



HAL
open science

Circuits mixtes de régulation entre petits ARN régulateurs et systèmes à deux composants chez *Escherichia coli*

Anaïs Brosse

► **To cite this version:**

Anaïs Brosse. Circuits mixtes de régulation entre petits ARN régulateurs et systèmes à deux composants chez *Escherichia coli*. Microbiologie et Parasitologie. Université Sorbonne Paris Cité, 2017. Français. NNT : 2017USPCC245 . tel-02145666

HAL Id: tel-02145666

<https://theses.hal.science/tel-02145666>

Submitted on 3 Jun 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.

Thèse de doctorat
de l'Université Sorbonne Paris Cité
Préparée à l'Université Paris Diderot
Ecole doctorale BioSPC ED562

CNRS UMR8261 Equipe « Contrôle de l'expression génétique par les ARN »

Circuits mixtes de régulation entre petits ARN régulateurs et systèmes à deux composants chez *Escherichia coli*

Par Anaïs Brosse

Thèse de doctorat de Microbiologie

Dirigée par Maude Guillier

Présentée et soutenue publiquement à Paris le 3 octobre 2017

Président du jury : Dussurget, Olivier / Professeur des Universités Paris Diderot / Institut Pasteur – Paris
Rapporteur : Soutourina, Olga / Professeur des Universités Paris Sud / CNRS I2BC – Orsay
Rapporteur : Repoila, Francis / Chargé de recherche / INRA Micalis – Jouy-en-Josas
Examineur : De Reuse, Hilde / Directrice de recherche / Institut Pasteur – Paris
Examineur : Collet, Jean-François / Professeur des Universités / Institut de Duve – Bruxelles
Examineur : Juste, Catherine / Chargée de recherche / INRA Micalis – Jouy-en-Josas
Directeur de thèse : Guillier, Maude / Chargée de recherche / IBPC CNRS UMR8261 – Paris
Membre invité : Espéli, Olivier / Directeur de recherche / CIRB Collège de France – Paris

Titre : Circuits mixtes de régulation entre petits ARN régulateurs et systèmes à deux composants chez *Escherichia coli*

Résumé :

Les petits ARN régulateurs et les systèmes à deux composants sont des régulateurs très répandus de l'expression des gènes chez les bactéries. Dans la plupart des cas, les systèmes à deux composants agissent comme des régulateurs transcriptionnels. Un grand nombre de petits ARN agissent quant à eux au niveau post-transcriptionnel en modulant la traduction et/ou la stabilité de leur(s) ARN messenger(s)-cible(s). Des connexions entre ces deux systèmes ont récemment pu mettre en lumière des circuits de régulations complexes aux propriétés encore peu connues.

Mon travail a tout d'abord porté sur la connexion entre le système à deux composants EnvZ-OmpR et les petits ARN OmrA et OmrB chez *Escherichia coli*. Dans un premier temps, nous avons montré qu'OmpR activait directement la transcription d'*omrA* et d'*omrB* en se fixant à leur promoteur. Cette activation permet la production des petits ARN OmrA et OmrB qui, via leur extrémité 5' conservée, ciblent à leur tour plusieurs ARN messagers-cibles et notamment le messenger *ompR-envZ*. En accord avec des études précédentes, le contrôle d'*ompR-envZ* par les Omr n'affecte pas le niveau de forme phosphorylée d'OmpR. Ce phénomène posait donc la question de l'intérêt d'une telle régulation. Nous avons ensuite pu montrer que la régulation d'*omrA* et d'*omrB* est assez unique car leurs promoteurs répondent non seulement à la forme phosphorylée mais aussi à la forme non phosphorylée d'OmpR. Ce phénomène permet à ces ARN régulateurs de limiter leur propre synthèse en ayant un effet limité sur l'expression des autres cibles d'OmpR comme les porines OmpC et OmpF.

Ce travail nous a conduits à chercher à caractériser d'autres exemples de modulation des systèmes à deux composants par des petits ARN régulateurs. Nous avons notamment étudié la régulation du système NarQ-NarP. En effet, nos travaux ont montré que la synthèse de *narP* était contrôlée par le petit ARN RprA. Cette régulation semble affecter les cibles de NarP et en particulier l'opéron *napFDAGHBC*. De plus, RprA répondrait au même stimulus que le système NarQ-NarP créant ainsi un lien physiologique entre le petit ARN et sa cible.

Pour finir, un autre aspect de ce travail de thèse a été de s'intéresser à la régulation d'OmrA/B dans un contexte d'infection des macrophages par une souche d'*Escherichia coli* pathogène, la souche LF82. En effet, des données suggéraient que ces petits ARN étaient induits au cours de l'infection. J'ai pu valider ces données et montrer que cette induction était dépendante de la présence du système EnvZ/OmpR.

En conclusion, j'ai pu montrer par diverses approches que les circuits de régulation intégrant des systèmes à deux composants et des petits ARN régulateurs possédaient des propriétés assez inédites permettant à la bactérie de s'adapter à divers stress.

Mots clefs : ARN régulateurs, Hfq, OmrA/B, RprA, Systèmes à deux composants, EnvZ-OmpR, NarQ-NarP

Title : Mixed regulatory circuits between small RNAs and two-component systems in *Escherichia coli*

Abstract :

Small regulatory RNA (sRNAs) and two component systems (TCS) are both widespread regulators of gene expression in bacteria. While TCS are mostly transcriptional regulators, a large class of sRNAs acts as post-transcriptional regulators of gene expression by modulating translation and/or stability of target-mRNAs. Many connections have recently been unraveled between these two types of regulators, resulting in mixed regulatory circuits with poorly characterized properties.

First, we have investigated in details the negative feedback circuit that exists between the EnvZ-OmpR TCS and the OmrA/B sRNAs in *Escherichia coli*. We have found that OmpR directly activates transcription from *omrA* and *omrB* promoters, allowing production of OmrA/B sRNAs that target multiple mRNAs through their conserved 5' end, including the *ompR-envZ* mRNA. In agreement with previous reports, we have found that this control of *ompR-envZ* by OmrA/B sRNAs does not affect the amount of OmpR-P i.e. the presumably active form of the regulator. This phenomenon therefore raised the question of the possible interest of such a regulation. Thereafter, we found that OmrA/B regulation is really unique because they respond to the phosphorylated form but also to the unphosphorylated form of OmpR. As a result, OmrA/B limit their own synthesis while they have only a limited effect on others targets of OmpR, such as the OmpC or OmpF porins.

This work led us to try to characterize other examples of modulation of two-component systems synthesis by small regulatory RNAs. In particular, we studied the regulation of the NarQ-NarP system. Indeed, our work showed that the synthesis of *narP* is controlled by the RprA sRNA. This regulation appears to affect NarP targets and in particular the *napFDAGHBC* operon. Moreover, RprA would respond to the same stimulus as the NarQ-NarP system, thus creating a physiological link between the small RNA and its target.

Finally, another aspect of this work was to study the regulation of OmrA/B in a context of infection of macrophages by an *Escherichia coli* pathogenic strain, LF82. Indeed, data suggested that these small RNAs were induced during infection. I was able to validate these data and showed that this induction was dependent on the presence of the EnvZ-OmpR system.

In conclusion, I have been able to show by various approaches that the control circuits integrating two-component systems and small regulatory RNAs have quite unique properties allowing the bacteria to adapt to various stresses.

Keywords : Regulatory RNAs, Hfq, OmrA/B, RprA, Two component systems, EnvZ-OmpR, NarQ-NarP

*« The process of scientific discovery is, in effect, a
continual flight from wonder. »*

Albert Einstein

Remerciements

Le travail présenté dans ce mémoire a été réalisé au sein du laboratoire d'expression génétique microbienne au sein de l'Institut de Biologie Physico-Chimique sous la direction de Maude Guillier.

Je tiens avant tout à remercier Maude pour la confiance dont elle me témoigne depuis plusieurs années en me confiant notamment ce projet de thèse. Merci pour ces nombreuses discussions scientifiques passionnantes qui m'ont permis d'apprendre beaucoup sur le plan théorique et expérimental. Tu as su canaliser mon énergie débordante et la mettre au service des projets de recherche que nous avons partagés. Ton esprit vif et ta gentillesse m'ont permis de m'épanouir intellectuellement dans un cadre de travail agréable et pour cela, je tiens à te dire merci. Merci également pour toutes ces belles opportunités que tu m'as donné. Ton soutien sans faille dans mes choix professionnels (cours du soir, concours, enseignement) me permet d'avancer avec beaucoup de sérénité. L'alchimie particulière qui se crée entre un doctorant et son directeur m'a permis d'une certaine façon de mieux te connaître et de découvrir des facettes inattendues de ta personnalité : je saurai maintenant que rien ne vaut les Soprano et que tu adores les polars et les Foo Fighters !

Je remercie sincèrement Olga Soutourina et Francis Repoila d'avoir gentiment accepté d'être rapporteur de ma thèse. Merci aussi à Hilde de Reuse, à Catherine Juste et à Jean-Francois Collet d'examiner mon travail, mais aussi à Olivier Dussurget d'avoir accepté de présider mon jury. Une mention particulière à Catherine qui m'a donné l'envie d'entrer dans le monde passionnant de la recherche en microbiologie il y a maintenant 10 ans. Je pense que ta curiosité scientifique et ta bonne humeur permanente y sont pour beaucoup. Je remercie également Olivier Espéli de participer à mon jury de thèse. Merci pour tous tes conseils et pour m'avoir donné l'opportunité d'apprendre beaucoup sur les relations hôte-pathogène en m'accueillant au sein de ton laboratoire.

J'adresse également tous mes remerciements à tous les membres de mon groupe de recherche présents et passés ainsi qu'aux membres de l'unité de recherche « UMR8261/FRE3630 /UPR9073 » ou d'autres unités de l'IBPC et plus spécialement à Jackie, Joël, Audrey C, Marie C, Mark, Maxence, Claude, Jonathan, Marie C, Agnès, Kyle, Josette (merci pour les cours de tricot !), Hilal, Bruna, Caroline L, Laetitia G, Assia, Marjorie, Carine, Frédérique, Ciaran, Mickael, Lionel, Claire T, Julien H, Christophe, Irène H, Isabelle K qui ont fait et font de cet institut un endroit très

sympathique pour travailler. Je remercie également Marc Dreyfus pour son soutien quand il était directeur de notre unité et plus généralement pour ses gentilles paroles. Merci aussi à toi Eliane pour tes conseils scientifiques et non-scientifiques qui m'ont permis de mener ce projet dans la bonne humeur au sein de ton groupe de recherche.

Bien évidemment je remercie Mathias Springer. Merci de m'avoir recrutée ce jour de septembre 2008, de m'avoir soutenue dans mes projets professionnels, merci pour tes nombreuses discussions scientifiques et bien sûr, je te remercie pour m'avoir conseillée dans l'écriture de ce manuscrit. Ta rigueur scientifique et tes innombrables questions (notamment en séminaire) ont fait de moi quelqu'un de plus exigeant et me permettent de toujours repousser mes limites.

Je remercie également les membres du groupe d'Olivier Espéli avec qui j'ai le plaisir de collaborer ou que je croise au Collège de France : Gaëlle, Sylvie, Elise, Victoria, Charlène, Hafez.

Merci également à vous Antonia et Olivier de me faire confiance depuis trois ans en me confiant les TD au CNAM. Cette expérience d'enseignement m'a permis de m'ouvrir à d'autres domaines scientifiques et ainsi prendre du recul sur mon travail de thèse.

Un grand merci à tous mes amis du labo qui m'ont soutenu durant toutes ces années d'école d'ingénieur CNAM et de thèse : Céline G d'avoir été une partenaire de labo au top, mais aussi une amie avec qui j'apprécie toujours autant de partager des moments en dehors du labo ; Alex pour nos nombreuses pause-café à discuter de tout et de rien, d'être capable de m'écouter parler pendant des heures, de m'avoir donné plein de conseils de bioinfo et bien sûr d'avoir donné ton avis franc et sincère sur ce manuscrit ; Céline A pour nos nombreux trajets en RER B parfois mythiques en temps de grève, tes conseils tricot ou les petits ciné du vendredi ; Nathalie L pour nos cours partagés au CNAM et ensuite nos cours de yoga (merci au passage à Anne et Marie-Hélène pour ces cours merveilleux qui m'ont accompagné dans cette dernière année de thèse !), nos pauses thé du midi à l'IBPC, tes conseils sur Pymol, la vie de thésard ou tout autre sujet scientifique ou non sur lesquels nous pourrions discuter pendant des heures ! Je pense avoir trouvé une alter-égo en terme de papotage ; Sabine pour ta gentillesse et ton humour : même à 15 000 km tu es un petit rayon de soleil ; Fanette pour nos discussions scientifiques, culinaires, sportives et j'en passe, pour m'avoir donné l'opportunité de goûter le meilleur tiramisu de ma vie (à ce jour) ; Emmeline d'être toujours là pour rigoler et envoyer des liens internet surprenants ; Sylvain R pour tes inventions rigolotes quand tu étais à l'IBPC et pour ces vacances en Ecosse mémorables ; Flore pour ton

humour décapant et tous ces apéros partagés ; Sylvain D pour les soirées Just-Dance et pizza chez toi ; Erin ma petite américaine pour ta gentillesse et pour m'aider à améliorer mon anglais ; Nguyen ma copine de labo depuis 2016 toujours dispo pour parler fashion et voyages. Vos encouragements, les pauses café/thé/goûter /glaces, les pique-nique au jardin du Luxembourg, les apéros et soirées que l'on a partagés, m'ont permis de mener à bien ce grand projet.

Je tiens à remercier aussi tous mes amis qui pour certains me soutiennent et supportent depuis longtemps : Amandine H-M, Jessica C, Claire T, Magalie B ; vos petits messages de soutien me redonnent toujours autant la pêche ; Clémence ma partenaire de rilage depuis plus de 15 ans (et oui ça ne nous rajeunit pas !) merci pour cette soupape de décompression ; Judite et William mille mercis pour nous avoir soutenus et supportés durant ce long processus de thèse (merci pour cette journée à Disney plus que bienvenue en fin de rédaction !) ; Guillaume l'H nos rendez-vous au Green Linet avec une bonne Guinness ont été d'une grande aide ; Damien , Maud , Claire et Julien merci pour toutes ces soirées à philosopher sur de nombreux sujets que ce soit autour d'une bonne raclette, d'un bon verre de vin, d'un petit barbecue ou à déambuler dans les rues de Paris ces petits bols d'air m'ont permis de décompresser et de me rebooster dans les moments de creux ; Audrey, Jérémy, Jessica, Medhi, Antoine, Coralie, Nicolas, Seb, Florie, Sylvain D merci pour tous ces moments partagés et pour nous avoir fait sortir du labo les week-ends !

Et enfin Isa, ma poule, ma biche, ma #bff, merci pour tout ; c'est un bonheur que d'être à tes côtés. Je me souviens encore de notre rencontre dans cette salle d'attente de concours CNRS. On a échangé quelques mots et on ne s'est plus quittées. Merci d'être aussi spontanée, drôle, franche, politiquement incorrecte, de me soutenir sans aucune réserve dans mes projets les plus fous.

Enfin, j'adresse mon affection à toute ma famille, et en particulier mes parents et mes grands-parents, leur tendresse, leurs conseils et leur confiance me portent et me guident chaque jour. Papa, Maman, Florine, Juju, Hugo, Pépé, Mémé, Nathalie, Eric, Caroline, Franck, Catherine, Manu, Julie, Justine, Théau, j'ai de la chance de vous avoir. Je remercie également Denise et Jean-Marie pour ces instants ressourçants sous le soleil de Verdun-sur-Garonne et de m'accueillir avec autant de gentillesse.

Enfin difficile de trouver les mots pour te remercier Julien mon partenaire dans la vie et dans la science. Sans toi cette thèse n'aurait pas eu la même saveur. Tu me pousses à toujours me dépasser et me fais grandir chaque jour. Merci d'être toi.

Abréviations

3'UTR : région 3' non traduite

5'UTR : région 5' non traduite

AA : acide aminé

ADN : acide désoxyribonucléique

AIEC : *Escherichia coli* adhérentes et invasives

AIP : Auto-Inducing peptide (peptide auto-inducteur)

AMPc : adénosine monophosphate cyclique

ARN : acide ribonucléique

ARNm : ARN messenger

ARNnc : ARN non codant

ARNr : ARN ribosomique

ARNt : ARN de transfert

Asp : Aspartate

ATP : adénosine triphosphate

B. subtilis : *Bacillus subtilis*

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

Csr : Carbon storage regulator

DAEC : *Escherichia coli* à adhésion diffuse

DEPC : diéthyl pyrocarbonate

DMSO : diméthyl-sulphoxide

EAggEC : *Escherichia coli* entéroaggrégatives

E. coli : *Escherichia coli*

EHEC : *Escherichia coli* entérohémorragiques

EIEC : *Escherichia coli* entéroinvasives

EnvZ-P : forme phosphorylée d'EnvZ

ETEC : *Escherichia coli* entérotoxigéniques

ExPEC : *Escherichia coli* pathogènes extraintestinales

GDP : guanosine diphosphate

GlcN6P : glucosamine-6-phosphate
GTP : guanosine triphosphate
IHF : Integration Host Factor
IPTG : isopropyl- β -D-thiogalactopyranoside
His : Histidine
HK : histidine kinase
H-NS : Histone-like nucleoid structuring protein
HPT : histidine phosphotransférase
LB : Luria Bertani
Lrp : Leucine-responsive regulatory protein
Met : méthionine
NarL-P : forme phosphorylée de NarL
NarP-P : forme phosphorylée de NarP
NTP : nucléotide triphosphate
Nts : nucléotides
OmpR-P : forme phosphorylée d'OmpR
ONPG : Ortho-nitrophényl- β -galactoside
ORF : Open reading frame (phase ouverte de lecture)
Pi : Phosphate inorganique
PNPase : PolyNucléotide Phosphorylase
RBS : Ribosome binding site (site de fixation du ribosome)
RNase : ribonucléase
RR : régulateur de réponse
S. aureus : *Staphylococcus aureus*
S. enterica : *Salmonella enterica*
SD : Shine Dalgarno
UPEC : *Escherichia coli* uropathogènes

Table des matières

1. Introduction	1
1.1. L'expression génétique et son contrôle chez les bactéries	1
1.1.1. La transcription bactérienne	2
1.1.1.1. Les différentes étapes de la transcription et ses principaux acteurs	2
1.1.1.1.1. Le promoteur	2
1.1.1.1.2. L'ARN polymérase	3
1.1.1.1.3. Les étapes de la transcription	4
1.1.1.2. Régulation de la transcription	6
1.1.1.2.1. Régulation du démarrage par des facteurs globaux	7
1.1.1.2.2. Régulation du démarrage par des facteurs spécifiques	9
1.1.1.3. Régulation de l'allongement et de la terminaison de la transcription	10
1.1.2. La traduction bactérienne	12
1.1.2.1. Les différentes étapes de la traduction et ses principaux acteurs	12
1.1.2.1.1. Le ribosome	12
1.1.2.1.2. L'ARN messager	14
1.1.2.1.3. Les ARN de transfert	15
1.1.2.1.4. Les différentes étapes de la traduction	15
1.1.2.2. Régulation du démarrage de la traduction	17
1.1.3. Stabilité des ARN messagers	20
1.1.3.1. Généralités sur les ribonucléases	21
1.1.3.2. La RNase E	23
1.1.3.3. La dégradation, un mécanisme de contrôle à part entière	26
1.2. Les systèmes à deux composants	27
1.2.1. Définitions – distribution génomique	27
1.2.2. Mode de transduction du signal	31
1.2.3. Communication entre les systèmes à deux composants	33
1.2.4. Pléiotropie des systèmes à deux composants et diversité des signaux	37
1.2.5. Le système EnvZ-OmpR : un système classique avec des propriétés inédites	39
1.2.5.1. La protéine kinase senseur : EnvZ	40
1.2.5.1.1. Caractéristiques d'EnvZ	40
1.2.5.1.2. Structure cristallographique d'EnvZ	43
1.2.5.1.3. Signal capté par EnvZ	44
1.2.5.1.4. Mécanisme d'activation d'EnvZ	45

1.2.5.2.	La protéine régulatrice : OmpR	47
1.2.5.2.1.	Caractéristiques d'OmpR.....	47
1.2.5.2.2.	Structure cristallographique d'OmpR.....	48
1.2.5.2.3.	Activation non classique d'OmpR.....	50
1.2.5.3.	Modèle de régulation d'OmpC et OmpF par le système EnvZ-OmpR	50
1.2.5.4.	Autres cibles d'EnvZ-OmpR.....	52
1.2.5.5.	Séquence consensus de fixation d'OmpR.....	55
1.2.5.6.	Robustesse du système EnvZ-OmpR.....	56
1.2.6.	Les systèmes NarQ-NarP et NarX-NarL	60
1.2.6.1.	Généralités sur la respiration chez <i>Escherichia coli</i>	60
1.2.6.2.	Régulation de la transition aérobie-anaérobie.....	63
1.2.6.3.	Les kinases senseurs NarX et NarQ.....	64
1.2.6.3.1.	Caractéristiques communes.....	64
1.2.6.3.2.	Perception du signal.....	65
1.2.6.3.3.	Affinité différente pour les régulateurs.....	67
1.2.6.4.	Les régulateurs NarL et NarP.....	68
1.3.	Les ARN régulateurs bactériens	74
1.3.1.	Classification des ARN régulateurs bactériens	75
1.3.1.1.	Les riboswitches.....	76
1.3.1.2.	Les ARN modulant l'activité d'une protéine.....	77
1.3.1.3.	ARN antisens.....	79
1.3.1.4.	ARN agissant par complémentarité imparfaite de séquence.....	80
1.3.1.4.1.	Effet sur la traduction.....	81
1.3.1.4.2.	Effet sur la dégradation.....	82
1.3.2.	Focus sur les ARN Hfq-dépendants	84
1.3.2.1.	La protéine Hfq.....	85
1.3.2.2.	Rôles physiologiques des ARN Hfq-dépendants.....	89
1.3.2.3.	Circuits mixtes de régulation.....	91
1.3.3.	Les petits ARN OmrA et OmrB	95
1.3.3.1.	Découverte des Omr.....	95
1.3.3.2.	Régulation des Omr par EnvZ-OmpR.....	97
1.3.3.3.	Cibles des Omr.....	98
1.3.4.	RprA	102
1.3.4.1.	Les cibles de RprA.....	102
1.3.4.2.	Régulateurs de RprA.....	104

1.4.	La bactérie <i>Escherichia coli</i>	108
1.4.1.	Généralités	108
1.4.2.	Les différentes espèces d' <i>Escherichia coli</i>	109
1.4.3.	<i>Escherichia coli</i> et la flore intestinale : entre commensalisme et pathogénie	110
1.4.3.1.	Le microbiote intestinal	110
1.4.3.2.	La maladie de Crohn.....	111
1.4.4.	La souche d' <i>Escherichia coli</i> LF82	114
1.5.	Objectifs de la thèse.....	121
2.	Matériels et méthodes	123
2.1.	Construction des souches d' <i>E. coli</i> non pathogènes	123
2.2.	Conditions de culture des souches d' <i>E. coli</i> non pathogènes.....	127
2.3.	Souches de LF82 et infection des macrophages	128
2.3.1.	Construction des souches.....	128
2.3.2.	Infection des macrophages.....	128
2.3.3.	Extraction des ARN au Trizol®.....	129
2.4.	Northern blot	130
2.5.	Western blot « Phostag »	130
3.	Résultats	133
3.1.	Circuit mixte de régulation entre les petits ARN OmrA et OmrB et le système à deux composants EnvZ-OmpR chez <i>Escherichia coli</i>	133
3.1.1.	Résultats publiés	133
3.1.2.	Autres régulateurs des Omr	166
3.1.2.1.	Partenaires d'OmpR	166
3.1.2.2.	Régulateurs indépendants d'OmpR.....	170
3.2.	La régulation des Omr par OmpR chez la bactérie pathogène <i>Escherichia coli</i> LF82	174
3.2.1.	Analyse des données de RNAseq	174
3.2.2.	Contrôle des Omr par OmpR chez LF82 au cours de l'infection	176
3.3.	NarQ-NarP, un autre système régulé par des petits ARN Hfq-dépendants	179
3.3.1.	Identification des petits ARN régulant l'expression de <i>narP</i>	179
3.3.2.	Le messenger <i>narP</i> , une cible directe de RprA	185
3.3.3.	Effet de la modulation du niveau de <i>narP</i> sur les cibles du système NarQ-NarP.....	189
3.3.4.	Réponse de RprA aux nitrates.....	193
3.3.5.	Dépendance à RprA des autres petits ARN régulant <i>narP</i>	194
4.	Conclusion et perspectives	197
4.1.	Réseau de régulation entre les Omr et EnvZ-OmpR	198

4.1.1.	Quel mécanisme pour expliquer la régulation des Omr par OmpR ?.....	198
4.1.2.	Quels sont les autres régulateurs des Omr ?.....	199
4.1.3.	Y a-t-il d'autres cibles d'OmpR non robustes et quel en est l'intérêt ?.....	201
4.1.4.	Les Omr ont-ils un rôle direct dans la virulence de LF82 ?	202
4.2.	Réseau de régulation entre NarP et RprA	203
4.2.1.	Quel est le mécanisme de contrôle de <i>narP</i> par RprA ?.....	203
4.2.2.	Comment la régulation de NarP par RprA se met-elle en place en réponse aux nitrates ?.....	204
4.2.3.	Comment fonctionne la régulation de NarP par les autres petits ARN ?.....	205
4.3.	Les circuits mixtes de régulation.....	206
4.3.1.	Quels déterminants pour expliquer la robustesse de certains systèmes à deux composants ?.....	206
4.3.2.	Quelles autres propriétés pour les circuits mixtes ?.....	207
5.	Annexes	211
	<i>Annexe 1 : Revue sur les méthodes utilisées pour l'étude des interactions entre ARN régulateurs et messagers.</i>	211
	<i>Annexe 2 : Composition détaillée des gels SDS-PAGE « Phos-tag »</i>	222
6.	Bibliographie.....	223

Table des illustrations

Figure 1 : Organisation d'une unité transcriptionnelle bactérienne	3
Figure 2 : La transcription chez les bactéries	4
Figure 3 : Mécanisme de terminaison intrinsèque.....	5
Figure 4 : Mécanisme de la terminaison Rho dépendante.....	6
Figure 5 : Vue d'ensemble du réseau de régulation transcriptionnelle chez <i>Escherichia coli</i>	7
Figure 6 : Schéma de la régulation de l'opéron lactose.....	10
Figure 7 : Régulation de l'expression de l'opéron tryptophane par atténuation de la transcription.....	11
Figure 8 : Structure en cryo-microscopie électronique du ribosome d' <i>E. coli</i> associé au facteur EF-Tu.....	13
Figure 9 : Structure d'un ARN de transfert.....	15
Figure 10 : Les principales étapes de la traduction procaryote	17
Figure 11 : Mécanismes de régulation de la traduction par des facteurs protéiques	18
Figure 12 : Autocontrôle traductionnel de la protéine ribosomique L20	19
Figure 13 : Voies de dégradation des ARN chez <i>E. coli</i>	22
Figure 14 : Schéma de l'organisation en domaines de la RNase E.....	24
Figure 15 : Mécanismes de clivage par la RNase E.....	25
Figure 16 : Schéma présentant les systèmes à deux composants classiques	27
Figure 17 : Diversité des systèmes à deux composants en fonction de la taille de différents génomes bactériens.....	29
Figure 18 : Cinétique de phospho-transfert de la kinase EnvZ et des différents régulateurs de réponse de la bactérie <i>E. coli</i>	34
Figure 19 : Comparaison entre cross-régulation ou cross-talk et systèmes branchés.....	35
Figure 20 : Représentation schématique des trois mécanismes de perception du signal par les histidines kinase.....	39
Figure 21 : Comparaison des modèles de régulation du système EnvZ-OmpR	42
Figure 22 : Structure primaire d'EnvZ et représentation des deux domaines d'EnvZ	43
Figure 23 : Modèle d'organisation des dimères d'EnvZ permettant l'autophosphorylation de la kinase ...	44
Figure 24 : Structure du domaine N-terminal d'OmpR modélisée à partir de celle de CheY.....	49
Figure 25 : Représentation du domaine de fixation à l'ADN d'OmpR.....	49
Figure 26 : Modèle de régulation d'OmpC et OmpF par OmpR-P.....	51
Figure 27 : Premier modèle d'osmorégulation d' <i>ompC</i> et d' <i>ompF</i> par OmpR.....	51
Figure 28 : Second modèle d'osmorégulation d' <i>ompC</i> et d' <i>ompF</i> par OmpR.....	52
Figure 29 : Cibles du système EnvZ-OmpR chez <i>E. coli</i> K12.....	55
Figure 30 : Comparaison des sites de fixations d'OmpR présents sur le promoteur d' <i>ompC</i> et d' <i>ompF</i> ..	56
Figure 31 : Description cinétique du système à deux composants EnvZ-OmpR	57
Figure 32 : Effet de la modulation du niveau d'OmpR sur la transcription d' <i>ompC</i> et d' <i>ompF</i>	59
Figure 33 : Vue d'ensemble de la chaîne respiratoire d' <i>E. coli</i>	60
Figure 34 : Organisation structurale des différents types de nitrate réductases bactériennes.....	62
Figure 35 : Résumé des différentes nitrate et nitrite réductases retrouvées chez <i>E. coli</i>	62
Figure 36 : Représentation schématique des protéines NarX et NarQ et de leurs domaines conservés.....	65
Figure 37 : Structure du domaine senseur de NarX à 1.7Å.....	66
Figure 38 : Modèle de régulation croisée asymétrique pour les systèmes à deux composants NarX-NarL et NarQ-NarP	67
Figure 39 : Positions des heptamères de NarL sur différents promoteurs	70
Figure 40 : Mécanisme d'activation du promoteur P_{nir}	71
Figure 41 : Agencement des régions de contrôle de l'opéron <i>napF</i>	73

Figure 42 : Schéma des deux catégories d'ARN régulateurs agissant par appariement de séquences.....	75
Figure 43 : Schéma d'un riboswitch et exemple de régulations associées	77
Figure 44 : Mode d'action des petits ARN CsrB et CsrC.	78
Figure 45 : Mécanisme d'action de GadY.....	80
Figure 46 : Mécanismes de régulation des ARN agissant par complémentarité imparfaite de séquence ...	81
Figure 47 : Modèle de coopération entre petit ARN régulateur, Hfq et la RNase E pour la dégradation d'un messager.....	83
Figure 48 : Structure homohexamérique de la protéine Hfq.	86
Figure 49 : Mode de fixation des ARN régulateurs sur la protéine Hfq.	87
Figure 50 : Cycle d'activité des ARN Hfq-dépendants.....	89
Figure 51 : Différents motifs retrouvés dans les circuits de régulation.....	92
Figure 52 : Simulation de vitesse de réponse au stress	93
Figure 53 : Modèle de régulation basé sur des ARN.....	94
Figure 54 : Organisation et conservation de la région chromosomique <i>omrA-omrB</i>	96
Figure 55 : Le système à deux composants EnvZ-OmpR active la transcription d'OmrA et OmrB	98
Figure 56 : L'extrémité 5' d'OmrA/B est impliquée dans la régulation de multiples cibles	99
Figure 57 : Appariement des Omr avec le 5'UTR des messagers <i>csgD</i> et <i>ydaM</i>	101
Figure 58 : Régulon des Omr.	102
Figure 59 : Sites d'interaction entre RpoS et les ARN régulateurs DsrA et RprA.....	102
Figure 60 : Sites d'interaction entre RprA et <i>csgD</i>	103
Figure 61 : Le phosphorelais Rcs d' <i>E. coli</i>	105
Figure 62 : Mode d'activation du système Rcs : le rôle majeur de RcsF et d'IgaA.....	107
Figure 63 : Image de microscope électronique à balayage d' <i>Escherichia coli</i> K12	108
Figure 64 : Représentation schématique de l'arbre phylogénétique des phyla majoritaires retrouvés dans le microbiote intestinal humain	111
Figure 65 : Cliché de microscopie électronique à transmission de cellules épithéliales Hep-2 infectées par la souche LF82	114
Figure 66 : Modèle de colonisation des cellules épithéliales par une bactérie de type AIEC.....	117
Figure 67 : Effet de la délétion d' <i>ompR</i> sur les capacités d'adhésion et d'invasion de LF82.....	118
Figure 68 : Rôle potentiel du système à deux composants EnvZ/OmpR et d'OmpC dans la virulence de LF82..	119
Figure 69 : Test de motilité sur agar mou des souches LF82 ou MG1655 (<i>E. coli</i> K12).....	120
Figure 70 : Activation des <i>omrA/B</i> et d' <i>ompC</i> par des mélanges OmpR/OmpR-P.....	167
Figure 71 : Différents activateurs en multcopie des Omr agissant via EnvZ-OmpR.....	168
Figure 72 : Activation des Omr par OmpR dans un contexte Δ <i>fadR</i> ou Δ <i>gcvA-ygdD</i>	169
Figure 73 : Expression des Omr en milieu mimimum ou en LB en l'absence d'OmpR.....	170
Figure 74 : Effets du plasmide pRstAB sur le promoteur d' <i>omrA</i> ou d' <i>omrB</i> dans un contexte sauvage ou mutant pour <i>ompR</i>	172
Figure 75 : Effets du plasmide pOmpR sur le promoteur d' <i>omrA</i> ou d' <i>omrB</i> dans un contexte sauvage ou mutant pour <i>rstAB</i>	173
Figure 76 : Profil d'expression des petits ARN régulateurs de LF82 au cours de l'infection.....	175
Figure 77 : Profil d'expression des messagers les plus exprimés dans le séquençage et de certains gènes issus du régulon OmpR chez la bactérie LF82 au cours de l'infection	176
Figure 78 : Induction des Omr au cours de l'infection chez LF82.....	177
Figure 79 : Expression d'OmrA et d'OmrB en phase exponentielle ou en phase stationnaire chez LF82 ou K12.....	178
Figure 80 : Construction des différentes fusions traductionnelles	182

Figure 81 : Expression des différentes fusions des régulateurs des systèmes à deux composants dans un contexte sauvage ou muté pour Hfq.....	183
Figure 82 : Activité spécifique de la souche AB1004 portant la fusion P_{tet} - <i>narP-lacZ</i> et transformée avec la banque de petits ARN Hfq-dépendants.....	184
Figure 83 : Mutations compensatoires dans RprA et dans la fusion P_{tet} - <i>narP-lacZ</i>	185
Figure 84 : <i>narP</i> est une cible directe de RprA.....	186
Figure 85 : Niveaux de surproduction de RprA et RprAmut2.....	187
Figure 86 : Les 105 premiers nucléotides de la phase codante de <i>narP</i> sont requis pour la régulation par RprA.....	188
Figure 87 : Contrôle des opérons <i>napFDAGHBC</i> et <i>nirBDC</i> par les régulateurs NarL et NarP.....	189
Figure 88 : L'opéron <i>napFDAGHBC</i> est bien spécifiquement induit par NarP en micro-anaérobiose et en présence de nitrates.....	191
Figure 89 : Le contrôle de <i>narP</i> par RprA diminue le niveau de l'opéron <i>napFDAGHBC</i>	192
Figure 90 : RprA est activé par la présence de nitrates indépendamment de NarP.....	193
Figure 91 : L'activation de RprA par les nitrates passe par le système Rcs.....	194
Figure 92 : L'effet d'OmrA, MicA, Spot42 et DicF est indépendant de RprA.....	195
Figure 93 : L'effet d'OmrA, MicA, Spot42 et DicF est indépendant de SdsN.....	195
Figure 94 : Résumé du travail sur le système EnvZ-OmpR.....	197
Figure 95 : Résumé du travail sur le système NarQ-NarP et son lien avec le petit ARN RprA.....	198
Figure 94 : Un site de fixation d'ArcA pourrait exister en amont du promoteur d' <i>omrA</i>	200
Figure 95 : Régulation directe et indirecte de <i>narP</i> par RprA.....	205
Figure 96 : Taux de synthèse des molécules appartenant aux systèmes à deux composants d' <i>E. coli</i>	207

Index des tableaux :

Tableau 1 : Les facteurs sigma chez <i>Escherichia coli</i>	3
Tableau 2 : Les ribonucléases d' <i>Escherichia coli</i>	23
Tableau 3 : Nombre de systèmes à deux composants répertoriés chez différents organismes modèles.....	28
Tableau 4 : Systèmes à deux composants chez <i>E. coli</i>	30
Tableau 5 : Cibles de NarL et/ou de NarP validées expérimentalement.....	69
Tableau 6 : Exemple de cibles directes de CsrA chez <i>E. coli</i>	78
Tableau 7 : ARN Hfq-dépendants présents chez <i>E. coli</i> utilisés au cours de cette étude.....	90
Tableau 8 : Souches et plasmides utilisés au cours du projet « Autres régulateurs des Omr ».....	124
Tableau 9 : Souches et plasmides utilisés au cours du projet « Le système à deux composants NarQ-NarP ».....	125
Tableau 10 : Oligonucléotides utilisés dans le projet « Le système à deux composants NarQ-NarP ».....	126
Tableau 11 : Composition du milieu MOPS pour les cultures en anaérobiose.....	127
Tableau 12 : Sondes biotinylées pour les Northern.....	130
Tableau 13 : Résultats du crible génétique.....	171
Tableau 14 : Systèmes à deux composants choisis pour cette étude.....	180

1. Introduction

1.1. L'expression génétique et son contrôle chez les bactéries

Dès le début du XXème siècle, avec l'étude des « enzymes adaptatives » telles que la β -galactosidase, il est apparu que les bactéries étaient capables de s'adapter à leur environnement en synthétisant préférentiellement les éléments qui leur étaient nécessaires. Depuis les années 1940, de nombreux travaux réalisés notamment chez l'organisme modèle *Escherichia coli* (*E. coli*) ont pu mettre en évidence de nombreux circuits de régulation qui dirigent, en fonction de stimuli externes ou internes, l'expression de certains gènes et la synthèse de protéines spécifiques. Nous savons aujourd'hui que la régulation peut s'exercer à toutes les étapes de l'expression d'un gène, aussi bien au niveau transcriptionnel que post-transcriptionnel notamment au niveau de l'étape de traduction ou encore à un niveau post-traductionnel. Cette régulation doit être d'autant plus fine et réactive chez les bactéries que celles-ci sont soumises à de brusques changements environnementaux. Ainsi, l'expression génétique doit s'adapter rapidement pour faire face à ces variations soudaines et multiples.

