



HAL
open science

Identification et caractérisation d'ARN régulateurs chez Escherichia coli régulés par RpoS

Demol Julien

► **To cite this version:**

Demol Julien. Identification et caractérisation d'ARN régulateurs chez Escherichia coli régulés par RpoS. Génomique, Transcriptomique et Protéomique [q-bio.GN]. 2014. hal-01979005

HAL Id: hal-01979005

<https://ephe.hal.science/hal-01979005>

Submitted on 16 Jan 2019

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

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

Public Domain

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE

ÉCOLE PRATIQUE DES HAUTES ÉTUDES

Sciences de la Vie et de la Terre

MÉMOIRE

Présenté

Par

DEMOL Julien

Pour l'obtention du diplôme de l'École Pratique des Hautes Études,

Soutenu le 10 juillet 2014

Identification et caractérisation d'ARN régulateurs chez *Escherichia coli* régulés par RpoS

Devant le jury suivant :

Ronot Xavier – Président

Kay Elisabeth – Rapporteur

Dorel Corinne – Examineur

Lacour Stéphan – Tuteur scientifique

Lasbleiz Christelle – Tuteur pédagogique

Mémoire préparé sous la direction de :

LACOUR Stéphan

Laboratoire CNRS-UJF UMR 5163, Laboratoire Adaptation et Pathogénie des Microorganismes

Équipe Contrôle de l'expression génique

Directeur : *Hans Geiselman*

Et de

LASBLEIZ Christelle

Laboratoire UMR Inserm U710-UM2-EPHE, Mécanisme Moléculaire dans les Démences Neurodégénératives

Directeur : *Jean-Michel Verdier*

Table des matières

I. INTRODUCTION.....	1
A. Escherichia coli.....	1
1. Généralités.....	1
2. Propriété du modèle bactérien.....	2
B. Régulation de l'expression génique.....	3
1. Le génome.....	3
2. L'ARN polymérase.....	3
3. Du gène à la protéine.....	5
4. Autres catégories d'ARN.....	6
C. Les ARN régulateurs.....	7
1. Généralités.....	7
2. Les ARN non codant.....	7
3. Hfq, protéine chaperonne des ARNnc.....	11
4. A la recherche des ARNnc.....	11
D. La protéine RpoS, facteur σ^{38} de l'ARN polymérase.....	12
E. Régulation de l'expression de protéines d'adhérences : les curli, par le facteur sigma RpoS et les ARNnc	13
1. Régulation transcriptionnelle des opérons <i>csg</i>	15
2. Régulation traductionnelle de l'ARNm <i>csgDEFG</i>	16
F. Objectifs.....	17
1. Identification et caractérisation de nouveaux ARNnc régulés par RpoS.....	17
2. Etude globale du réseau régulateur des opérons <i>csg</i>	18
II. MATÉRIELS ET MÉTHODES.....	19
A. Milieux, souches, plasmides.....	19
1. Les milieux de culture.....	19
2. Souches bactériennes.....	20
3. Fusion transcriptionnelle et traductionnelle.....	21
4. Autres plasmides.....	22
5. Lecture d'absorbance et de fluorescence en microplaque.....	22
B. Biologie moléculaire.....	23
1. Amplification d'ADN par PCR.....	23
2. Analyse par gel d'électrophorèse.....	24
3. Purification d'ADN synthétisé et d'ADN plasmidique.....	24
4. Extraction d'ARN totaux au phénol chaud.....	24
5. Northern Blot.....	25
6. Puce à ADnc.....	26
7. qRT-PCR.....	27
8. Transcription <i>in vitro</i>	27
9. Extension d'amorce.....	27
C. Transformation bactérienne et clonage.....	28
1. Préparation des cellules compétentes et transformation des bactéries.....	28
2. Transduction bactérienne.....	29
3. Recombinaison plasmidique et chromosomique.....	29
4. Clonage par restriction – ligation.....	30
D. Outils informatiques.....	30
III. Résultats.....	31
A. Identification d'ARNnc régulés par RpoS par étude transcriptomique.....	31
1. Introduction et motivations.....	31
2. Etude des données de puces Affymetrix.....	32
3. RNA-seq, vers l'identification d'ARNnc régulés par RpoS.....	33

4.	Détection de site de liaison de RpoS par ChIP-seq.....	41
5.	Identification d'un nouvel ARNnc : Djs.....	46
6.	Conclusions du chapitre	51
B.	Etude de la régulation de l'expression des curli par RpoS et les ARNnc.....	52
1.	Introduction.....	52
2.	Les fusions transcriptionnelles et traductionnelles.....	55
3.	Régulation de la transcription des gènes <i>csg</i>	57
4.	Régulation de la traduction de CsgD	60
5.	Effet des ARNnc sur la transcription des gènes <i>csgD</i> et <i>csgB</i>	63
6.	Régulation de la transcription des ARNnc modulant la traduction de CsgD et RpoS	68
7.	Stabilité de l'ARNnc McaS	75
8.	Conclusions du chapitre	79
IV.	Conclusions et Perspectives	80
A.	Identification d'ARNnc régulés par RpoS	80
1.	Les ARNnc régulés par RpoS.....	80
2.	Régulation négative des ARNnc par RpoS.....	80
3.	Un nouvel ARNnc identifié : Djs	81
B.	Les acteurs majeurs de la régulation de l'expression des curli.....	82
1.	La transcription des gènes <i>csg</i> est activée par l'ARNnc GcvB	82
2.	Répression transcriptionnelle et traductionnelle de CsgD par l'ARNnc McaS.....	83
3.	Modulation de la traduction du facteur OmpR.....	83
4.	Comparaison des méthodologies entre études	83
5.	Le rôle versatile et de thermorégulateur du facteur RpoS	84
6.	RpoS et l'ARNnc CsgD déterminent l'accumulation de l'ARNnc McaS	84
7.	Un nouvel ARNnc, RydC, régulateur traductionnel de CsgD.....	85
8.	Intuition personnelle	85
V.	Bibliographie.....	86
	Résumé.....	1

Liste des figures

Figure 1 : Schéma représentant la croissance bactérienne.	2
Figure 2 : Les facteurs sigma et leur distribution en fonction de la phase de croissance.....	4
Figure 3 : Mécanisme de transcription à la traduction.	6
Figure 4 : Différentes interactions ARN régulateurs/protéines.	8
Figure 5 : Mécanisme d'action des ARNnc <i>cis</i> -encodés.	9
Figure 6 : Mode d'action des ARNnc <i>trans</i> -encodés.	10
Figure 7 : Libération du RBS de l'ARNm RpoS par l'ARNnc DsrA.....	12
Figure 8: Région intergénique entre les opérons <i>csgDEFG</i> et <i>csgBAC</i> et sites de fixation de leurs régulateurs transcriptionnelles ou traductionnelles.....	14
Figure 9 : Réseau de régulation contrôlant la différenciation entre cellule adhérente ou motile.	16
Figure 10 : Recouvrement de la région intergénique par les sondes Affymetrix.....	33
Figure 11 : Mini-puce à ADNc comparant les expressions différentielles des ARN putatifs et connus entre une souche sauvage (WT) et <i>ΔrpoS</i>	37
Figure 12 : Courbe d'amplification de l'ARNnc GlmY par qPCR après RT.	40
Figure 13 : Alignement des régions promotrices des gènes codant les ARNnc OmrA et OmrB	42
Figure 14 : Profils transcriptionnels des ARNnc ciblés par le CHIP-seq.	43
Figure 15 : Northern Blot et quantification de l'accumulation des ARNnc OmrA, RyeA et SibC.	44
Figure 16 : Homologie de séquence des 15 copies de Djs.	46
Figure 17 : Distribution des différentes copies de Djs sur le chromosome d' <i>E. coli</i> K-12	47
Figure 18 : La surexpression de l'ARN putatif Djs affecte la liaison au Rouge Congo.	49
Figure 19 : Northern Blot visant à détecter l'ARN putatif Djs	50
Figure 20 : Extension d'amorces visant à démontrer l'existence de Djs.....	51
Figure 21: Phénotype de liaison au Rouge Congo à 28°C et 37°C.....	53
Figure 22: Expression des opérons <i>csg</i> à 28°C et 37°C.....	53
Figure 23: Réseau de régulation contrôlant l'expression des curli chez <i>E. coli</i>	54
Figure 24: Fusions transcriptionnelles et traductionnelles des gènes <i>csgD</i> et <i>csgB</i>	55
Figure 25: Expression des gènes <i>csgB</i> et <i>csgD</i> dans le mutant conditionnel <i>ompR234</i> comparé à la souche sauvage.	56
Figure 26: Régulation de la transcription du gène <i>csgD</i>	57
Figure 27 : Régulation de la transcription du gène <i>csgB</i>	59
Figure 28 : Activité des fusions traductionnelles <i>csgD::gfp</i> avec et sans 5'UTR.	61
Figure 29 : Bilan des régulations transcriptionnelles et traductionnelles observées.	62
Figure 30 : Transcription des gènes <i>csgD</i> et <i>csgB</i> : effet des mutations des ARNnc.....	63
Figure 31 : Liaison au Rouge Congo des dérivés de la souche BW25113.....	64
Figure 32 : La régulation transcriptionnelle et traductionnelle est plus visible avec la fusion <i>csgD</i> IG.66	
Figure 33 : État des connaissances sur les interactions entre membres du réseau de régulation de l'expression des curli.	67
Figure 34 : Transcription des ARNnc GcvB et DsrA.	69
Figure 35 : Transcription et accumulation de l'ARNnc RprA.	70
Figure 36 : Transcription et accumulation de l'ARNnc OmrA.....	71
Figure 37 : Transcription et accumulation de l'ARNnc McaS.	72
Figure 38 : Bilan sur la régulation transcriptionnelle des ARNnc modulant la traduction de CsgD.....	74
Figure 39 : Complémentarité entre les ARNnc McaS et OmrA.	75
Figure 40 : Détermination de la demi-vie de l'ARNnc McaS.	76
Figure 41 : Comparaison de l'accumulation de l'ARNnc McaS et de la région 5'UTR <i>csgD</i>	77
Figure 42 : Structure secondaire de la région 5'UTR de l'ARNm de CsgD.....	78
Figure 43 : Appariement de la partie 5'UTR de l'ARNm CsgD et de l'ARNnc McaS	78

Liste des tableaux

Tableau 1 : Liste des sondes biotinyllées pour les Northern Blot	26
Tableau 2 : Amorces pour les gènes ciblés en qRT-PCR.....	27
Tableau 3 : Analyse des ARN séquencés	34
Tableau 4: ARNnc connus et putatifs fortement régulés par RpoS en milieu minimum M9.....	35
Tableau 5 : ARN ciblés par la puce à ADNc.....	36
Tableau 6 : ARNnc détectés par les trois logiciels d'analyse du RNA-seq.....	39
Tableau 7 : Distribution du nombre de copies de Djs au sein des souches bactériennes de la famille des <i>Enterobacteriaceae</i>	48

Liste des abréviations

°C	Température Celsius
µg	Microgramme(s)
µl	Microlitre(s)
µM	Micromolaire
A.A	Acide Aminé
ADN	Acide DésoxyriboNucléique
ADNc	ADN complémentaire
ARN	Acide Ribonucléique
ARNm	ARN messenger
ARNnc	ARN non codant
ARNr	ARN ribosomale
ARNt	ARN transfert
b	Base(s)
DO	Densité Optique =(OD)
EDTA	Acide ethylène Diamine Tétracétique
IG	région intergénique
IPTG	IsoPropyl-β-D-ThioGalactoside
Kb	Kilo bases
LB	Luria Both
M	Molaire
M9	Milieu Minimum M9
mg	Milligramme(s)
min	Minute(s)
mL	Millilitre(s)
mM	Millimolaire(s)
ng	Nanogramme(s)
nM	Nanomolaire(s)
nt(s)	Nucléotide(s)
ORF	= (Open Reading Frame), CDS , cadre de lecture
pb	Paire de bases
pM	Picomolaire
RBS	Site de fixation du ribosome (Reading Binding Site)
RC	Rouge Congo
Tm°	Température d'hybridation des amorces
UTR	= (UnTranslated Region), région non traduite
V	Volt
Δ	Delta = délété

I. INTRODUCTION

A. *Escherichia coli*

1. Généralités

En 1885, Theodor Escherich, un bactériologue et pédiatre, identifie et caractérise lors d'une étude des diarrhées infantiles, une bactérie commensale issue des fèces. En 1919, celle-ci prit le nom *Escherichia coli* en son honneur. Cette *Gamma-proteobacterieae* appartenant à l'ordre des *Enterobacteriales*, de la famille des *Enterobacteriaceae* et du genre des *Escherichia* à une forme de bâtonnet, et une taille de l'ordre du micromètre. Longtemps supposée majoritaire au sein du microbiote intestinal, elle est en réalité largement minoritaire. *Escherichia coli* est majoritairement une bactérie commensale et non pathogène, néanmoins il existe de nombreuses souches d'*Escherichia coli* pathogènes dont les mécanismes de la pathogénicité sont multiples (Lukjancenko et al., 2010).

Cette bactérie facilement cultivable a été l'un des premiers modèles d'étude chez les bactéries à Gram négatif. Elle possède un temps de division de l'ordre de la vingtaine de minutes en milieu riche à 37°C, faisant partie des bactéries dont le taux de croissance est élevé, ce qui en fait un organisme modèle de choix. Entre 1988 et 1998, la souche *E. coli* MG1655 fut l'une des premières souches bactériennes dont le génome a été intégralement séquencé et rentré dans les bases de données (Kraev, 1988; Kroger, 1989; Kroger and Wahl, 1996, 1997, 1998; Wahl et al., 1994). Cet organisme comporte environ 4,6 millions de paires de bases codant environ 4300 protéines. L'accès à cette séquence génomique a permis l'étude et l'identification des gènes et des protéines leur étant associées, mais également de décrire les mécanismes de régulation génique et métabolique.

Au cours de ces dernières décennies, la compréhension des mécanismes biologiques a fortement progressé grâce au développement d'outils adaptés. Ces techniques ont permis l'identification des ARN non codants ayant un rôle sur la régulation des gènes et l'adaptation de la bactérie à son environnement. Ces petites molécules nucléotidiques sont des outils clés dans l'adaptation et la régulation de l'expression des gènes chez la bactérie.

2. Propriété du modèle bactérien

Escherichia coli est une bactérie à croissance rapide avec un temps de division cellulaire d'environ vingt minutes dans des conditions optimales de croissance. La croissance bactérienne est caractérisée par la succession de phases de croissance : la phase exponentielle de croissance, où la bactérie se divise à un rythme constant (ressources carbonées non limitantes) ; l'entrée en phase stationnaire (transition) où le rythme de division décroît progressivement (appauvrissement en ressources carbonées) et la phase stationnaire de croissance caractérisée par une division cellulaire faible (ressources carbonées épuisées) (Figure 1).

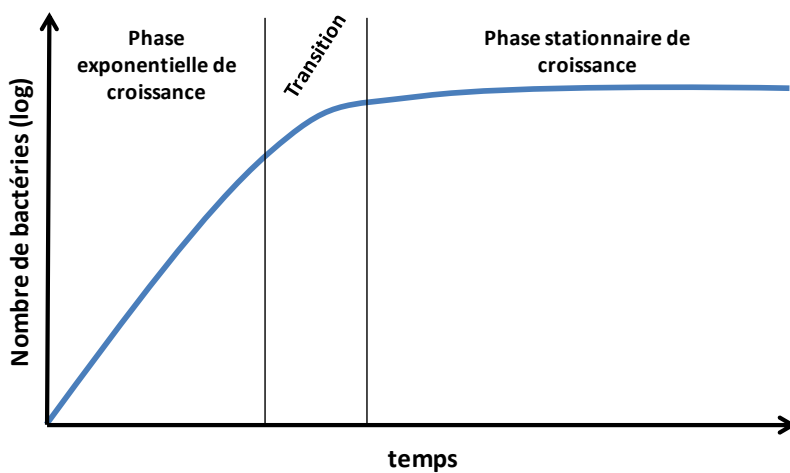


Figure 1 : Schéma représentant la croissance bactérienne.

Le nombre de bactéries augmente exponentiellement pendant la phase exponentielle de croissance ; leur division ralentie à l'entrée de la phase stationnaire de croissance où les ressources carbonées se raréfient pour atteindre la phase stationnaire de croissance où la population bactérienne croît très lentement.

B. Régulation de l'expression génique

1. Le génome

Le génome correspond à l'ensemble du matériel génétique d'un individu. Il est caractérisé par une séquence nucléique représentée par les lettres ATGC, correspondant aux quatre acides désoxyribonucléiques constituant la double hélice d'ADN. Chez *E. coli*, le génome est contenu dans un chromosome circulaire qui contient de nombreuses régions régulatrices précédant des régions appelées gènes. Les gènes sont transcrits par l'ARN polymérase en ARN. Chez les procaryotes, certains gènes sont organisés en opérons, lesquels sont transcrits en un ARN messenger polycistronique codant plusieurs protéines.

2. L'ARN polymérase

Contrairement aux eucaryotes, la transcription du génome procaryote, et notamment celui d'*E. coli*, est assurée par une ARN polymérase unique. Cette ARN polymérase est composée de cinq sous-unités ($\alpha_2\beta\beta'\omega$) formant le cœur de l'enzyme. Il se couple à un facteur sigma pour former l'holoenzyme fonctionnelle $E\sigma$. Cette holoenzyme couplée à son facteur sigma reconnaît un set spécifique de promoteurs permettant ainsi la transcription d'un set de gènes particuliers. Cette spécificité passe par leur capacité à se fixer sur une séquence particulière du promoteur appelée boîte. Par exemple, le facteur sigma majoritaire (RpoD ou σ^{70}) a une affinité particulière pour les boîtes TTGACA et TATAAT, respectivement boîtes -35 et -10 (par rapport au site d'initiation de la transcription).

Les facteurs sigma sont au nombre de sept au sein du génome d'*E. coli* (Figure 2); leur spécificité à réguler des sets de gènes en font des clés majeures dans l'adaptation de la bactérie à son environnement. Les sept facteurs sigma (Ishihama, 2000) sont les suivants :

- RpoD (σ^{70} ou σ^D) facteur végétatif de la bactérie responsable de la majorité de la transcription et spécifique des gènes de « ménage ».
- RpoS (σ^{38} ou σ^S) deuxième facteur majoritaire de la bactérie responsable de l'adaptation à l'entrée en phase stationnaire et de la survie aux stress environnementaux.
- RpoN (σ^{54} ou σ^N) impliqué dans l'expression des gènes lors de croissance en faible concentration d'azote.

- RpoH (σ^{32} ou σ^H) exprimant les gènes induits par un choc thermique.
- FliA (σ^{28} ou σ^F) intervenant dans l'expression des gènes liés à la motilité et la chimiotaxie de la bactérie.
- RpoE (σ^{24} ou σ^E) impliqué dans les chocs thermiques extrêmes.
- Fecl (σ^{19}) affectant l'expression des gènes liés à l'homéostasie du fer.

Parfois, l'holoenzyme seule ne permet pas d'initier la transcription d'un gène et l'intervention d'un cofacteur transcriptionnel est alors nécessaire. Cet exemple peut être illustré pour le gène *csgD* (activateur de la transcription des gènes *csgBA* codant pour la synthèse des curli, appendices extracellulaires impliqués dans l'adhérence, voir partie I.E) dont la transcription est dépendante du facteur OmpR (voir partie III.B.3.a). À l'inverse, les cofacteurs transcriptionnels peuvent également bloquer la transcription des gènes, tels que le cofacteur CpxR qui inhibe la transcription des gènes *csg* (Jubelin et al., 2005).

Intracellular Concentrations of RNA Polymerase Sigma Subunits in *Escherichia coli* W3110

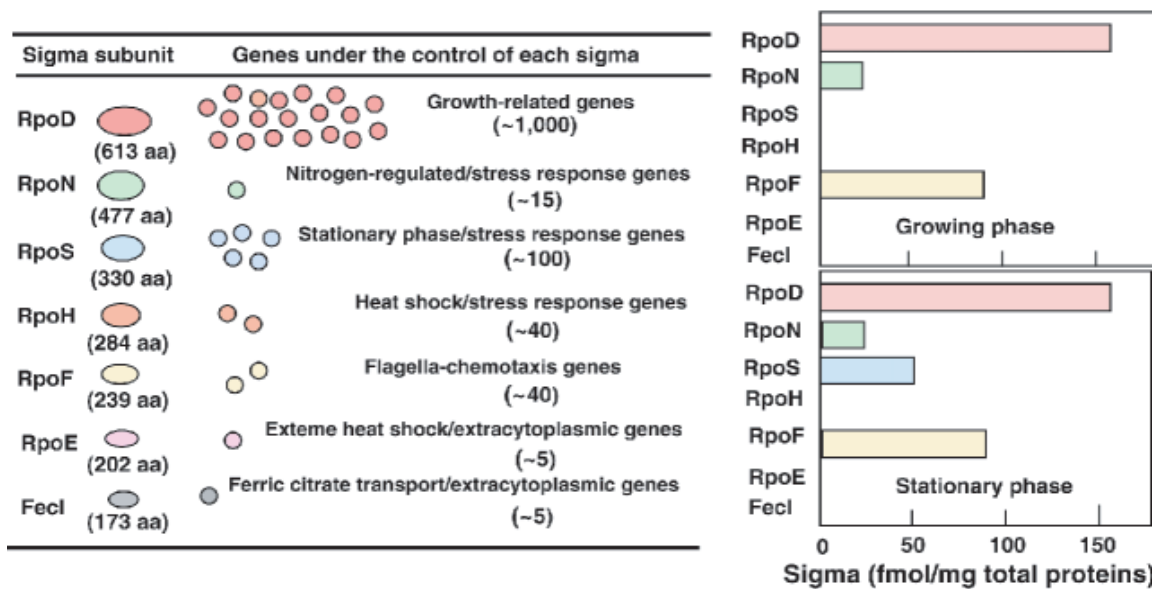


Figure 2 : Les facteurs sigma et leur distribution en fonction de la phase de croissance

Extrait de la revue annuelle (Ishihama, 2000). Le tableau de gauche liste les sept facteurs sigma présents chez *E.coli*, leur nombre d'acides aminés et il indique un nombre approximatif de gènes contrôlés ainsi que leur classe fonctionnelle. Les histogrammes sur la droite montrent l'abondance de chaque facteur sigma en fonction de la phase de croissance (exponentielle et stationnaire).

3. Du gène à la protéine

La transcription d'un gène débute au site d'initiation de la transcription, nommé +1 de transcription ou encore TSS pour Transcription Start Site et se poursuit jusqu'au terminateur transcriptionnel. Les terminateurs peuvent être de deux natures : Rho-dépendants ou Rho-indépendants. Un terminateur Rho-dépendant fait intervenir la protéine Rho qui possède un domaine de liaison à l'ARN. Lors de la fixation de la protéine à une séquence spécifique, cette protéine provoque l'arrêt de la transcription et le décrochement de l'ARN polymérase (Henkin, 1996). Un terminateur Rho-indépendant est constitué d'une région appariée en tige-boucle suivie d'une séquence riche en nucléotides T. Lorsque l'ARN polymérase atteint cette structure, la transcription est stoppée et une légère pause est observée, l'ARN néo-synthétisé se décroche de l'ADN provoquant l'arrêt de la transcription (Henkin, 1996).

L'ARNm comporte différentes régions traduites (CDS pour Coding DNA Sequence, ou ORF pour Open Reading Frame) encadrées par les régions non traduites 5' UTR (pour UnTranslated Region) et 3' UTR. Le CDS comporte l'ensemble des codons composant la séquence en acides aminés permettant la synthèse de la protéine. L'ARNm produit est traduit par une machinerie ribonucléoprotéique nommée ribosome, constituée d'un ensemble de protéines et d'ARN ribosomiques (5S, 16S et 23S). Le ribosome permet l'assemblage de la protéine en catalysant la réaction de liaison peptidique entre le groupement carboxylique et amine des acides aminés. Les régions 5' UTR et 3' UTR comportent elles les régions régulatrices de la traduction et de la stabilité de l'ARNm.

L'initiation de la traduction est assurée par la fixation du ribosome sur l'ARNm au niveau du RBS (pour Ribosome Binding Site, également appelé SD pour séquence de Shine Dalgarno) qui précède la séquence codante (CDS). Cette séquence riche en adénosines et guanines possède une séquence dite consensus complémentaire à l'ARNr 16S du ribosome. Cette séquence peut être dégénérée du consensus, le ribosome se fixera avec une affinité moins bonne au RBS, ce qui réduira le taux de traduction de l'ARNm. La traduction commence au niveau du codon initiateur (ou codon « START »), proche du RBS et se termine au codon « STOP ». Il existe un unique codon « START » (AUG pouvant être substitué très occasionnellement par le codon GUG) et trois codons « STOP » (UGA, UAG et UAA). Les soixante autres codons correspondent aux vingt acides aminés constituant les protéines. Le code génétique est dit redondant, car certains codons codent le même acide aminé. En conséquence, la majorité des mutations affectant les CDS n'entraîne pas de modification d'acide aminé (mutation silencieuse), garantissant ainsi une certaine stabilité du génome. Les processus de transcription à la synthèse protéine sont schématisés dans la Figure 3.

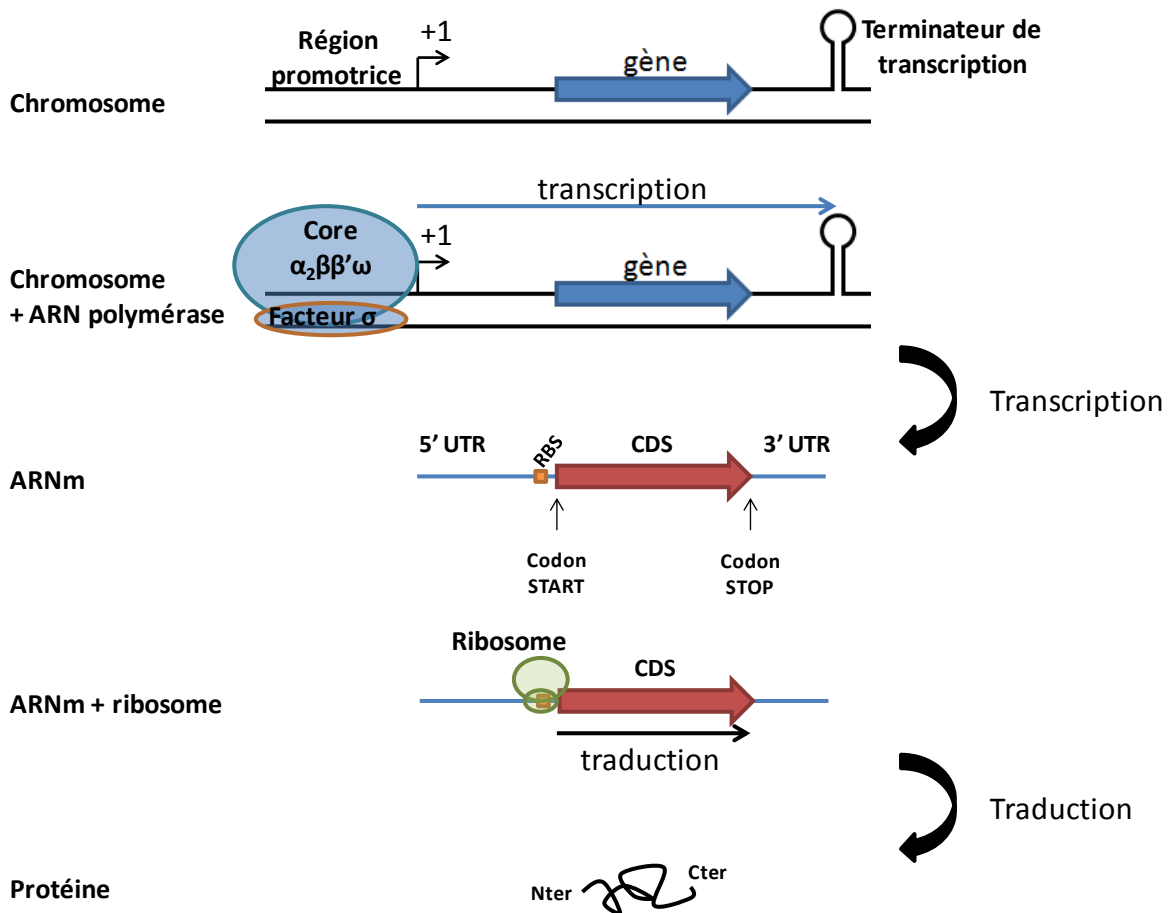


Figure 3 : Mécanisme de transcription à la traduction.

L'ARN polymérase (composée du core : $\alpha_2\beta\beta'\omega$ et du facteur σ) se fixe sur la région promotrice proche du +1 de transcription (TSS). La transcription de l'ARNm débute à partir du TSS et s'arrête au terminateur de transcription. La fixation du ribosome au RBS entraîne l'initiation de la traduction de l'ARNm au codon START se terminant au codon STOP par le détachement du ribosome. La protéine néo-synthétisée comporte un groupement amine libre (appelé Nter, N terminal) correspondant au premier acide aminé de la protéine et d'un groupement carboxylique libre (appelé Cter, C terminal) correspondant au dernier acide aminé de la protéine.

4. Autres catégories d'ARN

Tous les ARN produits par la bactérie ne sont pas traduits. Il existe une multitude d'ARN non traduits : les ARNr (ribosomiques) complexés aux protéines ribosomiques pour former le ribosome fonctionnel, les ARNt (de transfert) nécessaires lors de la synthèse protéique, portant l'anticodon et l'acide aminé, et les ARN régulateurs modulant la traduction et la stabilité des ARNm.

C. Les ARN régulateurs

1. Généralités

La découverte des ARN non codants (ARNnc) a permis de mieux comprendre les mécanismes de la régulation génique et le mode de la régulation traductionnelle (Waters and Storz, 2009). Les ARNnc étudiés jusqu'à présent interviennent dans la plupart des fonctions métaboliques et des fonctions liées aux stress environnementaux. Le nombre d'ARNnc identifiés et caractérisés chez *Escherichia coli* est inférieur à la centaine (http://cbmp.nichd.nih.gov/segr/ecoli_rnas.html). Néanmoins, les données de séquençage haut débit et les programmes de prédictions d'ARN laissent présager plus de deux cents autres ARNnc non caractérisés (Shinhara et al., 2011) au sein du génome d'*E. coli*, soit un vaste sous-ensemble du transcriptome d'*E. coli* à explorer. Ces molécules d'ARN de taille petite à moyenne, d'environ 50 à 200 nucléotides (nts) dans la grande majorité des cas, agissent par appariement avec des ARNm. Cette interaction ARN-ARN a pour conséquence une régulation traductionnelle notamment en agissant sur la stabilité de l'ARNm ciblé ou modulant l'accès du ribosome à son site de fixation (RBS) (Frohlich and Vogel, 2009). Il existe plusieurs catégories d'ARN régulateurs (Waters and Storz, 2009) dont les ARN à action en *trans* et les ARN interagissant avec des protéines dont le mécanisme d'action sera illustré dans les exemples de la partie suivante. Les cas des ARN du CRISPR (mécanisme de défense contre les bactériophages) (Waters and Storz, 2009) et du riboswitch (changement de structure secondaire de la 5'UTR d'un ARNm par des petites molécules) ne seront pas abordés par la suite (Waters and Storz, 2009).

a) Les ARN non codants

b) Les ARN régulateurs d'une activité protéique

Une partie des ARN régulateurs ont pour particularité d'agir avec des protéines pour modifier leur activité (Waters and Storz, 2009). Cette interaction ARN-protéine a pour principale conséquence la séquestration d'une protéine afin de réduire son activité. L'un des premiers ARN régulateurs identifiés ayant cette capacité à séquestrer une protéine est CsrB (Holmqvist and Vogel, 2013; Liu, 1997a). Cet ARN, d'une taille d'environ 300 nucléotides, comporte le motif de fixation de la protéine CsrA responsable de sa titration. Cette protéine a pour particularité de cibler les RBS des ARNm liés

au catabolisme de la néoglucogénèse en empêchant leur traduction (Holmqvist and Vogel, 2013; Liu, 1997b). Par ce mécanisme de séquestration de CsrA, les cibles de la protéine sont ainsi libérées de cette répression traductionnelle (Figure 4).

Un autre exemple du détournement de la machinerie d'expression génique par les ARNnc est l'ARN 6S. Il possède la caractéristique de bloquer l'holoenzyme σ^{70} en s'insérant dans celle-ci (Figure 4). Cet ARN en grande quantité lors de l'entrée en phase stationnaire permet de séquestrer l'holoenzyme σ^{70} au profit de l'holoenzyme σ^S (RpoS) afin de favoriser l'expression des gènes RpoS-dépendant liés à l'adaptation à l'entrée en phase stationnaire (Wassarman, 2007).

Un des autres mécanismes d'action des ARNnc est le mimétisme, le parfait exemple de cette fonctionnalité étant le duo d'ARN régulateurs : GlmY et GlmZ (Figure 4). L'ARN GlmZ est nécessaire à la stabilisation de l'ARNm GlmS permettant la traduction en Glucosamine-6-phosphate synthase (GlcN-6-P). Cet ARNnc (GlmZ) est la cible de deux protéines YhbJ et PcnB connues pour déstabiliser et augmenter sa dégradation (Reichenbach et al., 2008). Lors d'une chute de la concentration en GlcN-6-P, l'ARNnc mimétique GlmY est surproduit. Cet ARN a pour particularité d'avoir une grande similarité avec GlmZ, mais ne possède pas les fonctions de stabilisation de l'ARNm GlmS. GlmY permet la titration des deux enzymes (YhbJ et PcnB) pour permettre à l'ARNnc GlmZ de stabiliser sa cible l'ARNm glmS qui pourra alors être traduit (Reichenbach et al., 2008).

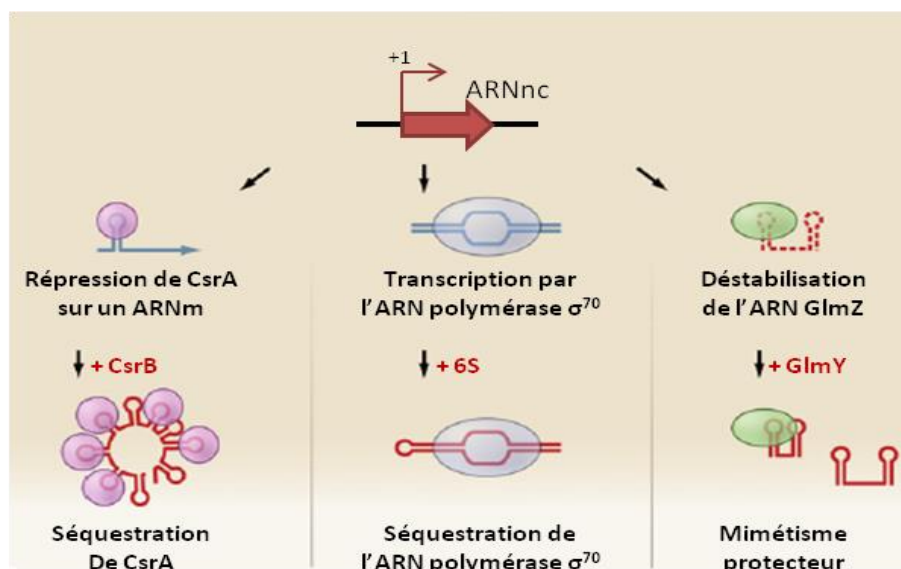


Figure 4 : Différentes interactions ARN régulateurs/protéines.

Adapté de la revue (Waters and Storz, 2009). L'ARNnc CsrB comporte plusieurs sites de fixation de la protéine CsrA qu'il séquestre. Ainsi csrB empêche CsrA de moduler la traduction des ARNm auxquels se fixe la protéine. L'ARNnc 6S s'associe à l'ARN polymérase σ^{70} en phase stationnaire de croissance afin de favoriser la transcription des gènes d'adaptation à la phase stationnaire de croissance par l'ARN polymérase σ^S . L'ARNnc GlmZ, homologue de l'ARNnc GlmY, permet de séquestrer les protéines dégradant GlmY par appariement avec celles-ci (mimétisme protecteur), permettant ainsi à l'ARNnc GlmY de moduler la traduction de l'ARNm GlmZ.

c) Les ARNnc à action en trans

Les ARNnc à action en *trans* sont des ARNnc s'appariant à un ARNm afin de moduler sa demi-vie et/ou sa traduction. Il existe deux catégories d'ARN à action en trans, les ARNnc *cis*-encodés qui sont des ARNnc anti-sens à appariement parfait, et les ARNnc *trans*-encodés, anti-sens à appariement partiel.

