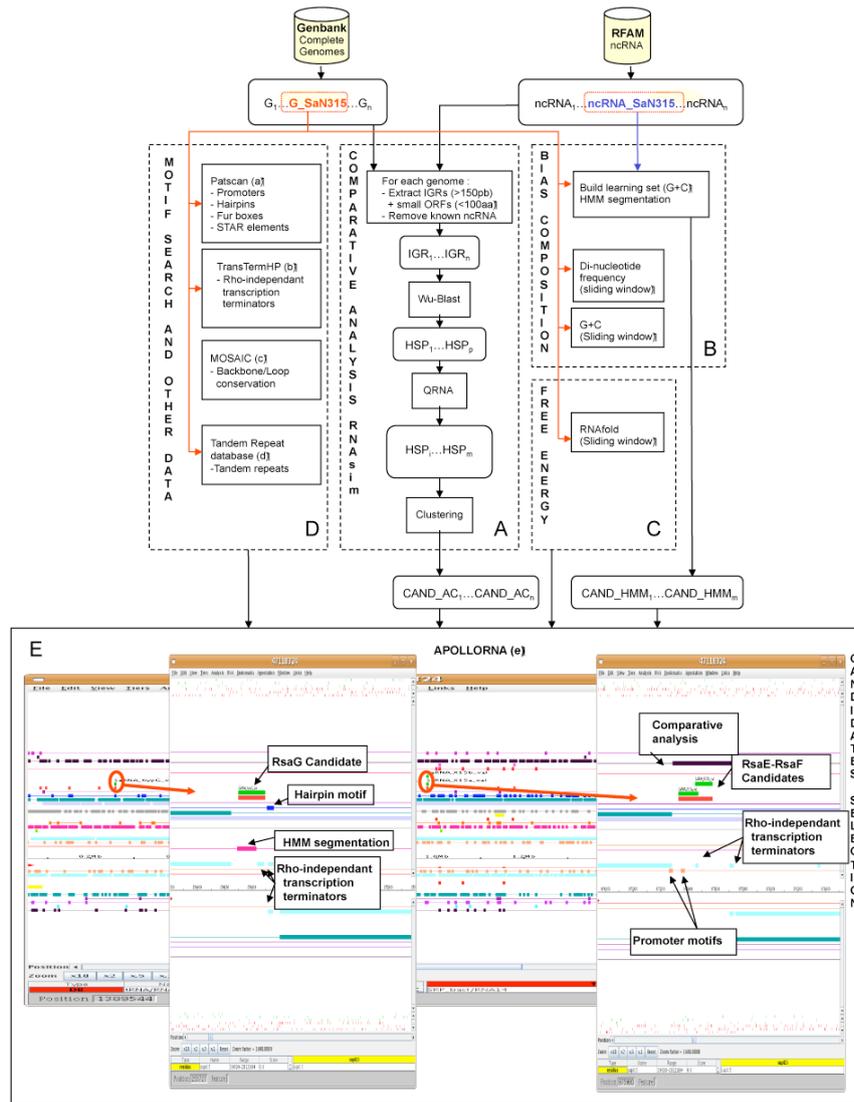
The RNAs were labeled at their 3' end using T4 RNA ligase and [ $\alpha$ - $^{32}$ P]pCp. The reaction was done at 4°C overnight in 20  $\mu$ l containing 10  $\mu$ g total RNAs, 2  $\mu$ l of commercial buffer 10x, 5 U of T4 RNA ligase (Ambion) and 5  $\mu$ l [ $\alpha$ - $^{32}$ P]pCp (3000 Ci/mmol, 10 mCi/ml, Amersham). Labeled RNAs were purified on a gel filtration column (Micro-Bio Spin P-6, Bio-Rad) to remove the excess of pCp. The quality of the RNAs was controlled by electrophoresis on a 8% polyacrylamide- 8 M urea gel. Labeled RNAs were kept at -20°C.

### **Home-made macroarrays**

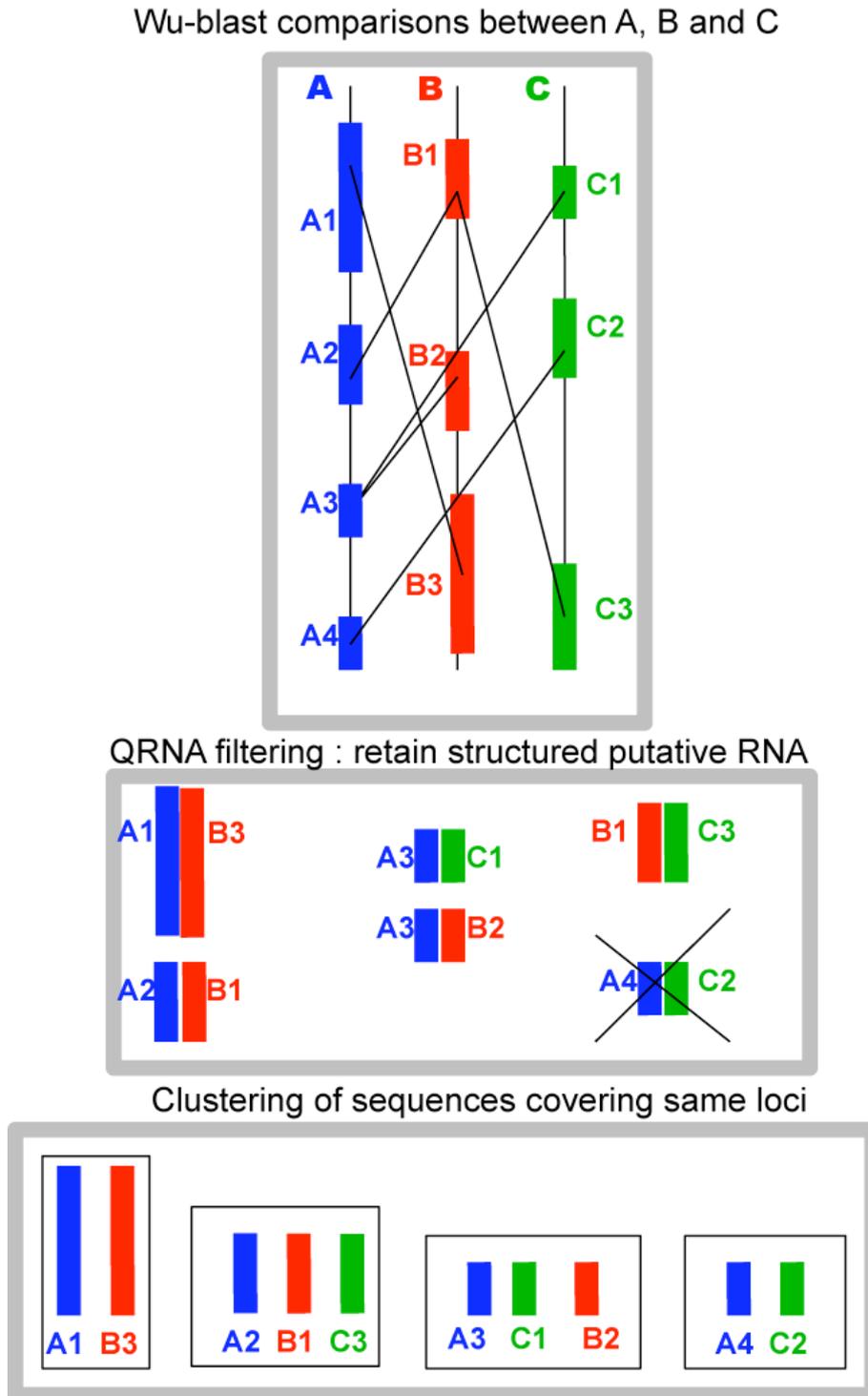
Macroarrays were done with dsDNAs corresponding to the 36 selected IGRs. The IGRs were PCR amplified from *S. aureus* RN6390 genomic DNA using a set of oligonucleotides (Table S4). The PCR reaction was carried out in a total volume of 50  $\mu$ l containing 5  $\mu$ l of commercial buffer 10x, 100  $\mu$ M dNTP mix, 20 ng genomic DNA, 1 U of Taq DNA polymerase, and 50 pmol of each oligonucleotide. The hybridization temperature varies from 45 to 55°C depending on the stability of the duplexes. After 35 cycles of 30 sec denaturation at 94°C, 30 sec hybridization at 45-55°C, and 1 min elongation at 72°C, the quality of the amplified dsDNA fragments was determined by electrophoresis on a 1% agarose gel. Then, 5  $\mu$ l of a PCR reaction were mixed in 5  $\mu$ l of a solution containing 95 % formamide, 18 mM EDTA, and 0.025% SDS, and the dsDNA was denatured 3 min at 90°C followed by cooling on ice. The denatured DNA (2  $\mu$ l) was then spotted in duplicate on a nitrocellulose membrane and dried. The membrane was incubated 1 h at 42°C in the hybridization buffer (50 mM sodium phosphate pH 7.2, 7 % SDS, 1 mM EDTA). The hybridization buffer was changed and the 3' end-labeled total RNAs, which were previously denatured, were added to the membrane. Hybridization was performed at 42°C overnight. The membrane was washed twice in buffer containing 40 mM sodium phosphate, 5 % SDS, and 1 mM EDTA for 15 min at 42°C. The radioactive spots were visualized by autoradiography of the hybridized membrane, which was previously exposed overnight.

**REFERENCES**

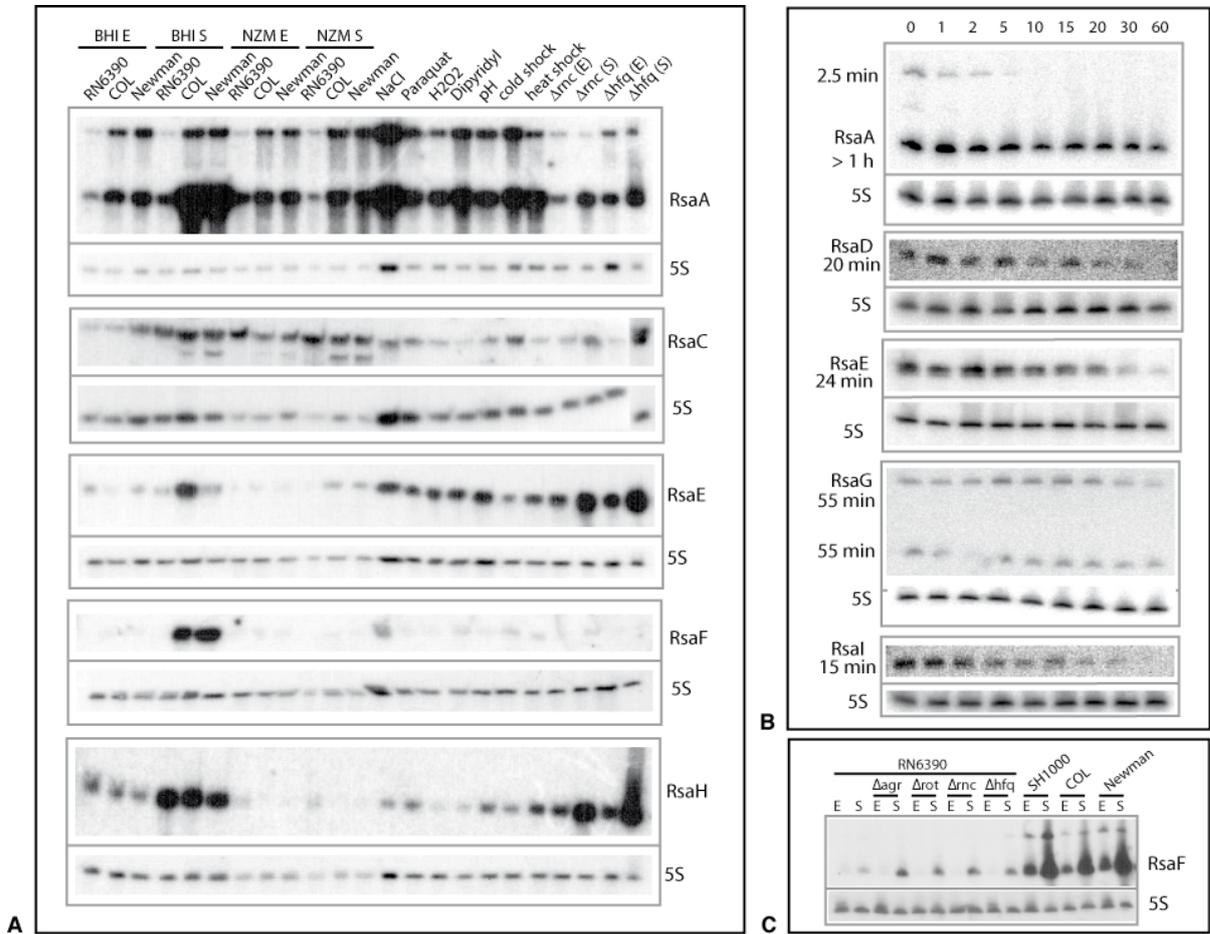
1. Lu, Y., Turner, R.J. and Switzer, R.L. (1996) Function of RNA secondary structures in transcriptional attenuation of the *Bacillus subtilis* *pyr* operon. *Proc. Natl. Acad. Sci. USA*, **93**, 14462-14467.
2. Schuster, P., Fontana, W., Stadler, P.F. and Hofacker, I.L. (1994) From sequences to shapes and back: a case study in RNA secondary structures. *Proc Biol Sci*, **255**, 279-284.
3. Bischoff, M., Dunman, P., Kormanec, J., Macapagal, D., Murphy, E., Mounts, W., Berger-Bachi, B. and Projan, S. (2004) Microarray-based analysis of the *Staphylococcus aureus*  $\sigma^B$  regulon. *J. Bacteriol.*, **186**, 4085-4099.
4. Pané-Farré, J., Jonas, B., Forstner, K., Engelmann, S. and Hecker, M. (2006) The  $\sigma^B$  regulon in *Staphylococcus aureus* and its regulation. *Int. J. Med. Microbiol.*, **296**, 237-258.
5. Livny, J., Teonadi, H., Livny, M. and Waldor, M.K. (2008) High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. *PLoS ONE*, **3**, e3197.
6. Pichon, C. and Felden, B. (2005) Small RNA genes expressed from *Staphylococcus aureus* genomic and pathogenicity islands with specific expression among pathogenic strains. *Proc. Natl. Acad. Sci. USA*, **102**, 14249-14254.
7. Roberts, C., Anderson, K.L., Murphy, E., Projan, S.J., Mounts, W., Hurlburt, B., Smeltzer, M., Overbeek, R., Disz, T. and Dunman, P.M. (2006) Characterizing the effect of the *Staphylococcus aureus* virulence factor regulator, SarA, on log-phase mRNA half-lives. *J. Bacteriol.*, **188**, 2593-2603.
8. Marchais, A., Naville, M., Bohn, C., Bouloc, P. and Gautheret, D. (2009) Single-pass classification of all non-coding sequences in a bacterial genome using phylogenetic profiles. *Genome Res.*, 10.1101/gr.089714.108.



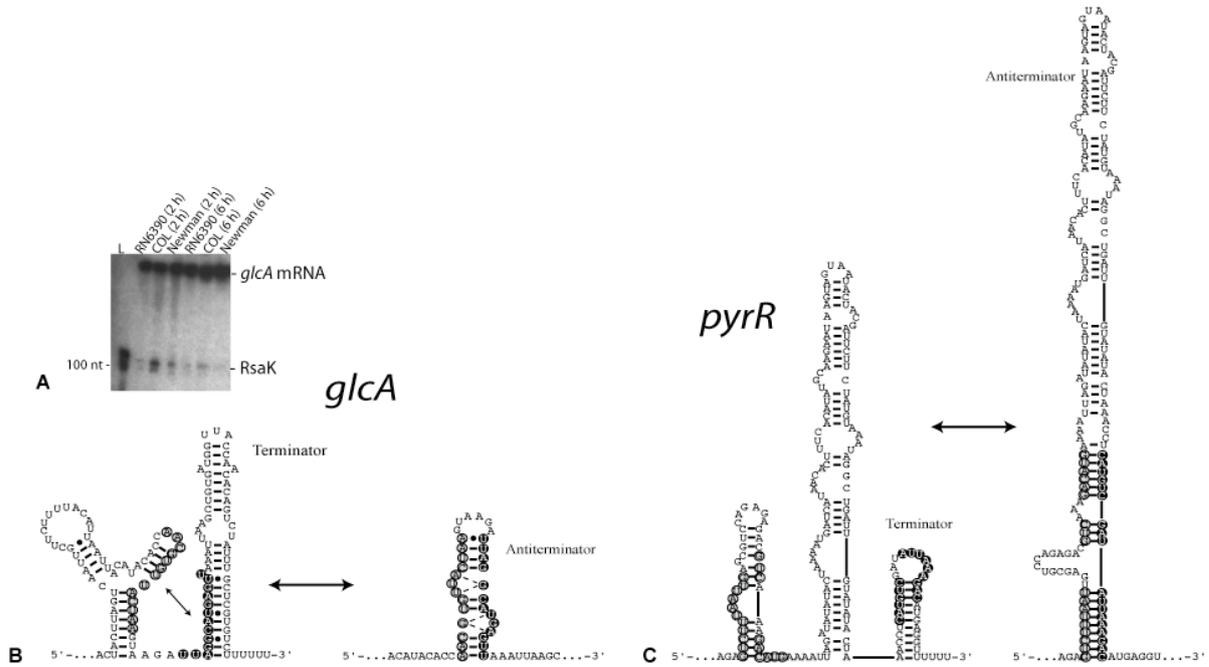
**Figure S1.** Strategy used to identify the functional intergenic regions of the *Staphylococcus aureus* N315 genome. Our strategy integrates data generated from complementary approaches (A-C) and data obtained from different sources (D) with the goal to select the most promising candidates. (A) Description of the comparative analysis implemented in a pipeline called RNAsim. Only clusters containing at least one IGR of *S. aureus* N315 genome longer than 150 nt were kept for further analysis. (B) The biased composition analysis relies on three approaches (HMM, di-nucleotide and G+C frequencies). A simple Hidden Markov Model (HMM) was used as an automatic segmentation approach, and was computed by considering known ncRNAs and riboswitches available in RFAM and Genbank databases. Di-nucleotides, G+C frequencies (B) and free energy values (C) were computed on a 90 nt sliding window along the genome. (D) Other data (promoters, STAR, different hairpin motifs) were computed using PatScan. Rho-independent terminators were obtained using TransTermHP. MOSAIC database was used to identify conserved and strain-specific regions among available *S. aureus* genomes. Tandem repeats were extracted from the Tandem Repeat database. (D) The data and the quantitative variables (free energy, G+C content) were integrated into ApolloRNA to select the IGR for functional analysis. Two examples are illustrated: RsaG was selected based on the HMM segmentation, the presence of a promoter and a Rho-independent terminator, and the orientation of the flanking genes; RsaE and RsaF were selected based on RNAsim, the presence of promoters and Rho-independent terminators, and the orientation of the flanking genes. (a-e) web sites from which softwares and/or data were obtained : (a) PatScan (<http://www-new.mcs.anl.gov/compbio/PatScan/HTML/scanner.html>), (b) TransTermHP (<http://transterm.cbcb.umd.edu/index.php>), (c) MOSAIC (<http://genome.jouy.inra.fr/mosaic>), (d) the tandem Repeat Database (<http://minisatellites.u-psud.fr>), (e) ApolloRNA (<http://carlit.toulouse.inra.fr/ApolloRNA/index.html>).



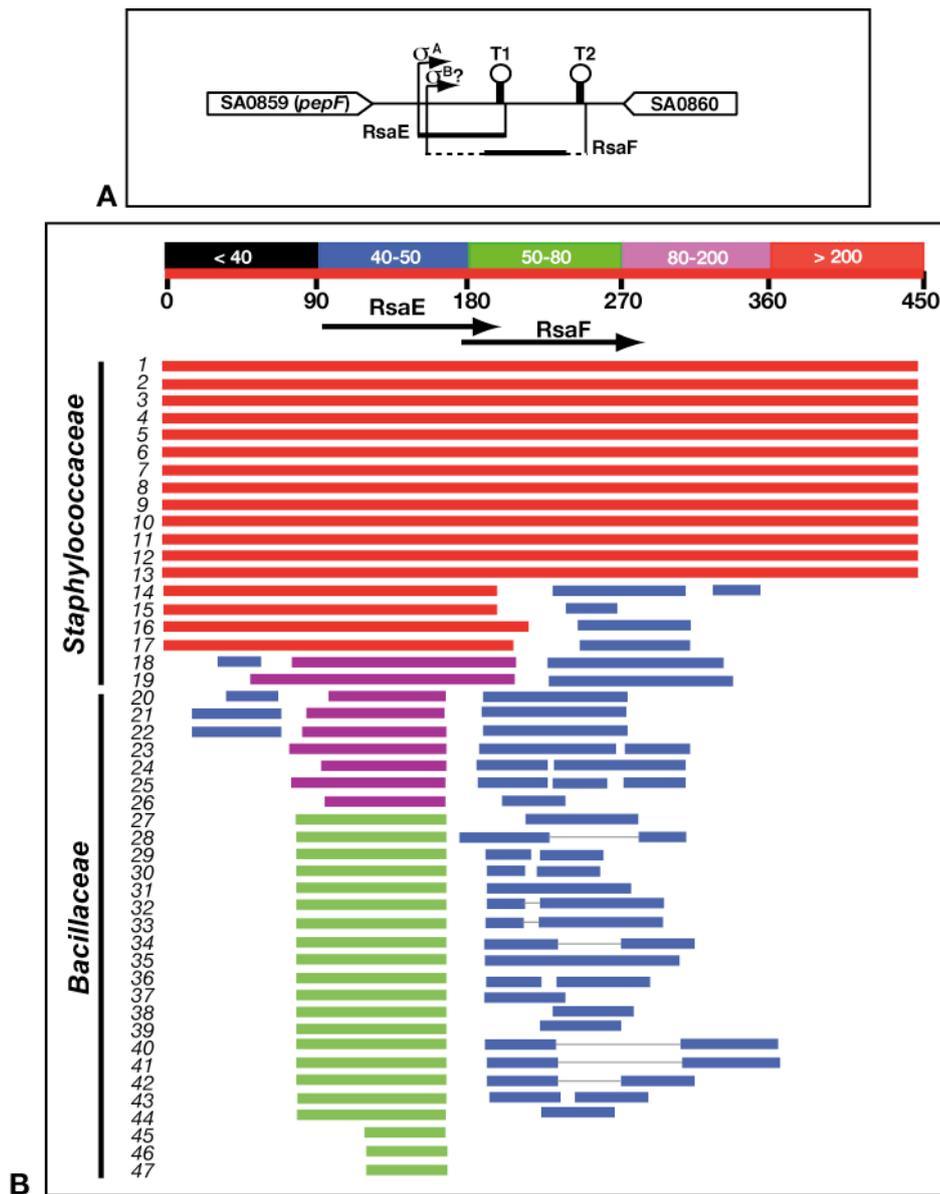
**Figure S2.** Strategy implemented in RNAsim. (A) Wu-blast comparisons are realized between intergenic regions of genomes A, B and C (including *S. aureus* strains, Staphylococci or firmicutes). (B) QRNA is used to filter pairwise alignments in order to retain only putative structured RNAs. In this example, the alignment between A4 and C2 is removed from the set of pairwise alignments. (C) Pairwise alignments are clustered when at least one sequence in an alignment overlaps one of the two sequences in another alignment. In this example, A2 aligns with B1 and B1 aligns with C3. A2, B1 and C3 are thus clustered as well as A3, B2 and C1.



**Figure S3.** Analysis of Rsa RNA expression. (B) Northern blot analysis of RsaA, RsaC, RsaE, RsaF, and RsaH in various *S. aureus* strains (RN6390, COL, and Newman), LUG774 (RN6390- $\Delta rnc$ ), and LUG911 (RN6390- $\Delta hfq$ ) grown in BHI (rich medium) or NZM (stringent medium) at the exponential (E) and stationary (S) phases of growth. COL strain was grown under various stress conditions: osmotic stress (NaCl), oxidative stress ( $H_2O_2$ , paraquat), iron chelating agent (dipyridyl), acidic pH, cold shock (25°C), and heat shock (42°C). (B) Determination of the half-lives of RsaA, RsaD, RsaE, RsaG, and RsaI. The cells were treated with rifampicin and total RNAs were extracted after 1, 2, 5, 10, 20, 30, and 60 min at 37°C in BHI medium. 5S rRNA was probed to quantify the yield of RNA in each lane. (C) Northern blot analysis of RsaF in *S. aureus* strains: RN6390, RN6911 ( $\Delta agr$ ), LUG1160 (RN6390- $\Delta rot$ , *rot* encodes the repressor of toxins), LUG911 (RN6390- $\Delta hfq$ ), SH1000, COL, and Newman. Total RNAs were prepared from cells grown in BHI medium at the exponential (E) and stationary (S) phases of growth.

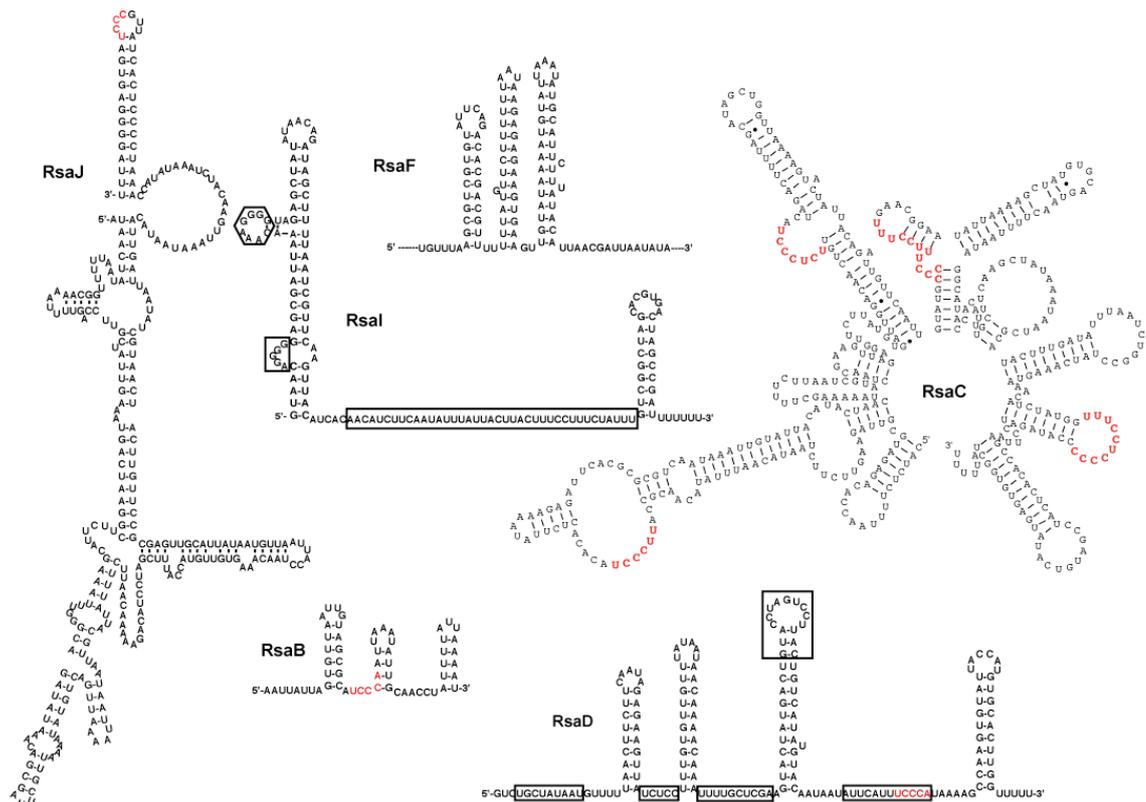


**Figure S4.** Examples of conserved *cis*-regulatory leader regions of mRNAs in *S. aureus*. (A) Northern blot analysis of RsaK in *S. aureus* strains (RN6390, COL, and Newman). Total RNAs were prepared from cells grown in BHI at the exponential (2 h) or stationary (6 h) phases of growth. Lane L: RNA ladder (Ambion). (B) Sequence of *S. aureus* RsaK corresponding to the leader region of *glcA* mRNA. The leader contains a conserved RAT (ribonucleic antiterminator) motif which is recognized by the glucose sensing regulatory protein GlcT. The stabilization of this structure allows the RNA polymerase to proceed transcription of *glcA* mRNA. (C) The regulatory leader regions of *S. aureus* *pyrR* and *pyrP* mRNAs. The two regions contain a AGAGAG hairpin loop which was proposed to be the main recognition element for the regulatory protein PyrR in *B. subtilis* (1). In the presence of UMP, PyrR binds to the target mRNAs, and favors the formation of a terminator to induce premature termination of transcription. Alternative pairings leading to the formation of a terminator or of an anti-terminator structure are schematized by encircled grey and black residues.



**Figure S5.** Genomic organization and sequence conservation of RsaE among *Staphylococcaceae* and *Bacillaceae*. (A) The *rsaE-rsaF* locus. RsaE is made from an autonomous IGR which contains a  $\sigma^A$ -consensus binding site located at the appropriate distance from the start site of *rsaE*, and a typical Rho-independent terminator of transcription (T1). In contrast to RsaE, RsaF is only present at high levels in strains expressing a fully active  $\sigma^B$  factor. A divergent  $\sigma^B$ -consensus binding site was predicting overlapping the  $\sigma^A$ - promoter. The size of RsaF (104 nts) was evaluated from Northern blots and the 5' start was determined by RACE. The presence of a longer RNA species on Northern blots suggests that RsaF may be the result of an endoribonuclease processing event. It remains to be determined whether the *S. aureus*-specific RsaF is functional, and the significance of the differential expression pattern of the two Rsa RNAs. (B) A schematic view of the sequence alignments of the intergenic region containing *rsaE* and *rsaF* from various *Staphylococcaceae* and *Bacillaceae*: 1-13, *S. aureus* Mu3, JH1, JH9, Mu50, MSSA476, MW2, USA300\_TCH1516, Newman, NCTC8325, USA300\_FPR3757, COL, MRSA252, and RF122, respectively; 14-15, *S. epidermidis* ATCC12228 and RP62A, respectively; 16, *S. haemolyticus* JCSC1435; 17, *S.*

*saprophyticus* ATCC 15305; 18, *S. carnosus* TM300; 19, *Micrococcus caseolyticus* JCSC5402; 20, *Bacillus pumilus* SAFR-032; 21-22, *B. licheniformis* ATCC 14580 and DSM 13, respectively; 23, *B. subtilis* 168; 24, *B. amyloliquefaciens* FZB42; 25, *Anoxybacillus flavithermus* WK1; 26, *Lysinibacillus sphaericus* C3-41; 27-28, *B. anthracis* A0248 and CDC 684; 29-34, *B. cereus* 03BB102, Q1, G9842, AH820, B4264, and AH187, respectively; 35, *B. weihenstephanensis* KBAB4; 36-37 and 41-42, *B. cereus* cytotoxis NVH, ATCC 14579, ATCC 10987 and E33L, respectively; 38 and 40, *B. thuringiensis* Al Hakam and serovar konkukian str. 97-27, respectively; 43-44, *B. anthracis* Ames Ancestor and Ames, respectively; 45, *Geobacillus* sp. WCH70; 46, *G. thermodenitrificans* NG80-2; 47, *G. kaustophilus* HTA426. The alignment was performed using BlastN (<http://www.ncbi.nlm.nih.gov/blast/>). The five colors represent the alignment scores. The red line below the color code corresponds to the whole IGR comprises *rsaE-rsaF* of *S. aureus* N315 used as the reference sequence. The numbering of the nucleotides is given and the black arrows correspond to *rsaE* and *rsaF* genes. Only *rsaE* (except the terminator) shows a high degree of conservation in *Staphylococcaceae* and *Bacillaceae*, and the C-rich sequence motifs are strictly conserved in all strains (Fig. 3).



**Figure S6.** Secondary structure of several Rsa RNAs. The secondary structures were predicted using RNAFold from the Vienna package (2). Strictly conserved nucleotides of RsaD, RsaH, and RsaI in four staphylococcal species (*S. aureus* N315, *S. epidermidis*, *S. saprophyticus*, and *S. haemolyticus*) are boxed. The conserved C-rich boxes are shown in red characters in RsaB, RsaC, RsaD, RsaH, and RsaJ.



**Table S1.** List of predicted conserved IGRs selected by RNAsim.

Start	End	strand	Conservation	Comments
114	510	+	Staphylococcus	Leader <i>dnaA</i>
1956	2155	+	Firmicutes	Leader <i>dnaN</i>
3417	3585	+	<i>S. aureus</i>	
10609	10859	+	Staphylococcus	
12441	12792	+	Firmicutes	T-box upstream of <i>serS</i>
14417	14692	+	Staphylococcus	
15776	16092	+	Firmicutes	S-box-SAM riboswitch
22218	22447	+	Staphylococcus	
24336	24625	+	<i>S. aureus</i>	Livny et al. (5)
30822	30996	+	<i>S. aureus</i>	
35750	36434	+	Staphylococcus	Includes small ORF
39893	40055	-	Firmicutes	Upstream of <i>pre</i> +Term, HMM, C-Box
42881	43180	+	Staphylococcus	HMM, RsaX01
43481	43716	+	Staphylococcus	RsaX01
49266	49722	+	Staphylococcus	
53428	53638	+	Firmicutes	Term upstream of SA0045
54468	54805	+	<i>S. aureus</i>	Livny et al. (5)
55552	56001	+	<i>S. aureus</i>	Cis-acting leader of <i>ermA</i>
62881	63210	-	Staphylococcus	Livny et al. (5)
66426	66629	-	Staphylococcus	Livny et al. (5)
68424	68723	-	Staphylococcus	
71351	71650	-	Firmicutes	Upstream of <i>tnp</i>
72895	73149	-	Firmicutes	Long IGR
76849	77115	+	Staphylococcus	Includes small ORF
81407	81949	+	Staphylococcus	RsaX02, Livny et al. (5)
87180	87668	+	Staphylococcus	
89956	90119	-	Staphylococcus	STAR
91349	91520	+	Staphylococcus	Term upstream of SA0081
95179	95634	+	Staphylococcus	RsaX03, Marchais et al. (8)
96951	97164	+	<i>S. aureus</i>	
97632	97851	+	<i>S. aureus</i>	
98489	98727	+	<i>S. aureus</i>	
102443	102644	+	<i>S. aureus</i>	
103633	103851	+	Staphylococcus	
115157	115385	-	<i>S. aureus</i>	
120468	120672	+	<i>S. aureus</i>	
122442	122644	+	<i>S. aureus</i>	Leader of <i>spa</i> , Roberts et al. (7)
124209	124406	-	<i>S. aureus</i>	
125182	125542	-	Staphylococcus	
141953	142143	+	<i>S. aureus</i>	
143771	143961	+	<i>S. aureus</i>	

147787	148049	+	Staphylococcus	
148698	148998	+	<i>S. aureus</i>	
149745	149927	+	Staphylococcus	RsaX04
150689	150922	+	<i>S. aureus</i>	
158609	158807	+	Staphylococcus	
161970	162120	+	<i>S. aureus</i>	
164074	164428	+	<i>S. aureus</i>	
167221	167408	+	<i>S. aureus</i>	
184882	185158	+	<i>S. aureus</i>	
186732	186949	+	Staphylococcus	
188293	188619	+	Staphylococcus	STAR
189594	189781	+	<i>S. aureus</i>	Term upstream of SA0166
193338	193534	+	<i>S. aureus</i>	
195127	195484	+	<i>S. aureus</i>	
196748	197105	+	<i>S. aureus</i>	
210408	210644	-	<i>S. aureus</i>	
212086	212272	-	<i>S. aureus</i>	
217068	217417	+	Firmicutes	RsaK, leader of <i>glcA</i>
222171	222426	+	<i>S. aureus</i>	
229090	229339	+	<i>S. aureus</i>	
232633	232800	+	<i>S. aureus</i>	
232830	233059	+	<i>S. aureus</i>	
243288	243477	+	<i>S. aureus</i>	
244236	244446	+	<i>S. aureus</i>	
252268	252525	+	<i>S. aureus</i>	
253038	253267	+	Staphylococcus	
258499	258846	+	<i>S. aureus</i>	
264490	264674	+	<i>S. aureus</i>	
275317	275537	+	<i>S. aureus</i>	
276925	277225	+	<i>S. aureus</i>	
278704	278900	-	<i>S. aureus</i>	
279310	279599	-	<i>S. aureus</i>	Small ORF (SAS007)
280772	281015	+	<i>S. aureus</i>	
281088	281255	+	<i>S. aureus</i>	Roberts et al. (7)
284212	284494	+	<i>S. aureus</i>	
285503	285741	+	<i>S. aureus</i>	
292549	292777	+	<i>S. aureus</i>	
296522	296764	+	<i>S. aureus</i>	
297966	298168	+	<i>S. aureus</i>	
304334	304490	+	<i>S. aureus</i>	
311529	311757	+	Staphylococcus	STAR, HMM
312788	312948	-	<i>S. aureus</i>	
315234	315409	-	<i>S. aureus</i>	

316427	316782	+	<i>S. aureus</i>	
325672	325930	-	<i>S. aureus</i>	
326827	327089	+	<i>S. aureus</i>	
341093	341259	+	<i>S. aureus</i>	
345753	346269	+	<i>S. aureus</i>	
348343	348530	-	<i>S. aureus</i>	
349847	350002	+	<i>S. aureus</i>	
353312	353489	+	<i>S. aureus</i>	
354601	354868	+	<i>S. aureus</i>	
360696	360850	+	Staphylococcus	
365096	365455	+	Staphylococcus	
374306	374590	+	<i>S. aureus</i>	
379765	379926	+	Staphylococcus	
383854	384128	+	<i>S. aureus</i>	
387801	388115	+	Staphylococcus	STAR, HMM
388699	388870	+	<i>S. aureus</i>	
394432	394587	+	<i>S. aureus</i>	
407471	407816	+	Staphylococcus	T-box upstream of SA0347, Marchais et al. (8)
407899	408121	+	Staphylococcus	
412795	413226	+	Firmicutes	End of <i>rpsR</i> +Term and downstream of IGR+Term
417756	417939	+	<i>S. aureus</i>	
418043	418608	+	Firmicutes	Includes SA0360 and downstream of IGR
423119	423530	+	Staphylococcus	
428000	428312	+	<i>S. aureus</i>	
430626	430930	+	Firmicutes	Purine riboswitch upstream of <i>xprT</i>
438863	439249	-	<i>S. aureus</i>	
440633	440824	+	<i>S. aureus</i>	
442870	443053	+	<i>S. aureus</i>	
445513	445727	+	<i>S. aureus</i>	
453907	454123	+	<i>S. aureus</i>	
456769	457081	+	<i>S. aureus</i>	
464508	464684	+	<i>S. aureus</i>	
476027	476316	+	<i>S. aureus</i>	
484610	484907	+	<i>S. aureus</i>	
490504	490655	+	Staphylococcus	
496737	496910	+	<i>S. aureus</i>	Upstream of tRNA
497009	497168	+	<i>S. aureus</i>	Downstream of tRNA
501512	501994	+	Firmicutes	4.5S RNA (SRP)
505754	506160	+	Staphylococcus	Upstream of rRNA
507976	508186	+	Staphylococcus	Between tRNA and 5.8S rRNA
511304	511472	+	Staphylococcus	Downstream of tRNA+term, Livny et al. (5)
511510	511749	+	Staphylococcus	IGR+Term