En effet, les bactéries sont des organismes pouvant vivre dans des environnements très différents. Dans la nature, non seulement elles colonisent des niches très diverses (lac salins, océans, volcans, intestins des mammifères...) mais elles entrent aussi en compétition avec d'autres espèces bactériennes. A cet égard, les bactéries sont exposées constamment à un grand nombre de stress du fait des changements rapides et constants de leur milieu naturel. On définit par « stress » toute modification chimique ou physique de ce milieu. Ainsi, les variations de température, de pH, de concentrations en nutriments et en sels, mais aussi la concentration en oxygène sont autant de stress que l'environnement peut exercer sur ces organismes et qui modifient leurs conditions de croissance. Dans le but de survivre et de s'adapter à leur environnement, les bactéries ont développé un grand nombre de systèmes leur permettant de « sentir » l'environnement et d'adapter l'expression de leurs gènes en fonction des conditions environnementales, et notamment en condition de stress.

Un des moyens mis au point par les bactéries pour répondre à ces stimuli est l'expression de certains types d'acides ribonucléiques (ARN) que sont les petits ARN régulateurs, mais également l'utilisation de systèmes dits « systèmes à deux composants » sur lesquels je reviendrai par la suite dans ce manuscrit.

Je vais tout d'abord présenter les principaux mécanismes de régulation de la transcription et notamment la plus commune à savoir la régulation du démarrage de la transcription. Ensuite, les mécanismes de régulation post-transcriptionnelle seront passés en revue.

1.1.1. La transcription bactérienne

La transcription est la première étape permettant l'expression du génome. Le but est de copier sélectivement certaines parties du génome (unité de transcription) et de produire d'un à plusieurs milliers de copies de molécules d'ARN dont la séquence est dictée par la section du génome considérée ou matrice d'acide désoxyribonucléique (ADN). Par la suite, la matrice ARN permettra la traduction en protéine de cette section. L'ARN polymérase, l'acteur majeur de la transcription, va reconnaître des séquences spécifiques (ou promoteurs) situées le long du chromosome bactérien puis la transcription se déroulera en suivant les trois étapes que sont le démarrage, l'allongement et la terminaison que je détaillerai par la suite.

Il a été montré que lors de la phase exponentielle de croissance les ARN polymérases sont préférentiellement localisées au niveau des gènes codant les ARN ribosomiques (ARNr) et les ARN de transfert (ARNt) (Cabrera and Jin 2003). Elles sont ainsi peu disponibles pour la transcription d'autres gènes. Les bactéries ont donc développé différentes stratégies permettant d'exprimer des gènes spécifiques en condition exponentielle de croissance. Je propose d'aborder la régulation de la transcription dans la seconde partie de ce chapitre.

1.1.1.1. Les différentes étapes de la transcription et ses principaux acteurs

1.1.1.1.1. Le promoteur

Les promoteurs bactériens peuvent être définis comme des séquences conservées et constituées de deux éléments : des séquences correspondant à la boîte -10 ou Pribnow box et celles définissant la boîte -35, la numérotation étant relative au +1 de démarrage de la transcription (figure 1). Le rôle du promoteur est de servir de site de reconnaissance pour l'ARN polymérase et de site de démarrage de la transcription. Chez les bactéries, les gènes sont le plus souvent organisés en opérons et souvent regroupés de manière fonctionnelle. De plus, un gène ou un opéron peuvent être transcrits à partir d'un ou plusieurs promoteurs ce qui donne naissance à divers ARN messagers (ARNm).

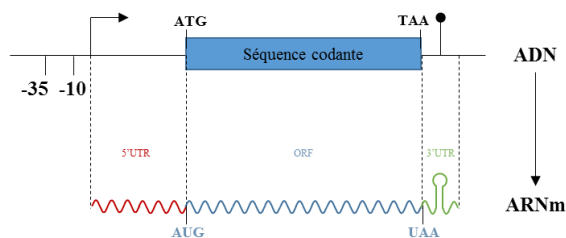


Figure 1 : Organisation d'une unité transcriptionnelle bactérienne. Sur l'ADN, on peut définir les positions -35 et -10 du promoteur. La séquence codante de l'ARNm possède un codon de démarrage (le plus souvent AUG) et se termine par un codon stop (ici UAA). La région 5' non traduite de l'ARNm (5'UTR : 5' UnTranslated Région) est suivie de la phase codante (ORF : Open reading Frame) puis de la région 3' non traduite (3'UTR).

1.1.1.2. L'ARN polymérase

L'ARN polymérase est l'enzyme responsable de la polymérisation des nucléotides (nts) pour donner à partir d'une matrice ADN une molécule d'ARN. Il existe un seul type d'ARN polymérase chez les bactéries pour transcrire tous les ARN. Cette polymérase est composée de cinq sous-unités : deux sous-unités α , une sous-unité β , une sous-unité β' et une sous-unité ω . A cette enzyme « cœur » s'ajoute une sous-unité σ variable. Le facteur majoritaire retrouvé chez *E. coli* est le facteur sigma 70 (σ^{70}). Il permet notamment la transcription des gènes impliqués dans les fonctions dites de « ménage » et dont les produits sont nécessaires au maintien des fonctions cellulaires de base. L'expression de ces gènes de « ménage » est souvent constitutive. Outre ce facteur σ^{70} , il existe chez *E. coli* six autres facteurs sigma (tableau I) appelés facteurs sigma alternatifs (Osterberg, del Peso-Santos et al. 2011). Ces facteurs entrent en compétition avec le facteur σ^{70} pour se fixer à l'ARN polymérase de façon à la rediriger vers des promoteurs spécifiques.

Tableau 1 : Les facteurs sigma chez *Escherichia coli*

Facteur σ	Gène	Rôle	Séquence consensus du promoteur			Référence
			-35	Espacement	-10	
σ^{70}/σ^D	<i>rpoD</i>	Fonctions essentielles en phase exponentielle	TTGACA	16-18	TATAAT	(Harley and Reynolds 1987)
σ^{54}/σ^N	<i>rpoN</i>	Métabolisme et transport d'azote	TGGCAC	5	TTGC(A/T)	(Merrick 1993)
σ^{38}/σ^S	<i>rpoS</i>	Entrée en phase stationnaire et stress	TTGACA	16	CTATACT	(Lacour, Kolb et al. 2003)
σ^{32}/σ^H	<i>rpoH</i>	Choc thermique et stress (cytoplasmique)	CTTGA	14	GNCCCCATNT	(Wang and deHaseth 2003)
σ^{28}/σ^F	<i>fliA</i>	Synthèse du flagelle /chimiotactisme	TAAAGTTT	11	GCCGATAA	(Koo, Rhodius et al. 2009)
σ^{24}/σ^E	<i>rpoE</i>	Stress membranaire	GGAACTT	16	TCNAA	(Rezuchova, Miticka et al. 2003)
$\sigma^{19}/\sigma^{FecI}$	<i>fecI</i> *	Transport du citrate de fer	GAAAAT	15	TGTCCT	(Enz, Mahren et al. 2003)

N : A, T, C ou G

* : le seul promoteur connu à ce jour pour être transcrit par σ^{19} est l'opéron *fecABCDE*

A l'exception de sigma E, les facteurs sigma alternatifs ne sont pas essentiels (De Las Penas, Connolly et al. 1997; Hayden and Ades 2008).

1.1.1.1.3. Les étapes de la transcription

Le processus de transcription s'effectue en trois étapes : le démarrage, l'allongement et enfin la terminaison (figure 2).

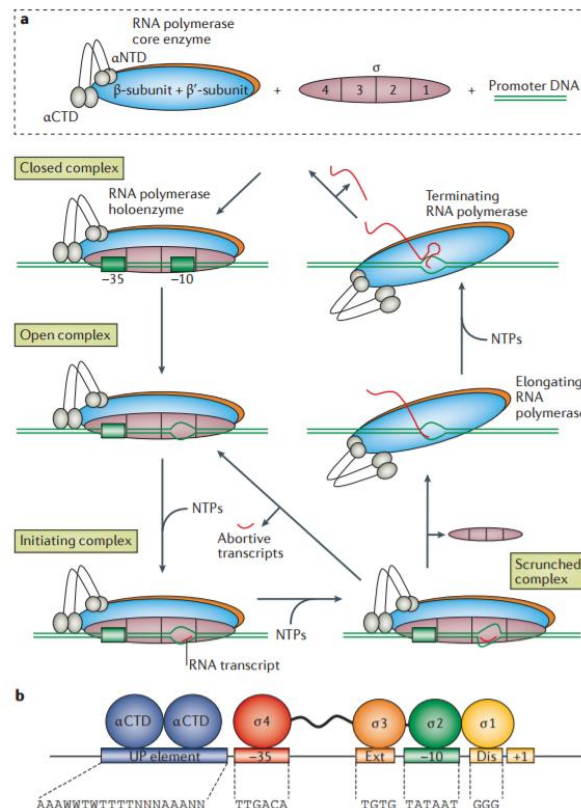


Figure 2 : La transcription chez les bactéries. a) L'ARN polymérase active contient l'enzyme « cœur » et un facteur sigma qui, sous la forme dite d'holoenzyme, va se fixer sur le promoteur pour former le complexe fermé. Dans ce complexe, les brins de la matrice ADN sont appariés et l'enzyme se localise sur le côté de l'hélice ADN. Le complexe fermé subit une transition vers le complexe ouvert où cette fois les brins d'ADN se séparent sur environ 17 paires de bases pour former la bulle de transcription. Un seul des deux brins sert de matrice et la nature du brin (sens ou antisens) est définie par le promoteur qui aura fixé la polymérase dans une orientation définie. L'addition de nucléotides triphosphates (NTPs) déclenche une nouvelle transition où le complexe ouvert va se transformer en complexe de démarrage. Au site actif de l'enzyme les deux premiers nucléotides sont alignés sur la matrice puis assemblés l'un à l'autre. La chaîne va s'allonger dans le sens $5' \rightarrow 3'$ par incorporation de nucléotides à l'extrémité 3' de la chaîne naissante. Bien souvent l'incorporation des 10 premiers nucléotides est inefficace et l'enzyme libère des transcrits abortifs. On parle de complexe abortif (scrunched complex) pour désigner cette étape où la polymérase peine à s'échapper du promoteur. Passé ces 10 nucléotides, l'enzyme se libère du promoteur et forme un complexe stable permettant l'allongement du transcrit ARN et conduisant à la libération du facteur sigma. La réaction se poursuit jusqu'à ce que la polymérase rencontre un terminateur. Cette étape de terminaison consiste en la libération du transcrit, la dissociation du brin d'ADN et de l'enzyme qui peut être recyclée en s'associant à un nouveau facteur sigma. b) Principaux éléments reconnus par l'ARN polymérase : les séquences UP des promoteurs les plus puissants, les régions -35 et -10 et la séquence Dis pour élément discriminatoire parfois retrouvée. Issue de (Browning and Busby 2016).

Dans un premier temps, l'ARN polymérase va s'associer à un facteur sigma, le plus souvent σ^{70} et va reconnaître spécifiquement la séquence promotrice d'un gène. Elle va ensuite dissocier les deux brins d'ADN sur environ 17 paires de bases pour former la bulle de transcription. Une matrice simple brin est donc disponible pour débiter la transcription. L'incorporation des premiers nucléotides et la formation de la première liaison phosphodiester est une étape délicate car le complexe ainsi formé est relativement instable. Il est possible à ce stade que la transcription s'arrête relarguant des ARN très courts, entre 2 à 8 nucléotides (Vo, Hsu et al. 2003). On appelle ce phénomène la transcription abortive. Si cette étape est franchie la transcription continue et c'est l'allongement de l'ARNm qui se fait dans le sens $5' \rightarrow 3'$ par ajout successif de nucléotides complémentaires de la matrice ADN à raison de 50 nucléotides par seconde. Après l'ajout d'une dizaine de nucléotides il a été rapporté que le facteur sigma se détachait et que l'enzyme core poursuivait seule le processus jusqu'au site de terminaison (Hansen and McClure 1980).

Cette dernière étape peut se faire de façon dépendante ou indépendante du facteur de transcription Rho selon la nature du site de terminaison. En effet, si l'ARN polymérase transcrit une séquence palindromique riche en guanosine (G) et en cytidine (C) puis suivie d'une séquence riche en adénosine (A), c'est la terminaison spontanée (Brendel, Hamm et al. 1986). La transcription de cette séquence va générer une tige boucle stable dans l'ARNm. La séquence répétée d'uridine (U) en fin d'ARNm appariés aux A de la matrice ADN est a contrario assez instable ce qui aura pour conséquence le décrochage de l'ARNm et la fermeture de la bulle de transcription (figure 3). Chez *E. coli*, la transcription d'environ la moitié des ARN codants et 64% des ARN non codants (ARNr, ARNt, petits ARN régulateurs) a une terminaison spontanée ou intrinsèque (Lesnik, Sampath et al. 2001). Les autres terminaisons sont médiées par le facteur Rho.

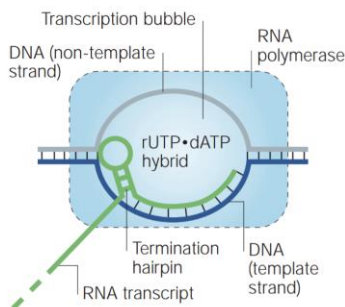


Figure 3 : Mécanisme de terminaison intrinsèque. La tige boucle des terminateurs intrinsèques est immédiatement suivie par une séquence répétée en uridine qui va s'apparier avec la séquence correspondante riche en A sur la matrice ADN. La formation de ce duplex ARN-ADN relativement instable va avoir pour conséquence le décrochage de l'ARNm et la fermeture de la bulle de transcription. Issue de (Greive and von Hippel 2005).

Les terminateurs Rho dépendants présentent une grande diversité de séquence mais partagent néanmoins quelques caractéristiques communes à savoir la présence de séquences nommées « *rut* » (pour « Rho utilisation ») et « *tsp* » (pour « transcription stop point ») séparées par une séquence qui peut s'étendre jusqu'à plus de 200 paires de bases (Nudler and Gottesman 2002). Rho est une hélicase ATP-dépendante qui se fixe sur l'ARN naissant au niveau du site « *rut* » et le parcourt jusqu'à son extrémité 3' (Boudvillain, Figueroa-Bossi et al. 2013). Lorsque Rho rencontre la polymérase qui pause au voisinage du site « *tsp* » cela aboutit au décrochage de l'ARN et la fermeture de la bulle de transcription sur la matrice ADN (figure 4).

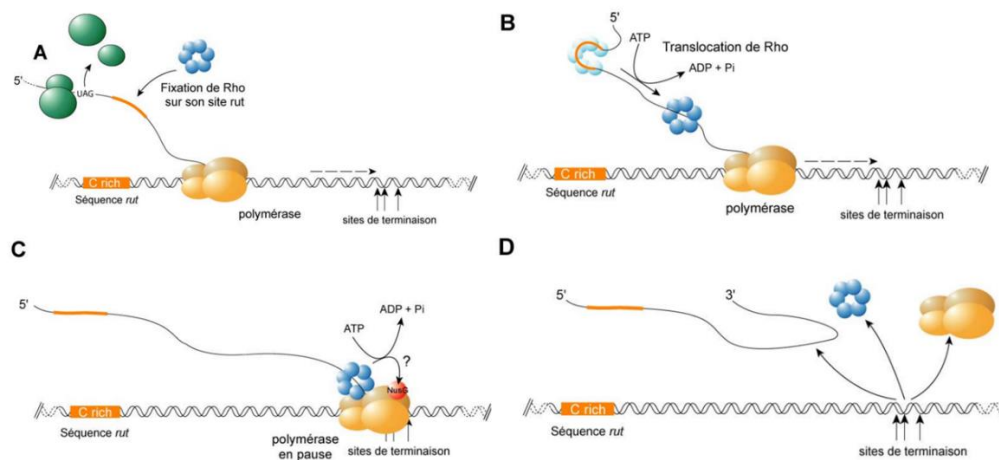


Figure 4 : Mécanisme de la terminaison Rho dépendante. Rho reconnaît et se fixe au site « *rut* » (A) puis se déplace de 5' en 3' le long de l'ARN naissant. Rho est arrêté au niveau de la polymérase (C) puis le complexe se dissocie et l'ARN est relargué (D). D'après (Vieu 2004).

1.1.1.2. Régulation de la transcription

De manière générale, l'expression des gènes doit répondre à des signaux extracellulaires comme par exemple la présence de glucose ou de magnésium dans le milieu de culture. Ces signaux sont ensuite intégrés par des protéines dites « protéines régulatrices » qui auront soit une action positive (activateurs) soit une action négative (répresseurs). Ces protéines sont des protéines de liaison à l'ADN qui reconnaissent spécifiquement des sites situés à proximité des gènes qu'elles régulent. Un activateur augmentera la transcription d'un gène en facilitant la transcription de celui-ci par l'ARN polymérase alors qu'un répresseur diminuera voire bloquera la transcription par cette même ARN polymérase. La plupart des régulateurs transcriptionnels agissent au moment du démarrage de la transcription en modulant l'affinité de la polymérase ou bien en entrant en

compétition avec elle. En effet, cette étape est la moins coûteuse énergétiquement pour exercer un contrôle puisque dans ce cas seuls les gènes dont les produits sont nécessaires sont transcrits en ARNm.

1.1.1.2.1. Régulation du démarrage par des facteurs globaux

Outre les facteurs sigma qui exercent une forme de régulation globale en dirigeant la polymérase sur des promoteurs spécifiques, il existe d'autres régulateurs globaux chez *E. coli* qui sont capables de réguler plusieurs dizaines voire centaines de gènes. Sept régulateurs globaux décrits chez *E. coli* (Lrp, FNR, CRP, ArcA et les protéines de nucléoïde bactérien IHF, HNS, Fis) sont suffisants pour moduler directement l'expression d'environ la moitié des gènes de cet organisme (Martinez-Antonio and Collado-Vides 2003) formant ainsi un réseau de régulation très complexe comme l'illustre la figure 5.

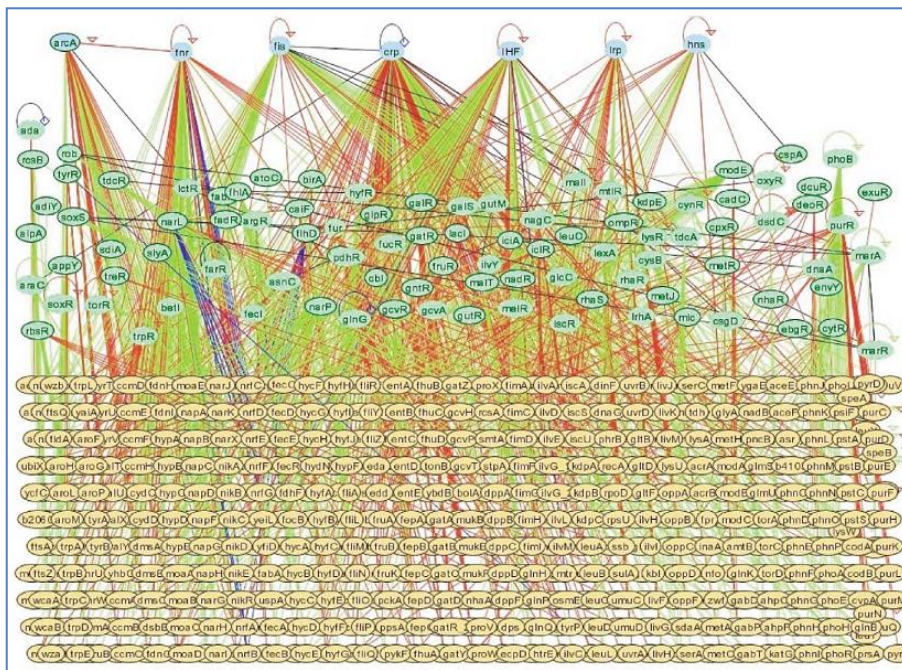


Figure 5 : Vue d'ensemble du réseau de régulation transcriptionnelle chez *Escherichia coli*. Les gènes régulés sont indiqués dans des bulles jaunes, les facteurs de transcription sont symbolisés par des bulles vertes, enfin les facteurs de transcription considérés comme des régulateurs globaux sont indiqués dans des bulles bleues. Les lignes vertes montrent une activation et les lignes rouges une répression. Les lignes bleues symbolisent une double régulation (activation et répression). Issue de (Martinez-Antonio and Collado-Vides 2003).

CRP ou CRP-AMPC est le premier facteur de transcription global à avoir été purifié (Zubay, Schwartz et al. 1970). La protéine CRP interagit avec l'adénosine monophosphate cyclique (AMPC) puis se fixe sur le promoteur de ces gènes cibles. A ce jour on décrit environ 100

promoteurs activés ou réprimés par CRP selon la position de son site de fixation sur l'ADN (Kolb, Busby et al. 1993). Son régulon est majoritairement composé de gènes ou opérons codant des enzymes du métabolisme des sucres ou pour divers systèmes de transport des sucres.

Un autre des régulateurs globaux important pour le métabolisme est FNR. Il régule plus de 100 gènes en condition de carence en oxygène (Kiley and Beinert 1998) impliqués notamment dans la résistance au stress acide, le chimiotactisme ou encore la structure de l'enveloppe. La concentration de FNR intracellulaire est relativement constante mais son activité est modulée par la présence d'oxygène qui va directement jouer sur l'état d'oxydation du centre fer-soufre présent dans chacun des monomères de FNR. En effet, la forme active de FNR est un homodimère dont chaque sous-unité possède un centre fer-soufre [4Fe-4S]. FNR peut ainsi reconnaître des séquences de type TTGATNNNNATCAA. En revanche en présence d'oxygène le centre [4Fe-4S] est converti en centre [2Fe-2S] ce qui empêche la dimérisation de FNR, de ce fait l'inactivant.

ArcA est également important lors de carence en oxygène. Ce régulateur fait partie du système à deux composants ArcB/ArcA. En condition anoxique, la kinase ArcB va s'autophosphoryler puis phosphoryler le régulateur transcriptionnel ArcA, lui permettant de réprimer de nombreux gènes impliqués dans des fonctions aérobies (Pena-Sandoval and Georgellis 2010). De plus, la présence de certains métabolites anaérobies comme le D-lactate ou encore le pyruvate stimule ArcB (Georgellis, Kwon et al. 1999). La plupart du temps ArcA agit de concert avec FNR.

Outre ces quatre facteurs de transcription, trois autres facteurs globaux sont associés spécifiquement au nucléoïde : Fis, H-NS et IHF. Le nucléoïde correspond à la région où se localise le chromosome bactérien. Les protéines qui lui sont associées jouent un rôle important pour la structure de l'ADN, notamment pour sa compaction, et peuvent, en fonction de leur positionnement, inhiber ou activer la transcription de certaines régions de l'ADN. Bien que leurs mécanismes d'action soient peu connus on sait cependant qu'IHF (Integration Host Factor) aide au surenroulement de l'ADN et assure des fonctions de régulation. En effet, IHF facilite la formation de la bulle de transcription, promouvant ainsi la formation du complexe de démarrage de la transcription. On compte environ 20 000 molécules d'IHF par cellule (Ishihama, Kori et al. 2014). H-NS (Histone-like nucleoid structuring protein) quant à elle condense et surenroule l'ADN. Cette capacité lui permet « d'éteindre » l'expression d'environ 5% des gènes d'*E. coli*, notamment ceux impliqués dans l'enveloppe cellulaire et jouant ainsi un rôle clé dans les

changements environnementaux et l'adaptation au stress (Hommais, Krin et al. 2001). Il est intéressant de préciser que l'expression des protéines nucléoïdes varie en fonction des conditions de croissance. Par exemple, en conditions optimales de croissance, le régulateur Fis est la protéine dominante du nucléoïde avec 75 000 copies par cellule ; en revanche, en phase stationnaire, son niveau n'est plus que de 400 copies (Ishihama, Kori et al. 2014). Fis régule de très nombreux gènes dont les fonctions vont de la virulence au métabolisme central en passant par la biosynthèse des acides aminés, la formation des biofilms ou la motilité. Selon des données de puces à ADN environ 21% des gènes seraient régulés par Fis (Cho, Knight et al. 2008).

Enfin, environ 10% des gènes d'*E. coli* sont affectés positivement ou négativement par Lrp (Leucine-responsive regulatory protein) et ce sous la dépendance ou non de la leucine. On compte notamment de nombreux gènes impliqués dans la biosynthèse et le catabolisme des acides aminés, le métabolisme carboné, le transport des nutriments ou la formation des pili (Cho, Barrett et al. 2008). La fixation de Lrp induit une courbure de l'ADN, la classant dans la littérature comme protéine nucléoïde bien que ceci soit débattu.

1.1.1.2.2. Régulation du démarrage par des facteurs spécifiques

Il existe plus d'une centaine de régulateurs spécifiques qui, en se fixant au site opérateur, vont jouer un rôle soit activateur soit répresseur (Martinez-Antonio and Collado-Vides 2003). Avec la découverte de la régulation de l'opéron lactose, Jacob et Monod (Jacob and Monod 1961) ont montré que l'action d'un régulateur transcriptionnel pouvait souvent dépendre du milieu de croissance, phénomène qui a été vérifié maintes fois avec d'autres régulateurs transcriptionnels. Comme présenté sur la figure 6, l'opéron lactose est composé de trois gènes : *lacZ*, *lacY*, *lacA*. Ces trois gènes codent respectivement une protéine nécessaire à la dégradation du lactose (β -galactosidase), une protéine nécessaire à son transport (perméase) et enfin une protéine de type galactoside O-acetyltransférase dont la fonction est encore peu connue. Le répresseur LacI, codé par le gène *lacI* en amont des gènes *lacZYA*, réprime la transcription de ces derniers lorsque le lactose est absent du milieu de culture. Lorsque la bactérie se trouve en présence de lactose comme source de carbone, une petite quantité de cette molécule pénètre dans la cellule et est dégradée en allolactose. Ce produit de dégradation va fixer LacI, modifier sa structure et diminuer son affinité pour le site opérateur situé en amont des gènes *lacZYA* libérant ainsi le promoteur reconnu par l'ARN polymérase.

Il a également été montré que la transcription de l'opéron lactose était activée par le régulateur global CRP uniquement dans les conditions où le glucose est absent du milieu de culture (Wilson, Zhan et al. 2007). Ceci met en évidence l'intégration de plusieurs informations disponibles par un seul et même gène.

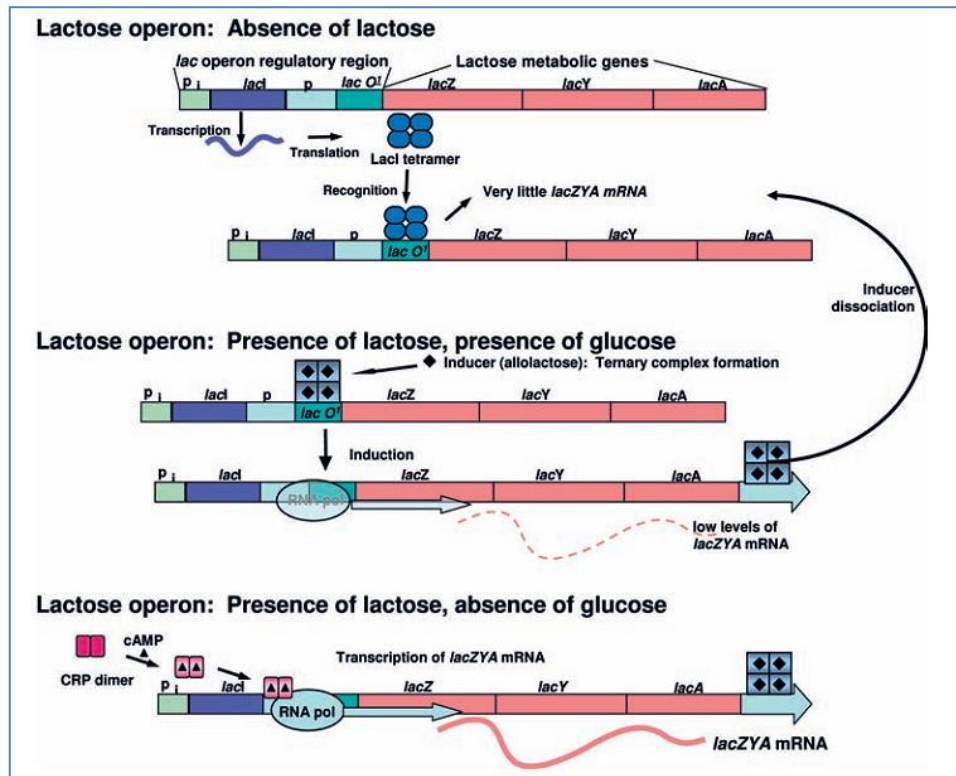


Figure 6 : Schéma de la régulation de l'opéron lactose. En absence de lactose, l'ARNm *lacI* est transcrit à partir du promoteur P_i et traduit en protéine LacI. Cette protéine possède une forte affinité pour l'opérateur lactose O^1 . Le site de fixation de LacI recouvrant le promoteur *lac*, sa présence inhibe la transcription des gènes de l'opéron *lac* que sont *lacZ*, *lacY* et *lacA* codant respectivement les enzymes que sont la β -galactosidase, la lactose perméase et la thiogalactoside transacétylase impliquées dans l'utilisation du lactose comme source de carbone. En présence de lactose et de glucose, l'1,6-allolactose, un métabolite secondaire est produit et se fixe sur LacI diminuant son affinité pour l'opérateur O^1 et conduisant à une faible transcription de l'opéron *lac*. En présence de lactose mais cette fois en absence de glucose, le régulateur transcriptionnel CRP facilite la transcription de l'opéron *lac* augmentant ainsi fortement la synthèse des enzymes permettant de métaboliser le lactose. Issue de (Wilson, Zhan et al. 2007).