Les ARNnc *cis*-encodés

Les ARNnc *cis*-encodés sont des ARN n'ayant qu'une cible unique. L'ARNm cible et l'ARNnc sont situés sur le même locus, l'un sur le brin positif, l'autre sur le brin négatif (Figure 5). Ainsi l'ARNnc est un anti-sens dont l'appariement est total avec son ARNm cible. Ce type d'ARN a été identifié pour la première fois pour son rôle dans la régulation du nombre de copies de plasmide avec une origine de réplication ColE1 (Tomizawa, 1981). Cette configuration est très largement répandue dans le cas de système « toxine/antitoxine », avec l'exemple de l'ARNnc SibC et de la toxine IbsC (Fozo et al., 2008; Han et al., 2010), où l'ARNnc empêche la traduction de l'ARNm en protéine toxique, par la séquestration du RBS ou par dégradation du complexe ARN-ARN par les RNases (Waters and Storz, 2009). Néanmoins, cette configuration n'est pas exclusive à ce système. L'un des mécanismes d'action des ARNnc est l'induction du clivage d'un opéron et la stabilisation des ARNm. L'exemple illustrant ce procédé est l'ARNnc GadY situé dans l'opéron *gadX gadW* (Brantl, 2007; Opdyke et al., 2011).

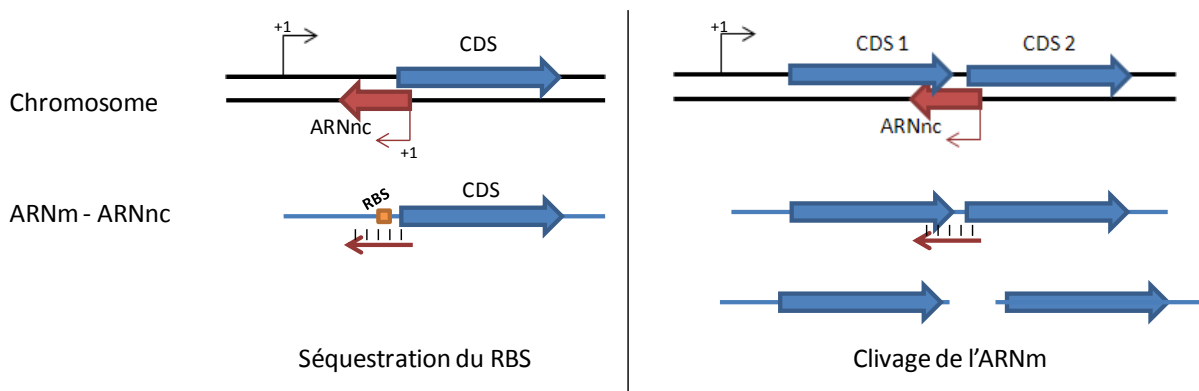


Figure 5 : Mécanisme d'action des ARNnc *cis*-encodés.

Adaptée de la revue (Waters and Storz, 2009). L'ARNnc est transcrit à partir du brin opposé du même locus que l'ARNm ciblé avec lequel il s'apparie complètement ; cet appariement peut entraîner une séquestration du RBS et empêcher la traduction de l'ARNm (exemple : SibC/IbsC) ou peut entraîner le clivage de l'ARNm polycistronique (l'ARNnc GadY sur l'ARNm polycistronique GadXW).

Les ARNnc trans-encodés

Les ARNnc *trans*-encodés sont des ARN régulateurs caractérisés par une structure secondaire composée d'une succession de régions appariées en tige-boucle et de régions simple brin. Ces ARN régulateurs ont la particularité d'avoir une multitude de cibles, mais l'appariement sur celles-ci est partiel et imparfait sur environ 10 à 25 bases, faisant intervenir dans la majorité des cas une protéine chaperonne : Hfq. Ces ARNnc sont généralement présents dans des réseaux complexes. Ils permettent une régulation différentielle d'un réseau comme par exemple le choix d'une bactérie dans son processus d'adhérence ou de motilité (Boehm and Vogel, 2012). Certains ARNnc, comme McaS, qui réprime la traduction des gènes liés à l'adhérence pour augmenter celles des gènes liés à la motilité (Boehm and Vogel, 2012). Les différents modes d'action de ces ARN sont les suivants (Figure 6):

- La libération du RBS : certains ARNm possèdent une région 5' UTR relativement structurée pouvant par repliement sur elle-même séquestrer le RBS, empêchant ainsi le ribosome de venir s'y fixer. L'interaction d'un ARNnc avec l'ARNm permet un changement de conformation de cette région 5' UTR, et la libération du RBS. L'ARNnc agit ici comme un activateur traductionnel (exemple l'ARNm RpoS et l'ARNnc DsrA (McCullen et al., 2010)).
- La séquestration du RBS par appariement de l'ARNnc sur celui-ci a pour conséquence de masquer le RBS au ribosome, et empêche donc la traduction de l'ARNm. Ce processus peut être couplé à la dégradation de l'ARNm pour réprimer la production de la protéine (exemple l'ARNnc RprA sur l'ARNm CsgD (Mika et al., 2012). Ces deux processus sont associés à la répression traductionnelle.

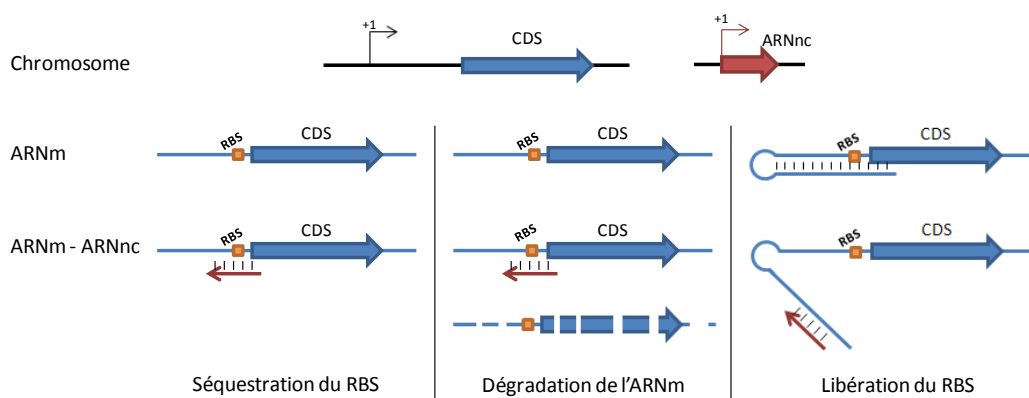


Figure 6 : Mode d'action des ARNnc trans-encodés.

Adaptée de la revue (Waters and Storz, 2009). L'ARNnc s'apparie au niveau de RBS de l'ARNm ciblé empêchant sa traduction (exemple McaS/ARNm CsgD). L'appariement entre l'ARNnc et l'ARNm peut également conduire à la dégradation du duplex. Enfin, la 5'UTR de l'ARNm peut se replier sur elle-même ne permettant pas au ribosome de se fixer sur le RBS (séquestré), l'appariement de l'ARNnc sur l'ARNm permet l'ouverture de la boucle et la libération du RBS, pour obtenir la traduction de l'ARNm (DsrA/ ARNm RpoS).

2. Hfq, protéine chaperonne des ARNnc

Hfq est une protéine appartenant à la famille des protéines LSm qui ont pour propriété de se lier aux ARN. Cette famille possède une structure tridimensionnelle particulière en forme d'anneau, composée de 6 à 7 sous-unités appelées tore. Chez *Escherichia coli*, cette protéine hexamérique lie les ARNnc pour stabiliser leur structure secondaire. Ce complexe ribonucléique permet aux ARNnc une meilleure interaction avec leurs cibles dont l'appariement est imparfait et porte sur des séquences de petites tailles. Dans certains cas, le complexe Hfq-ARNnc-ARNm entraîne le recrutement des RNAses qui dégraderont le duplex ARNm-ARNnc ou uniquement l'ARNm. Cette protéine se lie aux séquences riches en AU. Chez *E. coli*, la grande majorité des ARNnc sont Hfq-dépendants (De Lay et al., 2013; Huttenhofer and Vogel, 2006; Papenfort et al., 2010; Zhang et al., 2003) (Soper T, 2010).

3. À la recherche des ARNnc

Les outils d'étude des transcriptomes ont fortement évolué au cours de la dernière décennie. Les premières études réalisées grâce aux puces à ADN ont évolué vers le séquençage des ARN (convertis en ADNc) permettant de mettre en évidence des transcrits alors insoupçonnés. La résolution de cette technique permet de montrer les variations alléliques (SNP : Single Nucleotide Polymorphism), les sites d'initiation de la transcription (TSS, +1) et avec une grande précision la taille des transcrits. En parallèle de ces techniques de séquençage, les programmes de bio-informatique analysent les séquences génomiques afin de mettre en évidence TSS et terminateurs, ont permis aussi d'enrichir les banques de données en ARN non traduits. D'autres techniques, telles que la co-immunoprécipitation dirigée contre Hfq complexée à des ARNnc, permettent de compléter la base de données. L'étude du mutant *hfq* met en évidence l'importance de sa présence dans la régulation d'ARNm par les ARNnc.

D. La protéine RpoS, facteur σ^{38} de l'ARN polymérase

RpoS est l'un des facteurs sigma secondaires de la bactérie. Il est responsable de la transcription des gènes de l'adaptation à l'entrée en phase stationnaire, mais également de l'adaptation de la bactérie à un environnement stressant. L'ARNm RpoS possède une 5' UTR très structurée en tige-boucle dans laquelle le RBS est séquestré par l'appariement de régions avoisinantes. L'ARNm RpoS est la cible de plusieurs ARNnc *trans*-encodés. Les ARNnc DsrA, RprA et ArcZ activent la traduction de l'ARNm rpoS par appariement et libération du RBS (Figure 7). L'ARNnc OxyS agirait en répresseur en séquestrant Hfq, empêchant ainsi les trois autres ARNnc de se fixer à celle-ci (Henderson et al., 2013).

La protéine Hfq est fortement liée à la régulation traductionnelle de cet ARNm en se liant aux trois ARNnc responsables de sa ribo-régulation (libération du RBS). Une étude du mutant *hfq* sur une fusion traductionnelle de l'ARN RpoS montre une diminution de l'activité β -galactosidase de cette fusion d'un facteur cinq. La présence de Hfq pour les ARNnc ArcZ et RprA augmente l'efficacité de l'activité traductionnelle d'un facteur de vingt. L'ARNnc DsrA semble lui indépendant de Hfq pour pouvoir activer la traduction de RpoS, car la mutation Δhfq a peu d'impact sur sa capacité à lier l'ARNm RpoS (Mika et al., 2012). On observe une thermorégulation de l'ARNm RpoS préférentielle à 37°C par l'ARNnc ArcZ, ou aux températures inférieures à 28°C par DsrA. (Soper T, 2010), cependant la protéine RpoS est préférentiellement accumulée à 28°C.

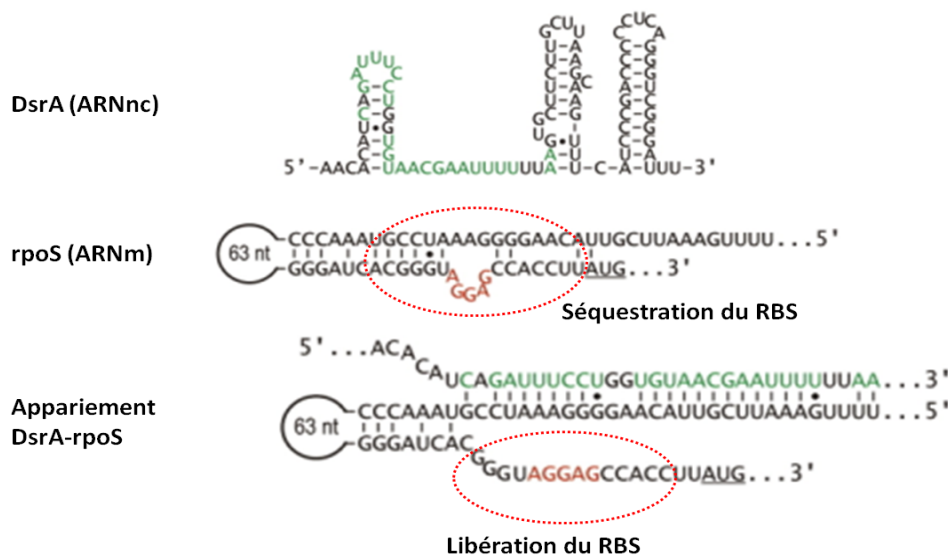


Figure 7 : Libération du RBS de l'ARNm RpoS par l'ARNnc DsrA.

Adaptée de la revue (Frohlich and Vogel, 2009). La 5'UTR de l'ARNm RpoS se replie sur elle-même pour former une structure secondaire entraînant la séquestration du RBS. En s'appariant à l'ARNm RpoS, l'ARNnc DsrA permet de libérer le RBS de cette structure secondaire, pour permettre au ribosome de venir se fixer et de traduire l'ARNm RpoS.

E. Régulation de l'expression de protéines d'adhérences : les curli, par le facteur sigma RpoS et les ARNnc

Les bactéries ont la capacité de s'organiser en communauté mono ou pluri-espèce nommée biofilm. Cette communauté bactérienne est encapsulée dans une matrice riche en polysaccharides, en protéines d'adhérence, et autres macromolécules, d'ADN notamment. Cette matrice confère à ces occupants une capacité de résistance accrue dans l'environnement aux nombreux stress (osmotique, thermique, mécanique, oxydatif...etc.) et aux antibiotiques. La souche utilisée au laboratoire, *Escherichia coli* BW25113, possède cette capacité à former des biofilms. Ce phénotype nécessite la présence d'adhésines pour la formation de celui-ci. Les études préliminaires au sein du laboratoire démontrent le besoin de l'expression des gènes responsable de l'expression des curli pour la formation du biofilm. Les curli sont des appendices extracellulaires de type fimbriae organisés en feuillets β . L'expression des curli et leur sécrétion sont codés par deux opérons *csgDEFG* (gène régulateur et système d'excrétion) et *csgBAC* (gènes structuraux). Ces opérons ont fait l'objet de nombreuses études montrant un nombre élevé de protéines régulatrices (Figure 8 B et C), néanmoins le réseau semble encore incomplet et les résultats diffèrent en fonction de la souche utilisée.

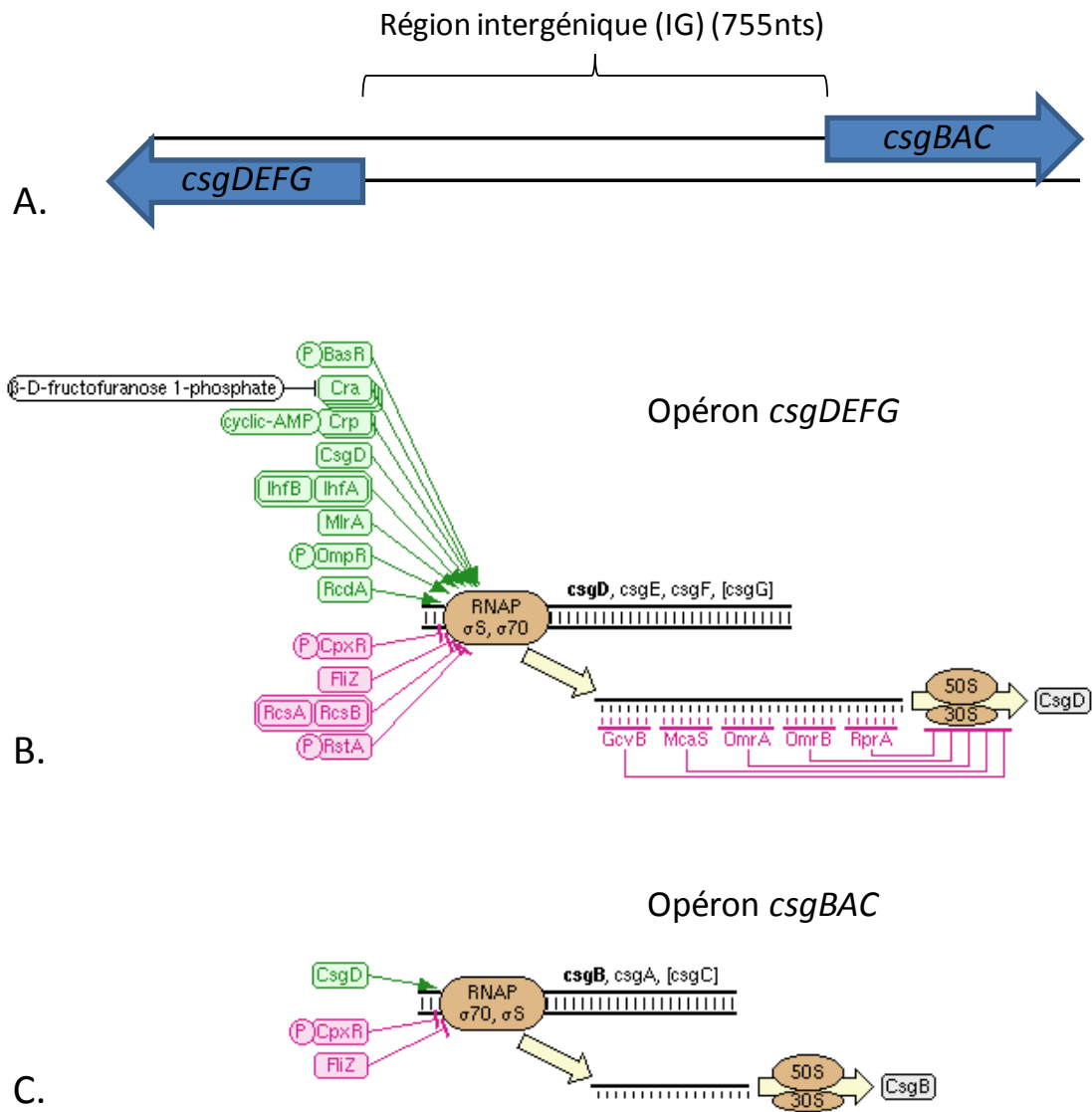


Figure 8: Région intergénique entre les opérons *csgDEFG* et *csgBAC* et sites de fixation de leurs régulateurs leurs transcriptionnelles ou traductionnelles.

A. Schéma représentant les opérons *csgDEFG* et *csgBAC* et leur région intergénique de 755nts **B** et **C** Les régulations transcriptionnelles et traductionnelles des deux opérons *csg* sont représentées en vert pour les régulations positives et en rose pour les répressions. Schémas extraits du site ecocyc.org

1. Régulation transcriptionnelle des opérons *csg*

Les opérons *csg* partagent la même région intergénique et sont encodées dans le sens opposé (Figure 8 A). Ils sont la cible de nombreuses protéines régulatrices rendant le réseau complexe à étudier. Une quarantaine de sites de fixation protéique sont trouvés au sein de cette région intergénique faisant intervenir un grand nombre de régulateurs transcriptionnels : CpxR (Dorel et al., 1999; Ogasawara et al., 2010; Prigent-Combaret et al., 2001) , Crp (Zheng et al., 2004), IHF (Ogasawara et al., 2010), OmpR (Prigent-Combaret et al., 2001), Cra (FruR) (Reshamwala and Noronha, 2011), et lui-même (Ogasawara et al., 2011). Toutes ces protéines régulent positivement l'expression de l'opéron *csgDEFG* à l'exception de CpxR.

Un cas intéressant est celui du régulateur OmpR qui a été mis en évidence par une étude mettant en avant le rôle de la mutation *ompR234* (CTG → CGG : L43R) (en réalité *ompR128*, la taille du gène étant revu à la baisse). Le mutant a été isolé en fermenteur pour sa capacité à former plus de biofilm qu'une souche sauvage (Prigent-Combaret et al., 2000; Vidal et al., 1998).

La transcription de l'opéron *csgDEFG* a pu être détectée pour une densité optique correspondant à la phase de transition d'une phase exponentielle à celle d'une phase stationnaire appelée entrée en phase stationnaire (DO = 4) et une température de 28°C. L'expression de l'ARNm à 37°C n'a pas été détectée (Mika et al., 2012). Une étude, via des puces à ADN, sur l'expression des gènes à la faible température de 23°C révèle que l'expression des opérons *csg* est RpoS-dépendante (White-Ziegler et al., 2008). L'expression de RpoS est également dépendante de la présence de l'ARN DsrA pour les faibles températures (Sledjeski et al., 1996), induisant l'accumulation de ce facteur à basse température.

L'opéron *csgBAC* est la cible d'un nombre restreint de protéines régulatrices : CpxR (Ogasawara et al., 2010), CsgD (Ogasawara et al., 2011) et Cra (Reshamwala and Noronha, 2011). CpxR régulateur négatif, CsgD et Cra régulateur positif de la transcription de l'opéron.

L'ensemble de la régulation transcriptionnelle des curli affecte principalement l'expression du gène régulateur et des gènes codants leur système d'excrétion. La régulation importante de la traduction des gènes *csgDEFG* par des ARNnc caractérise ces gènes comme des facteurs déterminants pour l'expression des curli (Figure 9)

2. Régulation traductionnelle de l'ARNm *csgDEFG*

L'ARN messager *csgDEFG* comporte une grande région 5' UTR où vient se fixer plusieurs ARNnc. Ces ARNnc régulateurs sont actuellement au nombre de six : RprA (Mika et al., 2012), McaS (Thomason et al., 2012), OmrAB (Holmqvist et al., 2010), GcvB (Jorgensen et al., 2012) et le dernier identifié très récemment, RydC (Bordeau and Felden, 2014). Ils répriment tous la traduction de l'ARNm *csgDEFG*. RprA est un ARNnc qui possède deux sites de fixation sur l'ARNm *csgDEFG*, l'un situé au milieu de la région 5' UTR et l'autre proche du RBS (Mika et al., 2012). Le système *rcs* lié à l'expression de la capsule, régule négativement l'expression des curli via l'expression de l'ARNnc RprA répresseur traductionnel de l'ARNm CsgD. Les ARN OmrAB bien que régulés positivement par OmpR régulent négativement CsgD contrairement à OmpR régulant positivement l'expression de l'ARNm *csgDEFG*. OmrA (88 nts) et OmrB (82 nts) ont une séquence similaire à 68% et leur extrémité 3' est identique. Ils s'apparient loin en amont du RBS sur 18-19 nts afin de déstabiliser une structure secondaire de l'ARNm *csgDEFG* en amont du RBS. Cette liaison est suffisante pour entraîner une diminution de la traduction de l'ARNm (Holmqvist et al., 2010).

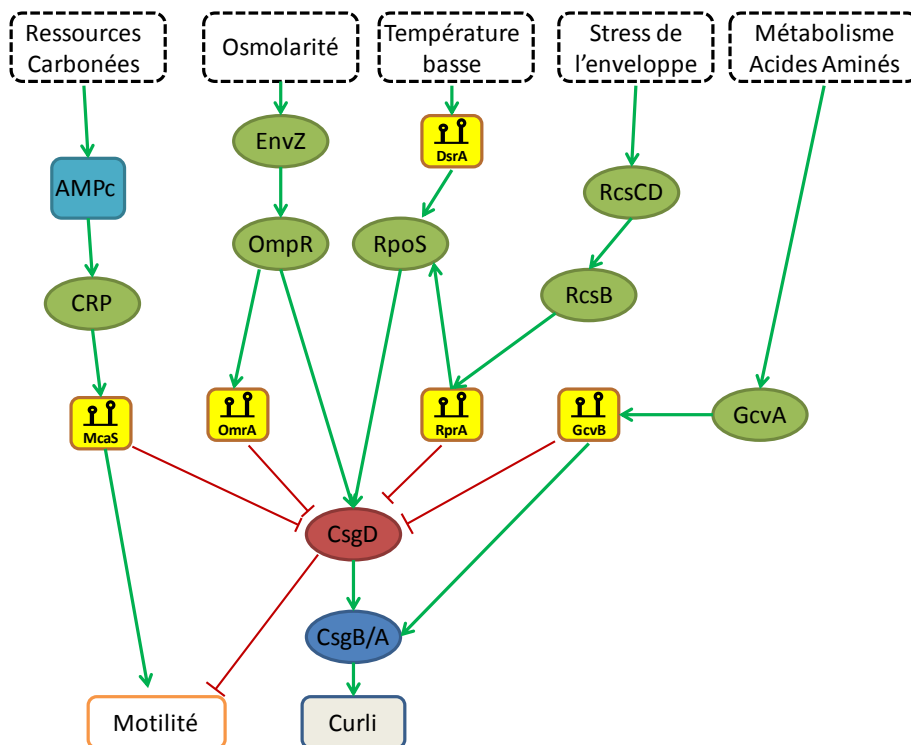


Figure 9 : Réseau de régulation contrôlant la différenciation entre cellule adhérente ou motile.

Adaptation du schéma de régulation contrôlant le choix entre motilité ou adhésion (Boehm and Vogel, 2012) Ce processus est induit par des stimuli environnementaux et des signaux du métabolisme. Les régulations positives sont indiquées par les flèches vertes, les répressions sont indiquées par les flèches rouges.

F. Objectifs

1. Identification et caractérisation de nouveaux ARNnc régulés par RpoS

RpoS peut être considéré comme le second facteur sigma végétatif d'*Escherichia coli* compte tenu de l'ampleur de son régulon qui comporte environ 500 gènes (Lacour and Landini, 2004; Patten et al. 2004; Weber et al., 2005; Dong 2008 and 2009). La régulation de son accumulation est parmi les plus complexes et sa traduction est régulée par quatre ARNnc. De plus en plus d'études démontrent le rôle important joué par ces ARNnc dans la régulation de l'expression génique et par extension dans de la cellule bactérienne. Le régulon de RpoS compte de nombreux facteurs de transcription ce qui démontre qu'une large partie des régulations par RpoS est probablement indirecte. Notre ambition était de démontrer que ce facteur sigma peut aussi largement affecter la traduction ou la fonction de certaines protéines chez *E. coli* en régulant l'expression d'ARN régulateurs ; c'est pourquoi nous avons entrepris d'identifier de nouveaux ARNnc régulés par RpoS. Depuis l'initiation de ce projet, on a découvert la régulation de quelques ARNnc par le facteur sigma RpoS : RyjA (Silva et al., 2013) et RyeB (Frohlich et al., 2012). D'autres régulations d'ARNnc ont aussi été suggérées par des expériences récentes de transcription *in vitro* (Maciag et al., 2011) et *in vivo* (Dong et al., 2011). Malheureusement, au moment de la rédaction de ce mémoire, un transcriptome de RpoS par séquençage d'ARN vient d'être publié chez *Salmonella* qui décrit une partie de nos découvertes, notamment la régulation de OmrA qui régule l'expression du régulateur CsgD et de porines (Levi-Meyrueis et al., 2014). Il décrit la régulation par RpoS d'ARNnc décrits chez *E.coli* ou spécifiques aux Salmonelles. Nous nous sommes d'abord appuyés sur les données générées par les puces Affymetrix de première génération (Dong et al., 2008; Patten et al., 2004) des prédictions bio-informatiques d'ARNnc sur le génome bactérien (Sridhar et al., 2010). Nous nous sommes ensuite résolus à initier une nouvelle étude transcriptomique à haute résolution par RNA-seq, dans le but d'identifier de nouvelles cibles de RpoS, en particulier des ARNnc connus, mais surtout inconnus.

2. Étude globale du réseau régulateur des opérons *csg*

Les opérons *csg* sont la cible de nombreuses protéines régulatrices, mais également la cible d'ARNnc. La régulation de leur expression est aussi complexe que celle du régulateur majeur de la réponse au stress, RpoS. Nous avons reconstruit des fusions transcriptionnelles et traductionnelles pour revisiter leur régulation dans une souche unique d'*E. coli*. Ces fusions nous ont permis d'observer la sensibilité de l'expression des deux opérons *csg* à la température de croissance imposée aux bactéries. Nous avons tenté d'identifier le facteur génétique responsable de cette thermorégulation. Nous avons pensé qu'il pouvait s'agir d'un ARNnc puisque la température influence les interactions ARN-ARN et avons entre autres tenté de démontrer si un petit ARN était impliqué. De plus, RpoS est connu pour s'accumuler préférentiellement à basse température en partie à cause de sa régulation par DsrA (Sledjeski et al., 1996). Nous avons donc voulu vérifier si RpoS était bien un régulateur majeur de la transcription des gènes *csg* et s'il était impliqué dans la thermorégulation par exemple en contrôlant l'expression des ARNnc qui module la traduction de RpoS.

II. MATÉRIELS ET MÉTHODES

A. Milieux, souches, plasmides

1. Les milieux de culture

La composition des différents milieux de culture est décrite ci-dessous

LB : Lysogeny Broth (pour 1 litre)

10 g Bacto Tryptone (BD)

5 g Bacto Yeast Extract (BD)

5 g Chlorure de Sodium

LA : LB Agar (pour 1litre)

Constituant du LB +12g de Bacto Agar (BD)

Gélose molle

200ml de LB sont supplémentés par 30ml d'Agar 2X (24g d'Agar par litre)

LA RC : LB-Agar Rouge Congo

Une dilution au 1/200^{ème} d'une solution de Rouge Congo à 0.8% est appliquée au volume de culture désiré (concentration finale de Rouge Congo 0.004%)

La solution de Rouge Congo est préparée à partir 0.8g de Congo Red (Sigma-Aldrich) compléter par de l'eau stérile à 100 ml, la solution est ensuite filtrée (filtre 0,2µm).

2. Souches bactériennes

Les expériences ont été réalisées sur la souche *E. coli* BW25113 (dérivée de la souche K12) et de l'utilisation d'une collection de mutants (délétion d'un gène non essentiel) émanant de l'université Keio (Japon) décrite dans (Baba et al., 2006). Certains mutants de cette librairie ont été utilisés pour notre étude et sont listés ci-dessous sous la référence « Collection Keio ». Les mutations des gènes sont créés par insertion d'une cassette de résistance à l'antibiotique Kanamycine (Kan) qui a pu être amené à être excisée par recombinaison des éléments FRT présent dans la cassette (distingué par l'annotation « ::Kan » pour la présence de la cassette ou « Δ » si excisée). Nous avons par ailleurs pu bénéficier de mutants déjà construits au sein du laboratoire, notamment par Caroline Ranquet sous la référence « Laboratoire_Caro ». Enfin, certains mutants de cette étude ont été réalisés par transduction ou par recombinaison chromosomique.

Souche	Mutation	Référence
JD87	<i>mcaS::Kan</i>	Cette étude
JD88	<i>ompR234 malT::Tn10</i>	Cette étude
JD90	<i>ompR234 malT::Tn10 ΔrpoS</i>	Cette étude
JD92	<i>ompR234 malT::Tn10 ΔrpoS mcaS::Kan</i>	Cette étude
JD91	<i>ompR234 malT::Tn10mcaS::Kan</i>	Cette étude
JD122	<i>gcvB::Kan</i>	Cette étude
JD123	<i>ompR234 malT::Tn10 gcvB::Kan</i>	Cette étude
JD124	<i>ompR234 malT::Tn10 rprA::Kan</i>	Cette étude
JD125	<i>ompR234 malT::Tn10 omrAB::Kan</i>	Cette étude
SL30	Aucune (WT)	Collection Keio
SL128	<i>ΔrpoS</i>	Collection Keio
SL121	<i>ΔcsgB</i>	Collection Keio
SL122	<i>ΔcsgD</i>	Collection Keio
SL167	<i>ompR::Kn</i>	Collection Keio
Caro3_59	<i>rprA::Kan</i>	Laboratoire
Caro3_62	<i>omrAB::Kan</i>	Laboratoire
Caro3_63	<i>dsrA::Kan</i>	Laboratoire
SL67	<i>ompR234 malT::Tn10</i>	Pringent-Combaret C. (1999)

3. Fusion transcriptionnelle et traductionnelle

Pour réaliser notre étude, nous avons construit des fusions transcriptionnelles et traductionnelles par insertion des séquences dans le vecteur pZE GFP (~20 copies/cellule, résistance spec^R)(Vecteur utilisé au laboratoire pour l'étude de la traduction et transcription d'un gène). Les fusions transcriptionnelles ont été construites en fusionnant le promoteur du gène au RBS et CDS du gène codant la GFP (voir protocole II.C.5) (Green Fluorescente Protein, GFPmut3 ASV), et les fusions traductionnelles par insertion du promoteur et de la région 5'UTR comportant le RBS endogène du gène fusionné au CDS de la GFP. Seules les constructions présentées dans ce mémoire sont listées dans le tableau suivant :

Souche	Construction	Type	Position
JD62	IG <i>csgD</i> ::GFP	TRAD	-620 à +163
JD77	p <i>rprA</i> ::GFP	TXN	-187 à +4
JD78	p <i>mcaS</i> ::GFP	TXN	-232 à +8
JD79	p <i>gcvB</i> ::GFP	TXN	-198 à +1
JD81	p <i>omrA</i> ::GFP	TXN	-114 à +8
JD82	p <i>omrB</i> ::GFP	TXN	-179 à +8
JD98	p <i>sibC</i> ::GFP	TXN	-200 à +8
JD100	p <i>dsrA</i> ::GFP	TXN	-179 à +9
JD101	p <i>ryeA</i> ::GFP	TXN	-200 à +8
JD129	p <i>csgD</i> ::GFP	TXN	-620 à +2
JD130	p <i>csgD</i> ΔUTR::GFP	TRAD	-620 à +137 (+2 à +110 délété)
JD131	p <i>csgD</i> UTR+::GFP	TRAD	-620 à +137
JD132	p <i>csgB</i> ::GFP	TXN	-668 à +2
JD133	p <i>ryeB</i> ::GFP	TXN	-154 à +8
JD134	p <i>ibsC</i> ::GFP	TXN	-23 à -173 de l'ATG *

Les abréviations TXN pour transcriptionnelle et TRAD pour traductionnelle correspondent à la spécificité de la fusion. Le terme « Position » correspond à la séquence clonée dans le vecteur pZE GFP, cette position fait référence au positionnement des nucléotides par rapport au TSS. Note pour l'astérisque (*), le promoteur du gène *ibsC* n'est pas référencé dans les bases de données.

4. Autres plasmides

En plus des fusions transcriptionnelles et traductionnelles, des vecteurs de surexpression (passive du au haut nombre de copies (environ 200/cellule)) ont été construits par insertion de la séquence d'intérêts dans le vecteur pTOPO (kit TOPO TA Cloning Kit Dual Promoter (Invitrogen)). Ces constructions ont été utilisées pour observer l'effet de la surexpression de l'entité insérée sur des phénotypes, tels que la croissance ou la liaison au Rouge Congo, mais également comme matrice de transcription *in vitro*. Les différentes constructions sont listées sous le tableau ci-dessous.

Souche	Construction	Souche	Construction
JD28	pTOPO Djs	JD97	pTOPO C5221
JD30	pTOPO IG1027	JD110	pTOPO ECS074
JD31	pTOPO IG1105	JD111	PTOPO NC014
JD32	pTOPO tpke11	JD113	pTOPO C0042
JD33	pTOPO C0465	JD114	pTOPO C0614
JD34	pTOPO contrôle	JD115	pTOPO ryhB
JD85	pTOPO gcvB	JD116	pTOPO isrC
JD86	pTOPO mcaS	JD118	pTOPO ECS066
JD95	pTOPO sibC	JD120	PTOPO oxyS
JD96	pTOPO ryeA	JD121	pTOPO IO41

Les différents plasmides (fusions transcriptionnelles, traductionnelles et vecteurs de surexpression) sont susceptibles d'avoir été transformés dans les souches listées en début de partie.

5. Lecture d'absorbance et de fluorescence en microplaque

Certaines souches, notamment les souches contenant les fusions transcriptionnelles ou traductionnelles ont été cultivées en microplaque (cinétique)(microplaque 96 puits Greiner), 150µl de milieu de culture, une nuit à température d'étude. Les bactéries sont ensuite inoculées à une DO finale de 0,05 et cultivées dans le lecteur de microplaque, pour obtenir l'absorbance (OD 600nm) et la fluorescence (excitation 485nm, émission 535nm) de chacune des souches bactériennes par le lecteur Tecan Infinite pro F200 (Tecan Group Ltd., Männedorf, Suisse), les données acquises sont générées par le logiciel associé Tecan i-control dans un fichier Excel.

B. Biologie moléculaire

1. Amplification d'ADN par PCR

L'amplification d'ADN par PCR peut être distinguée par l'utilisation de deux polymérase différentes : la « Phusion® High-Fidelity DNA Polymerase » (Biolabs) pour une amplification à haute fidélité (clonage, séquençage, fragments d'intérêts), ou par la « Taq DNA Polymerase » (Biolabs) pour une utilisation routinière. Les amorces utilisées ont été synthétisées chez Sigma-Aldrich. Les tampons 10X (Taq DNA Polymerase) ou 5X (Phusion® High-Fidelity DNA Polymerase) sont fournis avec les DNA polymérase. La solution de dNTPs est préparée à partir de dATP, dGTP, dCTP et dTTP (Biolabs). L'eau utilisée pour les mix de biologie moléculaire et une eau DNase et RNase free (distilled water, Gibco). Les composants et proportions du mix réactionnel de PCR ainsi que le programme d'amplification du thermocycleur sont détaillés ci-dessous.