511895	512130	+	Staphylococcus	Term upstream of SA0439
526716	526951	+	<i>S. aureus</i>	
530335	530625	+	<i>S. aureus</i>	
538614	538786	+	<i>S. aureus</i>	
542978	543170	+	<i>S. aureus</i>	
544110	544263	+	Staphylococcus	
545202	545410	+	<i>S. aureus</i>	
547095	547406	+	Staphylococcus	Term upstream of <i>lysS</i> , T-box
555815	556094	+	Staphylococcus	Upstream of tRNA, IGR+Term
560474	560625	+	Staphylococcus	
565362	565614	+	<i>S. aureus</i>	
568157	568636	+	<i>S. aureus</i>	
570146	570317	+	Staphylococcus	Term upstream of <i>cysE</i> , T-box
576540	576690	+	<i>S. aureus</i>	Leader of <i>rplA</i> , Roberts et al. (7)
577440	577710	+	Staphylococcus	L10 leader+Term, upstream of <i>rplJ</i>
578633	578795	+	Staphylococcus	Term upstream of SA0499
579409	579564	+	Staphylococcus	
586955	587112	+	Staphylococcus	Small ORF
587214	587416	+	Firmicutes	Leader of <i>rpsL</i> , Roberts et al. (7)
590574	590789	+	Staphylococcus	Term upstream of <i>tufA</i>
591975	592251	+	Staphylococcus	
594790	595053	+	<i>S. aureus</i>	
597734	597896	+	Staphylococcus	
604788	605214	+	<i>S. aureus</i>	
608160	608441	+	<i>S. aureus</i>	
612630	612992	+	<i>S. aureus</i>	
619632	619904	-	<i>S. aureus</i>	
621878	622099	+	<i>S. aureus</i>	
625046	625392	+	Staphylococcus	
630855	631008	+	<i>S. aureus</i>	
631050	631324	+	<i>S. aureus</i>	
632189	632341	+	<i>S. aureus</i>	
636758	637304	+/-	Staphylococcus	RsaA (+ strand)
640216	640541	+	Staphylococcus	
646069	646253	+	<i>S. aureus</i>	
646366	646552	+	<i>S. aureus</i>	
650163	650446	+	<i>S. aureus</i>	
651569	651954	+	Staphylococcus	STAR, HMM, Livny et al. (5)
652533	652685	+	<i>S. aureus</i>	
655059	655507	+	<i>S. aureus</i>	
656598	656828	+	<i>S. aureus</i>	Upstream of SA0563- <i>argS</i> , T-box ?
659828	660066	+	<i>S. aureus</i>	RsaX06
665140	665294	+	<i>S. aureus</i>	

666098	666298	+	Staphylococcus	RsaX07
666842	667117	-	Staphylococcus	
667125	667630	-	Staphylococcus	
669192	669369	+	<i>S. aureus</i>	
675825	676113	+	<i>S. aureus</i>	
679723	680024	-	<i>S. aureus</i>	RsaC
680509	680738	-	<i>S. aureus</i>	RsaC
686710	687007	+	<i>S. aureus</i>	
692039	692383	+	<i>S. aureus</i>	
694133	694433	+	<i>S. aureus</i>	
695728	696143	+	Staphylococcus	RsaD, Livny et al. (5)
697192	697355	+	Staphylococcus	
703029	703495	+	<i>S. aureus</i>	
712969	713456	+	Staphylococcus	
715250	715686	+	<i>S. aureus</i>	
716497	716842	+	Staphylococcus	
717493	718174	+	<i>S. aureus</i>	
720870	721025	+	<i>S. aureus</i>	
722073	722559	+	Staphylococcus	
728170	728344	+	Staphylococcus	
731774	731963	+	Staphylococcus	
735892	736086	+	Staphylococcus	
738044	738220	+	Staphylococcus	
741648	741935	+	Staphylococcus	
742699	742879	-	<i>S. aureus</i>	
744682	744959	+	Staphylococcus	
746121	746316	+	Staphylococcus	
749975	750264	+	<i>S. aureus</i>	
751463	751682	+	Staphylococcus	
753034	753240	+	<i>S. aureus</i>	
757175	757551	-	Staphylococcus	
758259	758527	-	<i>S. aureus</i>	
761087	761229	+	Staphylococcus	RsaX08, pre-Q1 riboswitch upstream of SA0668 Marchais et al. (8)
765821	766226	+	Staphylococcus	
768200	768422	+	Staphylococcus	
772138	772335	+	<i>S. aureus</i>	
776137	776453	+	Staphylococcus	STAR, HMM, Livny et al. (5)
777120	777557	+	Staphylococcus	STAR, HMM, Livny et al. (5)
780748	780920	-	<i>S. aureus</i>	
782415	782637	+	<i>S. aureus</i>	
782659	783086	+	Staphylococcus	Purine riboswitch upstream of <i>nrdI</i>
790883	791038	+	Staphylococcus	STAR

794592	794959	+	<i>S. aureus</i>	
798801	798968	-	Staphylococcus	
800069	800355	+	Staphylococcus	
803253	803584	+	<i>S. aureus</i>	
805998	806284	+	<i>S. aureus</i>	
808967	809331	+	Staphylococcus	Includes small ORF
810429	810772	+	<i>S. aureus</i>	Includes small ORF
812626	812888	+	<i>S. aureus</i>	
817939	818247	+	Staphylococcus	STAR, HMM
823225	824090	+	Staphylococcus	STAR, HMM
827069	827237	+	<i>S. aureus</i>	
827447	827597	+	Staphylococcus	
828251	828448	+	<i>S. aureus</i>	
829368	829873	+	Staphylococcus	RsaH, Livny et al. (5)
830837	831322	+	Staphylococcus	STAR, HMM
838883	839113	+	<i>S. aureus</i>	
839881	840052	+	Staphylococcus	
844742	845141	+	Staphylococcus	Downstream of tmRNA
846434	847030	+	<i>S. aureus</i>	
848235	848481	+	<i>S. aureus</i>	
851455	851614	+	<i>S. aureus</i>	
853205	853501	+	<i>S. aureus</i>	
854562	854863	+	<i>S. aureus</i>	
855441	855747	+	Firmicutes	Upstream of SA0746 (staphylococcal nuclease)
856473	856765	+	Staphylococcus	
857190	857552	+	Staphylococcus	Term far upstream of SA0748
859310	859768	+	<i>S. aureus</i>	
865573	865726	+	Staphylococcus	Roberts et al. (7)
872157	872606	+	<i>S. aureus</i>	
873353	873699	+	<i>S. aureus</i>	Livny et al. (5)
873882	874206	+	Staphylococcus	STAR, HMM
874260	874502	+	Staphylococcus	STAR, HMM
875429	875597	+	Firmicutes	SAM riboswitch upstream of SA0769
878278	878529	+	<i>S. aureus</i>	
879854	880109	+	Staphylococcus	Includes small ORF
886286	886588	+	<i>S. aureus</i>	
893715	893905	+	Staphylococcus	
897895	898263	+	<i>S. aureus</i>	Includes small ORF
904608	904758	+	<i>S. aureus</i>	
912118	912341	-	<i>S. aureus</i>	
920248	920429	+	<i>S. aureus</i>	
921586	921871	+	<i>S. aureus</i>	
924454	924635	+	Staphylococcus	STAR, HMM

928516	928811	+	Staphylococcus	
930204	930359	+	<i>S. aureus</i>	
939508	939672	+	Staphylococcus	
941308	941464	+	<i>S. aureus</i>	
944084	944612	+	Staphylococcus	STAR
953726	953966	+	<i>S. aureus</i>	
954414	954644	+	<i>S. aureus</i>	
954859	955148	+	<i>S. aureus</i>	Term upstream of <i>fabh</i> , riboswitch ?
957808	957990	+	Staphylococcus	
963726	963926	+	<i>S. aureus</i>	
970542	970834	+	Staphylococcus	
971369	971600	+	Staphylococcus	
975288	975460	+	Firmicutes	RsaE, Livny et al. (5)
975460	975740	+	<i>S. aureus</i>	RsaF
977641	977812	+	<i>S. aureus</i>	
983797	984032	+	<i>S. aureus</i>	
986113	986457	+	<i>S. aureus</i>	
988928	989089	+	Staphylococcus	
992059	992486	+	Staphylococcus	
996902	997115	+	<i>S. aureus</i>	
1002460	1002628	+	<i>S. aureus</i>	upstream of tRNA gene
1003508	1003661	-	<i>S. aureus</i>	
1003804	1004015	-	Staphylococcus	In antisense orientation to SA0883
1006559	1006849	+	<i>S. aureus</i>	
1010523	1010773	+	<i>S. aureus</i>	Long IGR
1011366	1011547	+	<i>S. aureus</i>	
1013265	1013422	+	<i>S. aureus</i>	
1019781	1019947	+	<i>S. aureus</i>	Downstream of <i>memb</i> , Potential sRNA with Term
1022764	1022962	-	<i>S. aureus</i>	
1023012	1023200	-	<i>S. aureus</i>	
1032642	1033043	+	Staphylococcus	
1039067	1039361	-	<i>S. aureus</i>	
1039680	1040504	-	Staphylococcus	STAR, HMM
1041406	1041601	+	Staphylococcus	
1052839	1053096	+	Staphylococcus	
1055907	1056798	+	Firmicutes	Antisense orientation of TPP, repeat elements
1058320	1058707	+	<i>S. aureus</i>	
1062659	1063092	+	Staphylococcus	
1066487	1066727	+	Staphylococcus	Livny et al. (5)
1068543	1069063	-	Firmicutes	
1069685	1069846	+	<i>S. aureus</i>	
1083867	1084046	-	Staphylococcus	

1085788	1085953	+	Staphylococcus	
1087945	1088099	+	<i>S. aureus</i>	
1089128	1089413	+	<i>S. aureus</i>	
1090887	1091225	+	Staphylococcus	
1095856	1096173	+	Staphylococcus	
1097586	1097756	+	<i>S. aureus</i>	
1106608	1106790	-	<i>S. aureus</i>	
1107856	1108058	+	<i>S. aureus</i>	
1112944	1113182	+	<i>S. aureus</i>	
1113949	1114289	+	Firmicutes	T-box upstream of <i>pheS</i>
1119148	1119545	+	<i>S. aureus</i>	
1127040	1127317	+	Staphylococcus	
1130625	1130822	+	<i>S. aureus</i>	
1132777	1133220	+	<i>S. aureus</i>	Includes small ORF
1134351	1134993	+	<i>S. aureus</i>	Term upstream of SA1002
1135503	1135759	+	<i>S. aureus</i>	
1136760	1137488	+	Staphylococcus	STAR, HMM, Livny et al. (5)
1137644	1137875	+	<i>S. aureus</i>	Long IGR
1137937	1138593	+	<i>S. aureus</i>	Long IGR
1138630	1139083	+	Staphylococcus	RsaX10
1141537	1142098	-	<i>S. aureus</i>	
1142166	1143006	-	Staphylococcus	STAR, HMM
1145459	1145817	+	Staphylococcus	
1149556	1149966	+	Staphylococcus	Includes small ORF
1151613	1151839	-	Staphylococcus	STAR
1151978	1152191	-	<i>S. aureus</i>	Upstream of tRNA gene
1152270	1152713	+	Staphylococcus	Upstream of small ORF
1154325	1154496	+	<i>S. aureus</i>	
1166960	1167207	+	<i>S. aureus</i>	
1171092	1171307	+	Staphylococcus	T-box upstream of <i>ileS</i>
1175179	1175939	+	Staphylococcus	STAR, HMM, RsaX11
1177413	1177766	+	Staphylococcus	Cis-regulatory leader <i>pyrR</i>
1178313	1178505	+	Staphylococcus	Cis-regulatory leader <i>pyrP</i> (PyrR)
1187782	1188280	+	<i>S. aureus</i>	
1188815	1188998	+	<i>S. aureus</i>	
1191840	1192077	+	<i>S. aureus</i>	
1195742	1196199	+	<i>S. aureus</i>	
1197660	1197816	+	Staphylococcus	
1204379	1204587	+	<i>S. aureus</i>	
1207056	1207252	+	Firmicutes	Long IGR, antisense orientation to <i>rpmB</i>
1207368	1207623	+	Staphylococcus	Long IGR
1209822	1210010	+	<i>S. aureus</i>	
1212081	1212259	+	Staphylococcus	

1215618	1215904	+	Staphylococcus	Repeats, Term 80 nt upstream of <i>hmrB</i>
1223881	1224165	+	Firmicutes	Leader of <i>rpsP</i>
1229115	1229512	+	<i>S. aureus</i>	
1235976	1236140	+	<i>S. aureus</i>	
1237018	1237171	+	Staphylococcus	
1240749	1241012	+	<i>S. aureus</i>	
1245038	1245198	+	Staphylococcus	Leader of <i>rpsB</i>
1245984	1246147	+	Firmicutes	Small ORF, leader of EF-TS
1248617	1248800	+	Staphylococcus	Term upstream of <i>uppS</i> , Livny et al. (5)
1250407	1250591	+	<i>S. aureus</i>	
1253679	1253869	+	<i>S. aureus</i>	
1263619	1263776	+	<i>S. aureus</i>	Leader of <i>truB-ribC-rpsO</i>
1265989	1266427	+	Staphylococcus	Leader of <i>pnpA</i>
1270493	1270665	+	<i>S. aureus</i>	
1279039	1279226	+	<i>S. aureus</i>	
1280406	1280564	+	Staphylococcus	
1281617	1281966	+	Staphylococcus	
1283591	1283812	+	<i>S. aureus</i>	
1283831	1284194	+	Staphylococcus	
1290733	1291002	+	<i>S. aureus</i>	
1296413	1296565	+	<i>S. aureus</i>	
1306539	1306761	+	<i>S. aureus</i>	
1309256	1309601	+	<i>S. aureus</i>	
1309685	1310025	+	Staphylococcus	
1310031	1310492	+	<i>S. aureus</i>	Includes small ORF
1310499	1311089	+	<i>S. aureus</i>	Includes small ORF
1311163	1311668	+	<i>S. aureus</i>	
1311817	1312415	+	<i>S. aureus</i>	Includes small ORF
1312805	1313249	+	Staphylococcus	Includes small ORF
1313887	1314054	+	Staphylococcus	Includes small ORF
1314277	1314447	+	<i>S. aureus</i>	
1315518	1315773	+	<i>S. aureus</i>	
1317511	1317692	+	<i>S. aureus</i>	
1322977	1323131	+	<i>S. aureus</i>	
1325311	1325491	+	Staphylococcus	
1329618	1329858	+	<i>S. aureus</i>	
1331898	1332087	+	Staphylococcus	
1336837	1337115	+	<i>S. aureus</i>	Includes small ORF
1337755	1337964	+	Staphylococcus	Includes small ORF
1338420	1338619	+	Staphylococcus	Includes small ORF
1340640	1340865	+	<i>S. aureus</i>	
1346496	1346695	+	Staphylococcus	
1348594	1349059	+	Staphylococcus	STAR, HMM

1352000	1352173	+	<i>S. aureus</i>	
1352697	1352869	+	<i>S. aureus</i>	
1360589	1360804	+	<i>S. aureus</i>	
1363256	1363611	+	<i>S. aureus</i>	
1366136	1366335	+	<i>S. aureus</i>	
1368171	1368382	+	Firmicutes	
1378888	1379152	+	<i>S. aureus</i>	
1381748	1382022	+	<i>S. aureus</i>	
1387702	1387901	-	<i>S. aureus</i>	
1388353	1388520	+	<i>S. aureus</i>	
1395079	1395520	-	Staphylococcus	STAR, HMM
1395695	1395951	-	Firmicutes	Includes small ORF
1398843	1399123	+	Staphylococcus	RsaX12 upstream of lysine riboswitch
1399339	1399740	+	Staphylococcus	Lysine riboswitch upstream of <i>lysC</i>
1408636	1408787	-	<i>S. aureus</i>	
1408937	1409164	-	Firmicutes	Leader of <i>cspA</i>
1409542	1409751	+	Staphylococcus	
1413176	1413388	-	<i>S. aureus</i>	
1417320	1417878	-	Staphylococcus	
1424188	1424971	-	Staphylococcus	Upstream of truncated <i>arIR</i>
1427366	1427562	-	<i>S. aureus</i>	
1432388	1432583	-	<i>S. aureus</i>	
1433547	1433788	-	<i>S. aureus</i>	
1436007	1436171	+	<i>S. aureus</i>	
1437004	1437246	+	Staphylococcus	STAR, HMM, Livny et al. (5)
1469803	1470083	-	<i>S. aureus</i>	
1475356	1475822	-	<i>S. aureus</i>	
1480194	1480447	+	<i>S. aureus</i>	
1483726	1484114	-	Firmicutes	RNase P, HMM
1485561	1485911	+	<i>S. aureus</i>	
1491105	1491401	-	<i>S. aureus</i>	
1492707	1492871	-	Staphylococcus	
1499069	1499225	-	<i>S. aureus</i>	
1499634	1499839	-	<i>S. aureus</i>	
1506684	1507446	-	<i>S. aureus</i>	
1510729	1511138	-	Firmicutes	Leader of <i>hu</i>
1514878	1515160	-	Staphylococcus	
1515273	1515551	-	<i>S. aureus</i>	
1523885	1524137	-	Firmicutes	FMN upstream of SA1316, HMM
1530598	1530818	-	<i>S. aureus</i>	
1544160	1544364	+	Staphylococcus	
1550247	1550438	-	<i>S. aureus</i>	
1553370	1553696	-	<i>S. aureus</i>	

1562522	1562927	-	<i>S. aureus</i>	
1568050	1568228	-	Staphylococcus	
1571964	1572117	+	<i>S. aureus</i>	
1577475	1577690	-	<i>S. aureus</i>	
1583172	1583325	-	Staphylococcus	
1585629	1585811	-	Firmicutes	Leader of <i>rpmG</i>
1588657	1588896	-	<i>S. aureus</i>	
1600077	1600363	+	Staphylococcus	T-Box upstream of <i>glyS</i> , Marchais et al. (8)
1605769	1606026	-	<i>S. aureus</i>	
1608584	1609165	-	Firmicutes	Leader of <i>rpsU</i>
1618277	1618638	-	Staphylococcus	STAR, HMM
1620964	1621273	-	Firmicutes	In antisense orientation to <i>rpsT</i>
1631726	1631958	-	<i>S. aureus</i>	
1632072	1632346	-	<i>S. aureus</i>	Includes part of SA1429
1632385	1632595	-	<i>S. aureus</i>	Includes part of SA1429
1633477	1633790	-	<i>S. aureus</i>	Upstream of STAR, HMM, Livny et al. (5)
1640858	1641060	-	<i>S. aureus</i>	
1645154	1645434	-	Staphylococcus	
1649154	1649422	-	Firmicutes	T-box upstream of <i>alaS</i>
1652768	1653061	-	Staphylococcus	STAR, HMM, Livny et al. (5)
1660490	1660646	-	Staphylococcus	Roberts et al. (7)
1663947	1664110	-	Staphylococcus	T-Box upstream of <i>hisS</i> , HMM, Livny et al. (5)
1667827	1668128	-	<i>S. aureus</i>	
1671068	1671269	-	<i>S. aureus</i>	
1682744	1682923	-	Staphylococcus	
1683462	1683761	-	Staphylococcus	
1685414	1685863	+	Firmicutes	Cis-regulatory leader of <i>ermA</i>
1696574	1696752	+	Firmicutes	T-box upstream of <i>valS</i>
1705181	1705379	-	Staphylococcus	
1711867	1712034	-	Firmicutes	L20 leader upstream of <i>infC</i>
1713606	1713993	-	Firmicutes	Lysine riboswitch upstream of <i>lysP</i>
1715961	1716354	-	Firmicutes	T-box upstream of <i>thrS</i> , HMM
1720401	1720569	-	Staphylococcus	
1724751	1725005	-	<i>S. aureus</i>	
1726510	1726993	-	<i>S. aureus</i>	
1732201	1732547	+	Staphylococcus	
1734098	1734256	+	<i>S. aureus</i>	
1736967	1737251	-	Staphylococcus	Livny et al. (5)
1740661	1740878	-	<i>S. aureus</i>	
1745150	1745402	-	Staphylococcus	STAR, HMM, Term upstream of SA1526
1746749	1746903	+	<i>S. aureus</i>	
1758029	1758239	-	<i>S. aureus</i>	
1761531	1761713	+	Staphylococcus	

1771001	1771159	+	Staphylococcus	
1773791	1774166	+	Staphylococcus	T-box upstream of <i>tyrS</i>
1775089	1775274	+	Staphylococcus	
1777967	1778142	+/-	<i>S. aureus</i>	RsaB (+ strand)
1780001	1780272	-	<i>S. aureus</i>	
1785343	1785657	-	<i>S. aureus</i>	
1787090	1787505	-	<i>S. aureus</i>	
1799158	1799538	-	<i>S. aureus</i>	
1801010	1801172	-	Staphylococcus	RsaX14
1801193	1801556	-	Staphylococcus	RsaX14
1803836	1804170	-	Staphylococcus	STAR, HMM
1804196	1804363	-	Staphylococcus	STAR, HMM
1807199	1807350	+	<i>S. aureus</i>	
1815565	1815759	-	<i>S. aureus</i>	
1822306	1822720	-	<i>S. aureus</i>	
1823612	1823809	+	Staphylococcus	
1828395	1828720	-	Firmicutes	FMN upstream of <i>ribD</i> , HMM
1830320	1830823	+	<i>S. aureus</i>	Term upstream of SA1591
1833394	1833678	+	Staphylococcus	
1833892	1834060	+	Staphylococcus	
1835663	1835820	+	<i>S. aureus</i>	
1839186	1839704	+	Staphylococcus	RsaX15, HMM
1839706	1839931	+	Firmicutes	Truncated transposase
1844412	1844781	+	Firmicutes	Leader of <i>pckA</i> , in antisense orientation to SAM riboswitch
1850782	1850936	-	<i>S. aureus</i>	
1851426	1851728	+	<i>S. aureus</i>	
1856369	1856798	+	Staphylococcus	<i>sprA</i> , Pichon and Felden (6), HMM
1857482	1857823	-	Staphylococcus	
1860738	1861018	-	<i>S. aureus</i>	
1865118	1865490	+	<i>S. aureus</i>	Term upstream of <i>splA</i>
1865836	1866079	+	<i>S. aureus</i>	
1866672	1867028	+	<i>S. aureus</i>	<i>sprB</i> , Pichon and Felden (6)
1867248	1867528	+	<i>S. aureus</i>	Downstream of <i>sprB</i>
1871172	1871931	+	Staphylococcus	<i>sprC</i> , Pichon and Felden (6)
1873386	1873642	+	Staphylococcus	
1873858	1874112	+	Staphylococcus	
1877123	1877283	-	<i>S. aureus</i>	
1881673	1881834	-	<i>S. aureus</i>	Downstream of tRNA cluster
1882658	1882992	-	<i>S. aureus</i>	Upstream of tRNA cluster
1883713	1883866	-	<i>S. aureus</i>	Term upstream of SA1649, riboswitch ?
1887313	1887560	+	Staphylococcus	
1899115	1899440	-	<i>S. aureus</i>	

1899478	1899762	-	<i>S. aureus</i>	HMM, Term
1901329	1901482	-	<i>S. aureus</i>	
1901976	1902324	-	Staphylococcus	
1907967	1908121	-	Staphylococcus	Upstream of STAR, HMM
1914230	1914658	+	Staphylococcus	RsaX16
1916129	1916289	+	<i>S. aureus</i>	3' UTR of <i>tnp</i> and upstream of tRNA cluster
1921856	1922066	-	Staphylococcus	Upstream of rRNA
1923904	1924623	-	Staphylococcus	Upstream of rRNA
1929336	1929486	+	Staphylococcus	
1931363	1931530	-	<i>S. aureus</i>	
1938347	1938528	-	<i>S. aureus</i>	
1939381	1939606	-	<i>S. aureus</i>	
1941930	1942087	+	<i>S. aureus</i>	
1944228	1944386	+	Staphylococcus	
1945105	1945410	+	<i>S. aureus</i>	
1950480	1950858	+	Staphylococcus	STAR, HMM
1951996	1952301	+	Staphylococcus	STAR, HMM, Livny et al. (5)
1952892	1953043	-	<i>S. aureus</i>	
1955115	1955395	+	Staphylococcus	
1955903	1956286	+	<i>S. aureus</i>	
1961321	1962071	-	Staphylococcus	STAR, HMM, Livny et al. (5)
1965286	1965636	+	Staphylococcus	
1975263	1975577	+	<i>S. aureus</i>	
1975655	1976120	+	<i>S. aureus</i>	
1977802	1978014	+	<i>S. aureus</i>	
1978693	1978946	-	Staphylococcus	
1991902	1993014	+	Staphylococcus	HMM, RsaX17 Detection of a stable RNA in antisense, Roberts et al. (7), Livny et al. (5)
2000792	2001169	+	Staphylococcus	HMM
2002762	2003025	+	<i>S. aureus</i>	
2004766	2004935	-	<i>S. aureus</i>	
2005435	2006027	-	<i>S. aureus</i>	
2006124	2006498	-	<i>S. aureus</i>	
2006879	2007560	+	<i>S. aureus</i>	sprD, Pichon and Felden (6)
2009578	2009756	-	<i>S. aureus</i>	
2010950	2011172	-	Staphylococcus	sprG, antisense to sprF, Pichon and Felden (6) ; Roberts et al. (7), C-rich motif
2012392	2012550	-	<i>S. aureus</i>	
2023371	2023533	-	Staphylococcus	
2044276	2044498	+	<i>S. aureus</i>	
2044531	2044689	+	<i>S. aureus</i>	
2046589	2046936	+	<i>S. aureus</i>	

2050407	2050637	+	<i>S. aureus</i>	
2052735	2053160	+	<i>S. aureus</i>	
2054442	2054834	+	Staphylococcus	STAR, HMM
2056712	2056982	+	<i>S. aureus</i>	
2057885	2058037	+	<i>S. aureus</i>	
2058852	2059635	+	<i>S. aureus</i>	
2060672	2061075	+	Staphylococcus	
2066059	2066357	-	<i>S. aureus</i>	
2074274	2074451	+	Firmicutes	
2076511	2076703	+	<i>S. aureus</i>	
2077355	2077666	+	Staphylococcus	STAR and Term upstream of SA1841
2078721	2078937	+	Staphylococcus	RNAIII-C-rich motif
2079188	2079445	+	Staphylococcus	RNAIII-C-rich motif
2082286	2082571	+	<i>S. aureus</i>	
2086113	2086282	-	<i>S. aureus</i>	
2089116	2089363	-	<i>S. aureus</i>	
2090087	2090243	+	Staphylococcus	
2092236	2092595	+	Staphylococcus	
2097090	2097527	+	Staphylococcus	Term upstream of <i>ilvD</i> , riboswitch ?
2112153	2112350	-	Staphylococcus	Upstream of rRNA
2112351	2112621	-	Staphylococcus	Upstream of rRNA
2114673	2115087	-	<i>S. aureus</i>	Upstream of tRNA
2117717	2118136	-	<i>S. aureus</i>	
2120826	2121111	-	<i>S. aureus</i>	
2125946	2126096	-	<i>S. aureus</i>	
2140430	2140694	+	<i>S. aureus</i>	
2142896	2143098	+	<i>S. aureus</i>	
2145816	2145986	+	Staphylococcus	
2150134	2150353	-	<i>S. aureus</i>	THI riboswitch upstream of <i>tenA</i>
2151848	2152017	+	<i>S. aureus</i>	
2154589	2154831	-	Staphylococcus	
2170438	2170674	-	Firmicutes	Leader of <i>rpmE</i>
2172033	2172272	-	Staphylococcus	
2173713	2173918	-	<i>S. aureus</i>	
2175623	2175838	-	<i>S. aureus</i>	
2176973	2177125	+	Staphylococcus	
2179408	2179559	-	<i>S. aureus</i>	
2183309	2183669	-	<i>S. aureus</i>	
2186103	2186358	+	Staphylococcus	
2191191	2191375	-	Staphylococcus	
2193769	2194327	-	<i>S. aureus</i>	
2195501	2195682	-	<i>S. aureus</i>	
2198126	2198299	+	Staphylococcus	

2198782	2198996	+	<i>S. aureus</i>	Livny et al. (5)
2199743	2200192	+	<i>S. aureus</i>	Cis-regulatory leader of <i>ermA</i>
2211959	2212517	+	Staphylococcus	Antisense to GlmS riboswitch
2217927	2218109	+	<i>S. aureus</i>	
2228994	2229181	-	Staphylococcus	
2233852	2234049	-	Staphylococcus	
2234050	2234320	-	Staphylococcus	Upstream of rRNA
2235891	2236330	-	Staphylococcus	Upstream of rRNA
2241225	2241396	-	<i>S. aureus</i>	
2244132	2244347	-	Staphylococcus	
2245730	2245955	-	Staphylococcus	
2256957	2257130	-	Staphylococcus	
2258694	2258999	-	<i>S. aureus</i>	
2269141	2269392	-	<i>S. aureus</i>	
2270166	2270385	+	<i>S. aureus</i>	
2271593	2271847	+	<i>S. aureus</i>	
2273257	2273473	+	<i>S. aureus</i>	
2278693	2279037	+	<i>S. aureus</i>	
2282469	2282662	-	Staphylococcus	
2283771	2283931	-	<i>S. aureus</i>	
2285904	2286114	-	<i>S. aureus</i>	
2287009	2287178	-	<i>S. aureus</i>	
2288114	2288407	+	<i>S. aureus</i>	
2290679	2290831	-	Staphylococcus	L13 leader upstream of <i>rplM</i>
2294193	2294358	-	<i>S. aureus</i>	End of <i>rplQ</i> , Roberts et al. (7)
2297295	2297786	-	Firmicutes	Cis-regulatory leader of <i>infA</i>
2315523	2315785	+	Staphylococcus	STAR, HMM
2317119	2317270	+	<i>S. aureus</i>	
2322654	2322811	-	<i>S. aureus</i>	
2326152	2326412	-	Staphylococcus	STAR, HMM
2326455	2326612	-	Staphylococcus	STAR, HMM
2335098	2335370	+	<i>S. aureus</i>	
2336178	2336334	+	Staphylococcus	
2338812	2339179	-	<i>S. aureus</i>	
2342611	2342890	+	Staphylococcus	
2347937	2348096	+	<i>S. aureus</i>	
2348445	2348807	-	Staphylococcus	
2351998	2352576	+	Staphylococcus	
2353589	2353825	+	Staphylococcus	
2356433	2356657	+	<i>S. aureus</i>	
2357081	2357300	+	Staphylococcus	
2365373	2365625	-	<i>S. aureus</i>	
2367831	2368208	+	Staphylococcus	Rsal, Marchais et al. (8)

2370043	2370414	+	<i>S. aureus</i>	
2371369	2371738	+	<i>S. aureus</i>	
2373129	2373403	+	Staphylococcus	
2375948	2376119	+	Staphylococcus	
2378340	2378664	+	Staphylococcus	
2384868	2385044	-	<i>S. aureus</i>	
2389095	2389322	+	<i>S. aureus</i>	
2389754	2389995	+	<i>S. aureus</i>	
2390997	2391225	+	Staphylococcus	
2393252	2393404	+	<i>S. aureus</i>	
2398087	2398249	-	<i>S. aureus</i>	
2399072	2399253	+	<i>S. aureus</i>	
2399983	2400209	+	<i>S. aureus</i>	
2405538	2405697	+	<i>S. aureus</i>	
2406372	2406582	+	<i>S. aureus</i>	
2409885	2410108	+	<i>S. aureus</i>	
2411321	2411574	-	<i>S. aureus</i>	
2412971	2413192	-	<i>S. aureus</i>	
2414829	2415053	+	<i>S. aureus</i>	
2425527	2425678	+	<i>S. aureus</i>	
2426297	2426582	-	<i>S. aureus</i>	
2429796	2430015	-	<i>S. aureus</i>	
2430501	2430698	-	<i>S. aureus</i>	
2431985	2432135	+	Staphylococcus	
2435998	2436269	+	<i>S. aureus</i>	RsaX18, stable RNA in antisense orientation, Roberts et al. (7)
2437028	2437317	+	Staphylococcus	Downstream of RsaX18
2439937	2440194	+	<i>S. aureus</i>	
2440561	2440869	+	Staphylococcus	HMM, repeats, RsaX19
2442167	2442330	-	<i>S. aureus</i>	
2442698	2443032	+	<i>S. aureus</i>	
2443583	2443744	+	<i>S. aureus</i>	
2445484	2445745	+	Staphylococcus	
2447564	2447719	+	<i>S. aureus</i>	
2456483	2456710	-	<i>S. aureus</i>	
2461228	2461426	-	<i>S. aureus</i>	
2461971	2462198	-	<i>S. aureus</i>	
2463308	2463504	-	<i>S. aureus</i>	
2466457	2466755	-	<i>S. aureus</i>	
2467832	2468006	+	<i>S. aureus</i>	
2473362	2473553	-	Staphylococcus	RsaX20
2474277	2474452	-	<i>S. aureus</i>	
2475454	2475778	+	<i>S. aureus</i>	

2475795	2476058	+	<i>S. aureus</i>	
2477641	2477908	+	<i>S. aureus</i>	
2478860	2479149	+	<i>S. aureus</i>	
2479216	2479421	+	<i>S. aureus</i>	
2486709	2486881	-	<i>S. aureus</i>	RsaJ
2490706	2491116	-	Staphylococcus	STAR, HMM
2497788	2498049	+	Staphylococcus	
2503698	2503855	+	<i>S. aureus</i>	
2509313	2509478	-	<i>S. aureus</i>	
2514267	2514440	+	Staphylococcus	
2516120	2516273	+	<i>S. aureus</i>	
2519098	2519353	-	<i>S. aureus</i>	
2525577	2525767	+	<i>S. aureus</i>	
2526558	2526857	+	Staphylococcus	
2536141	2536452	-	<i>S. aureus</i>	
2536625	2536820	-	<i>S. aureus</i>	
2544138	2544290	-	Staphylococcus	
2546465	2546691	-	Staphylococcus	
2549419	2549630	-	<i>S. aureus</i>	
2551315	2551599	-	<i>S. aureus</i>	
2556804	2557057	+	<i>S. aureus</i>	
2557826	2558087	+	<i>S. aureus</i>	
2562110	2562398	-	<i>S. aureus</i>	HMM
2563833	2564100	-	<i>S. aureus</i>	
2564499	2564800	+	<i>S. aureus</i>	
2568115	2568322	-	<i>S. aureus</i>	
2571236	2571482	-	<i>S. aureus</i>	
2571601	2571886	-	<i>S. aureus</i>	
2575158	2575471	-	<i>S. aureus</i>	
2580982	2581163	-	<i>S. aureus</i>	
2583795	2583980	-	<i>S. aureus</i>	
2585432	2585672	+	<i>S. aureus</i>	
2586700	2586952	+	<i>S. aureus</i>	RsaX22
2590828	2591139	+	Staphylococcus	RsaX23, sprA3, Pichon and Felden (6), HMM
2596738	2596915	+	<i>S. aureus</i>	
2599037	2599326	+	<i>S. aureus</i>	
2600725	2600877	+	<i>S. aureus</i>	
2601516	2601689	+	<i>S. aureus</i>	Livny et al. (5)
2605324	2605790	+	<i>S. aureus</i>	
2609352	2609550	-	<i>S. aureus</i>	
2611619	2611889	-	<i>S. aureus</i>	RAT motif in leader of <i>ptsG</i>
2614788	2614993	-	Staphylococcus	
2616463	2616619	+	Staphylococcus	

2618504	2618750	+	<i>S. aureus</i>	
2619921	2620083	+	Staphylococcus	
2620690	2621010	+	<i>S. aureus</i>	
2629104	2629317	+	<i>S. aureus</i>	
2635052	2635382	+	Staphylococcus	Includes part of repeats, Livny et al. (5)
2644680	2644981	-	<i>S. aureus</i>	
2646952	2647361	-	<i>S. aureus</i>	
2647851	2648018	-	<i>S. aureus</i>	
2648968	2649154	-	<i>S. aureus</i>	
2649183	2649430	-	<i>S. aureus</i>	RsaX24
2650531	2650761	+	<i>S. aureus</i>	
2651470	2651647	+	Staphylococcus	
2651696	2651846	+	<i>S. aureus</i>	
2652667	2652901	+	Staphylococcus	STAR, HMM
2661386	2661635	-	<i>S. aureus</i>	
2663861	2664068	+	<i>S. aureus</i>	
2666489	2666751	+	<i>S. aureus</i>	
2671087	2671338	-	Staphylococcus	Livny et al. (5)
2672085	2672534	+	<i>S. aureus</i>	Cis-regulatory leader of <i>ermA</i>
2678310	2678645	-	<i>S. aureus</i>	
2683676	2683834	+	<i>S. aureus</i>	
2684016	2684324	+	Staphylococcus	
2687272	2687440	+	Staphylococcus	
2688908	2689084	+	<i>S. aureus</i>	
2690843	2691222	-	<i>S. aureus</i>	
2697740	2697970	+	Staphylococcus	
2698571	2698723	+	Staphylococcus	
2700573	2700856	-	<i>S. aureus</i>	
2708048	2708199	-	<i>S. aureus</i>	
2708791	2709614	-	<i>S. aureus</i>	
2716436	2716702	+	Staphylococcus	
2721001	2721275	-	<i>S. aureus</i>	
2726839	2727174	-	<i>S. aureus</i>	
2727805	2727992	-	<i>S. aureus</i>	
2729572	2729953	-	<i>S. aureus</i>	
2730482	2730708	+	<i>S. aureus</i>	
2731227	2731379	+	<i>S. aureus</i>	
2739394	2739588	+	Staphylococcus	
2762069	2762602	-	<i>S. aureus</i>	
2763342	2763605	+	<i>S. aureus</i>	
2766556	2766758	-	<i>S. aureus</i>	
2768925	2769664	-	<i>S. aureus</i>	In antisense orientation to RsaX25
2773960	2774226	+	Staphylococcus	STAR

2776631	2777268	-	<i>S. aureus</i>	
2784309	2784518	-	<i>S. aureus</i>	
2792452	2792677	+	<i>S. aureus</i>	
2795645	2795809	+	<i>S. aureus</i>	
2796286	2796475	+	<i>S. aureus</i>	
2798095	2798255	+	<i>S. aureus</i>	
2800056	2800210	+	<i>S. aureus</i>	
2806131	2806408	+	Staphylococcus	Leader of <i>cspB</i>
2808484	2808774	-	Staphylococcus	
2814540	2814811	-	Firmicutes	Leader of <i>rpmH</i>

The boundaries of the conserved regions and their locations on the DNA strand are given. Comments include the known regulatory regions as defined by RFAM, the Rsa RNAs, the IGRs which have been predicted (5), and stable RNAs which were previously identified (6-8). Interesting features of conserved regions are mentioned. Term is for terminator of transcription.