Dans certains cas, l'activité des régulateurs transcriptionnels est elle-même régulée par des modifications post-traductionnelles comme la phosphorylation dans le cas des systèmes à deux composants qui seront abordés dans le chapitre 1.2.

1.1.1.3. Régulation de l'allongement et de la terminaison de la transcription

Une fois la transcription démarrée le complexe formé est relativement stable, néanmoins on note que l'allongement est un processus discontinu interrompu par des « pauses » jouant un rôle clé

dans sa régulation. En effet, il a été montré que les pauses transcriptionnelles, en modulant la structure de l'ARN et la vitesse de l'ARN polymérase, pouvaient la rendre plus sensible aux modifications structurales ou à l'action de régulateurs. Par exemple, il a été montré que la polymérase était capable de « pauser » sur un site « *ops* ». Cette pause permet le recrutement du facteur RfaH qui va modifier le complexe d'élongation, empêcher la terminaison et permettre ainsi la transcription des gènes en aval du site *ops* (Artsimovitch and Landick 2002). Les gènes contenant ces séquences sont particulièrement impliqués dans la pathogénèse.

Dans le cas des régulations par terminaison prématurée de la transcription on parle d'atténuation transcriptionnelle. Chez les procaryotes, l'atténuation est basée souvent sur le couplage entre transcription et traduction. En effet, la traduction de l'ARNm débute alors que celui-ci n'est pas entièrement synthétisé. Le phénomène d'atténuation a pour la première fois été décrit par l'équipe de Charles Yanofsky dans les années 1970 dans le cadre de l'étude de l'opéron tryptophane *trp* chez *E. coli*. Néanmoins, le terme d'atténuation n'a été employé qu'à partir de 1974 par un chercheur japonais, T. Kasai.

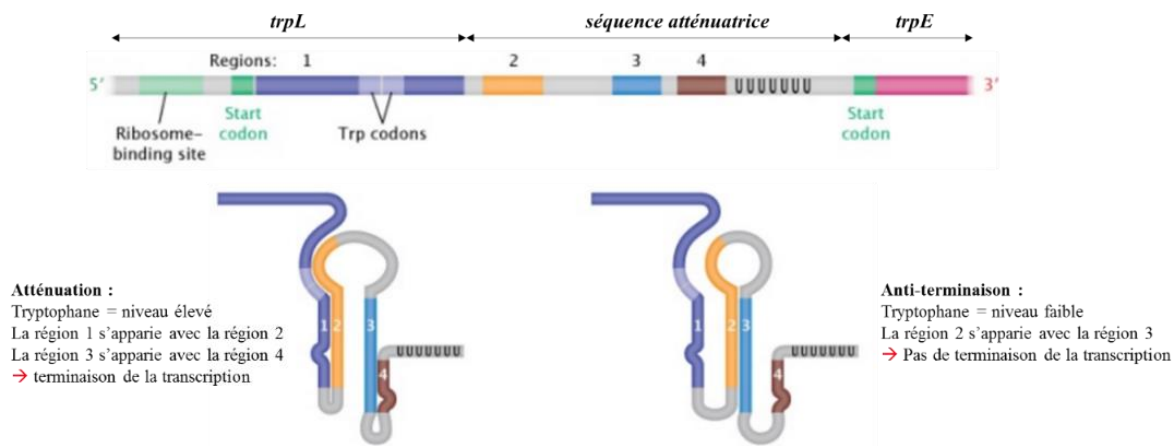


Figure 7 : Régulation de l'expression de l'opéron tryptophane par atténuation de la transcription. La concentration en tryptophane contrôle la transcription de l'opéron *trp*. A forte concentration en tryptophane, l'appariement entre les régions 1 et 2 d'une part et 3 et 4 d'autre part aboutit à la formation d'une structure permettant la terminaison de la transcription. A faible concentration en tryptophane, les régions 2 et 3 s'apparient laissant une structure qui cette fois ne permet pas la terminaison de la transcription et laisse donc la possibilité de transcrire les gènes en aval. Dans ce cas, une région du 5'UTR peut être traduite. Adaptée de (Ralston 2008).

L'opéron *trp* porte les gènes codant les enzymes requises pour la synthèse du tryptophane avec dans l'ordre *trp E*, *trp D*, *trp C*, *trp B* et *trp A*. Une séquence de 141 nucléotides sépare l'extrémité 5' de l'ARNm du codon de démarrage de la traduction du gène *trpE* (Yanofsky 1981). La région 5' de cette séquence code un peptide, TrpL, de 14 acides aminés dont deux tryptophanes et est

suivie de la séquence d'atténuation. L'ARN naissant obtenu après transcription des séquences « leader » et « atténuatrices » peut adopter deux structures secondaires différentes en fonction de la concentration d'ARN de transfert chargé en tryptophane (figure 7). Si cette concentration est suffisante, le peptide leader est traduit en entier, ce qui a pour conséquence la formation de la structure terminatrice (formée par les régions 3 et 4) et donc l'arrêt de la transcription. A l'inverse, si la concentration en ARN de transfert chargé en tryptophane est faible, le ribosome va s'arrêter au niveau de l'un ou l'autre des codons tryptophane du peptide leader. Dans ces conditions, la formation d'une structure anti-terminatrice (formée par les régions 2 et 3) est favorisée empêchant la formation de la tige-boucle terminatrice et permettant ainsi l'expression des gènes de l'opéron *trp*.

La régulation de la transcription a été le premier niveau de contrôle mis en évidence. Cependant, c'est loin d'être le seul. En effet, il existe également une régulation traductionnelle des ARNm qui sont traduits en protéines. Cependant, tous les ARN ne sont pas traduits en protéines notamment les ARN ribosomiques, les ARN de transfert et d'autres ARN qui peuvent également exercer des fonctions régulatrices comme nous le verrons dans le chapitre 1.3. Par ailleurs, la stabilité des ARNm peut être également fortement contrôlée.

1.1.2. La traduction bactérienne

La traduction est le processus permettant de synthétiser des protéines grâce aux ribosomes à partir de l'information génétique portée par les ARNm. Nous aborderons dans ce chapitre les principaux acteurs de ce processus, ses principales étapes mais également sa régulation.

1.1.2.1. Les différentes étapes de la traduction et ses principaux acteurs

1.1.2.1.1. Le ribosome

Le ribosome est un complexe formé de deux sous-unités nommées 50S et 30S en référence à leur coefficient de sédimentation dans un gradient de sucrose (figure 8). La petite-sous unité (30S) est constituée de l'ARN ribosomique 16S associé à 21 protéines (Green and Noller 1997). Son rôle est essentiel dans la fidélité du processus de traduction. En effet, elle contrôle l'interaction correcte entre les codons de l'ARNm et ceux des anti-codons portés par les ARN de transfert. 34 protéines ainsi que les ARN ribosomiques 23S et 5S forment la grande sous-unité (50S). Elle est notamment responsable de la formation des liaisons peptidiques entre la chaîne d'acides aminés (aa) naissante

et l'acide aminé correspondant au codon suivant présent sur la matrice ARNm. Il est intéressant de préciser que chaque protéine ribosomique est présente en un seul exemplaire à l'exception de L7/L12 où L7 est une forme acétylée de L12 à son extrémité N-terminale et qu'il existe deux dimères de L7/L12 dans un ribosome. Enfin L26 est identique à S20 mais elles sont présentes dans des sous-unités différentes.

A partir du début des années 2000, de nombreuses analyses de la structure du ribosome à plus forte résolution ont permis de répondre à de nombreuses interrogations sur le déroulement du processus de traduction (Ban, Nissen et al. 2000; Nissen, Hansen et al. 2000; Noeske, Wasserman et al. 2015). Par exemple, elles ont pu mettre en évidence le rôle fondamental des ARNr dans les étapes d'association des sous-unités, de positionnement de l'ARNt aux sites A/P/E ou encore dans la lecture des codons mais surtout sur la formation de la liaison peptidique (Nissen, Hansen et al. 2000). Les protéines ribosomiques seraient donc surtout présentes pour assurer le repliement correct et complexe des ARNr.

Il existe au sein du ribosome trois sites où peut se positionner l'ARNt (figure 8) :

- Site A : c'est dans ce site que se positionne l'ARNt aminoacylé (associé à son acide aminé spécifique) en entrant dans le ribosome
- Site P : il renferme l'ARNt portant la chaîne polypeptidique en cours de synthèse
- Site E : l'ARNt est présent dans ce site sous sa forme déacylée avant d'être dissocié de l'ARNm et du ribosome

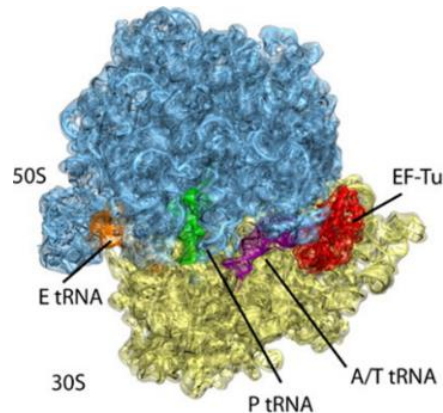


Figure 8 : Structure en cryo-microscopie électronique du ribosome d'*E. coli* associé au facteur EF-Tu à une résolution de 6.7 Å. Cette structure met en évidence les trois sites E, P et A où peut se localiser l'ARNt au cours du processus de traduction. D'après (Villa, Sengupta et al. 2009).

1.1.2.1.2. L'ARN messenger

L'ARNm porte les différents signaux essentiels à la reconnaissance par le ribosome. On peut diviser l'ARNm en trois portions : la région 5' non traduite (5'UTR : « 5' Untranslated Region ») contenant le site de fixation du ribosome (RBS : « Ribosome Binding Site »), la phase ouverte de lecture (ORF : « Open Reading Frame ») qui constitue la région codante et enfin la région 3' non traduite (3'UTR : « 3' Untranslated Region ») comprenant notamment le terminateur. Le ribosome se fixe sur l'ARNm au niveau du RBS où deux éléments assurent son recrutement et le positionnement du ribosome à ce site. Ces deux éléments sont la séquence Shine Dalgarno (SD) et le codon de démarrage. La séquence SD est complémentaire de la séquence dite « anti-SD » présente en 3' de l'ARN ribosomique 16S (Shine and Dalgarno 1974). Le consensus défini pour la séquence SD est UAAGGAGGUGA dont les résidus les plus conservés sont marqués en gras (Barrick, Villanueva et al. 1994). Ensuite, entre 5 et 8 nucléotides après le SD se trouve sur le messenger le codon de démarrage. Chez *E. coli*, il existe trois codons de démarrage : AUG, GUG et UUG. AUG est présent chez plus de 90% des messagers alors que GUG et UUG sont retrouvés respectivement dans 8% et 1% des ARNm. On peut noter qu'outre ces trois codons de démarrage canoniques il arrive que des codons alternatifs soient utilisés. C'est le cas du gène *infC* chez *E. coli* qui démarre avec un codon AUU. Enfin, il est intéressant de mentionner que la distance entre le SD et le codon de démarrage peut affecter l'efficacité de traduction (Chen, Bjerknes et al. 1994).

1.1.2.1.3. Les ARN de transfert

Les ARN de transfert ont un rôle clé dans le décodage de la matrice ARNm. Ils assurent la reconnaissance et la liaison entre un codon et un acide aminé précis. En effet, l'ARNt possède trois nucléotides complémentaires du codon avec lequel il s'apparie d'où le nom d'anti-codon. Ces ARN non codants d'environ 75 nucléotides sont reconnaissables à leur structure secondaire caractéristique en forme de « feuille de trèfle ». Leur structure tertiaire en forme de « L » est également très conservée. L'extrémité 3' de l'ARNt qui se termine toujours par un CCA, porte l'acide aminé lorsqu'il est sous sa forme aminoacylé. Spatialement, la boucle anticodon se situe à l'opposé de l'extrémité 3' de l'ARNt (figure 9).

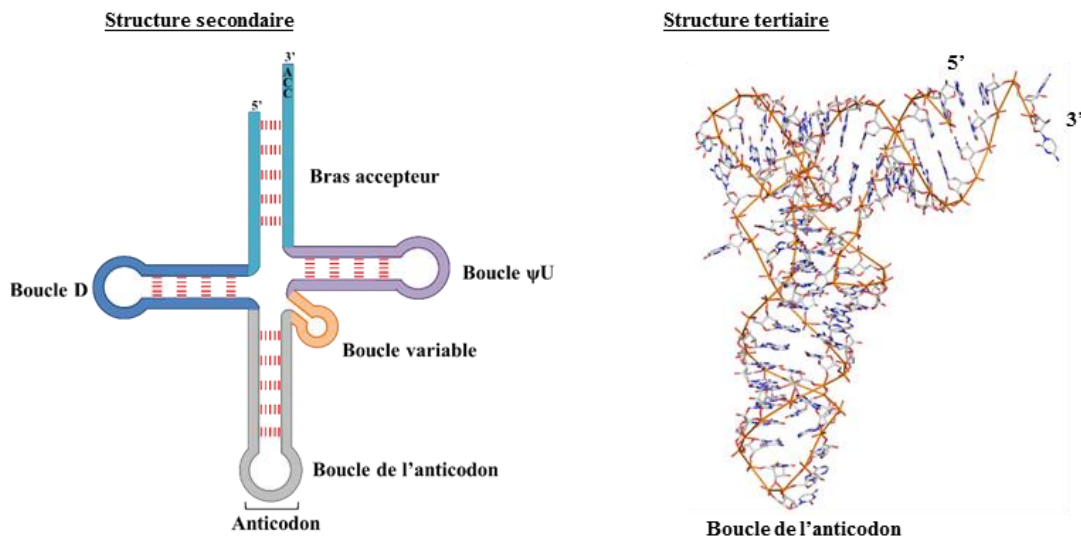


Figure 9 : Structure d'un ARN de transfert. A gauche de la figure, une représentation schématique de la structure secondaire montrant la forme caractéristique de ces ARN en forme de « feuille de trèfle ». A droite de la figure, la structure cristallographique de l'ARNt^{fmet} d'*E. coli* met en évidence la structure tertiaire en forme de « L » retrouvée chez tous les ARNt. La représentation a été faite sur Pymol à partir de la structure 5L4O disponible sur la PDB (résolution 2.8Å).

1.1.2.1.4. Les différentes étapes de la traduction

La traduction est tout comme la transcription un processus en trois étapes : l'initiation ou démarrage, l'allongement et enfin la terminaison.

Le démarrage de la traduction est l'étape clé du processus. Au cours de cette étape, l'ARNt initiateur ou ARNt^{fmet}, l'ARNm et la sous unité 30S s'assemblent pour former le complexe de pré-démarrage puis de démarrage (figure 10). Cette étape requiert l'intervention des facteurs de démarrage IF-1, IF-2 et IF-3. Dans un premier temps seuls IF-1 et IF-3 sont associés à la sous-

unité 30S. IF1 est localisé au site A du ribosome et empêche a priori la fixation de l'ARNt^{fmet} à cette position. IF3 quant à lui maintient les sous-unités ribosomiques séparées. Ensuite, IF-2 associé à une molécule de GTP (guanosine triphosphate) va reconnaître l'ARNt^{fmet} et permettre son entrée au site P du ribosome. Le complexe formé peut ainsi interagir avec l'ARNm par reconnaissance de la séquence SD et du codon de démarrage. De plus, IF-2 va stimuler l'association de la sous-unité 50S au reste du complexe déjà formé via son activité GTPase. En effet, après la fixation de l'ARNt^{fmet}, la sous unité 50S est recrutée, ce qui occasionne le relargage d'IF1 et IF3. Le GTP fixé à IF2 est hydrolysé et ce facteur quitte le complexe de démarrage. A ce stade l'ARNt^{fmet} est positionné au site P du ribosome et le site A est encore inoccupé.

L'allongement débute avec la fixation au site A de l'ARNt aminoacylé spécifique du second codon de l'ARNm associé avec le facteur EF-Tu et une molécule de GTP. EF-Tu catalyse la fixation de l'ARNt aminoacylé et est relargué sous forme EF-Tu-GDP. La formation de la liaison peptidique est quant à elle catalysée par l'ARNr 23S. S'en suit une étape de translocation, catalysée par le facteur EF-G, afin de libérer le site d'entrée. Ces étapes se répètent pour chaque acide aminé incorporé dans la chaîne. Dès la fixation du second acide aminé l'allongement est lancé et le ribosome va pouvoir catalyser la formation des liaisons peptidiques à une vitesse d'environ 20 acides aminés par seconde. Cette vitesse est évidemment fonction de la concentration en acides aminés disponibles, de celle des ARNt et enfin de la structure de l'ARNm traduit. Ce cycle d'allongement se termine lorsqu'un codon de terminaison occupera le site A du ribosome. Chez *E. coli*, il existe trois codons stop : UAA, UAG et UGA. Selon le codon ce dernier sera reconnu par un facteur de terminaison qui entraînera la dissociation entre le peptide formé et l'ARNt situé au site P. RF1 reconnaîtra les codons UAA et UAG et RF2 les codons UAA et UGA. Enfin, le ribosome se dissociera sous l'action du facteur de recyclage RRF et du facteur EF-G.

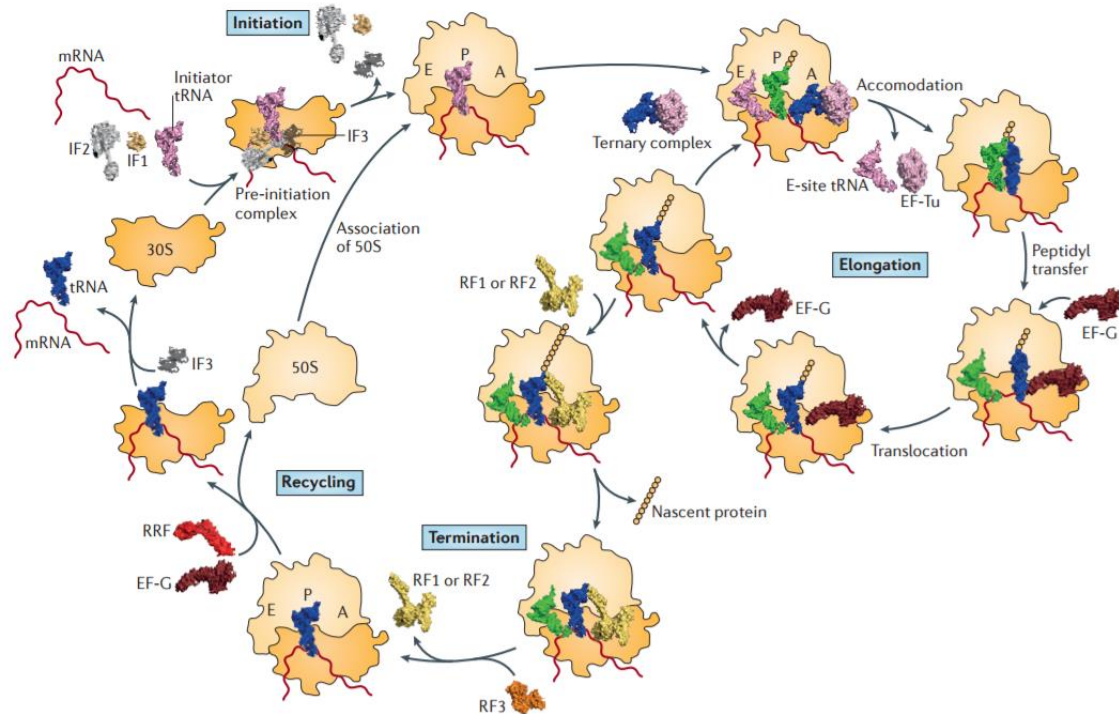


Figure 10 : Les principales étapes de la traduction procaryote. La traduction est initiée par l'assemblage des facteurs de démarrage IF1, IF2 et IF3 avec l'ARNm, la sous-unité 30S du ribosome et l'ARNt initiateur ou ARNt^{fmet} (fMet-tRNA^{fMet}). L'ARNt^{fmet} se positionne au site P du ribosome et contacte le codon de démarrage pour former le complexe de pré-démarrage. Ensuite, les facteurs IF1, IF2 et IF3 se dissocient, la sous-unité 50S du ribosome se fixe à son tour pour former le complexe de démarrage. L'ajout séquentiel des acides aminés suivants à la chaîne polypeptidique naissante nécessite la formation d'un complexe ternaire composé de l'ARNt porteur de l'acide aminé à intégrer, du facteur d'allongement EF-Tu et d'une molécule de GTP. Ce complexe entre au site A du ribosome, si la boucle anti-codon de l'ARNt reconnaît le codon présent sur l'ARNm, EF-Tu hydrolyse son GTP et est relâché du ribosome. L'ARNt correct positionné au site A subit une rotation au site P ou accommodation. C'est au site P qu'a lieu la formation de la première liaison peptidique. Pour qu'un nouvel allongement puisse se produire, l'ARNt déacétylé doit se déplacer du site P au site E. En même temps, l'ARNm doit se décaler de trois nucléotides pour présenter le codon suivant. On appelle ces deux mouvements coordonnés la translocation. Celle-ci est stimulée par le facteur EF-G qui, lorsqu'il se libère, permet à un nouveau cycle d'allongement de débuter. Lorsqu'un codon stop entre au site A, il est reconnu par les facteurs RF1 ou RF2. Ces facteurs vont promouvoir le relargage de la chaîne polypeptidique synthétisée. Le facteur RF3 va ensuite permettre la dissociation des facteurs RF1 et RF2 du ribosome. Pendant la phase dite de « recyclage » les facteurs RRF et EF-G aident à la dissociation des sous-unités 30S et 50S. A ce stade la sous-unité 30S est encore fixée au dernier ARNt et à la matrice ARNm. C'est la fixation du facteur IF3 qui permettra la libération de ces deux éléments. Issue de (Keiler 2015).

1.1.2.2. Régulation du démarrage de la traduction

Comme pour la régulation transcriptionnelle, il existe des contrôles traductionnels au niveau du démarrage, de l'allongement et de la terminaison de la traduction. Le démarrage est l'étape clé du contrôle car c'est une étape limitante et, de plus, réguler cette étape constitue une économie d'énergie. On peut noter que l'avantage du contrôle traductionnel est de répondre plus rapidement aux stimuli extérieurs. En effet, ce contrôle ciblant directement les messagers, il peut y avoir une

modification plus rapide du niveau protéique que dans le cas d'une répression transcriptionnelle par exemple où les ARNm préexistants à la régulation peuvent encore être traduits.

La régulation traductionnelle est souvent basée sur la possibilité de démarrer la traduction et donc sur la structure de la région 5' UTR de l'ARNm et la disponibilité du RBS (Romby and Springer 2007). La régulation traductionnelle fait intervenir des protéines qui se fixent le plus souvent à proximité du RBS et plus récemment il a été montré qu'elle pouvait être effectuée par des ARN régulateurs. Enfin l'ARNm peut servir lui-même de régulateur ; c'est le cas des riboswitches. Dans ce chapitre seule la répression de la traduction par les protéines sera détaillée. La régulation opérée par des ARN fera l'objet du chapitre 1.3.

Les gènes codant les protéines ribosomiques et les facteurs de traduction sont les principales cibles du contrôle traductionnel médié par des protéines. En effet, en condition de croissance rapide, les ribosomes représentent 50% de la masse sèche de la bactérie (Zengel and Lindahl 1994). Ainsi, selon la phase de croissance le nombre de ribosomes par cellule peut varier de 6700 à 71000 (Bremer and Dennis 1996). On comprend donc aisément l'intérêt de réguler la synthèse des différents constituants ribosomiques. Ainsi, dans chacun des opérons codants des protéines ribosomiques, une des protéines possède la capacité de se fixer à l'ARNm de l'opéron pour le réguler.

Deux mécanismes classiques de régulation du démarrage de la traduction existent et consistent en l'inhibition de celui-ci (figure 11) : la compétition avec le ribosome ou le piégeage de celui-ci.

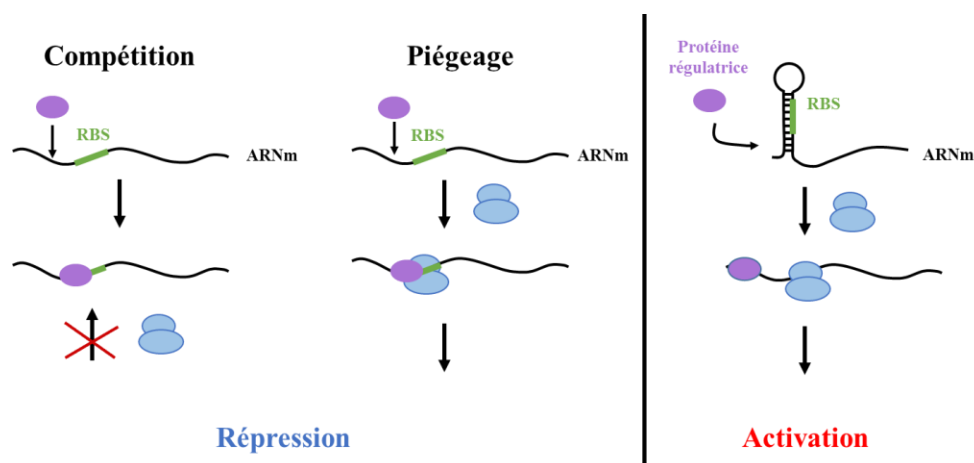


Figure 11 : Mécanismes de régulation de la traduction par des facteurs protéiques. Classiquement, la fixation du ribosome peut être inhibée par des mécanismes de compétition ou de piégeage. Il est également possible d'activer le démarrage de la traduction.

Le premier mécanisme est basé sur la compétition entre la protéine régulatrice et le ribosome pour la fixation à la région de démarrage de la traduction. La protéine régulatrice étant fixée à proximité du RBS, celle-ci empêche l'accès du ribosome au RBS par simple encombrement stérique.

Je me propose, pour illustrer ce propos, de détailler l'autocontrôle de la protéine ribosomique L20 qui a été étudié par des membres de notre laboratoire. L20 est une protéine de la sous-unité 50S du ribosome. Cette protéine est codée par le gène *rplT* qui fait partie de l'opéron IF3 constitué dans l'ordre des gènes *thrS* codant la thréonyl-t-RNA-synthétase, *infC* codant le facteur d'initiation de la traduction IF3, *rpmI* codant la protéine ribosomique L35 et *rplT* (figure 12A). L20 assure deux fonctions principales dans la cellule : elle est essentielle à l'assemblage de la sous-unité 50S et réprime la traduction des gènes *rpmI* et *rplT*.

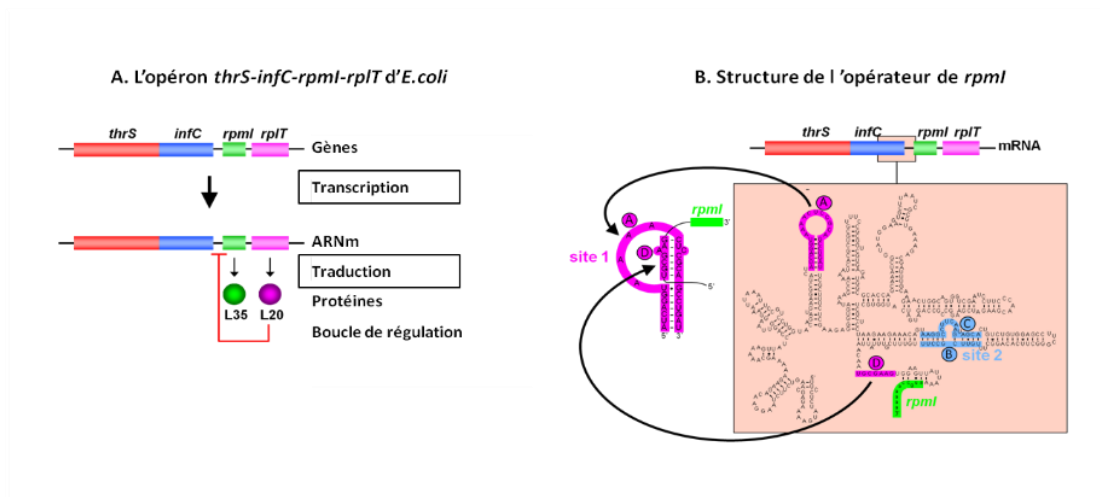


Figure 12 : Autocontrôle traductionnel de la protéine ribosomique L20 A : Schéma de l'opéron IF3 d'*E. coli* B : Structure de l'opérateur de *rpmI*, le trait noir représente l'interaction à longue distance qui permet de former le pseudonœud.

L'expression des gènes *rpmI* et *rplT* est contrôlée négativement au niveau de leur traduction par L20. Ce contrôle passe par la fixation de la protéine à deux sites de l'ARNm de *rpmI* : les sites 1 et 2 (figure 12B). Le site 1 résulte d'une interaction entre deux régions distantes d'environ 300 nucléotides formant ainsi un pseudonœud. Le site 2 est une hélice irrégulière. Il a été démontré *in vitro* par des expériences de toeprint que L20 pouvait entrer en compétition avec la sous-unité 30S du ribosome pour la fixation à la région de démarrage de l'ARNm de *rpmI*, inhibant de ce fait la traduction de ce dernier (Haentjens-Sitri, Allemand et al. 2008). Comme les synthèses de *rpmI* et de *rplT* sont couplées, L20 autorégule donc également sa propre traduction. Il a été également montré que la présence des deux sites est indispensable à l'autocontrôle même si L20 ne se fixe

qu'à un seul des deux sites et pas aux deux simultanément (Allemand, Haentjens et al. 2007). Ainsi une seule molécule de L20 se fixe à l'opérateur de l'ARNm de *rpmI*.

Le second mécanisme d'inhibition du démarrage de la traduction consiste à piéger le ribosome (figure 11). Dans ce cas, la fixation de la protéine régulatrice induit une modification structurale dans l'ARNm qui aura pour conséquence de bloquer le ribosome dans une étape précoce du démarrage. Contrairement au premier mécanisme le ribosome se fixe au RBS. Ce mécanisme a notamment été étudié lors de la mise en évidence de la régulation de l'ARNm *rpsO* par son propre produit, la protéine ribosomique S15 (Portier, Dondon et al. 1990).

Dans de rares cas, la fixation de la protéine régulatrice libère le RBS d'une structure secondaire permettant ainsi sa traduction par les ribosomes (figure 11). Il s'agit donc plus d'une levée d'inhibition que d'une réelle activation.

1.1.3. Stabilité des ARN messagers

Chez *E. coli*, les ARN représentent 20.5 % de la masse sèche de la bactérie en phase exponentielle de croissance (Neidhardt and Umbarger 1996). Parmi tous ces ARN on distingue, entre autres, d'une part les ARN stables, que sont les ARN ribosomiques (5S, 16S, 23S) et les ARN de transfert, et d'autre part les ARNm. Ces trois catégories d'ARN sont estimées respectivement à 18 700 copies pour chacun des ARNr, 198 000 copies pour la totalité des ARNt et 1380 copies pour l'ensemble des ARNm dans une cellule en phase exponentielle à 37°C c'est-à-dire lorsque la vitesse de croissance est importante.

La stabilité des ARN et en particulier celle des ARNr et ARNt provient de leur forte structuration mais également de leur association avec des protéines qui les protègent des ribonucléases (RNases). Ces enzymes spécifiques sont chargées des processus de dégradation des ARN. On peut préciser que les RNases jouent aussi un rôle majeur dans la maturation des ARN stables afin qu'ils adoptent leur structure fonctionnelle (Kushner 1996). En cas de défaillance de ces transcrits elles sont également chargées de les dégrader. De même, en condition de carence métabolique, les RNases sont capables de dégrader les ARN stables.

La modulation du niveau des ARNm est un processus cellulaire important et est à part entière un mécanisme de régulation post-transcriptionnelle chez les bactéries mais également un formidable avantage adaptatif. En effet, les ARNm bactériens ont des temps de demi-vie moyens de 1,3 min

à 37°C ce qui est notamment dû au couplage entre la transcription et la traduction (Arraiano, Andrade et al. 2010). Le « turnover » de ces ARN permet ainsi aux bactéries d'adapter relativement rapidement et facilement leur expression génétique aux conditions environnementales et notamment dans des situations de stress.

Par ailleurs, la stabilité des messagers bactériens est très dépendante de la traduction et participe au contrôle en suivant le contrôle traductionnel. En effet, un messenger moins traduit ne sera plus protégé par les ribosomes et donc plus exposé à la dégradation par des RNases. Il est intéressant de préciser que la stabilité peut cependant être affectée indépendamment de la traduction et être un mécanisme primaire de contrôle comme nous le verrons avec l'exemple de l'autocontrôle de la RNase III.

1.1.3.1. Généralités sur les ribonucléases

Les bactéries possèdent une grande variété de ribonucléases dont certaines ne sont présentes que chez certaines espèces bactériennes. On distingue deux grandes classes de ribonucléases : les endoribonucléases et les exoribonucléases (Rochat, Bouloc et al. 2013; Hui, Foley et al. 2014). Les principales endoribonucléases bactériennes sont la RNase E, la RNase Y et la RNase III. Leur rôle est de cliver les ARN au niveau de sites spécifiques présents au sein de l'ARN cible. Ces sont des enzymes hydrolytiques c'est-à-dire qu'elles utilisent une molécule d'eau afin de rompre la liaison phosphodiester et libèrent un ion monophosphate. Comme nous le verrons par la suite, la RNase E est essentielle dans la dégradation des messagers notamment chez *E. coli*. Chez les espèces ne possédant pas cette enzyme, la RNase Y assure la dégradation des messagers par voie endonucléolytique. Enfin, la RNase III est particulièrement importante car elle possède la capacité de dégrader des doubles brins d'ARN, structures qui bloquent en général l'activité des autres RNases.