Mix réactionnel	
Eau	qsp 50µl
Tampon 10X ou 5X	5µl ou 10µl
dNTPs 5mM	2µl
amorce 1, 10µM	2µl
amorce 2, 10µM	2µl
DMSO 100%(optionnel)	1.5µl
Taq polymérase	0.5µl

Programme du thermocycleur		
préchauffe	95°C	2min
Dénaturation	95°C	30s
Hybridation	selon Tm*	30s
Élongation	72°C	1kb/min
Élongation finale	72°C	5min
Pause	4°C	∞

La matrice est négligée dans le calcul du mix réactionnel. Elle peut être issue de colonies bactériennes, d'une culture cellulaire ou encore d'ADN plasmidique. Selon Tm* : température d'hybridation variable entre les amorces (45°C-72°C)(Annexe1). Le DMSO est utilisé afin de réduire les structures secondaires des amorces de grandes tailles (supérieures à 40nt)

2. Analyse par gel d'électrophorèse

Les fragments d'ADN obtenus lors de digestions enzymatiques ou par PCR sont observés par migration électrophorétique en gel d'agarose (Agarose standard, Eurobio)(0.5 à 3% selon la taille des ADN). 5µl d'échantillon est complété par 1µl de bleu de charge 6X (0.25% bromophenol blue, 0.25% xylene cyanol FF, 30% glycérol, élué dans de l'eau) sont déposés dans le gel d'agarose. Deux marqueurs de taille sont utilisés en fonction de la taille des ADN à observer, le SmartLadder (200bp à 10000bp, Eurogentec) ou le SmartLadderSF (100bp à 1000bp, Eurogentec). La migration est effectuée dans une cuve contenant du tampon TAE 1X à un voltage constant de 100V. Les gels sont ensuite baignés dans solution de Bromure d'éthidium (Eurobio 0.7mg/ml, dilution au 1/10000^{ème} dans de l'eau) . La révélation du gel est réalisée par exposition au rayon UVs.

3. Purification d'ADN synthétisé et d'ADN plasmidique

Les purifications de produit de PCR et les extractions sur gel sont réalisées par l'utilisation du kit NucleoSpin®Gel and PCR Clean-up. L'extraction d'ADN plasmidique est réalisée par l'utilisation du kit NucleoSpin® Plasmid QuickPure (Macherey-Nagel). Ces différentes extractions sont réalisées selon le protocole du fournisseur.

4. Extraction d'ARN totaux au phénol chaud

Une préculture de souche souhaitée est effectuée en milieu minimum M9 ou en LB sur la nuit. Elle est relancée au 100^{ème} ou au 50^{ème} selon la richesse du milieu. Une fois la DO souhaitée obtenue les cellules sont centrifugées (12000rpm, 2min). Le surnageant est éliminé et les cellules resuspendues dans 200µl de tampon de lyse 1X préchauffé à 65°C. 500µl de phénol-water préchauffé sont ajoutés, le mix est incubé dans un thermoagitateur (Eppendorf, Thermomixer 5436, 1400rpm, 5min, 65°C). 200µl de chloroforme est ajouté avant centrifugation (12000rpm, 10min, 4°C). La phase aqueuse est collectée dans un nouveau tube et complétée par le même volume de phénol water, une incubation au thermoagitateur est réalisée (1400rpm, 5min, 65°C) avant d'ajouter à nouveau 200µl de chloroforme. Le mélange est centrifugé (12000, 5min, 4°C), puis la phase aqueuse est transférée dans un nouveau tube auquel 3 volumes d'isopropanol 100% froid sont ajoutés. La précipitation est réalisée sur une nuit à -20°C. La précipitation est finalisée par une centrifugation (12000rpm, 10min, 4°C). L'isopropanol est intégralement enlevé, le culot est lavé par 500µl d'éthanol 75% froid avant centrifugation (12000rpm,10min, 4°C). Le surnageant est intégralement enlevé et les culots séchés afin d'éliminer les résidus d'éthanol. Les culots sont finalement resuspendus dans le volume d'eau « RNase free » désiré. Les échantillons sont conservés à -80°C.

Solution de lyse 8X (160µM Acetate de Sodium, 4% SDS, 8µM EDTA)

5. Northern Blot

Les Northern blot ont été effectués à partir des échantillons d'ARN totaux issus de la technique d'extraction au phénol chaud. Après dosage au Nanodrop (Thermo Scientific, NanoDrop2000 Spectrophotometer), ceux-ci sont normalisés à une concentration de $1\mu\text{g}/\mu\text{l}$ d'ARN totaux avant d'être dilués de moitié par du bleu de charge 2X (Invitrogen, Novex® TBE-Urea Sample Buffer 2X) pour être déposés sur un gel polyacrylamide dénaturant de 6% ou 10% (Novex® TBE-Urea Gels, 1mm). La migration est effectuée dans un tampon TBE1X baignant dans la cuve d'électrophorèse verticale (XCell SureLock™) soumis à un voltage constant de 180V. La durée de migration dépend de la résolution du gel et des fragments que l'on souhaite observer.

Le transfert des ARN totaux s'effectue sur une membrane de nylon chargée positivement (BrightStar®-Plus Positively Charged Nylon Membrane) par la technique d'électrotransfert dans un tampon TBE 0,5X baignant la cuve d'électrotransfert (Biorad, Mini Trans-Blot® Electrophoretic Transfer Cell) soumis à un voltage constant de 30V pendant 1h.

Les ARN totaux sont immobilisés sur la membrane de nylon par soumission aux UV à $120000\mu\text{l}$ (fonction autocrosslink) dans le crosslinker (Stratagene, Stratalinker® UV Crosslinker). Les membranes sont ensuite pré hybridées (1h) dans le tampon d'hybridation (Ambion® ULTRAhyb®-Oligo) préalablement tempéré à 68°C. La membrane est placée dans un tube en verre et mise sous agitation rotative dans un four d'hybridation (Stuart, Hybridisation oven, SI30H) (45°C, 4rpm). Les sondes biotynilées en 5' (Sigma Alderich)(Tableau 1) sont préparées pour une quantité finale de 100pmol dans 50 μl de tampon ULTRAhyb puis maintenu à 45°C le temps de la pré hybridation de la membrane. La sonde est alors dénaturée à 95°C avant d'être ajoutée à la membrane (1 nuit, 45°C, 4rpm).

Les membranes sont brièvement rincées par la solution 1 (2X SCC, 0.1% SDS), puis lavées deux fois par cette même solution (15min, 4rpm, 45°C). L'opération est renouvelée par un lavage d'une solution 2 (0.1X SCC, 0.1% SDS) pendant 15min répétées une seconde fois.

La révélation du Northern blot est entreprise par l'utilisation du kit Chemiluminescent Nucleic Acid Detection Module (Thermo Scientific) et des recommandations du fournisseur. Les membranes traitées sont exposées sous caméra luminescence pour l'acquisition d'image de la membrane.

[Btn]omrA	[Btn]TTGGTGCAAGAGACAGGGTACGAAG
[Btn]sibC	[Btn]CATCATACTGATTGTACTGTTACTC
[Btn]ryeA	[Btn]AGAGCCATTTCCCTGGACCGAATAC
[Btn]ryeB	[Btn]CCTTGCCCTTTAAGAATAGATGACG
[Btn]ibsC	[Btn]GAGTAACAGTACAATCAGTATGATG
[Btn]5S	[Btn]CGGCGCTACGGCGTTTCACTTCTG
[Btn]mcaS	[Btn]CCGCCAGACTCTACAGTACACAC
[Btn]-ARNcsgD	[Btn]TACTGACAGATGTTGCACTGCTGTGTGTAG

Tableau 1 : Liste des sondes biotinylées pour les Northern blot

Une molécule de biotine ([Btn]) a été greffée à l'extrémité de l'oligo en 5'.

6. Puce à ADNc

Les reverse transcription sont effectuées à partir du kit SuperScript™II Reverse Transcriptase (Invitrogen) et les instructions du fournisseur sur 5µg d'ARN totaux. La reverse transcription est purifiée pour éliminer l'alpha dCTP ³²P libre par l'utilisation d'une colonne de purification (Microcon-30kDa Centrifugal Filter Unit with Ultracel-30 membrane) selon les recommandations du fournisseur.

Les membranes de capture sont générées par dépôt de produit de PCR sur une membrane de nylon (BrightStar®-Plus Positively Charged Nylon Membrane). Les sondes normalisées à 1pmol dans 30µl de SSC2X sont déposées par l'utilisation d'un « spoter » (Biorad, Bio-Dot Apparatus) sur une membrane préalablement traitée par une solution de SSC 10X. Les sondes sont immobilisées sur la membrane, en déposant celle-ci sur du papier Wathman imbibé d'une solution de NaOH 0.4M.

Les membranes sont préhybridées par une la solution ULTRAhyb-oligo. L'intégralité du produit de reverse transcription est versé dans les tubes d'hybridation. L'hybridation est faite sur la nuit à 60°C.

Les lavages des membranes dépendent du bruit de fond émis par la membrane. Cependant, une première série de lavage est réalisée par deux solutions, la solution 1(SSC 0.05X, SDS 0.05%) pendant 5min à 60°C (fonction rôtissoire, 4rpm), puis par la solution 2 (SSC 0.1X, SDS 0.1%) pendant 5min à 60°C (fonction rôtissoire, 4rpm).

La membrane est placée dans un film Saran puis mise en contact avec une membrane IP et révélée après exposition de 3h à 48h par le phospho-imager (FUJIFILM FLA8000).

7. qRT-PCR

Les qRT PCR sont réalisées à partir des échantillons d'ARN totaux rétro-transcrits (SuperScript™II Reverse Transcriptase (Invitrogen)). La réaction, de qRT-PCR utilisant le kit MESA GREEN qPCR Master mix plus for SYBR Asay dTTP (Eurogentec) et instructions du fournisseur, est réalisée dans thermocycleur de qPCR (StepOne Plus, Applied Biosystem). Les sondes utilisées en qRT-PCR sont listées ci-dessous.

Cible	Amorce foward (5' -> 3')	Amorce reverse (5' -> 3')
RttR	CCTGAACTTCCCAACGAATCC	GCTCGAGCGAACCTGG
GlmY	CGAATGACGCACAACAAGG	CCGGCTTTGTTATGGAATAAGG
RyfD	CAAGACGATCCGGTACGC	CGATCCAGACGCATAACTCC
GlmZ	GCCTTAGTGCCTCATAAACTCC	TGGGTGCTTCACTCAACG
CrpT	AATTGACGCAGTGGAGTAGC	TCGGACTTGGTGCTTTATCG
RyeB	GACCACCACGCTTTTTATTGACC	

Tableau 2 : Amorces pour les gènes ciblés en qRT-PCR

8. Transcription *in vitro*

Les transcriptions *in vitro* sont réalisées sur les vecteurs pTOPO des ARN connus et putatifs. Ces vecteurs ont été préalablement traités par le protocole d'extraction au phénol chaud pour enlever toute trace de RNase. Les vecteurs pTOPO sont normalisés à une concentration de 3 nM et utilisées avec 80nM d'ARN polymérase ($E\sigma^{70}$: E. coli RNA Polymerase Holoenzyme (Sigma-Saturated)(Épicentre) et $E\sigma^5$: E. coli RNA Polymerase Core Enzyme (Épicentre) complexé au facteur Sigma RpoS (purification HIS-TAG)) pour être transcrit *in vitro* dans un tampon K-glutamate (40nM HEPES, pH 8.0, 10nM $MgCl_2$, 200nM $C_5H_8KNO_4$, 4mM dithiothreitol, et 500 μ g/ml de BSA) en présence d' α - ^{32}P -UTP, 10min à 37°C (Lacour et al., 2003). Les produits de transcription sont déposés sur gel de séquençage dénaturant à 7% puis quantifiés.

9. Extension d'amorce

Les extensions d'amorces sont réalisées à partir 200ng d'ARN totaux extraits au phénol chaud (voir II.B.4) et de du kit SuperScript™II Reverse Transcriptase (Invitrogen) en utilisant une amorce (Sigma-Aldrich) spécifique au gène à reverse transcrire. Les produits de transcription sont déposés sur gel de séquençage dénaturant à 7% puis quantifiés.

C. Transformation bactérienne et clonage

1. Préparation des cellules compétentes et transformation des bactéries

La souche estensemencée à partir d'une culture de nuit et diluée au 100^{ème} dans 100ml de LB pour atteindre environ une DO = 0.6. Les bactéries sont rendues chimio-compétentes par le CaCl₂ ou électro-compétentes une série de lavage au glycérol 10% froid.

Cellules chimio-compétentes au CaCl₂

La culture bactérienne est centrifugée (10min, 4°C), le milieu de culture est intégralement éliminé pour resuspendre les bactéries dans une solution stérile de CaCl₂ glycérol (50mM, 15% glycérol) placées sur glace (30min) avant d'être centrifugées (10min, 4°C) puis concentrées dans volume final (équivalent à 1/50^{ème} du volume initiale) d'une solution de CaCl₂ glycérol. Les cellules sont aliquotées et peuvent être conservées à -80°C pour une utilisation ultérieure.

100µl de cellule sont mis en contact avec 50ng de plasmide à transformer, et incubés (15-30 min) sur la glace avant le choc thermique à 42°C (1min30). Les cellules sont immédiatement replacées dans la glace (5min) avant régénération par ajout d'un ml de LB et mise sous agitation (température adéquate, 200rpm, 1h). Les cellules sont étalées sur LA plus antibiotique de sélection du plasmide et cultivées sur la nuit à température adéquate.

Cellules électro-compétentes

La culture bactérienne est centrifugée (10min, 4°C), puis le milieu de culture est intégralement éliminé. Les cellules sont rendues compétentes par deux séries de lavage par une solution de glycérol 10% froid et de centrifugation (10min, 4°C). Les cellules sont finalement resuspendues dans un 1/50^{ème} du volume initial (de la solution de glycérol 10%). Les cellules sont aliquotées et peuvent être conservées à -80°C pour une utilisation ultérieure.

50µl de cellule sont mis en contact avec 50ng du plasmide à transformer ou d'ADN à recombinaison dans une cuvette d'électroporation froide (Eurogentec, Electroporation cuvettes 2 mm (short) yellow cap). Cet ADN a été préalablement dialysé sur membrane 0.2µm (Millipore MF-membrane filters) pour le dessaler. L'électroporateur (Biorad capacitance extender plus, puls controller plus, gene pulser II) est préréglé à 200ohms, 25µF et 2,5kV pour faire le choc électrique. Immédiatement après le choc les cellules sont placées dans un tube eppendorf contenant 1ml de LB, puis mises à régénérer (température adéquate, 200rpm, 1h). Les cellules sont étalées sur LA plus antibiotique de sélection du plasmide et cultivées sur la nuit à température adéquate.

2. Transduction bactérienne

3. **Les transductions des souches utilisées dans notre étude ont été réalisées selon le protocole donné par (Thomason et al., 2007). La bactérie porteuse de la mutation (insertion de la cassette de résistance à un antibiotique (Kanamycine) dans le gène ciblé) est lysée par le phage P1. Le lysat filtré est mis en contact avec la souche receveuse afin de réaliser la transduction de la mutation. La souche transduite est sélectionnée sur milieu solide contenant l'antibiotique spécifique à la mutation transduite (recombinaison homologue sur le chromosome de la cassette de résistance à l'antibiotique insérer dans le gène muté).**

4. Recombinaison plasmidique et chromosomique

La souche NM1100, qui possède sur son chromosome le système mini- λ red est utilisée pour la recombinaison plasmidique. En revanche, pour la recombinaison sur chromosome, nous utilisons la souche désirée transformée par le plasmide pSIM5 comportant le système mini- λ red dont l'induction est réalisée par un court passage à 42°C (Système thermosensible). Le système mini- λ red est une modification du système d'intégration du phage λ au chromosome bactérien, afin d'intégrer des séquences d'intérêts générés par exemple par PCR. Ce système contient les protéines nécessaires à la protection du fragment PCR électroporé dans les cellules et des protéines nécessaires à la recombinaison homologue au sein du chromosome (Serra-Moreno et al., 2006).

La souche NM1100 ou la souche transformée par le plasmide pSIM5 est ensemencée au 100^{ème} d'une préculture de nuit dans 20ml de LB à 32°C jusqu'à l'obtention d'une culture à une DO de 0.6. Les cellules sont ensuite placées dans un bain agitant à 42°C pendant 15min (activation des fonctions mini- λ red). Les cellules sont immédiatement placées et refroidies dans de l'eau glacée. Les cellules sont ensuite rendues compétentes en appliquant le protocole de cellules-électro-compétentes vu précédemment.

5. Clonage par restriction – ligation

D. Le vecteur et l'insert (dans un ratio molaire de 1/3) sont digérés ensemble dans un volume final de 20µl : contenant les enzymes de restriction (NEB), le tampon 1X (NEB), et la BSA 1X (optionnel, NEB), le mix est incubé à 37°C durant 1h. La digestion est inactivée par une exposition de 20min à 80°C ou purifiée par le kit NucleoSpin®Gel and PCR Clean-up. La ligation est réalisée avec 10µl de la digestion (inactivée ou purifiée) dans un mix contenant la T4 DNA ligase (Fermentas, 1µl), le tampon de ligation 10X (Fermentas, 2µl) pour un volume final qsp 20µl d'eau pure) La ligation est réalisée à température ambiante pendant 30min. La ligation peut être immédiatement transformée dans les cellules chimio-compétentes (5µl de ligation pour 100µl de cellules compétentes).

E. Outils informatiques

La génération des graphiques de cinétique d'expression des gènes est réalisée par des scripts sous e-Python développés par Valentin Zulkower (Thésard de l'équipe).

Les manipulations des séquences nucléotidiques (alignement, construction génomique et plasmidique, structure secondaire des ARNnc etc.) ont été réalisées à l'aide du logiciel CLC workbench.

Les quantifications des Northern Blot ont été réalisées sous le logiciel Image J selon les manipulations suggérées par <http://www.di.uq.edu.au/sparqimagejblots>

III. Résultats

A. Identification d'ARNnc régulés par RpoS par étude transcriptomique

1. Introduction et motivations

Au cours de cette dernière décennie, de nombreuses études ont permis l'identification d'ARNnc dans les génomes bactériens. Ces études reposent sur l'utilisation de la génétique et de logiciels de prédiction d'ARN, mais un nombre restreint d'ARNnc putatifs ont pu être validés par Northern Blot. L'existence d'environ deux cents entités prédites n'a pas encore été vérifiée physiquement par électrophorèse ; de plus, nous ignorons encore l'identité du facteur sigma responsable de leur transcription.

L'identification des facteurs de transcription induisant l'expression des ARNnc devient aujourd'hui accessible grâce au séquençage à haut débit (RNAseq) et du CHIP Seq. Un des objectifs de notre étude visait à mieux apprécier l'étendue du régulon du facteur RpoS en démontrant son implication dans la régulation de certains ARNnc. Pour concrétiser ce projet, nous avons tout d'abord utilisé les données de transcriptome présentes dans la littérature essentiellement générées par les puces à ADN et les logiciels de prédiction d'ARNnc. Nous nous sommes concentrés sur les données générées par les puces Affymetrix comparant l'expression génique d'une souche sauvage et du mutant *ΔrpoS* (Dong et al., 2008; Dong and Schellhorn, 2009). Après l'analyse ces données et suite aux informations complémentaires de l'auteur, nous nous sommes malheureusement aperçu que les sondes des puces de premières générations n'étaient absolument pas adaptées à l'étude de l'expression des ARN transcrits à partir des régions intergéniques.

Notre équipe est impliquée dans un projet d'actualisation du régulon de RpoS, avec d'autres laboratoires européens, qui visent à identifier les déterminants de la reconnaissance sélective des promoteurs par RpoS. Une des motivations majeures de ce projet dans laquelle s'inscrit ce projet consiste en l'identification des ARNnc régulés par RpoS. Nous avons ainsi participé à la validation de cibles de RpoS identifiées grâce aux techniques de RNA-seq et de CHIP-seq, et bénéficier de l'expertise de bio-informaticiens pour l'étude de ces nouveaux jeux de données.

2. Étude des données de puces Affymetrix

Les données de puces Affymetrix de (Dong et al., 2008; Dong and Schellhorn, 2009) comportent le niveau d'expression différentielle des gènes d'une souche WT et d'une souche *ΔrpoS* réalisée dans deux milieux distincts : en milieu riche (LB) et en milieu minium (M63), mais également en phase exponentielle ou phase stationnaire de croissance. Ces sondes couvrent l'ensemble du génome d'*E.coli* avec environ dix à vingt sondes par cible. Ces jeux de données regroupent le niveau d'expression de 4403 gènes et de 2883 régions intergéniques. Nous nous sommes principalement intéressés aux données des régions intergéniques qui sont principalement les zones contenant les ARNnc.

Afin d'étudier les jeux de données des régions intergéniques, nous avons fait un tri informatique pour conserver les régions intergéniques régulées positivement ou négativement par RpoS avec un seuil de 20% d'expression différentielle entre la souche référence WT et le mutant *ΔrpoS*. Nous avons ainsi retenu 402 des 2883 régions intergéniques (IG) du chromosome d'*E. coli*. Dans ces IGs, le logiciel de prédiction d'ARNnc décrit par Sridhar 2010 a détecté trois ARNnc putatifs contenus dans les IG 1027, 1105 et 1907 (selon l'annotation des IGs par Dong et al., 2009). Ces trois régions intergéniques ainsi que celle contenant trois ARNnc référencés dans la littérature (DsrA, C0465 et tpke11) ont été clonées dans un vecteur pTOPO afin de réaliser des tests phénotypiques tels que des tests de croissance, de coloration au Rouge Congo et de motilité sur gélose molle pour voir si la surexpression de l'entité affectait ces caractéristiques. Ces plasmides ont également été utilisés comme matrice pour réaliser des transcriptions *in vitro*. Seul l'ARN détecté dans l'IG1907 a retenu notre attention, car sa surexpression présentait un phénotype. Nous l'avons nommé « Djs », et lui avons consacré toute une partie de ce chapitre (III.A.5), car il représente une séquence redondante du génome d'*E. coli*. Les autres ARNs n'ont pas été validés en transcription *in vitro* et ne confèrent aucunes modifications des phénotypes étudiés. Après obtention d'informations complémentaires sur la construction des puces Affymetrix, nous nous sommes aperçus que les puces n'étaient pas adaptées pour l'étude des régions intergéniques. La répartition des sondes ne couvre ni l'ensemble de chaque région intergénique ni les deux brins du chromosome dans ces régions (Figure 9). Il est donc impossible d'étudier l'expression des ARN transcrits à partir des régions intergéniques au moyen de ces puces. Cette version initiale des puces n'était pas plus adaptée à l'étude des ARNnc connus, car peu d'entre eux possèdent des sondes spécifiques pour les détecter. Elle ne permet donc pas d'étudier les régions intergéniques, mais seulement les gènes codant des protéines ainsi que leur extrémité 5' et 3' UTR (non traduites). Nous avons alors décidé de changer de méthodologie et de jeux de données en réalisant un nouveau transcriptome par séquençage d'ARN.

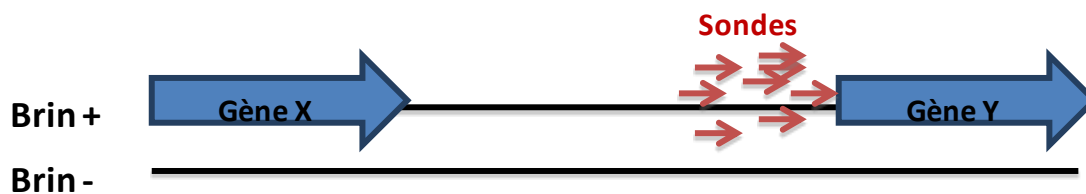


Figure 10 : Recouvrement de la région intergénique par les sondes Affymetrix.

Les gènes flanquant l'IG sont matérialisés par des flèches bleues : les petites flèches rouges correspondent aux sondes dessinées pour détecter les 3' et 5' UTR contenues dans les régions intergéniques. Ces sondes ne couvrent pas l'ensemble de la région intergénique ni les deux brins d'ADN et permettent par exemple dans cette illustration de mettre évidence une région 5'UTR pour le gène Y. La séquence des sondes de la puce est anti-sens de la séquence dont on veut détecter le transcrit.

3. RNA-seq, vers l'identification d'ARNnc régulés par RpoS

a) Conditions du RNA-seq et analyse sous EdgeR

Dans le but d'identifier de nouveaux ARN transcrits par RpoS, nous avons généré des échantillons d'ARN totaux issus d'une souche WT et d'une souche $\Delta rpoS$ (collection Keio). Ces souches ont été cultivées en milieu minimum glucosé (M9 Glucose 0,2%) jusqu'à l'entrée en phase stationnaire (DO=1,4) où RpoS est le plus actif. Les ARN totaux ont alors été extraits par la technique du phénol chaud. Nous avons choisi ces conditions de culture en espérant induire l'expression d'un plus grand nombre de gènes qu'en culture en milieu riche LB (comme par exemple la transcription des gènes permettant la synthèse des acides aminés qui sont déjà présents dans le milieu de culture). La constitution des bibliothèques de séquençage et le séquençage d'ARN lui-même (RNAseq) ont été réalisés par le laboratoire partenaire (CNR-ITB, Milan, Italie) qui nous a aussi fourni les données d'expression différentielle. L'analyse d'expression différentielle des gènes a d'abord été réalisée avec le logiciel EdgeR.

A.

Répartition des ARN séquencés			
Echantillon	ARNm	ARNr	ARN Intergénique
WT	92,0%	0,1%	7,8%
<i>ΔrpoS</i>	90,7%	1,3%	8,0%

B.

Echantillon	ARN détectées	ARN annotés communs	Gènes communs annotés détectés (4403)	ARN non annotés	Total ARN détectés
WT	4284	3725	84,60%	1124	4849
<i>ΔrpoS</i>	4290				

Tableau 3 : Analyse des ARN séquencés

A. Types d'ARN séquencés après ribo-déplétion. **B.** Couverture du génome. 85% des ARNs séquencés correspondent à l'expression de gènes annotés sur le chromosome d'*E. coli*.

L'ensemble des ARN séquencés dans nos deux extraits correspondant à 4849 annotations différentes du génome (Tableau 3). Mais en réalité l'analyse a porté sur environ 6000 annotations après ajout manuel des pseudo-gènes codants des ARNnc putatifs prédits par bio-informatique et répertoriés dans la base de données de l'Université de Keio au Japon (<http://rna.iab.keio.ac.jp/ECS/>).

Beaucoup de gènes annotés étaient différenciellement exprimés dans les deux souches, dont de nombreux gènes connus pour être régulés par RpoS. Nous avons donc appliqué une série de filtres sous Excel afin d'extraire une liste d'ARNnc putatifs fortement régulés par RpoS. Ces filtres visaient à conserver les ARN putatifs:

- d'une taille comprise entre 50 et 400 nucléotides, correspondant aux tailles des ARNnc décrits dans la littérature
- dont le ratio d'expression d'une souche à l'autre diffère d'un facteur supérieur à 2
- dont l'expression différentielle est strictement différente de celle des gènes flanquants l'ARNnc prédit, ceci afin d'éliminer les entités présentes dans un opéron.

Nous avons ainsi sélectionné 18 ARNnc prédits potentiellement régulés par RpoS. À ces derniers, nous avons rajouté 7 ARNnc connus dont le ratio d'expression d'une souche à l'autre diffère d'un facteur supérieur à 2. Les ARN sélectionnés (Tableau 4) ont fait l'objet d'une confirmation de leur expression différentielle (entre une souche WT et une souche *ΔrpoS*) par la technique d'hybridation en Dot Blot dont les résultats sont présentés dans la partie suivante.

	Nom	b-number	Position sur le chromosome		Orientation	Régulation par RpoS	Taille (nts)
ARNnc connus	RybC	b4585	506428	506511	+	-3,0	83
	RdIC	b4424	1269616	1269683	+	-2,5	67
	IsrC	b4435	2069339	2069542	+	-2,6	203
	SibA	b4436	2151333	2151475	+	-3,7	142
	RyhB	b4451	3578950	3579039	-	-2,3	89
	OxyS	b4458	4156308	4156417	-	3,3	109
	RyjA	b4459	4275950	4276089	-	2,4	139
ARNnc putatifs	C0042	∅	174939	175081	+	3,2	142
	NC014	∅	569952	570045	-	-5,4	93
	ECS074	∅	585114	585202	+	-3,3	88
	ECS198	∅	712067	712116	+	-3,2	49
	C0168	∅	712103	712177	-	-3,2	74
	C0198	∅	847276	847423	-	-2,4	147
	ECS066	∅	925327	925395	+	3,1	68
	I041	∅	931569	931643	-	2,9	74
	C0224	∅	950327	950399	-	5,7	72
	Contig_357	∅	1173101	1173217	-	4,3	116
	C0343	∅	1407387	1407461	+	4,3	74
	ECS081	∅	1848740	1848862	-	-3,7	122
	C0614	∅	2651474	2651560	-	-4,9	86
	ECS057	∅	3362087	3362160	-	-2,9	73
	ECS043	∅	3764296	3764361	-	-3,9	65
	C0950	∅	4104348	4104432	-	-4,5	84
	ECS168	∅	4434586	4434673	+	4,5	87
Contig_624	∅	4462877	4463002	-	2,9	125	

Tableau 4: ARNnc connus et putatifs fortement régulés par RpoS en milieu minimum M9.

Une valeur positive indique une activation par le facteur sigma RpoS ; à l'inverse, une valeur négative indique une répression par le facteur sigma RpoS.

b) Réalisation d'une mini Puce à ADNc (dot plot) de validation

Afin de sélectionner les ARNnc putatifs et connus à étudier en Northern Blot, nous avons créé une mini-puce à ADNc pour cribler nos différents candidats et vérifier leur régulation par RpoS. Pour fabriquer les sondes de notre mini puce à ADN, nous avons sélectionné des ARNnc connus et 8 ARNnc prédits, car 10 d'entre eux sont positionnés dans une région codante (CDS) du chromosome, sinon les sondes dessinées n'auraient pas permis de les distinguer des transcrits des gènes codants des protéines qu'ils chevauchent. Pour remédier à des problèmes d'amplification, d'ADN de petite taille, nous avons choisi des amorces en dehors des gènes codants les ARNnc ciblés pour générer des amplicons de plus grande taille. Les sondes générées par PCR pour mesurer le niveau d'expression des ARNnc connus et putatifs ont été greffées sur une membrane nylon. Elles servent à capturer les ADNc radio-marqués produits par transcription reverse à partir d'ARN totaux extraits des souches WT et mutante $\Delta rpoS$ au même stade de croissance que pour le RNAseq : DO=1.4 en milieu minimum M9 glucosé (0,2%). Les ARNnc connus et putatifs ainsi que des gènes contrôles de la puce sont listés dans le Tableau 5.

ARNnc candidats		CDS contrôles	ARNnc contrôles
ARN putatifs	ARNnc		
C0042	IscR	hdeA	AcrZ
C0224	OxyS	osmE	DsrA
C0343	RdIC	rplU	GcvB
C0614	RybC	rpoS	McaS
ECS066	RyjA	wrbA	MicC
ECS074		5S	OmrA
IO41			OmrB
NC014			RprA
			RyeA
			SibC

Tableau 5 : ARN ciblés par la puce à ADNc.

Les ARNnc ciblés par la puce à ADN, sont les ARNnc connus ou putatifs détectés par le RNAseq comme différentiellement exprimés entre une souche WT et $\Delta rpoS$ (ARNnc candidats). La catégorie « CDS contrôle » correspond à des gènes connus codant des protéines et régulés par RpoS (hdeA,osmE,wrbA), à un contrôle négatif (rplU, non régulé par RpoS) et le contrôle interne de normalisation, l'ARNr 5S. Les « ARNnc contrôles » sont les ARNnc connus, mais dont la régulation par RpoS n'est pas encore décrite.

L'analyse quantitative des résultats des hybridations dévoile un certain nombre de problèmes techniques. La variabilité des résultats et la présence de faux positifs nous ont indiqué que la puce n'était pas fiable et qu'il fallait davantage de mise au point notamment pour optimiser le dépôt, augmenter la spécificité des sondes et réduire le bruit de fond.

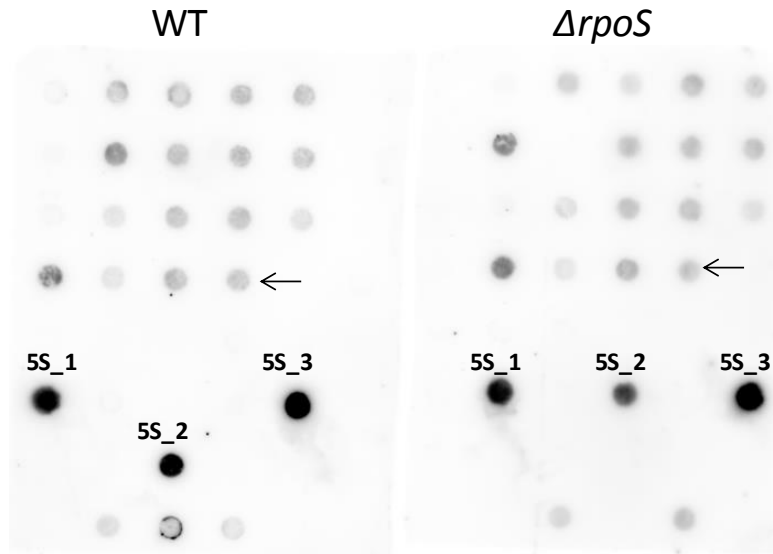


Figure 11 : Mini-puce à ADNc comparant l'expression différentielle des ARN putatifs et connus entre une souche sauvage (WT) et $\Delta rpoS$.

Le placement des spots des ARN putatifs et connus est identique entre la puce hybridant les ADNc issus de la reverse transcription des ARN totaux extraits d'une souche sauvage (WT) ou d'une souche mutée pour le gène *rpoS* ($\Delta rpoS$). Trois amplicons de l'ARNr 5S (5S_1, 5S_2, et 5S_3) sont disposés de façon asymétrique pour distinguer les deux puces. Les flèches désignent le spot de détection de l'ADNc du gène *rpoS*

La variabilité du signal d'hybridation pour le standard interne de normalisation (ARNr 5S) a été l'un des premiers problèmes soulevés. Le « spotage » multiple de ce standard sur les membranes suivantes a montré une forte variabilité du signal d'hybridation (5S_1, 5S_2 et 5S_3 (Figure 11)). Il était donc impossible de normaliser les valeurs des spots correspondant aux ARNc putatifs sur les deux membranes hybridées avec des ADNc provenant de la souche sauvage ou mutée pour *rpoS*.

D'autres problèmes ont également été repérés, comme la non-homogénéité des spots (Figure 11, visible sur le spot en dessous de celui du 5S_2 de la puce WT), malgré l'utilisation d'un système de dépôts dédié à cet usage, la variabilité des résultats d'une expérience à l'autre ou encore une régulation inverse de celle démontrée par le RNAseq. Mais le contrôle le plus troublant fut la détection du gène *rpoS* (Figure 11) à partir des extraits d'ARN totaux issus d'une souche $\Delta rpoS$, nous suggérant une contamination par de l'ADN génomique malgré un traitement à la DNase de nos échantillons d'ARN totaux avant la transcription inverse des ARNm en ADNc.

Nous avons essayé d'améliorer cette puce en changeant différents paramètres, tels que la quantité d'amplicons déposée, l'utilisation d'oligomères plus ou moins longs (30 à 70 nucléotides) à la place d'amplicons double brin permettant ainsi d'avoir un spot de capture spécifique pour la capture l'ARN. Malgré ces efforts, les résultats n'ont pas été grandement améliorés. La puce demandant une

trop longue période de mise au point pour devenir quantitative, nous avons décidé d'utiliser une autre approche de validation : la RT-qPCR.

c) Nouvelle analyse du RNA-seq et validation par qRT-PCR

Le logiciel EdgeR est assez performant pour l'analyse des transcriptomes classiques, mais il semble générer beaucoup de faux positifs lorsqu'il s'agit d'identifier des transcrits différentiellement exprimés parmi les ARN de petite taille. Nos collaborateurs bio-informaticiens de Milan ont réalisé une nouvelle analyse par l'utilisation des logiciels TruHmm (Li et al., 2013) et DetrProk (Toffano-Nioche et al., 2013) pour générer une nouvelle liste d'ARNnc candidats régulés par RpoS à partir du même jeu de données que celui utilisé sous le logiciel EdgeR. La liste des gènes codants des ARNnc connus et sélectionnés par les deux nouveaux logiciels (TruHmm et DetrProk) est comparée à celle générée par l'analyse sous edgeR dans le Tableau 6. Pendant que nous réalisons cette étude, un nouveau transcriptome par RNA-seq du régulon de RpoS a été décrit pour *Salmonella* (Levi-Meyrueis et al., 2014) qui renseigne sur l'expression différentielle d'ARNnc. Nous avons donc surligné en gris les ARNnc connus dont l'expression est régulée par RpoS d'après cette étude.