**Table S2.** List of predicted IGRs found by a Hidden Markov Model (HMM).

HMM	Start-End	Flanking genes	Comments
1	24159-24341	<i>purA/vicR</i>	tRNA
2	39699-40058	<i>Pre-bleO</i>	RNA <sub>sim</sub>
3	42881-43180	SA0035-SA0036	RsaX01, RNA <sub>sim</sub>
4	233122-233343	SA0197- <i>oppF</i>	
5	254692-254833	<i>uhpT</i> -SA0215	RsaG
6	311527-311657	<i>bglA</i> - SA0257	RNA <sub>sim</sub>
7	387896-388065	SA0329- <i>appF</i>	RNA <sub>sim</sub>
8	407757-407873	SA0347-SA0348	T-Box, RNA <sub>sim</sub> , Marchais et al. (8)
9	469644-469972	SA0410- <i>ndhF</i>	Long IGR with small ORF
10	496930-497008	<i>gltD-treP</i>	tRNA
11	501419-501690	SA0434-SA0435	SRP
12	506179-507975	<i>recR</i> -rRNA	rRNA
13	508203-511102	rRNA-rRNA	rRNA
14	511185-511878	rRNA-SA0439	5S rRNA
15	549617-550468	<i>lysS</i> -16S rRNA	tRNA
16	550607-552312	tRNA-23S rRNA	rRNA
17	555476-555860	23S rRNA-SA0476	rRNA
18	651668-651853	SA0557-SA0558	RNA <sub>sim</sub> , Livny et al. (5)
19	659030-659146	<i>argS</i> - SA0565	
20	777116-777256	SA0680- SA0681	RNA <sub>sim</sub> , Livny et al. (5)
21	777392-777519	SA0680- SA0681	RNA <sub>sim</sub> , Livny et al. (5)
22	818082-818207	<i>uvrA</i> - <i>hprK</i>	RNA <sub>sim</sub>
23	830889-831241	SA0725- <i>gapR</i>	RNA <sub>sim</sub>
24	843796-844174	<i>ssrP</i> -SA0738	tmRNA
25	873970-874214	truncated-SA*-SA0768	RNA <sub>sim</sub>
26	874300-874421	truncated-SA*- SA0768	RNA <sub>sim</sub>
27	875413-875457	SAS023-SA0769	RNA <sub>sim</sub> , SAM riboswitch
28	924504-924632	<i>gudb*</i> - <i>glpQ</i>	RNA <sub>sim</sub>
29	1002623-1002696	SA0881*-SA0882	tRNA
30	1040000-1040140	SA0914-foID	RNA <sub>sim</sub>
31	1056349-1056701	SA0929-SAS030	Repeats downstream of a TPP riboswitch
32	1137073-1137371	SA1004-SAS035	RNA <sub>sim</sub> , Livny et al. (5)
33	1142669-1142793	SA1008-SA1009	RNA <sub>sim</sub>
34	1152072-1152271	SA1016-SA1018	tRNA
35	1175550-1175624	SA1037- <i>lsp</i>	RNA <sub>sim</sub> , RsaX11
36	1348702-1348948	<i>opuD</i> - <i>citB</i>	RNA <sub>sim</sub>
37	1353221-1351302	SA1186-SA1187	STAR
38	1395440-1395567	SA1221-truncated SA	RNA <sub>sim</sub>

39	1399134-1399210	SA1224- <i>lysC</i>	RsaX12, RNAsim, upstream of Lysine riboswitch
40	1437057-1437239	SA1265-SA1266	RNAsim, Livny et al. (5)
41	1483795-1483974	SA1277-SA1279	RNase P, RNAsim
42	1523747-1523924	SA1316-SA1317	FMN riboswitch, RNAsim
43	1618576-1618787	<i>hemN-lepA</i>	RNAsim
44	1633168-1633298	SA1430- SA1431	RNAsim, Livny et al. (5)
45	1649265-1649307	<i>alaS</i> - SA1447	T-Box
46	1652906-1653242	SA1448- SA1449	Livny et al. (5)
47	1660504-1660682	SA1455- <i>aspS</i>	6S RNA
48	1663953-1664007	<i>hisS-lytH</i>	RNAsim, Livny et al. (5)
49	1716110-1716151	<i>thrS-dnaI</i>	RNAsim
50	1745253-1745539	SA1526- SA1527	RNAsim
51	1804057-1804207	SA1572- SA1573	RNAsim
52	1828454-1828572	<i>ribD</i> - SA1590	FMN riboswitch, RNAsim
53	1839422-1839492	SA1602- truncated-SA	RsaX15, RNAsim
54	1856483-1856685	truncated-SA- <i>tnp</i>	sprA, Pichon and Felden (6), RNAsim
55	1881838-1882594	<i>seo</i> -SA1649	tRNA
56	1899470-1899757	SA1662-SA1663	Long IGR + Term, RNAsim
57	1907748-1907903	SA1670- SAS053	RNAsim
58	1916292-1918962	<i>tnp</i> -23S	rRNA
59	1918961-1921962	tRNA- tRNA	tRNA
60	1923883-1925438	23S- SA1678	rRNA
61	1950616-1950792	<i>map</i> - SA1705	Leader of <i>map</i> , Marchais et al. (8), RNAsim
62	1952064-1952292	SA1705- SA1706	RNAsim, Livny et al. (5)
63	1961413-1961597	SA1714- SA1715	RNAsim, Livny et al. (5)
64	1992295-1992365	SA1738- SAS056	RsaX17, RNAsim
65	2001219-2001316	SAS057- SA1749	RNAsim
66	2011128-2011190	SA1760- <i>sep</i>	downstream of sprG, Pichon and Felden (6)
67	2054610-2054739	SA1814- SA1815	RNAsim
68	2094306-2094381	SA1853- SA1854	
69	2108489-2109159	<i>ilvA</i> -23S	rRNA
70	2109258-2112139	5S-16S	rRNA
71	2112585-2114171	23S-2tRNA	rRNA-tRNA
72	2114295-2114494	16S- SA1867	rRNA
73	2145394-2145685	SA1892- SA1893	Long IGR + Term
74	2212146-2212213	<i>glmS- mtlF</i>	glmS riboswitch
75	2230198-2230858	Arg-5S	rRNA
76	2230957-2232838	5S-16S	rRNA
77	2234884-2237870	23S- SA1969	rRNA
78	2282200-2282274	<i>alsS</i> - SA2010	long IGR
79	2315574-2315703	SA2053- SA2054	RNAsim
80	2326436-2326506	SA2062- <i>moaA</i>	RNAsim

81	2440668-2440861	SA2171- <i>gltT</i>	RNASim, RsaX19
82	2490989-2491119	SA2217- SA2219	RNASim
83	2491406-2491543	SA2217- SA2219	RNASim, sprA2, Pichon and Felden (6)
84	2513401-2513603	<i>opuCA</i> - SA2238	long IGR + Term
85	2544500-2544573	SA2267- SA2268	Upstream of RsaX28
86	2560653-2560793	SA2283-SA2285	Long IGR
87	2561037-2561177	SA2283-SA2285	
88	2561421-2561521	SA2283-SA2285	
89	2561805-2561945	SA2283-SA2285	
90	2562189-2562329	SA2283-SA2285	Long IGR, RNASim
91	2590865-2590957	<i>Fbp</i> -SA2305	sprA3, Pichon and Felden (6), RNASim
92	2652623-2652866	SA2360SA2361	RNASim
93	2769241-2769328	SA2457- <i>icaR</i>	RsaX25

This analysis exploits the composition bias between known RNAs (tRNAs, rRNAs, tmRNA) and the rest of the genome. Known RNAs (underlined in violin), annotated in RFAM database, include stable and universally conserved bacterial RNAs (rRNAs, tRNAs, SRP, 6S RNA, tmRNA, RNase P) and several *cis*-acting leader of mRNAs regulated by metabolites (riboswitches: TPP, thiamine pyrophosphate; SAM, S-adenosyl methionine, FMN, flavin mononucleotide; Gln6P, glucosamine 6-phosphate). STAR and sequence repeats are underlined in orange. This HMM analysis predicts 28 new ncRNAs (no color). Other comments include the Rsa RNAs studied in this work, the IGRs which were predicted by RNASim, and previously identified or predicted ncRNAs.

**Table S3.** Genes showing differential expression between RN6390 and  $\Delta$ *rsaE* mutant strains analyzed by microarray.

Gene name	Function	Fold change	COG category
<b>Genes upregulated by RsaE</b>			
<i>arcA</i>	Arginine deiminase	3,87	E
<i>arcB</i>	Ornithine carbamoyltransferase	2,02	E
<i>arcC</i>	Carbamate kinase	2,08	E
<i>arcD</i>	Arginine/ornithine antiporter	2,51	E
<i>arcR</i>	Transcriptional regulator, Crp/Fnr family	1,96	
<i>atl</i>	Bifunctional autolysin precursor	2,63	MG
<i>beta</i>	Choline dehydrogenase	6,14	E
<i>betB</i>	Betaine aldehyde dehydrogenase	33,97	C
<i>carB</i>	Carbamoyl phosphate synthase large subunit	1,95	E
<i>comK</i>	Competence transcription factor	3,76	
<i>cudT</i>	Osmoprotectant transporter, BCCT family	8,14	M
<i>fdh</i>	Formate dehydrogenase	2,01	C
<i>fmtA</i>	Autolysis and methicillin resistant-related protein	2,29	V
<i>fnbA</i>	Fibronectin-binding protein precursor	1,88	M
<i>glpQ</i>	Glycerophosphoryl diester phosphodiesterase	1,92	M
<i>hlgC</i>	Gamma-hemolysin component C	2,19	V
<i>isdC</i>	Iron transport associated domain protein	1,92	M
SAOUHSC_02608	LysR family regulatory protein	1,8	K
<i>Plc</i>	1-phosphatidylinositol phosphodiesterase	2,17	I
<i>pnbA</i>	Para-nitrobenzyl esterase	1,85	I
<i>purK</i>	Phosphoribosylaminoimidazole carboxylase, ATPase subunit	1,97	F
<i>pyrC</i>	Dihydroorotase	3,93	F
<i>pyrR</i>	Pyrimidine regulatory protein	10,05	F
<i>ribB</i>	Riboflavin synthase subunit alpha	2,77	H
<i>ribD</i>	Riboflavin biosynthesis protein	3,67	H
<i>rnc</i>	Ribonuclease III	1,98	K
SA0423	Hypothetical protein	3,28	S
SA1598	Hypothetical protein	2,79	S
SA2309	Hypothetical protein	1,93	S
SAOUHSC_00025	Putative 5'-nucleotidase	2,16	F
SAOUHSC_00367	Putative sodium:dicarboxylate symporter protein	2,9	R
SAOUHSC_00484	MesJ/Ycf62 family protein	1,83	D
SAOUHSC_00765	Putative comf operon protein 1	2,40	L
SAOUHSC_01036	Hypothetical protein	2,57	S
SAOUHSC_02723	Putative glycerate kinase	2,32	S
SAOUHSC_02770	Similar to diaminopimelate epimerase	3,15	E
SAOUHSC_02794	Hypothetical protein	2,40	S
SAOUHSC_02842	Hypothetical protein	2,63	S

SAOUHSC_02855	LysM domain protein	3,44	R
SAR1150	Antibacterial protein	2,92	V
SAOUHSC_02129	Hypothetical protein	2,17	S
SAV0884	Hypothetical protein	2,46	S
<i>ssaA</i>	Secretory antigen precursor SsaA	5,66	R
<i>sspA</i>	Serine protease; V8 protease; glutamyl endopeptidase	2,65	V
<b>Genes downregulated by RsaE</b>			
<i>bioA</i>	Adenosylmethionine-8-amino-7-oxononanoate	0,33	H
<i>bioB</i>	Biotin synthase	0,18	H
<i>bioD</i>	Dethiobiotin synthase	0,18	H
<i>capI</i>	Capsular polysaccharide synthesis protein CapI	0,48	M
<i>capJ</i>	Capsular polysaccharide synthesis protein CapJ	0,46	M
<i>capL</i>	Capsular polysaccharide synthesis protein CapL	0,55	M
<i>capM</i>	Capsular polysaccharide synthesis protein CapM	0,51	M
<i>capN</i>	Capsular polysaccharide synthesis protein CapN	0,49	M
<i>capO</i>	Capsular polysaccharide synthesis protein CapO	0,53	M
<i>fadA</i>	Acyl-CoA acetyltransferase	0,27	I
<i>fadB</i>	3-hydroxyacyl-CoA dehydrogenase	0,32	I
<i>fadE</i>	Acyl-CoA synthetase	0,46	I
<i>ilvB</i>	Acetolactate synthase small subunit	0,19	E
<i>ilvC</i>	Ketol-acid reductoisomerase	0,18	E
<i>ilvD</i>	Dihydro-acid dehydratase	0,16	E
<i>ilvN</i>	Acetolactate synthase small subunit	0,18	E
<i>leuA</i>	2-isopropylmalate synthase	0,36	E
<i>leuB</i>	3-hydroxyacyl-CoA dehydrogenase	0,3	C
<i>lldp2</i>	L-lactate permease	0,5	C
SAS1649	Putative soluble hydrogenase subunit	0,34	S
<i>oppA</i>	Oligopeptide transport system ATP-binding protein OppA	0,39	E
<i>oppB</i>	Oligopeptide transport system ATP-binding protein OppB*	0,48	E
<i>oppC</i>	Oligopeptide transport system ATP-binding protein OppC	0,43	E
<i>oppD</i>	Oligopeptide transport system ATP-binding protein OppD	0,48	E
<i>oppF</i>	Oligopeptide transport system ATP-binding protein OppF	0,45	E
<i>pgk</i>	Phosphoglycerate kinase	0,53	G
<i>purF</i>	Phosphoribosylpyrophosphate amidotransferase	0,54	F
SACOL2620/SA2393	4-aminobutyrate aminotransferase*	0,17	F
SA0769	ABC transporter ATP-binding protein	0,41	P
SA2491	Hypothetical protein	0,55	S
SACOL1020/SA0873	Hypothetical protein*	0,45	S
<i>Sak</i>	Staphylokinase precursor	0,53	V
SAOUHSC_00401	Hypothetical protein	0,43	S
SAOUHSC_00949	Sodium:alanine symporter family protein	0,51	E
SAOUHSC_01138	Acetyltransferase (GNAT) family protein	0,51	
SAOUHSC_01319	Aspartate kinase	0,35	E

SAOUHSC_02713	8-amino-7-oxononanoate synthase	0,2	H
SAOUHSC_03019	ABC transporter ATP-binding protein	0,55	P
SAOUHSC_03020	Similar to integral membrane protein	0,49	S
SAOUHSC_00842	ABC transporter ATP-binding protein	0,42	P
<i>sucC</i>	Succinyl-CoA synthetase subunit alpha*	0,62	G
<i>sucD</i>	Succinyl-CoA synthetase subunit beta*	0,65	G
<i>thrB</i>	Homoserine kinase	0,42	E
<i>thrC</i>	Threonine synthase	0,23	E

COG category	Number of genes	Functional groups
C	4	Energy production and conversion
D	1	Cell cycle, cell division, chromosome partitioning
E	21	Amino acid transport and metabolism
F	6	Nucleotide transport and metabolism
G	4	Carbohydrate transport and metabolism
H	6	Coenzyme transport and metabolism
I	5	Lipid transport and metabolism
K	2	Transcription
L	1	Replication, recombination and repair
M	12	Cell wall/membrane/envelope biogenesis
P	3	Inorganic ion transport and metabolism
R	3	General function prediction only
S	14	Function unknown
V	3	Defense mechanisms

(Top) Fold change corresponds to RN6390 versus  $\Delta$ *rsaE* mutant data: values > 1 correspond to genes that are overexpressed in RN6390 strain whereas values < 1 correspond to genes that are overexpressed in the  $\Delta$ *rsaE* mutant strain. \*Predicted RsaE-mRNA base pairings. (Bottom) The number of RsaE-dependent regulated genes corresponding to a defined group of function (COG) is given. The complete microarray data set has been posted on the Gene Expression Omnibus database (<http://www.ncbi.nlm.nih.gov/geo/>) under accession number GPL7137 for the platform design and GSE17135 for the original data set.

**Table S4.** List of oligonucleotides.

<b>PCR intergenic regions, northern blot<sup>1</sup></b>			
rsaX01f	CATGCTTGCCTAGGGGTATG	rsaX01r	ATGTCCAAGCTCCATTTTG
rsaX02f	TTTTTCATCATTTACGATACATCATT	rsaX02r	GCCAAGTTACATTGCTTAATCTCC
rsaX03f	TGATCACGATGAAGTCATTCAA	rsaX03r	GGTGCAAATAAAAATACATCAGG
rsaX04f	GAAAAATCGTGTGATTATCTGAGC	rsaX04r	TCAGTCAACTACTGCCAATATAACAT
rsaKf	TTGGTAACCATCACAGCTTAATTT	rsaKr	TTGAAAACGCATACTTCACCA
rsaX05f	GGTCTCGTCTAGGCCAAAGCA	rsaX05r	AGACACTGCATCACGGTACG
rsaAf	CCATTACAAAAATTGTATAGAGTAGC	rsaAr	ACCCGAGTAGTCTTCCTTGG
rsaX06f	CGCTGTGAAGTTAGATAGATGAGTTT	rsaX06r	GCAGTTGCTGCACAAACAAA
rsaX07f	AGGCAATGTGAAAAAGCTGA	rsaX07r	AGGAAGATAGCCGCATAACG
rsaCf	CATCGGATGAGTGTGGCTTA	rsaCr	ACGCCATTCCCTACACTC
rsaDf	CTTGGCTTTTATGGGAAATGA	rsaDr	TCTGCTATAATGTTTTTAACCTTCTCA
rsaX08f	ACGAAAGAAATGGGTTGCTC	rsaX08r	TTCTAGCTGAAACCCTCTATAAAAA
rsaHf	TCATTATTGTTGTTCAAAGGTTATACA	rsaHr	ACGATTAACGTACGGGTCCA
rsaX09f	TGATCGTGATAGCGGATACATTT	rsaX09r	GCTTTTATGTTATATTTTCAAGCAAGA
rsaEf	CCCCTTTGTTTGAAGTGAAAA	rsaEr	AATCGTTAATGTATAAAAAGTGCAT
rsaE-RT	TTTTCACTTCAAACAAAGGGG	rsaE-Cmotif	AAAGGGGATGGGAGAAA
rsaX10f	TCTAGAAATCTTTTTCAATTATATG	rsaX10r	TTCCATGTTGTCAAACCCTTT
rsaX11f	GTATTAATGGCGAGACTCCTGA	rsaX11r	CTTCGTATTGAATGGCTTCG
rsaX12f	ATACTTCGTATTGAATGGCTTCG	rsaX12r	GGACACGAGGCTGGGATATT
rsaX13f	GCTAAAAAGAAGAATCAACAAAAGC	rsaX13r	AAACAGTGGTCTTTATCATTCTG
rsaX14f	AATCCTCAACCAATGCCTACA	rsaX14r	CCATGAAAATGGGATAGCAA
rsaX15f	GGCTGGGACATAAATCCCTAA	rsaX15r	AGACTCCTGAGGGAGCAGTG
rsaX16f	TTAACTTTAACAATGATTAGTTGGCTA	rsaX16r	CATGACATTGATTAGTGAGTAGATTG
rsaX17f	GAAGGGACCCAACACAGAAA	rsaX17r	CTCAGCCTTGGTCTTCGACT
rsalf	CACGTGCTAGCCGACAAATA	rsalr	TAACAGGGGGAGCGATTAAA
rsaX18f	AGTGCGATTTCAAGATGCTG	rsaX18r	TCAAATCAACAACAATAACCATCA
rsaX19f	GCATGCATAAAAGCCCCTAA	rsaX19r	GCCCCTTACAACCTTTTGGT
rsaX20f	TTTCGAGTATCCCTAGCTCGTT	rsaX20r	GCTGCATGTACACAAGGAGTG
rsaJf	GGGCAGATATAAACAGCGATTC	rsaJr	GGGCAATAGTGAGGGGATTT
rsaX21f	TTCCAAATAAATAAAGGGATTTGTGA	rsaX21r	ATGCGTTCATTCATGCCATT
rsaX22f	GCTGAAAGTCACCTAAAATCCTTT	rsaX22r	TGTGTGGAAGACAGGGTAAA
rsaX23f	TTGCGAGTAGTGAGGGGATT	rsaX23r	TTAGTTGATTTTGCATTTTT
rsaX24f	CGCAATTA AAAAGCAATCAACA	rsaX24r	AATGTTATTGTGAATTGCGAGAA
rsaX25f	TACCTTGCGTAACCCGTCAC	rsaX25r	TGACAATACAAATCGATATTGAGAA
rsaGf	GTGCGTGAAGAGATGAAAGATAC	rsaGr	CTTCATAACAGGGGAATTATG

rsaBf	GTGTTAATTGTAGCGCATCCC	rsaBr	GGTTGCAATATTTAATTGGG
RT-RNAIII5	AATACATAGCACTGAGTCCAAGG	RT-RNAIII6	CCTAGATCACAGAGATGTGATGG
6Sf	ATGGGTTTTCTTGCAGCGTA	6Sr	CGTGTACGCTGCAAGAAAACCC
Anti-rsaA	GTAAAGTTCTCCCAAGGAAG		
Anti-rsaC	GGTTTCTCCCCCATAGATC		
Anti-rsaH	TTGACTCCCTTTAGTAGTGG		
Anti-rsaE	ATTTCTCCCATCCCCTTTGT		
Anti-rsaG	GATACTGCTTCTACCTTGC		
5SrRNAr	AACAGGTGTGACCTCCTTGC		

***In vitro* transcription/toeprint**

rsaG-5'- <i>Stu</i> I	TATAGGCCTCTGCCCGACGGCATG TGCG	rsaG-3'- <i>Bam</i> HI	TATGGATCCACAATACTCCCTACAATA TTG
rsaA- <i>Bam</i> HI-T7	AAAGGATCCTAATACGACTCACTATA GGTTAACCATTACAAAAATTGTATAG	rsaA- <i>Dra</i> I- <i>Eco</i> RI	TTGAATTCTTTAAAGTACACTTTGCTC ATAGCAAAGTGTACCC
rsaH- <i>Bam</i> HI-T7	AAAGGATCCTAATACGACTCACTATA GGTACCTTCGATAACGAATAAACATC TC	rsaH- <i>Dra</i> I- <i>Eco</i> RI	TTGAATTCTTTAAAAACGACCCGCAC GATTAACGTACGGG
rsaE- <i>Bam</i> HI-T7	AAAGGATCCTAATACGACTCACTATA GGAAATTAATCACATAACAAACATAC CCCTT	rsaE- <i>Dra</i> I- <i>Eco</i> RI	TTGAATTCTTTAAAAAACGTCGTGTC TGAATACACG
<i>opp</i> B- <i>Bam</i> HI-T7	AAAGGATCCTAATACGACTCACTATA GGAATGATTTTTTAATTTCTTCAGC	<i>opp</i> B- <i>Eco</i> RI	AAAGAATTCTGCGTAGCTACAGGATC
opptoe1	GGCGAACCTGGCATTAAATTC		
sa0873- <i>Bam</i> HI-T7	AAAAAAGGATCCGATCCCGCGAAAT TAATACGACTCACTATAGGTTAAACA AAGGAGGAATTCAGA	sa0873- <i>sa</i> I	AAAAAAGAGCTCAATTTCAAATGGCG CTTTAATTGT

**Deletion/replacement  $\Delta$ rsaE/aphA-3 mutant of *S. aureus* RN6390**

rsaE-140- <i>sm</i> al	TTTCCCGGGAGGT TTA CTACATA TG
rsaE-2041r- <i>Eco</i> RV	TAAGATATCG TTCATAATAT AACATGCTAT C
rsaE-2163- <i>Sal</i> I	CGTGTGCGACA TGTTCTTTTT TAATAAGAGA G
rsaE-3214r- <i>Xho</i> I	TGACTCGAGA TGGCTGGAGA ATTAC
rsaE-70-CT	GTT CAGATAG AGGTAATGAC
rsaE-3283r-CT	GATTA ACTGC TAATCATATG GTG

<sup>1</sup>Oligonucleotides used for the PCR reactions: r is for reverse and f for forward; RT oligonucleotide used for primer extension.

---

# **Conclusion générale & Perspectives**

## L'évolution de l'ARNIII et ses multiples fonctions

L'opéron *agr* régule plus de cent gènes en fonction de la densité cellulaire par le mécanisme de *quorum sensing*. Récemment, il a été montré que cet opéron agit selon deux voies distinctes de régulation: une première dépendante du facteur AgrA qui contrôle l'expression des gènes du métabolisme et les PSM ("Phenol Soluble Modulins"), et une deuxième dépendante de l'ARNIII qui permet la régulation des facteurs de virulence (figure 45) (Queck et al., 2008).

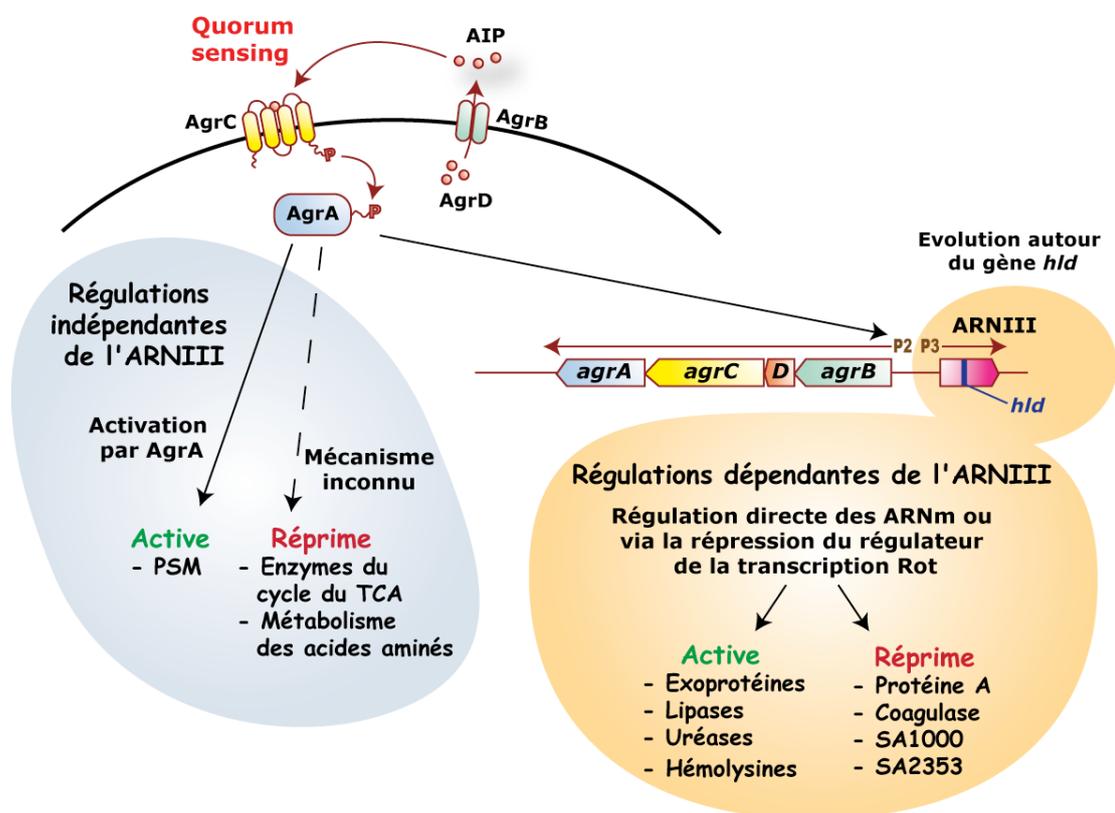


Figure 45: L'ARNIII et AgrA régulent de multiples gènes selon deux voies indépendantes. D'après Queck et al., 2008.

Il est difficile de dire laquelle de ces deux voies est apparue en premier mais l'architecture de l'opéron *agr* indique que la connexion des gènes de virulence au mécanisme de *quorum sensing* serait le résultat d'un événement secondaire de l'évolution. En effet, l'ARNIII code lui même pour une PSM (l'hémolysine  $\delta$ ). L'apparition des régions régulatrices de l'ARNIII autour de sa région codante est un événement plus probable que le réarrangement de l'ARNIII vers l'acquisition d'une phase ouverte de lecture et d'un

promoteur spécifique de AgrA. De plus, la fonction des différentes PSM (dont l'hémolysine  $\delta$ ) semble particulièrement importante dans les phases non infectieuses de la bactérie (i.e. formation des biofilms) (Kong et al., 2006) et ces peptides sont aussi exprimés dans *S. epidermidis* qui reste moins virulente que *S. aureus*. Ainsi, les fonctions régulatrices de l'ARNIII auraient évolué dans un deuxième temps vers la virulence afin de s'adapter à ses nouveaux habitats. Ceci pourrait expliquer pourquoi la plupart des espèces de staphylocoques sont faiblement pathogènes comparé à *S. aureus* qui est apparu plus tardivement dans l'arbre phylogénétique.

L'ARNIII est devenu l'un des principaux régulateurs des gènes de virulence de *S. aureus*. En régulant les protéines de surface et les exoprotéines il est directement responsable du balancement entre la phase de colonisation de l'hôte et la phase de dissémination. Ainsi, il participe entre autre à l'adhésion des bactéries mais également à lutte contre le système immunitaire. Nos derniers résultats montrent aussi que l'ARNIII réprime les ARNm *sa2353* et *lytM* qui codent pour une amidase et une peptidoglycane hydrolase, respectivement (Boisset et al., 2007; Dubrac et al., 2008). Par ailleurs, l'ARNm *sa2093*, une autre cible potentielle de l'ARNIII, code aussi pour une amidase. L'ARNIII pourrait donc également jouer un rôle dans le remodelage de la membrane bactérienne en influençant la synthèse du peptidoglycane en phase stationnaire. Par exemple, il pourrait participer à la régulation de la division cellulaire ou contrôler le processus de lyse dans des conditions défavorables. En effet, T. Geissmann a observé qu'en milieu minimum, des souches délétées de l'ARNIII sont complètement lysées au bout de quelques heures alors que ce phénomène ne s'observe pas avec la souche sauvage isogénique. Par ailleurs, le double mutant ARNIII/LytM (effectué par S. Boisset, INSERM Lyon) retrouve un phénotype sauvage ce qui montre que la dérégulation de *lytM* est directement responsable du phénomène de lyse. Ainsi, l'ensemble des résultats suggèrent que l'ARNIII contribuerait à moduler la composition de la surface bactérienne et à maintenir la structure du peptidoglycane, en accord avec la disponibilité des ressources du milieu extérieur.

### **Une signature possible pour les ARNnc agissant au niveau des ARNm**

Notre travail a révélé l'existence d'un motif conservé UCCC dans l'ARNIII et nombreux ARN Rsa. Bien que nous n'excluons pas totalement que ce motif soit reconnu

par une protéine, nous privilégions l'hypothèse qu'il serait important pour guider les ARNnc vers leurs ARNm cibles. Ce motif permet notamment à l'ARNIII et à RsaE de réprimer plusieurs gènes de virulence et du métabolisme, respectivement. Cette séquence se retrouve sur l'ARNnc SprA3 mais il a également été observé dans d'autres bactéries Gram-positif comme *L. monocytogenes* (Mandin et al., 2007) ou *B. subtilis* (Saito et al., 2009). De manière plus surprenante, ce motif UCCC se retrouve également sur l'ARNnc CyaR de *Salmonella* (Papenfort et al., 2008), et cette séquence interagit directement avec la séquence SD de l'ARNm cible. Ainsi, ce mécanisme de régulation pourrait être conservé entre bactéries Gram-positif et Gram-négatif. Toujours chez *Salmonella*, il a également été montré que l'ARN GcvB réprime la synthèse de toute une classe de protéines membranaires par l'intermédiaire d'un motif G/U conservé et organisé en simple brin (Sharma et al., 2007). Cette séquence s'hybride à des régions riches en A/C qui sont requises pour une initiation optimale de la traduction. Ces exemples montrent qu'un motif conservé en simple brin d'un ARN régulateur peut cibler une multitude d'ARNm aux fonctions reliées qui portent une signature commune. Ce mécanisme est un moyen efficace pour coordonner la synthèse de ces protéines.

Pour la première fois, ces différents motifs pourraient représenter un véritable critère de choix pour rechercher de nouveaux ARNnc mais également leurs cibles. Par exemple, le motif UCCC serait un excellent critère chez les bactéries dont le génome possède un faible pourcentage de bases G/C et dont l'initiation de la traduction est directement dépendante d'une forte séquence SD. Ainsi, il est probable que les années à venir voient l'émergence de plusieurs classes d'ARNm cibles, qui seraient réprimés par la signature spécifique d'un ARNnc.

### **Fonctions et mode d'action des ARN Rsa**

Nous avons identifié 11 nouveaux ARNnc chez *S. aureus*. Dans un futur proche, notre premier objectif sera de déterminer les cibles directes de ces ARN afin d'identifier leurs fonctions dans la bactérie. Si dans le cas de l'ARNIII, la RNase III est le seul cofacteur protéique qui ait été identifié jusqu'à présent, dans les bactéries Gram-négatif, la plupart des ARNnc sont aidés par l'action de la protéine Hfq et de la RNase E. Il est probable que certains des ARN Rsa fonctionnent selon des modes d'action différents de l'ARNIII et nécessitent des protéines pour optimiser leur action. Nous envisageons

d'exprimer des ARN portant une étiquette dans des souches de *S. aureus* délétées du gène codant pour l'ARN à étudier (Said et al., 2009). Après purification des complexes par chromatographie d'affinité, les ligands (protéines, ARNm) associés à cet ARN seront analysés par spectrométrie de masse (protéine) ou par séquençage (ARN). De plus, E. Lioliou dans l'équipe réalise actuellement des immunoprécipitations avec une RNase III inactive portant une étiquette FLAG. Les ARN coimmunoprécipités sont actuellement séquencés par l'approche de "deep sequencing". Ceci nous permettra d'identifier les ARNm substrats de la RNase III mais également d'éventuels duplexes ARNnc-ARNm cible.

L'existence biologique des 11 ARN Rsa a été validée dans des conditions de culture standard dans un milieu riche. Notre analyse bioinformatique suggère que de nombreuses régions intergéniques pourraient être fonctionnelles. Certaines de ces régions sont probablement exprimées selon des conditions particulières de stress; il ne fait donc aucun doute que de nombreux candidats aient échappé à notre attention. Ainsi, il serait intéressant d'étudier dans différentes souches de *S. aureus* (sauvage ou délété pour l'un des ARN Rsa) et en fonction de différents stress (choc thermique, stress acide, choix de la source carbonée, absence d'un métabolite, carence en acide aminé...) l'expression de ces ARN. Pour cela, il deviendra indispensable de réaliser des analyses à grande échelle en utilisant des techniques de séquençage haut débit. Cette étude sera réalisée en collaboration avec l'équipe de J. Schrenzel (Genève). Enfin, nous pourrions envisager la technique de "tiling arrays". Ces deux approches seront utiles pour répondre à une question donnée (comparaison entre deux souches sauvage et mutante, réponse à un stress particulier), et donneront des informations précieuses sur les différentes classes d'ARN produits et sur les opérons exprimés. Cependant ces techniques requièrent la synthèse d'ADNc ce qui introduit un biais. Par ailleurs des ARN peuvent être aussi générés par interférence transcriptionnelle, et ne pas être fonctionnel. Enfin, pour compléter les réseaux de régulation, il faudra caractériser les inducteurs de la synthèse des ARNnc, et définir comment leur synthèse et leur stabilité sont contrôlées. Existe t'il une hiérarchie entre ces ARNnc ?

Un autre déficit actuel est de définir l'impact de ces ARN régulateurs sur la virulence de *S. aureus* et sur les relations hôte-pathogène. Est-ce que ces ARN régulateurs sont essentiels pour la survie de la bactérie au sein de l'hôte ? Sont-ils exprimés dans les tissus infectés de patients ? On peut espérer que l'étude de ces nouveaux régulateurs et de leur

mécanisme d'action puisse révéler de nouvelles stratégies pour inhiber la croissance de *S. aureus* ou son pouvoir pathogène.

## Conclusion générale

Finalelement, l'état actuel des connaissances montre qu'il existe de nombreuses similarités entre les ARN régulateurs eucaryotiques et bactériens. En effet, il est intéressant de comparer les ARN antisens bactériens avec les miARN eucaryotiques. Ces deux types d'ARN se retrouvent impliqués dans la plupart des processus cellulaires et agissent principalement au niveau post-transcriptionnel en s'hybridant avec leurs ARNm cibles. Alors que ces hybridations se font dans la région 3'UTR des ARNm eucaryotiques (Filipowicz et al., 2008), elles se produisent majoritairement au niveau du site RBS des ARNm bactériens (Guillier et al., 2006; Storz et al., 2004; Vogel et Papenfort, 2006). Les régions ciblées sont donc clairement différentes, probablement liées aux mécanismes de traduction, mais ces interactions peuvent toutefois induire des effets analogues comme la répression traductionnelle et/ou la dégradation des ARNm. Enfin, ces deux types d'ARN sont également associés à des machineries protéiques spécifiques comme le complexe RISC chez les eucaryotes, le duo Hfq-RNase E chez *E. coli* ou encore la RNase III chez *S. aureus*. Dans chacun de ces cas, l'ARN antisens est lui directement responsable de la spécificité de reconnaissance des ARNm cibles.