Les exoribonucléases quant à elles dégradent les ARN majoritairement à partir de l'extrémité 3'. On distingue les 3' exonucléases à activité phosphorolytique (clivage par ajout d'un phosphate inorganique), telles que la PNPase et la RNase PH, des 3' exonucléases à activité hydrolytique telles que la RNase II et la RNase R (Arraiano, Andrade et al. 2010). Environ 90% de l'activité exoribonucléolytique est assurée par la RNase II. De plus, elle joue un rôle majeur dans la maturation des ARN de transfert mais aussi celle du 5S. Malgré l'efficacité de la RNase II et de la PNPase, ces enzymes sont inhibées par des structures secondaires dans l'ARNm, comme des

terminateurs Rho-indépendants (McLaren, Newbury et al. 1991), nécessitant l'intervention d'autres enzymes telles que la RNase R qui, elle, peut les dégrader (Cheng and Deutscher 2005). Enfin, la RNase II et la PNPase ne sont pas capables de dégrader des ARN inférieurs à environ 10 nucléotides (nts), ces derniers pouvant servir d'amorce à l'ARN polymérase et générant des transcrits alternatifs. Ainsi, chez *E. coli* l'oligoribonucléase codée par le gène *orn* est essentielle mais ce n'est pas le cas chez toutes les espèces bactériennes (Ghosh and Deutscher 1999). La présence de ces petits morceaux d'ARN ou nano-ARN est a priori délétère pour la cellule ou en tout cas est une étape nécessitant d'être régulée. Chez *E. coli* il a été proposé que ces nano-ARN pourraient former une nouvelle catégorie d'ARN régulateurs (Vvedenskaya, Sharp et al. 2012).

De manière générale les ARN bactériens sont dégradés par clivage endonucléolytique suivi d'une dégradation dans le sens 3'→5' médiée par une exoribonucléase (figure 13). La seule exception connue à ce jour est la RNase J1 de *Bacillus subtilis* (*B. subtilis*) pour laquelle une activité 5'→3' a été démontrée (Mathy, Benard et al. 2007). Une très faible activité 5'→3' (100 fois moins que J1 alors que l'activité endonucléolytique est équivalente) est également détectée pour la RNase J2 chez *B. Subtilis* (C. Condon, communication personnelle).

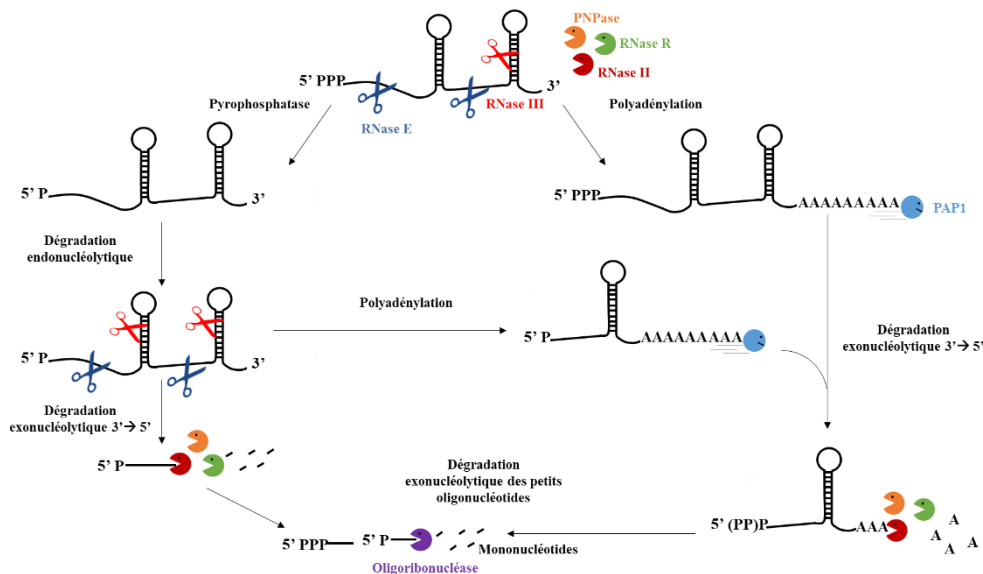


Figure 13 : Voies de dégradation des ARN chez *E. coli*. La dégradation de la majorité des transcrits débute par une coupure endonucléolytique par la RNase E. Cette enzyme préfère les extrémités 5' monophosphate mais cette règle n'est pas stricte. La RNase III est l'enzyme responsable des clivages endonucléolytiques initiaux des ARN structurés. Contrairement à la RNase E qui ne clive que des simples brins, la RNase III clive des doubles brins. Après le premier clivage endonucléolytique, le transcrit linéaire est rapidement dégradé par des exoribonucléases 3'→5' (RNase II, RNase R et PNPase). La RNase R, contrairement à la RNase II et la PNPase, est efficace sur les ARN hautement structurés. Une voie mineure dans la cellule est la dégradation exonucléolytique de transcrit de pleine taille.

La polyA-polymérase (PAP) ajoute une queue poly A à l'extrémité 3' de l'ARN ce qui crée une « plateforme » où peuvent se fixer les exoribonucléases. De multiples sites de polyadénylation et de dégradation exonucléolytique peuvent ainsi surmonter la présence de structures secondaires. Les petits oligonucléotides relargués par les exoribonucléases peuvent enfin être dégradés en mononucléotides par l'oligoribonucléase. Adaptée de (Arraiano, Andrade et al. 2010).

Il existe une vingtaine de ribonucléases chez *E. coli* (tableau 2) dont 3 seulement sont essentielles (Mackie 2013). Précisons néanmoins qu'elles ne sont pas toutes impliquées dans la dégradation des messagers. La RNase E est une endoribonucléase essentielle dans les processus de régulation post-transcriptionnels médiés par des ARN régulateurs chez les entérobactéries. Je vais donc dans la suite de ce chapitre me concentrer sur cette enzyme.

Tableau 2 : Les ribonucléases d'*Escherichia coli*

Endoribonucléases			Exoribonucléases			Systèmes toxine-antitoxine *		
Enzyme	Gène	Essentialité	Enzyme	Gène	Essentialité	Enzyme	Gène	Essentialité
RNase E	<i>rne</i>	oui	Oligo RNase	<i>orn</i>	oui	ChpB	<i>chpB</i>	non
RNase P	<i>rnpA, rnpB</i>	oui	RNase D	<i>rnd</i>	non	MazF	<i>mazF</i>	non
RNase G	<i>rng</i>	non	RNase PH	<i>rph</i>	non	RelE	<i>relE</i>	non
RNase HI	<i>rnhA</i>	non	RNase R**	<i>rnr</i>	non	RNase LS	<i>rnlA</i>	non
RNase HII	<i>rnhB</i>	non	RNase T	<i>rnt</i>	non	YhaV	<i>yhaV</i>	non
RNase BN	<i>rbn</i>	non	PNPase**	<i>pnp</i>	non	YoeB	<i>yoeB</i>	non
RNase I	<i>rna</i>	non	RNase II**	<i>rnb</i>	non			
RNase III	<i>rnc</i>	non						

* : systèmes toxine-antitoxine comprenant une RNase qui devient toxique pour la cellule en l'absence de son antitoxine protéique

** : devient essentielle en l'absence d'autre RNases (Cheng, Zuo et al. 1998)

1.1.3.2. La RNase E

La RNase E est une des ribonucléases essentielles d'*E. coli* (Mackie 2013). Elle est notamment impliquée dans les premières étapes de dégradation des ARNm mais également dans la maturation de certains ARN stables. Le rôle régulateur de la RNase E et plus précisément son lien avec les petits ARN sera détaillé dans le chapitre 1.3.1.4.2.

D'un point de vue structural, c'est une protéine constituée de 1061 acides aminés et organisée en deux grands domaines fonctionnels (figure 14). En N-terminal (acides aminés 1→529) on trouve le domaine catalytique de cette enzyme dont la structure est de forme globulaire. La portion C-terminale (acides aminés 530→1061) de la protéine est plutôt non structurée et sert de plateforme d'interaction avec ses nombreux partenaires du dégradosome que sont la PNPase (PolyNucléotide Phosphorylase) sous forme de trimère, l'hélicase RhlB et une émolase (Hui, Foley et al. 2014). Il

est intéressant de mentionner que seul le domaine catalytique est essentiel (Bouvier and Carpousis 2011).

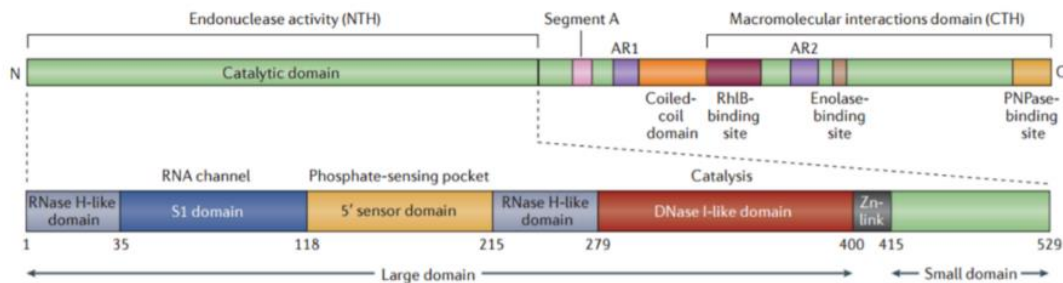


Figure 14 : Schéma de l'organisation en domaines de la RNase E. Issue de (Mackie 2013).

La forme active de la RNase E est un homotétramère localisé majoritairement au niveau de la membrane interne (Khemici, Poljak et al. 2008). La perte de cette localisation peut entraîner des défauts de croissance cellulaire. Cette endoribonucléase clive les séquences simple brin d'ARN riches en A et U et son inactivation entraîne une stabilisation très importante de l'ensemble des transcrits (Babitzke and Kushner 1991). Il a été montré qu'elle clive préférentiellement les ARNm possédant une extrémité 5' monophosphate (Mackie 1998) générées soit par des coupures endoribonucléolytiques soit par l'action de la pyrophosphatase RppH (Deana, Celesnik et al. 2008). On peut néanmoins indiquer que par un mécanisme dit « d'entrée directe » la RNase E peut cliver des ARNm avec une extrémité 5' triphosphate (Dreyfus 2009). C'est d'ailleurs via ce mécanisme qu'elle est capable de s'autoréguler en clivant son propre messenger (Jain and Belasco 1995). Le schéma présenté dans la figure 15 récapitule les mécanismes de clivage par la RNase E. On peut enfin conclure que l'état de phosphorylation des extrémités 5' et leur accessibilité est un facteur important de la stabilité des messagers.

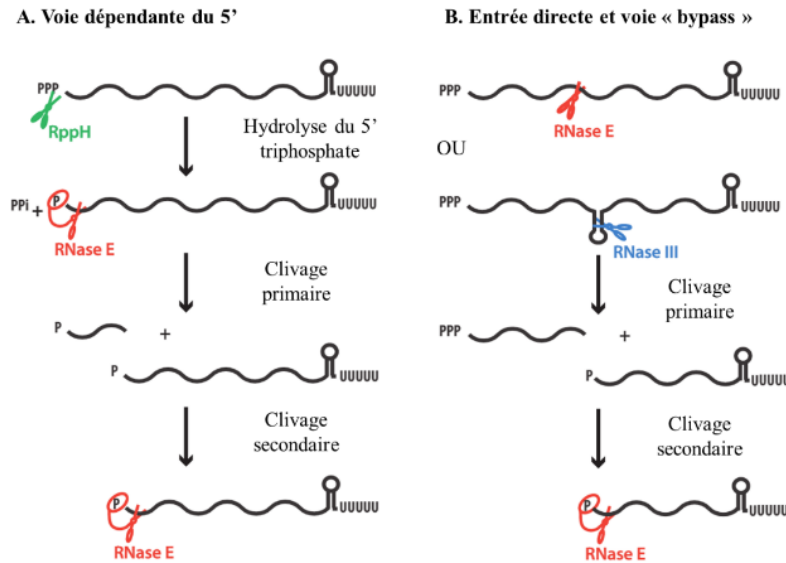


Figure 15 : Mécanismes de clivage par la RNase E. A. Dans cette voie de dégradation dépendante de l'extrémité 5', l'enzyme RppH hydrolyse le 5' triphosphate du transcrit primaire afin de libérer une extrémité 5' monophosphate. Cette extrémité monophosphate permettra de stimuler l'activité de la RNase E pour les clivages endonucléolytiques primaires et secondaires. B. Dans le mécanisme « d'entrée directe », le clivage endonucléolytique est indépendant de la nature de l'extrémité 5'. Enfin, dans la voie dite de contournement ou « bypass », le clivage de la tige boucle par la RNase III permet de libérer une extrémité 5' monophosphate qui stimulera la coupure de cet ARN par la RNase E. D'après (Bouvier and Carpousis 2011).

La RNase E a la capacité d'interagir avec d'autres enzymes. Elle s'associe notamment avec la PNPase, RhlB et l'énolase pour former le dégradosome. Il semblerait que l'association de la PNPase et de la RNase E facilite la dégradation exonucléolytique des ARN issus de clivage endonucléolytique. L'hélicase RhlB permet de dissocier les brins d'ARN et ainsi les exposer plus facilement aux RNases du complexe. Le rôle de l'énolase dans le dégradosome est en revanche mal compris. Il a été suggéré qu'elle puisse capter l'état métabolique cellulaire. Le regroupement spatial de ces enzymes au sein du dégradosome permettrait de rassembler toutes les étapes du processus dans un compartiment cellulaire et limiterait ainsi la diffusion des intermédiaires réactionnels. Par ailleurs, on peut noter que la RNase E interagit aussi avec une autre enzyme : la PolyA Polymérase (Raynal and Carpousis 1999). Cette enzyme ajoute une queue « poly A » aux ARN en cours de dégradation facilitant ainsi le processus. Contrairement aux ARN eucaryotes, les queues poly A bactériennes sont donc des éléments destabilisateurs. Une interaction entre Hfq et la RNase E a aussi été décrite (voir chapitre 1.3.1.4.2.).

1.1.3.3. La dégradation, un mécanisme de contrôle à part entière

Pour illustrer ce propos, j'ai choisi de présenter l'autocontrôle de la RNase III. Cette endonucléase très conservée dans le vivant, spécifique des régions double-brin, mature notamment les ARNr et les ARNt (Condon and Putzer 2002). Il a aussi été montré qu'elle dégradait les duplex petit ARN/ARNm (Opdyke, Kang et al. 2004; Vogel, Argaman et al. 2004). La RNase III est codée par le gène *rnc* situé dans un opéron portant en outre les gènes *era* et *recO* codant respectivement une GTPase et une protéine de réparation de l'ADN (Court, Gan et al. 2013 pour revue). L'extrémité 5' du transcrit *rnc-era-recO* porte une tige-boucle clivée par la RNase III elle-même (Bardwell, Regnier et al. 1989). Suite à cette coupure, l'ARNm est dégradé par la RNase E. Cette autorégulation, diminue d'environ 5 fois le niveau d'ARNm et influe directement sur l'expression d'*rnc* et d'*era*. L'autorégulation est possible car la quantité d'RNase III dans la cellule est faible (moins de 0.01% des protéines totales). En outre, la RNase III peut être positivement titrée par ses principaux substrats, les ARNr (en milieu riche surtout), qui l'éloignent ainsi de son propre ARNm. Comme la RNase III est impliquée dans la maturation des ARNr, l'autorégulation permet que son niveau soit cohérent avec les besoins de la cellule. Il a aussi été montré que l'autorégulation était présente indépendamment du niveau de traduction de *rnc* (Matsunaga, Simons et al. 1997). Ainsi la dégradation d'un messager peut être un mécanisme de contrôle à part entière.

1.2. Les systèmes à deux composants

1.2.1. Définitions – distribution génomique

Le terme de « système à deux composants » est apparu pour la première fois en 1986 pour désigner une nouvelle classe de protéines régulatrices chez les bactéries partageant des domaines conservés et un mécanisme commun de réponse à un signal environnemental (Nixon, Ronson et al. 1986). La fonction de base des systèmes à deux composants est donc de relayer un signal spécifique à travers la membrane jusqu'aux effecteurs qui vont déclencher la réponse cellulaire. Ce relais fait intervenir une série de transferts de phosphate et est médiée par deux types de protéines : une protéine kinase à histidine (HK), généralement membranaire qui sert de senseur et une protéine régulatrice ou « response regulator » (RR) (Stock, Robinson et al. 2000). La protéine HK possède un domaine central conservé et une extrémité N-terminale variable expliquant la spécialisation de ce domaine dans la détection d'un signal spécifique (figure 16). La région C-terminale contient le domaine histidine kinase qui s'autophosphoryle en réponse à son signal spécifique. Le groupement phosphate est ensuite transféré à un résidu aspartate conservé de la protéine régulatrice. La phosphorylation de cette dernière modifie ses propriétés biochimiques ce qui se traduit le plus fréquemment par l'activation de sa capacité à fixer l'ADN et dans certains cas par la fixation à une autre protéine. En effet, dans de nombreux cas la phosphorylation conduit à la dimérisation des RR qui sous cette forme ont la capacité de se fixer aux promoteurs de leurs gènes cibles. Il est donc communément admis que la forme active du régulateur est celle phosphorylée.

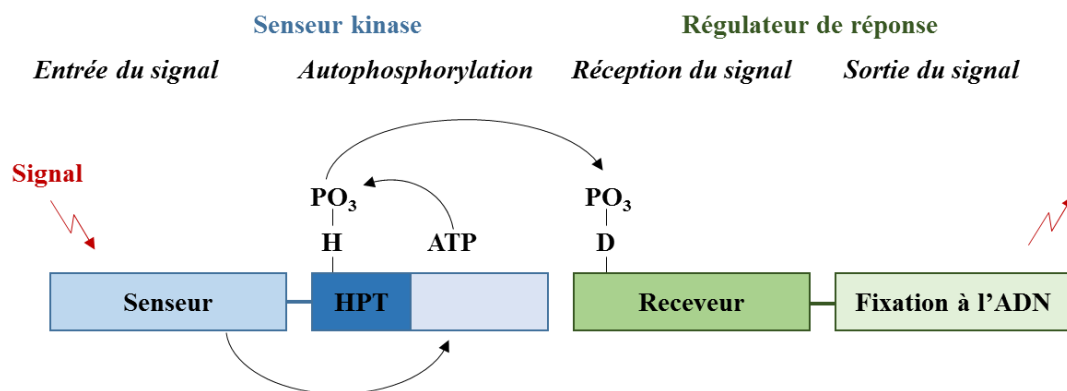


Figure 16 : Schéma présentant les systèmes à deux composants classiques La protéine HK est composée d'un domaine senseur et d'un domaine autokinase incluant la région histidine phosphotransférase HPT. La protéine RR est composée d'un domaine receveur et d'un domaine de fixation à l'ADN permettant la transduction du signal. H = histidine ; D = aspartate.

Nous savons aujourd'hui que les systèmes à deux composants sont présents à la fois chez les procaryotes et chez les eucaryotes mais sont particulièrement prédominants chez les bactéries (Stock, Robinson et al. 2000). Néanmoins, le nombre de systèmes à deux composants varie fortement d'une espèce bactérienne à l'autre. En effet, il n'y en a aucun chez *Mycoplasma genitalium* alors qu'il en existe plus de 150 chez *Nostoc punctiforme*. Le tableau 3 présente le nombre de systèmes à deux composants chez différents organismes modèles.

Tableau 3 : Nombre de systèmes à deux composants répertoriés chez différents organismes modèles (valeurs issues de la base de données MIST2.2 (Ulrich and Zhulin 2010))

Organisme considéré	Nombre d'histidines kinases	Nombre de régulateurs de réponse
<i>Escherichia coli</i> K12 substr. MG1655	29	31
<i>Caulobacter crescentus</i> CB15	58	44
<i>Helicobacter pylori</i> 26695	3	6
<i>Pseudomonas aeruginosa</i> PAO1	59	67
<i>Bacillus subtilis</i> 168	38	33
<i>Clostridium difficile</i> 630	49	54
<i>Enterococcus faecalis</i> V583	14	18
<i>Listeria monocytogenes</i> EGD	16	16
<i>Nostoc punctiforme</i> PCC 73102	160	98
<i>Staphylococcus aureus</i> NCTC8325	17	16

Il est intéressant de noter que le nombre de systèmes à deux composants croît globalement avec la taille du génome (Capra and Laub 2012) comme on peut le voir sur la figure 17.

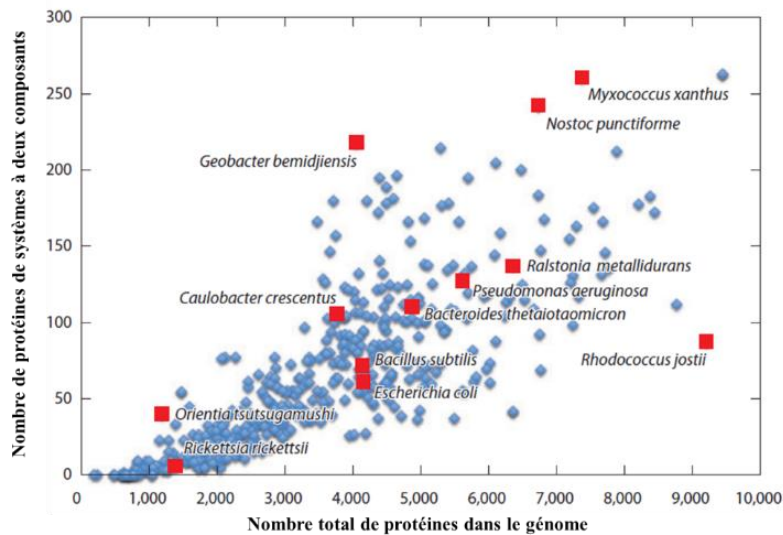


Figure 17 : Diversité des systèmes à deux composants en fonction de la taille de différents génomes bactériens. D'après (Capra and Laub 2012).

De plus, le nombre de systèmes à deux composants est étroitement corrélé aux niches écologiques et environnementales où se développe l'organisme considéré. En effet, les bactéries que l'on retrouve typiquement dans des environnements relativement stables auront moins de systèmes à deux composants dans leurs génomes que celles soumises à des changements brusques et constants de leur environnement et parfois indépendamment de la taille de leur génome. Par exemple, *Myxococcus xanthus* possède 136 histidines kinases et 127 régulateurs de réponse mais il est connu que cette bactérie résiste à de nombreux environnements notamment à des pH extrêmes où elle peut sporuler. Ce sont par exemple les systèmes à deux composants FrzCD et Che qui sont au cœur d'une cascade complexe de régulation permettant à cet organisme de se déplacer et modifier son cycle de vie en fonction de la disponibilité en nutriments (Zusman, Scott et al. 2007).

Chez *E. coli*, notre organisme d'étude, il a été décrit 29 histidines kinases et 31 régulateurs de réponse (Mizuno 1997). Ces éléments sont répartis en 24 systèmes à deux composants classiques, 5 phosphorelais et 2 régulateurs de réponse orphelins. Les phosphorelais sont une variation des systèmes à deux composants classiques où le groupement phosphate de l'histidine kinase est transféré intramoléculairement avant d'être transféré au régulateur de réponse (Laub and Goulian 2007). Le tableau 4 présente tous les systèmes à deux composants connus à ce jour chez *E. coli*.

Tableau 4 : Systèmes à deux composants chez *E. coli*. Le nom des systèmes précise en premier le nom de la protéine senseur et en second celui de la protéine régulatrice. Les phosphorelais complexes comprennent les systèmes avec une kinase hybride et les systèmes incluant un régulateur intermédiaire entre la kinase et le régulateur traductionnel. Liste réalisée à partir de la base de donnée en ligne EcoCyc (Keseler, Mackie et al. 2013).

Systèmes à deux composants classiques	Organisation génétique	Autres nomenclatures
AtoS-AtoC	<i>atoS-atoC</i>	
BaeS-BaeR	<i>baeS-baeR</i>	
BasS-BasR	<i>basR-basS</i>	PmrB/PmrA
CheA-CheB-CheY*	<i>cheA-cheW</i> et <i>cheB-cheY</i>	
CpxA-CpxR	<i>cpxR-cpxA</i>	CpxA (<i>rssE, ecf</i>) - CpxR (<i>yjiA</i>)
CreC-CreB	<i>creB-creC</i>	CreC (<i>phoM</i>) - CreB (<i>yjjE</i>)
CusS-CusR	<i>cusR-cusS</i>	CusR (<i>ylcA</i>) - CusS (<i>ybcZ</i>)
DcuS-DcuR	<i>dcuS-dcuR</i>	DcuS (<i>yjdH</i>) - DcuR (<i>yjdG</i>)
DpiB-DpiA	<i>dpiB-dpiA</i>	CitB/CitA
EnvZ-OmpR	<i>ompR-envZ</i>	
GlrK-GlrR	Codés séparément	GlrK (<i>yfhK, qseE</i>) GlrR (<i>yfhA, qseF</i>)
KdpD-KdpE	<i>kdpD-kdpE</i>	
NarQ-NarP	Codés séparément	
NarX-NarL	<i>narX-narL</i>	
NtrB-NtrC	<i>glnL-glnG</i>	NtrB (<i>glnR</i>) – NtrC (<i>glnG</i>)
PhoR-PhoB	<i>phoB-phoR</i>	
PhoQ-PhoP	<i>phoP-phoQ</i>	
QseC-QseB	<i>qseB-qseC</i>	QseC (<i>ygiY</i>) – QseB (<i>ygiX</i>)
RstB-RstA	<i>rstA-rstB</i>	RstB (<i>uspT</i>) – RstA (<i>urpT</i>)
UhpB-UhpA	<i>uhpA-uhpB</i>	
YedV-YedW	<i>yedW-yedV</i>	
YehU-YehT	<i>yehU-yehT</i>	
YpdA-YpdB	<i>ypdA-ypdB</i>	
ZraS-ZraR	<i>zraS-zraR</i>	ZraS (<i>hydH</i>) – ZraR (<i>hydG</i>)
Phosphorelais complexes	Organisation génétique	Autres nomenclatures
ArcA-ArcB	Codés séparément	
BarA-UvrY	Codés séparément	UvrY (<i>yecB</i>)
EvgS-EvgA	<i>evgA-evgS</i>	
TorS-TorR	Codés séparément	TorS (<i>yccI</i>)
RcsC-RcsD-RcsB	Codés séparément	RcsD (<i>yojN</i>)
Régulateurs orphelins	Organisation génétique	Autres nomenclatures
FimZ	Codé seul	<i>ybcA</i>
RssB	<i>rssA-rssB</i>	

* L'histidine kinase CheA peut phosphoryler CheY et CheB

Contrairement à la centaine de systèmes à deux composants identifiée chez les procaryotes, seul un nombre très limité de ces éléments a été identifié chez les eucaryotes. Par exemple, dans le génome de la levure de boulanger *Saccharomyces cerevisiae*, seul un système de phosphorelais impliqué dans l'osmorégulation a été mis en évidence (SLN1-YPD1-SSK1, SKN7). Les kinases

eucaryotes sont relativement faciles à identifier du fait de leurs homologues de séquence, leur substrats sont quant à eux bien plus difficiles à trouver par approche globale (Laub and Goulian 2007). Par ailleurs, les voies de transduction du signal impliquant des kinases eucaryotes appartiennent à des voies de transduction différentes du type « systèmes à deux composants » où cette fois la phosphorylation se fait sur des résidus tyrosine ou thréonine (ex : MAP kinases).

On peut mentionner qu'en 2015 un groupe de recherche a mis au point un algorithme (MetaPred2CS) permettant de prédire les systèmes à deux composants bactériens et plus précisément les paires « histidine kinase – régulateur », basé sur les homologues de séquence entre les différents systèmes à deux composants bactériens (Kara, Vickers et al. 2015). Différentes techniques expérimentales telles que le profil de phospho-transfert ou le crible double hybride sont possibles pour ces approches mais sont relativement chronophages. Aussi, une approche bioinformatique peut être très utile, notamment chez les organismes encore peu connus.

Les systèmes à deux composants sont les circuits de signalisation cellulaire prévalents chez les bactéries. Ces systèmes sont impliqués notamment dans les réponses adaptatives au stress. Ces réponses incluent le contrôle de la virulence, du métabolisme énergétique, de l'acquisition des nutriments mais aussi l'adaptation aux changements de conditions environnementales.

1.2.2. Mode de transduction du signal

Le « phosphorelais » peut être décomposé en trois étapes (Stock, Robinson et al. 2000).

1. Autophosphorylation: $HK-His+ATP \leftrightarrow HK-His\sim P+ADP$
2. Phosphotransfert: $HK-His\sim P+RR-Asp \leftrightarrow HK-His+RR-Asp\sim P$
3. Déphosphorylation: $RR-Asp\sim P+H_2O \leftrightarrow RR-Asp+P_i$

La première étape est le transfert du groupement γ -phosphoryl de l'adénosine triphosphate (ATP) à la chaîne latérale de l'histidine conservée de la protéine HK. Ensuite, la protéine HK catalyse le transfert de ce groupe phosphoryl du résidu histidine à la chaîne latérale d'un résidu aspartate conservé de la protéine RR situé dans son domaine receveur. La dernière étape de déphosphorylation de la protéine RR libère une molécule de phosphate inorganique (P_i). Ces trois réactions requièrent la présence d'ions métalliques divalents tels que le Mg^{2+} . Il est intéressant de

noter que, dans de nombreux cas, la protéine HK est bifonctionnelle et peut donc à la fois catalyser la phosphorylation et la déphosphorylation de la protéine régulatrice.

Dans de rares cas, le transfert du groupement phosphate ne se fait pas directement de l'histidine senseur à l'aspartate du régulateur mais passe par l'intermédiaire d'une ou plusieurs protéines appelées phosphorelais ou histidine phosphotransférase (HPT). Ces éléments additionnels constituent ainsi autant de points de contrôle et permettent également de moduler la réponse du système mais aussi de l'amplifier. Le système CheB-CheA impliqué dans la chimiotaxie en est un bon exemple (Kirby 2009). En fonction de la disponibilité en nutriments, la méthylation des chimiorécepteurs membranaires est contrôlée par les protéines CheB (méthylestérase) et CheR (méthyltransférase). La variation de leur état de méthylation est le signal inducteur pour la protéine HK CheA qui, via CheW, permet au régulateur CheY d'être phosphorylé. Ce dernier interagit avec la machinerie du flagelle permettant ainsi une modification de la trajectoire de la bactérie. Le système Rsc est également un phosphorelais très intéressant et sera détaillé dans le chapitre 1.3.4.2.

Dans la plupart des cas c'est la protéine régulatrice sous sa forme phosphorylée qui est responsable du contrôle. Néanmoins on peut citer quelques exceptions intéressantes où les promoteurs cibles peuvent être fixés par la forme non phosphorylée du régulateur. Par exemple, chez *Salmonella enterica* (*S. enterica*), le système SsrA-SsrB est responsable de l'expression de systèmes de sécrétion mais aussi de facteurs de virulence nécessaires à la survie de la bactérie dans les macrophages et ce grâce à la phosphorylation de SsrB. En revanche il a été montré que cette phosphorylation n'était pas requise pour l'activation de *csgD* (Desai, Winardhi et al. 2016). En effet, SsrB a la capacité d'activer *csgD* en levant l'inhibition exercée par le facteur de transcription H-NS sur ce gène. Par ailleurs, chez le même organisme, l'accumulation de la forme non phosphorylée de RcsB appartenant au phosphorelais Rcs a pour effet de promouvoir la formation de biofilm (Latasa, Garcia et al. 2012). Les auteurs ont pu montrer que la forme non phosphorylée de RcsB était requise pour l'expression de *csgD*. Lorsque RcsB est majoritairement sous sa forme phosphorylée, *csgD* n'est pas exprimé. La régulation dans ce cas est indirecte : elle passe par le petit ARN RprA. L'expression de celui-ci est activée par RcsB et par la suite, RprA inhibe post-transcriptionnellement *csgD*. Un autre exemple de choix de l'effet d'un RR sous forme non phosphorylée est le système DegS-DegU de *B. subtilis*. DegU possède deux conformations actives : phosphorylée pour promouvoir la synthèse d'enzymes de dégradation ou non

phosphorylée lorsqu'il s'agit d'activer la compétence bactérienne (Dahl, Msadek et al. 1992). Enfin, le dernier exemple intéressant est celui des régulateurs orphelins. Chez *Helicobacter pylori*, le régulateur HP1043 est essentiel à la croissance de la bactérie et n'est associé à aucune kinase (Beier and Frank 2000). Lorsque deux monomères non phosphorylés d'HP1043 s'associent cela mime la forme phosphorylée d'un régulateur de la sous-famille OmpR/PhoP permettant la fixation au promoteur du gène cible *tlpB* (Hong, Lee et al. 2007).

1.2.3. Communication entre les systèmes à deux composants

De manière générale les couples senseur-régulateur sont très spécifiques malgré la forte conservation entre les différentes histidines kinases et entre les divers régulateurs de réponse. Ceci est sans doute lié à la capacité de nombreuses histidines kinases de phosphoryler mais également de déphosphoryler leur régulateur associé (Stock, Robinson et al. 2000). Ainsi, en l'absence de stimulus, l'activité phosphatase de l'histidine kinase permet de maintenir le système à deux composants dans un mode « off » et donc d'éliminer toute phosphorylation parasite de son régulateur que ce soit via une autre histidine kinase ou via l'acétylphosphate présent dans la cellule.