ARNnc	Logiciels d'analyses		
	detrProk	TruHmm	EdgeR
C0067	-	WT	-
C0362	Commun	-	-
C0465	-	Commun	-
C0719	-	Commun	-
CrpT	WT	WT	-
CsrB	WT	Commun	-
DsrA	-	Commun	-
GcvB	Commun	WT	-
GlmY	WT	Commun	-
GlmZ	-	WT	-
IpeX	-	WT	-
McaS	-	Commun	-
MgrR	WT	WT	-
MicA	WT	-	-
MicC	-	Commun	-
MicF	-	Commun	-
MybC (ChiX)	-	-	<i>ΔrpoS</i>
OmrA	WT	Commun	-
OmrB	WT	Commun	-
OxyS	-	Commun	WT
RdlC	-	<i>ΔrpoS</i>	<i>ΔrpoS</i>
RseX	-	Commun	-
RttR	WT	Commun	-
RybC	-	-	<i>ΔrpoS</i>
RydB	Commun	-	<i>ΔrpoS</i>
RyeA	WT	-	<i>ΔrpoS</i>
RyeB (SdsR)	WT	WT	-
RyeF	WT	-	-
RyfD	WT	-	-
RyhB	-	-	<i>ΔrpoS</i>
RyjA	-	WT	WT
ryjB	-	Commun	-
SgsR	-	-	<i>ΔrpoS</i>
SibA	-	Commun	<i>ΔrpoS</i>
SokE	WT	-	-
SokX	WT	Commun	-
SraA	-	Commun	-
SroD	-	Commun	-
SroH	-	Commun	-
SymR	-	WT	-

Tableau 6 : ARNnc détectés par les trois logiciels d'analyse du RNA-seq.

« Commun » signifie que l'ARN est détecté dans les deux souches WT et *ΔrpoS*. « WT » ou « *ΔrpoS* » signifie que l'ARNnc a été principalement séquencé dans la souche WT ou la souche mutante *ΔrpoS*. Les ARNnc grisés sont des ARNnc décrits comme régulés par le facteur RpoS chez *Salmonella* (Levi-Meyrueis et al., 2014).

Nous avons dessiné des amorces pour amplifier ARNnc connus identifiés comme différentiellement régulés par les deux nouveaux logiciels detrProk et TruHmm, mais également par les constructions reportrices GFP (chapitre III.B) réalisées de ce projet. Afin de valider la nouvelle analyse des données du RNAseq, nous avons essayé de confirmer la régulation par RpoS par qRT-PCR des ARNnc suivants : RttR, GlmY, RyfD, GlmZ, CrpT, RprA, MicC, DsrA, MgrR, GcvB et RyeB. Après avoir vérifié que

la linéarité des couples d'amorces était correcte pour l'ensemble des cibles (c'est-à-dire que parfaitement proportionnelle à la quantité de matrice), nous avons étudiés l'expression différentielle de certaines cibles. Comme le montre l'illustration de l'amplification PCR de l'ARN GlmY (Figure 12), il n'y a pas différence du niveau d'ARN GlmY détecté dans la souche sauvage ou mutée pour *rpoS* et qu'aucune différence du niveau d'ARNr 16S n'a été détectée entre les deux extraits (Figure 12). L'amplification des ARNnc RttR, GlmY, GlmZ, CrpT, RyfD, et RyeB ne montre pas non plus d'expression différentielle de ces gènes. Pourtant la régulation par RpoS des ARNs GlmY et RyeB a bien été confirmée par l'étude chez *Salmonella* (Levi-Meyrueis et al., 2014).

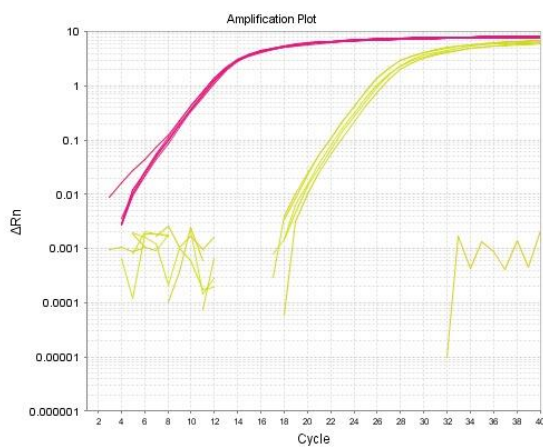


Figure 12 : Courbe d'amplification de l'ARNnc GlmY par qPCR après RT.

Amplification du standard interne 16S en rose, amplification de l'ARNnc GlmY en jaune. L'amplification sur les échantillons issus de la souche WT et $\Delta rpoS$ ne montre pas d'expression différentielle (même Ct et superposition des courbes).

d) Discussion et conclusion sur les résultats du RNA-seq

Il n'est pas exclu que nous découvriions des ARNnc régulés par RpoS et spécifiques à *E.coli* même si nos premiers résultats de validation par qRT-PCR ne sont pas encourageants. Nous avons essayé d'amplifier davantage de ces ARN, mais nous sommes confrontés à des problèmes de taille des cibles et de forte structuration des ARN. Il est donc probable que nous devons confirmer ces régulations par Northern blot. Nos collaborateurs doivent, de leur côté, essayer d'affiner leur analyse de détection des ARN non annotés sur le génome et de générer des quantifications précises de l'expression différentielle de ces ARNc dans les souches WT et $\Delta rpoS$. Enfin, il convient de noter que l'étude récemment publiée ne se focalise pas sur les prédictions d'ARN d'*E. coli* ce qui nous laisse espérer pouvoir identifier de nouveaux ARNnc et leur régulation. Il conviendra alors de trouver la cible de ces ARN dont nous aurons peut-être démontré l'existence (ARN putatifs).

4. Détection de site de liaison de RpoS par CHIP-seq

a) Objectif et principaux résultats du CHIP-seq

Parmi les 500 gènes régulés par RpoS, de nombreux gènes semblent être indirectement régulés (Lacour and Landini, 2004; Patten et al. 2004; Weber et al., 2005; Dong 2008 and 2009). Par exemple, RpoS régule certains cofacteurs de transcription tels que HNS, GadE, CsgD (etc...) qui modulent l'expression d'autres gènes. Des transcriptions *in vitro* réalisées à partir des holoenzymes $E\sigma^5$ et $E\sigma^{70}$ purifiées sur de l'ADN génomique extrait (Maciag et al., 2011) ont permis d'identifier une trentaine de gènes dont l'expression est induite par le facteur RpoS sans l'aide d'un cofacteur transcriptionnel. Pour compléter ce travail, une co-immunoprécipitation de RpoS avec de l'ADN génomique a été réalisée (ChIP-seq). Elle a permis d'identifier de nouveaux promoteurs directement liés par RpoS. Le ChIP-seq a en effet pour intérêt d'identifier les sites de fixation d'une protéine, donc les promoteurs liés directement par le facteur RpoS. 79 sites de fixation de RpoS ont été dénombrés, parmi ceux-ci trois régions intergéniques contenaient des ARNnc connus. Ma contribution à ce projet a été de mettre en évidence la régulation par RpoS de ces ARNnc par les techniques de Northern Blot et par des fusions transcriptionnelles.

Ces trois régions intergéniques comportent chacune deux gènes soit consécutifs sur un même brin d'ADN ou anti-sens, c'est-à-dire qu'ils se superposent sur les deux brins du locus du chromosome. La première IG contient deux ARNnc anti-sens l'un de l'autre, connue sous le nom de RyeA et RyeB (SdsR). La seconde contient un ARN codant pour une toxine et un ARNnc anti-sens (IbsC et SibC) (Fozo et al., 2008). La dernière région quant à elle contient deux ARNnc régulateurs co-orientés (OmrA et OmrB) qui régulent la traduction du facteur CsgD (Holmqvist et al., 2010).

Nous nous sommes intéressés à la régulation de ces entités, dont nous avons étudié la transcription et l'accumulation à l'aide de fusions transcriptionnelles et d'hybridations Northern Blot. Bien que le ChIP-seq ait été réalisé à 37°C, nous avons aussi complété ces études, en étudiant l'expression de ces ARNnc à 28°C, c'est-à-dire à une température où l'accumulation de RpoS est plus importante. Ainsi, à basse température de croissance, la régulation de ces ARNnc par RpoS devrait être plus importante.

b) Validation à l'aide de fusions transcriptionnelles

Pour étudier la transcription des ARNnc contenus dans les IG liées par RpoS, nous avons fusionné leur promoteur à la GFP et nous avons transformé ces fusions (plasmides) dans les souches WT et $\Delta rpoS$. Nous observons (Figure 14) une activation (d'un facteur deux au moins) de la transcription de l'ARNnc OmrA par le facteur RpoS aux deux températures. En revanche, la régulation des ARNnc OmrB, RyeA et SibC n'est pas significative voir inexistante. Nous avons également mesuré l'activité transcriptionnelle des gènes *ryeB(sdsR)* et *ibsC*, mais l'activité de ces fusions était trop faible, reflétant un faible niveau de transcription.

La température de croissance affecte peu l'activité transcriptionnelle de ces gènes et ne modifie pas la régulation d'OmrA par RpoS. Il est intéressant de noter que les ARNnc OmrA et B régulent les mêmes cibles cellulaires, ce qui est probablement lié à leur forte similarité de séquence (68%), mais ne sont pourtant pas régulés de la même manière par RpoS. Cette différence de régulation d'expression peut s'expliquer par une divergence de séquence de la région promotrice de ces deux gènes, illustrée ci-dessous (Figure 13).



Figure 13 : Alignement des régions promotrices des gènes codant les ARNnc OmrA et OmrB

Alignement de la région promotrice des ARNnc OmrA (OmrA prom) et OmrB (OmrB prom) de -40 à +1 par rapport au site d'initiation de la transcription (+1). « Conservation » correspond aux nucléotides communs aux deux régions promotrices indiquées par une barre pleine (100%) et du sur lignage rouge du nucléotide. Le « -35 » et le « -10 » correspondent respectivement aux boîtes -35 et -10 (I.B.2).

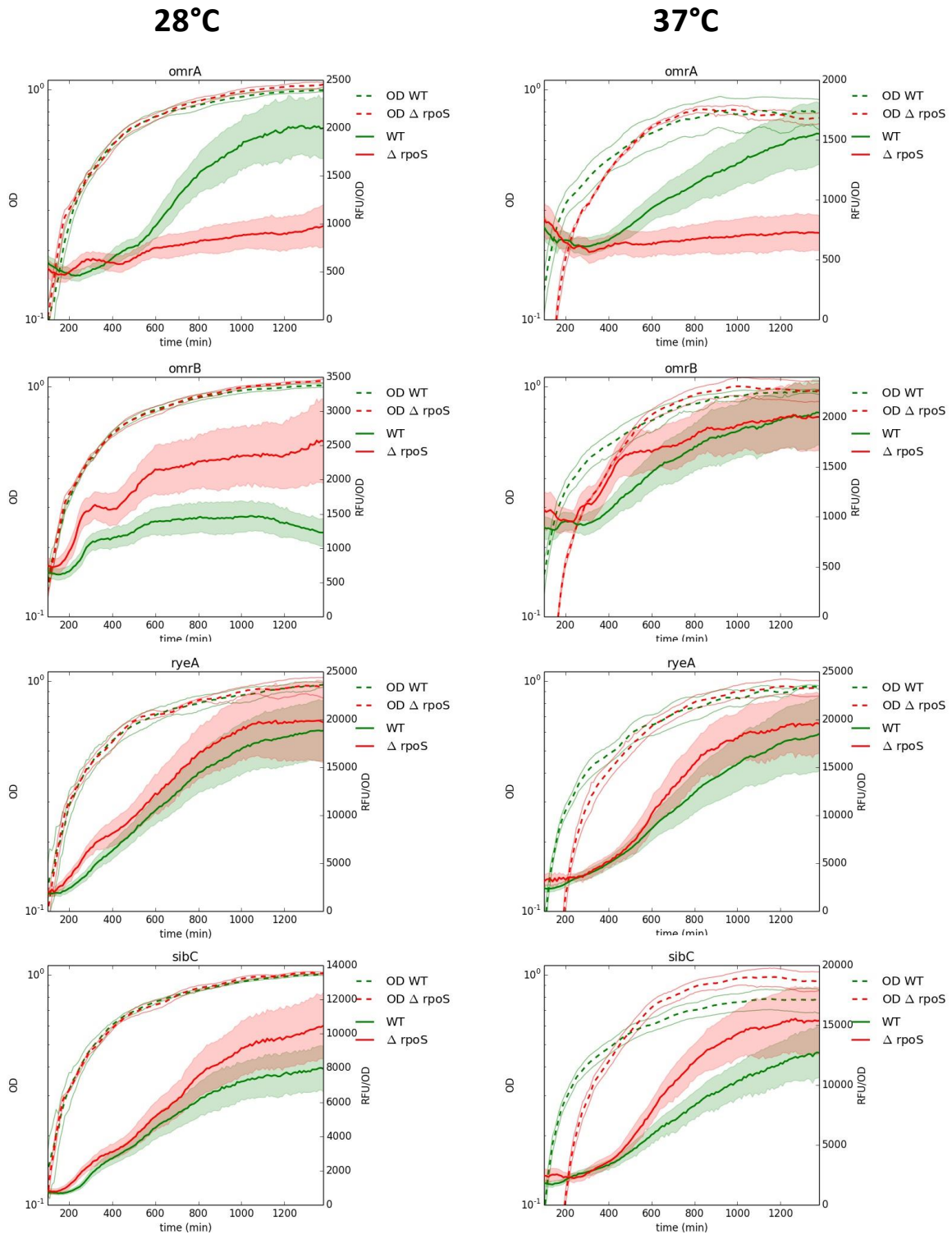


Figure 14 : Profils transcriptionnels des ARNnc ciblés par le ChIP-seq.

La fluorescence relative dans la souche sauvage est présentée en traits verts, celle dans le mutant *rpoS* en traits rouges. La zone colorée indique l'écart à la moyenne de triplicatas techniques. Les courbes en pointillé montrent la croissance de chaque souche. La moyenne des triplicatas d'une expérience typique est montrée aux deux températures de croissance : 28°C colonne de gauche et 37°C colonne de droite. Le nom de chaque gène étudié est précisé au-dessus de chaque graphique de suivi cinétique.

c) Validation par Northern Blot

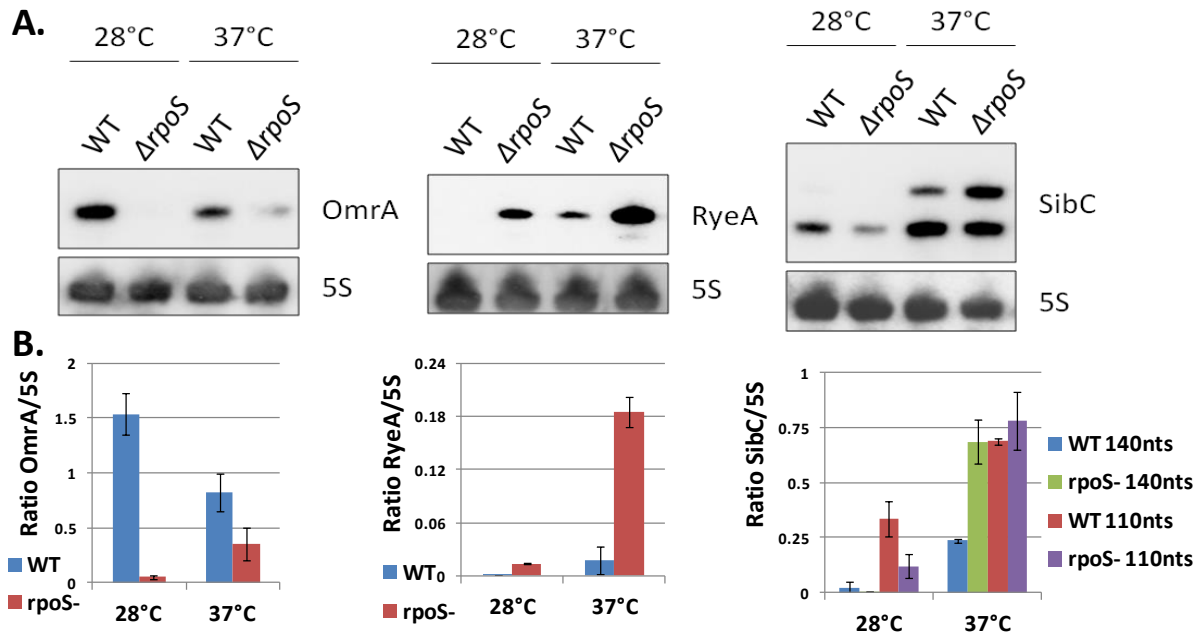


Figure 15 : Northern Blot et quantification de l'accumulation des ARNnc OmrA, RyeA et SibC.

(A.) Les ARN totaux ont été extraits à l'entrée de la phase stationnaire de croissance (DO=3) par la technique d'extraction des ARN au phénol chaud. 5µg d'ARN totaux ont été déposés par piste. (B) Quantification des ARN (duplicata biologiques et techniques).

La régulation positive de la transcription d'OmrA par RpoS a pu être confirmée par Northern blot (Figure 15), puisque l'ARNnc OmrA est préférentiellement accumulé dans la souche sauvage et ne s'accumule pas en l'absence de RpoS. Cependant, la régulation observée en Northern semble plus forte qu'avec la fusion transcriptionnelle ce qui suggère une possible régulation de la stabilité en plus de celle de l'initiation de la transcription d'*omrA*. L'accumulation de OmrA est aussi plus importante à 28°C qu'à 37°C ce qui pourrait être corrélé à l'accumulation préférentielle de RpoS aux températures inférieures à 30°C. Peut-être que cette observation sera utile pour expliquer la thermorégulation des curli dans le chapitre suivant.

La régulation négative de RyeA par RpoS est démontrée par le Northern Blot (Figure 15), à l'inverse de la fusion transcriptionnelle utilisant le promoteur de *ryeA*, le Northern Blot montre une plus forte accumulation de l'ARNnc RyeA dans le mutant *ΔrpoS* ce qui témoigne d'une régulation négative par le facteur sigma. Cependant, il se pourrait que celle-ci ne se fasse pas au niveau de l'initiation de sa transcription. Ceci suggère une régulation post-transcriptionnelle de cet ARNnc par l'activation par exemple d'une RNase régulée par le facteur RpoS. Nous n'avons pas pu étudier l'expression de l'autre gène contenu dans cet IG, RyeB, car l'hybridation Northern n'a produit aucun signal, cet ARNnc RyeB ne semble donc pas être exprimé ou trop faiblement dans nos conditions de cultures.

Néanmoins, ce gène est décrit comme régulé par RpoS sous le nom de SdsR (Frohlich et al., 2012). C'est donc probablement la régulation de RyeB plutôt que celle de RyeA qui explique la liaison (détecté en CHIP-seq) de RpoS à ce locus.

La régulation de SibC est complexe à analyser, car il y a deux transcrits et des régulations opposées selon la température. La sonde Northern dirigée contre SibC produit deux « bandes » à 37°C de 140 et 110 nt, si l'on en croit la littérature : en effet, la présence de deux transcrits SibC en Northern a déjà été observée par Han et ses co-auteurs (Han et al., 2010) qui suggéraient un processus de clivage de l'ARNnc SibC lorsqu'il est fortement exprimé. Mais nous n'avons pas la preuve dans nos expériences que ces deux transcrits sont issus d'un seul et même transcrits ; ils pourraient être transcrits à partir de deux promoteurs distincts dont l'un ne serait actif qu'à 37°C. Nous ne comprenons pas bien comment la régulation positive par RpoS à 28°C devient négative à 37°C. Notre hypothèse est que le transcrit complémentaire (lbsC) produit à partir de l'autre brin est exprimé à 37°C, de façon RpoS-dépendante, et que sa présence déstabilise l'ARNnc Sibc.

d) Discussion et conclusion des validations du CHIP-seq

Nous avons voulu expliquer la liaison de RpoS à trois loci codants des ARNnc identifiés par CHIP-seq et nous avons donc voulu vérifier leur régulation. Nous avons identifié une nouvelle régulation importante : l'activation OmrA par RpoS, qui pourrait renforcer le rôle du facteur sigma dans le contrôle de l'expression des curli. Cette régulation vient d'être confirmée dans une autre entérobactérie par l'étude du régulon de RpoS chez *Salmonella* (Levi-Meyrueis et al., 2014). Les autres loci liés par RpoS comportent des gènes codant pour les ARNnc RyeA et SibC dont la transcription ne semble pas régulée par RpoS. Nos hybridations Northern montrent toutefois une régulation négative par RpoS. Notre hypothèse est que RpoS pourrait stimuler la transcription des gènes anti-sens RyeB et lbsC. La régulation positive de l'expression de RyeB ayant déjà été démontrée (Frohlich et al., 2012). Ces ARN anti-sens s'apparient intégralement sur RyeA et SibC respectivement, créant ainsi des duplex ARN-ARN pouvant être la cible de la RNase III (Court et al., 2013; Doetsch et al., 2013). Cependant, l'accumulation de SdsR et d'lbsC n'a pas pu être détectée en Northern Blot malgré des dépôts d'ARN totaux de 15µg comparés aux 5µg déposés pour détecter les autres ARN. Ainsi, nous n'avons pas pu démontrer et expliquer la régulation négative par RpoS des transcrits SibC (de 140nt présents à 37°C) et de l'ARNnc RyeA par l'activation de l'expression de leur anti-sens par RpoS. Enfin, nous supposons que l'accumulation préférentielle de RpoS à 28°C explique

la plus forte expression d'OmrA à cette température. Elle n'explique en revanche pas pourquoi la régulation négative de RyeA est plus forte à 37°C.

5. Identification d'un nouvel ARNnc : Djs

a) Identification in silico

Djs a été identifié grâce à l'analyse des données de puce Affymetrix (Dong et al., 2008) et de l'utilisation du programme de prédiction d'ARNnc de (Sridhar et al., 2010). En recherchant une séquence spécifique pour détecter l'ARN putatif contenu dans l'IG1907 par Northern Blot, nous nous sommes aperçus que cette séquence était répétée au sein du génome de la bactérie.

L'ARN putatif d'une taille estimée à une centaine de nucléotides possède quatre homologues parfaits dont la séquence ne diffère que très peu (sept nucléotides de différence pour la copie la moins similaire). Cet ARN possède une séquence palindromique au milieu de l'ARN qui explique son repliement en une grande épingle à cheveu, structure secondaire capable de stabiliser l'ARNnc putatif. En nous concentrant sur ce motif palindromique, nous avons découvert une dizaine de séquences homologues supplémentaires dont la séquence, de part et d'autre du palindrome diverge davantage (Figure 16). Malgré cette variabilité de séquence, nous avons pu montrer que toute divergence au sein du palindrome était compensée pour maintenir la structure secondaire de l'ARN putatif. L'homologie de séquence entre copie varie de 47% à 99%, 13 copies ayant plus de 65% d'homologie et 4 plus de 95% d'homologie.

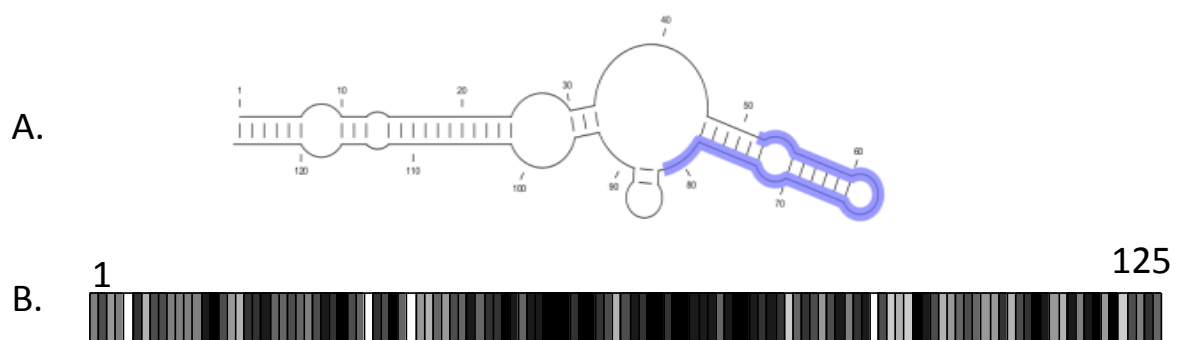


Figure 16 : Homologie de séquence des 15 copies de Djs.

(A.) Structure de l'ARN Djs prédite par le logiciel CLCworkbench, la zone bleutée sur la séquence correspond à la zone d'hybridation de nos sondes de Northern dirigée contre cet ARN (Figure 19). **(B.)** Conservation des 125 nucléotides de l'ARN Djs. Chaque case représente un nucléotide de l'ARN putatif. Le dégradé du blanc vers le noir correspond au pourcentage d'homologie entre les 15 homologues pour chaque position: le blanc signifie un faible degré d'homologie et noir une homologie élevée. La région centrale où l'homologie est la plus importante correspond à la séquence palindromique.

La répartition des homologues de l'ARN Djs sur le génome d'*E. coli* est présentée dans la Figure 17. Cette répartition des copies de l'ARN Djs est relativement homogène et présente une curieuse symétrie par rapport à l'axe entre origine (vers 500 méga bases) et le terminus de réplication (vers 2700 méga bases). Nous avons également recherché cette séquence dans d'autres organismes et avons constaté qu'elle était présente en nombre variable dans de nombreuses espèces bactériennes de la famille des *Enterobacteriaceae* qui compte beaucoup d'espèces de bactéries pathogènes (Tableau 7). Le nombre variable de copies de cette entité au sein des génomes suggère une duplication de la séquence dans certaines espèces.

La conservation et la duplication de cet élément génétique suggèrent que cet élément génétique a une fonction importante pour la physiologie de cette famille de bactéries. Son nombre variable pourrait suggérer une fonction de régulation de l'expression des gènes par titration, par exemple s'il interagit avec une ou des protéines régulatrices liant l'ADN ou l'ARN. La structure en épingle de cet élément rappelle celle d'un terminateur de transcription qui pourrait jouer le rôle d'atténuateur de la transcription de certains opérons s'il n'est pas transcrit en ARN individuel. S'il l'est, l'ARN résultant serait sûrement très stable au vu de sa forte structuration et pourrait jouer un rôle d'anti-sens des ARN transcrits à partir du brin opposé.

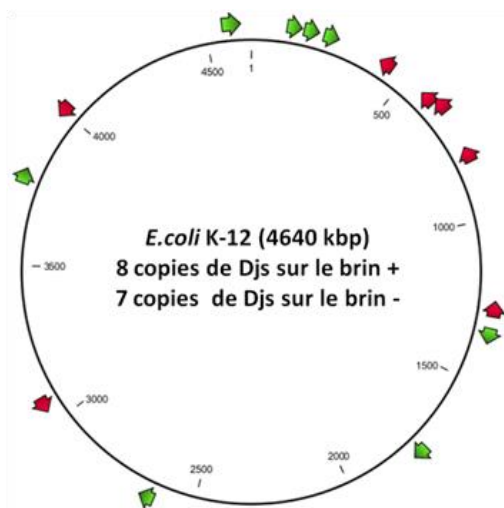


Figure 17 : Distribution des différentes copies de Djs sur le chromosome d'*E. coli* K-12

Les copies en vert sont codées sur le brin +, celles en rouge sur le brin - du chromosome circulaire.

Souche bactérienne	Nombre de copies
<i>Citrobacter sp.</i>	8
<i>Cronobacter sp.</i>	2
<i>Enterobacter sp.</i>	8
<i>Erwinia sp.</i>	de 1 à 48
<i>Escherichia sp.</i>	de 3 à 15
<i>Klebsillia sp.</i>	2
<i>Photorhabdus sp.</i>	1
<i>Salmonella sp.</i>	de 3 à 10
<i>Serratia sp.</i>	5
<i>Shigella sp.</i>	de 2 à 7
<i>Yersinia sp.</i>	de 8 à 28

Tableau 7 : Distribution du nombre de copies de Djs au sein des souches bactériennes de la famille des *Enterobacteriaceae*.

La recherche d'homologues de la séquence palindromique (partie centrale de 60 nt de l'ARN Djs) avec l'outil nblast sur le site NCBI : <http://blast.st-va.ncbi.nlm.nih.gov/Blast.cgi> a permis de détecter un nombre variable de copies de Djs dans les espèces bactériennes de la famille des *Enterobacteriaceae*. Ce nombre de copies varie de 1 à 48 en fonction des espèces, mais peut également varier au sein d'une même espèce.

b) Phénotypes conférés par la surexpression de Djs.

Afin de vérifier l'importance de l'entité répétée sur le génome d'*E.coli*, nous avons construit un vecteur de surexpression de Djs et testé son effet sur des phénotypes simples. Nous avons pu voir que la surexpression de Djs provoquait un retard de croissance sans pour autant modifier le taux spécifique de croissance de la souche en phase exponentielle (Figure 18 A). De façon intéressante, le retard de croissance induit par la surexpression de Djs n'est pas observé dans une souche $\Delta rpoS$, ce qui tend à prouver que le facteur sigma RpoS régule l'expression de Djs à partir du plasmide pTOPO. Sa surexpression affecte aussi la motilité (non montré) et l'expression des curli (Figure 18 B). Il semble donc que cette surexpression perturbe la cellule sans pour autant lui être fatale, mais que les régulations potentiellement perturbées ne le soient pas spécifiquement.

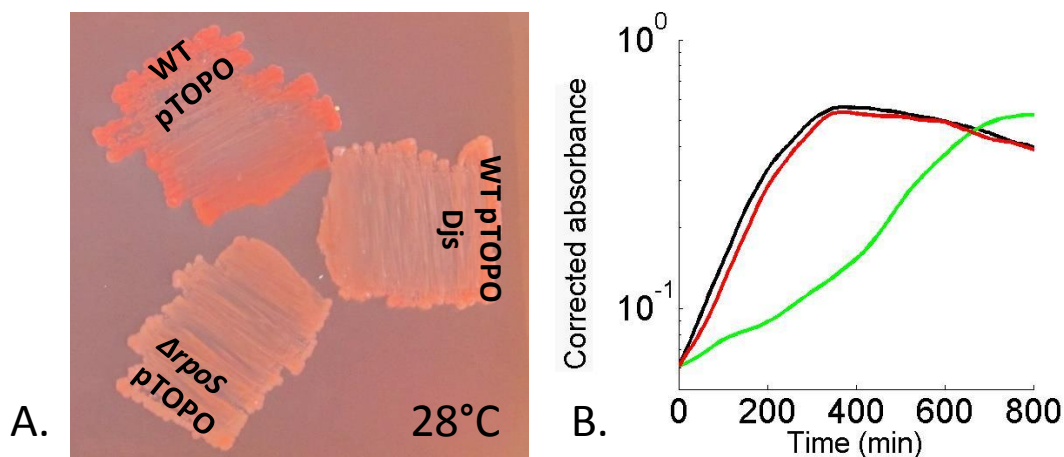


Figure 18 : La surexpression de l'ARN putatif Djs affecte la liaison au Rouge Congo.

(A.) Les souches sauvages (WT) et *ΔrpoS* transformées par le plasmide pTOPO (contrôle) ou transformée par le plasmide pTOPO contenant la région intergénique de Djs (pTOPO Djs) sont striées sur milieu riche LB-agar (LA) supplémenté d'une solution de Rouge Congo à la concentration finale de 0.004%. La boîte est incubée 48h à la température de 28°C La présence de curli est visible à 28°C dans une souche WT, car elle induit une coloration rouge de la strie de croissance bactérienne (voir Chapitre III.B.1). **(B.)** Cinétique de croissance bactérienne, en noir la souche WT (transformée par le plasmide pTOPO contrôle), en rouge la souche *ΔrpoS* (transformée par le plasmide pTOPO contrôle) et la courbe verte la souche WT contenant le vecteur de surexpression de Djs (28°C).

c) Détection en Northern Blot et extension d'amorce de Djs

Afin de démontrer l'existence physique de l'ARN Djs, c'est-à-dire l'expression d'au moins une des 15 copies sous forme d'un petit ARN, nous avons réalisé un Northern blot avec une sonde détectant la partie centrale et conservée de Djs (Figure 16 A). Nous avons pu détecter un ARN de courte taille (Figure 19), environ 120 nucléotides dans les bactéries cultivées en milieu minimum M9 et en milieu riche LB. Notre expérience ne nous permet pas de conclure sur son expression différentielle dans les souches sauvages et *ΔrpoS* et sa taille équivalente à celle de l'ARNr 5S nous empêche d'utiliser la sonde 5S comme témoin de charge pour les quantifications. Si nous avons prouvé l'existence physique de cet ARN, nous ne savons pas quelle copie est exprimée. D'autre part, il conviendrait de travailler dans un mutant dépourvu des 15 copies de Djs pour valider la spécificité de la sonde, ce qui est inenvisageable. En revanche nous pourrions tester différentes sondes biotinyllées complémentaires des régions variables de Djs, mais leur coût est prohibitif. Seule une analyse appropriée des données du RNAseq devrait permettre d'identifier quelles copies sont exprimées. Mais pour l'instant, nos collègues bio-informaticiens sont noyés dans les déductions de bruits de fond du séquençage et le tri des courtes séquences (« reads ») à « mapper » sur des parties du

génomique non annoté. Nous nous heurtons donc à des limites de technique pour pouvoir répondre à cette question. Cinq des séquences homologues de Djs ont été identifiées dans notre RNAseq dont une pourrait être différentiellement exprimée dans un mutant $\Delta rpoS$. Une fois tous les candidats identifiés, nous essayerons de valider l'existence des copies exprimées et leur régulation par RpoS par qRT-PCR.

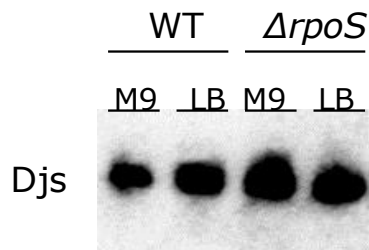


Figure 19 : Northern blot visant à détecter l'ARN putatif Djs

Les ARN totaux des souches WT et $\Delta rpoS$ ont été extraits à l'entrée de la phase stationnaire de croissance : (DO=1.4 pour le milieu minimum (M9 glucosé 0,2%) et DO=3 pour le milieu riche (LB) à 28°C. L'ARN détecté est approximé à une taille de 125 nts.

Nous avons aussi cloné 5 copies de Djs sur un vecteur pour tenter de transcrire certaine copie de Djs *in vitro* à l'aide des holoenzymes purifiées $E\sigma^S$ et $E\sigma^{70}$. Nous comptons ainsi mettre en évidence une reconnaissance préférentielle d'un promoteur par l'une des formes de polymérase : associée à RpoS. Malheureusement, aucune des 5 IGs utilisées comme matrice ne semble coder un ARN transcrit *in vitro* sans autre facteur de transcription. Pour valider l'approche expérimentale, nous avons pris comme contrôle positif la région intergénique codant l'ARNnc DsrA que nous avons pu détecter. Ce résultat suggère que la sonde Northern pourrait cross-hybrider une autre espèce d'ARN et que nous devrions répéter l'expérience de transcription *in vitro* avec les 10 autres copies de Djs.

Nous avons aussi réalisé une extension d'amorce pour voir si l'élément était transcrit comme petit ARN individuellement ou en tant que composante d'un ARN plus grand, 3' ou 5'UTR, ou composante d'un polycistron. Pour cela, nous avons utilisé des amorces sens et anti-sens chevauchant le palindrome central de Djs (Figure 16 A). Les résultats de cette expérience sont présentés dans la Figure 20. Curieusement, des produits d'extension n'ont pu être détectés qu'avec une des deux sondes (1907F). Ils mettent en évidence un ARN de courte taille et d'autres plus grands dont la taille demeure indéterminée. Cette dernière expérience démontre donc que l'ARN Djs pourrait bien exister et que plusieurs copies sont transcrites sous forme d'ARN plus longs. Compte tenu du fait que la sonde est complémentaire du centre de Djs, on peut imaginer montrer la production d'extensions de 120, 180, 300 et 400 nucléotides plus celle d'un très long transcrit (haut du gel Figure 20). Il conviendrait à présent de séquencer ces produits pour les identifier.

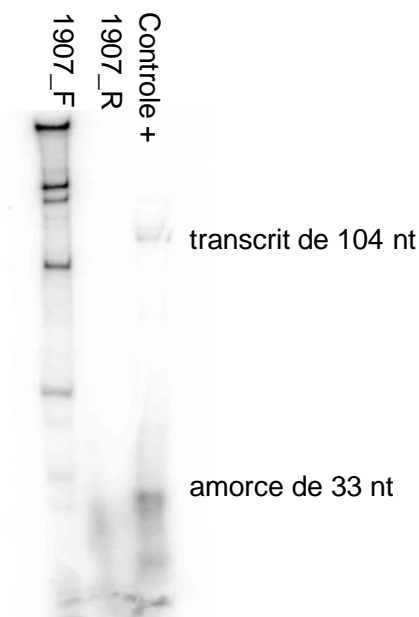


Figure 20 : Extension d’amorces visant à démontrer l’existence de Djs.