Une autre comparaison qu'il est également tentant de faire concerne les siARN et les ARNcr bactériens (ARN CRISPR). Ces deux types d'ARN sont tous deux maturés et guidés par des machineries protéiques analogues et constituent un moyen de défense efficace contre diverses infections virales. Cependant, alors que les siARN ciblent uniquement les ARN, de récentes études montrent que les ARNcr sont capables de s'hybrider à des molécules d'ARN et d'ADN (Marraffini et Sontheimer, 2008). Ainsi, ces deux systèmes sont proches mais auraient tout de même évolué de manière différente au cours des années. D'une certaine manière, ceci reflète la grande capacité de l'ARN à s'adapter à des besoins différents comme c'est le cas entre une cellule eucaryote et une cellule bactérienne.

Ainsi, l'ensemble des informations traitées dans ce manuscrit souligne l'importance que joue l'ARN dans le monde du vivant. Même si nos connaissances ont largement évolué

ces dernières années, de plus en plus, on peut s'attendre à identifier des ARN associés à des fonctions insoupçonnées. Aujourd'hui, le "monde à ARN" n'est plus seulement considéré comme l'échafaudage de la cellule mais plutôt comme l'un des acteurs principaux de la régulation des gènes qui fonctionne de manière coordonné avec le monde "moderne" à ADN. Bien sûr, l'ARN n'a pas encore dévoilé tous ses secrets, c'est pourquoi l'identification de ses fonctions et l'étude de ses mécanismes d'action sont des choses qui restent essentielles pour le futur. A terme, ces études nous permettrons de définir les rôles exacts de l'ARN dans les réseaux complexes de régulation présents au sein de chaque cellule.

---

# Bibliographie

- A**fonyushkin, T., Vecerek, B., Moll, I., Blasi, U., and Kaberdin, V. R. (2005). Both RNase E and RNase III control the stability of *sodB* mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic Acids Res* *33*, 1678-1689.
- Aiba, H. (2007). Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr Opin Microbiol* *10*, 134-139.
- Altuvia, S. (2007). Identification of bacterial small non-coding RNAs: experimental approaches. *Curr Opin Microbiol* *10*, 257-261.
- Altuvia, S., Locker-Giladi, H., Koby, S., Ben-Nun, O., and Oppenheim, A. B. (1987). RNase III stimulates the translation of the *cIII* gene of bacteriophage lambda. *Proc Natl Acad Sci U S A* *84*, 6511-6515.
- Amaral, P. P., and Mattick, J. S. (2008). Noncoding RNA in development. *Mamm Genome* *19*, 454-492.
- Argaman, L., and Altuvia, S. (2000). *fhlA* repression by OxyS RNA: kissing complex formation at two sites results in a stable antisense-target RNA complex. *J Mol Biol* *300*, 1101-1112.
- Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E. G., Margalit, H., and Altuvia, S. (2001). Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr Biol* *11*, 941-950.
- Arvidson, S., and Tegmark, K. (2001). Regulation of virulence determinants in *Staphylococcus aureus*. *Int J Med Microbiol* *291*, 159-170.
- Axmann, I. M., Kensche, P., Vogel, J., Kohl, S., Herzel, H., and Hess, W. R. (2005). Identification of cyanobacterial non-coding RNAs by comparative genome analysis. *Genome Biol* *6*, R73.
- B**abitzke, P., Baker, C. S., and Romeo, T. (2009). Regulation of Translation Initiation by RNA Binding Proteins. *Annu Rev Microbiol*
- Babitzke, P., and Romeo, T. (2007). CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Curr Opin Microbiol* *10*, 156-163.
- Bachellet, J. P., Cavaille, J., and Huttenhofer, A. (2002). The expanding snoRNA world. *Biochimie* *84*, 775-790.
- Ballal, A., Ray, B., and Manna, A. C. (2009). *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* *191*, 1656-1665.
- Barcelo, M., Chauvet, E., Boukhari, R., and Mbieleu, B. (2009). [Staphylococcal necrotizing pneumopathy due to Panton-Valentine leukocidin toxin with good outcome]. *Arch Pediatr* *16*, 32-36.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* *315*, 1709-1712.
- Barrick, J. E., Sudarsan, N., Weinberg, Z., Ruzzo, W. L., and Breaker, R. R. (2005). 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. *RNA* *11*, 774-784.
- Bassler, B. L. (2002). Small talk. Cell-to-cell communication in bacteria. *Cell* *109*, 421-424.
- Bejerano-Sagie, M., and Xavier, K. B. (2007). The role of small RNAs in quorum sensing. *Curr Opin Microbiol* *10*, 189-198.
- Beltran, M., Puig, I., Pena, C., Garcia, J. M., Alvarez, A. B., Pena, R., Bonilla, F., and de Herreros, A. G. (2008). A natural antisense transcript regulates *Zeb2/Sip1* gene

- expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev* 22, 756-769.
- Benito, Y., Kolb, F. A., Romby, P., Lina, G., Etienne, J., and Vandenesch, F. (2000). Probing the structure of RNAIII, the *Staphylococcus aureus* agr regulatory RNA, and identification of the RNA domain involved in repression of protein A expression. *RNA* 6, 668-679.
- Bhattacharyya, S. N., and Filipowicz, W. (2007). Argonautes and company: sailing against the wind. *Cell* 128, 1027-1028.
- Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I., and Filipowicz, W. (2006). Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell* 125, 1111-1124.
- Bischoff, M., Entenza, J. M., and Giachino, P. (2001). Influence of a functional sigB operon on the global regulators sar and agr in *Staphylococcus aureus*. *J Bacteriol* 183, 5171-5179.
- Blevins, J. S., Gillaspay, A. F., Rehtin, T. M., Hurlburt, B. K., and Smeltzer, M. S. (1999). The *Staphylococcal* accessory regulator (sar) represses transcription of the *Staphylococcus aureus* collagen adhesin gene (cna) in an agr-independent manner. *Mol Microbiol* 33, 317-326.
- Blomberg, P., Engdahl, H. M., Malmgren, C., Romby, P., and Wagner, E. G. (1994). Replication control of plasmid R1: disruption of an inhibitory RNA structure that sequesters the repA ribosome-binding site permits tap-independent RepA synthesis. *Mol Microbiol* 12, 49-60.
- Blomberg, P., Nordstrom, K., and Wagner, E. G. (1992). Replication control of plasmid R1: RepA synthesis is regulated by CopA RNA through inhibition of leader peptide translation. *EMBO J* 11, 2675-2683.
- Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M., Chevalier, C., Helfer, A. C., Benito, Y., Jacquier, A., Gaspin, C., Vandenesch, F., and Romby, P. (2007). *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev* 21, 1353-1366.
- Bouvier, M., Sharma, C. M., Mika, F., Nierhaus, K. H., and Vogel, J. (2008). Small RNA binding to 5' mRNA coding region inhibits translational initiation. *Mol Cell* 32, 827-837.
- Brantl, S. (2002). Antisense RNAs in plasmids: control of replication and maintenance. *Plasmid* 48, 165-173.
- Brennan, R. G., and Link, T. M. (2007). Hfq structure, function and ligand binding. *Curr Opin Microbiol* 10, 125-133.
- Brodersen, P., and Voinnet, O. (2009). Revisiting the principles of microRNA target recognition and mode of action. *Nat Rev Mol Cell Biol* 10, 141-148.
- Bronner, S., Stoessel, P., Gravet, A., Monteil, H., and Prevost, G. (2000). Variable expressions of *Staphylococcus aureus* bicomponent leucotoxins semiquantified by competitive reverse transcription-PCR. *Appl Environ Microbiol* 66, 3931-3938.
- Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuys, R. J., Snijders, A. P., Dickman, M. J., Makarova, K. S., Koonin, E. V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321, 960-964.
- Brunskill, E. W., and Bayles, K. W. (1996). Identification of LytSR-regulated genes from *Staphylococcus aureus*. *J Bacteriol* 178, 5810-5812.

- Caillet, J., Nogueira, T., Masquida, B., Winter, F., Graffe, M., Dock-Bregeon, A. C., Torres-Larios, A., Sankaranarayanan, R., Westhof, E., Ehresmann, B., Ehresmann, C., Romby, P., and Springer, M. (2003). The modular structure of *Escherichia coli* threonyl-tRNA synthetase as both an enzyme and a regulator of gene expression. *Mol Microbiol* 47, 961-974.
- Cam, H. P., Chen, E. S., and Grewal, S. I. (2009). Transcriptional scaffolds for heterochromatin assembly. *Cell* 136, 610-614.
- Carruthers, M. M., and Kabat, W. J. (1983). Mediation of staphylococcal adherence to mucosal cells by lipoteichoic acid. *Infect Immun* 40, 444-446.
- Carte, J., Wang, R., Li, H., Terns, R. M., and Terns, M. P. (2008). Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev* 22, 3489-3496.
- Chakrabarti, S. K., and Misra, T. K. (2000). SarA represses agr operon expression in a purified in vitro *Staphylococcus aureus* transcription system. *J Bacteriol* 182, 5893-5897.
- Chamberlain, N. R., and Imanoel, B. (1996). Genetic regulation of fatty acid modifying enzyme from *Staphylococcus aureus*. *J Med Microbiol* 44, 125-129.
- Chavakis, T., Hussain, M., Kanse, S. M., Peters, G., Bretzel, R. G., Flock, J. I., Herrmann, M., and Preissner, K. T. (2002). *Staphylococcus aureus* extracellular adherence protein serves as anti-inflammatory factor by inhibiting the recruitment of host leukocytes. *Nat Med* 8, 687-693.
- Chavakis, T., Wiechmann, K., Preissner, K. T., and Herrmann, M. (2005). *Staphylococcus aureus* interactions with the endothelium: the role of bacterial "secretable expanded repertoire adhesive molecules" (SERAM) in disturbing host defense systems. *Thromb Haemost* 94, 278-285.
- Chen, S., Lesnik, E. A., Hall, T. A., Sampath, R., Griffey, R. H., Ecker, D. J., and Blyn, L. B. (2002). A bioinformatics based approach to discover small RNA genes in the *Escherichia coli* genome. *Biosystems* 65, 157-177.
- Chendrimada, T. P., Finn, K. J., Ji, X., Baillat, D., Gregory, R. I., Liebhaber, S. A., Pasquinelli, A. E., and Shiekhattar, R. (2007). MicroRNA silencing through RISC recruitment of eIF6. *Nature* 447, 823-828.
- Cheung, A. L., Bayer, A. S., Zhang, G., Gresham, H., and Xiong, Y. Q. (2004). Regulation of virulence determinants in vitro and in vivo in *Staphylococcus aureus*. *FEMS Immunol Med Microbiol* 40, 1-9.
- Cheung, A. L., Bayer, M. G., and Heinrichs, J. H. (1997a). sar Genetic determinants necessary for transcription of RNAII and RNAIII in the agr locus of *Staphylococcus aureus*. *J Bacteriol* 179, 3963-3971.
- Cheung, A. L., Eberhardt, K., and Heinrichs, J. H. (1997b). Regulation of protein A synthesis by the sar and agr loci of *Staphylococcus aureus*. *Infect Immun* 65, 2243-2249.
- Cheung, A. L., and Fischetti, V. A. (1990). The role of fibrinogen in staphylococcal adherence to catheters in vitro. *J Infect Dis* 161, 1177-1186.
- Cheung, A. L., Koomey, J. M., Butler, C. A., Projan, S. J., and Fischetti, V. A. (1992). Regulation of exoprotein expression in *Staphylococcus aureus* by a locus (sar) distinct from agr. *Proc Natl Acad Sci U S A* 89, 6462-6466.
- Cheung, A. L., Nishina, K. A., Trottonda, M. P., and Tamber, S. (2008). The SarA protein family of *Staphylococcus aureus*. *Int J Biochem Cell Biol* 40, 355-361.

- Clemson, C. M., McNeil, J. A., Willard, H. F., and Lawrence, J. B. (1996). XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J Cell Biol* 132, 259-275.
- Collins, J. A., Irnov, I., Baker, S., and Winkler, W. C. (2007). Mechanism of mRNA destabilization by the glmS ribozyme. *Genes Dev* 21, 3356-3368.
- Condon, C. (2007). Maturation and degradation of RNA in bacteria. *Curr Opin Microbiol* 10, 271-278.
- Consortium (2004). Finishing the euchromatic sequence of the human genome. *Nature* 431, 931-945.
- Corrigan, R. M., Miajlovic, H., and Foster, T. J. (2009). Surface proteins that promote adherence of *Staphylococcus aureus* to human desquamated nasal epithelial cells. *BMC Microbiol* 9, 22.
- Cruz, J. A., and Westhof, E. (2009). The dynamic landscapes of RNA architecture. *Cell* 136, 604-609.

- D**ambach, M. D., and Winkler, W. C. (2009). Expanding roles for metabolite-sensing regulatory RNAs. *Curr Opin Microbiol* 12, 161-169.
- Daou-Chabo, R., Mathy, N., Benard, L., and Condon, C. (2009). Ribosomes initiating translation of the hbs mRNA protect it from 5'-to-3' exoribonucleolytic degradation by RNase J1. *Mol Microbiol* 71, 1538-1550.
- Darfeuille, F., Unoson, C., Vogel, J., and Wagner, E. G. (2007). An antisense RNA inhibits translation by competing with standby ribosomes. *Mol Cell* 26, 381-392.
- de Smit, M. H., and van Duin, J. (2003). Translational standby sites: how ribosomes may deal with the rapid folding kinetics of mRNA. *J Mol Biol* 331, 737-743.
- Desnoyers, G., Morissette, A., Prevost, K., and Masse, E. (2009). Small RNA-induced differential degradation of the polycistronic mRNA iscRSUA. *EMBO J* 28, 1551-1561.
- Dinger, M. E., Amaral, P. P., Mercer, T. R., Pang, K. C., Bruce, S. J., Gardiner, B. B., Askarian-Amiri, M. E., Ru, K., Solda, G., Simons, C., Sunkin, S. M., Crowe, M. L., Grimmond, S. M., Perkins, A. C., and Mattick, J. S. (2008). Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res* 18, 1433-1445.
- Doudna, J. A., and Cech, T. R. (2002). The chemical repertoire of natural ribozymes. *Nature* 418, 222-228.
- Drider, D., and Condon, C. (2004). The continuing story of endoribonuclease III. *J Mol Microbiol Biotechnol* 8, 195-200.
- Dubrac, S., Bisicchia, P., Devine, K. M., and Msadek, T. (2008). A matter of life and death: cell wall homeostasis and the WalKR (YycGF) essential signal transduction pathway. *Mol Microbiol* 70, 1307-1322.
- Dubrac, S., and Msadek, T. (2004). Identification of genes controlled by the essential YycG/YycF two-component system of *Staphylococcus aureus*. *J Bacteriol* 186, 1175-1181.
- Dufourc, E. J., Dufourcq, J., Birkbeck, T. H., and Freer, J. H. (1990). Delta-haemolysin from *Staphylococcus aureus* and model membranes. A solid-state 2H-NMR and 31P-NMR study. *Eur J Biochem* 187, 581-587.
- Dunman, P. M., Murphy, E., Haney, S., Palacios, D., Tucker-Kellogg, G., Wu, S., Brown, E. L., Zagursky, R. J., Shlaes, D., and Projan, S. J. (2001). Transcription profiling-based identification of *Staphylococcus aureus* genes regulated by the agr and/or sarA loci. *J Bacteriol* 183, 7341-7353.

- E**ntenza, J. M., Moreillon, P., Senn, M. M., Kormanec, J., Dunman, P. M., Berger-Bachi, B., Projan, S., and Bischoff, M. (2005). Role of sigmaB in the expression of *Staphylococcus aureus* cell wall adhesins ClfA and FnbA and contribution to infectivity in a rat model of experimental endocarditis. *Infect Immun* *73*, 990-998.
- Escolar, L., Perez-Martin, J., and de Lorenzo, V. (1999). Opening the iron box: transcriptional metalloregulation by the Fur protein. *J Bacteriol* *181*, 6223-6229.
- Espersen, F. (1985). Complement activation by clumping factor and protein A from *Staphylococcus aureus* strain E 2371. *Acta Pathol Microbiol Immunol Scand C* *93*, 59-64.
- Eulalio, A., Rehwinkel, J., Stricker, M., Huntzinger, E., Yang, S. F., Doerks, T., Dorner, S., Bork, P., Boutros, M., and Izaurralde, E. (2007). Target-specific requirements for enhancers of decapping in miRNA-mediated gene silencing. *Genes Dev* *21*, 2558-2570.
- F**eng, J., Bi, C., Clark, B. S., Mady, R., Shah, P., and Kohtz, J. D. (2006). The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev* *20*, 1470-1484.
- Figueroa-Bossi, N., Valentini, M., Malleret, L., and Bossi, L. (2009). Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Genes Dev*
- Filipowicz, W., Bhattacharyya, S. N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* *9*, 102-114.
- Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* *391*, 806-811.
- Forterre, P. (2006). The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res* *117*, 5-16.
- Forterre, P., and Gribaldo, S. (2007). The origin of modern terrestrial life. *HFSP J* *1*, 156-168.
- Foster, T. J. (2004). The *Staphylococcus aureus* "superbug". *J Clin Invest* *114*, 1693-1696.
- Foster, T. J. (2005). Immune evasion by staphylococci. *Nat Rev Microbiol* *3*, 948-958.
- Foster, T. J., and Hook, M. (1998). Surface protein adhesins of *Staphylococcus aureus*. *Trends Microbiol* *6*, 484-488.
- Fozo, E. M., Hemm, M. R., and Storz, G. (2008). Small toxic proteins and the antisense RNAs that repress them. *Microbiol Mol Biol Rev* *72*, 579-89, Table of Contents.
- Frith, M. C., Pheasant, M., and Mattick, J. S. (2005). The amazing complexity of the human transcriptome. *Eur J Hum Genet* *13*, 894-897.
- G**an, J., Tropea, J. E., Austin, B. P., Court, D. L., Waugh, D. S., and Ji, X. (2006). Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. *Cell* *124*, 355-366.
- Garst, A. D., Heroux, A., Rambo, R. P., and Batey, R. T. (2008). Crystal structure of the lysine riboswitch regulatory mRNA element. *J Biol Chem* *283*, 22347-22351.
- Garvis, S., Mei, J. M., Ruiz-Albert, J., and Holden, D. W. (2002). *Staphylococcus aureus* svrA: a gene required for virulence and expression of the agr locus. *Microbiology* *148*, 3235-3243.

- Gaskill, M. E., and Khan, S. A. (1988). Regulation of the enterotoxin B gene in *Staphylococcus aureus*. *J Biol Chem* *263*, 6276-6280.
- Geisinger, E., Adhikari, R. P., Jin, R., Ross, H. F., and Novick, R. P. (2006). Inhibition of rot translation by RNAIII, a key feature of agr function. *Mol Microbiol* *61*, 1038-1048.
- Geissmann, T., Possedko, M., Huntzinger, E., Fechter, P., Ehresmann, C., and Romby, P. (2006). Regulatory RNAs as mediators of virulence gene expression in bacteria. *Handb Exp Pharmacol* 9-43.
- Geissmann, T. A., and Touati, D. (2004). Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *EMBO J* *23*, 396-405.
- Genestier, A. L., Michallet, M. C., Prevost, G., Bellot, G., Chalabrysse, L., Peyrol, S., Thivolet, F., Etienne, J., Lina, G., Vallette, F. M., Vandenesch, F., and Genestier, L. (2005). *Staphylococcus aureus* Panton-Valentine leukocidin directly targets mitochondria and induces Bax-independent apoptosis of human neutrophils. *J Clin Invest* *115*, 3117-3127.
- Gerdes, K., and Wagner, E. G. (2007). RNA antitoxins. *Curr Opin Microbiol* *10*, 117-124.
- Gertz, S., Engelmann, S., Schmid, R., Ziebandt, A. K., Tischer, K., Scharf, C., Hacker, J., and Hecker, M. (2000). Characterization of the sigma(B) regulon in *Staphylococcus aureus*. *J Bacteriol* *182*, 6983-6991.
- Gilbert, S. D., Rambo, R. P., Van Tyne, D., and Batey, R. T. (2008). Structure of the SAM-II riboswitch bound to S-adenosylmethionine. *Nat Struct Mol Biol* *15*, 177-182.
- Gildehaus, N., Neusser, T., Wurm, R., and Wagner, R. (2007). Studies on the function of the riboregulator 6S RNA from *E. coli*: RNA polymerase binding, inhibition of in vitro transcription and synthesis of RNA-directed de novo transcripts. *Nucleic Acids Res* *35*, 1885-1896.
- Gill, S. R., Fouts, D. E., Archer, G. L., Mongodin, E. F., Deboy, R. T., Ravel, J., Paulsen, I. T., Kolonay, J. F., Brinkac, L., Beanan, M., Dodson, R. J., Daugherty, S. C., Madupu, R., Angiuoli, S. V., Durkin, A. S., Haft, D. H., Vamathevan, J., Khouri, H., Utterback, T., Lee, C., Dimitrov, G., Jiang, L., Qin, H., Weidman, J., Tran, K., Kang, K., Hance, I. R., Nelson, K. E., and Fraser, C. M. (2005). Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. *J Bacteriol* *187*, 2426-2438.
- Giraud, A. T., Calzolari, A., Cataldi, A. A., Bogni, C., and Nagel, R. (1999). The sae locus of *Staphylococcus aureus* encodes a two-component regulatory system. *FEMS Microbiol Lett* *177*, 15-22.
- Giraud, A. T., Cheung, A. L., and Nagel, R. (1997). The sae locus of *Staphylococcus aureus* controls exoprotein synthesis at the transcriptional level. *Arch Microbiol* *168*, 53-58.
- Giraud, A. T., Raspanti, C. G., Calzolari, A., and Nagel, R. (1994). Characterization of a Tn551-mutant of *Staphylococcus aureus* defective in the production of several exoproteins. *Can J Microbiol* *40*, 677-681.
- Gomez, M. I., Lee, A., Reddy, B., Muir, A., Soong, G., Pitt, A., Cheung, A., and Prince, A. (2004). *Staphylococcus aureus* protein A induces airway epithelial inflammatory responses by activating TNFR1. *Nat Med* *10*, 842-848.
- Gomez, M. I., O'Seaghda, M., Magargee, M., Foster, T. J., and Prince, A. S. (2006). *Staphylococcus aureus* protein A activates TNFR1 signaling through conserved IgG binding domains. *J Biol Chem* *281*, 20190-20196.

- Goodyear, C. S., and Silverman, G. J. (2004). Staphylococcal toxin induced preferential and prolonged in vivo deletion of innate-like B lymphocytes. *Proc Natl Acad Sci U S A* *101*, 11392-11397.
- Goodyear, C. S., Sugiyama, F., and Silverman, G. J. (2006). Temporal and dose-dependent relationships between in vivo B cell receptor-targeted proliferation and deletion-induced by a microbial B cell toxin. *J Immunol* *176*, 2262-2271.
- Gorke, B., and Vogel, J. (2008). Noncoding RNA control of the making and breaking of sugars. *Genes Dev* *22*, 2914-2925.
- Gottesman, S. (2004). The small RNA regulators of *Escherichia coli*: roles and mechanisms\*. *Annu Rev Microbiol* *58*, 303-328.
- Gottesman, S. (2005). Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends Genet* *21*, 399-404.
- Gregory, B. D., Yazaki, J., and Ecker, J. R. (2008). Utilizing tiling microarrays for whole-genome analysis in plants. *Plant J* *53*, 636-644.
- Grundmeier, M., Hussain, M., Becker, P., Heilmann, C., Peters, G., and Sinha, B. (2004). 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* *72*, 7155-7163.
- Grundy, F. J., and Henkin, T. M. (2003). The T box and S box transcription termination control systems. *Front Biosci* *8*, d20-31.
- Grundy, F. J., and Henkin, T. M. (2006). From ribosome to riboswitch: control of gene expression in bacteria by RNA structural rearrangements. *Crit Rev Biochem Mol Biol* *41*, 329-338.
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* *35*, 849-857.
- Guillier, M., and Gottesman, S. (2006). Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. *Mol Microbiol* *59*, 231-247.
- Guillier, M., and Gottesman, S. (2008). The 5' end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. *Nucleic Acids Res* *36*, 6781-6794.
- Guillier, M., Gottesman, S., and Storz, G. (2006). Modulating the outer membrane with small RNAs. *Genes Dev* *20*, 2338-2348.

- H**agggar, A., Hussain, M., Lonnie, H., Herrmann, M., Norrby-Teglund, A., and Flock, J. I. (2003). Extracellular adherence protein from *Staphylococcus aureus* enhances internalization into eukaryotic cells. *Infect Immun* *71*, 2310-2317.
- Halfmann, A., Kovacs, M., Hakenbeck, R., and Bruckner, R. (2007). Identification of the genes directly controlled by the response regulator CiaR in *Streptococcus pneumoniae*: five out of 15 promoters drive expression of small non-coding RNAs. *Mol Microbiol* *66*, 110-126.
- Halic, M., and Beckmann, R. (2005). The signal recognition particle and its interactions during protein targeting. *Curr Opin Struct Biol* *15*, 116-125.
- Hamilton, A., Voinnet, O., Chappell, L., and Baulcombe, D. (2002). Two classes of short interfering RNA in RNA silencing. *EMBO J* *21*, 4671-4679.
- Hansen, A. M., and Kaper, J. B. (2009). Hfq affects the expression of the LEE pathogenicity island in enterohaemorrhagic *Escherichia coli*. *Mol Microbiol* *73*, 446-465.
- Hansen, U., Hussain, M., Villone, D., Herrmann, M., Robenek, H., Peters, G., Sinha, B., and Bruckner, P. (2006). The anchorless adhesin Eap (extracellular adherence

- protein) from *Staphylococcus aureus* selectively recognizes extracellular matrix aggregates but binds promiscuously to monomeric matrix macromolecules. *Matrix Biol* 25, 252-260.
- Hantke, K. (2001). Iron and metal regulation in bacteria. *Curr Opin Microbiol* 4, 172-177.
- Hartleib, J., Kohler, N., Dickinson, R. B., Chhatwal, G. S., Sixma, J. J., Hartford, O. M., Foster, T. J., Peters, G., Kehrel, B. E., and Herrmann, M. (2000). Protein A is the von Willebrand factor binding protein on *Staphylococcus aureus*. *Blood* 96, 2149-2156.
- Hartz, D., McPheeters, D. S., Traut, R., and Gold, L. (1988). Extension inhibition analysis of translation initiation complexes. *Methods Enzymol* 164, 419-425.
- Heilmann, C., Niemann, S., Sinha, B., Herrmann, M., Kehrel, B. E., and Peters, G. (2004). *Staphylococcus aureus* fibronectin-binding protein (FnBP)-mediated adherence to platelets, and aggregation of platelets induced by FnBPA but not by FnBPB. *J Infect Dis* 190, 321-329.
- Henkin, T. M. (2008). Riboswitch RNAs: using RNA to sense cellular metabolism. *Genes Dev* 22, 3383-3390.
- Hindley, J. (1967). Fractionation of <sup>32</sup>P-labelled ribonucleic acids on polyacrylamide gels and their characterization by fingerprinting. *J Mol Biol* 30, 125-136.
- Hiramatsu, K. (2001). Vancomycin-resistant *Staphylococcus aureus*: a new model of antibiotic resistance. *Lancet Infect Dis* 1, 147-155.
- Hiramatsu, K., Cui, L., Kuroda, M., and Ito, T. (2001). The emergence and evolution of methicillin-resistant *Staphylococcus aureus*. *Trends Microbiol* 9, 486-493.
- Hjalt, T. A., and Wagner, E. G. (1995). Bulged-out nucleotides in an antisense RNA are required for rapid target RNA binding in vitro and inhibition in vivo. *Nucleic Acids Res* 23, 580-587.
- Hochhut, B., Dobrindt, U., and Hacker, J. (2005). Pathogenicity islands and their role in bacterial virulence and survival. *Contrib Microbiol* 12, 234-254.
- Holden, M. T., Feil, E. J., Lindsay, J. A., Peacock, S. J., Day, N. P., Enright, M. C., Foster, T. J., Moore, C. E., Hurst, L., Atkin, R., Barron, A., Bason, N., Bentley, S. D., Chillingworth, C., Chillingworth, T., Churcher, C., Clark, L., Corton, C., Cronin, A., Doggett, J., Dowd, L., Feltwell, T., Hance, Z., Harris, B., Hauser, H., Holroyd, S., Jagels, K., James, K. D., Lennard, N., Line, A., Mayes, R., Moule, S., Mungall, K., Ormond, D., Quail, M. A., Rabinowitsch, E., Rutherford, K., Sanders, M., Sharp, S., Simmonds, M., Stevens, K., Whitehead, S., Barrell, B. G., Spratt, B. G., and Parkhill, J. (2004). Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. *Proc Natl Acad Sci U S A* 101, 9786-9791.
- Hopkins, J. F., Panja, S., McNeil, S. A., and Woodson, S. A. (2009). Effect of salt and RNA structure on annealing and strand displacement by Hfq. *Nucleic Acids Res*
- 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. (2005). *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate spa gene expression. *EMBO J* 24, 824-835.
- I**ngavale, S., van Wamel, W., Luong, T. T., Lee, C. Y., and Cheung, A. L. (2005). Rat/MgrA, a regulator of autolysis, is a regulator of virulence genes in *Staphylococcus aureus*. *Infect Immun* 73, 1423-1431.

- J**acob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* 3, 318-356.
- Jobert, L., Pinzon, N., Van Herreweghe, E., Jady, B. E., Guialis, A., Kiss, T., and Tora, L. (2009). Human U1 snRNA forms a new chromatin-associated snRNP with TAF15. *EMBO Rep* 10, 494-500.
- Johansen, J., Rasmussen, A. A., Overgaard, M., and Valentin-Hansen, P. (2006). Conserved small non-coding RNAs that belong to the sigmaE regulon: role in down-regulation of outer membrane proteins. *J Mol Biol* 364, 1-8.
- Johansson, J., and Cossart, P. (2003). RNA-mediated control of virulence gene expression in bacterial pathogens. *Trends Microbiol* 11, 280-285.
- Johansson, J., Mandin, P., Renzoni, A., Chiaruttini, C., Springer, M., and Cossart, P. (2002). An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. *Cell* 110, 551-561.
- Jonsson, I. M., Arvidson, S., Foster, S., and Tarkowski, A. (2004). Sigma factor B and RsbU are required for virulence in *Staphylococcus aureus*-induced arthritis and sepsis. *Infect Immun* 72, 6106-6111.
- Jonsson, K., Signas, C., Muller, H. P., and Lindberg, M. (1991). Two different genes encode fibronectin binding proteins in *Staphylococcus aureus*. The complete nucleotide sequence and characterization of the second gene. *Eur J Biochem* 202, 1041-1048.
- Julio, S. M., Heithoff, D. M., and Mahan, M. J. (2000). *ssrA* (tmRNA) plays a role in *Salmonella enterica* serovar Typhimurium pathogenesis. *J Bacteriol* 182, 1558-1563.
- K**apoor, V., Travadi, J., and Braye, S. (2008). Staphylococcal scalded skin syndrome in an extremely premature neonate: a case report with a brief review of literature. *J Paediatr Child Health* 44, 374-376.
- Karlsson, A., Saravia-Otten, P., Tegmark, K., Morfeldt, E., and Arvidson, S. (2001). Decreased amounts of cell wall-associated protein A and fibronectin-binding proteins in *Staphylococcus aureus* *sarA* mutants due to up-regulation of extracellular proteases. *Infect Immun* 69, 4742-4748.
- Kawabata, S., Morita, T., Iwanaga, S., and Igarashi, H. (1985). Staphylocoagulase-binding region in human prothrombin. *J Biochem* 97, 325-331.
- Kawamoto, H., Morita, T., Shimizu, A., Inada, T., and Aiba, H. (2005). Implication of membrane localization of target mRNA in the action of a small RNA: mechanism of post-transcriptional regulation of glucose transporter in *Escherichia coli*. *Genes Dev* 19, 328-338.
- Kay, E., Humair, B., Denervaud, V., Riedel, K., Spahr, S., Eberl, L., Valverde, C., and Haas, D. (2006). Two GacA-dependent small RNAs modulate the quorum-sensing response in *Pseudomonas aeruginosa*. *J Bacteriol* 188, 6026-6033.
- King, B. F., and Wilkinson, B. J. (1981). Binding of human immunoglobulin G to protein A in encapsulated *Staphylococcus aureus*. *Infect Immun* 33, 666-672.
- Klenk, M., Koczan, D., Guthke, R., Nakata, M., Thiesen, H. J., Podbielski, A., and Kreikemeyer, B. (2005). Global epithelial cell transcriptional responses reveal *Streptococcus pyogenes* Fas regulator activity association with bacterial aggressiveness. *Cell Microbiol* 7, 1237-1250.
- Klocko, A. D., and Wassarman, K. M. (2009). 6S RNA binding to Esigma(70) requires a positively charged surface of sigma(70) region 4.2. *Mol Microbiol* 73, 152-164.

- Koenig, R. L., Ray, J. L., Maleki, S. J., Smeltzer, M. S., and Hurlburt, B. K. (2004). Staphylococcus aureus AgrA binding to the RNAlII-agr regulatory region. *J Bacteriol* 186, 7549-7555.
- Kolb, F. A., Engdahl, H. M., Slagter-Jager, J. G., Ehresmann, B., Ehresmann, C., Westhof, E., Wagner, E. G., and Romby, P. (2000a). Progression of a loop-loop complex to a four-way junction is crucial for the activity of a regulatory antisense RNA. *EMBO J* 19, 5905-5915.
- Kolb, F. A., Malmgren, C., Westhof, E., Ehresmann, C., Ehresmann, B., Wagner, E. G., and Romby, P. (2000b). 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* 6, 311-324.
- Kolb, F. A., Westhof, E., Ehresmann, C., Ehresmann, B., Wagner, E. G., and Romby, P. (2001). Bulged residues promote the progression of a loop-loop interaction to a stable and inhibitory antisense-target RNA complex. *Nucleic Acids Res* 29, 3145-3153.
- Komarova, A. V., Tchufistova, L. S., Dreyfus, M., and Boni, I. V. (2005). AU-rich sequences within 5' untranslated leaders enhance translation and stabilize mRNA in Escherichia coli. *J Bacteriol* 187, 1344-1349.
- Kong, K. F., Vuong, C., and Otto, M. (2006). Staphylococcus quorum sensing in biofilm formation and infection. *Int J Med Microbiol* 296, 133-139.
- Korostelev, A., Ermolenko, D. N., and Noller, H. F. (2008). Structural dynamics of the ribosome. *Curr Opin Chem Biol* 12, 674-683.
- Kreikemeyer, B., Boyle, M. D., Buttaro, B. A., Heinemann, M., and Podbielski, A. (2001). Group A streptococcal growth phase-associated virulence factor regulation by a novel operon (Fas) with homologies to two-component-type regulators requires a small RNA molecule. *Mol Microbiol* 39, 392-406.
- Kristiansen, S. V., Pascual, V., and Lipsky, P. E. (1994). Staphylococcal protein A induces biased production of Ig by VH3-expressing B lymphocytes. *J Immunol* 153, 2974-2982.
- Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* 31, 147-157.
- L**adhani, S. (2001). Recent developments in staphylococcal scalded skin syndrome. *Clin Microbiol Infect* 7, 301-307.
- Lang, S., Livesley, M. A., Lambert, P. A., Littler, W. A., and Elliott, T. S. (2000). Identification of a novel antigen from Staphylococcus epidermidis. *FEMS Immunol Med Microbiol* 29, 213-220.
- Lapouge, K., Schubert, M., Allain, F. H., and Haas, D. (2008). Gac/Rsm signal transduction pathway of gamma-proteobacteria: from RNA recognition to regulation of social behaviour. *Mol Microbiol* 67, 241-253.
- Lebeau, C., Vandenesch, F., Greenland, T., Novick, R. P., and Etienne, J. (1994). Coagulase expression in Staphylococcus aureus is positively and negatively modulated by an agr-dependent mechanism. *J Bacteriol* 176, 5534-5536.
- Lenz, D. H., Miller, M. B., Zhu, J., Kulkarni, R. V., and Bassler, B. L. (2005). CsrA and three redundant small RNAs regulate quorum sensing in Vibrio cholerae. *Mol Microbiol* 58, 1186-1202.
- Leontis, N. B., and Westhof, E. (2003). Analysis of RNA motifs. *Curr Opin Struct Biol* 13, 300-308.

- Li, H. L., Chelladurai, B. S., Zhang, K., and Nicholson, A. W. (1993). Ribonuclease III cleavage of a bacteriophage T7 processing signal. Divalent cation specificity, and specific anion effects. *Nucleic Acids Res* 21, 1919-1925.
- Liang, X., Yu, C., Sun, J., Liu, H., Landwehr, C., Holmes, D., and Ji, Y. (2006). Inactivation of a two-component signal transduction system, SaeRS, eliminates adherence and attenuates virulence of *Staphylococcus aureus*. *Infect Immun* 74, 4655-4665.
- Liang, X., Zheng, L., Landwehr, C., Lunsford, D., Holmes, D., and Ji, Y. (2005). Global regulation of gene expression by ArlRS, a two-component signal transduction regulatory system of *Staphylococcus aureus*. *J Bacteriol* 187, 5486-5492.
- Lina, G., Gillet, Y., Vandenesch, F., Jones, M. E., Floret, D., and Etienne, J. (1997). Toxin involvement in staphylococcal scalded skin syndrome. *Clin Infect Dis* 25, 1369-1373.
- Liu, M. Y., Gui, G., Wei, B., Preston, J. F. r., Oakford, L., Yuksel, U., Giedroc, D. P., and Romeo, T. (1997). The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in *Escherichia coli*. *J Biol Chem* 272, 17502-17510.
- Livny, J., Teonadi, H., Livny, M., and Waldor, M. K. (2008). High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. *PLoS One* 3, e3197.
- Livny, J., and Waldor, M. K. (2007). Identification of small RNAs in diverse bacterial species. *Curr Opin Microbiol* 10, 96-101.
- Long, J. P., Hart, J., Albers, W., and Kapral, F. A. (1992). The production of fatty acid modifying enzyme (FAME) and lipase by various staphylococcal species. *J Med Microbiol* 37, 232-234.
- Loughman, A., Fitzgerald, J. R., Brennan, M. P., Higgins, J., Downer, R., Cox, D., and Foster, T. J. (2005). Roles for fibrinogen, immunoglobulin and complement in platelet activation promoted by *Staphylococcus aureus* clumping factor A. *Mol Microbiol* 57, 804-818.
- Lucchetti-Miganeh, C., Burrowes, E., Baysse, C., and Ermel, G. (2008). The post-transcriptional regulator CsrA plays a central role in the adaptation of bacterial pathogens to different stages of infection in animal hosts. *Microbiology* 154, 16-29.
- Luong, T. T., Newell, S. W., and Lee, C. Y. (2003). Mgr, a novel global regulator in *Staphylococcus aureus*. *J Bacteriol* 185, 3703-3710.
- Lyon, G. J., Wright, J. S., Muir, T. W., and Novick, R. P. (2002). Key determinants of receptor activation in the agr autoinducing peptides of *Staphylococcus aureus*. *Biochemistry* 41, 10095-10104.