Néanmoins, on peut noter qu'*in vitro*, il a été montré que de nombreuses kinases étaient capables de phosphoryler un régulateur autre que celui qui lui est associé *in vivo* (Skerker, Prasol et al. 2005). Les auteurs ont pour cela utilisé la technique de «phosphotransfer profiling» qui consiste à incuber la portion cytoplasmique de l'histidine kinase considérée avec de l'ATP radiomarqué afin de pouvoir suivre l'autophosphorylation et le transfert du groupement phosphate. Dans un second temps, la kinase phosphorylée est incubée seule ou avec chacun des différents régulateurs de réponse présents dans le génome de l'organisme d'étude. Après migration sur gel et autoradiographie les phosphotransferts peuvent être visualisés (figure 18a). En absence de régulateur de réponse une simple bande correspondant à l'histidine kinase phosphorylée est présente. En revanche s'il y a eu phosphotransfert, on observe une bande correspondant au régulateur de réponse et/ou la disparition de la bande spécifique de la kinase. Dans cette expérience, EnvZ peut phosphoryler 15 régulateurs autres qu'OmpR si le temps d'incubation est de 60 min (figure 18b). En revanche, pour un temps d'incubation très court, entre la kinase et le régulateur purifiés, le seul phospho-transfert efficace est celui observé avec OmpR (Skerker, Prasol et al. 2005).

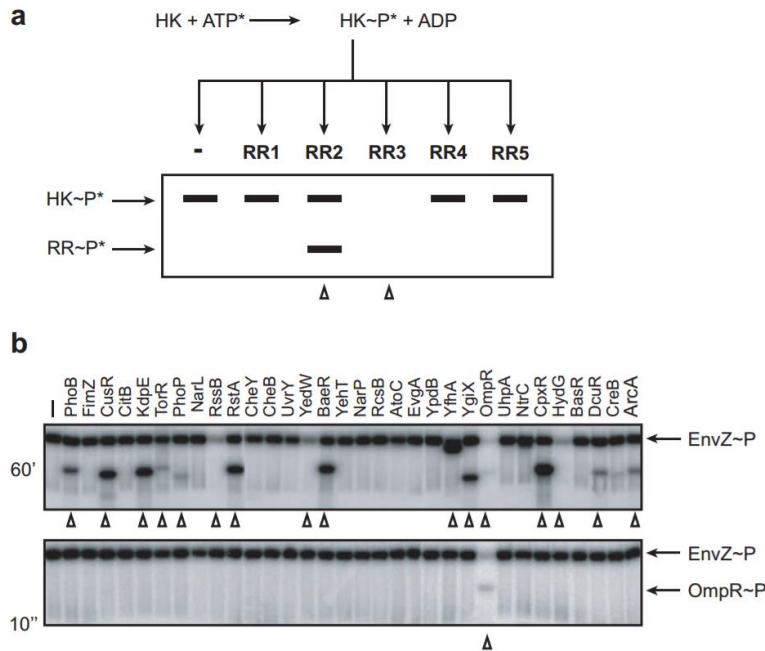


Figure 18 : Cinétique de phospho-transfert de la kinase EnvZ et des différents régulateurs de réponse de la bactérie *E. coli*. a. Schéma général de la technique de phosphotransfert. b. Résultats du phosphotransfert entre EnvZ et les 31 régulateurs de réponse d’*E. coli* (+CheY) pour des incubations de 60 min ou de 10 sec. Les triangles indiquent lorsqu’il y a eu phospho-transfert. Issue de (Skerker, Prasol et al. 2005; Laub and Goulian 2007).

Ces mêmes auteurs sont allés plus loin et ont pu démontrer que la spécificité reposait sur seulement quelques résidus présents dans la région C-terminale et qui, une fois mutés, modifiaient totalement la spécificité de la kinase (Skerker, Perchuk et al. 2008). Ils ont par exemple montré qu’en mutant les résidus T250, L254 et A255 d’EnvZ respectivement en V250, Y254 et R255 qui sont les résidus retrouvés sur RstB (système RstB-RstA), EnvZ devenait spécifique de RstA et non plus d’OmpR.

De façon plus globale, des travaux de Yamamoto *et al.* (Yamamoto, Hirao et al. 2005) ont pu montrer qu’en réalisant des tests d’associations *in vitro* entre les différentes kinases et régulateurs de réponse de la bactérie *E. coli* on pouvait identifier 21 « cross-talk » possibles c’est-à-dire qu’une kinase phosphoryle un autre régulateur que celui qui lui est classiquement associé.

Ce terme de « cross-talk » a été précisé par Mark Goulian et Michael Laub dans leur revue sur la spécificité des systèmes à deux composants (Laub and Goulian 2007). Ils indiquent qu’il est nécessaire d’avoir un contexte précis *in vivo* pour que des « cross-talk » émergent. Aussi l’absence de la kinase homologue est un des rares cas favorables à un cross-talk. De plus, ils précisent que ces interactions sont souvent délétères pour la cellule et on parle de « cross-talk » par opposition à la « cross-régulation » situation dans laquelle ces échanges sont plutôt bénéfiques (ex : système

NarX-NarL/ NarQ-NarP). Précisons que ces réactions de cross-talk ou cross-régulation sont différentes des voies branchées. Dans ce cas, ces systèmes font intervenir plusieurs régulateurs en réponse à une seule et même histidine kinase (comme le système Che d'*E. coli* où les régulateurs CheB et CheY répondent à une seule et même histidine kinase CheA) ou à l'inverse plusieurs histidines kinases peuvent activer un seul et même régulateur. Ainsi, le système Spo0A de *B. subtilis* où les kinases KinA et KinD peuvent activer le même régulateur SpoF, est un autre exemple de voie branchée. Dans ce type de voie de transduction, la délétion d'un des composés impacte la réponse au stimulus. Dans le cas des cross-régulations, où les deux voies agissent indépendamment, la délétion de l'une d'elles ne doit pas impacter l'autre. La figure 19 résume les différentes situations rencontrées.

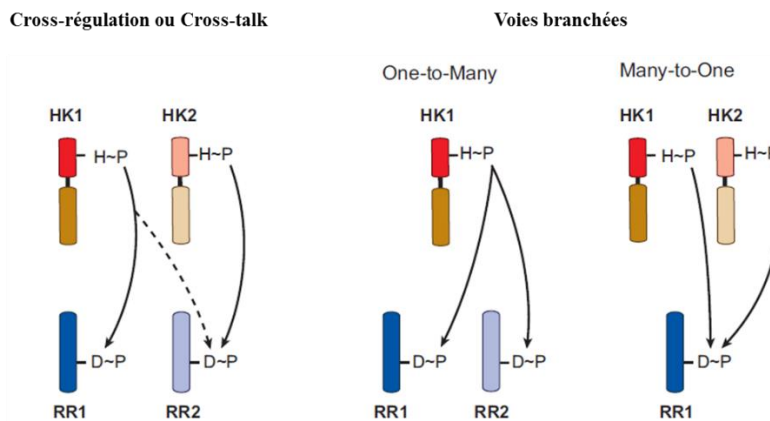


Figure 19 : Comparaison entre cross-régulation ou cross-talk et systèmes branchés. Dans le premier cas, les deux voies de transduction sont indépendantes mais il peut y avoir une kinase (HK1) qui active plusieurs régulateurs (RR1 et RR2) ou à l'inverse un régulateur (RR2) activé par plusieurs kinases (HK1 et HK2). Les pointillés indiquent qu'il y a en général une phosphorylation plus favorable du régulateur classiquement associé. S'il s'agit d'une interaction délétère pour la cellule on parle de cross-talk, si elle est bénéfique on parlera plutôt de cross-régulation. Dans le cas des voies branchées il y a une interdépendance des deux voies qui sont donc complémentaires. D'après (Laub and Goulian 2007).

Si l'on s'intéresse aux conditions permettant au cross-talk d'émerger on peut citer l'exemple des systèmes CpxA-CpxR et EnvZ-OmpR où il a été montré que non seulement l'absence de la kinase dédiée était requise mais aussi celle du régulateur croisé. En d'autres termes pour qu'EnvZ phosphoryle CpxR il est requis à la fois l'absence de la kinase CpxA mais aussi celle du régulateur OmpR (Siryaporn and Goulian 2008). Ces données suggèrent que la présence du régulateur associé à la kinase limite le cross-talk *in vivo* probablement parce que la kinase a une affinité plus importante pour son régulateur associé que pour un autre régulateur. Cet effet est amplifié par la stœchiométrie observée entre le senseur et le régulateur, à savoir une faible abondance des

molécules d'histidine kinase en comparaison de leur régulateur associé. En effet, chez *E. coli*, on compte environ 100 molécules d'EnvZ et 3500 molécules d'OmpR par exemple (Cai and Inouye 2002). Cette stœchiométrie est sans doute facilitée par le fait qu'*ompR* soit le premier cistron de l'opéron *ompR-envZ*, et aussi par la présence d'un SD sous-optimal pour le gène *envZ*. Il est intéressant de préciser que cette organisation est très souvent retrouvée dans les systèmes à deux composants et de ce fait est sans doute conservée pour limiter les cross-talk (Laub and Goulian 2007).

Malgré la rareté du phénomène on peut citer quelques exemples de régulations croisées *in vivo* dans un contexte sauvage qui sont donc bénéfiques pour la cellule. En effet, lorsque la bactérie *E. coli* se trouve en condition anaérobie, elle peut utiliser les nitrates ou les nitrites pour remplacer l'oxygène comme accepteur final d'électrons de la chaîne respiratoire. Les systèmes NarX-NarL et NarQ-NarP contrôlent les différents gènes impliqués dans ce processus. Une explication plus détaillée de ces systèmes se trouve par ailleurs dans le chapitre 1.2.6. Néanmoins on peut retenir que NarQ est capable de phosphoryler NarL en plus de son régulateur associé NarP et NarX peut non seulement phosphoryler NarL mais aussi NarP. Ceci permet à la cellule une régulation fine des différentes enzymes nécessaires à la respiration sur des nitrates en fonction de leur concentration extracellulaire (Noriega, Lin et al. 2010).

Un autre exemple intéressant est la cross-régulation qui existe entre le système EnvZ-OmpR et le système ArcB-ArcA. EnvZ-OmpR comme nous le verrons dans le chapitre 1.2.5. gouverne l'expression de nombreux gènes en réponse à l'osmolarité extracellulaire et notamment les porines majoritaires chez *E. coli* OmpC et OmpF (Cai and Inouye 2002). ArcB-ArcA module quant à lui les gènes impliqués dans la transition entre croissance aérobie et anaérobie. Il a été montré qu'en condition d'anaérobiose le profil d'expression des porines OmpC et OmpF était modifié à savoir que la quantité d'OmpC était augmentée et celle d'OmpF diminuée ce qui ressemble au profil d'expression obtenu en condition de haute osmolarité autrement dit quand le système EnvZ-OmpR est activé (Matsubara, Kitaoka et al. 2000). Les auteurs ont pu montrer que la kinase ArcB était capable dans les conditions d'anaérobiose d'activer OmpR créant ainsi une régulation croisée entre ces deux systèmes dans des conditions spécifiques mais physiologiques.

1.2.4. Pléiotropie des systèmes à deux composants et diversité des signaux

Grâce à l'essor du séquençage et l'arrivée des approches globales, des études ont pu montrer la variété des gènes régulés par les systèmes à deux composants. En 2002, une étude transcriptomique réalisée sur 36 mutants d'*E. coli* K12, dépourvus d'une protéine HK et/ou d'une protéine RR provenant d'un système à deux composants, a permis de mettre en évidence quatre mutants affectant significativement l'expression de plus de 100 gènes (Oshima, Aiba et al. 2002) : *ΔarcA*, *ΔarcB*, *ΔompR-envZ* et *ΔuvrY*. Chez 19 autres mutants, une modification de la transcription a été détectée affectant 21 à 100 gènes, sept autres mutants affectent l'expression de moins de 10 gènes. Cette étude montre clairement l'impact des systèmes à deux composants dans la régulation des gènes bactériens même si certains des effets observés sont bien sûr indirects. Les processus touchés dans plusieurs mutants sont la synthèse du flagelle, l'expression du régulon RpoS (facteur sigma majeur de la phase stationnaire de croissance), le système de transport du maltose et enfin la synthèse d'entérobactine, une protéine impliquée dans l'import du fer dans la cellule. Il est important de mentionner que les résultats dépendent des conditions expérimentales utilisées (37°C, en milieu riche et sous atmosphère aérobie) suggérant ainsi un rôle probablement plus important des systèmes à deux composants.

Il est intéressant de préciser que nombre de systèmes à deux composants sont impliqués dans la virulence bactérienne faisant ainsi de ces éléments une cible de choix pour éliminer les pathogènes (Barrett and Hoch 1998). Ces systèmes confèrent dans certains cas une résistance à un antibiotique comme par exemple le système VanS-VanR d'*Enterococcus faecalis* ou peuvent activer directement la transcription de gènes associés à la capacité d'invasion du pathogène comme c'est le cas pour PhoQ-PhoP mais aussi d'EnvZ-OmpR. En effet, EnvZ-OmpR est directement impliqué dans la modulation des gènes *vir* de *Shigella flexneri* (Bernardini, Fontaine et al. 1990). PhoQ-PhoP intervient quant à lui dans l'homéostasie du magnésium et la virulence notamment chez *Salmonella* (Groisman, 2001). Il a en effet été démontré qu'il était capable de répondre à la présence de peptides antimicrobiens mais il régule aussi *mgtC*, un facteur de virulence essentiel à la survie de *S. enterica* dans les macrophages (Blanc-Potard and Groisman 1997).

Bien qu'ayant de nombreux effets, les délétions des systèmes à deux composants ne sont pas dans la majorité des cas létales. On peut cependant citer une exception intéressante : le système à deux composants WalK-WalR. Ce système est essentiel chez *Staphylococcus aureus* (*S. aureus*) et est

impliqué dans le métabolisme de la membrane mais aussi dans le processus de formation des biofilms (Dubrac et al., 2007).

La singularité des systèmes à deux composants réside non seulement dans le fait que les fonctions régulées soient très diverses mais également dans le fait que l'activation des histidines kinases peut se faire via de très nombreux signaux y compris des paramètres physiques (température, pH, lumière..), chimiques (détection des nitrates, des métaux tels que le cuivre ...) ou encore en réponse à des peptides.

Il a par exemple été montré que chez *B. subtilis* la perception du froid se faisait notamment via le système à deux composants DesK-DesR percevant la fluidité membranaire (Mansilla, Cybulski et al. 2004). La perception de la lumière chez *Caulobacter crescentus* se fait également via un système à deux composants : LovK-LovR (Purcell and Crosson 2008). Le signal perçu par LovK à savoir la lumière bleue permettrait à cet organisme vivant essentiellement dans les lacs et rivières de se situer par rapport à la colonne d'eau, le renseignant indirectement sur la disponibilité en nutriments. En effet, la quantité de nutriments présente dans les sédiments du fond d'un lac est plus importante qu'en surface de celui-ci.

Pour finir, les systèmes à deux composants sont également très importants dans la communication des populations bactériennes via le « quorum sensing ». Ce mode de signalisation typiquement bactérien repose sur la production d'auto-inducteurs (peptides, homosérine lactones acyclées ...) permettant la synchronisation du comportement de toute la population en réponse aux changements environnementaux. Par exemple, la production du peptide auto-inducteur AIP (Auto-Inducing Peptide) chez *S. aureus* est directement sous le contrôle du système à deux composants AgrC-AgrA (Novick and Geisinger 2008).

En ce qui concerne le mécanisme de perception des signaux, trois mécanismes ont été décrits (Mascher, Helmann et al. 2006) et sont présentés dans la figure 20. La majorité des histidines kinases perçoivent des signaux extracellulaires ou périplasmiques (chez les bactéries à Gram négatif). Dans ce cas un large domaine extracellulaire senseur est encadré d'au moins deux hélices transmembranaires puis d'un domaine de transmission cytoplasmique (ex: NarX-NarL). Les domaines « senseur » et « transmetteur » se situent donc dans des compartiments cellulaires distincts. Cette topologie est généralement retrouvée chez les systèmes capables de sentir des variations en solutés ou nutriments. Le second groupe d'histidines kinases perçoit le signal grâce

aux hélices transmembranaires (entre 2 et 20 connectées par de courtes boucles) seules ou par une combinaison entre les hélices membranaires et le segment N-terminal cytoplasmique (ex : DesK-DesR). On parle dans ce cas de signal intra-membranaire. Dans un certain nombre de cas des protéines accessoires aident à la perception du signal. Enfin, le dernier groupe d'histidines kinases est constitué par les kinases percevant les signaux cytoplasmiques et on les retrouve soit solubles (ex : CheA) dans le cytoplasme ou ancrées à la membrane (ex : KdpD-KdpE, ArcB-ArcA). Dans tous les cas, une fois le signal perçu, la transduction du signal se fait de la même façon.

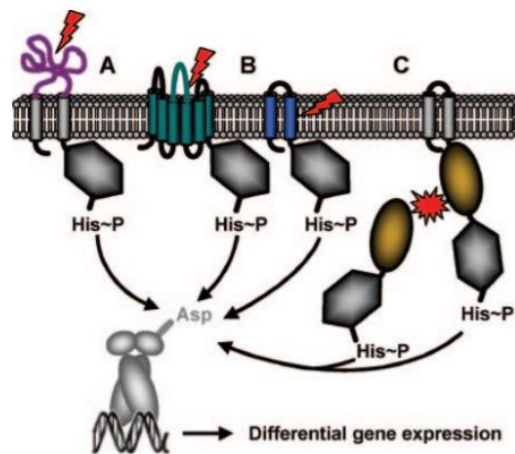


Figure 20 : Représentation schématique des trois mécanismes de perception du signal par les histidines kinases. (A) Senseur extracytoplasmique (B) Senseur intra-membranaire (C) senseur cytoplasmique. Issue de (Mascher, Helmann et al. 2006).

Les systèmes à deux composants répondent donc à de nombreux types de signaux et régulent une grande variété de gènes. Dans la suite de ce manuscrit, je vais revenir en détails sur les systèmes à deux composants qui m'ont principalement intéressée durant ma thèse : EnvZ-OmpR et NarX-NarL/NarQ-NarP. Nous verrons que chacun à leur manière ces systèmes présentent des propriétés particulières.

1.2.5. Le système EnvZ-OmpR : un système classique avec des propriétés inédites

Le système à deux composants EnvZ-OmpR, l'un des plus étudiés chez *E. coli*, est un système à deux composants classique. EnvZ-OmpR a été mis en évidence pour la première fois lors d'études génétiques qui ont permis d'isoler des mutants affectés dans l'expression des porines majeures d'*E. coli* que sont OmpC et OmpF (Sarma and Reeves 1977). Les porines OmpC et OmpF ne sont pas exprimées au même niveau selon les conditions d'osmolarité. En effet, à faible osmolarité la porine OmpF est prédominante alors qu'à forte osmolarité c'est la porine OmpC. Il est

communément admis que la régulation de ces porines a une importance majeure en ce qui concerne la diffusion passive des molécules hydrophobes et chargées à travers la membrane. Ainsi les pores plus larges d'OmpF (1,16 nm) permettraient d'augmenter la diffusion des nutriments dans la cellule à faible osmolarité alors que les pores plus petits d'OmpC (1,09 nm) limiteraient leur entrée. La clé de la régulation de ces deux porines se situe au locus d'*ompC* et *ompF* mais aussi à celui de l'opéron *ompB*. Des expériences de fusions entre *ompC* et *lacZ* et entre *ompF* et *lacZ* ont montré en effet que des mutations dans le locus *ompB* avaient un impact sur l'activité β -galactosidase de ces fusions (Hall and Silhavy 1981). Par la suite, des expériences de complémentation et de cartographie structurale ont suggéré que le locus *ompB* était en fait constitué de deux gènes, *ompR* et *envZ*. Cette hypothèse a été confirmée en 1982 par le séquençage de l'opéron *ompB* d'*E. coli* où deux cadres ouverts de lecture ont été mis en évidence (Wurtzel, Chou et al. 1982). Sur la base des analyses de séquences, il a été prédit que le gène *ompR* codait une protéine d'environ 29 kDa et qu'*envZ* codait une protéine d'environ 50 kDa. Les expériences de complémentation ont montré que ces deux gènes étaient co-transcrits avec *ompR* en amont d'*envZ* (Taylor, Hall et al. 1983).

1.2.5.1. La protéine kinase senseur : EnvZ

1.2.5.1.1. Caractéristiques d'EnvZ

La protéine de membrane interne EnvZ, composée de 450 acides aminés, est présente durant la phase exponentielle de croissance à environ 100 copies par cellule (Cai and Inouye 2002). A la suite de nombreux travaux, EnvZ a été définie comme une protéine senseur de type kinase à histidine.

Le domaine C-terminal de la protéine EnvZ possède une forte homologie de séquence avec les autres kinases senseurs à histidines retrouvées dans les systèmes à deux composants (Nixon, Ronson et al. 1986). Ce domaine conservé a la capacité de s'autophosphoryler et de catalyser la phosphorylation de la protéine RR associée, OmpR. De plus, EnvZ possède une activité phosphatase permettant de déphosphoryler OmpR.

L'activité d'autophosphorylation a été montrée pour la première fois lors d'expériences utilisant la forme tronquée d'EnvZ, EnvZ115, qui pouvait être phosphorylée *in vitro* en présence d'ATP (Igo and Silhavy 1988). En 1989, Delgado *et al.* ont pu montrer que l'activité était localisée dans le domaine C-terminal (Forst, Delgado et al. 1989). L'autophosphorylation intervient plus particulièrement au niveau du résidu histidine His-243. Ce résidu est très conservé dans tous les

senseurs des systèmes à deux composants. Une mutation du résidu His-243 en valine empêche toute autophosphorylation d'EnvZ (Forst, Delgado et al. 1989). L'utilisation de mutants d'*envZ* a par la suite permis de montrer que l'autophosphorylation d'EnvZ avait lieu en *trans* (phosphorylation effectuée entre monomères où un monomère catalyse la phosphorylation du second) et donc qu'une dimérisation d'EnvZ était nécessaire pour cette activité (Cai and Inouye 2003).

A la suite de son autophosphorylation, EnvZ est capable de transférer son groupement phosphate de l'His-243 au résidu aspartate d'OmpR. C'est le domaine cytoplasmique d'EnvZ qui est requis pour ce transfert de phosphate (Forst, Delgado et al. 1989).

Enfin, en présence d'ATP, EnvZ peut également déphosphoryler la forme phosphorylée d'OmpR, OmpR-P. La plupart des substitutions du résidu His-243 n'abolissent pas complètement l'activité de déphosphorylation. Aussi, une hypothèse serait que ce résidu ne soit pas directement impliqué dans l'activité phosphatase mais plutôt dans le maintien de la conformation d'EnvZ active pour déphosphoryler OmpR-P (Hsing and Silhavy 1997).

La quantité de forme phosphorylée d'OmpR résulte de la balance des fonctions kinase et phosphatase d'EnvZ (Pratt and Silhavy 1995). A faible osmolarité, le faible niveau d'OmpR-P résulte d'une activité phosphatase plus élevée que l'activité kinase. A l'inverse, à forte osmolarité, la quantité d'OmpR-P est plus importante car c'est l'activité kinase qui est prédominante. Le signal inducteur précis d'EnvZ étant encore inconnu, il est difficile de déterminer avec exactitude l'impact du stimulus sur le ratio activité kinase/phosphatase. Néanmoins, l'analyse de divers mutants d'EnvZ a permis de proposer deux modèles expliquant la variation des niveaux d'OmpR-P. Le modèle soutenu par le laboratoire d'Inouye propose que l'activité kinase d'EnvZ soit constante quelle que soit l'osmolarité du milieu. Le niveau de phosphorylation de OmpR serait alors complètement dépendant de l'activité phosphatase qui seule répondrait aux variations d'osmolarité (Utsumi, Brissette et al. 1989; Jin and Inouye 1993; Zhu, Qin et al. 2000). Le modèle du groupe de Thomas Silhavy suggère quant à lui une liaison étroite entre les sites responsables de ces deux activités enzymatiques. En effet, plusieurs séries de mutations isolées dans des régions conservées d'EnvZ affectent les deux activités (Hsing, Russo et al. 1998). De plus, des mutants de l'histidine 243 ont logiquement non seulement perdu leur capacité d'autophosphorylation et de transfert du groupement phosphate mais ils ont également une activité phosphatase diminuée voire

nulle (Hsing and Silhavy 1997). Ainsi, le modèle prévoit que l'activité auto-kinase d'EnvZ augmente avec le niveau d'osmolarité. A forte osmolarité, il y aurait donc plus de forme phosphorylée d'EnvZ (EnvZ-P). L'activité kinase serait assurée uniquement par EnvZ-P et l'activité phosphatase par la forme non phosphorylée d'EnvZ. A faible osmolarité, EnvZ serait majoritairement présente sous forme non phosphorylée et agirait comme une phosphatase expliquant ainsi la faible concentration d'OmpR-P. A l'inverse, à forte osmolarité, EnvZ-P serait majoritaire et donc l'activité kinase dominerait d'où l'accumulation de la forme phosphorylée d'OmpR. La figure 21 compare ces différents modèles. De nombreuses études sont encore requises pour valider l'un ou l'autre de ces deux modèles.

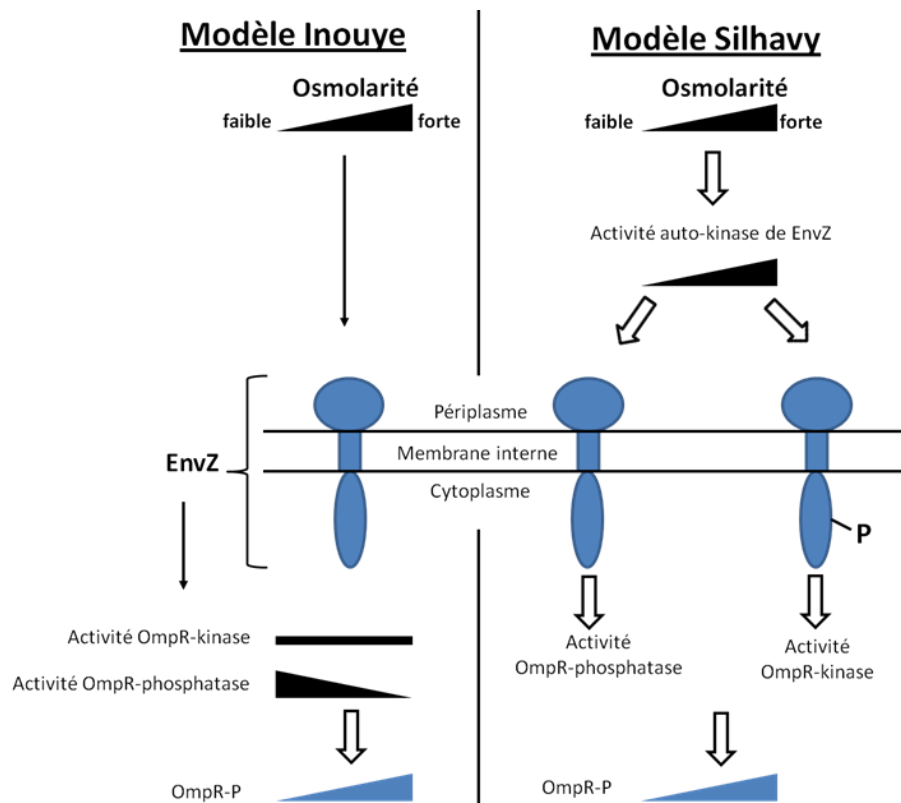


Figure 21 : Comparaison des modèles de régulation du système EnvZ-OmpR.

1.2.5.1.2. Structure cristallographique d'EnvZ

A ce jour, la structure complète d'EnvZ n'a pas été résolue, notamment en raison du défi que représente la cristallisation d'une protéine transmembranaire. Néanmoins, les deux domaines fonctionnels composant la partie cytoplasmique d'EnvZ ont été résolus séparément (figure 22).

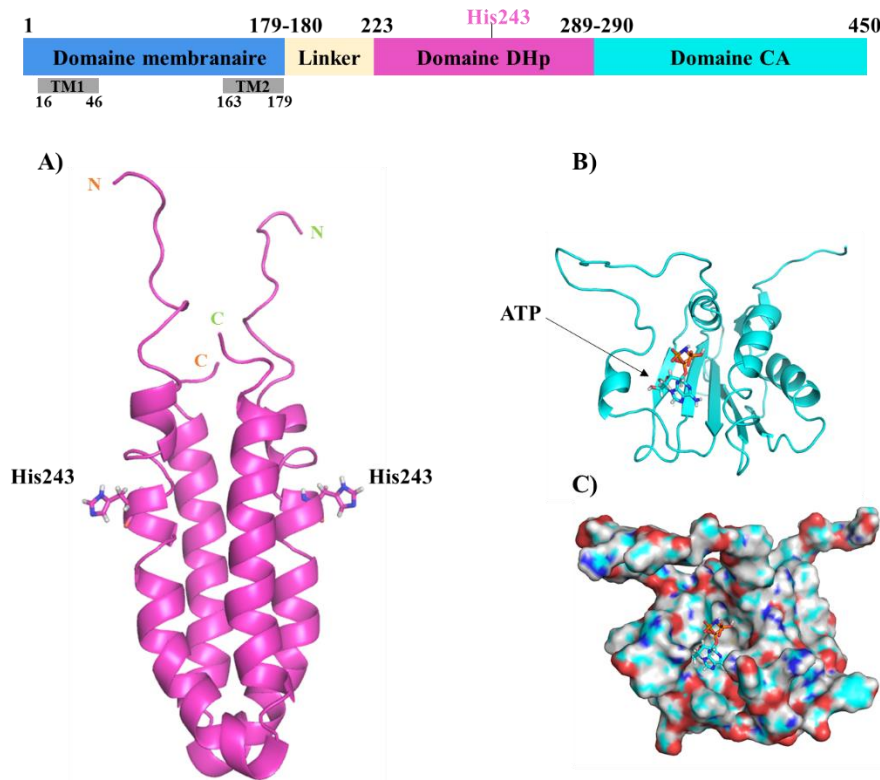


Figure 22 : Structure primaire d'EnvZ et représentation des deux domaines d'EnvZ effectuée avec Pymol à partir des structures des domaines DHp (1JOY) et CA (1BXD) cytoplasmiques d'EnvZ disponibles dans la base de données PDB. La représentation linéaire mentionne les différents domaines et précise l'emplacement des boucles périplasmique (TM1 et TM2). A) Domaine DHp d'EnvZ sous forme d'homodimère B) Domaine CA d'EnvZ en représentation « cartoon » ou C) « surface » pour visualiser la poche où se loge l'ATP.

Le domaine DHp pour « Dimerization and Histidine phosphotransfer domain » (résidus 223-289) contient le résidu histidine 243, le site d'autophosphorylation d'EnvZ ainsi que celui de transfert du phosphate vers le résidu D55 du régulateur OmpR (Tomomori, Tanaka et al. 1999). Le domaine CA pour « Catalytic ATP-binding domain » (résidus 290-450) contient quant à lui différentes régions très conservées permettant la phosphorylation du résidu His243 : c'est le cœur catalytique d'EnvZ (Tanaka, Saha et al. 1998). Les motifs conservés sont annotés N, G1, D, F et G2. La conformation tridimensionnelle de ce domaine de 161 aa permet la formation d'une « poche » dans laquelle est capturée une molécule d'ATP (Dutta, Qin et al. 1999). Sous sa forme dimérisée, le

domaine CA d'un monomère d'EnvZ se retrouve en face du site de phosphorylation du second monomère permettant ainsi l'autophosphorylation d'EnvZ (figure 23).

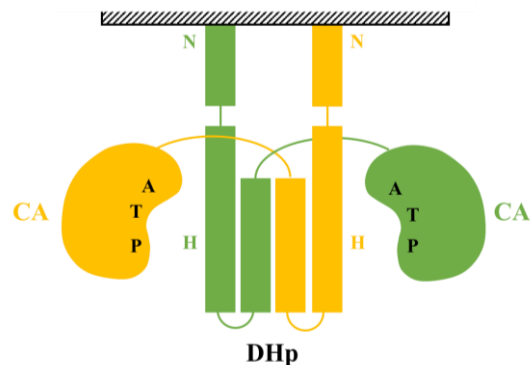


Figure 23 : Modèle d'organisation des dimères d'EnvZ permettant l'autophosphorylation de la kinase. Seule la portion cytoplasmique d'EnvZ est représentée.

1.2.5.1.3. Signal capté par EnvZ

Le signal exact reconnu par EnvZ est encore peu compris. Différentes études ont pu montrer comme nous l'avons précédemment mentionné qu'EnvZ répondait à des variations d'osmolarité ce qui constitue des variations d'un environnement complexe et ne dépend pas de la concentration d'une seule et unique molécule. De plus, différentes études ont pu mettre en évidence d'autres possibilités d'activation du système EnvZ-OmpR.

En accord avec l'osmolarité, la reconstitution *in vitro* du système EnvZ-OmpR grâce à des protéines purifiées a pu montrer que l'ajout de cations et plus précisément d'ion potassium (K⁺) favorisait l'augmentation de la forme phosphorylée d'OmpR via la stimulation de l'activité auto-kinase d'EnvZ (Jung, Hamann et al. 2001).

Deux études ont pu montrer qu'un des signaux perçu par EnvZ-OmpR est l'acidité. Ceci est également compatible avec l'activation à haute osmolarité d'EnvZ puisqu'une solution aqueuse acide contient en général plus de cations (H₃O⁺) que d'anions (OH⁻).