Extension d’amorce à partir des amorces 1907_F et 1907_R (complémentarité en bleu dans la Figure 16 A) sur les échantillons d’ARN totaux utilisés en RNAseq (M9 glucosé 0,2%, DO=1.4, souche sauvage). Les produits radiomarqués (avec de l’ α^{32} ATP) de cette extension d’amorces sont déposés sur gel de séquençage dénaturant à 7%.

6. Conclusions du chapitre

L’analyse d’expression différentielle d’ARNnc s’est avérée laborieuse et demande de nombreuses mises au point techniques. Si la technique du RNA-seq est parfaitement adaptée à cet objectif, son analyse est très longue et ses différents processus d’analyse (« pipeline ») peuvent exclure nombre de données intéressantes pour un biologiste. Pour identifier de nouveaux ARNnc, il faut des pipelines personnalisés et une excellente couverture du génome (ici 50 millions de reads). Nous avons été devancés pour la diffusion de notre seul résultat incontestable : la régulation positive de l’expression d’OmrA par RpoS. Nous espérons à présent confirmer la régulation de l’expression d’autres ARNnc par RpoS et mettre en évidence des divergences entre espèces apparentées. Notre mauvaise piste initiale, les données des premières puces Affymetrix, nous a permis de mettre en évidence une séquence répétée sur le génome d’*E. coli* qui pourrait avoir un rôle de régulation de l’expression génique jamais décrit chez les bactéries.

B. Étude de la régulation de l'expression des curli par RpoS et les ARNnc

1. Introduction

Nous avons observé au laboratoire que la coloration au Rouge Congo des colonies d'*E. coli* formées sur milieu LB est liée à l'expression des curli. L'absence des gènes structuraux (*csgBA*), de leur activateur CsgD ou du facteur sigma RpoS entraîne une perte de la capacité de la souche à pouvoir lier le Rouge Congo entraînant une décoloration de la colonie sur boîte (Figure 21). Ce phénotype de liaison du Rouge Congo s'exprime uniquement à des températures inférieures à 30°C. Lorsque les bactéries sont incubées, à une température supérieure à 30°C, par exemple à sa température optimale de croissance 37°C, la différence de phénotype entre une souche sauvage et les mutants $\Delta csgD$ et $\Delta csgBA$ ne sont plus visibles (Figure 21). On remarque même qu'à 37°C, ces mutants présentent le même degré de coloration que la souche sauvage, ce qui signifie que la faible coloration observée (comparativement à l'incubation à 28°C) n'est peut-être plus dépendante de la présence de curli à la surface des cellules.

Nous nous sommes intéressés à l'expression du régulateur CsgD et des protéines structurales CsgB et CsgA dans la souche BW25113 d'*E. coli*, nouveau modèle d'étude puisque l'on dispose notamment d'une collection de mutants systématiques de chaque gène non essentiel (émanant de l'université de Keio, Baba *et al.*, 2006). Leur expression a été étudiée à l'aide de fusions transcriptionnelles et traductionnelles dans différents contextes génétiques. Contrairement à certains dogmes trouvés dans la littérature, nous observons la transcription du gène *csgD* à 37°C dans des données de puce à ADN (de l'équipe de Herb Schellhorn) et nous avons pu confirmer cette induction grâce à notre fusion transcriptionnelle. La Figure 22 montre que malgré la transcription du gène *csgD* à 37°C, nous n'observons pas la transcription du gène *csgB*. Ceci suggère soit une répression transcriptionnelle des gènes *csgBA* à 37°C ou une répression de la traduction de l'ARNm *csgD* après sa transcription à cette température. Hors la plupart des études sur la régulation de la traduction de CsgD ont été faites à 28°C certainement sur le postulat que CsgD n'est transcrit qu'à cette température, ce qui n'est pas le cas, tout au moins dans la souche BW25113.

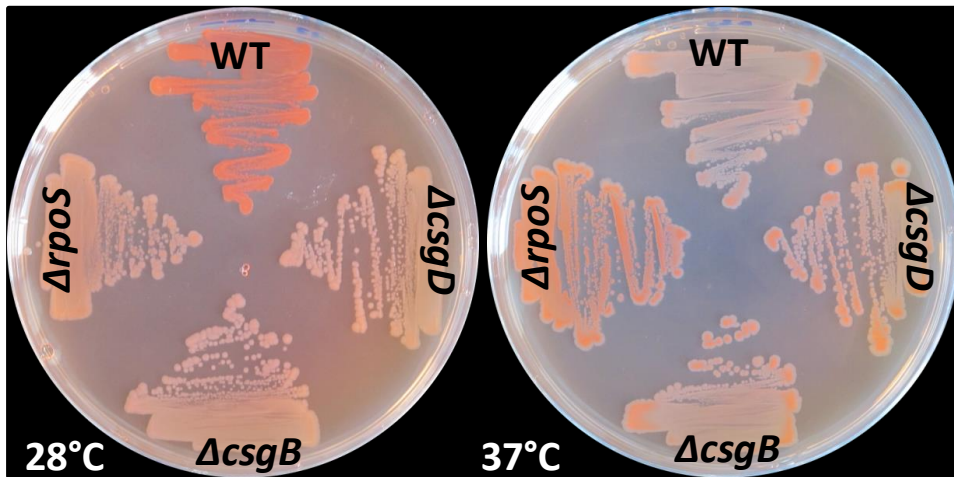


Figure 21: Phénotype de liaison au Rouge Congo à 28°C et 37°C.

Les souches sauvages (WT), $\Delta csgD$, $\Delta csgB$ et $\Delta rpoS$ sont striées sur milieu riche LB-agar (LA) supplémenté d'une solution de Rouge Congo à une concentration finale de 0,004%. Les boîtes sont incubées 48h aux températures de 28°C et 37°C. La présence de curli est visible à 28°C dans une souche WT, car elle induit une coloration rouge de la strie de croissance bactérienne. À 37°C les souches WT, $\Delta csgD$ et $\Delta csgB$ ont la même coloration faible, qui semble curli indépendante. Au contraire, la souche $\Delta rpoS$ présente une plus forte coloration que la souche WT à 37°C. Cette observation pourrait suggérer que le contrôle par RpoS de l'expression des gènes *csg* est différent aux deux températures, ou que RpoS réprime l'expression d'un autre déterminant génique de la coloration.

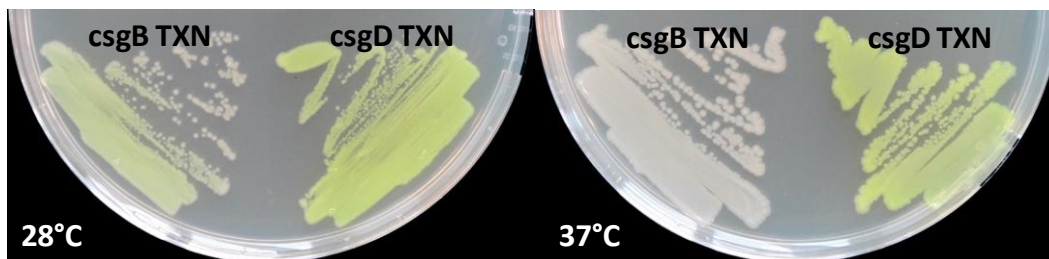


Figure 22: Expression des opérons *csg* à 28°C et 37°C.

La souche WT, transformée par les plasmides porteurs d'une fusion transcriptionnelle du promoteur de *csgD* ou *csgB* à la green fluorescent protein (GFP) ont été striées sur milieu riche (LB) et incubées 48h à une température de 28°C ou 37°C. La photo montre que l'expression de ces gènes est visible à l'œil nu.

La régulation de l'expression des curli est un sujet relativement documenté dans la littérature (Figure 23) puisque leur expression détermine la formation d'un biofilm (Prigent-Combaret et al., 2001). De nombreuses études ont révélé un réseau génique complexe déterminant le « choix » entre l'adhésion ou la motilité pour la bactérie ; il est présenté dans la Figure 23. Mais aucune étude n'a été réalisée à ce jour pour tenter de hiérarchiser les interactions entre ces facteurs génétiques essentiels pour l'expression des deux opérons *csg*. Nous nous sommes particulièrement appliqués à revisiter le rôle des ARNnc et de RpoS dans cette régulation, afin de déterminer les interactions manquantes et majeures entre les différents constituants du réseau de régulation. Si de nombreux

ARNnc de ce réseau répriment la traduction du régulateur CsgD, la capacité de RpoS à réguler l'expression de ces ARNnc n'a pas été étudiée. Enfin, nous avons aussi essayé de déterminer si certains des régulateurs connus de ce réseau sont responsables de la thermorégulation de l'expression des curli. En effet, le phénotype de la souche mutée pour le facteur sigma RpoS est assez intrigant puisqu'il change selon la température (Figure 21). Pourquoi le mutant $\Delta rpoS$ favorise la liaison au Rouge Congo à 37°C alors que sa présence détermine le phénotype à 28°C vraisemblablement en induisant l'expression des curli? Il est possible que la coloration des colonies sur Rouge Congo à 37°C ne soit plus déterminée par la présence de curli. Il nous est donc apparu indispensable de revisiter le rôle de RpoS pour la régulation de l'expression des curli (Figure 23). En somme, à travers cette étude, nous avons tenté de répondre aux questions suivantes:

RpoS régule-t-il directement la transcription des gènes *csg* ?

RpoS régule-t-il les ARNnc régulant la traduction de CsgD ?

Quels sont les éléments clés de l'expression des curli ?

Comment les gènes *csgBA* sont-ils thermo-régulés ?

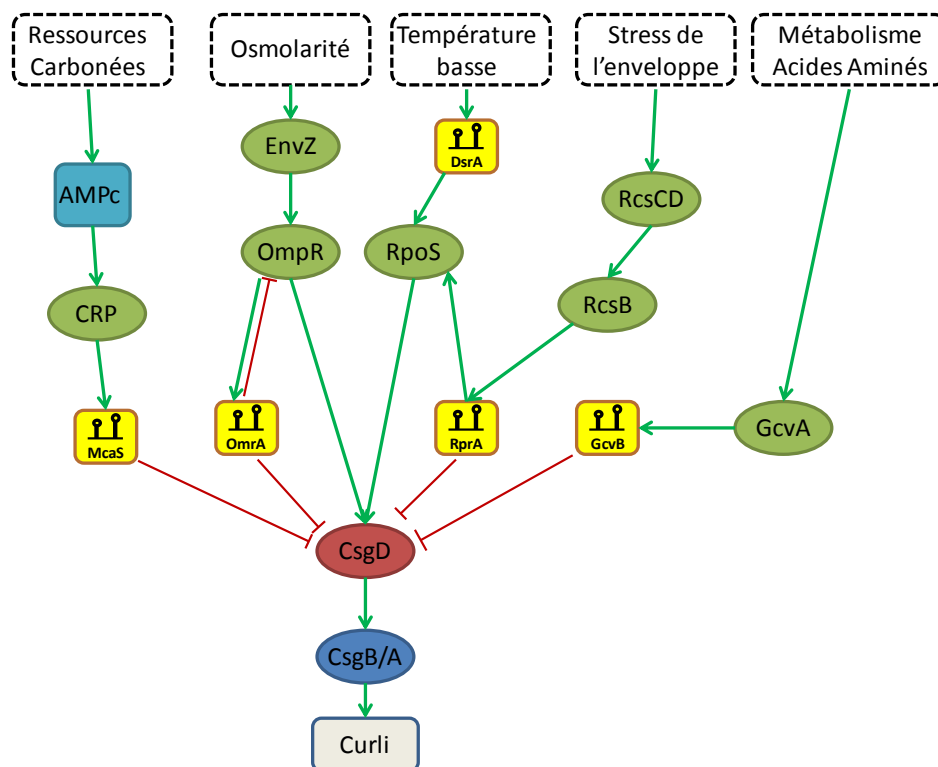


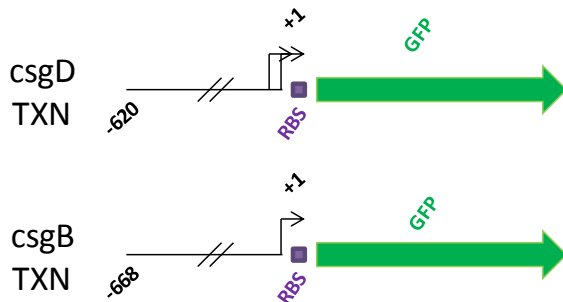
Figure 23: Réseau de régulation contrôlant l'expression des curli chez *E. coli*.

(Adaptation de l'illustration proposée par (Boehm and Vogel, 2012)) Le processus d'expression des curli est induit par des stimuli environnementaux et de signaux du métabolisme. Les régulations positives sont indiquées par les flèches vertes, les répressions sont indiquées par les flèches « aplaties » rouges.

2. Les fusions transcriptionnelles et traductionnelles

Nous avons construit différentes fusions transcriptionnelles et traductionnelles des promoteurs des gènes *csgD* et *csgB* à la Green Fluorescente Protein (GFP) sur un plasmide à faible nombre de copies pour rester proche des conditions physiologiques en termes de nombre de cibles par régulateur (Figure 24). Ces fusions ont été transformées dans différentes souches mutantes afin de revisiter chaque interaction du réseau de régulation de l'expression des curli.

Fusions transcriptionnelles



Fusions traductionnelles

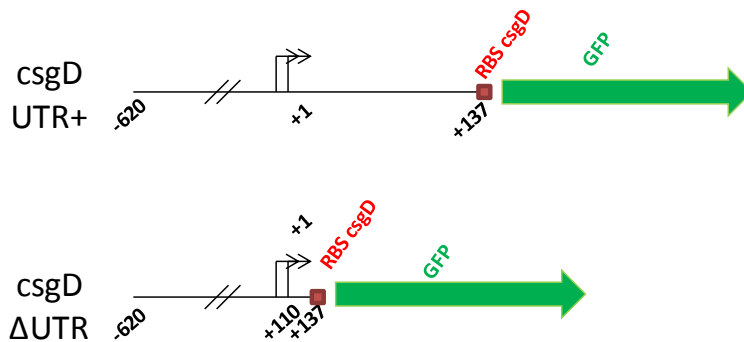


Figure 24: Fusions transcriptionnelles et traductionnelles des gènes *csgD* et *csgB*.

La fusion *csgD TXN* est une fusion transcriptionnelle de *csgD* comportant les deux promoteurs décrits dans la base de données ecocyc.org. Elle permettra d'étudier les régulations transcriptionnelles du gène *csgD*. La fusion *csgD UTR+* est une fusion traductionnelle de *csgD* comportant les deux promoteurs et la partie 5'UTR de *csgD* fusionnée au CDS de la GFP. Elle permettra d'étudier la transcription et la traduction du gène *csgD*. La fusion *csgD ΔUTR* est une fusion traductionnelle de *csgD* où la partie 5'UTR entre le +2 de transcription jusqu'au +110 a été supprimée en conservant uniquement la région du RBS endogène de *csgD*. Elle permettra d'étudier la transcription et la traduction du gène sans la répression des ARNnc qui ne peuvent plus se fixer sur l'ARNm de *csgD*. La fusion *csgB TXN* est une fusion transcriptionnelle de *csgB* comportant l'unique promoteur décrit dans la base de données ecocyc.org. Elle permettra d'étudier la transcription du gène *csgB*.

En plus des souches mutées pour les régulateurs de la transcription (RpoS, CsgD, OmpR), nous avons aussi utilisé le mutant *ompR234* qui favorise l'expression des curli. En effet, cette mutation de la protéine OmpR permet au facteur transcriptionnel de se lier à l'ADN sans phosphorylation préalable, il devient donc constitutivement actif. Cette mutation a été transduite dans la souche BW25113 pour vérifier si elle stimulait bien la production des curli dans ce contexte génétique. De façon surprenante, la mutation *ompR234* ne semble pas avoir d'impact sur l'expression des curli dans notre souche (Figure 25) : la fusion transcriptionnelle de *csgB* est active à 28°C dans les deux contextes génétiques et inactive à 37°C. La fusion transcriptionnelle de *csgD* est elle active aux deux températures. L'activité constitutive de OmpR ne semble donc pas augmenter l'expression des curli, elle n'augmente pas le taux de transcription de l'opéron *csgBA* et n'en modifie pas la thermo-sensibilité. Ceci contraste avec les observations faites par l'équipe ayant identifié la mutation *ompR234*, qui a observé l'augmentation de l'expression des curli en présence de cette protéine mutée (Prigent-Combaret et al., 2001). En particulier, la mutation *ompR234* ne lève pas la possible répression de la transcription de *csgBA* à 37°C. L'observation visuelle sur boîte ne permet pas non plus de conclure à une activation de la transcription de *csgD*, pourtant OmpR possède bien un site de fixation en amont du promoteur du gène *csgD* s'étendant des positions -40 à -59 par rapport à son TSS (Figure 8). Ainsi la mutation *ompR234* pourrait affecter la traduction de *csgD*, voire même celle de *csgB* ou *A* plutôt que seulement la transcription de *csgD*. Nous avons construit et étudié des fusions traductionnelles de *csgD* et transcriptionnelles des petits ARN régulant sa traduction. Compte tenu de la faible activité de certaines de ces fusions, nous avons choisi de montrer le niveau de fluorescence relative (Florescence brute /Absorbance) sans déduction préalable du bruit de l'auto-fluorescence généré par la souche contenant le vecteur pZE sans promoteur (figures 26 à 30).

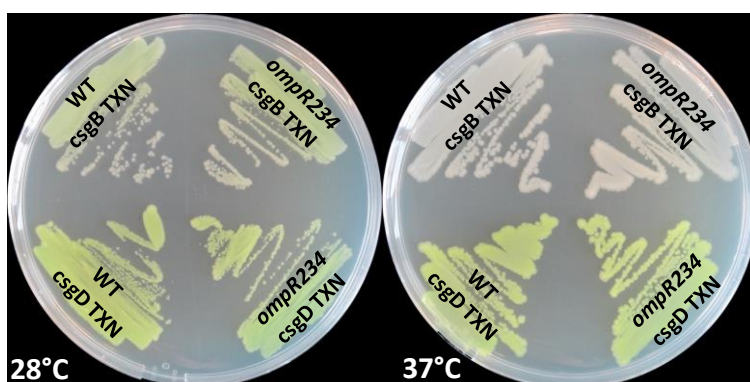


Figure 25: Expression des gènes *csgB* et *csgD* dans le mutant conditionnel *ompR234* comparé à la souche sauvage.

Les souches WT et *ompR234* transformées par le plasmide portant les fusions transcriptionnelles *csgD::gfp* ou *csgB::gfp* sont été striées sur milieu riche (LB) et incubées 48h à une température de 28°C ou 37°C. La mutation *ompR234* ne modifie pas l'expression des gènes *csg* visible à l'œil nu.

3. Régulation de la transcription des gènes *csg*

a) Transcription du gène *csgD*

Nous avons mesuré le niveau de fluorescence reflétant la transcription du gène *csgD* dans les mutants des protéines régulatrices : RpoS, OmpR, et CsgD. Ces mesures sont présentées dans la Figure 26 et doivent être comparées au niveau de fluorescence obtenus pour la souche sauvage (WT).

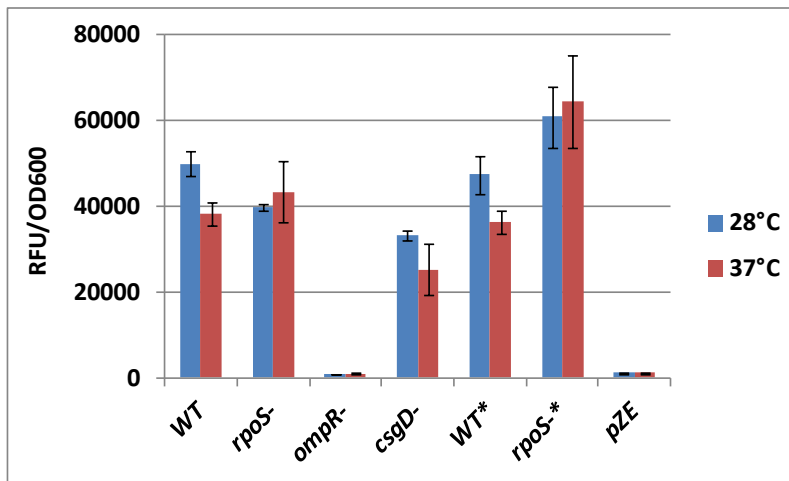


Figure 26: Régulation de la transcription du gène *csgD*.

La RFU/OD600 correspond à la valeur de fluorescence mesurée divisée par l'absorbance. pZE correspond au vecteur vide servant de contrôle pour le niveau d'auto-fluorescence de la souche WT. Les souches ont été cultivées en microplaque dans 150µl de milieu LB. Un aliquot de 100µl de ces cultures a été prélevé après 22h de culture pour être déposé dans une microplaque permettant la mesure de l'absorbance et de la fluorescence par le lecteur Tecan. L'astérisque (*) indique la présence de la mutation *ompR234* dans la souche (Moyennes et écarts-types sont calculés sur des triplicatas biologiques et techniques).

Nous pouvons observer une forte transcription du gène *csgD* légèrement supérieure (20%) à une température de 28°C comparée à celle mesurée à 37°C dans la souche sauvage (WT) et le mutant *ompR234* (WT*). La transcription du gène *csgD* est bien strictement dépendante du régulateur OmpR, car nous pouvons observer l'absence d'un signal de fluorescence dans la souche délétée $\Delta ompR$, comparable à celui du contrôle négatif pZE. Ce résultat en milieu liquide (Figure 26, comparer WT et WT*) ou solide (Figure 25) contraste avec l'absence d'effet de la mutation *ompR234* sur la transcription de *csgD*. De plus, comme nous l'avons observé (Figure 25), l'activité de ce promoteur est peu affectée par la température de culture.

L'expression de *csgD* est très faiblement affectée par l'absence du facteur sigma RpoS à 28°C qui induit une perte d'expression de 20% comparée à la souche WT. En revanche à 37°C, le facteur RpoS semble très légèrement réprimer sa transcription (peu significatif (15%), mais très reproductible et plus visible en mesures cinétiques (résultats non montrés)).

Si la mutation *ompR234* ne modifie pas la transcription du gène *csgD*, sa présence modifie l'activité de RpoS et souligne une action répressive de RpoS sur la transcription de *csgD* aux deux températures, 28°C et 37°C.

Dans la littérature, la protéine CsgD régule sa propre transcription. La région promotrice du gène *csgD* présente quatre sites de fixation de CsgD (Ogasawara et al., 2011). Deux sites de liaisons sont décrits comme distaux (de -188 à -159 et de -126 à -106 par rapport au TSS) et deux autres sont chevauchants au site d'initiation de la transcription (de +2 à +41). CsgD pourrait donc à la fois s'auto-activer et se réprimer. Dans nos conditions, nous observons une diminution de 25% de la transcription de *csgD* dans le mutant $\Delta csgD$, l'effet est donc globalement une auto-activation.

En conclusion, la transcription de *csgD* est bien dépendante du facteur de transcription OmpR et CsgD active globalement sa propre transcription. Le facteur RpoS n'est en revanche pas indispensable pour la transcription de *csgD*. Dans un contexte génétique *ompR234*, RpoS semble réprimer la transcription de *csgD* quelque soit la température. Enfin, la transcription de *csgD* est peu modulée par la température qui influence peu les régulations observées.

b) Transcription du gène *csgB*

Comme nous l'avons mentionné, l'expression des curli semble nulle à 37°C puisque le mutant $\Delta csgB$ est aussi peu coloré que la souche WT (Figure 21). Leur expression semble positivement modulée par RpoS à 28°C puisqu'un mutant $\Delta rpoS$ est décoloré sur Rouge Congo (Figure 21). Nous avons vérifié si ces déductions sont compatibles avec l'activité promotrice de l'opéron *csgBA* mesurés dans différentes souches (Figure 27).

Comme démontré par les figures 19, 25, 28 et 29, l'expression de curli est thermo-dépendante. Or l'activité de la fusion transcriptionnelle *csgB::gfp* est uniquement mesurable à 28°C dans la souche sauvage. Les protéines régulatrices RpoS, OmpR et CsgD sont bien essentielles à l'expression de *CsgB* car leur délétion entraîne l'absence d'activité de la fusion transcriptionnelle (Figure 27 A). Ceci confirme l'importance du facteur sigma RpoS et de l'activateur CsgD pour l'induction de *csgB*. Les sites de fixations de CsgD sur le promoteur de *csgB* sont connus (Ogasawara et al., 2010) et on sait que RpoS interagit avec *csgB* (Bougdoor et al., 2004). Ces deux régulations sont vraisemblablement

directes, car leur délétion entraîne une perte totale de l'activité de la fusion transcriptionnelle. En revanche, puisqu'il n'existe pas de site de fixation du régulateur OmpR en amont du promoteur de *csgBA*, nous supposons que l'effet activateur d'OmpR sur la transcription de *csgB* est la conséquence de la forte activation de la transcription de *csgD* par OmpR (Figure 26).

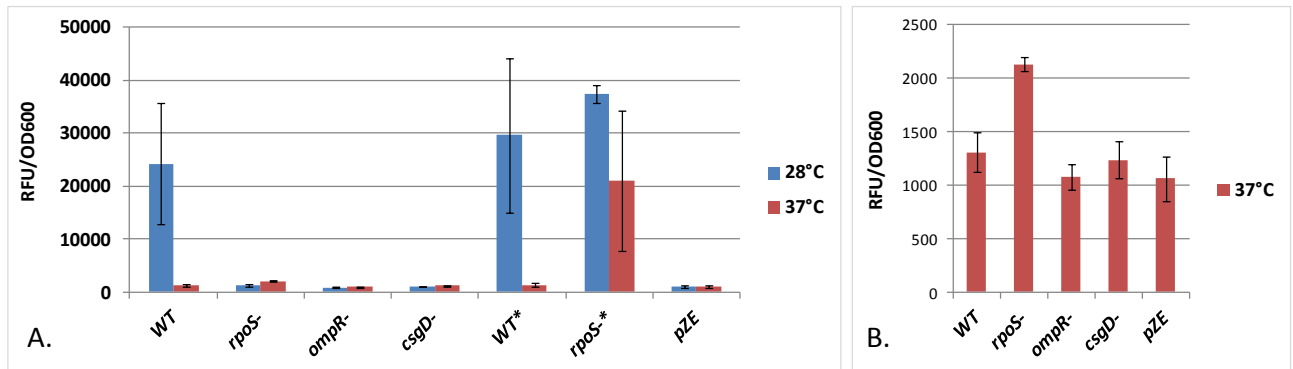


Figure 27 : Régulation de la transcription du gène *csgB*.

(A.) La RFU/OD600 correspond à la valeur de fluorescence mesurée divisée par l'absorbance. pZE correspond au vecteur vide servant de contrôle pour le niveau d'auto-fluorescence de la souche WT. Les souches ont été cultivées en microplaque dans 150µl de milieu LB. Un aliquot de 100µl de ces cultures a été prélevé après 22h de culture pour mesurer l'absorbance et de la fluorescence par le lecteur Tecan. L'astérisque (*) indique la présence de la mutation *ompR234* dans la souche (Tripliquats biologiques et techniques). (B.) « Zoom » de la Figure 27A montrant uniquement la valeur de RFU/OD600 à 37°C pour les souches WT, $\Delta rpoS$, $\Delta ompR$, $\Delta csgD$, et le contrôle pZE.

De façon intéressante, la mutation *ompR234*, stimule modestement l'activité transcriptionnelle de *csgB* (Figure 27), la encore très probablement de façon indirecte sans affecter le niveau d'expression de *csgB* à 37°C. En revanche, tout comme c'était le cas de *csgD*, elle amplifie un peu l'effet répressur de RpoS qui est très probablement indirect. Curieusement, cette répression par RpoS n'est visible qu'à la température de 37°C ce qui pourrait être la conséquence de la régulation négative de la transcription de *csgD*. En effet, l'activité de la fusion *csgB::gfp* est uniquement mesurable dans la souche $\Delta rpoS$ à 37°C (Figure 27 B). Ce résultat s'accorde donc avec la coloration au Rouge Congo de la souche $\Delta rpoS$ à 37°C (par opposition à la souche WT, Figure 21) et semble démontrer l'expression des curli à 37°C seulement en absence de RpoS. Cette répression est encore une fois amplifiée dans le contexte *ompR234*. À ce stade, nous nous demandons quel est le mécanisme qui permet à RpoS d'être à la fois activateur ou répressur selon la température ? Ces observations suggèrent elles que RpoS puisse avoir un effet sur la traduction de CsgD ?

En somme, nos résultats obtenus sur la souche BW25113, sont en accord avec le réseau de régulation suggéré à partir de l'étude de différentes souches d'*E. coli* et *Salmonella*. En particulier, OmpR active bien la transcription de *csgD* dont le produit active celle de *csgB*. En revanche, le rôle de RpoS n'est pas clair : il semble indispensable pour l'induction de *csgB* à 28°C, mais constituerait

plutôt un déterminant négatif de la transcription de *csgB* et *csgD* à 37°C. Il pourrait donc être impliqué dans la thermorégulation de l'expression des curli, comme cela avait été suggéré par (Bougdoor et al., 2004) qui décrit la fixation du facteur sigma RpoS au promoteur des gènes *csgBA* plus importante à une température inférieure à 30°C comparés à celle observée pour une température de 37°C.

4. Régulation de la traduction de CsgD

a) Étude de l'activité des fusions traductionnelles *csgD* TRAD UTR+ et *csgD* TRAD ΔUTR

Afin d'étudier la traduction de l'activateur de l'expression des curli, nous avons construit deux fusions qui rapportent l'activité transcriptionnelle et traductionnelle de *csgD*. Toutes deux possèdent le RBS du gène avec (UTR+) ou sans (ΔUTR) la séquence non traduite 5'UTR (Figure 24). Dans la souche sauvage, les fusions traductionnelles *csgD* présentent une activité quasi nulle, c'est-à-dire similaire à celle du vecteur vide pZE quelque soit la température de croissance (Figure 28 A). Seule la fusion traductionnelle *csgD* ΔUTR dans le mutant *ΔrpoS* semble présenter une activité très légèrement supérieure à l'auto-fluorescence. Ceci signifierait que RpoS affecte négativement la traduction de *csgD* (observable uniquement dans en présence de la protéine OmpR mutée (*ompR234*)), possiblement en influençant le niveau d'expression des ARNnc. Pour tester cette hypothèse, nous avons étudié l'expression des deux fusions traductionnelles dans les souches délétées de ces ARN régulateurs (affectant la traduction de l'ARNm CsgD) et avons étudié la régulation de la transcription de ces ARN par RpoS (voir III.B.5.a).

Si RpoS modifie bien l'expression de certains de ces ARN (voir III.B.6), la mutation des ARNnc identifiés comme répresseurs de la traduction de CsgD, n'a pas d'effet sur l'activité des fusions dans la souche sauvage. Cette faible activité est semblable à l'auto-fluorescence des souches sauf pour la mutation *ΔmcaS*.

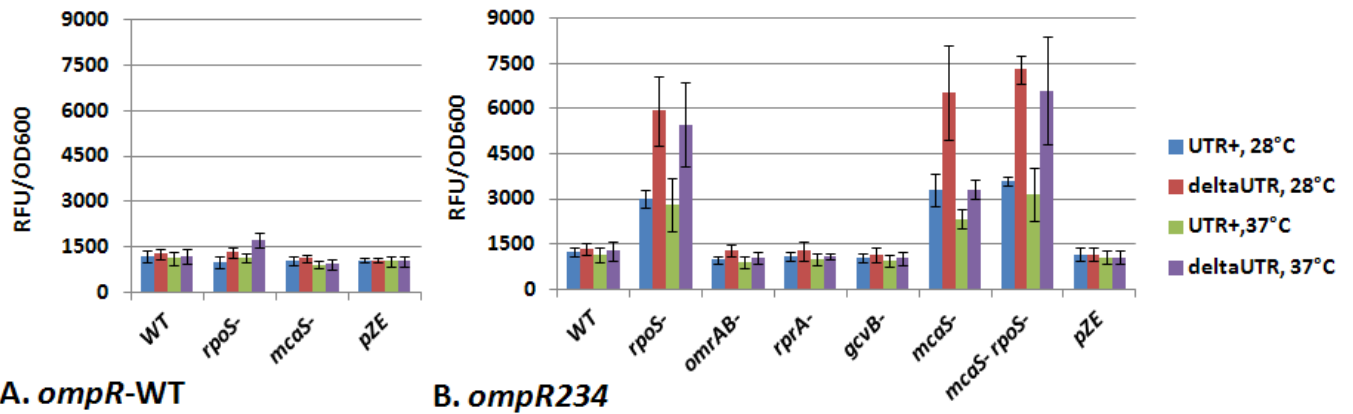


Figure 28 : Activité des fusions traductionnelles *csgD* ::*gfp* avec et sans 5'UTR.

La RFU/OD600 correspond à la valeur de fluorescence mesurée divisée par l'absorbance. pZE correspond au vecteur vide servant de contrôle pour le niveau d'auto-fluorescence de la souche WT. Les souches ont été cultivées en microplaque dans 150µl de milieu LB. Un aliquot de 100µl de ces cultures a été prélevé après 22h de culture pour mesurer de l'absorbance et de la fluorescence par le lecteur Tecan. L'astérisque (*) indique la présence de la mutation *ompR234* dans la souche (Tripliquatas biologiques et techniques). **A.** Allèle du gène sauvage *ompR* **B.** Allèle du gène muté *ompR234*

Dans le contexte génétique *ompR234* (Figure 28 B), les deux fusions traductionnelles (+/- 5'UTR) ne présentent pas plus d'activité que dans la souche sauvage. Ceci semble signifier que l'activation constitutive de OmpR n'a pas de conséquence directe sur la traduction de CsgD (en présence de tout les autres acteurs du réseau de régulation) et que ce régulateur affecte seulement la transcription de l'opéron *csgDEFG*. Par ailleurs, il est très surprenant que la suppression de la séquence des sites de fixation des ARNnc, laquelle est responsable de la répression de la traduction de *csgD* (Δ UTR), ne résulte pas en une augmentation de l'activité traductionnelle. Nous espérons que l'élévation du taux de transcription de *csgD* dans la souche *ompR234* nous permettrait de mesurer des différences d'activité entre les fusions avec et sans 5'UTR et ainsi d'apprécier l'effet régulateur des ARNnc.

La mutation des gènes codant les ARNnc n'augmente pas non plus cette activité traductionnelle, mais il est souvent difficile de mesurer l'effet de la suppression d'un gène codant un ARNnc. Seule la délétion du gène codant l'ARN McaS favorise la traduction de *csgD*, mais uniquement dans le contexte *ompR234* (comparaison des Figure 28 A et B). Il est intéressant de noter qu'à 28°C, la mutation de McaS affecte plus la fusion sans 5'UTR qu'avec 5'UTR, c'est-à-dire sans la séquence nécessaire à son appariement qui provoque la répression de la traduction. Ceci semble signifier que McaS agit aussi sur la transcription de ce gène, et c'est ce que nous montrons dans la partie 5 de ce chapitre (Figure 30 A). Cet effet répresseur de McaS est aussi visible à 37°C.

La répression de la traduction de *csgD* par RpoS est aussi amplifiée dans la souche *ompR234*. Nous pouvons également observer que la levée de répression dans la souche Δ *rpoS ompR234* est comparable à celle observée dans la souche Δ *mcaS ompR234*. Nous n'observons pas une activité

traductionnelle plus élevée dans la souche porteuse des deux délétions $\Delta rpoS \Delta mcaS ompR234$, comparativement à chaque délétion assortie de l'allèle $ompR234$. Il n'y a donc pas additivité des régulations et la régulation par RpoS pourrait être indirectement médiée par la régulation de McaS. Toutefois, il serait alors surprenant de ne pas avoir strictement les mêmes effets régulateurs aux deux températures de croissance. Or à 37°C, l'effet répresseur de RpoS est plus fort que celui de McaS sur la fusion traductionnelle $csgD \Delta UTR$ (Figure 28 B). Donc la levée de répression de la fusion TRAD $csgD \Delta UTR$ dans la souche $ompR234$ pourrait se faire par des voies indépendantes. Nous supposons que la protéine RpoS pourrait indirectement affecter la fusion traductionnelle de $csgD::gfp$ en contrôlant à la fois l'expression d'un répresseur de sa transcription (Figure 26) et/ou de sa traduction comme McaS (Figure 28). Le rôle de RpoS dans cette régulation demeure incompris et pourrait faire encore appel à une autre entité.