**M**akarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., and Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct* 1, 7.

Malmgren, C., Engdahl, H. M., Romby, P., and Wagner, E. G. (1996). An antisense/target RNA duplex or a strong intramolecular RNA structure 5' of a translation initiation signal blocks ribosome binding: the case of plasmid R1. *RNA* 2, 1022-1032.

Mandal, M., and Breaker, R. R. (2004). Gene regulation by riboswitches. *Nat Rev Mol Cell Biol* 5, 451-463.

Mandal, M., Lee, M., Barrick, J. E., Weinberg, Z., Emilsson, G. M., Ruzzo, W. L., and Breaker, R. R. (2004). A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science* 306, 275-279.

- Mandin, P., and Gottesman, S. (2009). A genetic approach for finding small RNAs regulators of genes of interest identifies RybC as regulating the DpiA/DpiB two-component system. *Mol Microbiol* 72, 551-565.
- Mandin, P., Repoila, F., Vergassola, M., Geissmann, T., and Cossart, P. (2007). Identification of new noncoding RNAs in *Listeria monocytogenes* and prediction of mRNA targets. *Nucleic Acids Res* 35, 962-974.
- Manna, A., and Cheung, A. L. (2001). Characterization of sarR, a modulator of sar expression in *Staphylococcus aureus*. *Infect Immun* 69, 885-896.
- Manna, A. C., Bayer, M. G., and Cheung, A. L. (1998). Transcriptional analysis of different promoters in the sar locus in *Staphylococcus aureus*. *J Bacteriol* 180, 3828-3836.
- Manna, A. C., and Cheung, A. L. (2003). sarU, a sarA homolog, is repressed by SarT and regulates virulence genes in *Staphylococcus aureus*. *Infect Immun* 71, 343-353.
- Manna, A. C., and Cheung, A. L. (2006a). Expression of SarX, a negative regulator of agr and exoprotein synthesis, is activated by MgrA in *Staphylococcus aureus*. *J Bacteriol* 188, 4288-4299.
- Manna, A. C., and Cheung, A. L. (2006b). Transcriptional regulation of the agr locus and the identification of DNA binding residues of the global regulatory protein SarR in *Staphylococcus aureus*. *Mol Microbiol* 60, 1289-1301.
- Marchais, A., Naville, M., Bohn, C., Bouloc, P., and Gautheret, D. (2009). Single-pass classification of all noncoding sequences in a bacterial genome using phylogenetic profiles. *Genome Res* 19, 1084-1092.
- Marraffini, L. A., and Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322, 1843-1845.
- Marraffini, L. A., and Sontheimer, E. J. (2009). Invasive DNA, chopped and in the CRISPR. *Structure* 17, 786-788.
- Martianov, I., Ramadass, A., Serra Barros, A., Chow, N., and Akoulitchev, A. (2007). Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* 445, 666-670.
- Martin, P. K., Li, T., Sun, D., Biek, D. P., and Schmid, M. B. (1999). Role in cell permeability of an essential two-component system in *Staphylococcus aureus*. *J Bacteriol* 181, 3666-3673.
- Marvin, M. C., and Engelke, D. R. (2009). RNase P: increased versatility through protein complexity? *RNA Biol* 6, 40-42.
- Marzi, S., Fechter, P., Chevalier, C., Romby, P., and Geissmann, T. (2008). RNA switches regulate initiation of translation in bacteria. *Biol Chem* 389, 585-598.
- Marzi, S., Myasnikov, A. G., Serganov, A., Ehresmann, C., Romby, P., Yusupov, M., and Klaholz, B. P. (2007). Structured mRNAs regulate translation initiation by binding to the platform of the ribosome. *Cell* 130, 1019-1031.
- Masse, E., and Gottesman, S. (2002). A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc Natl Acad Sci U S A* 99, 4620-4625.
- Masse, E., Salvail, H., Desnoyers, G., and Arguin, M. (2007). Small RNAs controlling iron metabolism. *Curr Opin Microbiol* 10, 140-145.
- Masse, E., Vanderpool, C. K., and Gottesman, S. (2005). Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J Bacteriol* 187, 6962-6971.
- Mathy, N., Benard, L., Pellegrini, O., Daou, R., Wen, T., and Condon, C. (2007). 5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell* 129, 681-692.

- Matsunaga, J., Simons, E. L., and Simons, R. W. (1996). RNase III autoregulation: structure and function of rncO, the posttranscriptional "operator". *RNA* 2, 1228-1240.
- Mattick, J. S. (2001). Non-coding RNAs: the architects of eukaryotic complexity. *EMBO Rep* 2, 986-991.
- Mattick, J. S. (2009). The genetic signatures of noncoding RNAs. *PLoS Genet* 5, e1000459.
- Mattick, J. S., and Makunin, I. V. (2005). Small regulatory RNAs in mammals. *Hum Mol Genet* 14 *Spec No 1*, R121-32.
- Mattick, J. S., and Makunin, I. V. (2006). Non-coding RNA. *Hum Mol Genet* 15 *Spec No 1*, R17-29.
- Mazmanian, S. K., Liu, G., Ton-That, H., and Schneewind, O. (1999). Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. *Science* 285, 760-763.
- McAleese, F. M., Walsh, E. J., Sieprawska, M., Potempa, J., and Foster, T. J. (2001). Loss of clumping factor B fibrinogen binding activity by Staphylococcus aureus involves cessation of transcription, shedding and cleavage by metalloprotease. *J Biol Chem* 276, 29969-29978.
- McGavin, M. J., Zahradka, C., Rice, K., and Scott, J. E. (1997). Modification of the Staphylococcus aureus fibronectin binding phenotype by V8 protease. *Infect Immun* 65, 2621-2628.
- McNamara, P. J., Milligan-Monroe, K. C., Khalili, S., and Proctor, R. A. (2000). Identification, cloning, and initial characterization of rot, a locus encoding a regulator of virulence factor expression in Staphylococcus aureus. *J Bacteriol* 182, 3197-3203.
- Meier, U. T. (2005). The many facets of H/ACA ribonucleoproteins. *Chromosoma* 114, 1-14.
- Menzies, B. E. (2003). The role of fibronectin binding proteins in the pathogenesis of Staphylococcus aureus infections. *Curr Opin Infect Dis* 16, 225-229.
- Mercer, T. R., Dinger, M. E., and Mattick, J. S. (2009). Long non-coding RNAs: insights into functions. *Nat Rev Genet* 10, 155-159.
- Merino, E., Jensen, R. A., and Yanofsky, C. (2008). Evolution of bacterial trp operons and their regulation. *Curr Opin Microbiol* 11, 78-86.
- Meyer, F., Girardot, R., Piemont, Y., Prevost, G., and Colin, D. A. (2009). Analysis of the specificity of Pantone-Valentine leucocidin and gamma-hemolysin F component binding. *Infect Immun* 77, 266-273.
- Mironov, A. S., Gusarov, I., Rafikov, R., Lopez, L. E., Shatalin, K., Kreneva, R. A., Perumov, D. A., and Nudler, E. (2002). Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* 111, 747-756.
- Mitarai, N., Benjamin, J. A., Krishna, S., Semsey, S., Csiszovszki, Z., Masse, E., and Sneppen, K. (2009). Dynamic features of gene expression control by small regulatory RNAs. *Proc Natl Acad Sci U S A*
- Mizuno, T., Chou, M. Y., and Inouye, M. (1984). A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc Natl Acad Sci U S A* 81, 1966-1970.
- Mockler, T. C., Chan, S., Sundaresan, A., Chen, H., Jacobsen, S. E., and Ecker, J. R. (2005). Applications of DNA tiling arrays for whole-genome analysis. *Genomics* 85, 1-15.

- Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* *155*, 733-740.
- Moll, I., Afonyushkin, T., Vytvytska, O., Kaberdin, V. R., and Blasi, U. (2003). Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA* *9*, 1308-1314.
- Moller, T., Franch, T., Udesen, C., Gerdes, K., and Valentin-Hansen, P. (2002). Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. *Genes Dev* *16*, 1696-1706.
- Montange, R. K., and Batey, R. T. (2006). Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature* *441*, 1172-1175.
- Montange, R. K., and Batey, R. T. (2008). Riboswitches: emerging themes in RNA structure and function. *Annu Rev Biophys* *37*, 117-133.
- Morfeldt, E., Taylor, D., von Gabain, A., and Arvidson, S. (1995). Activation of alpha-toxin translation in *Staphylococcus aureus* by the trans-encoded antisense RNA, RNAlII. *EMBO J* *14*, 4569-4577.
- Morita, T., Maki, K., and Aiba, H. (2005). RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev* *19*, 2176-2186.
- Morita, T., Mochizuki, Y., and Aiba, H. (2006). Translational repression is sufficient for gene silencing by bacterial small noncoding RNAs in the absence of mRNA destruction. *Proc Natl Acad Sci U S A* *103*, 4858-4863.
- Mulcahy, H., O'Callaghan, J., O'Grady, E. P., Adams, C., and O'Gara, F. (2006). The posttranscriptional regulator RsmA plays a role in the interaction between *Pseudomonas aeruginosa* and human airway epithelial cells by positively regulating the type III secretion system. *Infect Immun* *74*, 3012-3015.
- N**ahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L., and Breaker, R. R. (2002). Genetic control by a metabolite binding mRNA. *Chem Biol* *9*, 1043.
- Nair, S. P., Bischoff, M., Senn, M. M., and Berger-Bachi, B. (2003). The sigma B regulon influences internalization of *Staphylococcus aureus* by osteoblasts. *Infect Immun* *71*, 4167-4170.
- Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into *Petunia* Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell* *2*, 279-289.
- Narberhaus, F., Waldminghaus, T., and Chowdhury, S. (2006). RNA thermometers. *FEMS Microbiol Rev* *30*, 3-16.
- Neher, S. B., Bradshaw, N., Floor, S. N., Gross, J. D., and Walter, P. (2008). SRP RNA controls a conformational switch regulating the SRP-SRP receptor interaction. *Nat Struct Mol Biol* *15*, 916-923.
- Nguyen, T., Ghebrehiwet, B., and Peerschke, E. I. (2000). *Staphylococcus aureus* protein A recognizes platelet gC1qR/p33: a novel mechanism for staphylococcal interactions with platelets. *Infect Immun* *68*, 2061-2068.
- Ni Eidhin, D., Perkins, S., Francois, P., Vaudaux, P., Hook, M., and Foster, T. J. (1998). Clumping factor B (ClfB), a new surface-located fibrinogen-binding adhesin of *Staphylococcus aureus*. *Mol Microbiol* *30*, 245-257.
- Nicholson, A. W. (1999). Function, mechanism and regulation of bacterial ribonucleases. *FEMS Microbiol Rev* *23*, 371-390.

- Nikulin, A., Stolboushkina, E., Perederina, A., Vassilieva, I., Blaesi, U., Moll, I., Kachalova, G., Yokoyama, S., Vassylyev, D., Garber, M., and Nikonov, S. (2005). Structure of *Pseudomonas aeruginosa* Hfq protein. *Acta Crystallogr D Biol Crystallogr* *61*, 141-146.
- Noller, H. F., Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Lancaster, L., Dallas, A., Fredrick, K., Earnest, T. N., and Cate, J. H. (2001). Structure of the ribosome at 5.5 Å resolution and its interactions with functional ligands. *Cold Spring Harb Symp Quant Biol* *66*, 57-66.
- Nottrott, S., Simard, M. J., and Richter, J. D. (2006). Human let-7a miRNA blocks protein production on actively translating polyribosomes. *Nat Struct Mol Biol* *13*, 1108-1114.
- Novick, R. P. (2003a). Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Mol Microbiol* *48*, 1429-1449.
- Novick, R. P. (2003b). Mobile genetic elements and bacterial toxinoses: the superantigen-encoding pathogenicity islands of *Staphylococcus aureus*. *Plasmid* *49*, 93-105.
- Novick, R. P., and Geisinger, E. (2008). Quorum sensing in staphylococci. *Annu Rev Genet* *42*, 541-564.
- Novick, R. P., and Jiang, D. (2003). The staphylococcal saeRS system coordinates environmental signals with agr quorum sensing. *Microbiology* *149*, 2709-2717.
- Novick, R. P., Ross, H. F., Projan, S. J., Kornblum, J., Kreiswirth, B., and Moghazeh, S. (1993). Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J* *12*, 3967-3975.
- Nudler, E., and Mironov, A. S. (2004). The riboswitch control of bacterial metabolism. *Trends Biochem Sci* *29*, 11-17.
- O**htani, K., Bhowmik, S. K., Hayashi, H., and Shimizu, T. (2002). Identification of a novel locus that regulates expression of toxin genes in *Clostridium perfringens*. *FEMS Microbiol Lett* *209*, 113-118.
- Opdyke, J. A., Kang, J. G., and Storz, G. (2004). GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J Bacteriol* *186*, 6698-6705.
- Oscarsson, J., Harlos, C., and Arvidson, S. (2005). Regulatory role of proteins binding to the spa (protein A) and sarS (staphylococcal accessory regulator) promoter regions in *Staphylococcus aureus* NTCC 8325-4. *Int J Med Microbiol* *295*, 253-266.
- Oscarsson, J., Kanth, A., Tegmark-Wisell, K., and Arvidson, S. (2006). 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* *188*, 8526-8533.
- P**alma, M., Shannon, O., Quezada, H. C., Berg, A., and Flock, J. I. (2001). Extracellular fibrinogen-binding protein, Efb, from *Staphylococcus aureus* blocks platelet aggregation due to its binding to the alpha-chain. *J Biol Chem* *276*, 31691-31697.
- Palmer, M. (1998). Staphylococcal alpha toxin. *Symp Ser Soc Appl Microbiol* *27*, 125S-126S.
- Pang, K. C., Frith, M. C., and Mattick, J. S. (2006). Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends Genet* *22*, 1-5.
- Papenfors, K., Pfeiffer, V., Lucchini, S., Sonawane, A., Hinton, J. C., and Vogel, J. (2008). Systematic deletion of *Salmonella* small RNA genes identifies CyaR, a conserved CRP-dependent riboregulator of OmpX synthesis. *Mol Microbiol* *68*, 890-906.

- Papenfort, K., and Vogel, J. (2009). Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level. *Res Microbiol* 160, 278-287.
- Patel, A. H., Kornblum, J., Kreiswirth, B., Novick, R., and Foster, T. J. (1992). Regulation of the protein A-encoding gene in *Staphylococcus aureus*. *Gene* 114, 25-34.
- Pfeiffer, V., Papenfort, K., Lucchini, S., Hinton, J. C., and Vogel, J. (2009). Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* 16, 840-846.
- Pfeiffer, V., Sittka, A., Tomer, R., Tedin, K., Brinkmann, V., and Vogel, J. (2007). A small non-coding RNA of the invasion gene island (SPI-1) represses outer membrane protein synthesis from the *Salmonella* core genome. *Mol Microbiol* 66, 1174-1191.
- Pichon, C., and Felden, B. (2005). Small RNA genes expressed from *Staphylococcus aureus* genomic and pathogenicity islands with specific expression among pathogenic strains. *Proc Natl Acad Sci U S A* 102, 14249-14254.
- Pillai, R. S., Bhattacharyya, S. N., Artus, C. G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science* 309, 1573-1576.
- Ponting, C. P., Oliver, P. L., and Reik, W. (2009). Evolution and functions of long noncoding RNAs. *Cell* 136, 629-641.
- Prevost, G., Cribier, B., Couppie, P., Petiau, P., Supersac, G., Finck-Barbancon, V., Monteil, H., and Piemont, Y. (1995). Panton-Valentine leucocidin and gamma-hemolysin from *Staphylococcus aureus* ATCC 49775 are encoded by distinct genetic loci and have different biological activities. *Infect Immun* 63, 4121-4129.
- Prevost, K., Salvail, H., Desnoyers, G., Jacques, J. F., Phaneuf, E., and Masse, E. (2007). The small RNA RyhB activates the translation of shiA mRNA encoding a permease of shikimate, a compound involved in siderophore synthesis. *Mol Microbiol* 64, 1260-1273.
- Proft, T., and Fraser, J. D. (2003). Bacterial superantigens. *Clin Exp Immunol* 133, 299-306.
- Prokesova, L., Potuznikova, B., Potempa, J., Zikan, J., Radl, J., Hachova, L., Baran, K., Porwit-Bohr, Z., and John, C. (1992). Cleavage of human immunoglobulins by serine proteinase from *Staphylococcus aureus*. *Immunol Lett* 31, 259-265.
- Q**iu, R., Pei, W., Zhang, L., Lin, J., and Ji, G. (2005). Identification of the putative staphylococcal AgrB catalytic residues involving the proteolytic cleavage of AgrD to generate autoinducing peptide. *J Biol Chem* 280, 16695-16704.
- Queck, S. Y., Jameson-Lee, M., Villaruz, A. E., Bach, T. H., Khan, B. A., Sturdevant, D. E., Ricklefs, S. M., Li, M., and Otto, M. (2008). RNAIII-independent target gene control by the agr quorum-sensing system: insight into the evolution of virulence regulation in *Staphylococcus aureus*. *Mol Cell* 32, 150-158.
- R**achid, S., Ohlsen, K., Wallner, U., Hacker, J., Hecker, M., and Ziebuhr, W. (2000). Alternative transcription factor sigma(B) is involved in regulation of biofilm expression in a *Staphylococcus aureus* mucosal isolate. *J Bacteriol* 182, 6824-6826.
- Rasmussen, A. A., Eriksen, M., Gilany, K., Udesen, C., Franch, T., Petersen, C., and Valentin-Hansen, P. (2005). Regulation of ompA mRNA stability: the role of a small regulatory RNA in growth phase-dependent control. *Mol Microbiol* 58, 1421-1429.

- Rasmussen, A. A., Johansen, J., Nielsen, J. S., Overgaard, M., Kallipolitis, B., and Valentin-Hansen, P. (2009). A conserved small RNA promotes silencing of the outer membrane protein YbfM. *Mol Microbiol*
- Rechtin, T. M., Gillaspay, A. F., Schumacher, M. A., Brennan, R. G., Smeltzer, M. S., and Hurlburt, B. K. (1999). Characterization of the SarA virulence gene regulator of *Staphylococcus aureus*. *Mol Microbiol* 33, 307-316.
- Recsei, P., Kreiswirth, B., O'Reilly, M., Schlievert, P., Gruss, A., and Novick, R. P. (1986). Regulation of exoprotein gene expression in *Staphylococcus aureus* by agar. *Mol Gen Genet* 202, 58-61.
- Reed, S. B., Wesson, C. A., Liou, L. E., Trumble, W. R., Schlievert, P. M., Bohach, G. A., and Bayles, K. W. (2001). Molecular characterization of a novel *Staphylococcus aureus* serine protease operon. *Infect Immun* 69, 1521-1527.
- Regassa, L. B., Couch, J. L., and Betley, M. J. (1991). Steady-state staphylococcal enterotoxin type C mRNA is affected by a product of the accessory gene regulator (*agr*) and by glucose. *Infect Immun* 59, 955-962.
- Reichow, S. L., Hamma, T., Ferre-D'Amare, A. R., and Varani, G. (2007). The structure and function of small nucleolar ribonucleoproteins. *Nucleic Acids Res* 35, 1452-1464.
- Rentmeister, A., Mayer, G., Kuhn, N., and Famulok, M. (2007). Conformational changes in the expression domain of the *Escherichia coli* thiM riboswitch. *Nucleic Acids Res* 35, 3713-3722.
- Repoila, F., Majdalani, N., and Gottesman, S. (2003). Small non-coding RNAs, coordinators of adaptation processes in *Escherichia coli*: the RpoS paradigm. *Mol Microbiol* 48, 855-861.
- Rice, K., Peralta, R., Bast, D., de Azavedo, J., and McGavin, M. J. (2001). Description of staphylococcus serine protease (*ssp*) operon in *Staphylococcus aureus* and nonpolar inactivation of *sspA*-encoded serine protease. *Infect Immun* 69, 159-169.
- Rich, M. (2005). Staphylococci in animals: prevalence, identification and antimicrobial susceptibility, with an emphasis on methicillin-resistant *Staphylococcus aureus*. *Br J Biomed Sci* 62, 98-105.
- Rieneck, K., Renneberg, J., Diamant, M., Gutschik, E., and Bendtzen, K. (1997). Molecular cloning and expression of a novel *Staphylococcus aureus* antigen. *Biochim Biophys Acta* 1350, 128-132.
- Rivas, E., and Eddy, S. R. (2001). Noncoding RNA gene detection using comparative sequence analysis. *BMC Bioinformatics* 2, 8.
- Roberts, C., Anderson, K. L., Murphy, E., Projan, S. J., Mounts, W., Hurlburt, B., Smeltzer, M., Overbeek, R., Disz, T., and Dunman, P. M. (2006). Characterizing the effect of the *Staphylococcus aureus* virulence factor regulator, SarA, on log-phase mRNA half-lives. *J Bacteriol* 188, 2593-2603.
- Romby, P., Vandenesch, F., and Wagner, E. G. (2006). The role of RNAs in the regulation of virulence-gene expression. *Curr Opin Microbiol* 9, 229-236.
- Romeo, T., Gong, M., Liu, M. Y., and Brun-Zinkernagel, A. M. (1993). Identification and molecular characterization of *csrA*, a pleiotropic gene from *Escherichia coli* that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. *J Bacteriol* 175, 4744-4755.
- Said-Salim, B., Dunman, P. M., McAleese, F. M., Macapagal, D., Murphy, E., McNamara, P. J., Arvidson, S., Foster, T. J., Projan, S. J., and Kreiswirth, B. N.

- (2003). Global regulation of *Staphylococcus aureus* genes by Rot. *J Bacteriol* *185*, 610-619.
- Said, N., Rieder, R., Hurwitz, R., Deckert, J., Urlaub, H., and Vogel, J. (2009). In vivo expression and purification of aptamer-tagged small RNA regulators. *Nucleic Acids Res*
- Saito, S., Kakeshita, H., and Nakamura, K. (2009). Novel small RNA-encoding genes in the intergenic regions of *Bacillus subtilis*. *Gene* *428*, 2-8.
- Santos, J. M., Drider, D., Marujo, P. E., Lopez, P., and Arraiano, C. M. (1997). Determinant role of *E. coli* RNase III in the decay of both specific and heterologous mRNAs. *FEMS Microbiol Lett* *157*, 31-38.
- Saravia-Otten, P., Muller, H. P., and Arvidson, S. (1997). Transcription of *Staphylococcus aureus* fibronectin binding protein genes is negatively regulated by *agr* and an *agr*-independent mechanism. *J Bacteriol* *179*, 5259-5263.
- Sauter, C., Basquin, J., and Suck, D. (2003). Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from *Escherichia coli*. *Nucleic Acids Res* *31*, 4091-4098.
- Schmidt, K. A., Manna, A. C., and Cheung, A. L. (2003). SarT influences sarS expression in *Staphylococcus aureus*. *Infect Immun* *71*, 5139-5148.
- Schmidt, K. A., Manna, A. C., Gill, S., and Cheung, A. L. (2001). SarT, a repressor of alpha-hemolysin in *Staphylococcus aureus*. *Infect Immun* *69*, 4749-4758.
- Sendi, P., and Proctor, R. A. (2009). *Staphylococcus aureus* as an intracellular pathogen: the role of small colony variants. *Trends Microbiol* *17*, 54-58.
- Serganov, A., Huang, L., and Patel, D. J. (2008). Structural insights into amino acid binding and gene control by a lysine riboswitch. *Nature* *455*, 1263-1267.
- Serganov, A., and Patel, D. J. (2007). Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat Rev Genet* *8*, 776-790.
- Serganov, A., Yuan, Y. R., Pikovskaya, O., Polonskaia, A., Malinina, L., Phan, A. T., Hobartner, C., Micura, R., Breaker, R. R., and Patel, D. J. (2004). Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem Biol* *11*, 1729-1741.
- Sharma, C. M., Darfeuille, F., Plantinga, T. H., and 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.
- Sharp, P. A. (2009). The centrality of RNA. *Cell* *136*, 577-580.
- Shaw, L., Golonka, E., Potempa, J., and Foster, S. J. (2004). The role and regulation of the extracellular proteases of *Staphylococcus aureus*. *Microbiology* *150*, 217-228.
- Shearwin, K. E., Callen, B. P., and Egan, J. B. (2005). Transcriptional interference--a crash course. *Trends Genet* *21*, 339-345.
- Sheehan, B. J., Foster, T. J., Dorman, C. J., Park, S., and Stewart, G. S. (1992). Osmotic and growth-phase dependent regulation of the eta gene of *Staphylococcus aureus*: a role for DNA supercoiling. *Mol Gen Genet* *232*, 49-57.
- Silvaggi, J. M., Perkins, J. B., and Losick, R. (2005). Small untranslated RNA antitoxin in *Bacillus subtilis*. *J Bacteriol* *187*, 6641-6650.
- Simons, R. W., and Kleckner, N. (1983). Translational control of IS10 transposition. *Cell* *34*, 683-691.
- Singh, V. K., Schmidt, J. L., Jayaswal, R. K., and Wilkinson, B. J. (2003). Impact of sigB mutation on *Staphylococcus aureus* oxacillin and vancomycin resistance varies with parental background and method of assessment. *Int J Antimicrob Agents* *21*, 256-261.

- Sittka, A., Lucchini, S., Papenfort, K., Sharma, C. M., Rolle, K., Binnewies, T. T., Hinton, J. C., and Vogel, J. (2008). Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet* 4, e1000163.
- Smeltzer, M. S., Hart, M. E., and Iandolo, J. J. (1993). Phenotypic characterization of xpr, a global regulator of extracellular virulence factors in *Staphylococcus aureus*. *Infect Immun* 61, 919-925.
- Somerville, G. A., and Proctor, R. A. (2009). At the crossroads of bacterial metabolism and virulence factor synthesis in *Staphylococci*. *Microbiol Mol Biol Rev* 73, 233-248.
- Song, L., Hobaugh, M. R., Shustak, C., Cheley, S., Bayley, H., and Gouaux, J. E. (1996). Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science* 274, 1859-1866.
- Song, T., Mika, F., Lindmark, B., Liu, Z., Schild, S., Bishop, A., Zhu, J., Camilli, A., Johansson, J., Vogel, J., and Wai, S. N. (2008). A new *Vibrio cholerae* sRNA modulates colonization and affects release of outer membrane vesicles. *Mol Microbiol* 70, 100-111.
- Soper, T. J., and Woodson, S. A. (2008). The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA* 14, 1907-1917.
- Sorek, R., Kunin, V., and Hugenholtz, P. (2008). CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat Rev Microbiol* 6, 181-186.
- Stec-Niemczyk, J., Pustelny, K., Kisielewska, M., Bista, M., Boulware, K. T., Stennicke, H. R., Thogersen, I. B., Daugherty, P. S., Enghild, J. J., Baczynski, K., Popowicz, G. M., Dubin, A., Potempa, J., and Dubin, G. (2009). Structural and functional characterization of SplA, an exclusively specific protease of *Staphylococcus aureus*. *Biochem J* 419, 555-564.
- Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000). Two-component signal transduction. *Annu Rev Biochem* 69, 183-215.
- Storz, G., Altuvia, S., and Wassarman, K. M. (2005). An abundance of RNA regulators. *Annu Rev Biochem* 74, 199-217.
- Storz, G., Opdyke, J. A., and Zhang, A. (2004). Controlling mRNA stability and translation with small, noncoding RNAs. *Curr Opin Microbiol* 7, 140-144.
- Stougaard, P., Molin, S., and Nordstrom, K. (1981). RNAs involved in copy-number control and incompatibility of plasmid R1. *Proc Natl Acad Sci U S A* 78, 6008-6012.
- Sudarsan, N., Hammond, M. C., Block, K. F., Welz, R., Barrick, J. E., Roth, A., and Breaker, R. R. (2006). Tandem riboswitch architectures exhibit complex gene control functions. *Science* 314, 300-304.
- Sudarsan, N., Lee, E. R., Weinberg, Z., Moy, R. H., Kim, J. N., Link, K. H., and Breaker, R. R. (2008). Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science* 321, 411-413.
- Sun, X., Zhulin, I., and Wartell, R. M. (2002). Predicted structure and phyletic distribution of the RNA-binding protein Hfq. *Nucleic Acids Res* 30, 3662-3671.
- Svenningsen, S. L., Tu, K. C., and Bassler, B. L. (2009). Gene dosage compensation calibrates four regulatory RNAs to control *Vibrio cholerae* quorum sensing. *EMBO J* 28, 429-439.
- Svenningsen, S. L., Waters, C. M., and Bassler, B. L. (2008). A negative feedback loop involving small RNAs accelerates *Vibrio cholerae*'s transition out of quorum-sensing mode. *Genes Dev* 22, 226-238.

- Swem, L. R., Swem, D. L., O'Loughlin, C. T., Gatmaitan, R., Zhao, B., Ulrich, S. M., and Bassler, B. L. (2009). A quorum-sensing antagonist targets both membrane-bound and cytoplasmic receptors and controls bacterial pathogenicity. *Mol Cell* *35*, 143-153.
- T**amber, S., and Cheung, A. L. (2009). SarZ promotes the expression of virulence factors and represses biofilm formation by modulating SarA and agr in *Staphylococcus aureus*. *Infect Immun* *77*, 419-428.
- Tegmark, K., Karlsson, A., and Arvidson, S. (2000). Identification and characterization of SarH1, a new global regulator of virulence gene expression in *Staphylococcus aureus*. *Mol Microbiol* *37*, 398-409.
- Thomas, B., Pugalenti, A., and Chilvers, M. (2009a). Pleuropulmonary complications of PVL-positive *Staphylococcus aureus* infection in children. *Acta Paediatr*
- Thomas, D., Dauwalder, O., Brun, V., Badiou, C., Ferry, T., Etienne, J., Vandenesch, F., and Lina, G. (2009b). *Staphylococcus aureus* superantigens elicit redundant and extensive human Vbeta patterns. *Infect Immun* *77*, 2043-2050.
- Thomas, D. Y., Jarraud, S., Lemercier, B., Cozon, G., Echasserieau, K., Etienne, J., Gougeon, M. L., Lina, G., and Vandenesch, F. (2006). Staphylococcal enterotoxin-like toxins U2 and V, two new staphylococcal superantigens arising from recombination within the enterotoxin gene cluster. *Infect Immun* *74*, 4724-4734.
- Throup, J. P., Zappacosta, F., Lunsford, R. D., Annan, R. S., Carr, S. A., Lonsdale, J. T., Bryant, A. P., McDevitt, D., Rosenberg, M., and Burnham, M. K. (2001). The srhSR gene pair from *Staphylococcus aureus*: genomic and proteomic approaches to the identification and characterization of gene function. *Biochemistry* *40*, 10392-10401.
- Tinsley, R. A., Furchak, J. R., and Walter, N. G. (2007). Trans-acting glmS catalytic riboswitch: locked and loaded. *RNA* *13*, 468-477.
- 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, M. A., Soubigou, G., Regnault, B., Coppee, J. Y., Lecuit, M., Johansson, J., and Cossart, P. (2009). The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* *459*, 950-956.
- Toledo-Arana, A., Repoila, F., and Cossart, P. (2007). Small noncoding RNAs controlling pathogenesis. *Curr Opin Microbiol* *10*, 182-188.
- Toor, N., Rajashankar, K., Keating, K. S., and Pyle, A. M. (2008). Structural basis for exon recognition by a group II intron. *Nat Struct Mol Biol* *15*, 1221-1222.
- Torres-Larios, A., Dock-Bregeon, A. C., Romby, P., Rees, B., Sankaranarayanan, R., Caillet, J., Springer, M., Ehresmann, C., Ehresmann, B., and Moras, D. (2002). Structural basis of translational control by *Escherichia coli* threonyl tRNA synthetase. *Nat Struct Biol* *9*, 343-347.
- Trabucchi, M., Briata, P., Garcia-Mayoral, M., Haase, A. D., Filipowicz, W., Ramos, A., Gherzi, R., and Rosenfeld, M. G. (2009). The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature* *459*, 1010-1014.
- Tremaine, M. T., Brockman, D. K., and Betley, M. J. (1993). Staphylococcal enterotoxin A gene (sea) expression is not affected by the accessory gene regulator (agr). *Infect Immun* *61*, 356-359.
- Trotochaud, A. E., and Wassarman, K. M. (2004). 6S RNA function enhances long-term cell survival. *J Bacteriol* *186*, 4978-4985.

- Trotochaud, A. E., and Wassarman, K. M. (2005). A highly conserved 6S RNA structure is required for regulation of transcription. *Nat Struct Mol Biol* *12*, 313-319.
- Udekwi, K. I., Darfeuille, F., Vogel, J., Reimegard, J., Holmqvist, E., and Wagner, E. G. (2005). Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes Dev* *19*, 2355-2366.
- U**mbach, J. L., and Cullen, B. R. (2009). The role of RNAi and microRNAs in animal virus replication and antiviral immunity. *Genes Dev* *23*, 1151-1164.
- Urban, J. H., and Vogel, J. (2007). Translational control and target recognition by *Escherichia coli* small RNAs in vivo. *Nucleic Acids Res* *35*, 1018-1037.
- V**alentin-Hansen, P., Eriksen, M., and Udesen, C. (2004). The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Mol Microbiol* *51*, 1525-1533.
- van Schaik, W., and Abee, T. (2005). The role of sigmaB in the stress response of Gram-positive bacteria -- targets for food preservation and safety. *Curr Opin Biotechnol* *16*, 218-224.
- Vandenesch, F., Kornblum, J., and Novick, R. P. (1991). A temporal signal, independent of agr, is required for hla but not spa transcription in *Staphylococcus aureus*. *J Bacteriol* *173*, 6313-6320.
- Vasudevan, S., and Steitz, J. A. (2007). AU-rich-element-mediated upregulation of translation by FXR1 and Argonaute 2. *Cell* *128*, 1105-1118.
- Vecerek, B., Moll, I., and Blasi, U. (2007). Control of Fur synthesis by the non-coding RNA RyhB and iron-responsive decoding. *EMBO J* *26*, 965-975.
- Vecerek, B., Rajkowitsch, L., Sonnleitner, E., Schroeder, R., and Blasi, U. (2008). The C-terminal domain of *Escherichia coli* Hfq is required for regulation. *Nucleic Acids Res* *36*, 133-143.
- Vogel, J. (2009). A rough guide to the non-coding RNA world of *Salmonella*. *Mol Microbiol* *71*, 1-11.
- Vogel, J., Argaman, L., Wagner, E. G., and Altuvia, S. (2004). The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Curr Biol* *14*, 2271-2276.
- Vogel, J., and Papenfort, K. (2006). Small non-coding RNAs and the bacterial outer membrane. *Curr Opin Microbiol* *9*, 605-611.
- Vogel, J., and Sharma, C. M. (2005). How to find small non-coding RNAs in bacteria. *Biol Chem* *386*, 1219-1238.
- Vogel, J., and Wagner, E. G. (2007). Target identification of small noncoding RNAs in bacteria. *Curr Opin Microbiol* *10*, 262-270.
- Vulliamy, T., Marrone, A., Goldman, F., Dearlove, A., Bessler, M., Mason, P. J., and Dokal, I. (2001). The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. *Nature* *413*, 432-435.
- W**adler, C. S., and Vanderpool, C. K. (2007). 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* *104*, 20454-20459.
- Wagner, E. G. (2009). Kill the messenger: bacterial antisense RNA promotes mRNA decay. *Nat Struct Mol Biol* *16*, 804-806.
- Wagner, E. G., Altuvia, S., and Romby, P. (2002). Antisense RNAs in bacteria and their genetic elements. *Adv Genet* *46*, 361-398.