La première étude (Stincone, Daudi et al. 2011) a pu montrer que le système EnvZ-OmpR avait un rôle central dans la réponse au stress acide. Un mutant $\Delta ompR$ a en effet un taux de survie bien moins important qu'une souche sauvage en présence d'un choc acide. De plus, un certain nombre de gènes requis pour l'adaptation au stress acide ne sont plus modulés dans un mutant $\Delta ompR$.

La seconde étude (Quinn, Cameron et al. 2014) a pu montrer qu'un sous-ensemble de cibles d'OmpR était régulé en fonction du pH. Cette régulation est dépendante de la fixation d'OmpR à ses cibles car les niveaux de messenger *ompR* ou de la protéine OmpR entre pH 4,5 et pH 7 sont les mêmes chez *E. coli*. En revanche chez *Salmonella* il y a une augmentation du messenger *ompR* et de la protéine OmpR suggérant des mécanismes de régulation différents chez les deux organismes.

Le signal acidité est très intéressant d'un point de vue physiologique puisqu'il a été montré qu'OmpR possédait un rôle important dans l'infection. En effet, des mutants d'*ompR* chez *Salmonella* sont atténués dans la virulence chez la souris (Dorman, Chatfield et al. 1989). Ce point sera également développé dans la partie 1.4.4. consacrée à la bactérie pathogène LF82 où nous verrons qu'OmpR joue un rôle dans la pathogénie de cette souche d'*E. coli*.

1.2.5.1.4. Mécanisme d'activation d'EnvZ

Pour connaître quel domaine d'EnvZ reconnaissait le signal extérieur, différentes études avec des mutants d'EnvZ ont pu proposer des résultats à ce jour contradictoires.

Tokishita et ses collaborateurs ont montré que des délétions du domaine périplasmique d'EnvZ avaient pour conséquence une perte de la réponse aux variations d'osmolarité et plaçaient EnvZ dans une conformation analogue à celle obtenue à haute osmolarité (Tokishita, Kojima et al. 1991). Cette conformation n'étant pas réversible cela aboutit à l'insensibilité aux variations du milieu extérieur et un phénotype OmpF⁻ OmpC⁺ suggérant l'essentialité de ce domaine dans l'acquisition du signal par EnvZ.

Cependant, un autre groupe montre en 1996 que des délétions du domaine périplasmique permettaient de conserver la régulation des porines OmpF et OmpC en réponse aux variations d'osmolarité (Leonardo and Forst 1996). Il est intéressant de préciser que sur les 3 délétions testées 2 seulement donnent lieu à un phénotype sauvage. De plus, les délétions testées sont différentes des travaux de Tokishita. Néanmoins ils montrent que le remplacement du domaine périplasmique d'EnvZ par celui de PhoR permet de conserver la régulation des porines en fonction de l'osmolarité. Ces travaux iraient donc plutôt dans le sens de l'hypothèse contraire à savoir que la région périplasmique ne serait pas requise pour l'acquisition du signal par EnvZ. Ils proposent comme modèle que la haute osmolarité engendre des perturbations membranaires altérant l'interaction entre les dimères d'EnvZ et ayant pour conséquence l'augmentation de l'activité

kinase et la diminution de l'activité phosphatase. La région périplasmique sera elle requise pour sentir d'autres signaux encore inconnus.

Ces deux études contradictoires utilisent des plasmides portant des portions différentes d'EnvZ aussi la variabilité des résultats peut-être dûe au nombre de copies différent des plasmides mais aussi à une variation de la conformation des mutants qui s'inséreraient plus ou moins correctement dans la membrane.

Parallèlement à l'étude des boucles périplasmiques, différentes études évoquent une implication de la région dite « linker » d'EnvZ pour la transmission du signal d'osmolarité (Park and Inouye 1997). La région « linker », constituée d'environ 50 résidus, relie le deuxième segment transmembranaire d'EnvZ au domaine cytoplasmique de la protéine. Des mutations dans cette région aboutissent à un phénotype OmpC⁺ quel que soit l'osmolarité du milieu. L'ancrage à la membrane serait donc requis pour la transmission du signal via le domaine linker. De plus, des mutations dans le premier domaine transmembranaire d'EnvZ rendent celle-ci insensible aux variations d'osmolarité (Hsing, Russo et al. 1998). Ensemble, ces données suggèrent que le signal perçu a besoin d'être propagé à travers la membrane via les segments transmembranaire puis transmis au domaine cytoplasmique via la région linker. Cependant les mutations de la région linker sur une version d'EnvZ cytoplasmique n'abolissent pas l'osmorégulation ce qui sous-entend que la localisation membranaire ne serait pas forcément requise (Park and Inouye 1997).

Une étude plus récente va aussi dans ce sens et propose que le signal de stress osmotique soit transmis de manière passive au domaine cytoplasmique d'EnvZ. Les deux hélices présentes dans cette région par un mécanisme de contraction/relaxation stimuleraient l'autophosphorylation d'EnvZ et le transfert du phosphate à OmpR (Wang, Morgan et al. 2012). Cette étude se base sur une version d'EnvZ tronquée où seule est présente la région cytoplasmique du domaine C-terminal d'EnvZ. Les différents résultats contradictoires peuvent s'expliquer par l'utilisation de protéines de tailles différentes par rapport à EnvZ sauvage (270 aa au lieu 450 aa) suggérant des différences de conformation possibles. De plus, les mutants d'EnvZ tronquée utilisés sont produits à partir de plasmides et leurs niveaux pourraient être bien supérieurs à ceux observé dans un cadre physiologique.

En conclusion, de nombreuses études ont tenté de répondre à la question du mécanisme de perception du signal par EnvZ. Malheureusement, à ce jour et malgré les diverses hypothèses avancées, notre compréhension de la perception du signal par EnvZ reste incomplète.

1.2.5.2. La protéine régulatrice : OmpR

1.2.5.2.1. Caractéristiques d'OmpR

OmpR est une protéine cytoplasmique de 239 acides aminés. Elle est présente à environ 3500 copies par cellule durant la phase exponentielle de croissance (Cai and Inouye 2002). Deux domaines la composent : le domaine N-terminal (positions 1-124), site de la phosphorylation, et le domaine C-terminal (positions 137-239), site de fixation à l'ADN (Martinez-Hackert and Stock 1997). Ces deux domaines sont reliés par une région appelée également « linker » (positions 125-136).

Hall et Silhavy ont proposé en 1981 que le rôle d'OmpR serait de réguler l'expression d'*ompC* et d'*ompF* au niveau transcriptionnel (Hall and Silhavy 1981). Sur la base de comparaison de séquences, OmpR partage une homologie avec plusieurs autres protéines de fixation à l'ADN comme le répresseur *cro* du phage lambda suggérant qu'OmpR présente aussi cette propriété de liaison à l'ADN (Wurtzel, Chou et al. 1982). Cette hypothèse a été confirmée en 1986 par des expériences d'empreinte à la DNase I. Les auteurs ont ainsi déterminé les sites précis de fixation d'OmpR sur les promoteurs d'*ompC* et d'*ompF* (Norioka, Ramakrishnan et al. 1986). De nombreuses autres études ont pu mettre en évidence le nombre de monomères d'OmpR fixés sur chacun des promoteurs à basse et haute osmolarité (Maeda and Mizuno 1990; Huang, Schieberl et al. 1994; Harlocker, Bergstrom et al. 1995; Huang and Igo 1996).

Il est important de préciser que la phosphorylation d'OmpR par EnvZ est un élément-clé dans la régulation des cibles *ompC* et *ompF*. En effet, la phosphorylation d'OmpR augmente *in vitro* l'affinité de la protéine pour les régions promotrices d'*ompC* et *ompF* d'un facteur 10 (Head, Tardy et al. 1998). En phosphorylant sa protéine régulatrice, EnvZ module ainsi la capacité d'OmpR à se fixer à l'ADN et donc module le niveau de transcription des gènes cibles du système EnvZ-OmpR. Des expériences suggèrent qu'un changement de conformation entre OmpR et OmpR-P serait à l'origine de ce changement d'affinité pour l'ADN (Stock, Robinson et al. 2000). Plus récemment, il a été montré que la phosphorylation induisait la dimérisation d'OmpR (Barbieri, Wu et al. 2013).

La phosphorylation d'OmpR par EnvZ a lieu dans le domaine N-terminal d'OmpR et plus précisément au niveau du résidu aspartate Asp-55 (Delgado, Forst et al. 1993). Des expériences d'empreinte à la DNase I réalisées avec des fragments C-terminaux d'OmpR ont démontré que cette région était quant à elle impliquée dans la liaison à l'ADN et plus précisément était capable de se fixer aux régions régulatrices d'*ompC* et d'*ompF*. Par ailleurs, il a été montré que l'activation de la transcription d'*ompC* nécessitait l'interaction du fragment C-terminal d'OmpR avec la sous unité α de l'ARN polymérase. Il a été proposé que cette interaction stabilise celle qui a lieu entre les promoteurs et l'ARN polymérase ce qui *in fine* augmenterait la transcription des cibles positives d'OmpR (Kondo, Nakagawa et al. 1997).

Enfin, la région « linker » serait très importante dans la communication entre les deux domaines d'OmpR. En effet, des mutations dans cette zone empêchent l'activation du domaine C-terminal (Mattison, Oropeza et al. 2002). Contrairement à la version sauvage d'OmpR la phosphorylation du mutant G129D (dans le linker) ne permet pas l'augmentation de sa fixation au promoteur d'*ompC*. Le domaine C-terminal d'OmpR seul est également incapable d'activer la transcription d'*ompC* (Walthers, Tran et al. 2003). Ces différentes données sont en accord avec le fait que la phosphorylation du domaine N-terminal permet d'activer le changement de conformation du domaine C-terminal afin d'augmenter son affinité pour les promoteurs cibles. Cette activation serait en partie conduite par la région « linker ».

1.2.5.2.2. Structure cristallographique d'OmpR

La structure complète d'OmpR n'a pour le moment pas été résolue. Cristalliser une protéine telle qu'OmpR avec deux domaines reliés par un linker extrêmement flexible s'avère être une tâche très délicate. A ce jour, le domaine de fixation à l'ADN a été cristallisé séparément.

Quant au domaine receveur N-terminal sa structure a été modélisée grâce à celle de CheY. En effet, la structure des domaines receveurs est fortement conservée entre les différents régulateurs de réponse rendant cette stratégie *in silico* possible (figure 24). Elle se représente sous la forme de cinq feuilletts β parallèles entourés par cinq hélices α . Au centre se situe le résidu aspartate 55 phosphorylé par EnvZ.

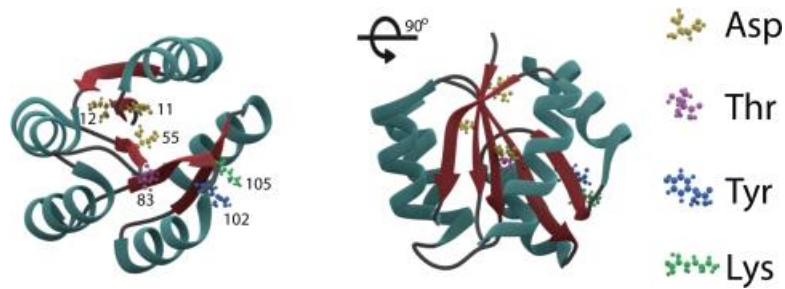


Figure 24 : Structure du domaine N-terminal d'OmpR modélisée à partir de celle de CheY (PDB 4QYW). Les résidus conservés requis pour la phosphorylation sont mis en évidence (aspartate 11/12/55, thréonine 83, lysine 102 et tyrosine 105). Les hélices α sont représentées en vert, les feuillets β en rouge et les boucles en gris. Issue de (Desai and Kenney 2017).

La structure du domaine C-terminal du régulateur OmpR (Martinez-Hackert and Stock 1997) met en évidence un motif de liaison à l'ADN qui a été nommé wHTH pour winged Helix Turn Helix (figure 25). Ce motif est formé de trois hélices α et de plusieurs feuillets β antiparallèles. L'hélice α_3 (résidus 201-212) semble particulièrement importante car c'est elle qui reconnaît l'ADN chez les autres régulateurs de la même famille qu'OmpR. La structure d'OmpR en présence d'un promoteur cible n'a pas été effectuée. Néanmoins des expériences de mutagenèse ont pu mettre en évidence le rôle clé des acides aminés « sérine 200 » et « valine 203 » pour la fixation d'OmpR à l'ADN (Russo, Slauch et al. 1993; Kato, Tsuzuki et al. 1995). Ces deux acides aminés se situant dans l'hélice α_3 cela conforte donc le rôle clé de cette hélice. Plus tard, des expériences de cristallographie proposent que les résidus clé soient en plus de ceux-ci l'arginine 182 et la thréonine 224 situés respectivement dans l'hélice α_3 et dans le feuillet β_6 (Martinez-Hackert and Stock 1997). La boucle reliant les feuillets β_6 et β_7 semble également importante pour la reconnaissance de l'ADN.

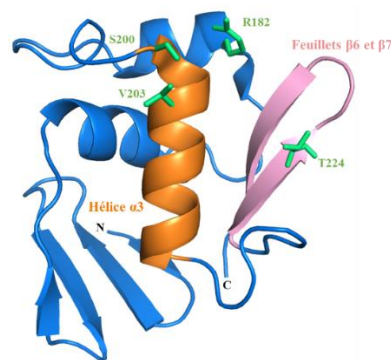


Figure 25 : Représentation du domaine de fixation à l'ADN d'OmpR effectuée avec Pymol à partir de la structure 1OPC du domaine de fixation à l'ADN d'OmpR disponible dans la base de données PDB. L'hélice α_3 et les feuillets β_6 et β_7 requis pour la fixation à l'ADN sont mis en évidence ainsi que les 4 résidus critiques pour cette fixation.

1.2.5.2.3. Activation non classique d'OmpR

Comme nous l'avons vu dans le chapitre 1.2.3., OmpR peut être phosphorylé par d'autres régulateurs qu'EnvZ. A ce jour, deux régulations ont pu être montrées *in vivo* et ce dans des conditions particulières à savoir la phosphorylation d'OmpR par ArcB en condition d'anaérobiose (Matsubara, Kitaoka et al. 2000) et par CpxA en absence d'EnvZ et de CpxR (Siryaporn and Goulian 2008) suggérant qu'il s'agisse d'un phénomène relativement limité. Par ailleurs, il a été démontré qu'*in vivo* OmpR pouvait être phosphorylé par l'acétylphosphate, un métabolite présent dans le cytoplasme de la cellule (Matsubara and Mizuno 1999).

1.2.5.3. Modèle de régulation d'OmpC et OmpF par le système EnvZ-OmpR

La régulation différentielle d'OmpC et OmpF par OmpR est un processus complexe. Plusieurs études génétiques ont été réalisées afin d'identifier des mutants d'OmpR exprimant des phénotypes différents. Le laboratoire de Thomas Silhavy a classé en trois catégories les mutations selon le phénotype obtenu sur la base de l'expression d'*ompC* et d'*ompF* (Hall and Silhavy 1981; Slauch and Silhavy 1989). Nous avons d'ailleurs utilisé ces mutants comme vous le verrez dans la partie « résultats ».

La régulation différentielle serait basée sur un gradient de la forme active d'OmpR (OmpR-P) avec EnvZ comme régulateur de ce gradient. Cette hypothèse a été testée par Russo et Silhavy qui ont développé un modèle permettant de décrire la relation entre les concentrations de OmpR-P et l'expression des porines et ce en examinant l'effet de mutations d'EnvZ sur l'expression d'*ompC* et d'*ompF* (Russo and Silhavy 1991). Pour cela, ils ont utilisé trois mutants : *envZ::Tn10*, *envZ473* et *env247* qui, respectivement, n'ont plus d'activité kinase ni phosphatase, plus d'activité phosphatase ou plus d'activité kinase. A faible concentration en OmpR-P, donc à faible osmolarité, l'expression d'*ompF* est favorisée. A l'inverse, lorsque la concentration en OmpR-P est forte, donc à haute osmolarité, l'expression d'*ompC* est activée et celle d'*ompF* réprimée (figure 26).

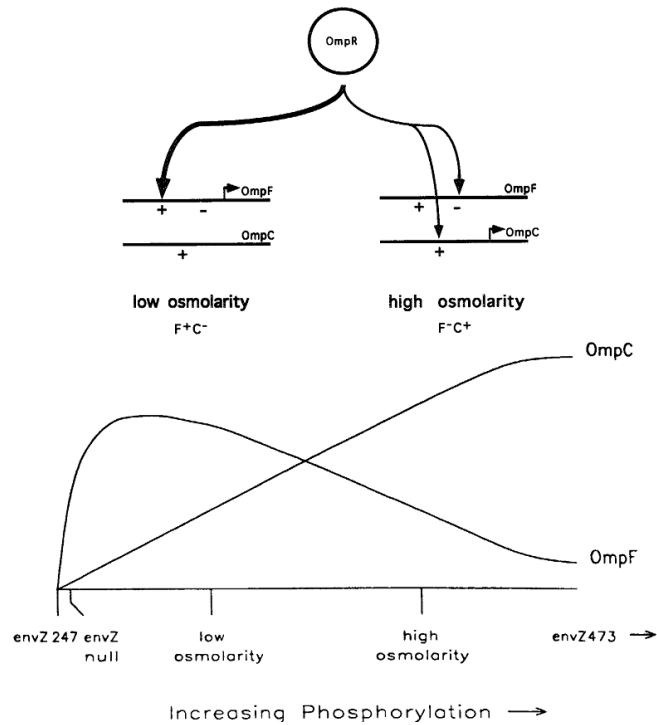


Figure 26 : Modèle de régulation d'OmpC et OmpF par OmpR-P. Issue de (Pratt and Silhavy 1995).

Des études ultérieures ont proposé que la régulation différentielle d'*ompC* et d'*ompF* soit liée à la présence de sites de faible et de forte affinité pour la forme phosphorylée d'OmpR au promoteur d'*ompF* (Kenney 2002). En effet, à faible osmolarité, seuls les sites de forte affinité pour OmpR seraient occupés et *ompF* serait activé. A forte osmolarité, un site supplémentaire serait fixé entraînant une courbure de l'ADN et donc une répression d'*ompF* (figure 27). Pour *ompC*, OmpR ne se fixerait qu'à forte osmolarité.

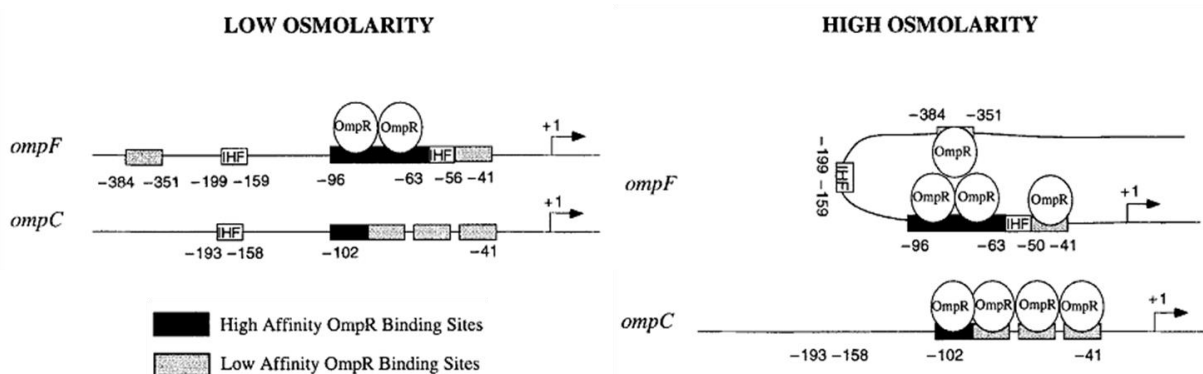


Figure 27 : Premier modèle d'osmorégulation d'*ompC* et d'*ompF* par OmpR. A faible osmolarité, seuls les sites de haute affinité d'*ompF* sont occupés par OmpR. A forte osmolarité, OmpR se fixe également sur un site de faible affinité entraînant la répression d'*ompF*. Pour *ompC* les sites de faible et forte affinité ne sont occupés qu'à forte osmolarité. Issue de (Pratt, Hsing et al. 1996).

Un autre modèle propose quant à lui que cette fois tous les sites d'*ompF* soient fixés mais qu'à forte osmolarité, un changement de conformation d'OmpR entraînerait ce même phénomène de courbure de l'ADN et donc une répression d'*ompF* (figure 28). Pour *ompC*, le modèle proposé est qu'à faible osmolarité, OmpR se fixerait sur le site de forte affinité C1 (site le plus éloigné du démarrage), *ompC* n'est donc pas transcrit. A forte osmolarité, OmpR se fixerait aux sites C1, C2 et C3 entraînant l'activation de la transcription d'*ompC* par un phénomène de coopérativité entre les sites. Ce modèle sous-entend que l'affinité d'OmpR et d'OmpR-P serait la même or nous savons qu'OmpR-P a une affinité bien plus importante qu'OmpR, pour le promoteur d'*ompC* notamment, aussi nous avons une préférence pour le premier modèle de régulation décrit par Pratt et Silhavy.

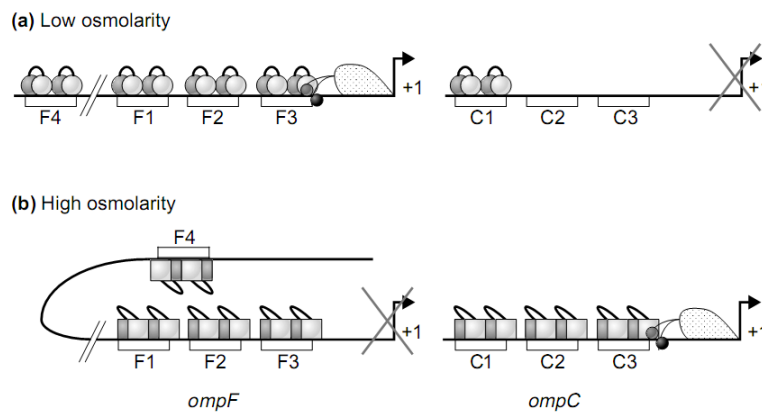


Figure 28 : Second modèle d'osmorégulation d'*ompC* et d'*ompF* par OmpR. OmpR se fixe hiérarchiquement aux trois sites entre -100 et -40 par rapport au démarrage de la transcription d'*ompC* et d'*ompF* et à un site en amont (-380 -350) pour *ompF*. Issue de (Kenney 2002).

1.2.5.4. Autres cibles d'EnvZ-OmpR

Outre les porines OmpC et OmpF, EnvZ-OmpR régule l'expression d'autres gènes impliqués notamment dans le transport des acides gras, la virulence, l'osmorégulation. L'étude transcriptomique d'Oshima *et al.* en 2002 prédit 125 cibles directes ou indirectes du système à deux composants EnvZ-OmpR (Oshima, Aiba *et al.* 2002). A ce jour, différentes études fonctionnelles et mécanistiques ont pu valider plus d'une dizaine de cibles d'OmpR.

En plus d'OmpC et d'OmpF, nombre de cibles d'OmpR sont liées à la membrane et bien souvent à la virulence. En effet, l'adhésion et la motilité bactérienne sont des fonctions importantes dans la virulence et nécessitent des appendices de surface permettant à la cellule de se mouvoir ou d'adhérer à des surfaces. Ces fonctions sont contrôlées par deux régulateurs majeurs : CsgD et

FlhDC respectivement. Chez *E. coli*, l'opéron *csgDEFG* code les protéines nécessaires à l'export et à la stabilisation des fibres « curli ». *csgD*, le premier gène de l'opéron, est un régulateur transcriptionnel (activateur et répresseur) des gènes associés à l'adhésion bactérienne ainsi qu'à la formation de biofilms. Les facteurs environnementaux pouvant influencer la transcription de *csgDEFG* sont la disponibilité en nutriments, la densité cellulaire, et les chocs osmotiques, acides et de température. Il a été montré en 2001 qu'OmpR était requis pour l'activation de l'opéron *csgDEFG* (Prigent-Combaret, Brombacher et al. 2001). Des expériences d'empreinte ont pu mettre en évidence six sites de fixation d'OmpR en amont de *csgD* (Gerstel, Park et al. 2003).

Des expériences avec des fusions *flhDC-lacZ* ont par ailleurs montré que la transcription de *flhDC* codant FlhDC, le principal régulateur de la synthèse des flagelles, était affectée par les changements d'osmolarité mais également par les variations de niveau d'acétylphosphate. Des expériences d'empreinte à la DNase I ont montré qu'OmpR, en se fixant aux régions -17 à +49 et -135 à -179, protège ces régions de *flhDC* et agirait comme un répresseur transcriptionnel (Shin and Park 1995).

OmpR régule donc directement les flagelles et les curli. De façon intéressante, il a été montré qu'OmpR régulait de façon plus indirecte ces structures via la répression de *bola* (Yamamoto, Nagura et al. 2000). Ce facteur de transcription est très important pour la transition entre le mode de vie planctonique ou en biofilm (Dressaire, Moreira et al. 2015). La surexpression de BOLA conduit notamment à la formation de biofilms.

Les pili de type I sont également des appendices requis dans la virulence de nombreuses souches et notamment chez les souches d'*E. coli* uropathogènes (Abraham, Babu et al. 1985). Chez ces organismes le passage entre une bactérie recouverte de pili (phase ON) et une bactérie sans aucun pili (phase OFF) implique l'inversion d'une portion du génome, *fimS*, comprenant le promoteur de *fimA*, la sous-unité majeure des pili. En aval de *fimA* se situent les gènes *fimB* et *fimE* codant les recombinases spécifiques à cette inversion (Abraham, Freitag et al. 1985). En 2012, il a été montré que *fimB* était sous le contrôle direct d'OmpR (Rentschler, Lovrich et al. 2013). La répression de *fimB* par OmpR en condition de pH acide et de forte osmolarité (tractus urinaire) conduit à une faible expression des pili de type I à la surface de la bactérie ce qui limiterait sa phagocytose.

Outre *csgD*, *flhDC* et *fimB*, EnvZ-OmpR contrôle l'expression d'autres facteurs impliqués dans la virulence. En effet, ce système intervient dans la synthèse de la microcine MccB (antibiotique produit par certaines entérobactéries à Gram négatif) chez *E. coli* (Moreno, Gonzalez-Pastor et al. 2002) mais également dans la synthèse du polysaccharide Vi, un élément de la capsule indispensable à la réplication de *Salmonella* dans les macrophages (Lee, Detweiler et al. 2000). Enfin il a été montré qu'EnvZ-OmpR, via la régulation du système à deux composants SsrA-SsrB, intervenait dans la mise en place du système de sécrétion de type III lors de l'infection des cellules eucaryotes par *Salmonella typhimurium* (Garmendia, Beuzon et al. 2003).

La membrane cellulaire est un lieu d'échange très important pour la cellule et il est intéressant de voir qu'OmpR régule également des gènes impliqués dans l'entrée des acides gras et des solutés mais également la capture du fer. Il a en effet été montré qu'*E. coli* était capable d'utiliser les acides gras du milieu extérieur comme unique source de carbone et d'énergie ou comme précurseurs dans la biosynthèse des phospholipides membranaires. Le transport de ces molécules est assuré par la protéine FadL principalement dans des conditions de faible osmolarité. A forte osmolarité, OmpR réprime l'expression de *fadL* en se fixant de part et d'autre du site de démarrage de la transcription (Higashitani, Nishimura et al. 1993). D'autre part, en 2004, Goh *et al.* ont pu montrer qu'une perméase, codée par le gène *ydgR* (*dtpA*) chez *E. coli* et *tppB* chez *Salmonella enterica serovar typhimurium*, était régulée transcriptionnellement par le système EnvZ-OmpR (Goh, Siino et al. 2004). Des expériences d'empreinte à la DNase I ont pu montrer qu'OmpR se fixait à un site spécifique. La particularité de ce système est qu'il n'est pas sensible à l'osmolarité. En effet, les sites resteraient occupés quelle que soit la concentration d'OmpR phosphorylée dans le milieu donc que ce soit à basse ou à haute osmolarité. Pour finir, la chélation du fer semble être un phénomène également régulé par le système EnvZ-OmpR. En effet, il a été montré que la synthèse de l'entérobactine, un sidérophore chez *E. coli*, est augmentée dans un mutant *ompB* (Oshima, Aiba et al. 2002). Le mécanisme, direct ou indirect, reste encore inconnu à ce jour.

Jusqu'ici nous avons vu qu' EnvZ-OmpR contrôle le niveau de diverses protéines mais ce système régule également la synthèse de plusieurs petits ARN régulateurs : MicC et MicF (Chen, Zhang et al. 2004), OmrA et OmrB (Guillier, Gottesman et al. 2006). MicC et MicF répriment l'expression d'*ompC* et d'*ompF* respectivement. Nous savons que la régulation de MicF par OmpR est directe. Il est intéressant de noter que les gènes *micF* et *ompC*, exprimés de manière divergente, sont

adjacents et partagent les sites de fixation d'OmpR (Delihias and Forst 2001). Le mécanisme de régulation de MicC par OmpR est encore inconnu. En ce qui concerne la régulation des Omr par OmpR, je détaillerai ce point dans la partie 1.3.3 consacrée aux ARN régulateurs.

En conclusion, OmpR régule de nombreuses cibles qui de façon intéressante sont majoritairement reliées à la membrane. La figure 29 récapitule les cibles d'OmpR retrouvées chez *E. coli* K12, notre organisme principal d'étude.

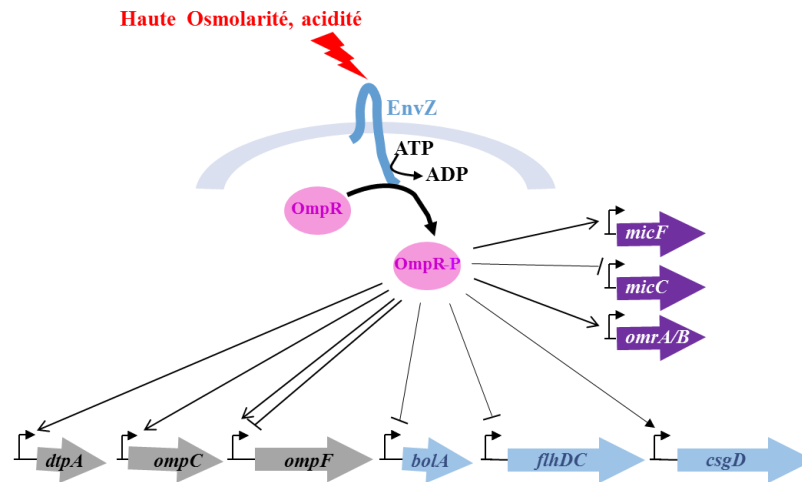


Figure 29 : Cibles du système EnvZ-OmpR chez *E. coli* K12. Les petits ARN sont indiqués en violet, les cibles reliées à la motilité en bleu et les protéines de membrane telles que les porines et les perméases en gris.

1.2.5.5. Séquence consensus de fixation d'OmpR

Bien qu'une séquence consensus de fixation d'OmpR ait été proposée plusieurs fois, il est néanmoins difficile de définir un consensus précis si l'on prend en compte tous les gènes connus à ce jour et régulés de façon directe par OmpR. Une première publication (Maeda, Takayanagi et al. 1991) indiquait qu'OmpR se fixait sur une séquence consensus asymétrique de 20 nucléotides « TTTTACTTTTTGTAACATAT ». Plus tard, en 1995, il a été proposé qu'OmpR se fixe sur une séquence de type « TTACATNTN » répétée deux fois (Harlocker, Bergstrom et al. 1995). Pour finir, en 2013, une étude de ChIP-seq portant sur l'étude du régulon OmpR chez *Salmonella* a proposé que le motif de fixation d'OmpR soit « TGTWACAW » où W peut être soit un A ou un T. Le ChIP-seq (Chromatin ImmunoPrecipitation Sequencing) consiste à cartographier les régions génomiques associées à une protéine spécifique, ici OmpR, et définir la séquence pour chacun des sites de fixation. L'alignement de tous les sites a permis aux auteurs de proposer un site consensus.

Au vu des différentes études, un consensus strict pour OmpR ne semble donc pas se dégager. La figure 30 présente l'alignement des sites de fixation d'OmpR sur certains promoteurs. On remarque que seuls 2 nucléotides sont totalement conservés. Par ailleurs, les paires « AC » souvent présentes sont requises pour la fixation d'OmpR à *ompC* et à *ompF* *in vivo* comme le montrent les expériences de mutagenèse de Pratt et Silhavy (Pratt and Silhavy 1995).

```
ompC1 : ACATTTTGAAACATCTAT
ompC2 : GATAAATGAAACATCTTA
ompC3 : AGTTTTAGTATCATATTC
ompF1 : ACTTTTGGTTACATATTT
ompF2 : TCTTTTTTGAAACAAATC
ompF3 : TATCTTTGTAGCACTTTC
ompF4 : TTACATTGCAACATTTAC
omrA  : ACACCTCGTTGCATTTCC
omrB  : ACCTTTGGTTACCTTTG
```

Figure 30 : Comparaison des sites de fixations d'OmpR présents sur le promoteur d'*ompC* (3 sites C1/C2/C3) et d'*ompF* (4 sites F1/F2/F3/F4) avec les sites putatifs de fixation sur le promoteur d'*omrA* et d'*omrB*. Les nucléotides en rose sont ceux totalement conservés entre les différents sites. Les bases A et C soulignées sont importantes pour la fixation d'OmpR *in vivo* (Pratt and Silhavy 1995).

1.2.5.6. Robustesse du système EnvZ-OmpR

Une étude de 2003 prédisait que la quantité de forme phosphorylée d'OmpR était robuste c'est-à-dire insensible aux variations d'EnvZ ou d'OmpR (Batchelor and Goulian 2003).