En somme, seule la mutation de l'ARNnc McaS affecte de façon mesurable la traduction de CsgD (Figure 24). RpoS semble aussi impliqué dans la régulation négative post-transcriptionnelle de CsgD. Les mesures de fluorescence dans un double mutant ($\Delta mcaS \Delta rpoS$) indiquent que ces régulateurs pourraient agir ensemble, nous estimons nécessaire de vérifier si les ARNnc n'affectent pas aussi, indirectement, la transcription des gènes csg . Nos résultats indiquent que McaS pourrait, comme RpoS, participer à la thermorégulation de l'expression des curli. L'ensemble des régulations observées dans les parties A, B, C et D sont illustrées dans la Figure 29.

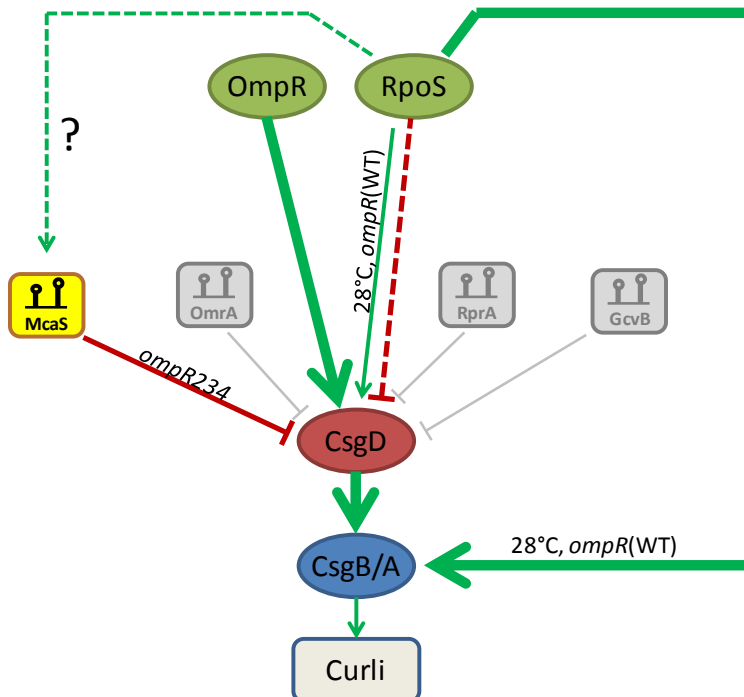


Figure 29 : Bilan des régulations transcriptionnelles et traductionnelles observées.

Les flèches vertes indiquent des activations, les rouges des répressions. Les interactions non observées sont indiquées en gris. La force des interactions est matérialisée par la grosseur des flèches. Les annotations sur les flèches correspondent à l'unique condition où cette régulation a pu être observée. Les éléments grisés sont des régulations non confirmées par notre étude. Lorsque l'allèle d'*ompR* n'est pas précisé, la régulation est observée dans les deux allèles d'*ompR*. Lorsque la température n'est pas précisée, la régulation est observée dans aux deux températures.

5. Effet des ARNnc sur la transcription des gènes *csgD* et *csgB*

a) Influence des ARNnc sur la transcription des gènes *csg*

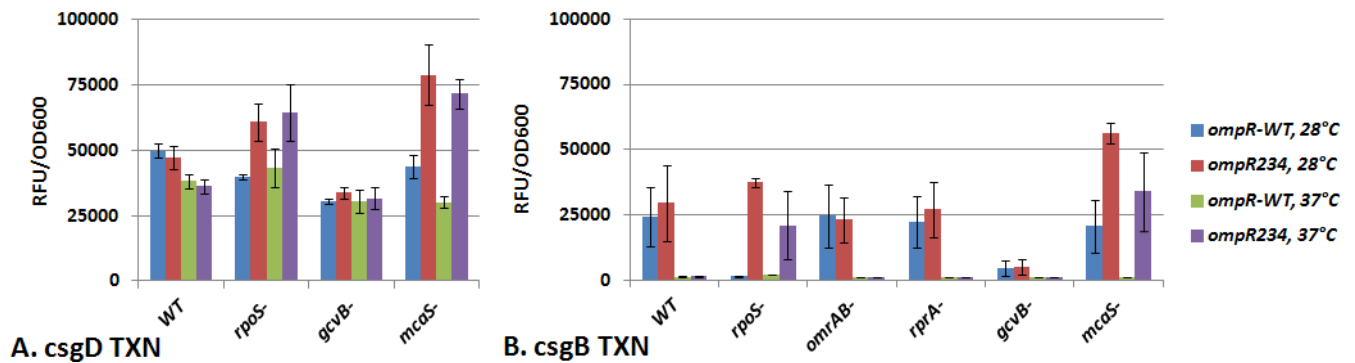


Figure 30 : Transcription des gènes *csgD* et *csgB* : effet des mutations des ARNnc

A. La RFU/OD600 correspond à la valeur de fluorescence mesurée divisée par l'absorbance. pZE correspond au vecteur vide servant de contrôle pour le niveau d'auto-fluorescence de la souche WT. Les souches ont été cultivées en microplaque dans 150µl de milieu LB. Un aliquot de 100µl de ces cultures a été prélevé après 22h de culture pour mesurer de l'absorbance et de la fluorescence par le lecteur Tecan. L'astérisque (*) indique la présence de la mutation *ompR234* dans la souche (Tripliquats biologiques et techniques). A. Fusion transcriptionnelle du gène *csgD* (*csgD* TXN) B. Fusion transcriptionnelle du gène *csgB* (*csgB* TXN)

Compte tenu du fait que certaines mutations affectaient les fusions traductionnelles de *csgD*, qui reflètent les régulations transcriptionnelles et traductionnelles, nous avons voulu vérifier si les ARNnc n'influençaient pas aussi la transcription des gènes *csg*.

La transcription du gène *csgD* est affectée (moins 40%) par la délétion de l'ARNnc GcvB à 28°C dans les deux contextes génétiques, allèle *ompR* ou *ompR234* (Figure 30 A). Au contraire, la mutation $\Delta mcaS$ stimule la transcription du gène *csgD* mais uniquement dans le contexte *ompR234*. Donc la régulation traductionnelle de *csgD* par McaS (observée dans la Figure 28) est en partie due à un effet indirect de McaS sur la transcription du gène. Il en est de même pour la régulation négative par RpoS dans le contexte *ompR234* : le facteur sigma RpoS réprime (indirectement) à la fois la transcription et la traduction de *csgD* aux deux températures de croissance.

L'absence de l'ARNnc GcvB réduit de plus de 80% le niveau de transcription de *csgB* pour une température de croissance de 28°C (Figure 30 B). Comme cela a été observé sur milieu solide (Figure 21, partie A), l'activité transcriptionnelle du gène *csgB* est aussi faible en milieu liquide à 37°C. Toutefois, il est facile d'observer l'augmentation de cette activité dans les souches codant un élément de répression. Nous observons ainsi une répression transcriptionnelle de *csgB* par McaS, uniquement visible dans un contexte *ompR234* et à 37°C (Figure 30 B).

En conclusion, nous confirmons dans cette partie que RpoS et McaS répriment à la fois *csgB* et *csgD* au niveau transcriptionnel, en plus de leur effet répresseur sur la traduction. Nos résultats montrent aussi l'effet stimulateur (indirect) de GcvB sur la transcription de *csgB* et dans une moindre mesure de la transcription de *csgD*.

b) Liaison au Rouge Congo des mutants des protéines régulatrices et des ARNnc

Afin de déterminer si les régulations revisitées dans les parties précédentes de ce chapitre ont un impact intuitif sur l'expression des curli, nous avons inoculé toutes les souches mutantes sur une boîte de culture solide LB Agar additionnée de Rouge Congo (RC, Figure 31).

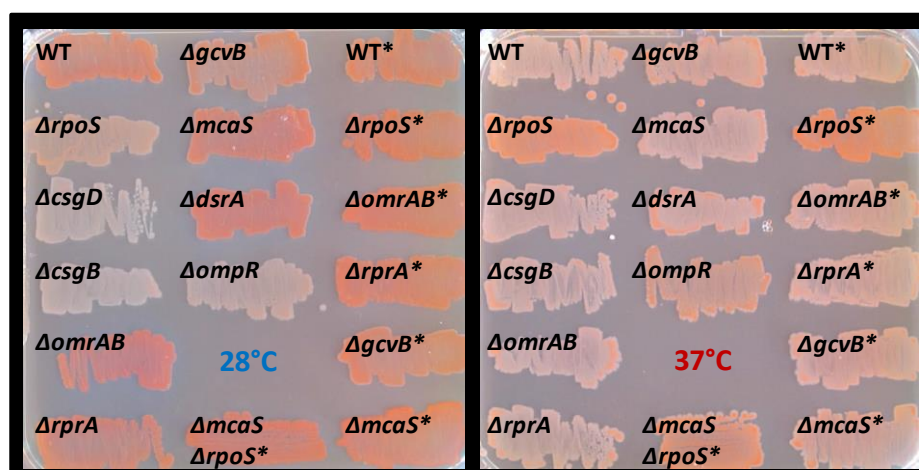


Figure 31 : Liaison au Rouge Congo des dérivés de la souche BW25113.

Les différentes souches ont été striées sur gélose riche complétée par 0,004% (p/v) de Rouge Congo. Les souches ont été incubées pendant 48h à 28 et 37°C. L'astérisque indique la présence de l'allèle *ompR234* dans les souches *ΔrpoS*, *ΔomrAB*, *ΔrprA*, *ΔgcvB*, *ΔmcaS* et *ΔmcaS/ΔrpoS*.

À 28°C, les mutants *ΔmcaS* et *ΔomrAB* ont une coloration plus intense que la souche sauvage ce qui confirme leur rôle de répresseur sur la traduction de CsgD. Par opposition, la mutation *ΔgcvB*, engendre une légère décoloration au RC (Rouge Congo), ce qui conforte son rôle d'activateur (indirect) de la transcription de *csgB* (Figure 27 B). Le mutant *ΔompR* est décoloré comme les souches *ΔcsgD* et *ΔcsgB* démontrant l'importance de ce régulateur pour l'expression des curli. Pourtant, l'allèle *ompR234* ne conduit pas à une coloration plus rouge, ce qui est en accord avec l'absence d'effet de cette mutation ponctuelle sur la transcription (Figure 25) et la traduction (Figure 28) du

gène *csgD* (WT comparé à WT*). Dans ce même contexte génétique *ompR234* (noté WT*), la mutation additionnelle d'un gène codant l'un des ARNnc répresseurs n'affecte plus le phénotype de coloration. À 37°C, la très faible liaison au RC des bactéries semble indépendante de la présence de curli à la surface des bactéries puisque les mutants *csg* sont aussi peu colorés que la souche sauvage. À cette température, la souche sauvage présente une aussi faible coloration que la souche *ompR234* (WT*) ; seules les souches porteuses de la délétion $\Delta rpoS$ sont plus colorées. Le mutant *rpoS* perd son phénotype décoloré indépendamment de l'allèle *ompR234* ou *ompR*-sauvage. Ceci contraste fortement avec certaines données de fusions reportrices, en particulier de la traduction de *csgD* (Figure 28) qui montre un fort effet répresseur de RpoS sur la traduction de *csgD* uniquement dans la souche *ompr234*.

Toujours à 37°C, les mutants d'ARNnc affectant la traduction de CsgD ont la même coloration que la souche sauvage et les mutants Δcsg . En revanche, la mutation $\Delta mcaS$ n'affecte plus la coloration. Le mutant $\Delta ompR$ présente une coloration partielle et intermédiaire à celle du sauvage et du mutant $\Delta rpoS$. Ceci indique qu'OmpR pourrait contribuer à la répression due à RpoS. Les régulateurs OmpR et RpoS semblent donc nécessaires à l'expression des curli à 28°C, mais pas à 37°C.

c) La fusion *csgD* IG : une fusion traductionnelle de plus forte activité

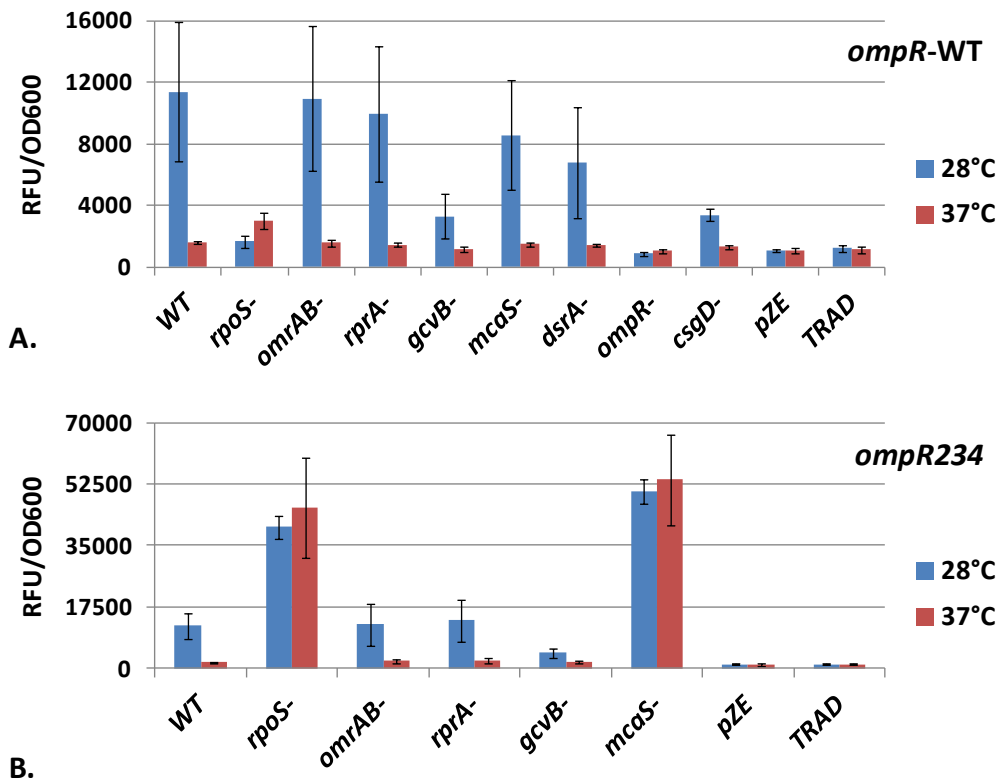


Figure 32 : La régulation transcriptionnelle et traductionnelle sont plus visibles avec la fusion *csgD* IG.

La RFU/OD600 correspond à la valeur de fluorescence mesurée divisée par l'absorbance. pZE correspond au vecteur vide servant de contrôle pour le niveau d'auto-fluorescence de la souche WT. Les souches ont été cultivées en microplaque dans 150µl de milieu LB. Un aliquot de 100µl de ces cultures a été prélevés après 22h de culture pour mesurer de l'absorbance et de la fluorescence par le lecteur Tecan. L'astérisque (*) indique la présence de la mutation *ompR234* dans la souche (Tripliquats biologiques et techniques). (**A.**) La fusion *csgD* IG dans l'allèle *ompR*. (**B.**) La fusion *csgD* ID dans l'allèle *ompR234*. L'intitulé « TRAD » correspond à la fusion *csgD* TRAD (UTR+) dans la souche sauvage ou *ompR234* permettant de faire le comparatif entre la fusions *csgD* IG et les fusions *csgD* TRAD (UTR+).

La fusion *csgD*-IG est une fusion traductionnelle comportant à la fois le RBS du gène *csgD* et celui du rapport GFP ; tout comme les fusions présentées dans la Figure 24 et jusque là utilisées, l'ensemble de l'IG est présente pour piloter l'expression du gène *csgD*. L'activité de cette fusion (Figure 28) est plus importante que celle des fusions traductionnelles étudiées précédemment (Figure 26). De plus, cette fusion semble être sensible à la température et son activité beaucoup plus importante à 28°C qu'à 37°C où l'activité est presque nulle. Cependant, la fusion *csgD*-IG semble se comporter comme la fusion transcriptionnelle avec une activité amoindrie dans les mutants $\Delta rpoS$, $\Delta gcvB$ et $\Delta ompR$ à une température de 28°C et elle permet également de confirmer l'auto-activation par CsgD. À 37°C, nous observons une légère répression de l'activité de la fusion par RpoS comparée à celle de la souche sauvage, ce qui n'était pas observé avec la fusion transcriptionnelle (Figure 26), mais avec la fusion traductionnelle avec ou sans 5'UTR (Figure 28).

Dans un contexte génétique *ompR234* (Figure 32-B), l'activité de cette fusion est plus importante dans les mutants $\Delta rpoS$ et $\Delta mcaS$ aux deux températures, comme c'était le cas des fusions traductionnelles (Figure 28). À l'inverse, nous observons une perte d'activité de la fusion dans un mutant $\Delta gcvB$ à 28°C comme c'était le cas avec la fusion transcriptionnelle (Figure 26). Au final, cette fusion n'est pas beaucoup plus active à 37°C que les fusions précédentes dans une souche *ompR234*.

d) Bilan sur les interactions majeures du réseau de régulation contrôlant l'expression des curli.

Pour conclure, cette fusion rapporte de façon plus marquée à la fois les régulations transcriptionnelles et traductionnelles et possède une activité supérieure à 28°C comparée à la précédente fusion traductionnelle. Elle permet donc de visualiser des effets régulateurs des ARNnc dans la souche sauvage et plus uniquement dans le contexte *ompR234*. Elle permet notamment de mieux voir que DsrA a un effet répresseur mesurable sur l'expression de *csgD*, que l'on peut corrélérer à sa coloration RC plus rouge que celle du sauvage à 28°C dans le contexte *ompR* sauvage (Figure 31).

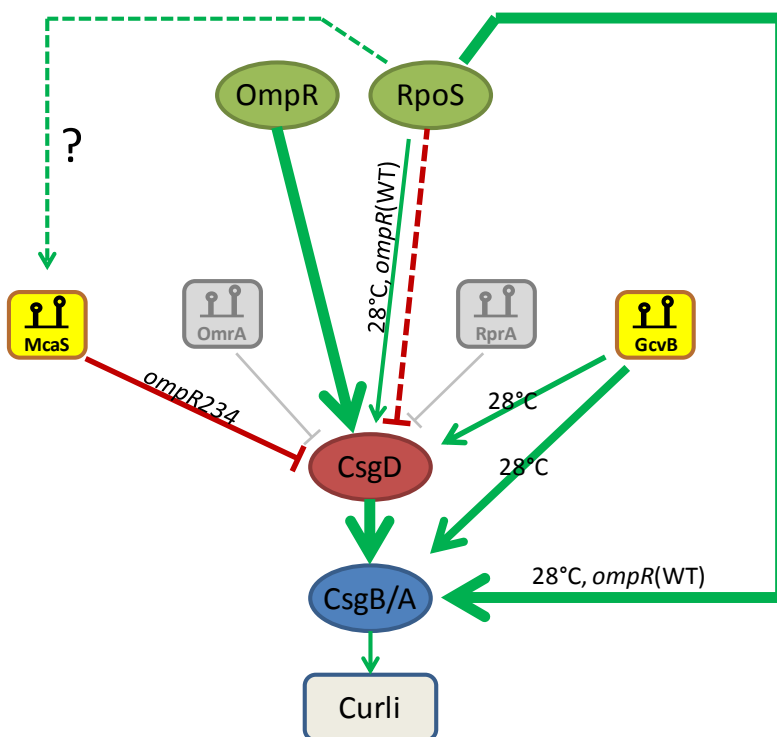


Figure 33 : État des connaissances sur les interactions entre membres du réseau de régulation de l'expression des curli.

Les flèches vertes indiquent des activations, les rouges des répressions. Les interactions non observées sont indiquées en gris. La force des interactions est matérialisée par la grosseur des flèches. Les annotations sur les flèches correspondent à l'unique condition où cette régulation a pu être observée. Les éléments grisés sont des régulations non confirmées par notre étude. Lorsque l'allèle d'*ompR* n'est pas précisé, la régulation est observée dans les deux allèles d'*ompR*. Lorsque la température n'est pas précisée, la régulation est observée dans aux deux températures.

Le réseau de régulation d'expression des curli est complexe certaines régulations ne sont visibles que dans un contexte particulier (Figure 33). Par exemple, le rôle activateur de RpoS n'est visible que

dans un contexte *ompR*-WT. Nous avons néanmoins pu vérifier l'activation du gène *csgD* par OmpR et du gène *csgB* par CsgD. Le rôle répresseur des ARNnc sur la traduction de CsgD n'a pas pu être confirmé sauf la répression par McaS dans un contexte génétique *ompR234*. Nous montrons également que l'ARNnc GcvB agit non pas sur la traduction de CsgD, mais plutôt sur la transcription de *csgB* et *csgD*. Cette régulation est sûrement indirecte est médiée par un facteur de transcription dont la traduction serait activée par GcvB. Le réseau de régulation présenté serait donc incomplet, et il est fort probable que les régulateurs soient interconnectés, car il pourrait interagir entre eux dans certaines conditions comme nous l'avons suggéré pour RpoS et OmpR. Afin de vérifier notamment si l'expression de certains ARNnc répresseurs de CsgD est sous le contrôle de RpoS, nous avons construit des fusions transcriptionnelles et les avons transformées dans les souches WT, $\Delta rpoS$, $\Delta csgD$ et $\Delta ompR$.

6. Régulation de la transcription des ARNnc modulant la traduction de CsgD et RpoS

Dans cette partie du chapitre, nous avons essayé de déterminer entre autres, si la répression indirecte de la traduction de CsgD par RpoS passait par la régulation de la transcription des ARNnc. Nous avons donc construit des fusions transcriptionnelles de la région promotrice de ces ARN (DsrA, RprA, OmrA, McaS, GcvB) avec la GFP et les avons transformées dans les souches mutées pour les facteurs de transcription RpoS, OmpR et CsgD. Nous espérons aussi à travers cette étude de l'expression des ARNnc, déterminer si leur niveau de transcription est modulé par la température auquel cas, certains pourraient être impliqués dans la thermorégulation de l'expression des curli.

a) Transcription des ARNnc GcvB et DsrA

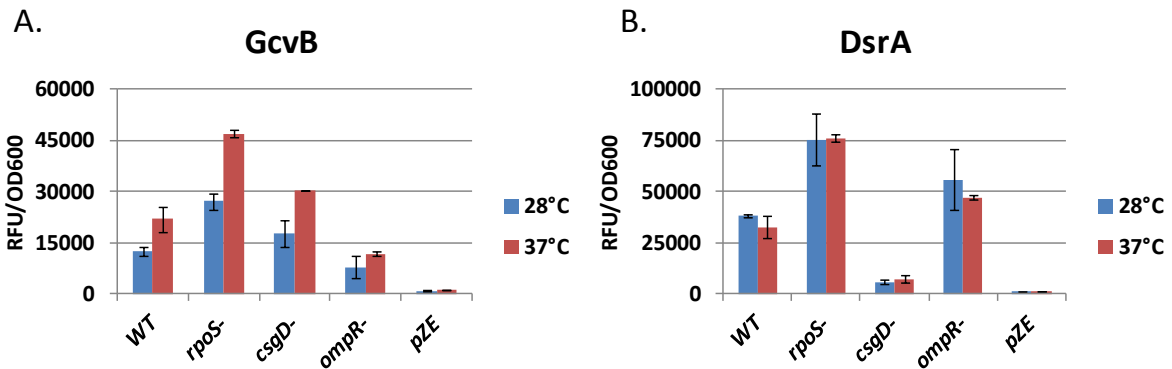


Figure 34 : Transcription des ARNnc GcvB et DsrA.

Les mesures de fluorescence des fusions (*gcvB*::GFP et *dsrA*::GFP) ont été effectuées après 22h de culture en milieu riche (LB) aux températures de 28°C et 37°C. Les fluorescences brutes ont été divisées par l'absorbance de la culture pour donner des valeurs de fluorescence relative (RFU/OD600). Les valeurs indiquées pour le plasmide vide pZE-Spec renseignent sur le niveau d'autofluorescence. Les souches utilisées sont les souches WT, $\Delta rpoS$, $\Delta csgD$ et $\Delta ompR$ dans un contexte génétique *ompR*-WT pour la fusion transcriptionnelle de GcvB (A.) et la fusion transcriptionnelle de DsrA. (B.). Les barres d'erreurs correspondent à l'écart-type entre les triplicatas biologiques et techniques.

La transcription de l'ARNnc GcvB est dépendante de la température (Figure 34 A) : elle est deux fois plus importante à 37°C comparée à 28°C. La transcription de *gcvB* est faiblement réprimée par CsgD, et plus fortement par RpoS (d'un facteur deux). OmpR semble activer la transcription de GcvB, nous observons une activité deux fois moins importante dans le mutant $\Delta ompR$ à 37°C. La transcription de l'ARNnc GcvB est donc régulée par les trois facteurs de transcription RpoS, OmpR et CsgD. Aucun de ces facteurs ne semble fortement impliqué dans sa sensibilité à la température de croissance. Si l'ARNnc GcvB pourrait contribuer à la thermorégulation de l'expression des curli, GcvB stimule (indirectement) la transcription des gènes *csgD* et *csgB* uniquement à 28°C. (Figures 28 et 30). Il n'explique donc pas la thermorégulation des curli.

La transcription de l'ARNnc DsrA (Figure 34 B) n'est quant à elle pas thermo-dépendante et elle est réprimée par RpoS et par OmpR, d'un facteur deux au moins. De façon intéressante, CsgD est essentiel à l'expression de DsrA puisque sa mutation entraîne une perte d'activité de la fusion de 90% comparée à la souche sauvage. Nous observons donc une boucle de rétrocontrôle négatif de RpoS sur l'expression de l'ARN DsrA qui stimule sa traduction.

b) Transcription de l'ARNnc RprA

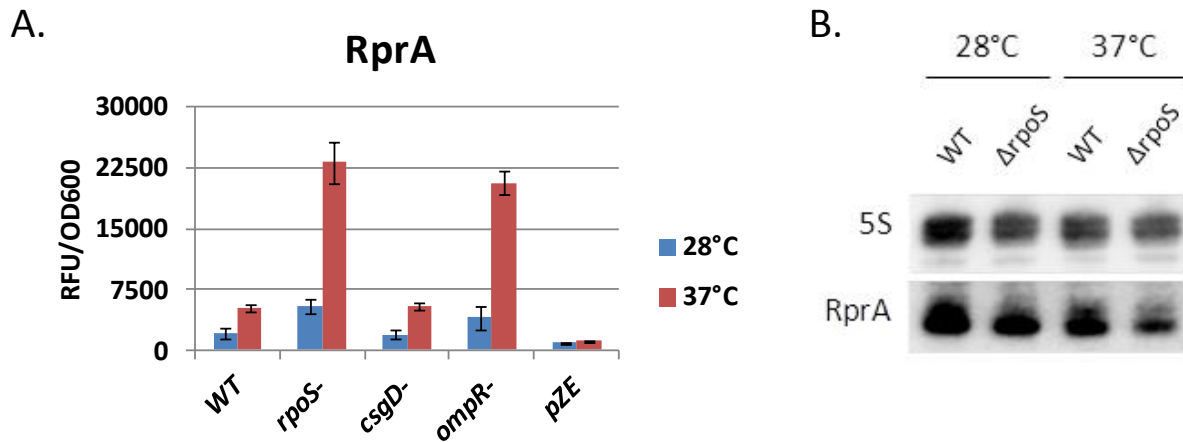


Figure 35 : Transcription et accumulation de l'ARNnc RprA.

A. Les mesures de fluorescence de la fusion *rprA::GFP* ont été effectuées après 22h de culture en milieu riche aux températures de 28°C et 37°C. Les fluorescences brutes ont été divisées par l'absorbance de la culture pour donner des valeurs de fluorescence relative (RFU/OD600). Les valeurs indiquées pour le plasmide vide pZE-Spec renseignent sur le niveau d'autofluorescence. Les souches utilisées sont les souches WT, $\Delta rpoS$, $\Delta csgD$ et $\Delta ompR$ dans un contexte génétique *ompR*-WT pour la fusion transcriptionnelle de RprA. Les barres d'erreurs correspondent à l'écart-type entre les triplicatas biologiques et techniques. **B.** Northern Blot de l'ARNnc RprA, les ARN totaux ont été extraits à une DO=3 par la technique du phénol chaud seulement dans la souche WT et délétée pour le gène *rpoS*. 5µg d'ARN totaux ont été déposés par piste du gel d'électrophorèse.

La Figure 35 A montre que la transcription de l'ARNnc RprA est également plus forte à une température de 37°C, comme c'était le cas pour GcvB. La transcription de l'ARNnc RprA est fortement réprimée par les facteurs de transcription RpoS et OmpR (d'un facteur 4) à 37°. Malgré sa sensibilité à la température de croissance, cet ARNnc ne semble pas participer à la thermorégulation de l'expression des curli, car nous n'avons pas observé d'effet de régulation de cet ARNnc sur la traduction de CsgD (Figure 28). Comme le montre l'hybridation Northern de la Figure 35 B, l'accumulation de l'ARNnc RprA ne semble pourtant ni affectée par la température ni par le facteur RpoS. Il est donc possible que RpoS ait un rôle différent dans l'initiation de la transcription et la stabilité de certains ARN régulateurs comme le montrera également l'étude l'expression de McaS (ci dessous). Au même titre que DsrA, RprA active la traduction de RpoS qui réprime l'initiation de la transcription de *rprA* par rétrocontrôle.

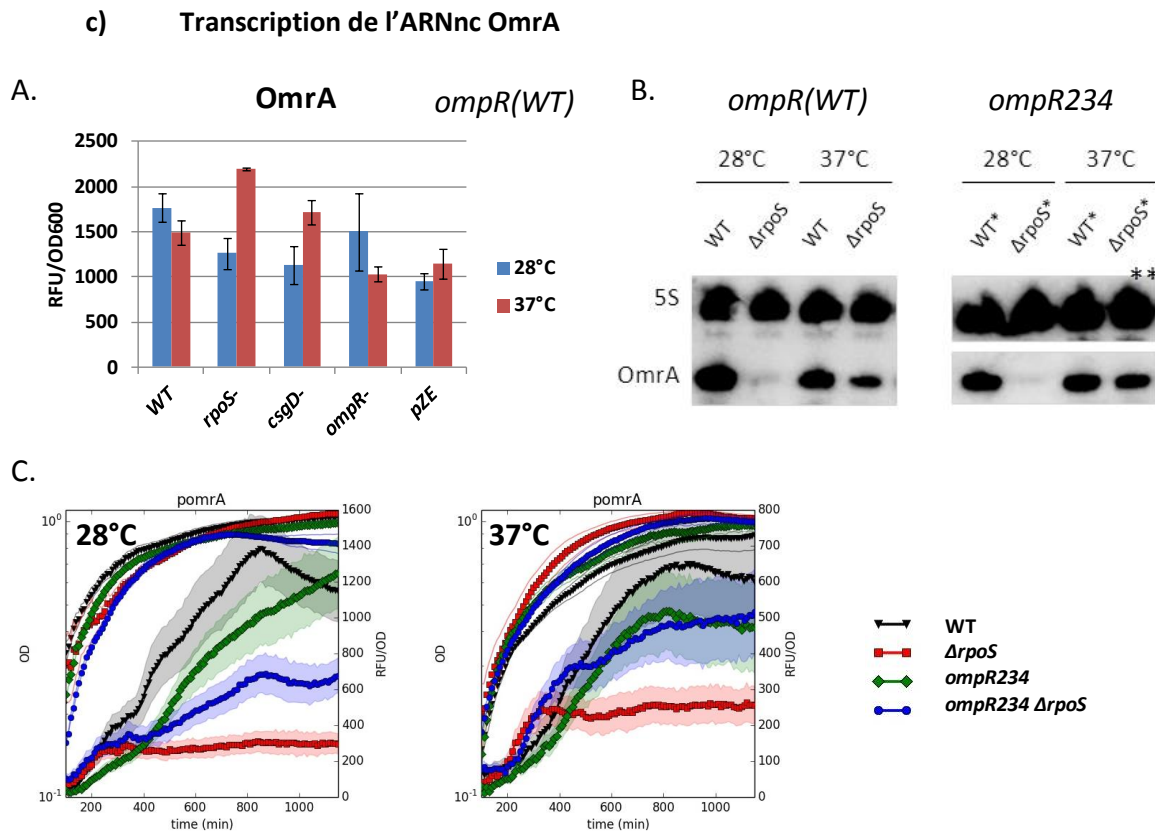


Figure 36 : Transcription et accumulation de l'ARNnc OmrA

A. Les mesures de fluorescence de la fusion *omrA::GFP* ont été effectuées après 22h de culture en milieu riche aux températures de 28°C et 37°C. Les fluorescences brutes ont été divisées par l'absorbance de la culture pour donner des valeurs de fluorescence relative (RFU/OD600). Les valeurs indiquées pour le plasmide vide pZE-Spec renseignent sur le niveau d'auto-fluorescence. Les souches utilisées sont les souches WT, $\Delta rpoS$, $\Delta csgD$ et $\Delta ompR$ dans un contexte génétique *ompR*-WT pour la fusion transcriptionnelle d'OmrA. Les barres d'erreurs correspondent à l'écart-type entre les triplicatas biologiques et techniques. **B.** Northern Blot de l'ARNnc OmrA, les ARN totaux ont été extraits à une DO=3 par la technique du phénol chaud seulement dans la souche WT et délétée pour le gène *rpoS*. 5 μ g d'ARN totaux ont été déposés par piste du gel d'électrophorèse. Les doubles astérisques ** signalent la réhybridation de la membre par la sonde OmrA, cette membrane ayant été précédemment co-hybridée par les sondes McaS et 5S (vu dans la sous-partie suivante) **C.** Cinétique d'expression de l'ARNnc OmrA en milieu riche (LB) dans les souches WT et $\Delta rpoS$ portant ou non la mutation *ompR234*.

L'activité de la fusion transcriptionnelle d'OmrA présente une faible activité (deux fois supérieure à l'auto-fluorescence) comparée aux autres ARNnc étudiés. En conséquence, il est difficile d'interpréter l'effet des mutations des régulateurs sur cette fusion. Toutefois, l'étude cinétique après déduction du bruit de fond (fluorescence de pZE), et le Northern blot confirme bien que le facteur sigma RpoS est déterminant pour la transcription d'OmrA à 28°C (

Figure 36 : Transcription et accumulation de l'ARNnc OmrA A et B). Curieusement à 37°C, la régulation positive d'OmrA par RpoS n'est visible que dans la souche sauvage et plus dans la souche *ompR234*. Si la mutation d'*ompR* (*ompR234*) n'affecte pas le niveau de transcription de cet ARNnc, elle modifie une fois de plus le rôle régulateur joué par RpoS. Nous serions tentés d'inclure cet ARN

parmi les régulateurs impliqués dans la thermorégulation de l'expression des curli lié à la présence des régulateurs RpoS et OmpR

d) Transcription de l'ARNnc McaS

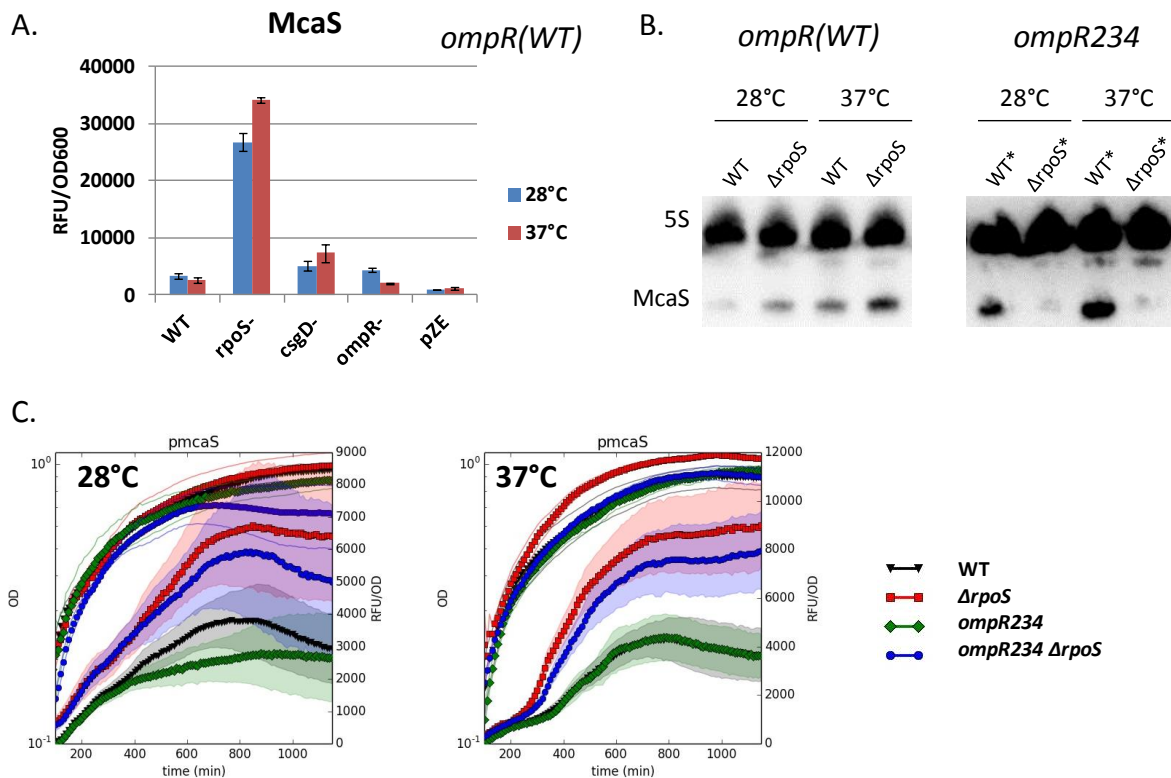


Figure 37 : Transcription et accumulation de l'ARNnc McaS.