- Wagner, E. G., and Brantl, S. (1998). Kissing and RNA stability in antisense control of plasmid replication. *Trends Biochem Sci* 23, 451-454.
- Wagner, E. G., and Simons, R. W. (1994). Antisense RNA control in bacteria, phages, and plasmids. *Annu Rev Microbiol* 48, 713-742.
- Wahl, M. C., Will, C. L., and Luhrmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. *Cell* 136, 701-718.
- Waldminghaus, T., Heidrich, N., Brantl, S., and Narberhaus, F. (2007). FourU: a novel type of RNA thermometer in *Salmonella*. *Mol Microbiol* 65, 413-424.
- Walev, I., Martin, E., Jonas, D., Mohamadzadeh, M., Muller-Klieser, W., Kunz, L., and Bhakdi, S. (1993). Staphylococcal alpha-toxin kills human keratinocytes by permeabilizing the plasma membrane for monovalent ions. *Infect Immun* 61, 4972-4979.
- Walev, I., Palmer, M., Martin, E., Jonas, D., Weller, U., Hohn-Bentz, H., Husmann, M., and Bhakdi, S. (1994). Recovery of human fibroblasts from attack by the pore-forming alpha-toxin of *Staphylococcus aureus*. *Microb Pathog* 17, 187-201.
- Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M. G., Glass, C. K., and Kurokawa, R. (2008). Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* 454, 126-130.
- Wassarman, K. M. (2007). 6S RNA: a regulator of transcription. *Mol Microbiol* 65, 1425-1431.
- Wassarman, K. M., Repoila, F., Rosenow, C., Storz, G., and Gottesman, S. (2001). Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev* 15, 1637-1651.
- Wassarman, K. M., and Saecker, R. M. (2006). Synthesis-mediated release of a small RNA inhibitor of RNA polymerase. *Science* 314, 1601-1603.
- Wassarman, K. M., and Storz, G. (2000). 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* 101, 613-623.
- Wassarman, K. M., Zhang, A., and Storz, G. (1999). Small RNAs in *Escherichia coli*. *Trends Microbiol* 7, 37-45.
- Waters, L. S., and Storz, G. (2009). Regulatory RNAs in bacteria. *Cell* 136, 615-628.
- Weilbacher, T., Suzuki, K., Dubey, A. K., Wang, X., Gudapaty, S., Morozov, I., Baker, C. S., Georgellis, D., Babitzke, P., and Romeo, T. (2003). A novel sRNA component of the carbon storage regulatory system of *Escherichia coli*. *Mol Microbiol* 48, 657-670.
- Wertheim, H. F., Melles, D. C., Vos, M. C., van Leeuwen, W., van Belkum, A., Verbrugh, H. A., and Nouwen, J. L. (2005). The role of nasal carriage in *Staphylococcus aureus* infections. *Lancet Infect Dis* 5, 751-762.
- Wiedenheft, B., Zhou, K., Jinek, M., Coyle, S. M., Ma, W., and Doudna, J. A. (2009). Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure* 17, 904-912.
- Wilderman, P. J., Sowa, N. A., FitzGerald, D. J., FitzGerald, P. C., Gottesman, S., Ochsner, U. A., and Vasil, M. L. (2004). Identification of tandem duplicate regulatory small RNAs in *Pseudomonas aeruginosa* involved in iron homeostasis. *Proc Natl Acad Sci U S A* 101, 9792-9797.
- Wilusz, J. E., Sunwoo, H., and Spector, D. L. (2009). Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev* 23, 1494-1504.
- Winkler, W. C., and Breaker, R. R. (2005). Regulation of bacterial gene expression by riboswitches. *Annu Rev Microbiol* 59, 487-517.

- Winkler, W. C., Cohen-Chalamish, S., and Breaker, R. R. (2002). An mRNA structure that controls gene expression by binding FMN. *Proc Natl Acad Sci U S A* 99, 15908-15913.
- Wutz, A., and Gribnau, J. (2007). X inactivation Xplained. *Curr Opin Genet Dev* 17, 387-393.
- Y**ang, Z., Zhu, Q., Luo, K., and Zhou, Q. (2001). The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* 414, 317-322.
- Yarwood, J. M., McCormick, J. K., and Schlievert, P. M. (2001). Identification of a novel two-component regulatory system that acts in global regulation of virulence factors of *Staphylococcus aureus*. *J Bacteriol* 183, 1113-1123.
- Yazaki, J., Gregory, B. D., and Ecker, J. R. (2007). Mapping the genome landscape using tiling array technology. *Curr Opin Plant Biol* 10, 534-542.
- Yin, R. L., Li, C., Yang, Z. T., Zhang, Y. J., Bai, W. L., Li, X., Yin, R. H., Liu, H., Liu, S., Yang, Q., Cao, Y. G., and Zhang, N. S. (2009). Construction and immunogenicity of a DNA vaccine containing clumping factor A of *Staphylococcus aureus* and bovine IL18. *Vet Immunol Immunopathol*.
- Z**hang, K., Mosch, K., Fischle, W., and Grewal, S. I. (2008). Roles of the Ctr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nat Struct Mol Biol* 15, 381-388.
- Zhang, S., and Stewart, G. C. (2000). Characterization of the promoter elements for the staphylococcal enterotoxin D gene. *J Bacteriol* 182, 2321-2325.
- Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J., and Lee, J. T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 322, 750-756.

---

# **Annexes**

---

## **Annexe 1: Rôle de l'ARN dans la régulation de l'initiation de la traduction chez les bactéries.**

Chez les bactéries, l'initiation de la traduction est l'étape limitante de la synthèse protéique. Pour cette raison, cette étape est régulée finement par de nombreux facteurs parmi lesquels des molécules d'ARN. Cette revue fait le point sur le rôle de l'ARN au cours de ces différentes régulations. La diversité des mécanismes identifiés est illustrée au travers d'exemples caractéristiques.

*Signalement bibliographique ajouté par :*

**l'UNIVERSITÉ DE STRASBOURG**  
**Service Commun de la Documentation**

**RNA switches regulate initiation of translation in bacteria**

Stefano MARZI, Pierre FECHTER, Clément CHEVALIER, Pascale ROMBY and  
Thomas GEISSMANN

**Biological Chemistry, 2008, vol. 389, pages 585-598**

**Annexe I : p. 289-302**

La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.

Les utilisateurs de l'UdS peuvent consulter cette publication sur le site de l'éditeur :

<http://dx.doi.org/10.1515/BC.2008.055>

La version imprimée de cette thèse peut être consultée à la bibliothèque ou dans un autre établissement via une demande de prêt entre bibliothèques (PEB) auprès de nos services :

<http://scd.unistra.fr/services/peb/>

---

## **Annexe 2: Techniques de cartographie en solution des ARN.**

Ce chapitre de "Methods in Molecular Biology" explique le principe et les applications des expériences de cartographie en solution des ARN par diverses sondes enzymatiques ou par l'acétate de plomb. Ici sont donnés les avantages et les inconvénients de cette méthode ainsi que les protocoles détaillés qui ont été utilisés au cours de mon travail de thèse.

# Chapter 16

## Probing mRNA Structure and sRNA–mRNA Interactions in Bacteria Using Enzymes and Lead(II)

Clément Chevalier, Thomas Geissmann, Anne-Catherine Helfer, and Pascale Romby

### Summary

Enzymatic probing and lead(II)-induced cleavages have been developed to study the secondary structure of RNA molecules either free or engaged in complex with different ligands. Using a combination of probes with different specificities (unpaired vs. paired regions), it is possible to get information on the accessibility of each nucleotide, on the binding site of a ligand (noncoding RNAs, protein, metabolites), and on RNA conformational changes that accompanied ligand binding or environmental conditions (temperature, pH, ions, etc.). The detection of the cleavages can be conducted by two different ways, which are chosen according to the length of the studied RNA. The first method uses end-labeled RNA molecules and the second one involves primer extension by reverse transcriptase. We provide here an experimental procedure that was designed to map the structure of mRNA and mRNA–sRNA interaction in vitro.

**Key words:** RNA, RNA–RNA interaction, Secondary structure, RNA structure probing, Ribonuclease, Lead(II)-induced cleavages

---

### 1. Introduction

The diversity of RNA functions is intimately linked to the capability of the RNA to adopt different conformations. RNA switches have the obvious advantage to promote diverse points of contacts for the selective recognition of various RNA ligands. For instance large RNA–protein complexes such as the ribosome require numerous proteins that coordinate RNA conformational changes during the assembly (1, 2). More recently it was shown

Chevalier et al.

that regulatory regions of bacterial mRNAs undergo conformational changes in response to the environmental cues (thermosensors) (3) or to metabolite concentration changes (riboswitches) (4, 5). In addition, many novel small noncoding RNAs (sRNA) have been identified in bacterial genomes (6). A large class of sRNA represses or activates gene expression through direct pairings with target mRNA (7). Mechanistic studies reveal that the interaction between sRNA and mRNA is usually constrained by the structures of both RNAs, and that fast pairings are required for efficient regulation (8, 9). Over the years the determination of high-resolution RNA and ribonucleoprotein particle (RNP) structures by X-ray analysis and NMR provided a guide for further ongoing mechanistic and functional studies (4, 10). While a huge progress has been made in this field with the discovery of many novel RNA structural motifs (11), it is however still difficult to predict accurately the structure of long RNA molecules such as mRNA simply based on the primary structure (12). Therefore enzymatic and chemical probing remains a useful approach to monitor the RNA structure of any size under a large variety of experimental conditions. Structure mapping in solution provides the accessibility or the reactivity of each nucleotide toward chemicals or enzymes, and identifies without ambiguity the unpaired RNA regions. The elaboration of a secondary structure model is obtained by coupling the mapping data with the help of several computer folding programs based on energy minimization (13), statistics (14), or stochastic simulations (15, 16), and if available with phylogenetic and sequence comparison (*see* Chapter “Structural Probing of RNA Thermosensors”). The effect of mutations at strategic positions and base compensatory changes can further validate the existence of helices. Enzymes are also extensively used to map the binding site of a ligand (noncoding RNAs, protein, metabolites), to study RNP assembly, and to monitor the mRNA conformational changes that accompanied ligand binding or biochemical activity.

Ribonucleases were first adapted for RNA sequencing methodology (17). Most of the enzymes induce cleavages within unpaired RNA regions (18), whereas RNase VI is the only probe that provides positive signal for the existence of helical regions (19, 20) (*see* **Table 1**). The enzymes are easy to handle and are useful to identify secondary structure RNA elements such as hairpin or pseudoknot motifs. Due to the bulky size of the RNases, they are however sensitive to steric hindrance. Another probe which provides information on unpaired regions is the divalent cation lead(II). Since stable RNA helices are resistant to lead(II)-induced cleavages, the probe was largely used to map prokaryotic mRNA–sRNA interaction *in vitro* (21–24). Lead(II) was also

## Probing mRNA Structure and sRNA–mRNA Interactions in Bacteria Using Enzymes

**Table 1**  
**Properties of the probes. *MW* is for molecular weight, *G* for guanine, *A* for adenine, *C* for cytosine, *U* for uracile, *N* for nucleotide**

Probes	MW	specificity	Product	Special considerations
RNase T1	1,1000	Unpaired G	...Gp3'	Active under a wide range of conditions: pH, <i>T</i> between 4 and 55°C, with or without magnesium ions and monovalent ion, active in urea
RNase T2	3,6000	Unpaired A > C, G, U	...Ap3'	Active under different conditions: <i>T</i> between 4 and 40°C, works with or without magnesium ions
RNase A	1,2500	Unpaired C and U	...C/Up3'	Active under a wide range of conditions: <i>T</i> °C between 4 and 55°C, with or without magnesium ion and monovalent ion, active in urea
Nuclease S1	3,2000	Unpaired N	...N <sub>OH</sub> 3'	Optimum pH at 4.5, requires Zn <sup>2+</sup> for its activity
RNase V1	1,5900	Paired or stacked N	...N <sub>OH</sub> 3'	Requires divalent cations and active under a wide range of <i>T</i> (from 4 to 45°C) and pH (4–9)
Pb(II) acetate	200	Specific divalent ion sites  Unpaired regions	...Np3'	Cleaves under a large variety of experimental conditions. Ratio Mg <sup>2+</sup> /Pb <sup>2+</sup> influences the reaction Appropriate for footprinting, can be used in vivo

found to be appropriate to map the structure of mRNAs and regulatory RNAs in living bacterial cells since it easily penetrates the membrane (25, 26). Thus, the comparison between in vivo and in vitro mapping determines the functional RNA structure.

In the present review, we provide an experimental guide of the most commonly used enzymes and lead(II)-induced cleavages for mapping RNA structure in vitro. Other detailed protocols used for chemical probing in vitro will be found in this issue (see Chapters “Structural Probing of RNA Thermosensors” and “Analysis of the RNA Backbone: Structural Analysis of Riboswitches by In-Line Probing and Selective 2'-Hydroxyl Acylation and Primer Extension”) and have been previously reported (27, 28).

---

## 2. Materials

### 2.1. RNA Preparation and Renaturation

1. Safety rules have to be applied for handling radioactive materials (*see Note 1*).
2. In vitro transcribed RNA of interest (e.g., *Staphylococcus aureus* regulatory RNAlII and its target mRNA), purified and dephosphorylated at 5' end.
3. T4 polynucleotide kinase (PNK), 10 U/ $\mu$ L, supplied with 10 $\times$  T4 PNK buffer (Fermentas Vilnius, Lithuania).
4. T4 RNA ligase (Ambion, Austin, TX).
5. Radiochemicals: [ $\gamma$ - $^{32}$ P]ATP (3,000 Ci/mmol); [5'- $^{32}$ P]pCp (3,000 Ci/mmol).
6. RNA elution buffer: 500 mM ammonium acetate, pH 6.5, 1 mM ethylenediaminetetraacetic acid (EDTA).
7. Phenol saturated with 0.1 M Na-acetate, pH 6.5.
8. Phenol/chloroform/isoamyl alcohol (25:24:1) mixture, pH adjusted to 8.0 with Tris-HCl.

### 2.2. Enzymatic and Lead(II) Hydrolysis

1. Basic laboratory materials and equipment are required such as microcentrifuge, vortex, thermoblock, water bath, radioactivity counter, autoradiography films, intensifying screens.
2. The buffer conditions for enzymatic hydrolysis and lead(II) have to be adapted according to the mRNA and the nature of the ligand. Composition of buffers is given at the final concentration in the assays. Stock solutions, that are usually prepared, are 5–10 times concentrated. All solutions, water, Eppendorf tubes and tips should be RNase free. The buffers are stored at  $-20^{\circ}\text{C}$ .
3. Buffer N1: 50 mM Tris-HCl, pH 7.5, 10 mM  $\text{MgCl}_2$ , 100 mM KCl.
4. Buffer N2: 50 mM Tris-HCl, pH 7.5, 10 mM  $\text{MgCl}_2$ , 100 mM KCl, 1 mM  $\text{ZnCl}_2$ .
5. Buffer N3: 50 mM Tris-acetate, pH 7.5, 5 mM Mg-acetate, 100 mM K-acetate (*see Note 2*).
6. Several ribonucleases are commercially available. RNase T1 (specific for unpaired guanines) can be purchased from Fermentas (ref. EN0541, 1,000 U/ $\mu$ L), RNase T2 (specific for unpaired residues with a preference for adenine) is from Invitrogen (Carlsbad, CA) (ref. 18031-013, 20 U/ $\mu$ L), RNase V1 (specific for double stranded regions) is from Pierce (Rockford, IL) (ref. MB092700, 0.9 U/ $\mu$ L) or from Ambion (Billerica, MA) (ref. 2275, 0.1 U/ $\mu$ L), nuclease S1 (specific for unpaired regions) is from (Madison, WI) (M5761, 100 U/ $\mu$ L), and RNase A (specific for unpaired C and U residues) is from Ambion (AM2274, 1  $\mu\text{g}/\text{mL}$ ).

## Probing mRNA Structure and sRNA–mRNA Interactions in Bacteria Using Enzymes

7. Lead(II) acetate (ref. 31723, Acros organics, Geel, Belgium).
8. 1 µg/mL total yeast tRNA (Sigma).
9. Precipitation/inactivation buffer (Ambion).
10. 0.3 M Na-acetate, pH 6.0.

**2.3. Fractionation  
of End-Labeled RNA  
Fragments**

1. Buffer ΔT1: 20 mM sodium citrate of pH 4.5, 1 mM EDTA, 7 M urea, 0.02% (w/v) xylene cyanol, 0.02% (w/v) bromophenol blue.
2. Ladder Buffer: 0.1 M Na<sub>2</sub>CO<sub>3</sub>/0.1 M NaHCO<sub>3</sub>, pH 9.
3. RNA-loading buffer: 0.02% xylene cyanol, 0.02% bromophenol blue in 8 M urea.

**2.4. Detection of  
Cleavages by Primer  
Extension**

1. 1× RTB buffer: 50 mM Tris–HCl of pH 8.0, 10 mM MgCl<sub>2</sub>, 50 mM KCl, 1 mM dithiothreitol (DTT).
2. dNTP mix: 2.5 mM of dATP, dGTP, dCTP, dTTP (Amersham).
3. Avian myeloblastosis virus reverse transcriptase (AMV RT) purchased from MP Biochemicals (France) or Life Sciences (USA).
4. DNA-loading buffer: 1 mM EDTA, 0.02% xylene cyanol, 0.02% bromophenol blue in formamide.
5. Oligodeoxyribonucleotide primer 5'-AGGGAATGTTTTA-CAGTAT-3'.
6. Micro Bio-spin 6 chromatography column (Biorad).
7. RNA hydrolysis buffer: 50 mM Tris–HCl of pH 7.5, 7.5 mM EDTA, 0.5% SDS.
8. Individual dideoxyribonucleotide nucleotide triphosphates (ddNTP).

**2.5. Fractionation of  
Cleaved Fragments  
by Polyacrylamide-  
Urea Gel Electro-  
phoresis**

1. Electrophoresis apparatus for slab gels (30 × 40 cm) and generator (2,000 W).
2. 1× TBE buffer: 0.09 M Tris–borate of pH 8.3, 1 mM EDTA.
3. 25% Polyacrylamide gel in 8 M urea: dissolve 480 g urea in 625 mL Rotiphorese 40 solution (acrylamide/bis-acrylamide ratio 19:1) (Roth, Karlsruhe, Germany), adjust the volume to 1 L with bidistilled water, filtrate the solution.
4. Other reagents for PAGE: 8 M urea, *N*, *N*, *N'*, *N'*-tetramethylethylenediamine (TEMED); ammonium persulfate (APS) should be prepared as a 10% (w/v) solution in water (*see Note 3*).
5. Gel-fixing solution: 10% ethanol, 6% acetic acid in water. Prepare 2 L before gel fixing.

---

### 3. Methods

#### 3.1. Setting the Experimental Conditions

1. Probing the conformation of RNAs with different enzymes requires defined buffer conditions (pH, ionic strength, magnesium concentration, temperature). The optimal conditions vary with the enzymes, and subtle conformational changes of the RNA may occur under different experimental conditions (**Table 1**). For instance, nuclease S1 requires  $Zn^{2+}$  for its activity and its optimal pH is 4.5. Some of the enzymes, as RNases T1 and A, do not require magnesium for hydrolysis and can work at different temperatures (29). Thus, the influence of monovalent or divalent ion (such as magnesium) can be tested on the RNA folding, and thermal transition of RNA molecules can be obtained by varying the temperature (*see Note 4*). Such experiments provide information on the stability of the secondary structure domains. They also allow the identification of tertiary elements since these interactions are the first to break during the melting of an RNA structure. The functioning of some of the riboswitches is dependent on the kinetics of ligand binding (30). In addition, the kinetics for complex formation between sRNA and target mRNAs is essential for regulation (9). In most of the cases, footprinting assays using enzymes are conducted under equilibrium conditions. In order to get a dynamic view of complex formation, time-resolved kinetic footprinting assays have been developed with probes that generate hydroxyl radicals in the ms range (2, 31, 32). Recent methods make use of a quench-flow apparatus and exploit reactions that are faster than the interactions between mRNA with their ligands (33). Such a method can be certainly adapted with lead(II) ion.
2. The probe to RNA ratio must be adapted so that the experiments are conducted under limited and statistical conditions in order to get less than one cut per molecule. For the first experiment, different concentrations of the probes and a time-scale dependence should be performed. This is also required when the commercial source of the probe has been changed. As mentioned earlier, defined mild buffer conditions have to be used for all the probes in order to be close to the *in vivo* conditions (neutral pH, presence of divalent and monovalent ions, temperature). In that respect, RNase T1 (specific for unpaired guanines), RNase A (specific for unpaired cytosines and uridines), RNase VI (specific for paired regions), and lead(II) (specific for unpaired regions) can be used under strictly identical experimental conditions and in addition provide complementary information. Since the purification of RNase T2 will be no more commercially available, nuclease S1

## Probing mRNA Structure and sRNA–mRNA Interactions in Bacteria Using Enzymes

can provide equivalent data. However the activity of nuclease S1 is considerably reduced at pH 7.5, and hence large amount of enzyme should be added in the assays. Reducing agents (DTT, or  $\beta$ -mercaptoethanol) should be included in the footprinting assays if the RNA ligand is a protein (*see Note 5*).

3. The protocols presented here have been used for the analysis of the *Staphylococcus aureus* regulatory RNAlII free or bound to target mRNAs (22, 34). This RNA is the intracellular effector of the quorum-sensing system and regulates in a coordinated way numerous virulence factors (35, 36). Typical experiments on end-labeled RNAlII are shown in **Fig. 1** (*see Note 6*). The probing experiments were useful to delimitate the base pairings between the *S. aureus* RNAlII and the target mRNA (*see Note 7*). They also suggested that noncanonical base pairs could form within the duplex and thus would contribute to enhance the stability of the pairings. Interestingly, in several cases the structures of the interacting RNAs could impose topological constraints that limit the propagation of the intermolecular base pairings even if the extent of base pairing between the two RNAs is predicted to be rather long (34, 37). Furthermore, probing the RNA structure was helpful to design mutations at strategic positions of the RNAs to evaluate their effects on *in vivo* regulation.

### 3.2. Detection Methods

1. The identification of the cleavages can be done by two different methodologies depending on the length of the RNA molecule. The first method, which uses end-labeled RNA, is limited to molecules containing less than 200 nucleotides due to the gel resolution limitation (17). This method can only detect cleavages and is well appropriate for many of the bacterial sRNAs, which have a size below 300 nucleotides. However, for long RNAs such as *S. aureus* RNAlII (514 nucleotides long), the primer extension approach is preferable (38). The latter method detects stops of reverse transcription (RT) at the residue preceding a cleavage or a modification at a Watson–Crick position. Thus, this approach can be used for many enzymes and chemical probes and can be applied to RNA of any size. The length of the primer varies usually from 12 to 18 nucleotides. For long RNA, primers are selected every 200 nucleotides. Before probing the RNA structure, assays should be performed to define the best concentration of the RNA, the choice of the primer sequence, and the hybridization conditions in order to get an efficient primer extension.

### 3.3. RNA Preparation

#### 3.3.1. RNA Transcription

RNA is typically transcribed *in vitro* with T7 RNA polymerase from a plasmid template carrying the T7 promoter fused to the gene of interest (39) (*see Note 8*). The RNA is then separated



## Probing mRNA Structure and sRNA–mRNA Interactions in Bacteria Using Enzymes

from shorter RNA fragments, DNA template, and the excess of NTP by using either gel filtration column (40), Mono Q column (41), or denaturing polyacrylamide-urea gel electrophoresis (PAGE) (39). If the RNA is not homogeneous after transcription due to the presence of abortive transcription or cleavage products, it is worthwhile to use PAGE instead chromatography for the purification procedure. However the elution process of the RNA from the gel may not be highly efficient for long RNAs (>500 nucleotides).

## 3.3.2. End Labeling of RNA

1. For 5' end labeling, the RNA should be dephosphorylated at its 5' end, and labeled using [ $\gamma$ - $^{32}$ P]ATP and T4 PNK according to the Ambion protocol ([http://www.ambion.com/techlib/misc/RNA5\\_labeling.html](http://www.ambion.com/techlib/misc/RNA5_labeling.html)). To avoid the 5' dephosphorylation of RNA, which is not highly efficient for structured RNAs, in vitro RNA transcription can be carried out in the presence of ApG.
2. The 3' end labeling is performed with [ $5'$ - $^{32}$ P]pCp and T4 RNA ligase as previously described (42). A detailed protocol is also given by Ambion ([http://www.ambion.com/techlib/misc/RNA3\\_labeling.html](http://www.ambion.com/techlib/misc/RNA3_labeling.html)). The labeled RNAs should be purified by denaturing PAGE on 8% polyacrylamide (0.5% bis-acrylamide)/8 M urea slab gels (for gel preparation *see Subheading 3.5.2, step 1*). Gel filtration column can also be used to separate the labeled RNA from the excess of unincorporated radioisotope. However it has to be checked that the RNA is homogeneous and is not cleaved after the labeling procedures.

## 3.3.3. Purification and Renaturation of RNA

1. After PAGE purification, labeled or cold RNAs are eluted from gel slices covered with the RNA elution buffer in the presence of 10% (v/v) phenol by gentle mixing at 4°C overnight. Add equal volume of phenol/chloroform/isoamyl alcohol mixture, mix the samples for 1 min, and centrifuge 1 min at high speed. Take carefully the aqueous phase containing the RNA, transfer the solution into a new sterile 1.5-mL micro tube, and add 2.5 volumes of cold ethanol for RNA precipitation. Incubate the mixture at -20°C overnight and collect RNA by centrifugation at  $13,000 \times g$  for 30 min. After two washing steps with 200  $\mu$ L of 70% cold ethanol, the pellet is vacuum dried and dissolved in sterile water. Since the RNA is purified under denaturing conditions, it is worth spending effort to carry out a renaturation process before the probing experiments (*see Note 9*). One protocol is as follows: the RNA is preincubated 1 min at 90°C in sterile water, quickly cooled on ice for 1 min and incubated at 20°C or at 37°C in the appropriate buffer containing MgCl<sub>2</sub> for 20 min.

Chevalier et al.

### **3.4. Enzymatic Hydrolysis and Lead(II)-Induced Cleavages**

All reactions are conducted in a total volume of 10  $\mu\text{L}$ . Appropriate dilutions of enzymes and of lead(II)-acetate are done in sterile water or in the commercial buffers just before use. The dilutions of enzymes are extemporaneously prepared. For footprinting assays, the mRNA–RNAIII complex is preformed before the enzymatic or Pb(II) reaction in the appropriate buffer optimal for binding. Incubation controls in the absence of the probes and in the presence (or in the absence) of the RNA ligand are always performed in order to detect nonspecific cleavages in RNA or RT pauses (*see Note 10*). In these controls, the enzyme or lead(II) are replaced by sterile water or buffer used for the dilution of enzymes.

#### **3.4.1. Enzyme Hydrolysis**

1. Labeled RNAIII or mRNA (50,000 cpm, 1  $\mu\text{L}$ ), or the cold RNA species (1 pmol, 1  $\mu\text{L}$ ) are denatured in 4  $\mu\text{L}$  sterile water at 90°C for 1 min and then cooled on ice for 2 min. Centrifuge briefly and keep the tubes in ice.
2. 2  $\mu\text{L}$  5 $\times$  buffer N1 is added and the samples are incubated at 20°C (or 37°C) for 15 min for renaturation.
3. mRNA–RNAIII duplex formation is performed at 25°C (or 37°C) for 10 min in the presence of increasing concentrations of RNAIII or mRNA (final concentrations 100, 200, 400 nM). The samples are mixed and centrifuged briefly.
4. 1  $\mu\text{L}$  total tRNA is added to all samples.
5. Enzymatic hydrolysis is performed by addition of 1  $\mu\text{L}$  RNase as follows:
  - RNase T1 (0.2 U/ $\mu\text{L}$ ), 10 min at 20°C or 5 min at 37°C
  - RNase T2 (0.05 U/ $\mu\text{L}$ ), 10 min at 20°C or 5 min at 37°C
  - RNase A (1  $\mu\text{g}/\text{mL}$ ), 10 min at 20°C or 5 min at 37°C
  - RNase V1 (0.05 U/ $\mu\text{L}$ ), 5 min at 25°C or 2 min at 37°C
6. The RNases are added on one side of the Eppendorf tube and reactions are initiated after brief mixing and centrifugation (1 s) of the samples.
7. For nuclease S1, cleavage reaction is carried out in Buffer N2 in the presence of 50 U of enzyme for 10 min at 20°C (or 5 min at 37°C).
8. In order to define the best conditions for the hydrolysis, it is important to try initially three different concentrations of the enzymes: RNase T1 (0.1, 0.2, 0.5 U), RNase T2 (0.01, 0.05, 0.1 U), RNase A (0.25, 1, 2  $\mu\text{g}$ ), RNase V1 (0.01, 0.05, 0.1 U), and nuclease S1 (25, 50, 100 U).

#### **3.4.2. Lead(II)-Induced Cleavages**

1. Labeled mRNA or RNAIII (1  $\mu\text{L}$ , 50,000 cpm), or the cold RNA species (1 pmol, 1  $\mu\text{L}$ ) are denatured (*see Subheading 3.4.1, step 1*) and renatured in the presence of 2  $\mu\text{L}$  of buffer N3 (5 $\times$ ) at 20°C (or 37°C) for 15 min.
2. mRNA–RNAIII duplex formation is performed at 25°C (or 37°C) for 10 min in the presence of increasing concentrations

## Probing mRNA Structure and sRNA-mRNA Interactions in Bacteria Using Enzymes

of RNAlIII or mRNA (final concentrations: 100, 200, 400 nM). Mix and centrifuge briefly the samples.

3. 1  $\mu\text{L}$  of total tRNA is added to all samples.
4. Hydrolysis is initiated with 2.5  $\mu\text{L}$  of different concentrations of lead(II)-acetate from 12, 40, 80 to 120 mM for 10 min at 20°C or 5 min at 37°C. Mix and centrifuge briefly the samples (*see Note 11*). The best results in our hands were with 40 mM.

**3.4.3. Reaction Termination**

## Enzymatic Hydrolysis

1. Enzymatic hydrolysis are arrested by adding 20  $\mu\text{L}$  of the precipitation/inactivation buffer.
2. After vigorous mixing, the samples are left in a dry-ice-ethanol bath for 10 min and centrifuged at  $13,000 \times g$  at 4°C for 15 min.
3. The supernatants are discarded (if using end-labeled RNA, check that no radioactivity is present) and the pellets are washed twice with 200  $\mu\text{L}$  of 70% cold ethanol. After a short centrifugation at  $13,000 \times g$  for 5 min at 4°C, the supernatants are discarded and the pellets are vacuum dried (not more than 5 min).
4. End-labeled RNA fragments are dissolved in 6  $\mu\text{L}$  of RNA-loading buffer, whereas cold RNA fragments are dissolved in 4  $\mu\text{L}$  of sterile water.
5. If a protein is used as a ligand in footprinting assays, it is advised to make a phenol extraction. In that case, add to all samples, 40  $\mu\text{L}$  of 0.3 M Na-acetate of pH 6.0 and 50  $\mu\text{L}$  of phenol/chloroform/isoamyl alcohol mixture. Mix the samples for 1 min and centrifuge 1 min at high speed. Take carefully the aqueous phase containing the RNA, transfer the solution into a new sterile 1.5-mL micro tube, and add 2.5 volumes of cold ethanol (~150  $\mu\text{L}$ ) for RNA precipitation. The samples are then precipitated and treated as in **Subheading 3.4.3.1, steps 2–4**.

## Lead(II)-Induced cleavages

1. The reactions are stopped by adding 5  $\mu\text{L}$  of 0.1 M EDTA.
2. Add 50  $\mu\text{L}$  of 0.3 M Na-acetate of pH 6.0 and 150  $\mu\text{L}$  of cold ethanol to all samples. After a vigorous mix, the samples are transferred in a dry ice-ethanol bath for 10 min. The samples are then precipitated and treated as in **Subheading 3.4.3.1, steps 2–4**.

**3.5. Fractionation of End-Labeled RNA Fragments****3.5.1. Ladders for Cleavage Assignments**

1. RNase T1 ladder: labeled mRNA (25,000 cpm) is preincubated at 50°C for 5 min in 5  $\mu\text{L}$  of the Buffer  $\Delta\text{T1}$  containing 1  $\mu\text{g}$  total tRNA. Reaction is then performed at 50°C for 10 min in the presence of 1  $\mu\text{L}$  of RNase T1 (0.5 U).
2. Alkaline ladder: labeled mRNA (100,000 cpm) is incubated at 90°C for 3 min in the presence of total tRNA (2  $\mu\text{g}$ ) in 5  $\mu\text{L}$  of the Ladder Buffer.
3. Both ladders can be prepared for several experiments and stored at -20°C. Do not heat the samples before loading on PAGE.

Chevalier et al.

### 3.5.2. Separation of End-Labeled RNA Fragments by PAGE

1. The end-labeled RNA fragments are separated by electrophoresis on 12 or 15% polyacrylamide-(0.5% bis) – 8 M urea slab gels (0.35 mm × 30 cm × 40 cm) in 1× TBE. To prepare 100 mL of 15% gel, mix 60 mL 25% polyacrylamide – 8 M urea solution, 10 mL 10× TBE buffer, 30 mL 8 M urea, 75 µL TEMED, and 750 µL 10% APS. The gel solution is poured slowly between two glass plates that are separated by one spacer on each side and placed horizontally on the bench. After polymerization (~30 min), the comb is removed and the wells are washed carefully.
2. Prerun the gel at 75 W for 30 min using 1× TBE as running buffer.
3. The samples are heated (except the RNase T1 and alkaline ladders) for 3 min at 90°C, centrifuged briefly and 3-µL aliquots are loaded per well. Before loading, be aware that each sample contains the same amount of radioactivity (except for the ladder that should have twice more radioactivity). Load in the following order: incubation controls, reactions on free RNA and in the presence of increasing concentrations of ligand, RNase T1 ladder, and alkaline ladder.
4. Run PAGE at 75 W to heat the gel and to avoid band compression. The migration time must be adapted to the length of the RNA, knowing that on 15% polyacrylamide gel, xylene cyanol migrates as 39 nucleotide- and bromophenol blue as 9 nucleotide-long RNA. Short migration on 15% gel is convenient to fractionate small-size fragments (1–50 nucleotide-long RNA fragments). For a 250 nucleotide-long RNA, a longer migration on 12% PAGE is necessary to assign the cleavages on the whole RNA molecule (*see Fig. 1*).
5. At the end of the run, remove carefully the upper glass, fix the 12% gel for 5–30 min in Fixing gel solution, transfer to Whatman 3-MM paper, and dry for 30 min at 80°C. The 15% gel can be transferred without drying on an old autoradiography film and wrapped with a plastic film. Overnight exposure is done at –80°C using an intensifying screen.
6. Several enzymatic properties have to be taken into account when reading the gels (*see Notes 12 and 13*).
7. Several technical problems may be revealed during PAGE (*see Notes 14–17*).

### 3.6. Detection of Cleavages by Primer Extension Analysis

#### 3.6.1.5' End Labeling of Oligodeoxyribonucleotide Primer

1. The following reagents are mixed: 10 µL of 5 µM oligodeoxyribonucleotide primer, 5 µL of [ $\gamma$ -<sup>32</sup>P]ATP, 2 µL of 10× T4 PNK buffer, 2.5 µL of sterile water, 0.5 µL T4 polynucleotide kinase. Incubate at 37°C for 1 h.
2. To the sample, add 30 µL sterile water, load on a Micro Biospin 6 chromatography column (Biorad), and centrifuge at 1,000 × *g* for 5 min at 20°C. The volume is then adjusted with

## Probing mRNA Structure and sRNA–mRNA Interactions in Bacteria Using Enzymes

sterile water to get 100,000 cpm/ $\mu\text{L}$  of end-labeled DNA primer. This rapid purification protocol is sufficient to remove the excess of radioactive ATP.

**3.6.2. Hybridization**

1. To the 4  $\mu\text{L}$  of the cleaved mRNA (1 pmol), add 1  $\mu\text{L}$  of 5' end-labeled DNA primer (around 100,000 cpm).
2. The samples are heated 1 min at 90°C and quickly cooled on ice after a brief centrifugation.
3. 1  $\mu\text{L}$  5 $\times$  RTB buffer is added and the samples are incubated for 15 min at 20°C.

**3.6.3. Primer Extension**

1. The reaction is done in 15  $\mu\text{L}$ .
2. To the hybridization mix, add 2  $\mu\text{L}$  5 $\times$  RTB, 2  $\mu\text{L}$  dNTP mix (2.5 mM of each dNTP), 4  $\mu\text{L}$  sterile water, 1  $\mu\text{L}$  RT (2 U/ $\mu\text{L}$ , diluted freshly in the commercial buffer). The samples are incubated for 30 min at 37°C.
3. To all samples, add 50  $\mu\text{L}$  of 0.3 M Na-acetate of pH 6.0 and 200  $\mu\text{L}$  of cold ethanol. The samples are then precipitated and treated as in **Subheading 3.4.3.1, steps 2–3**. The end-labeled DNA fragments are dissolved in 6  $\mu\text{L}$  of DNA-loading buffer. All samples are adjusted to the same amount of radioactivity per 1  $\mu\text{L}$ .
4. To improve the quality of the gels, the RNA template may be hydrolyzed by alkaline treatment. Just after primer extension, add 20  $\mu\text{L}$  of the RNA hydrolysis buffer and 3.5  $\mu\text{L}$  of 3 M KOH. The samples are heated at 90°C for 3 min and at 37°C for at least 1 h. To all samples, add 6  $\mu\text{L}$  3 M acetic acid, 100  $\mu\text{L}$  0.3 M Na-acetate of pH 6.0, and 300  $\mu\text{L}$  of cold ethanol. After precipitation (*see Subheading 3.4.3.1, steps 2–3*), the pellets are washed twice with 70% ethanol, vacuum dried, and dissolved in 6  $\mu\text{L}$  of DNA-loading buffer.

**3.6.4. Gel Fractionation of Labeled cDNA Fragments**

1. The cleavage positions are identified by running in parallel a sequencing reaction (*see* Chapter “Structural Probing of RNA Thermosensors”). The elongation step is performed as described in **Subheading 3.6.3** except that in the presence of one of the dideoxynucleotides ddXTP (2.5  $\mu\text{M}$ ), the corresponding deoxyribonucleotide dXTP (25  $\mu\text{M}$ ) and the three other deoxyribonucleotides (100  $\mu\text{M}$ ) are added.
2. All samples are heated at 90°C for 3 min, and centrifuged briefly.
3. 3- $\mu\text{L}$  Aliquots are loaded per well on 8% polyacrylamide-(0.4% bis)/8 M urea slab gels in 1 $\times$  TBE. The migration conditions must be adapted to the size of the fragments to be analyzed, knowing that on 8% polyacrylamide gel, xylene cyanol migrates as 81 nucleotide- and bromophenol blue as 19 nucleotide-long RNA. After migration, the gels are dried, and exposed with an X-ray film and intensifying screen overnight at –80°C.

Chevalier et al.

4. Technical problems that are revealed by the autoradiography may occur during the handling process (*see* **Notes 16–20**).