Ces résultats sont basés sur un modèle mathématique qui prend en compte le fait qu'*in vivo* la quantité de molécules d'OmpR est plus importante que celle d'EnvZ d'une part et d'autre part que la kinase EnvZ est bifonctionnelle c'est-à-dire qu'elle est à la fois capable de phosphoryler et de déphosphoryler OmpR.

Dans un premier temps les auteurs ont décrit le système EnvZ-OmpR en faisant apparaître les différentes réactions chimiques ainsi que les coefficients cinétiques notés « k » associés à ces transformations (figure 31). A ce stade on voit que la quantité d'OmpR-P ne dépend que de la concentration initiale d'EnvZ et d'OmpR.

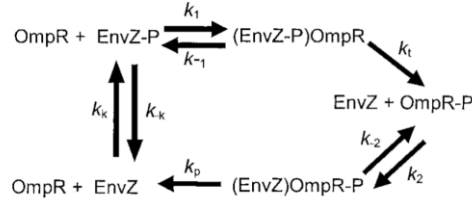


Figure 31 : Description cinétique du système à deux composants EnvZ-OmpR. Issue de (Batchelor and Goulian 2003).

L'évolution cinétique de chacune des espèces au cours du temps peut se définir selon les six équations différentielles indiquées ci-dessous :

$$\begin{aligned}
 [1] \quad & \frac{d[(EnvZP)OmpR]}{dt} = k_1[EnvZP][OmpR] - (k_{-1} + k_t)[(EnvZP)OmpR] \\
 [2] \quad & \frac{d[(EnvZ)OmpRP]}{dt} = -(k_p + k_{-2})[(EnvZ)OmpRP] + k_2[EnvZ][OmpRP] \\
 [3] \quad & \frac{d[EnvZP]}{dt} = k_k[EnvZ] - k_{-k}[EnvZP] + k_{-1}[(EnvZP)OmpR] - k_1[EnvZP][OmpR] \\
 [4] \quad & \frac{d[EnvZ]}{dt} = k_{-k}[EnvZP] - k_k[EnvZ] + (k_p + k_{-2})[(EnvZ)OmpRP] + k_t[(EnvZP)OmpR] - k_2[EnvZ][OmpRP] \\
 [5] \quad & \frac{d[OmpR]}{dt} = k_{-1}[(EnvZP)OmpR] - k_1[EnvZP][OmpR] + k_p[(EnvZ)OmpRP] \\
 [6] \quad & \frac{d[OmpRP]}{dt} = k_t[(EnvZP)OmpR] + k_{-2}[(EnvZ)OmpRP] - k_2[EnvZ][OmpRP]
 \end{aligned}$$

En sommant les variations relatives à EnvZ et celles relatives à OmpR, on obtient la variation totale d'EnvZ ou d'OmpR dans le système quelque soit sa forme. Les auteurs posent comme condition que le système soit à l'équilibre donc que les variations de concentration soient nulles. Par conséquent on peut écrire :

$$\begin{aligned}
 \frac{d[(EnvZP)OmpR]}{dt} + \frac{d[(EnvZ)OmpRP]}{dt} + \frac{d[EnvZP]}{dt} + \frac{d[EnvZ]}{dt} &= 0 \\
 \frac{d[(EnvZP)OmpR]}{dt} + \frac{d[(EnvZ)OmpRP]}{dt} + \frac{d[OmpR]}{dt} + \frac{d[OmpRP]}{dt} &= 0
 \end{aligned}$$

Ces deux équations donnent la concentration totale d'EnvZ, $[EnvZ]_T$ ou d'OmpR, $[OmpR]_T$, à l'équilibre. En intégrant ces équations il est possible d'écrire :

$$\begin{aligned}
 [7] \quad & [EnvZ]_T = [(EnvZP)OmpR] + [(EnvZ)OmpRP] + [EnvZP] + [EnvZ] \\
 [8] \quad & [OmpR]_T = [(EnvZP)OmpR] + [(EnvZ)OmpRP] + [OmpR] + [OmpRP]
 \end{aligned}$$

Les auteurs proposent de remplacer dans le système regroupant les équations 1-6, une équation dépendant d'EnvZ (équations 1-4) par l'équation [7] et une équation dépendant d'OmpR (équations 1/2/5/6) par l'équation [8] afin de faire intervenir $[EnvZ]_T$ et $[OmpR]_T$ car ce sont les données mesurables *in vivo*. Arbitrairement ils substituent l'équation [4] par l'équation [7] et

l'équation [6] par l'équation [8]. En prenant en compte que le système soit à l'équilibre et que les variations sont nulles on peut définir le nouveau système d'équations suivant :

$$\begin{aligned}
 [9] \quad & 0 = k_1[EnvZP][OmpR] - (k_{-1} + k_t)[(EnvZP)OmpR] \\
 [10] \quad & 0 = -(k_p + k_{-2})[(EnvZ)OmpRP] + k_2[EnvZ][OmpRP] \\
 [11] \quad & 0 = k_k[EnvZ] - k_{-k}[EnvZP] + k_{-1}[(EnvZP)OmpR] - k_1[EnvZP][OmpR] \\
 [12] \quad & 0 = k_{-1}[(EnvZP)OmpR] - k_1[EnvZP][OmpR] + k_p[(EnvZ)OmpRP] \\
 [7] \quad & [EnvZ]_T = [(EnvZP)OmpR] + [(EnvZ)OmpRP] + [EnvZP] + [EnvZ] \\
 [8] \quad & [OmpR]_T = [(EnvZP)OmpR] + [(EnvZ)OmpRP] + [OmpR] + [OmpRP]
 \end{aligned}$$

Afin de définir les paramètres influençant le niveau d'OmpR-P dans la cellule, les auteurs ont proposé une résolution mathématique de ce système. Pour cela, ils ont successivement réduit les équations deux à deux pour obtenir une seule équation exprimant la concentration d'OmpR-P en fonction de tous les autres termes. Ils obtiennent l'équation suivante :

$$\begin{aligned}
 & [OmpRP]^2 + (\varepsilon - C_t - C_p - [OmpR]_T) \cdot [OmpRP] + C_p \cdot [OmpR]_T - \varepsilon \cdot C_p = 0 \\
 \text{avec } & C_t = \frac{k_{-k}(k_t + k_{-1})}{k_t k_1} \quad C_p = \frac{k_k(k_p + k_{-2})}{k_p k_1} \quad \text{et } \varepsilon = [(EnvZP)OmpR] + [(EnvZ)OmpRP]
 \end{aligned}$$

L'équation obtenue est du second degré¹ avec un discriminant positif et donc admet deux solutions. La solution à racine négative est choisie puisque la concentration [OmpR-P] est supposée inférieure ou égale à [OmpR]_T. La solution peut s'écrire sous la forme :

$$[OmpRP] = \frac{1}{2}(C_t + C_p + [OmpR]_T - \varepsilon) - \frac{1}{2}\sqrt{(C_t + C_p + [OmpR]_T - \varepsilon)^2 - 4C_p[OmpR]_T - \varepsilon}$$

Dans la mesure où biologiquement nous savons que [EnvZ]_T << [OmpR]_T on peut en déduire que le ratio $\frac{[EnvZ]_T}{[OmpR]_T} \rightarrow 0$ et donc, d'après les équations [7] et [8], $\varepsilon \ll [OmpR]_T$. Les auteurs proposent donc de simplifier la solution de la manière suivante :

$$[OmpRP] = C_p + \dots$$

Les « ... » identifient tous les termes négligés par les auteurs d'après les conditions fixées ci-dessus. Ils concluent que la concentration d'OmpR-P est indépendante de celle d'EnvZ mais aussi de celle d'OmpR. Ainsi OmpR-P est seulement dépendant du terme C_p. Ce terme reflète la phosphorylation et la déphosphorylation d'OmpR par EnvZ donc OmpR-P dépend entièrement du

¹ Rappel : équation du second degré $\rightarrow ax^2 + bx + c = 0$ avec comme discriminant $\Delta = b^2/4ac$
 Si $\Delta > 0$ l'équation admet deux solutions : $x_{1,2} = \frac{-b \pm \sqrt{\Delta}}{2a}$

signal senti par le système à deux composants EnvZ-OmpR mais pas des concentrations d'EnvZ ou d'OmpR.

Pour valider ce modèle mathématique ils ont suivi la variation de la forme phosphorylée d'OmpR lorsqu'ils modulent artificiellement la concentration d'EnvZ ou la concentration d'OmpR. La variation d'OmpR-P a été suivie de manière indirecte en étudiant la transcription issue des promoteurs d'*ompC* et d'*ompF*. Ces expériences ont pu valider le modèle mathématique puisque la transcription d'*ompC* et d'*ompF* est constante dans une large gamme d'EnvZ ou d'OmpR (figure 32). Néanmoins on remarque que si on surproduit de façon trop importante OmpR (plus de 10 fois par rapport au niveau sauvage) on perd la robustesse.

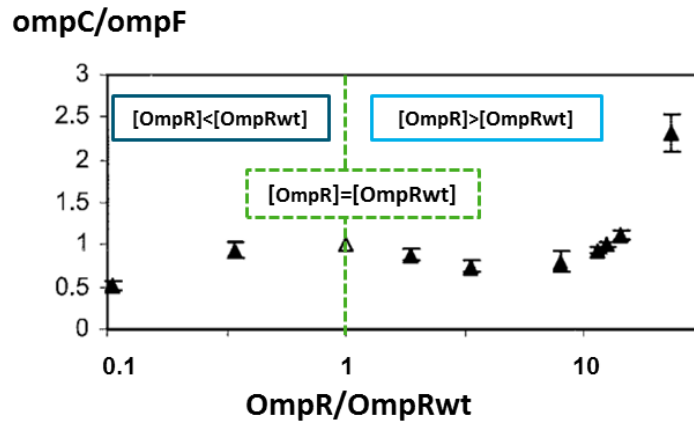


Figure 32 : Effet de la modulation du niveau d'OmpR sur la transcription d'*ompC* et d'*ompF* d'après (Batchelor and Goulian 2003). En abscisse on retrouve le niveau d'OmpR surproduit par rapport au niveau chromosomique d'OmpR (OmpRwt). En ordonnée est indiqué le ratio entre la transcription d'*ompC* et d'*ompF*. Pour obtenir ces ratios les auteurs ont utilisé des fusions transcriptionnelles entre *ompC-cfp* et *ompF-yfp*. Ce profil est valable à la fois en condition de basse et haute osmolarité mais également pour EnvZ quand c'est la concentration d'EnvZ qui varie. Les valeurs sont représentées en ratio mais ces données sont valables pour chaque fusion prise séparément (Communication personnelle M. Goulian - Siryaporn, Perchuk et al. 2010).

Ce modèle est relativement étonnant puisqu'on sait qu'il a été décrit des contrôles au niveau transcriptionnel du système EnvZ-OmpR par IHF (Tsui, Huang et al. 1991) et CRP (Huang, Tsui et al. 1992) et comme nous le verrons plus tard un contrôle post-transcriptionnel par les petits ARN OmrA et OmrB. De plus, OmpR s'autorégule a priori de façon positive (Quinn, Cameron et al. 2014). On peut donc légitimement se poser la question de l'intérêt de ces régulations si elles n'affectent pas la quantité de la forme supposé active du régulateur à savoir OmpR-P. Une grande partie de mon travail de thèse aborde cette question dans le cadre de la régulation par les Omr.

1.2.6. Les systèmes NarQ-NarP et NarX-NarL

Un autre système à deux composants sur lequel je me suis penchée dans ce travail de thèse est le système NarQ-NarP qui va de paire avec le système NarX-NarL chez *E. coli*. Tous deux sont capables de sentir la présence de nitrates et de nitrites dans le milieu extérieur. Après une brève introduction sur la respiration chez *E. coli*, je détaillerais ces deux systèmes.

1.2.6.1. Généralités sur la respiration chez *Escherichia coli*

La respiration cellulaire comprend l'ensemble des processus permettant aux organismes de générer de l'énergie sous forme d'ATP et via le gradient de protons à partir des nutriments présents dans leur environnement. Que ce soit en aérobiose ou en anaérobiose, les chaînes respiratoires des bactéries telles qu'*E. coli* sont composées de trois éléments principaux : les déshydrogénases primaires, les quinones et oxydoréductases terminales. Séquentiellement les déshydrogénases primaires vont oxyder un substrat spécifique (NADH le plus souvent) et transférer les électrons à un pool de quinones. Celles-ci sont des intermédiaires permettant le transport des électrons et des protons depuis les deshydrogénases jusqu'aux oxydoréductases. L'étape finale consiste en la réduction de l'accepteur final d'électrons par oxydation des quinones. Cette dernière étape est réalisée par les oxydoréductases terminales. La figure 33 présente les principaux donneurs et accepteurs d'électrons et types de quinones présentes chez *E. coli*.

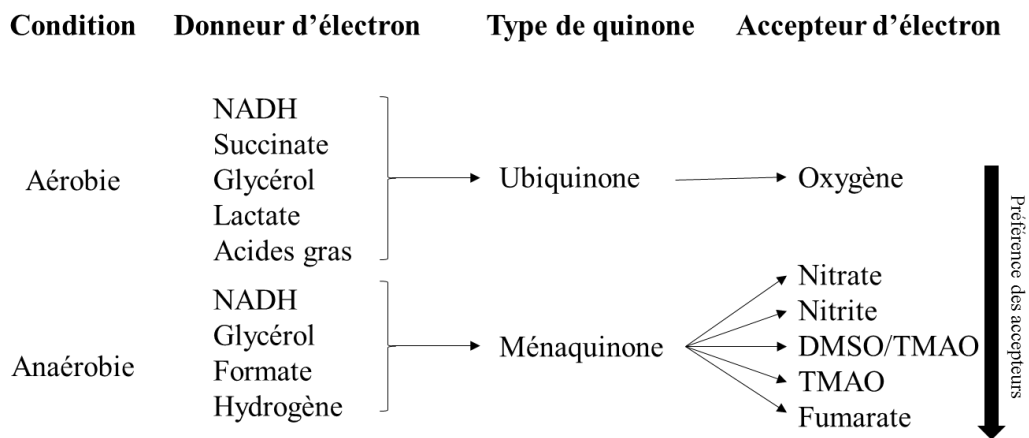


Figure 33 : Vue d'ensemble de la chaîne respiratoire d'*E. coli*. En condition de croissance aérobie où le glucose est la principale source de carbone, l'oxygène est l'accepteur majeur d'électron. Pour les autres accepteurs d'électrons, le flux d'électrons est conduit au pool de quinone par des déshydrogénases spécifique à chaque donneur. (TMAO) : trimethyl-amine-N-oxide ; (DMSO) : dimethyl-sulphoxide.

Ainsi, lorsque l'oxygène se raréfie, certaines espèces, dont *E. coli* sont capables de faire appel à d'autres accepteurs d'électrons tels que les nitrates, le fumarate ou se diriger vers la fermentation de différents sucres simples. Il est important de préciser que chez *E. coli*, les nitrates et nitrites peuvent jouer le rôle d'accepteurs finaux d'électron mais la bactérie ne les assimile pas ; elle les rejette dans le milieu extérieur (sous forme d'ammoniac).

La réduction des nitrates en nitrites ($\text{NO}_3^- + 2e^- + 2\text{H}^+ \rightarrow \text{NO}_2^- + \text{H}_2\text{O}$) est une étape catalysée par des enzymes de type nitrate réductase. En aérobiose, l'expression de ces enzymes est réprimée afin que les électrons de la chaîne respiratoire soient préférentiellement orientés vers l'oxygène possédant un potentiel redox plus élevé ($\text{O}_2/\text{H}_2\text{O}$ $E'^\circ = + 820$ mV) que celui du nitrate ($\text{NO}_3^-/\text{NO}_2^-$ $E'^\circ = + 420$ mV). On peut noter que le rendement énergétique de la respiration sur nitrates est plus faible que sur oxygène². Chez les bactéries à Gram positif, les nitrate réductases sont associées à la membrane (Potter, Angove et al. 2001). En revanche chez les bactéries à Gram négatif on peut trouver jusqu'à trois types de nitrate réductases qui ont comme point commun d'être associées à un cofacteur à molybdène (MoCo) de type bis-MGD (bis(molybdoptérine guanine dinucléotide)molybdène). On retrouve dans le cytoplasme des nitrate réductases de type NAS (assimilatory nitrate reductase) solubles. Celles associées à la membrane sont dites de type NAR (respiratory nitrate reductase). Et enfin, les nitrate réductases solubles dans le périplasma de type NAP (periplasmic nitrate reductase). La figure 34 présente l'organisation de ces trois types de nitrate réductases.

² glucose + 12 NO₃⁻ → 6 CO₂ + 6 H₂O + 12 NO₂⁻ ΔG^{'°} = -1766 kJ et glucose + 6 O₂ → 6 CO₂ + 6 H₂O ΔG^{'°} = -2870 kJ

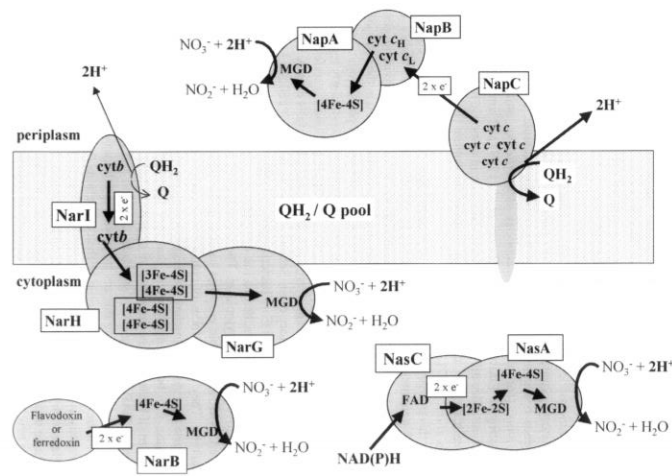


Figure 34 : Organisation structurale des différents types de nitrate réductases bactériennes. NarGHI est un exemple de nitrate réductase associée à la membrane. Dans ce cas le transport d'électrons du quinol (ex : ubiquinol UQH₂) au nitrate est couplé à une boucle redox pour un retour à l'état initial. NapABC est un exemple de nitrate réductase périplasmique et enfin NasAC est un exemple de nitrate réductase cytoplasmique. NarB est un exemple de nitrate réductase cytoplasmique à une seule sous-unité présente chez les cyanobactéries par exemple. Issue de (Potter, Angove et al. 2001).

Dans le but d'utiliser les nitrates comme accepteur final d'électron, *E. coli* synthétise deux enzymes : une nitrate réductase ancrée à la membrane, appelée NRA, codée par l'opéron *narGHJI* et une nitrate réductase périplasmique soluble NapA codée par l'opéron *napFDAGHBC* (Stewart 1988). Une seconde nitrate réductase membranaire est présente chez cet organisme, la nitrate réductase NRZ codée par l'opéron *narZYWV*. La synthèse de NRA est induite par la présence de nitrate uniquement en anaérobie alors que NRZ est synthétisée à un niveau cryptique rendant la transition aérobie-anaérobie plus aisée. NapA est quant à elle synthétisée en présence de faibles concentrations de nitrates que ce soit en présence ou en absence d'oxygène. Lorsque la concentration en nitrates augmente son expression est réprimée. A forte concentration en nitrates, c'est NRA qui est l'enzyme prédominante. Une fois les nitrates réduits en nitrites, ces derniers sont réduits en ammoniac (NH₃) par les enzymes de type nitrite réductase (figure 35). Les nitrites peuvent également directement servir d'accepteurs finaux d'électrons. *E. coli* possède deux nitrites réductases : NirB localisée dans le cytoplasme et NrfA située dans le périplasm.

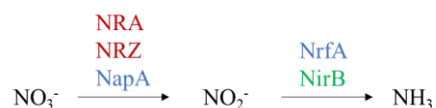


Figure 35 : Résumé des différentes nitrate et nitrite réductases retrouvées chez *E. coli*.

1.2.6.2. Régulation de la transition aérobie-anaérobie

Lors de la transition aérobie-anaérobie l'expression de nombreux gènes doit être régulée. Une étude a montré par une approche de puce à ADN, qu'un tiers des transcrits d'*E. coli* exprimés en condition aérobie sont altérés lors de cette transition (Salmon, Hung et al. 2003). Ce remodelage nécessitant une bonne coordination, il a ensuite été montré que 49 % de ces gènes sont directement ou indirectement modulés par un seul et même régulateur : FNR. L'activation synchronisée de nombreux gènes en anaérobie par FNR est donc accompagnée de contrôles secondaires en fonction des sources d'énergie disponibles. Après l'oxygène, les nitrates sont la source d'énergie préférentielle puis vient le fumarate et enfin l'acétate (Unden and Bongaerts 1997). Aussi en anaérobie, l'utilisation de la voie fumarate est réprimée tant qu'il y aura des nitrates dans le milieu. Le respect de cette règle est assuré par les systèmes à deux composants NarX-NarL et NarQ-NarP. Les kinases seront capables de « sentir » la présence des nitrates et les régulateurs vont, entre autre, inhiber les gènes codant des fumarates réductase et activer ceux pour certaines nitrate réductases. Il est intéressant de préciser que les nitrate et nitrite réductases d'*E. coli* sont pour la grande majorité également contrôlées par FNR. On peut par exemple citer le cas de NRA où le régulateur NarL et FNR doivent agir de concert pour activer efficacement la transcription de l'opéron *narGHJ* (Spiro and Guest 1990). On peut également noter que la transcription de l'opéron *narXL* pourrait être réprimée par FNR mais le résultat de différentes études est contradictoire (Takahashi, Hattori et al. 1994; Constantinidou, Hobman et al. 2006).

Par ailleurs, en absence d'oxygène, la transcription de nombreux gènes est également contrôlée par le système à deux composants ArcB-ArcA et en particulier les gènes du métabolisme central et notamment différentes enzymes du cycle de Krebs (*sdhCDAB*, *icd*, *fumA*, *mdh*, *gltA*, *acnA*, et *acnB*) mais aussi les gènes codant des éléments de la chaîne respiratoire comme les cytochromes (Iuchi and Lin 1988). ArcA agit le plus souvent comme un répresseur en condition anaérobie mais il est également capable d'activer la transcription (Malpica, Sandoval et al. 2006). Par ailleurs, on peut préciser que FNR active (très faiblement cependant) le système ArcB-ArcA et que certains gènes cibles d'ArcA sont également contrôlés par FNR suggérant une coordination entre les régulateurs (Iuchi and Lin 1993).

1.2.6.3. Les kinases senseurs NarX et NarQ

Chez *E. coli*, les gènes codants les systèmes NarX-NarL et NarQ-NarP sont situés sur trois loci différents. *narX* et *narL* sont codés en opéron alors que *narQ* et *narP* sont des gènes codés indépendamment à des loci distants de presque 300 kb.

Les kinases NarX et NarQ sont des paralogues c'est-à-dire que la séquence de ces deux protéines a divergé après une duplication. De façon intéressante on ne trouve que chez *E. coli* et chez *S. enterica* à la fois *narX* et *narQ* associés à *narL* et *narP* (Stewart 2003). Dans les autres génomes séquencés on ne retrouve qu'une seule version de la kinase sensible aux nitrates. Par homologie de séquence et conservation de la synténie il est possible de préciser si l'organisme possède une kinase de type « NarX » ou de type « NarQ ». Par exemple, chez *Vibrio cholerae* ou *Yersinia pestis* n'est présent que *narQ* alors que chez *Pseudomonas aeruginosa* on ne retrouve que *narX*. La présence simultanée de NarQ et de NarX a donc beaucoup été étudiée et a pu mettre en évidence le rôle subtil de ces deux kinases mais aussi leurs caractéristiques communes.

1.2.6.3.1. Caractéristiques communes

NarX et NarQ sont deux protéines associées à la membrane de 598 et 566 acides aminés respectivement. Elles possèdent une topologie similaire à savoir deux hélices transmembranaires délimitant le domaine N-terminal périplasmique essentiel à la perception du signal et relié au domaine cytoplasmique par un linker.

Leurs séquences en acides aminés partagent 32 % d'identité, essentiellement localisée dans 4 « motifs » de la région cytoplasmique et un « motif » de la région périplasmique. La séquence périplasmique conservée nommée « boîte P » est constituée de 18 acides aminés consécutifs dont 15 sont identiques (Chiang, Cavicchioli et al. 1992). Cette région étant unique à NarX et à NarQ, il a été proposé qu'elle soit à l'origine de la reconnaissance des nitrates. La partie cytoplasmique contient quant à elle trois domaines (H, N et G1) retrouvés chez d'autres systèmes à deux composants et un domaine « C » spécifique à NarX et NarQ (figure 36). Le domaine H porte notamment le site d'autophosphorylation à savoir le résidu histidine 399 pour NarX et histidine 370 pour NarQ. Sont également conservés une seconde histidine (NarX His-513 et NarQ His-484) et une asparagine (NarX Asn-509 et NarQ Asn-480) toutes deux essentielles à la transmission du phosphate au régulateur NarL (Cavicchioli, Schroder et al. 1995) et potentiellement au régulateur

NarP (non testé cependant). On peut noter que ces deux résidus sont séparés de quatre acides aminés ce qui est une caractéristique de la famille des kinases senseurs de type UhpB (Chiang, Cavicchioli et al. 1992).

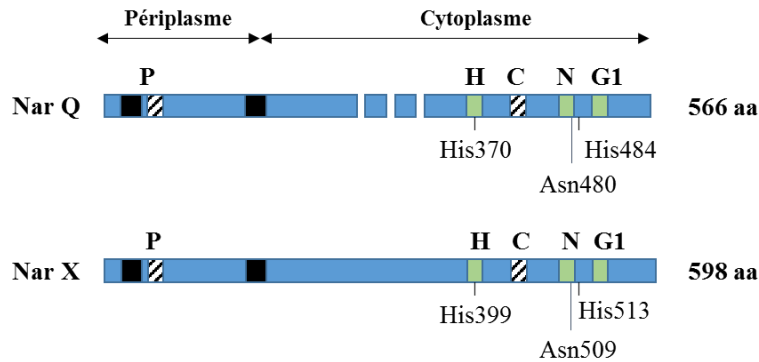


Figure 36 : Représentation schématique des protéines NarX et NarQ et de leurs domaines conservés. Les deux senseurs sont alignés par rapport à leur extrémité N-terminale (+1). En noir sont représentés les segments transmembranaires hydrophobes de 20 à 22 acides aminés de long. Les régions conservées avec les autres systèmes à deux composants à savoir les régions H, N et G1 sont indiquées en vert. Les régions conservées cytoplasmique (C) ou périplasmique (P) uniques aux protéines NarX et NarP sont représentées par des zones hachurées. La position des acides aminés conservés est également mentionnée.

1.2.6.3.2. Perception du signal

Il a été montré que la protéine NarQ reconnaissait de façon équivalente les nitrates et les nitrites alors que NarX avait une affinité différentielle pour ces deux ligands avec une préférence marquée pour les nitrates (Lee, Delgado et al. 1999). Cette différenciation serait due aux deux acides aminés (His-45 et Lys-49) de la boîte P de NarX différents de la boîte P de NarQ (Williams and Stewart 1997). En effet, une mutation du résidu Lys-49 de *narX* a comme conséquence d'obtenir un phénotype *narQ*⁺ à savoir qu'il n'y a plus de discrimination entre nitrite et nitrate dans une souche délétée pour *narQ* par ailleurs. Par ailleurs, une double mutation His-45/Lys-49 conduit à une insensibilité aux nitrites *in vivo*. Si ces résidus essentiels à la fonction de NarX sont mutés dans la séquence de *narQ* (ce qui correspond aux résidus Glu-41 et Arg-55) cela n'a aucun effet ce qui confirme la spécificité du senseur NarX.

Par ailleurs si l'histidine 45 et la lysine 49 semblent essentielles pour la discrimination entre nitrates et nitrites, ils ne sont a priori pas le site de fixation de ces ligands. En effet, le résidu critique serait l'arginine 54. Une étude a pu montrer qu'une mutation de ce résidu avait pour conséquence de rendre insensible NarX à la présence de nitrate ou de nitrite, laissant celle-ci bloquée dans un stade « activité phosphatase seule » (Cavicchioli, Schroder et al. 1995). Une

mutation du résidu Arg-50 de NarQ (qui correspond à l'Arg-54 de NarX) donne également un phénotype d'insensibilité (Chiang, Cavicchioli et al. 1997).

La structure cristallographique du domaine senseur de NarX en complexe avec du nitrate a pu confirmer l'importance des résidus de la boîte P et en particulier l'arginine 54. Elle a pu également préciser la structure du site de fixation du ligand (Cheung and Hendrickson 2009) à savoir que le nitrate se localise au milieu de deux dimères de NarX (figure 37). Très récemment, la structure cristallographique des domaines senseurs, transmembranaires et cytoplasmiques de NarQ a été résolue (Gushchin, Melnikov et al. 2017). Cette étude a pu mettre en évidence que la fixation des ligands (nitrate/nitrites) induisait des changements conformationnels significatifs responsables de la transduction du signal.

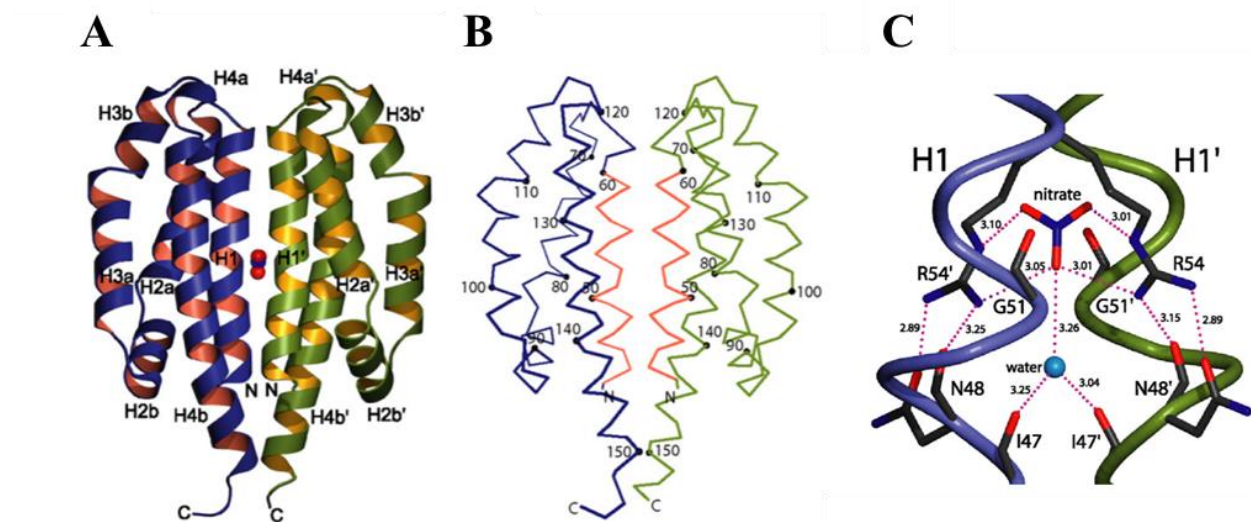


Figure 37 : Structure du domaine senseur de NarX à 1.7Å. La structure permet de mettre en évidence la localisation du nitrate entre les deux dimères de NarX (A) et plus précisément au niveau des boîtes P (en rouge sur la représentation du squelette carboné en B). Le nitrate établit des liaisons hydrogène avec les résidus N54, N48 des deux monomères de NarX ainsi qu'une molécule d'eau (C). Issue de (Cheung and Hendrickson 2009).

Un modèle propose qu'en présence de nitrate ou de nitrite dans le périplasma, ces derniers vont interagir avec les acides aminés de la boîte P de NarX ou NarQ (Chiang, Cavicchioli et al. 1997). Cette interaction aura pour effet d'engendrer un changement conformationnel du domaine périplasmique qui à la suite conduira à un changement conformationnel du domaine cytoplasmique des deux protéines. Ce dernier changement permettra aux protéines NarX et NarQ de passer d'un mode « OFF » à un mode « ON » leur permettant de s'autophosphoryler sur leur résidu histidine

(His-399 pour NarX et His-370 pour NarQ) et transmettre leur phosphate aux régulateurs. La mutation de l'arginine 54 de *narX* a pour conséquence de conserver la protéine dans un mode « OFF ». En revanche une mutation de l'histidine 45 en proline conserve la protéine dans un mode « ON ».

Enfin, l'activité autokinase de NarX est 100 fois plus sensible à la présence de nitrates que de nitrites (Lee, Delgado et al. 1999). En effet, l'autophosphorylation de NarX *in vitro* apparait à un seuil en nitrate de 5 μ M (50% d'autophosphorylation à 35 μ M). En présence de nitrite le seuil est de 500 μ M (50% autophosphorylation à 3,5 mM).

En conclusion, NarX a un seuil de détection relativement différent entre les nitrites et les nitrates ce qui est probablement dû à la nature des acides aminés présents dans la boîte P.

1.2.6.3.3. Affinité différente pour les régulateurs

Une fois que les kinases NarX et NarQ se sont autophosphorylées sur leur résidu histidine elles peuvent transmettre le phosphate au régulateur qui leur est associé. Cependant, la régulation croisée entre ces deux systèmes obéit à des règles bien particulières (figure 38).