A. Les mesures de fluorescence de fusion *mcaS::GFP* ont été effectuées après 22h de culture en milieu riche aux températures de 28°C et 37°C. Les fluorescences brutes ont été divisées par l'absorbance de la culture pour donner des valeurs de fluorescence relative (RFU/OD600). Les valeurs indiquées pour le plasmide vide pZE-Spec renseignent sur le niveau d'autofluorescence. Les souches utilisées sont les souches WT, $\Delta rpoS$, $\Delta csgD$ et $\Delta ompR$ dans un contexte génétique *ompR*-WT pour la fusion transcriptionnelle de McaS. Les barres d'erreurs correspondent à l'écart-type entre les triplicatas biologiques et techniques. **B.** Northern Blot de l'ARNnc McaS, les ARN totaux ont été extraits à une DO=3 par la technique du phénol chaud seulement dans la souche WT et délétée pour le gène *rpoS*. 5µg d'ARN totaux ont été déposés par piste du gel d'électrophorèse. **C.** Cinétique d'expression de l'ARNnc McaS en milieu riche (LB) dans les souches WT et $\Delta rpoS$ portant ou non la mutation *ompR234*.

Nous observons (Figure 37 A) que la transcription de l'ARNnc McaS est fortement réprimée par le facteur RpoS, et qu'essentiellement les régulations par CsgD et OmpR sont négligeables. La transcription de l'ARNnc McaS n'est pas affectée par la température de croissance, pas plus que l'effet répresseur de RpoS. L'étude de la transcription de l'ARNnc McaS en cinétique (Figure 37 C) confirme la régulation négative par RpoS, mais le niveau de répression est moins important que celui observé en lecture ponctuelle (d'un facteur de régulation de 30 en lecture ponctuel, contre un facteur 2 en cinétique) (Figure 37 A). La présence de la mutation *ompR234* n'affecte pas la transcription de McaS ni la répression de RpoS sur McaS observée grâce à l'utilisation de la fusion transcriptionnelle. Toutefois, en Northern Blot (Figure 37 B), si la répression de McaS par RpoS est visible aux deux températures (28°C et 37°C, d'un facteur 2) dans la souche sauvage, elle ne l'est plus dans le contexte *ompR234*. Contrairement aux données obtenues en cinétique (Figure 37 C), le Northern blot montre que l'accumulation de l'ARNnc McaS devient dépendante du facteur RpoS dans un contexte génétique *ompR234*. Nous observons donc une répression de l'initiation de la transcription de McaS, quels que soient l'allèle *ompR*, et une accumulation dépendante du facteur RpoS dans un contexte génétique *ompR234*. Ceci démontre donc que la régulation post-transcriptionnelle (stabilité de McaS) est dépendante de la présence de la mutation *ompR234*. et que celle-ci pourrait conjointement avec le facteur RpoS activer l'expression d'un élément qui empêche la dégradation de l'ARNnc McaS. Nous pouvons également imaginer qu'ils répriment une activité RNase ? Nous avons donc commencé par recherche des séquences d'ADN pouvant coder un anti-sens partiel de McaS qui affecterait sa stabilité. Le résultat de cette recherche est présenté dans la partie III.B.7.

e) **Bilan de régulation de l'expression des curli et des ARNnc**

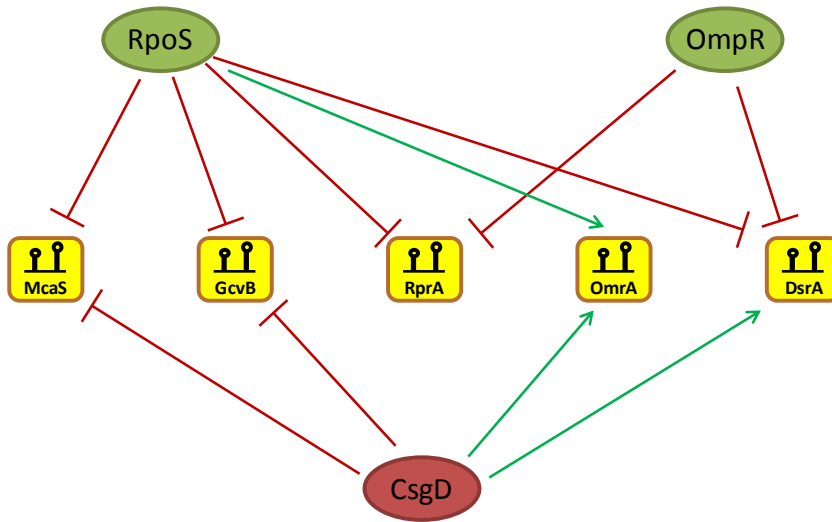


Figure 38 : Bilan sur la régulation transcriptionnelle des ARNnc modulant la traduction de CsgD.

Les facteurs RpoS, OmpR et CsgD régulent la transcription des ARNnc (répression représentée par une flèche rouge et activation représentée par une flèche verte). RpoS réprime la transcription des ARNnc McaS, GcvB, RprA et DsrA et active la transcription de l'ARNnc OmrA. OmpR réprime la transcription de l'ARNnc RprA et DsrA. Enfin le facteur CsgD réprime la transcription des ARNnc McaS et GcvB, et active la transcription des ARNnc OmrA et DsrA.

Dans notre étude, nous démontrons la régulation des ARNnc modulant la traduction de CsgD par les deux modulateurs de la transcription, CsgD et OmpR, et le facteur sigma RpoS. La protéine régulatrice CsgD exerce un rétrocontrôle négatif sur l'ARNnc qui réprime fortement sa traduction. Nous avons également observé une régulation négative de la transcription de l'ensemble de ces ARNnc par le facteur sigma RpoS, qui est très probablement le reflet d'une régulation indirecte. Ce mini sous réseau de régulation est incomplet puisqu'il n'explique pas notamment la régulation négative par le facteur sigma. De plus, des résultats de qRT-PCR (non montrés) laissent penser que RpoS active de la transcription du facteur OmpR en phase stationnaire de croissance ; or RpoS active faiblement la transcription de CsgD à 28°C (Figure 26). À ce stade, nous pensons que la régulation négative par RpoS de l'expression des ARNnc pourrait être indirectement médiée par l'activation, par RpoS, de la transcription des gènes *ompR* et *csgD* qui réprimerait à leur tour physiquement la transcription des ARNnc en se liant à leur promoteur.

7. Stabilité de l'ARNnc McaS

La Figure 37 montre que la transcription de l'ARNnc McaS est réprimée par RpoS, quelque soit l'allèle *ompR*, alors que l'hybridation Northern ne confirme pas cette régulation négative dans la souche *ompR234*. Nous pensons donc que la stabilité de cet ARNnc pourrait être modifiée dans ce contexte génétique par la présence (ou la surabondance par rapport à la souche sauvage) d'un autre ARN modifiant sa stabilité. Ainsi, la formation d'un duplex McaS-ARN-tiers pourrait conduire à sa dégradation par une RNase. Mais pourquoi ce complexe instable se formerait-il uniquement dans la souche $\Delta rpoS$ *ompR234*? Il convenait donc tout d'abord d'identifier un ARN complémentaire de McaS dont l'expression soit régulée par OmpR et RpoS. Nous avons recherché, parmi les partenaires du réseau de régulation, des complémentarités entre séquences d'ARN pour identifier des duplex potentiels. Nous avons ensuite étudié la stabilité de l'ARNnc McaS dans différentes souches mutantes.

En cherchant à expliquer la différence de stabilité de l'ARNnc McaS dans les contextes génétiques $\Delta rpoS$ et *ompR234* $\Delta rpoS$, nous avons identifié une forte complémentarité de séquence entre les ARNnc McaS et OmrA qui est présentée dans la Figure 39. Or la transcription de l'ARNnc OmrA est activée par OmpR (Figure 36) et par RpoS à 28°C (Figure 36).

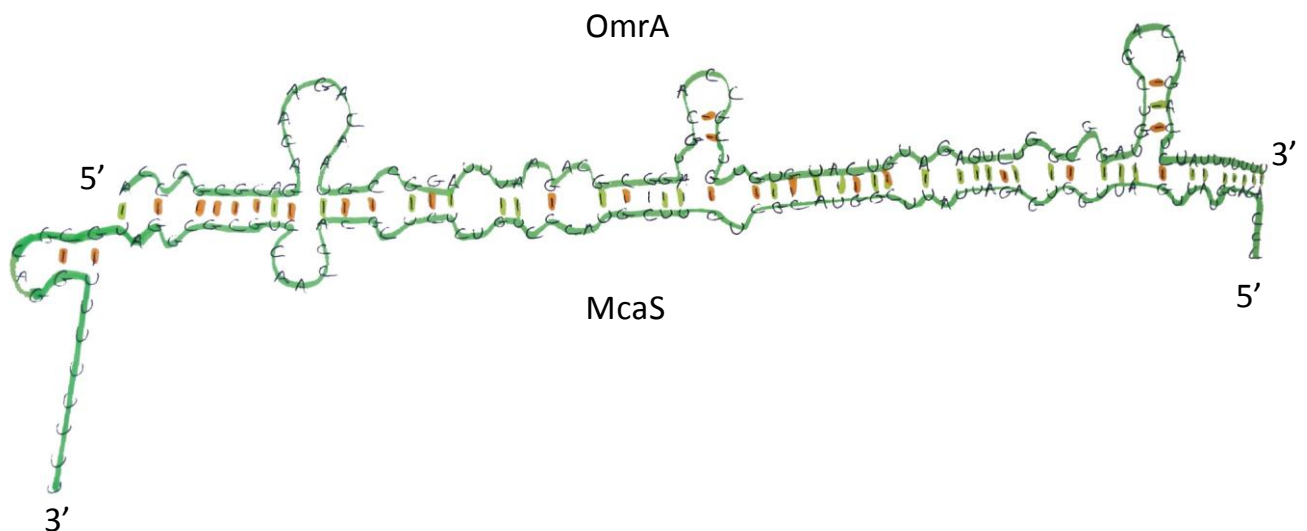


Figure 39 : Complémentarité entre les ARNnc McaS et OmrA.

La séquence des ARNnc McaS et OmrA est surlignée par un trait continu vert foncé. Les petits traits orange entre ces ARN correspondent à une interaction entre nucléotides G-C et les traits vert clair à une interaction entre nucléotides G-U ou A-U.

Pour confirmer notre hypothèse d'interaction entre ARN conduisant à la déstabilisation de McaS, nous avons étudié la demi-vie de McaS dans les contextes génétiques suivants compte tenu de sa complémentarité avec OmrA : WT, *ompR234*, $\Delta rpoS$ *ompR234* et $\Delta omrAB$ *ompR234* (Figure 40).

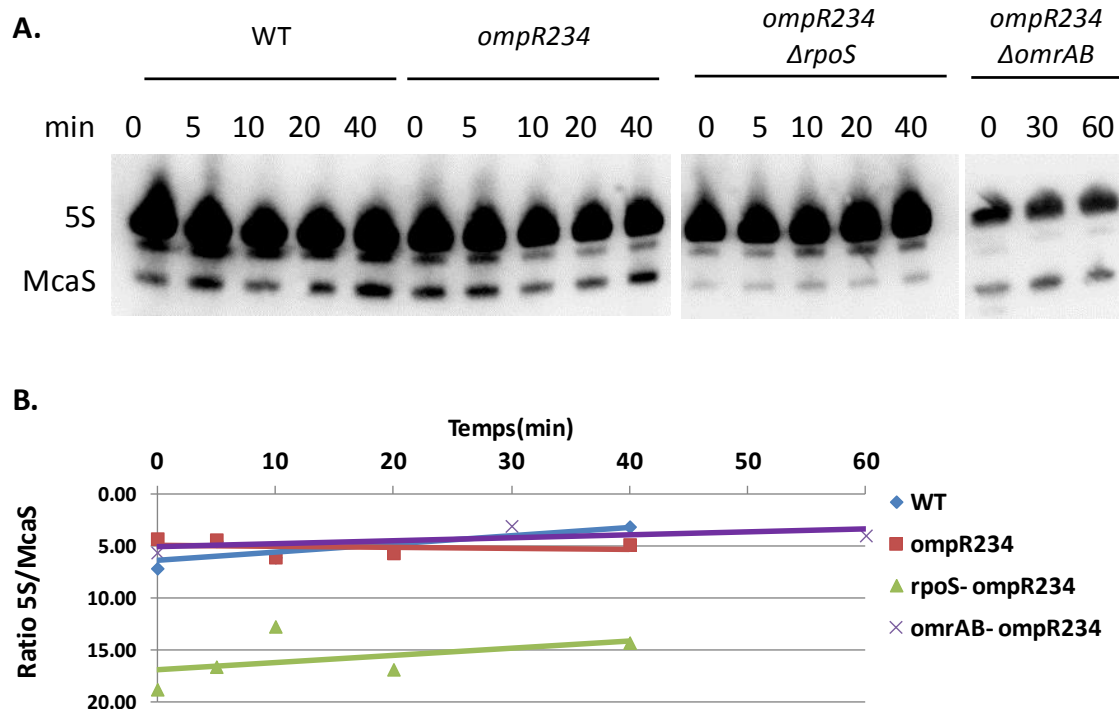


Figure 40 : Détermination de la demi-vie de l'ARNnc McaS.

A. Les souches WT, *ompR234*, *ompR234* $\Delta rpoS$ et *ompR234* $\Delta omrAB$ sont cultivées en milieu riche (LB) à 37°C. Les ARN totaux sont extraits à différents temps après l'ajout de rifampicine (300µg/ml final). 5µg d'ARN totaux sont déposés par piste. **B.** Quantification de l'ARNnc McaS. Les courbes représentent le ratio de la quantité des ARNr stables 5S et de l'ARNnc McaS mesurée à l'aide du logiciel ImageJ.

Dans la souche sauvage, le niveau d'ARN McaS reste stable pendant 40 minutes après l'ajout de la rifampicine, ce qui démontre que cet ARNnc est très stable. Cette stabilité de l'ARNnc McaS n'est pas affectée par la présence de la mutation *ompR234* ou de la mutation $\Delta omrAB$ (toujours dans le contexte génétique *ompR234*). La présence des ARNnc OmrA et OmrB ne semble donc pas influencer la stabilité de McaS. L'expérience pour déterminer la demi-vie confirme toutefois que l'accumulation de McaS dépend de la présence de RpoS dans la souche *ompR234*, et confirme notre précédente observation (Figure 37). L'absence de RpoS, dans la souche *ompR234* uniquement, favorise la transcription d'un élément qui déstabiliserait l'ARN McaS. L'expression de cet élément pourrait donc être négativement régulée par RpoS (indirectement) et positivement par OmpR, les deux régulations antagonistes étant liées à l'activité constitutive d'OmpR. L'alternative possible serait que l'absence de

RpoS, dans le contexte *ompR234* seulement, pourrait conduire à l'absence d'induction de l'élément déstabilisant McaS. Cette régulation n'est pas facile à appréhender compte tenu du fait que RpoS réprime aussi la transcription de McaS dans les souches WT et *ompR234* (Figure 37). De plus, ce n'est pas la stabilité de McaS qui est modifiée dans le contexte $\Delta rpoS$ *ompR234*, mais seulement le niveau d'expression de McaS, sa stabilité n'étant pas affectée 60 min. après ajout de l'antibiotique (rifampicine) bloquant la synthèse d'ARN.

L'interaction probable entre McaS et OmrA n'expliquerait donc pas l'absence de McaS dans la souche $\Delta rpoS$ *ompR234*. Nous avons étudié l'expression et l'accumulation d'une autre séquence largement complémentaire de McaS : la région 5'UTR de CsgD qui est une des cibles géniques de McaS. En effet, nous avons constaté un lien étroit entre l'accumulation de l'ARNnc McaS et celle de l'ARNm CsgD (Figure 41).

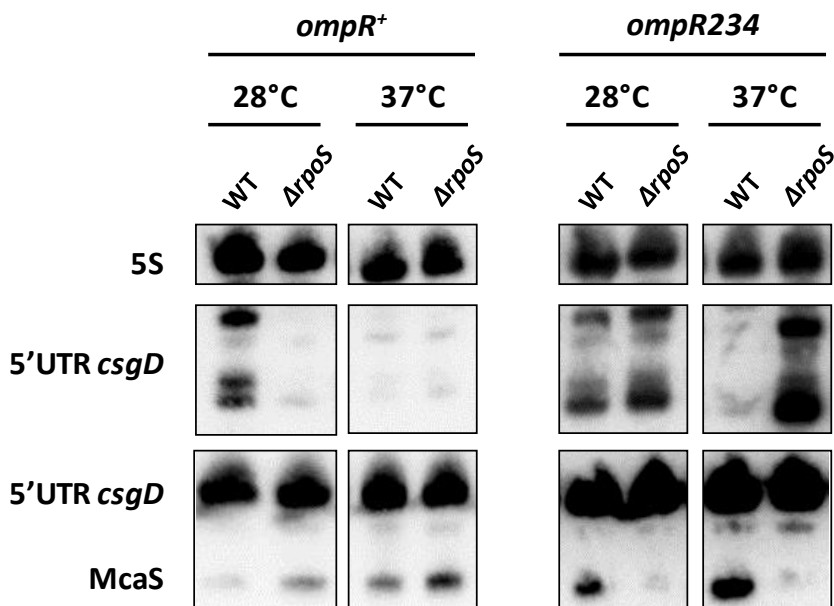


Figure 41 : Comparaison de l'accumulation de l'ARNnc McaS et de la région 5'UTR *csgD*

Northern Blot de l'ARNnc McaS et de la 5'UTR de l'ARNm *csgD*. Les ARN totaux ont été extraits à l'entrée de la phase stationnaire de croissance (DO=3) par la technique d'extraction des ARN au phénol chaud. 5µg d'ARN totaux ont été déposés par piste. Les ARN détectés pour la sonde [Btn] ARN-*csgD* (5'UTR de l'ARNm CsgD) ont une taille approximée à 90, 80 et 60 nts. L'ARNnc détecté par la sonde [Btn] McaS à une taille de 90nts

Nous observons (Figure 41) la présence d'ARN de petite taille (60 à 90 nucléotides) détectés grâce à une sonde anti-sens du 5'UTR de l'ARNm de CsgD. La présence de ces ARN peut soit correspondre à l'expression du gène *csgD* ou seulement celle de son 5'UTR qui pourrait s'accumuler indépendamment. En effet, la région 5'UTR possédant une forte structure secondaire (Figure 42), les fragments d'ARNm sont peut-être des produits d'un procédé de clivage des séquences non appariées et par conséquent du reste de l'ARNm polycistronique *csgDEFG*. De plus, les conditions favorables à

l'accumulation de cet ARN 5'UTR *csgD* ne reflètent pas à celles que nous avons pu observer au moyen de la fusion transcriptionnelle *csgD* ::GFP (TXN) (Figure 30): par exemple, ces ARN ne sont pas détectés à 37°C (*ompR*-WT) et sont fortement régulés par RpoS dans le contexte *ompR* sauvage. En comparant la présence de l'ARNnc McaS et de ces fragments d'ARNm de CsgD, nous pouvons observer une corrélation entre l'absence de l'un et la présence de l'autre espèce d'ARN dans chacune des souches ou conditions (hormis *ompR*-WT-37°C). Les fragments d'ARNm *csgD* sont majoritaires présents quand l'ARNnc McaS est minoritaire, et vice-versa. L'ARNnc McaS réprime la traduction de l'ARNm CsgD par un large appariement créant une structure secondaire sur la partie 5'UTR de l'ARNm CsgD (Figure 43). La formation du duplex entre l'ARNm CsgD et l'ARNnc McaS pourrait être reconnue par une RNase dégradant ce duplex d'ARN. Cette hypothèse est consolidée par la régulation de la transcription de *csgD* par RpoS et OmpR.

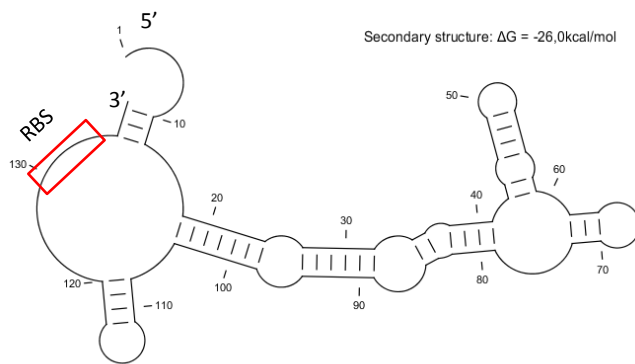


Figure 42 : Structure secondaire de la région 5'UTR de l'ARNm de CsgD.

Image générée par le logiciel CLC workbench par l'outil de prédiction de structure secondaire.

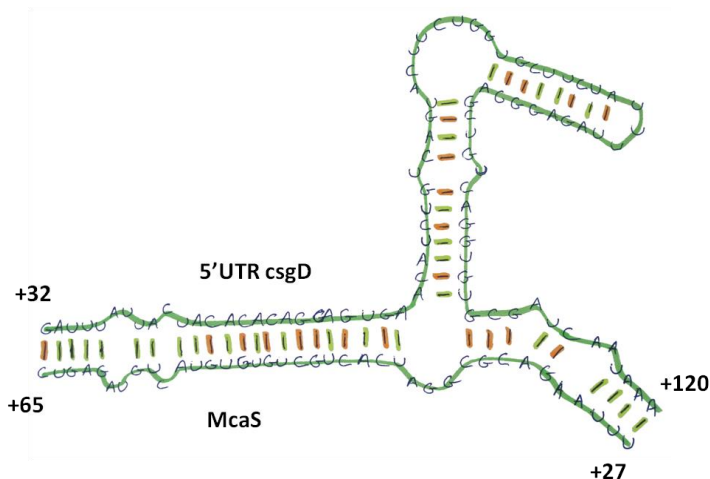


Figure 43 : Appariement de la partie 5'UTR de l'ARNm CsgD et de l'ARNnc McaS

Illustration adaptée de (Jorgensen et al., 2012) après corrections des différences de séquences observées entre la souche utilisée par ces auteurs (*E. coli* K-12 strain SØ928) et notre souche du laboratoire BW25113. Les traits continus vert foncé soulignent les ARN McaS et *csgD* 5'UTR; les petits traits orange correspondent à une interaction entre nucléotides G-C, les traits vert clair correspondent à une interaction entre nucléotides G-U ou A-U.

Il semblerait donc que le promoteur de *csgD* puisse contrôler l'expression de la région 5'UTR de CsgD qui s'accumulerait dans la cellule indépendamment de l'ARNm de *csgD* ou de l'opéron *csgDEFG*. Ce court ARN pourrait moduler l'effet répresseur prédominant de McaS sur la traduction de CsgD.

8. Conclusions du chapitre

Nous avons revisité la régulation de l'expression des curli dans une souche unique d'*E.coli*, BW25113 qui dérive de la souche MG1655 (souche de référence). Nous avons pu vérifier l'importance du facteur OmpR pour la transcription de *csgD* et l'importance de RpoS et CsgD pour la transcription de *csgB*. OmpR semble réguler la transcription de *csgB* indirectement en régulant celle de *csgD*. Nos résultats suggèrent que seul McaS régule fortement la traduction de CsgD et l'ARNnc GcvB influencerait la transcription des gènes *csg* indirectement plutôt que de réguler la traduction de CsgD comme suggéré par (Jorgensen et al., 2012). Les mêmes facteurs de transcription régulent l'expression des gènes codant les ARNnc qui répriment la traduction de CsgD. Enfin nous avons identifié des interactions probables sur plusieurs dizaines de nucléotides entre les ARNnc OmrA et McaS et le 5'UTR de *csgD*. Ceci suggère des interactions importantes entre ARN modulateurs de l'expression des curli. Enfin, nous n'expliquons pas encore la thermorégulation de l'expression des curli.

IV. Conclusions et Perspectives

A. Identification d'ARNnc régulés par RpoS

1. Les ARNnc régulés par RpoS

Nous avons pu identifier des ARNnc régulés par le facteur sigma RpoS. La régulation de certains de ces ARNnc par RpoS a été confirmée par une étude récente réalisée chez *Salmonella enterica serovar Typhimurium* (Levi-Meyrueis et al., 2014). Elle confirme notamment la régulation des ARNnc identifiés lors de notre RNAseq : DsrA, GcvB, GlmY, MicA, OmrA, RyeB, RyeF, RyfD et RyhB (Tableau 6, page 39). Nous confirmons la régulation de ces ARNnc par l'utilisation de fusions transcriptionnelles (DsrA, OmrA, GcvB) (pages 69 et 71) ou en Northern blot (OmrA) (page 71). Une partie des ARNnc identifiés dans notre RNAseq n'ayant pas été validés par l'étude chez *Salmonella*, nous pourrions examiner la régulation de ces ARNnc par ces différentes approches : qRT-PCR, Northern Blot, fusion transcriptionnelle et transcription *in vitro*. Il est aussi possible que nous identifions des ARNnc régulés par RpoS et spécifiques à *E. coli*. De nombreux ARNnc ont été prédits chez *E. coli* (Herbig and Nieselt, 2011; Shinhara et al., 2011; Sorek and Cossart, 2010), mais leur existence physique doit être validée par Northern Blot. Cette partie du projet est largement inachevée et il conviendra d'identifier les cibles cellulaires de ces nouveaux ARNnc.

2. Régulation négative des ARNnc par RpoS

À notre grande surprise, une majorité des ARNnc sont négativement régulés par RpoS et seul l'ARNnc OmrA est activé par le facteur sigma RpoS (Figure 38, page 74). Ces répressions par le facteur sigma RpoS sont sûrement indirectes et dues à l'induction de l'expression d'un ou plusieurs répresseurs intermédiaires non identifiés. Nous avons aussi observé une activation du facteur de transcription OmpR par RpoS (qRT-PCR, résultat à confirmer). Sachant que la mutation $\Delta ompR$ affecte l'expression de certains ARNnc (RprA, DsrA, Figure 38, page 74), ce régulateur pourrait être un répresseur médiant la répression des ARNnc par RpoS. OmpR n'est pas le seul régulateur transcriptionnel régulé par RpoS, sont dénombrés à environ 40 (réf de gène arrays). Nous disposons, au laboratoire, d'une collection de mutants non essentiels de la souche BW25113, ainsi nous pourrions étudier la transcription des ARNnc dans les souches mutantes de ces régulateurs globaux.

3. Un nouvel ARN identifié : Djs

Une des découvertes importantes de ce travail est certainement l'identification d'un nouvel ARNnc en multiples copies au sein du génome des bactéries de la famille des *Enterobacteriaceae*, baptisé Djs. Cet ARN a pu être détecté en Northern blot (Figure 19, page 50), mais nous n'avons pas pu déterminer quelle copie était exprimée. Des expériences d'extension d'amorces confirment l'expression de multiples copies de Djs de taille variable (Figure 20, page 51). Afin d'identifier les copies de Djs exprimées, nous pourrions faire séquencer les ARN générés par extension d'amorce. Pour mieux appréhender le rôle de cet ARN, nous pourrions déléter les copies exprimées du génome afin d'étudier leur rôle sur des phénotypes simples (exemples : l'impact de la délétion sur le taux de croissance, la motilité et la coloration au Rouge Congo). Nous avons détecté certaines de ces séquences en RNAseq (5 loci/15) et également par CHIP-seq, un site de fixation préférentielle de RpoS à proximité d'une de ces copies. Nous ne pensons pas que toutes les copies de cet ARN sont exprimées, certaines copies étant localisées dans des opérons. Nous postulons qu'une des copies exprimées puisse s'apparier par complémentarité avec un ARNm polycistronique et pour en réguler la stabilité ou la traduction du gène en aval du site d'appariement. Des essais de transcription *in vitro* de certaines de ces séquences par les holoenzymes $E\sigma^{70}$ et $E\sigma^S$ purifiées n'ont pas permis d'obtenir de transcrit, ce qui nous suggère deux hypothèses : les copies testées ne sont pas les copies qui sont exprimées dans la cellule, ou leur transcription est dépendante d'un cofacteur transcriptionnel de l'ARN polymérase absent dans nos essais. La surexpression de l'une des copies de cet ARN affecte également le phénotype de liaison au Rouge Congo et entraîne la décoloration des colonies formées par une souche sauvage (WT) transformée par le plasmide de sur-expression de Djs (Figure 18, page 49). Bien que le mécanisme et les gènes ciblés ne soient pas identifiés, une analyse *in silico* de l'appariement entre la séquence de Djs et l'ARNm du régulateur CsgD, ou de l'ARNm des gènes structuraux des curli, *csgBA*, serait utile pour pouvoir comprendre ce phénotype. Nous comptons, par cette étude, identifier un ARNnc régulé par RpoS et impliqué dans la régulation de l'expression des curli, ainsi Djs est potentiellement un bon candidat pour être l'un de ces régulateurs.

B. Les acteurs majeurs de la régulation de l'expression des curli

Nous avons revisité la régulation de l'expression des curli afin d'en identifier les régulateurs majeurs. Ne pouvant pas observer l'accumulation des protéines structurales extra-membranaires et du régulateur CsgD par Western blot faute de posséder des anticorps adéquats (non commercialisés), nous avons étudié la régulation transcriptionnelle et traductionnelle de ces gènes au moyen de fusions reportrices et d'hybridation en Northern Blot. Nous avons pu confirmer l'importance du facteur de transcription OmpR qui est essentiel à la transcription du régulateur des gènes structuraux *csgBA*, CsgD (Figure 26, page 57). Nous avons aussi montré une forte régulation traductionnelle de CsgD par l'ARNnc McaS (Figure 28, page 61).

1. La transcription des gènes *csg* est activée par l'ARNnc GcvB

A notre grande surprise nous avons observé que l'ARNnc, GcvB, favorise la transcription de l'ensemble des gènes *csg* (Figure 30, page 63), mais régule peu la traduction de CsgD (Figure 28, page 61). Le seul processus de régulation transcriptionnelle par les ARNnc décrits chez *E. coli* concerne l'ARNnc 6S qui inhibe l'activité de l'holoenzyme σ^{70} en imitant un promoteur et en titrant cette holoenzyme de par son abondance dans le cytoplasme (Sharma and Chatterji, 2010). Il est donc fort probable que GcvB influence plutôt indirectement la transcription des opérons *csg*. D'après la littérature, GcvB réprime la traduction d'un système à deux-composantes (PhoQP) (Coornaert et al., 2013) qui pourrait modifier la phosphorylation du facteur OmpR nécessaire à son activation. Cependant, nous observons la même activation, en absence de l'ARNnc GcvB ($\Delta gcvB$), des opérons *csg* dans la souche sauvage (WT) et la souche *ompR234* où la protéine OmpR est active sans phosphorylation (Figure 30, page 63). Ainsi, l'ARNnc GcvB ne semble pas activer la transcription des gènes *csg* en modulant la phosphorylation du facteur OmpR. En conséquence, il serait intéressant d'étudier la traduction des répresseurs de l'opéron *csgBA*, CpxR et FlhZ, dans la souche sauvage (WT) et le mutant $\Delta gcvB$, pour voir si l'activation de la transcription des gènes *csgBA* passe par la régulation de ces répresseurs par GcvB.

2. Répression transcriptionnelle et traductionnelle de CsgD par l'ARNnc McaS

Nous avons pu confirmer la répression traductionnelle de CsgD par l'ARNnc McaS uniquement dans les souches *ompR234* (Figure 28, page61), mais nous avons observé une forte répression transcriptionnelle des gènes *csg* par l'ARNnc McaS (souche $\Delta mcaS$ et *ompR234* $\Delta mcaS$) et (Figure 30, page63). Cette répression traductionnelle peut potentiellement être expliquée par l'activation traductionnelle des régulateurs FlhDC (dimère) (Thomason et al., 2012). Le dimère FlhDC active la transcription du gène *fliZ* (Lee et al., 2011) qui code un répresseur transcriptionnel des gènes *csg* (Pesavento et al., 2008). Ainsi la répression par l'ARNnc McaS de la transcription des gènes *csg* pourrait être médiée par FliZ. Une expérience comparant la transcription des gènes *csg* dans une souche WT, $\Delta mcaS$ et un double mutant $\Delta mcaS \Delta fliZ$ pourrait répondre à cette interrogation.

3. Modulation de la traduction du facteur OmpR

L'ARNnc OmrA est décrit comme répresseur traductionnelle d'OmpR (Guillier and Gottesman, 2006), or nous avons montré que la transcription d'OmrA est activée par le facteur RpoS (Figure 36, page 71) ; ainsi l'ARNnc OmrA pourrait-il médier la répression transcriptionnelle de CsgD, plutôt qu'une régulation traductionnelle, en affectant la traduction de OmpR, activateur essentiel à la transcription de *csgD*. Cependant notre étude ne montre aucun impact de la mutation $\Delta omrAB$ sur la transcription de *csgD*. Seuls les ARNnc GcvB et McaS affectent la transcription des gènes *csg* (Figure 30, page63). Les ARNnc GcvB et McaS pourraient aussi moduler la traduction du facteur OmpR. Ainsi l'étude de la traduction du facteur OmpR dans les mutants $\Delta mcaS$ et $\Delta gcvB$ pourrait mettre en évidence de nouvelles interactions et démontrer qu'OmpR est la cible de ces ARNnc pour réguler la transcription des gènes *csg* en régulant celle de CsgD. Ainsi, ces ARNnc n'influeraient pas la traduction de CsgD, mais plutôt celle du facteur de transcription OmpR qui stimule la transcription de CsgD.

4. Comparaison des méthodologies entre études

Nous n'avons pas confirmé l'ensemble des régulations traductionnelles de CsgD décrites dans la littérature (Figure 23, page54)(Boehm and Vogel, 2012) et qui ont été observées en sur-exprimant les ARNnc. Cette différence d'approche expérimentale peut expliquer les différences de régulation observées. Nous n'avons pas pu étudier l'effet de la surexpression des ARNnc sur nos fusions transcriptionnelles et traductionnelles des gènes *csg*, pour cause d'incompatibilité des plasmides. Nous pourrions néanmoins étudier l'influence sur la transcription des gènes *csg* de la surexpression

des ARNnc par la technique de qRT-PCR. Seul l'ARNnc McaS a pu être identifié comme répresseur traductionnelle de CsgD, mais ceci uniquement dans un contexte génétique *ompR234* (voir IV.B.2). Les ARNnc RprA, et OmrA/B n'ont apparemment qu'un rôle mineur dans la régulation de la traduction de CsgD. Afin de compléter nos observations, il nous paraît nécessaire de faire le parallèle entre les mutations de ces ARNnc et leur surexpression dans une souche WT pour confirmer les régulations et leurs variations dans notre souche d'étude, BW25113. Cette expérience pourra peut-être confirmer l'importance des ARNnc RprA et OmrA/B pour la régulation de l'expression des curli. Pour mieux appréhender les phénotypes observés, nous pourrions également comparer le niveau d'activité de nos fusions transcriptionnelles (*csgD* TXN et *csgB* TXN) dans différentes souches d'*E. coli* (notre souche modèle étant différente des souches utilisées dans les autres études). Si le niveau d'activation des fusions transcriptionnelles est significativement différent, nous pourrions faire l'hypothèse que les régulations que nous avons observées dépendent des souches d'*E. coli* et pourraient donc différer de celles observées chez *Salmonella*.

5. Le rôle versatile et de thermorégulateur du facteur RpoS

Le rôle d'activateur du facteur sigma RpoS a été confirmé uniquement dans le contexte *ompR*-WT à 28°C, et reste un déterminant majeur de l'expression des curli à basse température (Figure 21, page 53). RpoS joue le rôle de répresseur dans le contexte *ompR234* et semble aussi être l'un des facteurs déterminants dans la répression transcriptionnelle de l'opéron *csgD* à 37°C dans le contexte *ompR234* (Figure 26, page 57). RpoS est donc impliqué dans la thermorégulation de l'expression des curli qui ne dépend pas des ARNnc, même si la régulation de certains ARNnc par RpoS est observée à une seule température (exemple RprA, Figure 35, page 70).