---

#### 4. Notes

1. Safety rules: for manipulating radioactivity, work behind a plexiglass screen, and wear glasses and gloves. Perform phenol extraction under a fume hood.
2. For lead(II)-induced cleavages, buffers with chloride ions should be avoided since  $\text{PbCl}_2$  may precipitate in solution.
3. To keep high resolution of the gels, acrylamide, urea solutions, and in particular ammonium persulfate should be prepared freshly.
4. RNase T1 and RNase A hydrolysis can be conducted under a variety of experimental conditions. It is essential however to adapt the enzymatic hydrolysis for each condition in order to have less than one cut or modification per molecule, i.e., more than 80% of the RNA should not be modified or cleaved.
5. If no effect of ligand binding is observed in footprinting experiments, it is essential to define the experimental conditions (buffer, temperature, ions) that are required for efficient binding by other methodologies such as bandshift analysis. The footprinting experiments should also be performed in the presence of increasing concentrations of the ligand (*see* **Fig. 1**). Such experiments might reveal different ligand-binding sites on the RNA molecule.
6. Each experiment should be repeated at least twice, and only the reproducible cleavages will be considered. As mentioned previously, the elaboration of a secondary structure RNA model requires data from enzymes of complementary specificities. Only the combination will help to define helical and loop regions.
7. Results should be interpreted with care because protection does not necessarily result from a direct shielding effect, but could be due to a steric hindrance effect (particularly observed with the bulky RNases) or to a conformational change of the RNA. Contrarily, enhanced cleavages result from RNA conformational changes (**Fig. 1**).
8. In vitro RNA transcripts generated by T7 RNA polymerase suffer from a considerable degree of heterogeneity at 3' end and sometimes from heterogeneity at the 5' end. This can be revealed by doubled cleavages using end-labeled RNA. The heterogeneity poses a significant problem when analyzing the cleavage sites using long end-labeled RNAs. Heterogeneity can

## Probing mRNA Structure and sRNA–mRNA Interactions in Bacteria Using Enzymes

be eliminated by addition of self-cleaving ribozyme sequences at the 5'- and/or 3''-end of the target RNA sequence (43). PAGE is typically the method of choice for separation of ribozymes from the RNA of interest. For long RNA molecules, electro-elution might help to increase the elution efficiency.

9. During the purification, RNA can be partially denatured; therefore, it is essential to design “renaturation” protocols in order to obtain “conformationally homogeneous” RNA population. It is also necessary to test whether this conformation is biologically relevant (enzymatic activity for ribozyme, efficient ligand binding). Alternative RNA conformations may coexist, and can be revealed by the simultaneous presence of single-stranded and double-stranded specific cleavages. By varying the concentration of  $MgCl_2$ , one of the two conformers might be stabilized.
10. Appropriate incubation controls are essential to identify cleavages that are induced during the incubation treatments, and the pauses of reverse transcriptase that are due to stable secondary structures or cuts. Nucleotides for which strong bands are visible in the control lanes are not considered for interpretation. Incubation control in the presence of the ligand has to be done since the ligand might be contaminated with traces of RNase during its purification. If too many bands are observed in the incubation controls of the end-labeled RNA, repurify the RNA and prepare new sterile buffers. If too many RT pauses are observed in the incubation controls, it can be due to RNase contamination, strong secondary structure of RNA (the extension can be done at  $42^\circ C$ ), or primer location. For this experiment, AMV (Avian Myeloblastosis Virus) RT should be used rather than MMLV (Moloney Murine leukemia Virus) RT, the latter being much more sensitive to RNA secondary structure.
11. Since lead(II) competes with  $Mg^{2+}$  for RNA binding, the efficiency of cleavages will depend on the  $Pb^{2+}/Mg^{2+}$  ratio.
12. The RNase cleavages in the RNA can induce conformational rearrangements of the cleaved RNA that can potentially provide new targets for secondary cleavages. Thus, these secondary cleavages do not reflect the native structure of the RNA. Usually these cleavages are weak, are not reproducibly found in all experiments, and occur when the RNA digestion is too strong. They can be distinguished from the primary cuts by comparing the hydrolysis patterns obtained from the 5' or 3' end-labeled RNA.
13. RNase V1 and nuclease S1 hydrolysis generate RNA fragments, which end up with 3' OH and 5'-P groups in contrast to alkali and most of the RNases (Table 1). Therefore, 5' end-labeled

Chevalier et al.

fragments generated by alkali will migrate faster than the RNase VI/nuclease S1 fragments, and conversely the 3'-end-labeled RNA generated by alkali will migrate slower than the RNase VI/nuclease S1 fragments. This difference is only observed for the shortest RNA fragments (*see Fig. 1*).

14. If no full-length RNA is observed, the RNA hydrolysis is too strong. Time of the hydrolysis and/or enzyme concentrations have to be reduced.
15. Compression of bands due to stable secondary structure (in general rich in G–C base pairs) can be observed using end-labeled RNA. Heat the samples before loading on the gel and the gels should be warm before sample loading and during the migration.
16. If end-labeled RNA aggregates in the gel pockets and only fragments of small sizes can be visualized on the gel, the RNA pellet was not correctly dried after ethanol precipitation. Do not interpret the data.
17. Samples may not migrate correctly during electrophoresis due to the presence of salt. Add several washing steps with 80% ethanol at the end of the procedure.
18. Absence of radioactive signal after primer extension could mean that the modified RNA did not efficiently precipitate, or the hybridization conditions are not optimized. Since the modified RNA is not labeled, particular caution should be taken to prevent the loss of the pellet.
19. Reverse transcriptase stops at the nucleotide preceding the cleaved nucleotide. Thus, the resulting cDNA is one nucleotide shorter than the cDNA corresponding to the sequencing lane.
20. Too many reverse transcriptase stops in the control lanes may be due to several reasons: degradation of the RNA template, pauses of the enzyme due to stable RNA secondary structure (increase the temperature of elongation, change the primer), the conditions of elongation are not sufficiently optimized (adjust the enzyme or dNTP concentration).

---

### Acknowledgments

This work was supported by the Centre National de la Recherche Scientifique (UPR 9002 CNRS), the University Louis Pasteur of Strasbourg, the Ministère de la Recherche (ANR05-MIIME, ANR07-BLANC), and the European Community (FOSRAK, EC005120; BacRNA EC018618).

## Probing mRNA Structure and sRNA–mRNA Interactions in Bacteria Using Enzymes

## References

1. Kaczanowska, M., and Ryden-Aulin, M. (2007). Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* **71**, 477–494.
2. Nguyenle, T., Laurberg, M., Brenowitz, M., and Noller, H. F. (2006). Following the dynamics of changes in solvent accessibility of 16 S and 23 S rRNA during ribosomal subunit association using synchrotron-generated hydroxyl radicals. *J. Mol. Biol.* **359**, 1235–1248.
3. Narberhaus, F., Waldminghaus, T., and Chowdhury, S. (2006). RNA thermometers. *FEMS Microbiol. Rev.* **30**, 3–16.
4. Serganov, A., and Patel, D. J. (2007). Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat. Rev. Genet.* **8**, 776–790.
5. Barrick, J. E., and Breaker, R. R. (2007). The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol.* **8**, R239.
6. Gottesman, S., McCullen, C. A., Guillier, M., Vanderpool, C. K., Majdalani, N., Benhammou, J., Thompson, K. M., FitzGerald, P. C., Sowa, N. A., and FitzGerald, D. J. (2006). Small RNA regulators and the bacterial response to stress. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 1–11.
7. Storz, G., Altuvia, S., and Wassarman, K. M. (2005). An abundance of RNA regulators. *Annu. Rev. Biochem.* **74**, 199–217.
8. Wagner, E. G. H., Altuvia, S., and Romby, P. (2002). Antisense RNAs in bacteria and their genetic elements. *Adv. Genet.* **46**, 361–398.
9. Brantl, S. (2007). Regulatory mechanisms employed by cis-encoded antisense RNAs. *Curr. Opin. Microbiol.* **10**, 102–109.
10. Chen, Y., and Varani, G. (2005). Protein families and RNA recognition. *FEBS J.* **272**, 2088–2097.
11. Leontis, N. B., Lescoute, A., and Westhof, E. (2006). The building blocks and motifs of RNA architecture. *Curr. Opin. Struct. Biol.* **16**, 279–287.
12. Jossinet, F., Ludwig, T. E., and Westhof, E. (2007). RNA structure: bioinformatic analysis. *Curr. Opin. Microbiol.* **10**, 279–285.
13. Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415.
14. Do, C. B., Woods, D. A., and Batzoglou, S. (2006). CONTRAfold: RNA secondary structure prediction without physics-based models. *Bioinformatics* **22**, e90–e98.
15. Xayaphoummine, A., Bucher, T., and Isambert, H. (2005). Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Res.* **33**, W605–W610.
16. Gulyaev, A. P., van Batenburg, F. H., and Pleij, C. W. (1995). The computer simulation of RNA folding pathways using a genetic algorithm. *J. Mol. Biol.* **250**, 37–51.
17. Donis-Keller, H., Maxam, A. M., and Gilbert, W. (1977). Mapping adenines, guanines, and pyrimidines in RNA. *Nucleic Acids Res.* **4**, 2527–2538.
18. Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J. P., and Ehresmann, B. (1987). Probing the structure of RNAs in solution. *Nucleic Acids Res.* **15**, 9109–9128.
19. Lockard, R. E., and Kumar, A. (1981). Mapping tRNA structure in solution using double-strand-specific ribonuclease V1 from cobra venom. *Nucleic Acids Res.* **9**, 5125–5140.
20. Favorova, O. O., Fasiolo, F., Keith, G., Vasilenko, S. K., and Ebel, J. P. (1981). Partial digestion of tRNA–aminoacyl-tRNA synthetase complexes with cobra venom ribonuclease. *Biochemistry* **20**, 1006–1011.
21. Kolb, F. A., Malmgren, C., Westhof, E., Ehresmann, C., Ehresmann, B., Wagner, E. G. H., and Romby, P. (2000). 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* **6**, 311–324.
22. 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. (2005). *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate *spa* gene expression. *EMBO J.* **24**, 824–835.
23. Darfeuille, F., Unoson, C., Vogel, J., and Wagner, E. G. H. (2007) An antisense RNA inhibits translation by competing with standby ribosomes. *Mol. Cell* **26**, 381–392.
24. Sharma, C. M., Darfeuille, F., Plantinga, T. H., and 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.
25. Lindell, M., Romby, P., and Wagner, E. G. H. (2002). Lead(II) as a probe for investigating RNA structure *in vivo*. *RNA* **8**, 534–541.
26. Ivanova, N., Lindell, M., Pavlov, M., Holmberg Schiavone, L., Wagner, E. G. H., and Ehrenberg, M. (2007). Structure probing of tmRNA in distinct stages of trans-translation. *RNA* **13**, 713–722.
27. Huntzinger, E., Possedko, M., Winter, F., Moine, H., Ehresmann, C., and Romby, P. (2005). Probing RNA structures with enzymes and chemicals *in vitro* and *in vivo*, in *Handbook of RNA Chemistry* (Hartmann, R. K.,

Chevalier et al.

- Bindereif, A., Schön, A., and Westhof, E., eds), Wiley-VCH, Weinheim, pp. 151–171.
28. Marchand, V., Mougin, A., Méreau, A., and Branlant, C. (2005). Study of RNA–protein interactions and RNA structure in ribonucleo-protein particles, in *Handbook of RNA Chemistry* (Hartmann, R. K., Bindereif, A., Schön, A., and Westhof, E., eds), Wiley-VCH, Weinheim, pp. 172–228.
29. Waldminghaus, T., Heidrich, N., Brantl, S., and Narberhaus, F. (2007). FourU: a novel type of RNA thermometer in *Salmonella*. *Mol. Microbiol.* **65**, 413–424.
30. Coppins, R. L., Hall, K. B., and Groisman, E. A. (2007). The intricate world of riboswitches. *Curr. Opin. Microbiol.* **10**, 176–181.
31. Fabbretti, A., Pon, C. L., Hennelly, S. P., Hill, W. E., Lodmell, J. S., and Gualerzi, C. O. (2007). The real-time path of translation factor IF3 onto and off the ribosome. *Mol. Cell* **25**, 285–296.
32. Shcherbakova, I., Mitra, S., Beer, R. H., and Brenowitz, M. (2006). Fast Fenton footprinting: a laboratory-based method for the time-resolved analysis of DNA, RNA and proteins. *Nucleic Acids Res.* **34**, e48.
33. Fabbretti, A., Milon, P., Giuliodori, A. M., Gualerzi, C. O., and Pon, C. L. (2007). Real-time dynamics of ribosome–ligand interaction by time-resolved chemical probing methods. *Methods Enzymol.* **430**, 45–58.
34. Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M., Chevalier, C., Helfer, A. C., Benito, Y., Jacquier, A., Gaspin, C., Vandenesch, F., and Romby, P. (2007). *Staphylococcus aureus* RNAlII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev.* **21**, 1353–1366.
35. Novick, R. P. (2003). Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Mol. Microbiol.* **48**, 1429–1449.
36. Toledo-Arana, A., Repoila, F., and Cossart, P. (2007). Small noncoding RNAs controlling pathogenesis. *Curr. Opin. Microbiol.* **10**, 182–188.
37. Kolb, F. A., Engdahl, H. M., Slagter-Jäger, J. G., Ehresmann, B., Ehresmann, C., Westhof, E., Wagner, E. G. H., and Romby, P. (2000). Progression of a loop–loop complex to a four-way junction is crucial for the activity of a regulatory antisense RNA. *EMBO J.* **19**, 5905–5915.
38. Qu, H. L., Michot, B., and Bachelier, J. P. (1983). Improved methods for structure probing in large RNAs: a rapid ‘heterologous’ sequencing approach is coupled to the direct mapping of nuclease accessible sites. Application to the 5’ terminal domain of eukaryotic 28S rRNA. *Nucleic Acids Res.* **11**, 5903–5920.
39. Milligan, J. F., and Uhlenbeck, O. C. (1989). Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.* **180**, 51–62.
40. Romaniuk, P. J., de Stevenson, I. L., and Wong, H. H. (1987). Defining the binding site of *Xenopus* transcription factor IIIA on 5S RNA using truncated and chimeric 5S RNA molecules. *Nucleic Acids Res.* **15**, 2737–2755.
41. Jahn, M. J., Jahn, D., Kumar, A. M., and Söll, D. (1991). Mono Q chromatography permits recycling of DNA template and purification of RNA transcripts after T7 RNA polymerase reaction. *Nucleic Acids Res.* **19**, 2786.
42. England, T. E., Bruce, A. G., and Uhlenbeck, O. C. (1980). Specific labeling of 3’ termini of RNA with T4 RNA ligase. *Methods Enzymol.* **65**, 65–74.
43. Walker, S. C., Avis, J. M., and Conn, G. L. (2003). General plasmids for producing RNA *in vitro* transcripts with homogeneous ends. *Nucleic Acids Res.* **31**, e82.

---

## **Annexe 3: Technique de "toeprint"**

Ce deuxième chapitre de "Methods in Molecular Biology" explique le principe et les applications des expériences de "toeprint". Ce chapitre montre que ces expériences sont particulièrement appropriées pour surveiller l'effet des ARN régulateurs sur la régulation de l'initiation de la traduction chez les bactéries. Les protocoles qui ont été utilisés au cours de ma thèse sont également donnés.

# Chapter 18

## Ribosomal Initiation Complexes Probed by Toeprinting and Effect of trans-Acting Translational Regulators in Bacteria

Pierre Fechter, Clément Chevalier, Gulnara Yusupova, Marat Yusupov, Pascale Romby, and Stefano Marzi

### Summary

Toeprinting was developed to study the formation of ribosomal initiation complexes in bacteria. This approach, based on the inhibition of reverse transcriptase elongation, was used to monitor the effect of ribosomal components and translational factors on the formation of the active ribosomal initiation complex. Moreover, this method offers an easy way to study *in vitro* how mRNA conformational changes alter ribosome binding at the initiation site. These changes can be induced either by environmental cues (temperature, ion concentration), or by the binding of metabolites, regulatory proteins, and trans-acting RNAs. An experimental guide is given to follow the different steps of the formation of ribosomal initiation complexes in *Escherichia coli* and *Staphylococcus aureus*, and to monitor the mechanism of action of several regulators on translation initiation *in vitro*. Protocols to prepare the ribosome and the subunits are also given for *Thermus thermophilus*, *Staphylococcus aureus*, and *Escherichia coli*.

**Key words:** Ribosome, Ribosome purification, mRNA, Translation initiation, Toeprinting, Translational regulator, *Escherichia coli*, *Staphylococcus aureus*, *Thermus thermophilus*

---

### 1. Introduction

Translational regulation in bacteria results in fast adaptation of protein synthesis to environmental conditions. In *Escherichia coli*, the synthesis of many essential proteins that bind RNA, proteins involved in metabolism and in stress-related responses, are regulated at the translational level. Regulation occurs by a variety of mechanisms that influence the binding of the mRNA to the 30S ribosomal subunit at a step preceding the assembly of the active initiation complex, i.e. formation of the codon-anticodon

Fechter et al.

interactions. Many of these regulatory events involve structured elements in the 5' untranslated regions of mRNAs that are specifically recognized by a large variety of trans-acting factors ranging from metabolites and non-coding RNAs to proteins (1, 2). In other cases, structured mRNA regions can directly sense external signals like temperature (3), or the intracellular concentration of magnesium without the help of a protein (4). Conformational switches within mRNA can lead to translation activation by facilitating the accessibility of the ribosome to its Translation Initiation Region (TIR), or to repression by promoting the formation of an inhibitory structure that masks the TIR. In most of the known cases, competitive binding of the repressor and the ribosome to overlapping/adjacent mRNA-binding sites leads to the formation of mutually exclusive repressor-mRNA and ribosome-mRNA complexes (5, 6). However, in several examples, the repressor and the ribosome bind to distinct sites, raising the question of how repression may occur. Simultaneous binding of the repressor and the ribosome to the same mRNA was demonstrated for the translational regulation of the two ribosomal proteins S4 and S15 in *E. coli* (6, 7). The formation of the mRNA-repressor-ribosome complex causes ribosome stalling at the pre-initiation stage and prevents the formation of the active complex. Recent data show that repressor protein traps the mRNA structure on the platform of the 30S subunit outside the normal mRNA path and prevents the accommodation step required to form the anticodon-codon interaction in the peptidyl-tRNA (P site)-binding site (8). More recent data indicate that sRNAs bind far upstream of the TIR and prevent the formation of the active ribosomal initiation complex without the strong alteration of mRNA structure (9). Therefore, in addition to in vivo genetics and expression studies (10), it is of interest to monitor in vitro the step of the initiation process at which regulation occurs.

An easy way to inspect ribosome binding to mRNAs at the initiation step was devised by L. Gold and coworkers (11). This method, called toeprinting, involves cDNA synthesis by reverse transcriptase on a template mRNA to which the ribosome, together with the initiator tRNA, is bound. The position of the reverse transcriptase termination (toeprint) gives the exact position of the 3' edge of the mRNA with respect to the bound ribosome and the tRNA species (12). In the presence of the initiator tRNA, the toeprint was identified for different mRNAs to be located at position +16 from the AUG codon. This 3' boundary position of the ribosome was later on validated by footprinting assays (13, 14) and by X-ray analysis (15-17). Toeprinting assays were further used to show that the selection of the correct initiator tRNA is made by the three initiation factors on the 30S subunit (12). The ribosome-mRNA binary complex was also detected while being less stable than the ternary tRNA-mRNA-ribosome

## Ribosomal Initiation Complexes Probed by Toeprinting and Effect of trans-Acting

(18, 19). Thus, different steps of the formation of the initiation complex can be probed. Furthermore, a comparative study with different mRNAs indicated that the nucleotide distance between the Shine–Dalgarno sequence (SD), the P site, and the 3' edge of the mRNA is fixed while secondary structure elements can be accommodated within the mRNA track (16, 20). More recently, toeprinting was applied to monitor the effect of a novel ribosomal elongation factor LepA during the elongation process (21). Hence, it was obvious to apply this strategy to unravel various regulatory mechanisms. The technique was successfully used to prove the temperature dependence of ribosome binding to thermosensor mRNAs (22, 23) and to follow the formation of binary and ternary ribosome complexes in the presence of regulatory proteins (8) and trans-acting sRNAs (e.g. (9, 24, 25)). We present here an experimental guide for toeprinting assays and provide typical examples obtained on several mRNAs regulated by proteins or trans-acting RNAs. In addition, protocols for the purification of homogenous ribosomes are given for *Escherichia coli*, *Staphylococcus aureus*, and *Thermus thermophilus*.

---

## 2. Materials

### 2.1. Cell Culture and Lysis

1. French Press (Thermo Spectronic) and fermentor (Infors HT) are required for high-scale purification of ribosomes.
2. For *S. aureus*, ribosomes are obtained from the strain RN6390 that carries a deletion of the *rub* gene (see Note 1). *S. aureus* is grown in Brain Heart Infusion broth (AES laboratoire, Bruz, France). All bacterial cultures are stored in 25% glycerol at  $-80^{\circ}\text{C}$ .
3. *E. coli* MRE600 cells.
4. *T. thermophilus* HB8 cells.
5. LB medium: 1% (w/v) trypton, 0.5% (w/v) yeast extract, 1% (w/v) NaCl. The pH is adjusted to 7.5 with KOH.
6. BHI medium (AES laboratoire).
7. Blood-agar plates (Merck).
8. *T. thermophilus* medium: 0.5% (w/v) polypeptone; 0.2% (w/v) yeast extract; 0.2% (w/v) NaCl; vitamin-mineral solution (USBiological, Swampscott, MA), 1 mL/L media. The pH is adjusted with KOH to 7.2–7.4.

### 2.2. Ribosome Purification

All solutions and buffers are sterilized or filtrated before use. Buffers for toeprinting and reverse transcription can be stored at  $-20^{\circ}\text{C}$ .

Fechter et al.

1. Gradient maker (SG15, Amersham) connected to a peristaltic pump (Biorad).
2. Capillary tubes used to collect gradient (e.g. Corning).
3. FPLC system with fraction collector (Amersham).
4. Low-speed centrifuge (Sigma) and ultracentrifuge with rotors Ti45, Ti70.1, SW32Ti, and SW28 (Beckmann).
5. Buffer AE: 20 mM MgCl<sub>2</sub>, 200 mM NH<sub>4</sub>Cl, 20 mM Tris-HCl of pH 7.5, 0.1 mM ethylenediamine tetraacetic acid disodium salt (EDTA), 6 mM 2-mercaptoethanol.
6. Cushion 1E: 30% sucrose, 10 mM MgCl<sub>2</sub>, 500 mM NH<sub>4</sub>Cl, 20 mM Tris-HCl of pH 7.5, 0.1 mM EDTA, 6 mM 2-mercaptoethanol.
7. Buffer BE: 10 mM MgCl<sub>2</sub>, 50 mM NH<sub>4</sub>Cl, 20 mM Tris-HCl of pH 7.5, 0.1 mM EDTA, 6 mM 2-mercaptoethanol.
8. Buffer DE: 3 mM MgCl<sub>2</sub>, 300 mM NH<sub>4</sub>Cl, 20 mM Tris-HCl of pH 7.5, 0.15 mM Na<sub>2</sub>EDTA, 2 mM dithiothreitol (DTT).
9. Solution SE: 50% sucrose. To prepare 500 mL of 50% sucrose solution, dissolve 250 g of sucrose in sterile water and add one spoon of Bentonite (Sigma). Leave the solution stirring at 20°C for 1 h, and then remove Bentonite by centrifugation at 20,000 × *g* for 20 min.
10. Buffer S1: 20% sucrose (use solution SE) in Buffer DE (*see Note 2*).
11. Buffer S2: 5% sucrose (use solution SE) in Buffer DE.
12. Loading buffer for agarose gel: 40 mM Tris-acetate of pH 8.0, 1 mM EDTA, 0.25% bromophenol blue, 30% glycerol.
13. Running buffer for agarose gel: 40 mM Tris-acetate of pH 8.0, 1 mM EDTA.
14. Buffer CE: 10 mM MgCl<sub>2</sub>, 50 mM KCl, 20 mM Tris-HCl of pH 7.5, 0.1 mM EDTA, 1 mM DTT.
15. Buffer AS: 20 mM Tris-HCl of pH 7.3, 200 mM NH<sub>4</sub>Cl, 20 mM MgCl<sub>2</sub>, 1 mM EDTA, 2 mM β-mercaptoethanol.
16. Buffer BS: 20 mM Tris-HCl of pH 7.3, 500 mM NH<sub>4</sub>Cl, 10 mM MgCl<sub>2</sub>, 1 mM EDTA, 2 mM β-mercaptoethanol.
17. Buffer CS: 20 mM Tris-HCl of pH 7.3, 300 mM NH<sub>4</sub>Cl, 3 mM MgCl<sub>2</sub>, 0.15 mM EDTA, 2 mM β-mercaptoethanol. Buffer CS<sub>0.1</sub>: buffer CS with 0.1 mM MgCl<sub>2</sub>.
18. Buffer DS: 20 mM Tris-HCl of pH 7.3, 50 mM NH<sub>4</sub>Cl, 5 mM MgCl<sub>2</sub>, 0.15 mM EDTA, 2 mM β-mercaptoethanol.
19. Buffer AT: 150 mM MgCl<sub>2</sub>, 500 mM NH<sub>4</sub>Cl, 40 mM Tris-HCl of pH 7.5, 1.5 mM EDTA, 1 mM DTT.
20. Buffer BT: 50 mM MgCl<sub>2</sub>, 150 mM NH<sub>4</sub>Cl, 20 mM Tris-HCl of pH 7.5, 0.5 mM EDTA, 1 mM DTT.

## Ribosomal Initiation Complexes Probed by Toeprinting and Effect of trans-Acting

21. Cushion 1T: 1.5 M sucrose, 0.68 M CsCl, 150 mM MgCl<sub>2</sub>, 20 mM Tris-HCl of pH 7.5, 1.5 mM EDTA, 1 mM DTT.
22. Cushion 2T: 1.8 M sucrose, 0.8 M CsCl, 150 mM MgCl<sub>2</sub>, 20 mM Tris-HCl of pH 7.5, 1.5 mM EDTA.
23. Buffers CT<sub>0</sub>, CT<sub>0.8</sub> and CT<sub>1</sub>: 10 mM MgCl<sub>2</sub>, 400 mM NaCl, 20 mM Tris-HCl of pH 7.5, 0.5 mM EDTA, 1 mM DTT and 0.0, 0.8, and 1.0 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, respectively.
24. YM100 Centricon cartridges (Millipore).
25. DNase I RNase free (Roche Diagnostics GmbH, Germany).
26. Phenylmethanesulphonylfluoride (PMSF): dissolve in ethanol. Handle under hood with extreme care.
27. 10 mM Tris-HCl, pH 7.5.
28. *Staphylococcus aureus* lysostaphin (Sigma).
29. Buffer TE: 10 mM Tris-HCl of pH 8.0, 1 mM EDTA.
30. 1 M MgCl<sub>2</sub>.
31. 200-mL Toyopearl Butyl 650S column.
32. 3 M Na-acetate, pH 5.2.

**2.3. mRNA Preparation**

The RNA is transcribed in vitro with T7 RNA polymerase from a linearized plasmid template carrying the T7 promoter fused with the gene of interest (26). The RNA is then purified, as described in Chapter “Probing mRNA Structure and sRNA-mRNA Interactions in Bacteria Using Enzymes and Lead(II)” by Chevalier et al., by using gel filtration column (27) or Mono Q column (28). If RNA is not homogenous after purification due to the presence of abortive transcripts or cleavage products, a further purification using 8% polyacrylamide-(0.5% bis)/8 M urea slab gels in 1× TBE is recommended.

**2.4. Toeprinting Assays**

1. The symbols + and - in the buffer names refer to the presence or absence of MgCl<sub>2</sub>, respectively. Buffer TP5X+: 100 mM MgCl<sub>2</sub>, 100 mM Tris-HCl of pH 7.5, 300 mM NH<sub>4</sub>Cl, 5 mM DTT.
2. Buffer TP5X-: 100 mM Tris-HCl of pH 7.5, 300 mM NH<sub>4</sub>Cl, 5 mM DTT.
3. MgCl<sub>2</sub>100 solution: 100 mM MgCl<sub>2</sub>.
4. dNTP500 mix: 500 μM of each dATP, dGTP, dCTP, dTTP.
5. DNA primer for RT: the length of the primer usually varies from 12 to 18 nucleotides. The primer is complementary to a region of the mRNA located around 40 nucleotides downstream of the initiation codon. It is worthwhile to make efforts on establishing the optimal conditions for hybridization and extension.

Fechter et al.

6. DNA-loading buffer: 1 mM EDTA, 0.02% xylene cyanol, 0.02% bromophenol blue in formamide.
7. T4 polynucleotide kinase (PNK), 10 U/ $\mu$ L (Fermentas, Glen Burnie, MD) supplied with 10 $\times$  PNK buffer.
8. [ $\gamma$ - $^{32}$ P]ATP, 3,000 Ci/mmol (Amersham, France).
9. tRNA<sup>fMet</sup> (Sigma).
10. Reverse transcriptases (RT): MMLV RT (Gibco), AMV RT (MP Biochemicals or Life Sciences).
11. Total tRNA, 2  $\mu$ g/ $\mu$ L (Sigma).
12. DNA elution buffer: 500 mM ammonium acetate of pH 6.0, 1 mM EDTA.
13. Phenol: saturate with 20 mM Tris-HCl, pH 8.0.
14. Chloroform/isoamyl alcohol (24:1) mixture.

### **2.5. Buffers and Solutions for PAGE**

1. Electrophoresis apparatus for slab gels (30 cm  $\times$  40 cm, BRL) and generator (2,000 W).
2. 1 $\times$  TBE: 0.09 M Tris-borate of pH 8.3, 1 mM EDTA.
3. 25% Polyacrylamide/8 M urea solution: dissolve 480 g urea (Merck) in 625 mL Rotiphorese 40 gel solution with acrylamide/bis-acrylamide ratio 19:1 (Roth, Karlsruhe, Germany), adjust the volume to 1 L with bidistilled water, and filter the solution. Use this stock solution to prepare 8% polyacrylamide gel/8 M urea/1 $\times$  TBE gels.

---

## **3. Methods**

### **3.1. Setting the Experimental Conditions**

#### *3.1.1. Formation of Binary Complexes*

1. Efficient initiation of translation in bacteria requires the formation of a stable initiation complex where the initiation codon (mainly AUG) interacts with the initiator fMet-tRNA<sup>fMet</sup> in the ribosomal P site. This interaction is the result of the mRNA accommodation step that follows the mRNA-30S binding step (29). Prior to the toeprinting assay, the formation of the complex between the mRNA and the 30S subunit in the absence of the initiator tRNA can be monitored by different approaches, such as filter-binding assays on nitrocellulose membrane, analytical sucrose gradients, and gel filtration chromatography. These methods require labelled mRNA and can detect the formation of a stable binary mRNA-30S complex. The toeprinting assay, however, can additionally map the 3' edge of the mRNA segment shielded by the bound ribosome. It was previously shown that the reverse transcriptase stops close

## Ribosomal Initiation Complexes Probed by Toeprinting and Effect of trans-Acting

to the SD sequence of the unstructured mRNA bound to the ribosome (18). The formation of the SD–aSD helix is the most stable interaction that usually takes place in the binary complex. This complex can be detected using the reverse transcriptase (RT) from the Avian Myeloblastosis Virus (AMV) or from the Moloney Murine Leukemia Virus (MMLV). However, these RTs have different properties. Under suboptimal experimental conditions, MMLV RT is much more sensitive to the secondary structures of the mRNA than AMV RT and can detect rather weak mRNA–30S subunit interactions. The AMV RT is more processive than MMLV RT and is usually used if the binary complex is stable. An example is given in Fig. 1a with the *rpsO* mRNA. The weak interaction between the pseudoknot structure of the mRNA and the ribosome was detected only at low concentrations of MMLV RT. Hence, the use of MMLV at suboptimal conditions can be a valuable tool to monitor the effect of the structured regulatory region of the mRNA on 30S recognition in the absence of the initiator tRNA (Fig. 1a) (8).

## 3.1.2. Formation of Ternary Complexes

1. The presence of the initiator tRNA<sup>Met</sup> is necessary to form an irreversible and active mRNA–tRNA–30S initiation complex. In this complex, mRNA is placed into the 30S mRNA channel (15), and the mRNA–30S interactions are strengthened by the codon–anticodon interaction in the P site. In this case, both AMV and MMLV RT can be used to monitor the mRNA–ribosome interaction. The toeprint is observed at position +16 with AMV RT and usually at position +17 with MMLV RT (Fig. 1).

## 3.1.3. Formation of Regulatory Complexes

1. The regulatory protein traps the ribosome in the inactive pre-initiation complex. The transition from the pre-initiation complex to the active initiation complex is a key step under tight regulation. Structured regions present in the leader regions of mRNAs may delay the mRNA accommodation step. trans-Acting regulators or physical cues (pH, temperature) may prevent this transition from occurring and cause the accumulation of the pre-initiation complex. Ribosomal protein S15 regulates the translation of its own mRNA by stalling the 30S subunit in the pre-initiation complex (7). In this system, S15 stabilizes a pseudoknot structure that is also recognized by the 30S subunit and prevents the conformational changes that are required to accommodate the initiation codon in the P site. The toeprinting assay performed with MMLV RT shows that in the presence of S15, the ternary complex 30S–mRNA–tRNA



## Ribosomal Initiation Complexes Probed by Toeprinting and Effect of trans-Acting

is not formed anymore (toeprint at +17 disappears) while the binary 30S–mRNA complex is stabilized (toeprint at +10 appears) (**Fig. 1a**). These effects can be observed only if the initiator tRNA is added after the repressor protein; therefore, it is essential to mix the reaction compounds in a defined order.

2. Repression through direct masking of the RBS by regulatory RNA. Many translational repressors bind the 5' leader region of the mRNA that overlaps with the RBS. Such regulation is documented for mRNAs targeted by the regulatory RNAIII from *Staphylococcus aureus*. The RNAIII is the intracellular effector of the quorum sensing system that regulates numerous virulence factors in a coordinated way (30). The RNAIII base pairs with target mRNAs to form extended duplexes that hinder the SD sequence of the mRNA. This duplex is also a substrate for the double-strand-specific endoribonuclease III, which initiates the rapid degradation of the repressed mRNA (25). Since the mRNA signal disappears rapidly as soon as the RNAIII is produced, it is quite difficult to prove by in vivo experiments whether the RNAIII inhibits gene expression at the transcriptional or at the post-transcriptional level. On the other hand, a toeprinting assay allows one to monitor in vitro whether the formation of the RNAIII–mRNA duplex is sufficiently stable to prevent the formation of the binary and ternary complexes. A typical example is given for the 5' regulatory region of SA2353 mRNA that is recognized by RNAIII (**Fig. 1b**). *S. aureus* ribosomes provide a toeprint at the expected +16 position, which disappears upon the addition of RNAIII. Concomitantly, a toeprint appears at position +3 that corresponds to the formation of the RNAIII–mRNA duplex. A mutant of RNAIII that is not able to bind rapidly to the mRNA has no effect on ribosome binding. Again, efficient competition between RNAIII and the ribosome is only observed if the initiator tRNA is added after the regulatory RNA.
3. Translation initiation in Gram-positive and Gram-negative bacteria may slightly differ due to some variations in mRNAs and ribosomes. Would these differences influence the regulation of translation initiation? Toeprinting assays may help to answer this question if one can study heterologous complexes. Here we show that *S. aureus* SA2353 mRNA is equally well recognized by *S. aureus* and *E. coli* ribosomes and that the competition between ribosomes and RNAIII does not depend on the nature of the ribosome (**Fig. 1b**). *Thermus thermophilus* ribosomes also bind efficiently to *E. coli* mRNAs (results not shown).

Fechter et al.

### 3.2. Ribosome Preparation

#### 3.2.1. Preparation of the *E. coli* 70S Ribosome and 30S Subunits

1. Prepare a 1-L preculture of *E. coli* MRE600 cells by growing bacteria overnight at 37°C in LB medium using 1 mL of the frozen bacterial stock as a starting culture. Transfer the preculture to 100 L fermentor and continue growth with aeration in the same medium at 37°C. Collect cells at  $OD_{600} = 1$ .
2. Wash 200 g cells in 1 L of buffer AE, and resuspend in 200 mL of the same buffer. Add DNase I to 1 U/mL and PMSF to 1 µg/mL.
3. Disrupt cells in French Press and remove debris by centrifugation for 30 min at  $30,000 \times g$  at 4°C.
4. Collect ribosomes by centrifugation at 45,000 rpm in a Ti 45 rotor for 4 h at 4°C. The ribosome pellet forms two layers. On the bottom of the tube, ribosomes form a glass-type pellet. A loose ribosome pellet is formed on the top of the glass-type pellet. Wash upper layer from the surface of glass-type pellet with buffer AE.
5. Resuspend ribosomes in 50 mL buffer AE and load on the top of high-density Cushion 1E buffer in Ti 45 centrifuge tube (25 mL of ribosome suspension is loaded onto 25 mL of Cushion 1E) and centrifuged at 34,000 rpm for 19 h at 4°C. Ribosomes form pellet on the bottom of centrifuge tube.
6. Resuspend ribosomes in buffer BE to final concentration 50 mg/mL (~ 20 µM) and keep at -80°C (*see Note 3*) in small aliquots or in 1 mL aliquots for 30S subunit preparation.
7. Dialyse extensively the 70S ribosomes (1 mL from 50 mg/mL solution) in Buffer DE at 4°C.
8. Prepare sucrose gradient directly in ultra-clear Beckman tubes (25 × 89 mm) for the SW28 rotor. Mix 18 mL of Buffer S1 and 18 mL of Buffer S2 using a gradient maker in order to obtain a 36 mL 20% (bottom) to 5% (top) sucrose gradient. Six tubes with gradients are usually prepared. The gradients are kept on ice for 2 h prior to ultracentrifugation.
9. Carefully layer the dialysed 70S ribosomes on the top of the tube (no more than 8 mg/per gradient) and perform ultracentrifugation in SW28 rotor at 23,000 rpm ( $70,000 \times g$ ) for 10 h at 4°C.
10. Fractionate the sucrose gradients and check the lighter fractions containing 30S subunits by electrophoresis. Take small aliquots and mix them with the same volume of loading buffer for agarose gel. Load the samples on 1% agarose gel and carry out electrophoresis in running buffer for agarose gel (*see Note 4*).
11. Dialyse the fractions containing pure 30S subunits in Buffer CE at 4°C overnight. Concentrate the 30S subunits using

## Ribosomal Initiation Complexes Probed by Toeprinting and Effect of trans-Acting

YM100 Centricons to a final concentration of 10 mg/mL (~12  $\mu$ M) and store at  $-80^{\circ}\text{C}$  in 5- $\mu$ L aliquots.