6. RpoS et l'ARNm CsgD déterminent l'accumulation de l'ARNnc McaS

Dans un contexte génétique *ompR234*, l'accumulation de l'ARNnc McaS est étroitement liée à la présence de RpoS et de la 5'UTR de l'ARNm CsgD. L'absence du facteur sigma RpoS entraîne une faible accumulation de l'ARNnc McaS et une accumulation accrue de la 5'UTR de l'ARNm CsgD. À l'inverse, la présence du facteur sigma RpoS conduit à l'accumulation de l'ARNnc McaS et une baisse de la quantité de la 5'UTR de l'ARNm CsgD. Nous supposons que RpoS contrôle l'expression d'un répresseur de la transcription du gène *csgD* et donc de sa 5'UTR. Nous savons également que l'appariement de la 5'UTR de l'ARNm *csgD* à l'ARNnc McaS génère un duplex ARNm-ARNnc (Figure 43, page 78) qui pourrait être spontanément reconnu par une RNase qui le dégraderait, très

probablement la RNase III qui dégrade les ARN double brins (Court et al., 2013; Doetsch et al., 2013). Des essais *in vitro* ont démontré une protection de McaS en présence de l'ARNm de *csgD* (Jorgensen et al., 2012), alors que nous observons l'absence de l'ARNnc McaS quand l'accumulation de la 5'UTR de l'ARNm de *CsgD* est importante dans la souche *ompR234 ΔrpoS* (Figure 41, page 77). Pour confirmer que l'accumulation de l'ARNnc McaS dépend de la présence de RpoS et de l'absence de la 5'UTR de l'ARNm de *CsgD*, nous pourrions étudier l'accumulation de l'ARNnc McaS dans les souches *ompR234* et *ΔrpoS ompR234* délétée de la 5'UTR du gène *csgD*.

7. Un nouvel ARNnc, RydC, régulateur traductionnel de CsgD

Une étude récente montre que l'ARNnc RydC réprime la traduction de l'ARNm *CsgD* (Bordeau and Felden, 2014). Il nous paraît essentiel d'étudier les effets de la mutation de cet ARNnc sur nos fusions, car il pourrait s'agir d'un acteur majeur de la thermorégulation de l'expression des curli. De plus, nous n'avons pas réussi à identifier le thermorégulateur de l'expression des curli (qui semble être RpoS-dépendant) ; ainsi, nous pourrions compléter cette étude en étudiant la régulation transcriptionnelle de l'ARNnc RydC par le facteur RpoS, à différentes températures de croissance.

8. Intuition personnelle

Mon intuition personnelle est que la traduction de *CsgD* ou son accumulation est très faible à 37°C malgré une activité transcriptionnelle du gène équivalente à celle observée à 28°C (Figure 26, page 57). L'expression des curli étant dépendante de *CsgD*, seule une étude en Western blot pourra permettre de répondre à cette question. Cependant, il est difficile de se procurer un anticorps pour détecter cette protéine. Une alternative pourrait être d'étiqueter la protéine afin de pouvoir détecter son accumulation comme cela a été fait précédemment (Holmqvist et al., 2010).

V. Bibliographie

- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology* 2, 2006 0008.
- Boehm, A., and Vogel, J. (2012). The *csgD* mRNA as a hub for signal integration via multiple small RNAs. *Molecular microbiology* 84, 1-5.
- Bordeau, V., and Felden, B. (2014). Curli synthesis and biofilm formation in enteric bacteria are controlled by a dynamic small RNA module made up of a pseudoknot assisted by an RNA chaperone. *Nucleic acids research* 42, 4682-4696.
- Bougourd, A., Lelong, C., and Geiselmann, J. (2004). Crl, a low temperature-induced protein in *Escherichia coli* that binds directly to the stationary phase sigma subunit of RNA polymerase. *The Journal of biological chemistry* 279, 19540-19550.
- Brantl, S. (2007). Regulatory mechanisms employed by cis-encoded antisense RNAs. *Current opinion in microbiology* 10, 102-109.
- Coornaert, A., Chiaruttini, C., Springer, M., and Guillier, M. (2013). Post-transcriptional control of the *Escherichia coli* PhoQ-PhoP two-component system by multiple sRNAs involves a novel pairing region of GcvB. *PLoS genetics* 9, e1003156.
- Court, D.L., Gan, J., Liang, Y.H., Shaw, G.X., Tropea, J.E., Costantino, N., Waugh, D.S., and Ji, X. (2013). RNase III: Genetics and function; structure and mechanism. *Annual review of genetics* 47, 405-431.
- De Lay, N., Schu, D.J., and Gottesman, S. (2013). Bacterial small RNA-based negative regulation: Hfq and its accomplices. *The Journal of biological chemistry* 288, 7996-8003.
- Doetsch, M., Stampfl, S., Furtig, B., Beich-Frandsen, M., Saxena, K., Lybecker, M., and Schroeder, R. (2013). Study of *E. coli* Hfq's RNA annealing acceleration and duplex destabilization activities using substrates with different GC-contents. *Nucleic acids research* 41, 487-497.
- Dong, T., Kirchhof, M.G., and Schellhorn, H.E. (2008). RpoS regulation of gene expression during exponential growth of *Escherichia coli* K12. *Molecular genetics and genomics : MGG* 279, 267-277.
- Dong, T., and Schellhorn, H.E. (2009). Control of RpoS in global gene expression of *Escherichia coli* in minimal media. *Molecular genetics and genomics : MGG* 281, 19-33.
- Dong, T., Yu, R., and Schellhorn, H. (2011). Antagonistic regulation of motility and transcriptome expression by RpoN and RpoS in *Escherichia coli*. *Molecular microbiology* 79, 375-386.
- Dorel, C., Vidal, O., Prigent-Combaret, C., Vallet, I., and Lejeune, P. (1999). Involvement of the Cpx signal transduction pathway of *E. coli* in biofilm formation. *FEMS microbiology letters* 178, 169-175.
- Fozo, E.M., Kawano, M., Fontaine, F., Kaya, Y., Mendietta, K.S., Jones, K.L., Ocampo, A., Rudd, K.E., and Storz, G. (2008). Repression of small toxic protein synthesis by the Sib and OhsC small RNAs. *Molecular microbiology* 70, 1076-1093.
- Frohlich, K.S., Papenfort, K., Berger, A.A., and Vogel, J. (2012). A conserved RpoS-dependent small RNA controls the synthesis of major porin OmpD. *Nucleic acids research* 40, 3623-3640.
- Frohlich, K.S., and Vogel, J. (2009). Activation of gene expression by small RNA. *Current opinion in microbiology* 12, 674-682.
- Guillier, M., and Gottesman, S. (2006). Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. *Molecular microbiology* 59, 231-247.
- Han, K., Kim, K.S., Bak, G., Park, H., and Lee, Y. (2010). Recognition and discrimination of target mRNAs by Sib RNAs, a cis-encoded sRNA family. *Nucleic acids research* 38, 5851-5866.
- Henderson, C.A., Vincent, H.A., Casamento, A., Stone, C.M., Phillips, J.O., Cary, P.D., Sobott, F., Gowers, D.M., Taylor, J.E., and Callaghan, A.J. (2013). Hfq binding changes the structure of *Escherichia coli* small noncoding RNAs OxyS and RprA, which are involved in the riboregulation of *rpoS*. *Rna* 19, 1089-1104.
- Henkin, T.M. (1996). Control of transcription termination in prokaryotes. *Annual review of genetics* 30, 35-57.

Herbig, A., and Nieselt, K. (2011). nocoRNAC: characterization of non-coding RNAs in prokaryotes. *BMC bioinformatics* 12, 40.

Holmqvist, E., Reimegard, J., Sterk, M., Grantcharova, N., Romling, U., and Wagner, E.G. (2010). Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. *The EMBO journal* 29, 1840-1850.

Holmqvist, E., and Vogel, J. (2013). A small RNA serving both the Hfq and CsrA regulons. *Genes & development* 27, 1073-1078.

Huttenhofer, A., and Vogel, J. (2006). Experimental approaches to identify non-coding RNAs. *Nucleic acids research* 34, 635-646.

Ishihama, A. (2000). Functional modulation of Escherichia coli RNA polymerase. *Annual review of microbiology* 54, 499-518.

Jorgensen, M.G., Nielsen, J.S., Boysen, A., Franch, T., Moller-Jensen, J., and Valentin-Hansen, P. (2012). Small regulatory RNAs control the multi-cellular adhesive lifestyle of Escherichia coli. *Molecular microbiology* 84, 36-50.

Jubelin, G., Vianney, A., Beloin, C., Ghigo, J.M., Lazzaroni, J.C., Lejeune, P., and Dorel, C. (2005). CpxR/OmpR interplay regulates curli gene expression in response to osmolarity in Escherichia coli. *Journal of bacteriology* 187, 2038-2049.

Kraev, A.S. (1988). [A simple system of cloning in phage M13 and DNA sequencing with terminators]. *Molekuliarnaia biologija* 22, 1164-1197.

Kroger, M. (1989). Compilation of DNA sequences of Escherichia coli. *Nucleic acids research* 17 *Suppl*, r283-309.

Kroger, M., and Wahl, R. (1996). Compilation of DNA sequences of Escherichia coli K12 (ECD and ECDC; update 1995). *Nucleic acids research* 24, 29-31.

Kroger, M., and Wahl, R. (1997). Compilation of DNA sequences of Escherichia coli K12: description of the interactive databases ECD and ECDC (update 1996). *Nucleic acids research* 25, 39-42.

Kroger, M., and Wahl, R. (1998). Compilation of DNA sequences of Escherichia coli K12: description of the interactive databases ECD and ECDC. *Nucleic acids research* 26, 46-49.

Lacour, S., Kolb, A., and Landini, P. (2003). Nucleotides from -16 to -12 determine specific promoter recognition by bacterial sigmaS-RNA polymerase. *The Journal of biological chemistry* 278, 37160-37168.

Lee, Y.Y., Barker, C.S., Matsumura, P., and Belas, R. (2011). Refining the binding of the Escherichia coli flagellar master regulator, FlhD4C2, on a base-specific level. *Journal of bacteriology* 193, 4057-4068.

Levi-Meyrueis, C., Monteil, V., Sismeiro, O., Dillies, M.A., Monot, M., Jagla, B., Coppee, J.Y., Dupuy, B., and Norel, F. (2014). Expanding the RpoS/sigmaS-Network by RNA Sequencing and Identification of sigmaS-Controlled Small RNAs in Salmonella. *PloS one* 9, e96918.

Li, S., Dong, X., and Su, Z. (2013). Directional RNA-seq reveals highly complex condition-dependent transcriptomes in E. coli K12 through accurate full-length transcripts assembling. *BMC genomics* 14, 520.

Liu, M.Y. (1997a). The RNA Molecule CsrB Binds to the Global Regulatory Protein CsrA and Antagonizes Its Activity in Escherichia coli. *Journal of Biological Chemistry* 272, 17502-17510.

Liu, M.Y.R.T. (1997b). The Global Regulator CsrA of Escherichia coli Is a Specific mRNA-Binding Protein. *Journal of bacteriology* 179, 4639-4642.

Lukjancenکو, O., Wassenaar, T.M., and Ussery, D.W. (2010). Comparison of 61 sequenced Escherichia coli genomes. *Microbial ecology* 60, 708-720.

Maciag, A., Peano, C., Pietrelli, A., Egli, T., De Bellis, G., and Landini, P. (2011). In vitro transcription profiling of the sigmaS subunit of bacterial RNA polymerase: re-definition of the sigmaS regulon and identification of sigmaS-specific promoter sequence elements. *Nucleic acids research* 39, 5338-5355.

McCullen, C.A., Benhammou, J.N., Majdalani, N., and Gottesman, S. (2010). Mechanism of positive regulation by DsrA and RprA small noncoding RNAs: pairing increases translation and protects rpoS mRNA from degradation. *Journal of bacteriology* 192, 5559-5571.

Mika, F., Busse, S., Possling, A., Berkholtz, J., Tschowri, N., Sommerfeldt, N., Pruteanu, M., and Hengge, R. (2012). Targeting of *csgD* by the small regulatory RNA RprA links stationary phase, biofilm formation and cell envelope stress in *Escherichia coli*. *Molecular microbiology* *84*, 51-65.

Ogasawara, H., Yamada, K., Kori, A., Yamamoto, K., and Ishihama, A. (2010). Regulation of the *Escherichia coli* *csgD* promoter: interplay between five transcription factors. *Microbiology* *156*, 2470-2483.

Ogasawara, H., Yamamoto, K., and Ishihama, A. (2011). Role of the biofilm master regulator CsgD in cross-regulation between biofilm formation and flagellar synthesis. *Journal of bacteriology* *193*, 2587-2597.

Opdyke, J.A., Fozo, E.M., Hemm, M.R., and Storz, G. (2011). RNase III participates in GadY-dependent cleavage of the *gadX-gadW* mRNA. *Journal of molecular biology* *406*, 29-43.

Papenfort, K., Bouvier, M., Mika, F., Sharma, C.M., and Vogel, J. (2010). Evidence for an autonomous 5' target recognition domain in an Hfq-associated small RNA. *Proceedings of the National Academy of Sciences of the United States of America* *107*, 20435-20440.

Patten, C.L., Kirchoff, M.G., Schertzberg, M.R., Morton, R.A., and Schellhorn, H.E. (2004). Microarray analysis of RpoS-mediated gene expression in *Escherichia coli* K-12. *Molecular genetics and genomics* : MGG *272*, 580-591.

Pesavento, C., Becker, G., Sommerfeldt, N., Possling, A., Tschowri, N., Mehlis, A., and Hengge, R. (2008). Inverse regulatory coordination of motility and curli-mediated adhesion in *Escherichia coli*. *Genes & development* *22*, 2434-2446.

Prigent-Combaret, C., Brombacher, E., Vidal, O., Ambert, A., Lejeune, P., Landini, P., and Dorel, C. (2001). Complex regulatory network controls initial adhesion and biofilm formation in *Escherichia coli* via regulation of the *csgD* gene. *Journal of bacteriology* *183*, 7213-7223.

Prigent-Combaret, C., Prensier, G., Le Thi, T.T., Vidal, O., Lejeune, P., and Dorel, C. (2000). Developmental pathway for biofilm formation in curli-producing *Escherichia coli* strains: role of flagella, curli and colanic acid. *Environmental microbiology* *2*, 450-464.

Reichenbach, B., Maes, A., Kalamorz, F., Hajnsdorf, E., and Gorke, B. (2008). The small RNA GlmY acts upstream of the sRNA GlmZ in the activation of *glmS* expression and is subject to regulation by polyadenylation in *Escherichia coli*. *Nucleic acids research* *36*, 2570-2580.

Reshamwala, S.M., and Noronha, S.B. (2011). Biofilm formation in *Escherichia coli* *cra* mutants is impaired due to down-regulation of curli biosynthesis. *Archives of microbiology* *193*, 711-722.

Serra-Moreno, R., Acosta, S., Hernalsteens, J.P., Jofre, J., and Muniesa, M. (2006). Use of the lambda Red recombinase system to produce recombinant prophages carrying antibiotic resistance genes. *BMC molecular biology* *7*, 31.

Sharma, U.K., and Chatterji, D. (2010). Transcriptional switching in *Escherichia coli* during stress and starvation by modulation of sigma activity. *FEMS microbiology reviews* *34*, 646-657.

Shinhara, A., Matsui, M., Hiraoka, K., Nomura, W., Hirano, R., Nakahigashi, K., Tomita, M., Mori, H., and Kanai, A. (2011). Deep sequencing reveals as-yet-undiscovered small RNAs in *Escherichia coli*. *BMC genomics* *12*, 428.

Silva, I.J., Ortega, A.D., Viegas, S.C., Garcia-Del Portillo, F., and Arraiano, C.M. (2013). An RpoS-dependent sRNA regulates the expression of a chaperone involved in protein folding. *Rna* *19*, 1253-1265.

Sledjeski, D.D., Gupta, A., and Gottesman, S. (1996). The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in *Escherichia coli*. *The EMBO journal* *15*, 3993-4000.

Soper T, M.P., Majdalani N, Gottesman S, Woodson SA. (2010). Positive regulation by small RNAs and the role of Hfq. *PNAS* *107*, 9602-9607.

Sorek, R., and Cossart, P. (2010). Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nature reviews Genetics* *11*, 9-16.

Sridhar, J., Sambaturu, N., Sabarinathan, R., Ou, H.Y., Deng, Z., Sekar, K., Rafi, Z.A., and Rajakumar, K. (2010). sRNAscanner: a computational tool for intergenic small RNA detection in bacterial genomes. *PLoS one* *5*, e11970.

- Thomason, L.C., Costantino, N., and Court, D.L. (2007). E. coli genome manipulation by P1 transduction. *Current protocols in molecular biology* / edited by Frederick M Ausubel [et al] *Chapter 1*, Unit 1 17.
- Thomason, M.K., Fontaine, F., De Lay, N., and Storz, G. (2012). A small RNA that regulates motility and biofilm formation in response to changes in nutrient availability in *Escherichia coli*. *Mol Microbiol* *84*, 17-35.
- Toffano-Nioche, C., Luo, Y., Kuchly, C., Wallon, C., Steinbach, D., Zytnicki, M., Jacq, A., and Gautheret, D. (2013). Detection of non-coding RNA in bacteria and archaea using the DETR'PROK Galaxy pipeline. *Methods* *63*, 60-65.
- Tomizawa, J.I.T.S.G.S.T. (1981). Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. *Proceedings of the National Academy of Sciences of the United States of America* *78*, 1421-1425.
- Vidal, O., Longin, R., Prigent-Combaret, C., Dorel, C., Hooreman, M., and Lejeune, P. (1998). Isolation of an *Escherichia coli* K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new *ompR* allele that increases curli expression. *Journal of bacteriology* *180*, 2442-2449.
- Wahl, R., Rice, P., Rice, C.M., and Kroger, M. (1994). ECD--a totally integrated database of *Escherichia coli* K12. *Nucleic acids research* *22*, 3450-3455.
- Wassarman, K.M. (2007). 6S RNA: a small RNA regulator of transcription. *Current opinion in microbiology* *10*, 164-168.
- Waters, L.S., and Storz, G. (2009). Regulatory RNAs in bacteria. *Cell* *136*, 615-628.
- White-Ziegler, C.A., Um, S., Perez, N.M., Berns, A.L., Malhowski, A.J., and Young, S. (2008). Low temperature (23 degrees C) increases expression of biofilm-, cold-shock- and RpoS-dependent genes in *Escherichia coli* K-12. *Microbiology* *154*, 148-166.
- Zhang, A., Wassarman, K.M., Rosenow, C., Tjaden, B.C., Storz, G., and Gottesman, S. (2003). Global analysis of small RNA and mRNA targets of Hfq. *Molecular microbiology* *50*, 1111-1124.
- Zheng, D., Constantinidou, C., Hobman, J.L., and Minchin, S.D. (2004). Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic acids research* *32*, 5874-5893.

ANNEXE 1

Nom de l'amorce	Séquence 5' -> 3'	Utilisation(s)
mcaS_Fwd	TATCTCGAGTTCATGTTACATCCTCCCGC	pTOPO McaS et mcaS::GFP
mcaS_Prom	TATGAATTCGCGCCGGTGACTGTGTTATA	mcaS::GFP
mcsA_Rev	CGGCTATCTGCAAAGTTAAAAC	pTOPO McaS et mcaS::GFP
omrA_Fwd	TATCTCGAGGCCGGTCATCAATCTGTAAC	omrA::GFP
omrA_Prom	TATGAATTCCTCTGGGATCTTGATTGTG	omrA::GFP
omrB_Fwd	TATCTCGAGTCAGTGTTACGGAAAACGCC	omrB::GFP
omrB_Prom	TATGAATTCCTCTGGGATCACCACCTTTA	omrB::GFP
rprA_Fwd	TATCTCGAGAATTTCTCATCAGGCGGCTC	rprA::GFP
rprA_Prom	TATGAATTCCTGTAGCTAATAGTAGGCAT AC	rprA::GFP
gcvB_Fwd	TATCTCGAGTGAAACTTAAATGGCGTGTC	pTOPO GcvB et gcvB::GFP
gcvB_Prom	TATGAATTCGCCGCCACTATAGGTATTTG	gcvB::GFP
gcvB_Rev	CGTGAAAGAGATGGTCAAC	pTOPO GcvB
arcZ_Fwd	TATCTCGAGCAACACCGCTTCATGAATTT	pTOPO ArcZ et arcZ::GFP
arcZ_Rev	TAGAAGTGCTGAAAGCGTGG	pTOPO ArcZ
arcZ_Prom	TATGAATTCGCCGCACGCAGTTAAGATAG	arcZ::GFP
rybB_Fwd	TATCTCGAGATACAGCTCGCAGTAACATC	rybB::GFP
rybB_Prom	TATGAATTCTCAAAGAAAAGCAGTGGCAC	rybB::GFP
MicC_Fwd	TATCTCGAGGCCCTTTATCCCCATTTCCGG	micC::GFP
MicC_Prom	TATGAATTCGCCGCAGATAAACGTATAG	micC::GFP
X5_tpke11	AAACTGCTCGAGAAGGTGAAGAC	tpke11::GFP et pTOPO tpke11
E3_tpke11	ATCTCGAATTCATCTTGCTTAGCC	pTOPO tpke11
E3p_tpke11	TCAGTAGAATTCGCCGTTTATAGGGC	tpke11::GFP
X5_C0465	GCTGAACGACTCGAGTTAATGC	C0465::GFP et pTOPO C0465
E3_C0465	AAGAGAATTCGCTGTTAGATAGC	pTOPO C0465
E3p_C0465	GCATCAGCAATAAGAATTC	C0465::GFP
X5p_IG1105	GCGGCCTCGAGTAATTTGAGTTGTTAAA GCTACGATAAATATTATGTTTTTACGAATT CAGGATG	IG1105::GFP
E3P_IG1105	GATCCTGAATTCGTAAAAACATAATATTTA TCGTAGCTTTAACAACCTCGAAATTA CTCGA GGCCGC	IG1105::GFP
E3_IG1105	CACGATGAATTCATAGTTTCT	pTOPO IG1105
X5_IG1027	AACATGCTCGAGGATATTGCAGC	IG1027::GFP et pTOPO IG1027
E3_IG1027	GTAAGAATTCAGAACACCATTAAC	IG1027::GFP et pTOPO IG1027
X5_dsrA	AGTACTCGAGTTACCAGGATGATTC	dsrA::GFP
E3_dsrA	ACACCAGGAATTCTGATGTGTTTCATCACC	dsrA::GFP
csgB_5IG_X	GACTTCATTCTCGAGGATGAAACCCCGC	csgB::GFP TXN
ceg_E3p_csgB	ATTTGGAATTCATGTTGTCACCCTGGAC	csgB::GFP TXN
csgD_5IG_X	ATTTGTTCTCGAGGTTGTCACCCTGGAC	csgD::GFP (fwd)
Ju_E3_csgD	GACCATGAATTCATGGACTTCATTAAC	csgD IG::GFP (rev)
E3-PcsgD	TTTGAATTCGATTACATCTGATTTCAATC	csgD::GFP TXN (rev)
csgD_TRAD	ATGCGCATGCGGTAACCCCGCTTTTTTTAT TGATCGC	csgD::GFP TRAD UTR+ (rev)

csgD_TRAD_UTR-	ATGCGCATGCGGTAACCCCGCTTTTTTAT TGATCGCGATTACATCTGATTTCAATC	csgD::GFP TRAD dUTR (rev)
ryeB_X_fwd	ATGCCTCGAGTTGATGATCCTAACAATCCC	ryeB::GFP
ryeB_E_prom	ATGCGAATTCATCAGCTGAAATGTTAAAA ACCGC	ryeB::GFP
ibsC_X_fwd	ATGCCTCGAGCTTCTCCACACCGTTGCCGC	ibsC::GFP
ibsC_E_prom	ATGCGAATTCCTTCGGTGAGGGCTTTACCG	ibsC::GFP
XhoI_fwd_ryeA	AAACTCGAGTTGCGGCGCAACTGCTCAAG ACAACC	ryeA::GFP et pTOPO RyeA
EcoRI_prom_ryeA	AAAGAATTCCTGACTTTATAGTCAACCATA ACAC	ryeA::GFP
ryeA_rev	TACCTAAAATAGCCTTTTGATTTCC	pTOPO RyeA
XhoI_fwd_sibC	AAACTCGAGACCTTACAAAAGTGGCAGCA C	sibC::GFP et pTOPO SibC
EcoRI_prom_sibC	AAAGAATTCACCCCTTACTCACTAAATTAGG TC	sibC::GFP
sibC_rev	AGGGCAGCATGGGGCTGTAACG	pTOPO SibC
gcvB_KO_fwd	TGAGCTTCTACCAGCAAATACCTATAGTGG ATTCCGGGGATCCGTCGACC	mutant gcvB::Kn
gcvB_KO_rev	GCTACATTAATCACTATGGACAGACAGGG TTGTAGGCTGGAGCTGCTTCG	mutant gcvB::Kn
mcaS_KO_fwd	GTCAGTGAAGAAAATTGGCAACTAAAGGT TATTCCGGGGATCCGTCGACC	mutant mcaS::Kn
mcaS_KO_rev	CTGTCGACATCCGCCAGACTCTACAGTACA TGTAGGCTGGAGCTGCTTCG	mutant mcaS::Kn
[Btn]-ARNcsgD	[Btn]TACTGACAGATGTTGCACTGCTGTGT GTAG	Northern Blot
[Btn]mcaS	[Btn]CCGCCAGACTCTACAGTACACAC	Northern Blot
[Btn]ryeB	[Btn]CCTTGCCCTTTAAGAATAGATGACG	Northern Blot
[Btn]ibsC	[Btn]GAGTAACAGTACAATCAGTATGATG	Northern Blot
[Btn]omrA	[Btn]TTGGTGCAAGAGACAGGGTACGAAG	Northern Blot
[Btn]sibC	[Btn]CATCATACTGATTGTACTGTTACTC	Northern Blot
[Btn]ryeA	[Btn]AGAGCCATTTCCCTGGACCGAATAC	Northern Blot
[Btn]C5221	[Btn]AAGTGTAGCGAATGCCATCATTGCC	Northern Blot
[Btn]MicC	[Btn]AAATCAGTGGCAATGCAATGGCCCA	Northern Blot
[Btn]RprA	[Btn]GATTTCCATGCTTATAAATCAATAT	Northern Blot
[Btn]DsrA	[Btn]GCACTTAAAAAATTCGTTACACCAGG	Northern Blot
5BIO-5S probe	[Btn]CGGCGCTACGGCGTTTCACTTCTG	Northern Blot
ecs066-f	TAGGGAGTTGGAGTAAAACG	amplicon puce
ecs066-r	AGAACGAGTAAAAGGTCCG	amplicon puce
ryhB-f	TTCAAATGCGAGTCAAATGC	amplicon puce
ryhB-r	GATCAAAAAAAAAAGCCAGCAC	amplicon puce
isrC-f	CTGACTCTGAGTCTGCTCAC	amplicon puce
isrC-r	GTCGTTTCATCAGCTTTTCC	amplicon puce
rdlC-f	CAGGTCGTGCCAGAAAATC	amplicon puce
rdlC-r	CATTGGTTGCTGGAAAAGAG	amplicon puce
C0614-f	GTAATATCTTTTACCGGTTAG	amplicon puce

C0614-r	TTGACTTGTGTTACATGAGC	amplicon puce
C0343-f	TAAGAGTCCACCCAATGAC	amplicon puce
C0343-r	CTCACGATGCGACAATTTAATC	amplicon puce
C0224-f	GAAATCGTACAGTAAAAAGCG	amplicon puce
C0224-r	CGGTGTTTCTCCAGATGTG	amplicon puce
ECS074-f	ATTTGCGTTTTACTGTGTGG	amplicon puce
ECS074-r	TCCCCTACCTCAAACATG	amplicon puce
C0042-f	CTGCGCTGTAAAGATTTTC	amplicon puce
C0042-r	AGGGAGTATCAGTTTTTCATC	amplicon puce
NC017-f	GTGTTTTCATAAAAACCTCTTTG	amplicon puce
NC017-r	GGTGTATGTCACGTATTTTG	amplicon puce
oxyS-f	ACGATAGTTCATGGCGATAG	amplicon puce
oxyS-r	GCAAAAAAAGCGGATCCTG	amplicon puce
IO41-f	GAAATCATGAAGATTCACGG	amplicon puce
IO41-r	TGCGAAAACATTCGAGGAAG	amplicon puce
rybC-r	TGTTGCGCTAAAAAATGGC	amplicon puce
rybC-f	AAGCAAGCATAACGCTGATG	amplicon puce
ryjA-f	AAAAAATCGTCTGAGATGAC	amplicon puce
ryjA-r	TGCTGGAAGATGAACAAAAC	amplicon puce
5S fwd	TTGCCTGGCGGCAGTAGCGC	amplicon puce
rev 5S	CTGGCAGTTCCTACTCTCGC	amplicon puce
arcZ_puce_rev	AAAAAATGACCCCGGCTAGAC	amplicon puce
arcZ_puce_fwd	GTGCGGCCTGAAAAACAGTG	amplicon puce
gcvB_puce_fwd	CAAATACCTATAGTGGCGGC	amplicon puce
gcvB_puce_rev	AAGGTAAAAAAGCACCGC	amplicon puce
5s rev DB	AGATAAAACGAAAGGCCCA	amplicon puce
5S fwd DB	ATTAATCAGAACGCAGAAG	amplicon puce
rybB_puce_fwd	CACTGCTTTTCTTTGATGTC	amplicon puce
rybB_puce_rev	CAAAAAACCCATCAACCTTG	amplicon puce
dsrA_puce_fwd	ACACATCAGATTTCTG	amplicon puce
dsrA_puce_rev	AAAAATCCCGACCCTGAG	amplicon puce
rprA_puce_fwd	GCTCACGGTTATAAATCAAC	amplicon puce
rprA_puce_rev	AAAAAAGCCCATCGTGG	amplicon puce
omrAB_puce_fwd	CCAGAGGTATTGATTGGTGA	amplicon puce
omrAB_puce_rev	CGAAAAAACCCTGCGCATC	amplicon puce
sibC_puce_fwd	TAAGGGTAAGGGAGGATTG	amplicon puce
ryeAB_puce_fwd	TAAAGTCAGCGAAGGAAATGC	amplicon puce
micC_puce_fwd	TTCTGCGCGGGTTATATG	amplicon puce
micC_puce_rev	TCTAAAAAAGCCCGGAC	amplicon puce
omrA DB f	CGTGACAGCCACAATCAAGA	amplicon puce
omrA DB r	GAGCGACAGTAAATTAGGTG	amplicon puce
micC DB f	CTATACGTTATTCTGCGCGG	amplicon puce
micC DB r	CTCTGGATAAGGATTATCCAAT	amplicon puce
rprA DB f	TGATCGACGCAAAAAGTCC	amplicon puce
rprA DB r	GTAGCGAAGCGAAAAATG	amplicon puce
gcvB DB f	GAGCTTCTACCAGCAAATAC	amplicon puce

gcvB DB r	GATCGCAAGGTAACAAAAAGCACC	amplicon puce
ryjA DB f	CGGTGACATACAATTAAGC	amplicon puce
ryjA DB r	GCTGGAAGATGAACAAAAC	amplicon puce
ecs066-f	TAGGGAGTTGGAGTAAAACG	amplicon puce
ecs066-r	AGAACGAGTAAAAGGTCCG	amplicon puce
ryhB-f	TTCAAATGCGAGTCAAATGC	amplicon puce
ryhB-r	GATCAAAAAAAAAAGCCAGCAC	amplicon puce
isrC-f	CTGACTCTGAGTCTGCTCAC	amplicon puce
isrC-r	GTCGTTTCATCAGCTTTTCC	amplicon puce
rdlC-f	CAGGTCGTGCCAGAAAATC	amplicon puce
rdlC-r	CATTGGTTGCTGGAAAGAG	amplicon puce
C0614-f	GTAATATCTTTTACCGTTAG	amplicon puce
C0614-r	TTGACTTGTGTTACATGAGC	amplicon puce
C0343-f	TAAGAGTCCACCCAATGAC	amplicon puce
C0343-r	CTCACGATGCGACAATTTAATC	amplicon puce
C0224-f	GAAATCGTACAGTAAAAGCG	amplicon puce
C0224-r	CGGTGTTTCTCCAGATGTG	amplicon puce
ECS074-f	ATTTGCGTTTTACTGTGTGG	amplicon puce
ECS074-r	TCCACTACCTCAAACATG	amplicon puce
C0042-f	CTGCGCTGTAAAGATTTTC	amplicon puce
C0042-r	AGGGAGTATCAGTTTTTCATC	amplicon puce
NC017-f	GTGTTTTCATAAAAACCTCTTTG	amplicon puce
NC017-r	GGTGTATGTCACGTATTTTG	amplicon puce
oxyS-f	ACGATAGTTCATGGCGATAG	amplicon puce
oxyS-r	GCAAAAAAAAAAGCGGATCCTG	amplicon puce
IO41-f	GAAATCATGAAGATTCACGG	amplicon puce
IO41-r	TGCGAAAACATTCGAGGAAG	amplicon puce
rybC-r	TGTTGCGCTAAAAAATGGC	amplicon puce
rybC-f	AAGCAAGCATAACGCCTGATG	amplicon puce
ryjA-f	AAAAAATCGTCTGAGATGAC	amplicon puce
ryjA-r	TGCTGGAAGATGAACAAAAC	amplicon puce
PUCE_McaS	CTGTGTGTAAGTCTGAGAGTCTGGCGGATGT	oligo puce
PUCE_omrA	GGTACGCTCTTCGTACCCTGTCTCTTGC	oligo puce
PUCE_omrB	GGGTTGAGCACATGAATTACACCAGCCTG	oligo puce
PUCE_rprA	AGTGAAACAACGAATTGCTGTGTGTAGTC	oligo puce
	T	
PUCE_gcvB	ATGGTAGCAAAGCTACCTTTTTTCACTTCC	oligo puce
PUCE_dsrA	GCTTCTTGCTTAAGCAAGTTTCATCCCGAC	oligo puce
PUCE_arcZ	AACTCATCATAATAATTTACGGCGCAGCCA	oligo puce
PUCE_rybB	TTTGATGTCCCCATTTTGTGGAGCCCATC	oligo puce
PUCE_5S	CCCACCTGACCCCATGCCGAACCTCAGAA	oligo puce
PUCE_MicC	GGCCATTGCATTGCCACTGATTTTCCAAC	oligo puce
5S 70mer	GGCGCTACGGCGTTTTCACTTCTGAGTTCGG CATGGGGTCAGGTGGGACCACCGCGCTAC TGCCGCCAGGC	oligo puce
16S-70mer(rrnB)	GCTAGTAGGTGGGGTAACGGCTCACCTAG GCGACGATCCCTAGCTGGTCTGAGAGGAT	oligo puce

GACCAGCCACAC

5S-70mer(rrfB)	CGGTGGTCCCACCTGACCCCATGCCGAACT CAGAAGTGAAACGCCGTAGCGCCGATGGT AGTGTGGGGTC	oligo puce
rpoS-70mer	AGCAGAGCATCGTCAAATGGCTGTTCGAG CTGAACGCCAAACAGCGTGAAGTGCTGGC ACGTCGATTCGG	oligo puce
ryeA-70mer	GCAGGCTTAGTTGCCTTGCCCTTTAAGAAT AGATGACGACGCCAGGTTTTCCAGTTTGC GTGCAAAATGG	oligo puce
dsrA-full	CGCTTTTTATTGACCATTTGCACGCAA GGAAAACCTGGCGTCGTCATCTATTCTTAA AGGGCAAGG	oligo puce
gcvB_70mer	TGCGTGTCTGGTGAACCTTTGGCTTACGG TTGTGATGTTGTGTTGTTGTGTTTGAATT GGTCTGCGAT	oligo puce

Résumé

Les bactéries ont la capacité de s'adapter à leur environnement par une régulation fine de leur expression génique. Cette régulation génique s'opère par les facteurs sigma, sous-unités de l'ARN polymérase permettant la discrimination des promoteurs, et de l'aide éventuelle de cofacteurs pour permettre la transcription de sets de gènes spécifiques aux besoins de la bactérie. Les bactéries contrôlent également l'accumulation de leurs protéines, au niveau post-transcriptionnel, à l'aide notamment d'ARNnc permettant de moduler la traduction d'ARNm ciblés, par l'appariement : ARNm-ARNnc. Nous avons montré, grâce aux données obtenues en RNA-seq, CHIP-seq, Northern Blot et de fusions reportrices, que le facteur sigma RpoS induit ou module la transcription de ces ARNnc. Nous avons notamment observé que le facteur sigma RpoS est réprime (indirectement) la transcription de nombreux ARNnc et induisait celle d'au moins un ARNnc. . Afin d'étudier un exemple de conséquences de la régulation transcriptionnelle d'ARNnc par RpoS , nous avons choisi d'étudier la régulation de l'expression des curli, appendices extracellulaires impliqués dans l'adhérence et la formation de biofilm. Nous avons observé que l'expression des gènes *csg* est régulée à la fois transcriptionnellement et traductionnellement par le facteur sigma RpoS et les ARNnc. Nous avons également pu observer une thermorégulation de l'expression des gènes structuraux des curli qui pourrait impliquer ces deux types de régulateurs.

Mots clés : RpoS, ARNnc, Curli, Expression génique, RNA-seq, CHIP-seq.