3.2.2. Preparation of the *S. aureus* 70S Ribosome and 30S Subunits

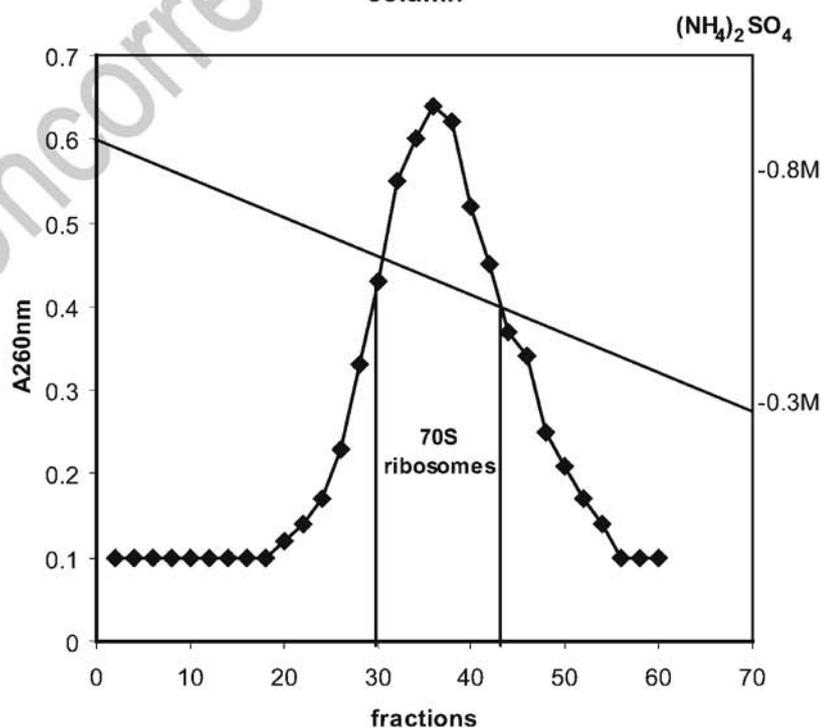
1. Grow preculture of *S. aureus* cells RN6390 in 10 mL of BHI medium at  $37^{\circ}\text{C}$  overnight using a single colony grown on a blood-agar plate. Transfer the preculture to 1 L of BHI medium and continue growth at  $37^{\circ}\text{C}$ . Collect cells at  $\text{OD}_{600} = 5$  and wash them with 10 mM Tris-HCl, pH 7.5.
2. Lyse 5 g of *S. aureus* cells for 30 min at  $37^{\circ}\text{C}$  in 10 mL of TE supplemented with 200  $\mu$ g of lysostaphin and 100 U of DNase I. Pellet cell debris by centrifugation for 20 min at  $30,000 \times g$  at  $4^{\circ}\text{C}$ .
3. Load the supernatants onto 0.5 mL 35% sucrose in buffer AS, in Ti70.1 centrifuge tubes. Pellet the ribosomes for 5 h at  $100,000 \times g$  at  $4^{\circ}\text{C}$ . The ribosomes form a glass-type pellet covered by cell debris. Carefully wash the pellet with 200  $\mu$ L of buffer BS to remove the thin upper layer.
4. Resuspend ribosomes in 2 mL buffer 2 with smooth agitation at  $4^{\circ}\text{C}$  overnight. Load the solution onto 0.5 mL 35% sucrose in buffer BS, in Ti70.1 centrifuge tubes. Centrifuge for 4 h at  $100,000 \times g$  at  $4^{\circ}\text{C}$ .
5. Carefully resuspend the translucent pellets in 7 mL buffer CS and incubate for 30 min at  $37^{\circ}\text{C}$  to dissociate ribosomes that are not tightly formed. Precipitate the ribosomes by addition of 170  $\mu$ L 1 M  $\text{MgCl}_2$  and 2/3 volumes of ethanol. Centrifuge for 30 min at  $3,500 \times g$ , and resuspend the pellets in 1 mL buffer DS.
6. Load the solution (0.5 mL, not more than 5 mg of ribosomes per gradient) on top of the 15 mL 10–40% sucrose gradient prepared with buffer DS, and centrifuge in SW32Ti rotor for 17 h at 19,000 rpm at  $4^{\circ}\text{C}$ .
7. Collect the 70S ribosome fractions and precipitate by 2/3 volumes of ethanol. The ribosomes are then recovered by centrifugation for 30 min at  $3,500 \times g$  at  $4^{\circ}\text{C}$ . Resuspend the pellets in 1 mL of buffer  $\text{CS}_{0.1}$  and incubate for 30 min at  $37^{\circ}\text{C}$  for the dissociation of the ribosomal subunits.
8. Layer the solution (0.5 mL/per gradient, 1 mg) onto the top of the 15 mL 10–30% sucrose gradient prepared with buffer 3 and centrifuge in a SW32Ti rotor for 15 h at 19,000 rpm at  $4^{\circ}\text{C}$ .
9. Collect and pool the 30S subunit fractions. Adjust  $\text{MgCl}_2$  concentration to 20 mM. Precipitate the 30S subunits with 2/3 volumes of ethanol for 2 h on ice, centrifuge for 30 min at  $3,500 \times g$  at  $4^{\circ}\text{C}$ , wash with ethanol 80%, and dry under vacuum. Dissolve the pellet in a minimal volume of buffer 1 and store in small aliquots at  $-80^{\circ}\text{C}$ .

Fechter et al.

3.2.3. Preparation of the *T. thermophilus* Ribosomes

1. Add 1 mL of *T. thermophilus* HB8 stock culture to 1 L of the *T. thermophilus* medium and grow bacteria at 75°C overnight. Continue growth in 100 L fermentor and collect cells at  $OD_{600} = 1$ .
2. Wash cells (100 g) with 1 L of buffer AT, resuspend the pellet in 100 mL of the same buffer. Add DNaseI and PMSF to 1 U/mL and 1 µg/mL, respectively.
3. Disrupt cells in French Press and remove debris by centrifugation for 30 min at  $30,000 \times g$  at 4°C.
4. Layer 29 mL supernatant on top of 7 mL Cushion 1T and centrifuge in SW28 rotor at 27,000 rpm for 20 h at 4°C.
5. Collect a 5-mL fraction from the bottom of the cushion and dilute it three times with buffer BT. Layer 29 mL ribosomes on 7 mL Cushion 2T and centrifuge in SW28 rotor at 27,000 rpm for 40 h at 4°C. Collect a 4-mL fraction from the bottom of cushion and dialyse against buffer CT.
6. Load ribosomes (900 mg) on 200 mL Toyoperl Butyl 650S column equilibrated in buffer CT<sub>0</sub>, and CT<sub>1</sub>. Wash the column by 2 volumes of buffer CT<sub>0.8</sub>. Elute ribosomes by 900 mL reverse gradient of ammonium sulphate (from 80 to 40%) in buffers CT<sub>0</sub>/CT<sub>1</sub> using the flow rate 6 mL/min and the fraction volume 12 mL. The 70S peak is collected as shown in Fig. 2 and ribosomes are dialysed against buffer BT.

### Ribosome chromatography on Toyoperl Butyl 650S column

Fig. 2. Elution profile of *T. thermophilus* ribosomes from Toyoperl Butyl 650S column.

## Ribosomal Initiation Complexes Probed by Toeprinting and Effect of trans-Acting

**3.3. Toeprinting Assays****3.3.1. 5' End Labelling of Oligodeoxyribonucleotide**

1. In order to prepare the radiolabelled DNA primer complementary to the 3' end of the mRNA, incubate 1  $\mu\text{L}$  100  $\mu\text{M}$  oligodeoxyribonucleotide at 37°C for 1 h in the following mixture: 5  $\mu\text{L}$  [ $\gamma$ - $^{32}\text{P}$ ]ATP, 1  $\mu\text{L}$  10 $\times$  T4 PNK buffer, 2  $\mu\text{L}$  sterile water, 1  $\mu\text{L}$  T4 PNK.
2. Add 10  $\mu\text{L}$  DNA loading buffer to the labelling reaction and perform PAGE in a 12%/8 M urea gel to remove the excess of [ $\gamma$ - $^{32}\text{P}$ ]ATP.
3. Cut the DNA band and elute DNA in DNA elution buffer at 4°C under continuous shaking for 12 h. Perform a phenol/chloroform extraction and ethanol precipitation (*see Subheading 3.3.8, step 4*).
4. Wash the pellet with 70% ethanol, dry under vacuum and dissolved radioactive primer in sterile water at concentration  $\sim$ 200,000 cpm/ $\mu\text{L}$ .

**3.3.2. mRNA-Primer Hybrid Formation and Renaturation**

1. Mix 5 pmol of mRNA and 5  $\mu\text{L}$  of labelled primer in Buffer TP1X- (final volume: 20  $\mu\text{L}$ ). Denature mRNA and primer for 1 min at 90°C and cool the mixture for 1 min on ice.
2. Adjust  $\text{MgCl}_2$  concentration to 10 mM with  $\text{MgCl}_2$ 100 solution (2.5  $\mu\text{L}$ ). Incubate the sample at 20°C for 20 min for formation of the mRNA/primer hybrid and mRNA renaturation.

**3.3.3. Ribosome Activation**

1. Reactivate both 70S ribosomes and 30S subunits stored at -80°C (*see Note 5*) by incubation at 37°C for 15 min (the storage buffer should contain 10 mM  $\text{MgCl}_2$ ).
2. After reactivation, dilute the 70S ribosomes 10 times in Buffer TP1X- to obtain final  $\text{MgCl}_2$  concentration of 1 mM. Incubate the ribosomes for 15 min at 37°C to facilitate subunit dissociation.

**3.3.4. Formation of the Binary Complex**

Mix the mRNA-primer mixture (1  $\mu\text{L}$ , 0.25 pmol of mRNA) with 2 pmol of ribosomes or 30S subunits, and 2  $\mu\text{L}$  Buffer TP15X- in the final volume of 10  $\mu\text{L}$ . Adjust the  $\text{MgCl}_2$  concentration to 8 mM with  $\text{MgCl}_2$ 100 solution and form the mRNA-ribosome complex at 37°C for 15 min. Incubate free mRNA as a control.

**3.3.5. Formation of the Ternary Complex**

Add 0.5  $\mu\text{L}$  (40  $\mu\text{M}$ ) of uncharged initiator  $\text{tRNA}^{\text{Met}}$  to the binary mRNA-ribosome complex. The stable ternary mRNA-tRNA-ribosome complex is then formed by incubation at 37°C for 5 min. Incubate free mRNA as a control.

**3.3.6. Formation of Pre-initiation Complex Trapped by Repressor Protein**

1. The *E. coli* S15 protein represses its own synthesis by stabilizing the complex between the *rpsO* mRNA and ribosome at the pre-initiation step. This regulation can be visualized by toeprinting (**Fig. 1a**).
2. Mix the mRNA-DNA mixture (1  $\mu\text{L}$ , 0.25 pmol of mRNA) with 2 pmol of ribosomes or 30S subunits, 10 pmol of purified

Fechter et al.

S15 (*see Note 6*), 2  $\mu\text{L}$  Buffer TP5X- in a final volume of 10  $\mu\text{L}$ . Adjust  $\text{MgCl}_2$  concentration to 8 mM with  $\text{MgCl}_2$ 100 solution, and carry out formation of the mRNA-S15-ribosome complex at 37°C for 15 min.

3. To monitor the effect of ribosomal protein S15 on the formation of the mRNA-tRNA-ribosome complex, add 0.5  $\mu\text{L}$  of uncharged initiator  $\text{tRNA}^{\text{Met}}$  to the pre-formed S15-mRNA-ribosome complex and incubate at 37°C for 5 min. As shown in **Fig. 1a**, if tRNA is added after the repressor protein, the ternary mRNA-tRNA-ribosome complex cannot be formed. Conversely, if tRNA is added before the repressor protein, the stalled mRNA-S15-ribosome complex cannot be formed (8).
4. Incubation controls should be done under strictly identical conditions using the mRNA which is either free or bound to the repressor protein.

### 3.3.7. Effect of trans-Acting Antisense RNA on the Formation of the Ribosomal Initiation Complex

1. Toeprinting assays can be used to visualize formation of the regulatory mRNA-antisense RNA complexes that cause inhibition of the mRNA-ribosome binding (**Fig. 1b**).
2. Refold the mRNA-DNA primer hybrid as described in **Subheading 3.3.2**. In parallel, denature and refold the regulatory RNA (here *S. aureus* RNAIII) as described in **Subheading 3.3.2**. Mix the mRNA-DNA mixture (1  $\mu\text{L}$ , 0.25 pmol of mRNA) with 2 pmol of ribosomes, 1  $\mu\text{L}$  of the regulatory RNA (perform several reactions with regulatory RNA concentration in the range of 20–250 nM), 2  $\mu\text{L}$  of Buffer TP5X- in a final volume of 10  $\mu\text{L}$ . Adjust  $\text{MgCl}_2$  concentration to 8 mM with  $\text{MgCl}_2$ 100 solution and carry out the binding reaction at 37°C for 15 min.
3. Add uncharged initiator  $\text{tRNA}^{\text{Met}}$  (0.5  $\mu\text{L}$ , 40  $\mu\text{M}$ ) to the binary complex and incubate for 5 min at 37°C.
4. Incubation controls should be done as in **Subheading 3.3.6, step 4**.

### 3.3.8. Reverse Transcriptase Extension and Sequencing

1. Just prior reverse transcription, increase  $\text{MgCl}_2$  concentration in the samples to 15 mM with the  $\text{MgCl}_2$ 100 solution. This step is essential for stabilizing the complexes formed with the 70S ribosomes from *T. thermophilus*.
2. Add to all samples (incubation controls and reactions) 1.5  $\mu\text{L}$  dNTP500 mix (50  $\mu\text{M}$  final concentration), 1  $\mu\text{L}$  Buffer TP5X+, 1.5  $\mu\text{L}$  sterile water and 1  $\mu\text{L}$  MMLV RT (3 U) or 1  $\mu\text{L}$  AMV RT (2 U). The extension reactions (final volume 15  $\mu\text{L}$ ) are incubated for 15 min at 37°C. At high concentration, MMLV enzyme is less sensitive to secondary structures which might be important for the gene expression control (**Fig. 1a**).

## Ribosomal Initiation Complexes Probed by Toeprinting and Effect of trans-Acting

3. To help ethanol precipitation, add 1  $\mu\text{L}$  total tRNA to the samples that do not contain tRNA or ribosome.
4. Carry out phenol extraction of all samples by addition 110  $\mu\text{L}$  of 300 mM Na-acetate and 120  $\mu\text{L}$  phenol/chloroform/isoamyl alcohol (24:24:1) mixture. Vortex the samples for at least 1 min at 20°C, and centrifuge at 15,000  $\times g$  for 3 min at 4°C. Carefully take the aqueous upper phase, and add 110  $\mu\text{L}$  of chloroform/isoamyl alcohol mixture. Mix the samples vigorously for 1 min at 20°C, and centrifuge for 3 min at 15,000  $\times g$  at 4°C (*see* Note 7).
5. Take the upper phase and add 2.5 volumes of cold ethanol. Mix solutions and keep them in dry ice for 30 min. Centrifuge the samples at 4°C at 15,000  $\times g$  to precipitate the labelled DNA fragments. Wash the pellets twice with 70% cold ethanol, vacuum dry and dissolve in 10  $\mu\text{L}$  of DNA loading buffer.

**3.4. Gel Fractionation**

1. The RT pauses are identified by running a sequencing reaction in parallel with toeprinting samples (*see* Chapter “Structural Probing of RNA Thermosensors”). The primer extension step of the sequencing reactions is performed with AMV reverse transcriptase (*see* Subheading 3.3.8, step 2) in the presence of one dideoxynucleotide ddXTP (2.5  $\mu\text{M}$ ), the corresponding deoxynucleotide dXTP (25  $\mu\text{M}$ ) and three other deoxynucleotides (100  $\mu\text{M}$ ).
2. After RT primer extension, heat all samples at 90°C for 3 min and centrifuged briefly.
3. Load 3  $\mu\text{L}$  per well on a 8% polyacrylamide-(0.4% bis)/8 M urea slab gel and perform PAGE in 1 $\times$  TBE. Store the remaining samples at -20°C. The migration time should be adjusted to the size of the analyzed fragments, assuming that during 8% PAGE, xylene cyanol, and bromophenol blue migrate as 81 nucleotide and 19 nucleotide long DNA, respectively. After PAGE, transfer the gel on a dispensable thin support (i.e. an old autoradiograph), and expose at -80°C with X-ray film and intensifying screen overnight.

**4. Notes**

1. For safety reason, manipulation with pathogenic bacteria such as *S. aureus* must be carried out in a Biosafety Level 2 laboratory (according to the country rule) that contains equipment required for cell growth, lysis, and centrifugation of bacteria.

Fechter et al.

2. For convenience, prepare a 2× stock solution of Buffer DE and mix it with 50% sucrose solution to prepare Buffer S1 and Buffer S2.
3. Ribosomes and ribosome subunits are flash frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .
4. Electrophoresis of ribosomes on 1% agarose gel allows the visualization of 23S rRNA (component of 50S subunits) and 16S rRNA (component of 30S subunits) which should appear as distinct bands.
5. No more than three cycles of flash freezing–defreezing should be performed for the 30S subunits.
6. *E. coli* S15 was purified according to ref. 8, stored at  $4^{\circ}\text{C}$  and reactivated for 15 min at  $37^{\circ}\text{C}$  in a buffer containing 20 mM  $\text{MgCl}_2$ , 270 mM KCl, 50 mM Tris–HCl, pH 7.5, 3 mM DTT, 0.02 mg/mL bovine serum albumin.
7. After primer extension, phenol extraction followed by ethanol precipitation is not usually necessary. Therefore, 10  $\mu\text{L}$  of DNA loading Buffer can be directly added to all samples. The assays are then heated for 3 min at  $90^{\circ}\text{C}$  and 6  $\mu\text{L}$  are loaded per well on 8% gel. However, the migration of the labelled cDNA fragments might be sometimes perturbed (*see* Fig. 1b, left part). Phenol extraction helps to get a better resolution on the gel.

---

## Acknowledgements

We are grateful to Anne-Catherine Helfer for the toeprinting assays performed on *E. coli rpsO* mRNA. This work was supported by financial support from the Centre National de la Recherche Scientifique (UPR 9002 CNRS), from the University Louis Pasteur of Strasbourg, from the Ministère de la Recherche (ANR05-MIIME, ANR07-BLANC), from the European Community (FOSRAK, EC005120; BacRNA EC018618).

## References

1. Serganov, A. and Patel, D. J. (2007). Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat. Rev. Genet.* **8**, 776–790.
2. Romby, P. and Springer, M. (2007). Translational control in prokaryotes. In *Translational Control in Biology and Medicine* (Hershey, J, Sonnenberg, N, Matthews, M, eds), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 807–832.
3. Narberhaus, F., Waldminghaus, T. and Chowdhury, S. (2006). RNA thermometers. *FEMS Microbiol. Rev.* **30**, 3–16.
4. Coppins, R. L., Hall, K. B. and Groisman, E. A. (2007). The intricate world of riboswitches. *Curr. Opin. Microbiol.* **10**, 176–181.
5. Romby, P. and Springer, M. (2003). Bacterial translational control at atomic resolution. *Trends Genet.* **19**, 155–161.

## Ribosomal Initiation Complexes Probed by Toeprinting and Effect of trans-Acting

- Schlxax, P. J. and Worhunsky, D. J. (2003). Translational repression mechanisms in prokaryotes. *Mol. Microbiol.* **48**, 1157–1169.
- Ehresmann, C., Ehresmann, B., Ennifar, E., Dumas, P., Garber, M., Mathy, N. et al. (2004). Molecular mimicry in translational regulation: the case of ribosomal protein S15. *RNA Biol.* **1**, 66–73.
- Marzi, S., Myasnikov, A. G., Serganov, A., Ehresmann, C., Romby, P., Yusupov, M. et al. (2007). Structured mRNAs regulate translation initiation by binding to the platform of the ribosome. *Cell* **130**, 1019–1031.
- Darfeuille, F., Unoson, C., Vogel, J. and Wagner, E. G. (2007). An antisense RNA inhibits translation by competing with standby ribosomes. *Mol. Cell* **26**, 381–392.
- Vogel, J. and Wagner, E. G. (2007). Target identification of small noncoding RNAs in bacteria. *Curr. Opin. Microbiol.* **10**, 262–270.
- Hartz, D., McPheeters, D. S., Traut, R. and Gold, L. (1988). Extension inhibition analysis of translation initiation complexes. *Methods Enzymol.* **164**, 419–425.
- Hartz, D., McPheeters, D. S. and Gold, L. (1989). Selection of the initiator tRNA by *Escherichia coli* initiation factors. *Genes Dev.* **3**, 1899–1912.
- Huttenhofer, A. and Noller, H. F. (1994). Footprinting mRNA–ribosome complexes with chemical probes. *EMBO J.* **13**, 3892–3901.
- Sacerdot, C., Caillet, J., Graffe, M., Eyermann, F., Ehresmann, B., Ehresmann, C. et al. (1998). The *Escherichia coli* threonyl-tRNA synthetase gene contains a split ribosomal binding site interrupted by a hairpin structure that is essential for autoregulation. *Mol. Microbiol.* **29**, 1077–1090.
- Yusupova, G. Z., Yusupov, M. M., Cate, J. H. and Noller, H. F. (2001). The path of messenger RNA through the ribosome. *Cell* **106**, 233–241.
- Jenner, L., Romby, P., Rees, B., Schulze-Briese, C., Springer, M., Ehresmann, C. et al. (2005). Translational operator of mRNA on the ribosome: how repressor proteins exclude ribosome binding. *Science* **308**, 120–123.
- Korostelev, A., Trakhanov, S., Asahara, H., Laurberg, M., Lancaster, L. and Noller, H. F. (2007). Interactions and dynamics of the Shine Dalgarno helix in the 70S ribosome. *Proc. Natl Acad. Sci. U. S. A.* **104**, 16840–16843.
- Hartz, D., McPheeters, D. S., Green, L. and Gold, L. (1991). Detection of *Escherichia coli* ribosome binding at translation initiation sites in the absence of tRNA. *J. Mol. Biol.* **218**, 99–105.
19. Philippe, C., Eyermann, F., Benard, L., Portier, C., Ehresmann, B. and Ehresmann, C. (1993). Ribosomal protein S15 from *Escherichia coli* modulates its own translation by trapping the ribosome on the mRNA initiation loading site. *Proc. Natl Acad. Sci. U. S. A.* **90**, 4394–4398.
20. Ringquist, S., MacDonald, M., Gibson, T. and Gold, L. (1993). Nature of the ribosomal mRNA track: analysis of ribosome-binding sites containing different sequences and secondary structures. *Biochemistry* **32**, 10254–10262.
21. Qin, Y., Polacek, N., Vesper, O., Staub, E., Einfeldt, E., Wilson, D. N. et al. (2006). The highly conserved LepA is a ribosomal elongation factor that back-translocates the ribosome. *Cell* **127**, 721–733.
22. Waldminghaus, T., Heidrich, N., Brantl, S. and Narberhaus, F. (2007). FourU: a novel type of RNA thermometer in Salmonella. *Mol. Microbiol.* **65**, 413–424.
23. Brunel, C., Romby, P., Moine, H., Caillet, J., Grunberg-Manago, M., Springer, M. et al. (1993). Translational regulation of the *Escherichia coli* threonyl-tRNA synthetase gene: structural and functional importance of the thrS operator domains. *Biochimie* **75**, 1167–1119.
24. Sharma, C. M., Darfeuille, F., Plantinga, T. H. and 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.
25. Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M. et al. (2007). *Staphylococcus aureus* RNAlII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev.* **21**, 1353–1366.
26. Milligan, J. F. and Uhlenbeck, O. C. (1989). Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.* **180**, 51–62.
27. Romaniuk, P. J., de Stevenson, I. L. and Wong, H. H. (1987). Defining the binding site of Xenopus transcription factor IIIA on 5S RNA using truncated and chimeric 5S RNA molecules. *Nucleic Acids Res.* **15**, 2737–2755.
28. Jahn, M. J., Jahn, D., Kumar, A. M. and Soll, D. (1991). Mono Q chromatography permits recycling of DNA template and purification of RNA transcripts after T7 RNA polymerase reaction. *Nucleic Acids Res.* **19**, 2786.
29. Marzi, S., Fechter, P., Chevalier, C., Romby, P. and Geissmann, T. RNA switches regulate initiation of translation in bacteria. *Biol. Chem.* **389**, 585–598.
30. Novick, R. P. and Jiang, D. (2003). The staphylococcal saeRS system coordinates environmental signals with agr quorum sensing. *Microbiology* **149**, 2709–2717.

---

## **Annexe 4: Posters**

**Poster 1: The 3' domain of *Staphylococcus aureus* RNAIII, a virulence-specific regulatory domain.** Présenté au “5<sup>th</sup> Colmar Scientific Symposium” en 2007

**Poster 2: L'ARNIII de *Staphylococcus aureus* réprime le facteur coagulase par un mécanisme de type antisens.** Présenté au sifrARN "meeting" ainsi qu'au 35e Forum des Jeunes Chercheurs à Montpellier en 2008.

Ces posters résument les résultats obtenus concernant l'ARNIII et l'identification de ses cibles. Ces deux posters ont été récompensés par le prix du meilleur poster.



# The 3'-domain of *Staphylococcus aureus* RNAIII, a virulence-specific regulatory domain

Clément Chevalier<sup>1</sup>, Sandrine Boisset<sup>2</sup>, Pierre Fechter<sup>1</sup>, Christine Gaspin<sup>3</sup>, François Vandenesch<sup>2</sup>, Pascale Romby<sup>1</sup> & Thomas Geissmann<sup>1</sup>

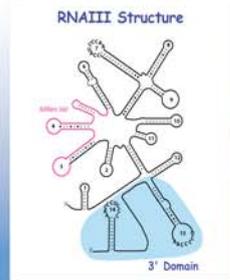


<sup>1</sup>Architecture et réactivité de l'ARN, Université Louis Pasteur, IBMC-CNRS, 15 rue René Descartes, 67084 Strasbourg, France.  
<sup>2</sup>Centre National de Référence des Staphylocoques, Université Lyon 1, 69008 Lyon, France; INSERM, E0230, 69008 Lyon, France  
<sup>3</sup>Unité de Biométrie et Intelligence Artificielle, INRA, Chemin de Borde-Rouge, 31326 Castanet-Tolosan, France

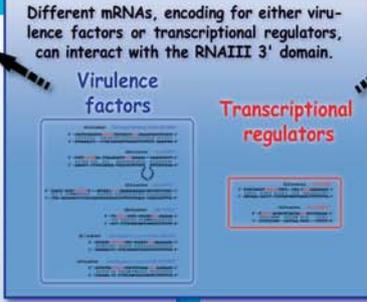
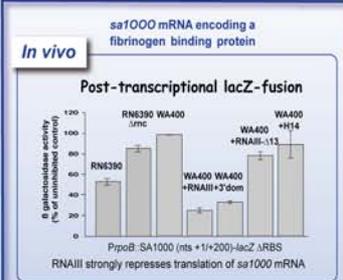


*Staphylococcus aureus* is a versatile pathogen, whose pathogenicity depends largely on the expression of virulence factors (exoproteins and surface proteins). Their expression depends on a complex network of regulatory genes, among them a two-component signal transduction system, the global regulon *agr*. The effector of the *S. aureus agr* system is the RNAIII. The mechanism by which RNAIII controls the expression of its target genes is still largely unknown.

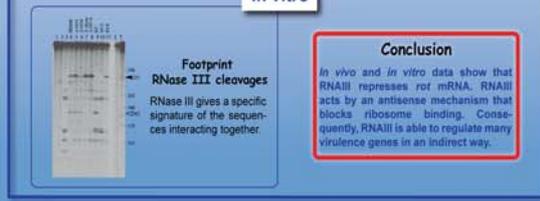
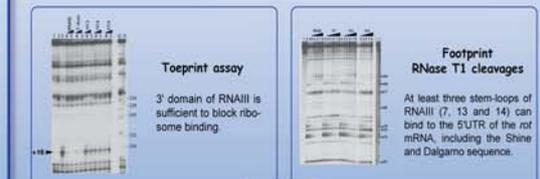
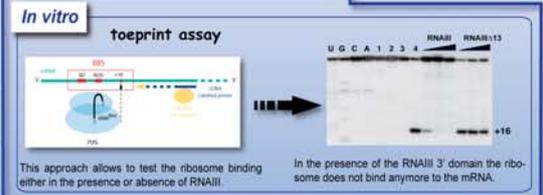
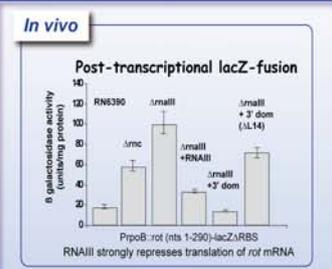
Previously, we have shown that the 3'-domain (nts 388-514) of this RNA was sufficient to exert a regulatory activity on the protein A, a virulence gene expressed at the surface of the bacteria. The aim of this study was to test whether this domain could form a intrasc regulatory domain, looking for others potential targets.



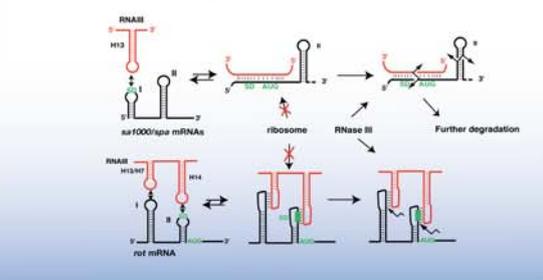
## RNAIII represses multiple virulence factors



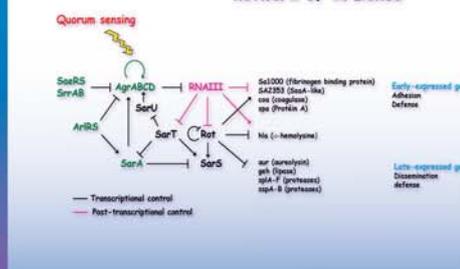
## RNAIII represses the transcriptional regulator Rot (repressor of toxin)



## RNAIII binds to its target mRNAs via one or two loop-loop interactions.



## RNAIII, a keyplayer of a regulatory network of virulence





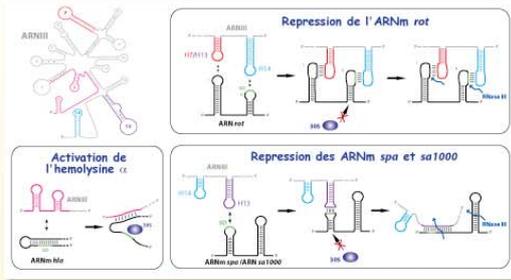
# L'ARNIII de *Staphylococcus aureus* réprime le facteur coagulase par un mécanisme de type antisens

Clément Chevalier<sup>1</sup>, Sandrine Boisset<sup>2</sup>, Pierre Fechter<sup>1</sup>, Benoit Masquida<sup>1</sup>, Thomas Geissmann<sup>1</sup>, Christine Gaspin<sup>1</sup>, François Vandenesch<sup>3</sup> & Pascale Romyb<sup>1</sup>

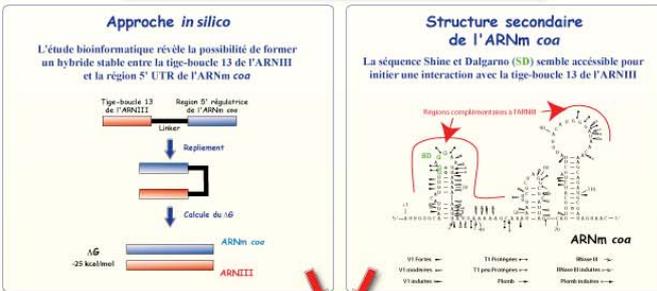
<sup>1</sup>Architecture et réactivité de l'ARN, Université Louis Pasteur, IBMC-CNRS, 15 rue René Descartes, 67084 Strasbourg, France.  
<sup>2</sup>Centre National de Référence des Staphylocoques, Université Lyon 1, 69008 Lyon, France; INSERM, E0230, 69008 Lyon, France  
<sup>3</sup>Unité de Biométrie et Intelligence Artificielle, INRA, Chemin de Borde-Rouge, 31326 Castanet-Tolosan, France



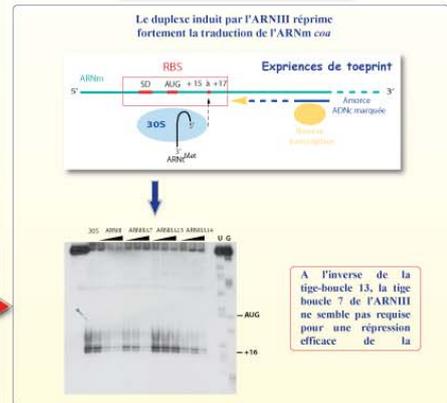
*Staphylococcus aureus* est une bactérie gram positive dont le pouvoir pathogène est lié à l'expression de nombreux facteurs de virulence. Ces facteurs codent généralement pour des protéines de surface et des exoprotéines, et leur expression est finement contrôlée par un réseau de régulation complexe. Parmi les régulateurs, on distingue l'opéron *agr*, un système à deux composants dont l'effecteur intracellulaire est l'ARNIII. Cet ARN régule de manière temporelle plus d'une centaine de gènes et agit principalement par des mécanismes de type antisens. L'ARNIII réprime d'une part toute une classe de facteurs de virulence induit lors des premières phases de l'infection (protéine A, facteur SA1000, antigène sécrété SsaA) et d'autre part le facteur de transcription Rot. Les mécanismes de reconnaissance varient en fonction des ARNm cibles. Dans les deux cas, les hybrides ARN sont suffisamment stables pour bloquer la fixation du ribosome et induire une dégradation rapide initiée par l'endoribonucléase III. Au cours de ce travail nous avons validé une autre cible codant pour un facteur majeur de virulence, le facteur coagulase.



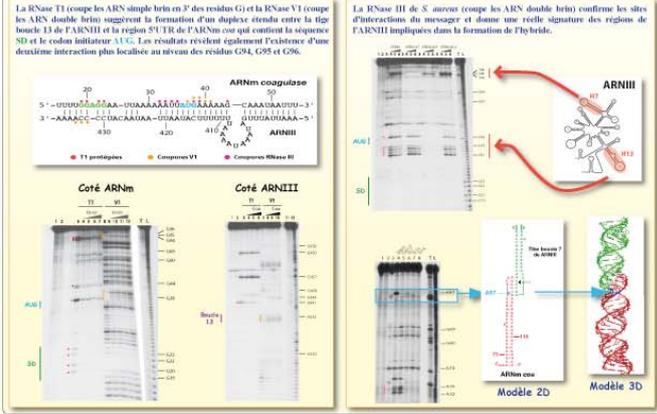
## Formation d'un hybride ARNIII/ARNm *coa*



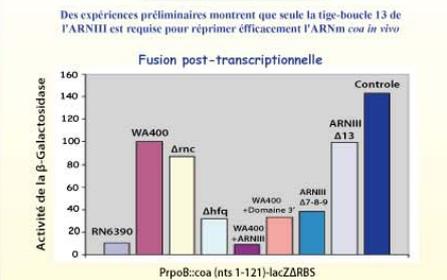
## Effet de l'ARNIII sur la fixation du ribosome



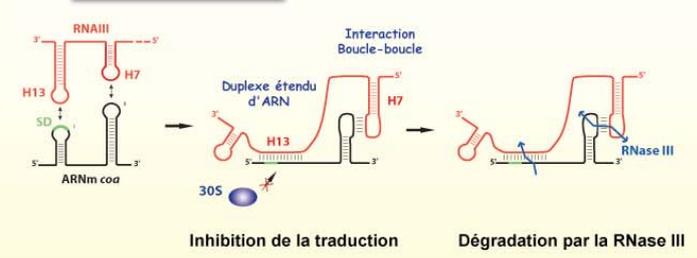
## Deux domaines de l'ARNIII interagissent avec l'ARNm *coa*



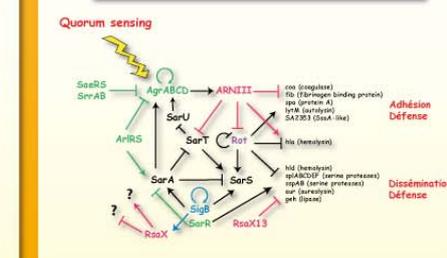
## Validation in vivo



## Modèle de régulation



## Un réseau de régulation complexe



---

## Annexe 5: Communications scientifiques

**26-31 juin 2009, RNA meeting, Madison (Etats-Unis).** Communication orale: Novel regulatory ARNs in *Staphylococcus aureus*: **Chevalier C**, Geissmann T, Cros MJ, Boisset S, Fechter P, Noirot C, Schrenzel J, François P, Vandenesch F, Gaspin C & Romby P

**15-16 septembre 2008, European meeting, “BACRNAs”, Strasbourg (France).** Communication orale: *Staphylococcus aureus* RNAIII represses *lytM* and *coa* mRNAs: **Chevalier C**, Boisset S, Fechter P, Masquida B, Geissmann T, Gaspin C, Vandenesch F & Romby P

**8-12 septembre 2008, 35e Forum des Jeunes Chercheurs et sifrARN meeting, Montpellier (France).** Poster (**Prix du meilleur poster**): L'ARNIII de *Staphylococcus aureus* réprime le facteur coagulase par un mécanisme de type antisens : **Chevalier C**, Boisset S, Fechter P, Masquida B, Geissmann T, Gaspin C, Vandenesch F & Romby P

**8-9 novembre 2007, International meeting “5<sup>th</sup> Colmar Scientific Symposium”, Colmar (France).** Communication orale et Poster (**Prix du meilleur poster**): The 3' domain of *Staphylococcus aureus* RNAIII, a virulence-specific regulatory domain: **Chevalier C**, Boisset S, Fechter P, Gaspin C, Vandenesch F, Romby P & Geissmann T

**15-17 octobre 2007, 34<sup>e</sup> Forum des Jeunes Chercheurs, Ile des Embiez (France).** Communication orale : Rôles multiples de l'ARNIII dans la régulation des gènes de virulence chez *Staphylococcus aureus*. **Chevalier C**, Geissmann T, Boisset S, Huntzinger E, Fechter P, Gaspin C, Vandenesch F & Romby P

**26-27 octobre 2006, European meeting “Rencontre Bâle-Strasbourg”, Mittelwhir (France).** Poster : The 3'-domain of *Staphylococcus aureus* RNAIII, a virulence-specific regulatory domain. Geissmann T, Fechter P, **Chevalier C**, Huntzinger E, Boisset S, Benito Y, Jacquier A, Gaspin C, Vandenesch F & Romby P