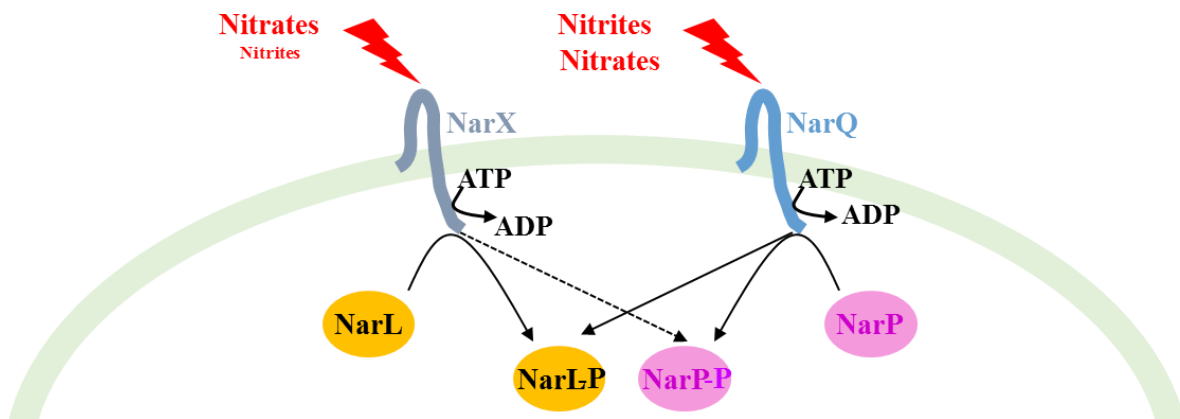


Figure 38 : Modèle de régulation croisée asymétrique pour les systèmes à deux composants NarX-NarL et NarQ-NarP. Alors que NarQ ne présente pas de préférence pour la phosphorylation des deux régulateurs, NarX phosphoryle préférentiellement NarL.

Des études avec des mutants nuls montrent que, seuls, NarX ou NarQ ne sont pas suffisants pour obtenir une réponse au nitrate normale vis-à-vis des opérons NarL-dépendants (Rabin and Stewart 1992). En revanche, NarX seul permet d'avoir un phénotype sauvage pour les gènes NarP-dépendants en absence de NarL (Rabin and Stewart 1993). Une étude en 2009 a permis de mieux comprendre le lien entre ces deux systèmes et d'expliquer l'intérêt de la redondance. En effet, il a

pu être montré que NarX avait une préférence cinétique pour la phosphorylation de NarL par rapport à celle de NarP (figure 38). En revanche NarQ montre une cinétique de phosphorylation semblable vis-à-vis des deux régulateurs (Noriega, Lin et al. 2010). Une étude plus récente a précisé que ces préférences cinétiques seraient liées à des affinités différentes dans les couples senseurs-régulateurs (Huynh, Chen et al. 2015).

Enfin, en présence de nitrite, NarX se comporte de façon prédominante comme une phosphatase pour NarL ce qui a pour conséquence d'avoir un faible niveau de forme phosphorylée de NarL et donc une faible activation des promoteurs NarL-dépendants (Rabin and Stewart 1993).

Nous pouvons donc conclure que NarX semble être un senseur relativement spécialisé répondant de façon différente aux nitrites et aux nitrates et ayant une préférence pour le régulateur NarL. A l'inverse, NarQ semble être plus généraliste n'exerçant à priori aucune préférence de ligand ou de régulateur associé.

1.2.6.4. Les régulateurs NarL et NarP

Les régulateurs NarL et NarP sont des protéines de respectivement 23,9 kDa et 23,6 kDa.

NarL active l'expression de différents opérons dont les produits permettent la respiration sur des nitrates (*fdnG*, *narG*, *narK* et *nirB*) et réprime l'expression d'opérons codant des enzymes permettant la respiration sur des sources alternatives à l'oxygène et au nitrate (*fdrA* ou *dmsA*). Le tableau 5 présente toutes les cibles connues de NarL à ce jour et si ces dernières sont également contrôlées par NarP.

Tableau 5 : Cibles de NarL et/ou de NarP validées expérimentalement

Gène/Opéron cible	Fonction	Effet de NarP	Référence
Activées par NarL			
<i>cydDC</i>	ABC transporteur	ND	(Cook, Membrillo-Hernandez et al. 1997)
<i>fdnGHI</i>	Formate déshydrogénase-N	-	(Berg and Stewart 1990; Wang and Gunsalus 2003)
<i>hcp</i>	Oxyde nitrique réductase	+	(Filenko, Spiro et al. 2007)
<i>moeAB</i>	Molybdoptérine molybdène transférase	ND	(Hasona, Self et al. 2001)
<i>narGHJI</i>	Nitrate réductase	0	(Stewart and Parales 1988)
<i>narK</i>	Nitrate-nitrite anti-porteur	0	(Bonney and DeMoss 1992)
<i>narXL</i>	Senseur (autorégulation)	0	(Stewart and Parales 1988)
<i>nirBDC</i>	NADH-nitrite réductase	+	(Page, Griffiths et al. 1990; Wang and Gunsalus 2000)
<i>nuoABCEFGHIJKLMNOP</i>	NADH déshydrogénase	ND	(Bongaerts, Zoske et al. 1995)
<i>ogt</i>	Méthyltransférase	ND	(Squire, Xu et al. 2009)
<i>yeaR-yoaG</i>	inconnue	0	(Lin, Bledsoe et al. 2007)
<i>ydeP</i>	Protéine de résistance au stress acide	0	(Yoshida, Ishihama et al. 2015)
Réprimées par NarL			
<i>adhE</i>	Aldéhyde alcool déshydrogénase	ND	(Membrillo-Hernandez and Lin 1999)
<i>aspA-dcuA</i>	Aspartate ammoniac lyase-transporteur de dicarboxylate	-	(Goh, Bledsoe et al. 2005)
<i>caiF</i>	Activateur transcriptionnel	ND	(Eichler, Buchet et al. 1996)
<i>dcuB-fumB</i>	Transporteur de dicarboxylate-fumarase B	ND	(Golby, Kelly et al. 1998)
<i>dcuS</i>	Histidine kinase du système à deux composants DcuS/DcuR	ND	(Goh, Bledsoe et al. 2005)
<i>hyaABCDE</i>	Hydrogénase 1	-	(Richard, Sawers et al. 1999)
<i>hybOABCDEFG</i>	Hydrogénase 2	-	(Richard, Sawers et al. 1999)
<i>napFDAGHBC</i>	Nitrate réductase périplasmique	++	(Choe and Reznikoff 1993; Darwin and Stewart 1995)
<i>dmsABC</i>	Diméthylsulfoxyde réductase	ND	(Cotter and Gunsalus 1989)
<i>fdrABCD</i>	Fumarate réductase	0	(Kalman and Gunsalus 1989)
<i>nrfABCDEFG * </i>	Formate-nitrite réductase	+	(Wang and Gunsalus 2000)
<i>focA-pflBA</i>	Pyruvate-formate lyase	-	(Kaiser and Sawers 1995)
<i>norVW</i>	Flavorubredoxine	-	(da Costa, Teixeira et al. 2003)
<i>fdhF</i>	Formate déshydrogénase H	0/-	(Wang and Gunsalus 2003)
<i>nikABCDE</i>	Transporteur de nickel	ND	(Rowe, Starnes et al. 2005)
<i>ompW</i>	Protéine de membrane externe	0/-	(Xiao, Lai et al. 2016)
<i>torCAD</i>	Triméthylamine N-oxide réductase	ND	(Iuchi and Lin 1987)
<i>ydhYVWXUT</i>	Oxydoréductase	-	(Al Ibrahim and Green 2017)
<i>ynfEFGHI</i>	Diméthylsulfoxyde réductase	ND	(Xu, Busby et al. 2009)

+: active faiblement ; ++: active fortement ; -: inhibe ; 0: non régulé par NarP
 ND : non déterminé * A faible concentration en nitrate NarL est activateur

D'un point de vue structurel, c'est le résidu Asp-59 de NarL qui est phosphorylé par NarX ou NarQ. Le domaine C-terminal de NarL reconnaît ses promoteurs-cibles en se fixant sous forme de dimères antiparallèles (Maris, Kaczor-Grzeskowiak et al. 2005) sur des séquences répétées inversées de type heptamère ayant un motif TACYYMT (Y = C ou T, M = A ou C) (Darwin,

Tyson et al. 1997). Précisément ce sont dans un premier temps les résidus Lys-188, Val-189 et Lys-192 qui permettent l'interaction protéine-ADN, et dans un second temps le squelette carboné de la protéine (Maris, Kaczor-Grzeskowiak et al. 2005). Les heptamères se trouvent le long des régions de contrôle des gènes régulés avec différents arrangements et peuvent être plus ou moins répétés, le plus souvent sous la forme de deux heptamères séparés par deux paires de bases ou motif « 7 – 2 – 7 ». Selon l'organisation NarL réprime ou active la transcription de ses gènes cibles. Dans de nombreux cas, la présence concomitante de FNR et celle de NarL est requise pour la régulation. Néanmoins, une étude globale a pu montrer que certains gènes pouvaient être régulés par NarL indépendamment de FNR (Constantinidou, Hobman et al. 2006). La figure 39 présente pour différents promoteurs la position des heptamères de NarL et s'il y a lieu le site de fixation de FNR.

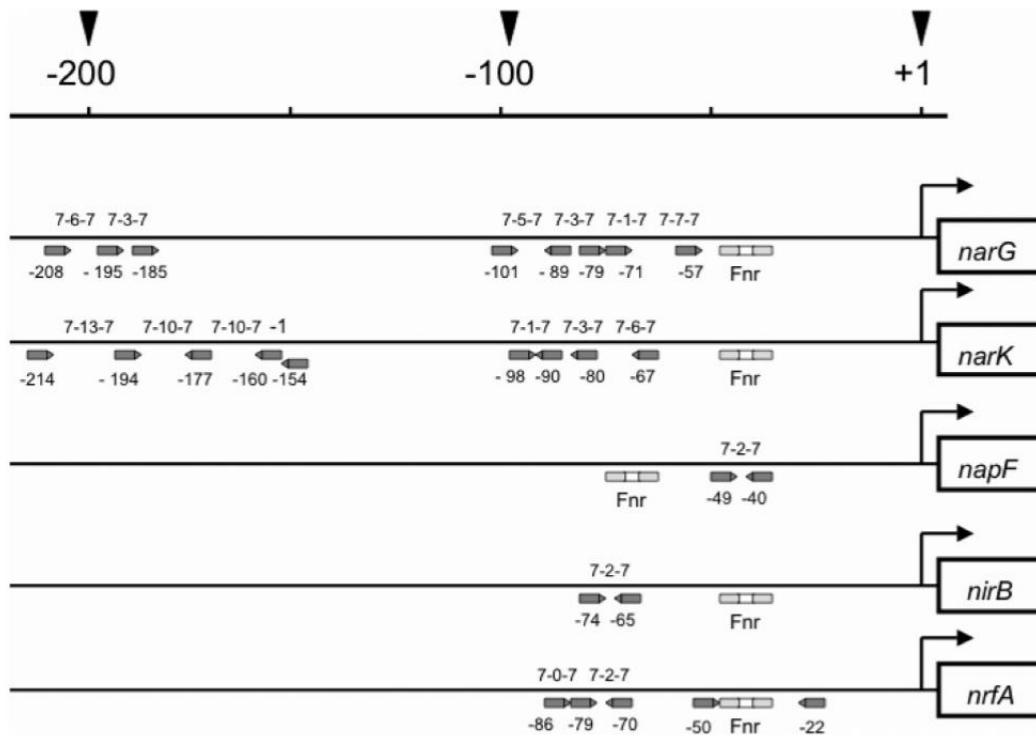


Figure 39 : Positions des heptamères de NarL sur différents promoteurs. Issue de (Darwin, Tyson et al. 1997).

Contrairement à NarL, NarP ne reconnaît que les heptamères sous forme de motifs de type « 7 – 2 – 7 » (Darwin, Tyson et al. 1997). A ce jour, toutes les cibles connues de NarP sont aussi des cibles de NarL ; en revanche NarL possède des cibles non régulées par NarP telles que les opérons *narG* et *fdrA* (tableau 5).

Afin d'illustrer les mécanismes de régulation par NarP et NarL j'ai choisi deux exemples que sont les opérons cibles *nirBDC* et *napFDAGHBC*.

L'opéron *nirBDC* code dans l'ordre les deux sous-unité de la nitrite réductase NADH-dépendante de *E. coli* (gènes *nirB* et *nirD* pour la grande et la petite sous-unité respectivement) et le transporteur de nitrite NirC (Peakman, Crouzet et al. 1990). Le site de démarrage de la transcription de *nirB* est précédé par un site de fixation de FNR centré en position -41.5 (figure 40). Un second site de fixation à FNR est également présent et centré en position -89.5. Il est intéressant de préciser que FNR seul n'est pas capable d'activer l'opéron et nécessite la présence de NarL ou NarP (Wu, Tyson et al. 1998). De la même façon NarP et NarL seuls ne sont pas suffisants. Sous forme phosphorylée ces deux régulateurs se fixent sur un motif heptamérique de type 7-2-7 centré en position 69.5 (Tyson, Cole et al. 1994).

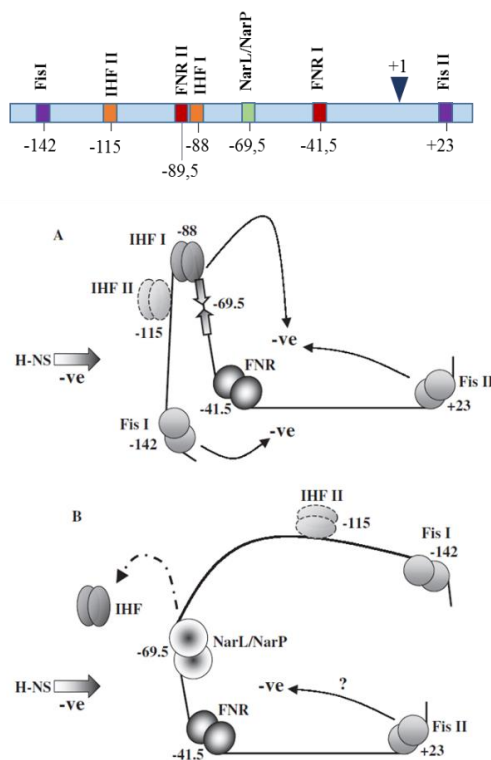


Figure 40 : Mécanisme d'activation du promoteur P_{nir} . Représentation linéaire des sites de fixation de FNR (en rouge), de Fis (en violet), d'IHF (en orange) et de NarL/NarP (en vert) sur le promoteur de l'opéron *nirBDC*. Mécanisme d'action proposé par (Browning, Cole et al. 2004) pour illustrer la régulation au promoteur *nir* en condition anaérobie et en absence (A) ou en présence (B) de nitrites ou de nitrates. Dans les deux conditions l'activité générale du promoteur est réprimée par HNS. A. La fixation de Fis et IHF inhibe transcription dépendante de FNR (-ve). B. la fixation de NarL/NarP sous forme phosphorylée déplace IHF du site I et remodèle l'architecture du promoteur de *nir*, rendant possible l'activation de la transcription par FNR.

Pour expliquer ce phénomène de co-dépendance, une étude a pu montrer l'implication d'IHF et de Fis dans la régulation du promoteur de *nirB*. D'une part Fis possède un site de fixation (Fis I) centré en position -142 du promoteur de *nirB* (P_{nirB}) et d'autre part IHF peut se fixer autour de la position -88 (site IHF I) (figure 40). La présence de ces molécules a pour effet d'empêcher l'activation de l'expression de *nir* par FNR, sans pour autant affecter sa fixation (Browning, Cole et al. 2004). Les formes phosphorylées de NarL/NarP en se fixant sur leur site vont avoir pour effet de déplacer IHF et de modifier le repliement du promoteur P_{nirB} (Browning, Cole et al. 2000). Cette modification laisse la possibilité à FNR d'activer la transcription probablement à cause de la conformation différente du promoteur. En conclusion, NarL/NarP aurait pour rôle de déplacer IHF du site I. Fis possède également un autre site de fixation en position +23 (Fis II). Une fixation sur ce site réprime le promoteur P_{nirB} probablement en interférant avec la fixation de l'ARN polymérase. Le rôle exact de Fis dans la régulation n'est pas encore élucidé. Une possibilité serait que Fis et H-NS coopèrent (site de fixation de H-NS recouvrant partiellement celui de NarL/NarP) pour réprimer l'opéron mais il n'y a aucune évidence dans ce sens.

Pour finir il est intéressant de préciser que la nitrite réductase est également soumise à un contrôle post-transcriptionnel. En effet, le messageur *nirB* est une cible de l'ARN régulateur RyhB (Wright, Richter et al. 2013).

La nitrate réductase périplasmique d'*E. coli* et ses protéines accessoires sont quant à elles codées par l'opéron *napFDAGHBC* où *napA* code la sous-unité catalytique, *napB* et *napC* le cytochrome *c*, *napD* une protéine requise pour l'assemblage de l'enzyme et enfin *napF*, *napG* et *napH* des protéines permettant le couplage entre l'oxydation de l'ubiquinol et la réduction des nitrates. La nitrate réductase est requise pour la respiration en nitrates en présence de faible concentration de ce substrat. A plus forte concentration en nitrates c'est la nitrate réductase codée par l'opéron *narG* qui prend le relais (Wang, Tseng et al. 1999). En revanche, il a été montré chez la bactérie dénitrifiante *Thiosphaera pantotropha*, qu'en l'absence de l'opéron *narG*, un mutant pour la région régulatrice de l'opéron *napF* était capable d'assurer la respiration nitrate anaérobie via la surexpression de ce système (Bell, Page et al. 1993).

Durant la croissance anaérobie, l'opéron *napF* est induit par FNR et NarP alors qu'il est réprimé par NarL. Le site de fixation de FNR est centré entre les positions -64 et -65 par rapport au +1 de transcription (notation -64.5). Seul, FNR stimule à un faible niveau la transcription. Les formes

phosphorylées de NarP et de NarL (NarP-P et NarL-P) se fixent quant à elles, aux deux séquences heptamériques arrangées dans un motif 7-2-7 centré en position -44.5. Seule la fixation de NarP-P permet d'activer efficacement la transcription. La fixation de NarL-P empêche la fixation de NarP-P et donc agit comme un antagoniste. A forte concentration en nitrate, les niveaux élevés de NarL-P, limitent la fixation des molécules de NarP-P. Il en résulte un faible niveau de transcription. A l'inverse, lorsque la concentration en nitrate est faible, la phosphorylation préférentielle de NarP active l'opéron *napF*. De la même façon, en présence de nitrites, l'activité phosphatase prédominante de NarX conduit à de faibles niveaux de NarL-P et donc l'activation par NarP-P de *napF* (Wang, Tseng et al. 1999). Une étude en 2003, a permis de préciser que l'opéron *napF* possédait deux promoteurs dont les -10 sont séparés de seulement 2 nucléotides (Stewart, Bledsoe et al. 2003). Le promoteur P1 est activé par le régulateur FNR et NarP-P alors que le promoteur P2 est très faiblement activé par FNR et réprimé par NarL-P. La figure 41 résume l'agencement des régions de contrôle de l'opéron *napF*.

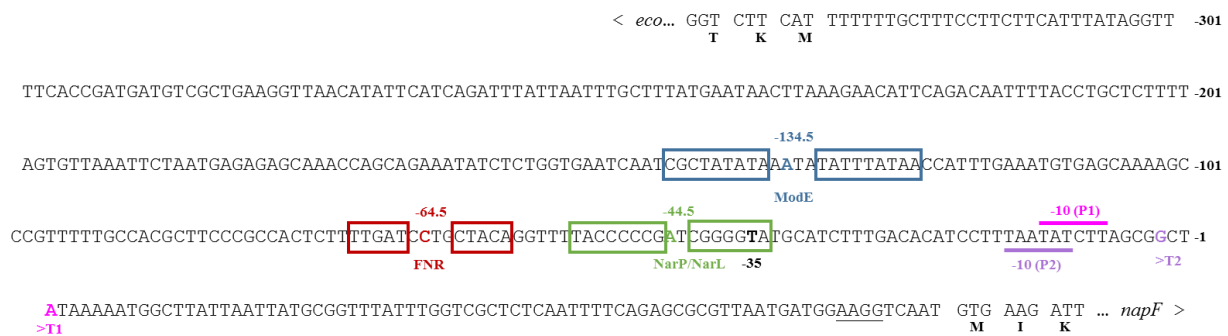


Figure 41 : Agencement des régions de contrôle de l'opéron *napF*. La numérotation est relative au +1 de transcription (T1). Les séquences de fixation de ModE, FNR, NarP/NarL sur *napF*. Les deux promoteurs, ainsi que leur -10 sont indiqués en rose et mauve ainsi que le -35 commun aux deux promoteurs. La séquence soulignée correspond au SD de *napF*. Les trois premiers codons et les acides aminés qui leur correspondent sont également indiqués pour *napF* et le gène *eco* transcrit de façon divergente.

1.3. Les ARN régulateurs bactériens

Durant de nombreuses années la communauté scientifique pensait que les ARN avaient pour unique rôle d'être un support intermédiaire de l'information génétique (ARNm) et d'assurer la synthèse protéique via les ARNt et les ARNr. La régulation des gènes était quant à elle dédiée aux protéines régulatrices pouvant agir au niveau transcriptionnel ou post-transcriptionnel. Même si dès 1961, Jacob et Monod ont avancé l'hypothèse selon laquelle des ARN pourraient jouer un rôle dans la régulation (Jacob and Monod 1961), il aura fallu attendre 1981 pour que le premier petit ARN régulateur, le RNA I, impliqué dans la régulation de la réplication des plasmides soit mis en évidence (Tomizawa, Itoh et al. 1981). Au cours des années qui ont suivi un nombre croissant de petits ARN régulateurs ont été découverts. Leur seule caractéristique commune étaient d'être de petite taille (<150 nts) et de ne posséder aucun cadre ouvert de lecture. Ils ont donc été regroupés sous le nom de petits ARN non codants (ARNnc). Aujourd'hui la terminologie utilisée pour parler de cette classe est plutôt sRNA pour « small RNA » (ou petit ARN) puisqu'il a pu être montré que certains petits ARN codaient pour de petits peptides. Les premiers petits ARN régulateurs ont été identifiés de manière fortuite. Par exemple, en 1997, alors que le laboratoire de Gisela Storz s'intéressait à l'adaptation au stress oxydatif des bactéries, son équipe a pu mettre en évidence l'accumulation spécifique d'un transcrit de 109 nts dans ces conditions particulières qui par la suite a été nommé OxyS (Altuvia, Zhang et al. 1998). Jusqu'au début des années 2000, une douzaine de petits ARN ont donc été mis en évidence par hasard, et ce n'est qu'avec l'arrivée du séquençage des génomes, que plusieurs équipes ont pu montrer qu'il y avait potentiellement bien plus de petits ARN pouvant s'avérer régulateurs. En 2001, trois travaux majeurs ont pu mettre en évidence une soixantaine de nouveaux petits ARN potentiellement régulateurs dans le génome d'*Escherichia coli*. (Argaman, Hershberg et al. 2001; Rivas, Klein et al. 2001; Wassarman, Repoila et al. 2001; Argaman, Elgrably-Weiss et al. 2012). Depuis, le développement des techniques d'identification (« tiling arrays », « RNA-sequencing ») a permis la découverte de centaines d'entre eux chez de nombreuses autres bactéries et archées. En parallèle de leur identification, de nombreuses études fonctionnelles ont mis en évidence la variété des processus biologiques dans lesquels peuvent intervenir des petits ARN comme le métabolisme du carbone, le quorum sensing ou encore la virulence mais également la diversité des mécanismes mis en jeu. Durant ma thèse j'ai d'ailleurs pu contribuer à la rédaction d'une revue permettant de résumer les différentes techniques

permettant d'étudier à la fois *in vivo* et *in vitro* les mécanismes de régulation mis en jeu entre les petits ARN et leur cible ARNm. Celle-ci se trouve en annexe 1 (Jagodnik, Brosse et al. 2017).

1.3.1. Classification des ARN régulateurs bactériens

Les ARN régulateurs bactériens sont un ensemble hétérogène de molécules impliquées dans la régulation de l'expression des gènes. Deux grands groupes se distinguent : ceux ayant une action en *cis* et ceux ayant une action en *trans*. L'action est dite « en *cis* » lorsque c'est la même molécule qui joue à la fois le rôle de régulateur et de cible. C'est le cas des riboswitches. A l'inverse, l'action est dite « en *trans* » lorsque le régulateur contrôle l'expression d'un gène autre que le sien. C'est le groupe le plus répandu. Il comprend les ARN régulateurs modulant l'activité d'une protéine (ex : 6S, CsrB) et les ARN agissant par appariement avec un ARNm-cible. Dans ce groupe, sont distingués les ARN codés en *cis* appelés aussi ARN antisens et ceux codés en *trans* (figure 42).

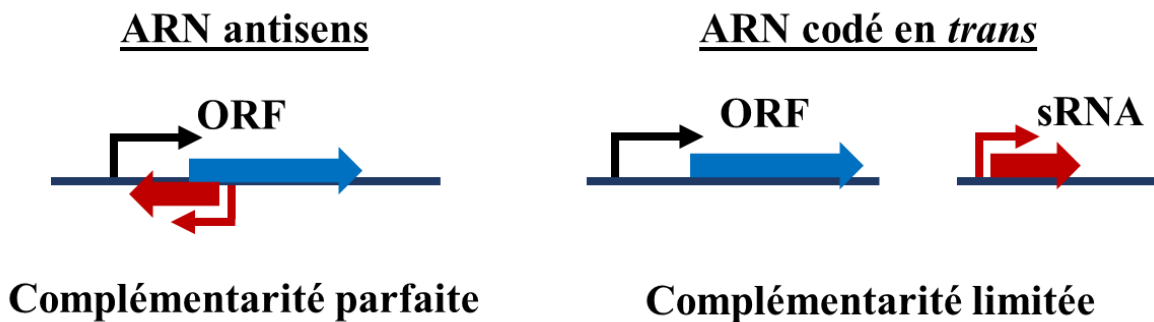


Figure 42 : Schéma des deux catégories d'ARN régulateurs agissant par appariement de séquences.

Par définition, les ARN antisens possèdent une complémentarité parfaite avec leur cible car ils sont codés sur le brin opposé à celle-ci. Les ARN codés en *trans* ne partagent quant à eux qu'une complémentarité limitée avec leur ARNm-cible.

Un cas particulier d'ARN régulateur agissant par complémentarité de séquence est apparu récemment : les ARN appartenant au système CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR associated protein). Ce système fascinant confère à certaines bactéries et archées une immunité adaptative envers les plasmides et virus que l'on peut apparenter au système immunitaire que nous connaissons chez l'homme (Sorek, Lawrence et al. 2013). Il s'agit de séquences répétées séparées par de courtes séquences nommées « spacers ». Les CRISPR sont transcrits sous forme d'un long précurseur ARN appelé pré-cr-RNA. Ce précurseur est clivé

par des endoribonucléases spécifiques donnant ainsi les cr-RNA qui sont les unités fonctionnelles. Les cr-RNA vont ensuite s'associer à des protéines de la famille Cas pour former des complexes de surveillance. Les cr-RNA sont capables de s'apparier à des ARN/ADN exogènes pour lesquelles ils possèdent une séquence complémentaire acquise lors d'une précédente infection. Ils vont ensuite diriger ces molécules exogènes vers les systèmes de dégradation. Le système le plus étudié provient de *Streptococcus pyogenes* pour lequel l'endonucléase est la Cas9. Il est intéressant de mentionner que les CRISPR sont utilisées actuellement comme outil de « genome editing » pour réaliser des modifications sur le génome telles que de l'extinction de gènes ou des mutations ponctuelles. Cette catégorie étant un peu à part, elle ne sera pas détaillée dans la suite de ce manuscrit. De même les riboswitches, les ARN modulant l'activité d'une protéine et les ARN antisens ne seront abordés que brièvement. En effet, j'ai choisi de me focaliser sur les ARN régulateurs codés en *trans* et agissant par appariement imparfait de séquence qui ont été au cœur de mon sujet de thèse.

1.3.1.1. Les riboswitches

Les riboswitches sont des structures présentes le plus souvent dans la région 5' non traduite de certains ARNm qui permettent de réguler l'expression génique en réponse à un ligand précis. Ce système de régulation a été montré pour un bon nombre de gènes, notamment ceux impliqués dans le métabolisme ou la capture de petits ligands tels que des acides aminés, des cofacteurs, des nucléotides ou encore des ions métalliques (Smith, Fuchs et al. 2010).

Les riboswitches s'organisent en général de la façon suivante : une séquence dite « aptamère », c'est-à-dire qui fixe le métabolite de façon extrêmement spécifique et avec une forte affinité, est située en amont de la « plateforme d'expression », dont la structure secondaire va être modifiée suite à la fixation du métabolite. Selon la nouvelle conformation adoptée il y aura soit terminaison, anti-terminaison, ou alors un contrôle post-transcriptionnel par inhibition ou activation de la traduction (figure 43).

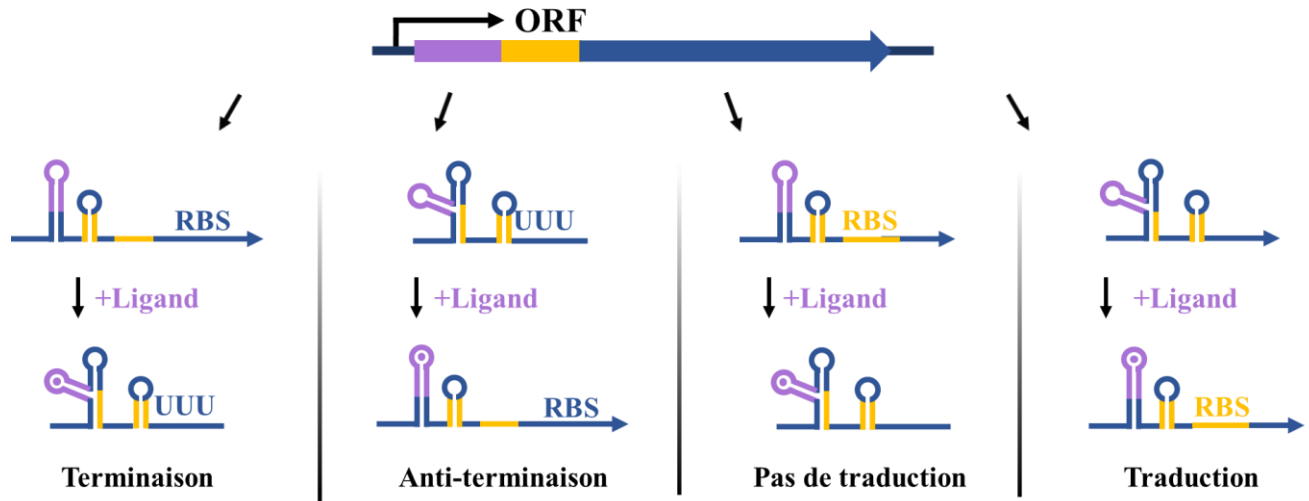


Figure 43 : Schéma d'un riboswitch et exemple de régulations associées. Les riboswitches sont composés d'une région aptamère (en violet) et une plateforme d'expression (en jaune) présente généralement dans le 5'UTR d'un ARNm (en bleu). La fixation d'un ligand spécifique peut aboutir à une régulation transcriptionnelle de la synthèse de l'ARNm (à gauche) où un contrôle traductionnel (à droite). D'après (Waters and Storz 2009).

Il est intéressant de noter qu'un mécanisme non canonique a été décrit par Winkler *et al* (Winkler, Nahvi et al. 2004) qui couple sensibilité au métabolite, grâce à un riboswitch, et clivage d'un ARN messager. Il s'agit du riboswitch-ribozyme *glmS*, mis en évidence chez la bactérie à Gram positif *B. subtilis*. Cet ARNc fixe la glucosamine-6-phosphate (GlcN6P) et ensuite clive l'ARNm *glmS* dans la séquence du ribozyme. La RNase J peut ensuite dégrader le produit de clivage en aval à partir de l'extrémité 5'-OH libérée ce qui empêche la traduction de l'ARNm (Collins, Irnov et al. 2007).

Une catégorie particulière de riboswitches est à mentionner : les ARN thermosenseurs qui répondent à des changements de température plutôt qu'à la fixation d'un ligand (Righetti and Narberhaus 2014).

1.3.1.2. Les ARN modulant l'activité d'une protéine

Cette famille d'ARN régulateurs a comme particularité de mimer le substrat ARNm ou ADN de la protéine qu'ils ciblent. En général le site de fixation est présent en multiples copies sur la séquence de l'ARN régulateur. Un exemple très bien caractérisé chez *E. coli* est le système Csr (Carbon Storage Regulator) qui est un régulateur global du métabolisme central. Il contrôle une grande variété de mécanismes adaptatifs incluant la formation de biofilms, la motilité ou encore la virulence. Le système Csr est composé de quatre éléments. Le composant principal est le

régulateur post-transcriptionnel CsrA, une protéine qui se lie à l'ARN et qui influe, le plus souvent négativement, sur la traduction. Le tableau 6 présente des exemples de cibles directes de la protéine CsrA identifiées chez *E. coli*.

Tableau 6 : Exemple de cibles directes de CsrA chez *E. coli*.

Gène cible	Fonction	Type de régulation	Référence
<i>glgCAP</i>	Métabolisme du glycogène	-	(Baker, Morozov et al. 2002)
<i>hfq</i>	Chaperonne à ARN	-	(Baker, Eory et al. 2007)
<i>pgaABCD</i>	Attachement des biofilms	-	(Wang, Dubey et al. 2005)
<i>flhDC</i>	Régulateur de la motilité	+	(Wei, Brun-Zinkernagel et al. 2001)

L'activité de CsrA est régulée négativement par deux ARN non codants CsrB et CsrC qui séquestrent ce régulateur. CsrB (366 nts) et CsrC (245 nts) possèdent respectivement 22 et 13 sites de fixation à CsrA. Ces sites de fixation sont définis par un motif « GGA » conservé situé au sommet d'une tige-boucle. En temps normal, CsrA se fixe à ses ARNm cibles au niveau de la séquence Shine Dalgarno, bloquant ainsi l'accès du ribosome et empêchant la traduction. Lorsque les ARN régulateurs CsrB et CsrC sont synthétisés, ils titrent de nombreuses molécules de CsrA déréprimant de ce fait les cibles précédemment régulées négativement par CsrA (figure 44).

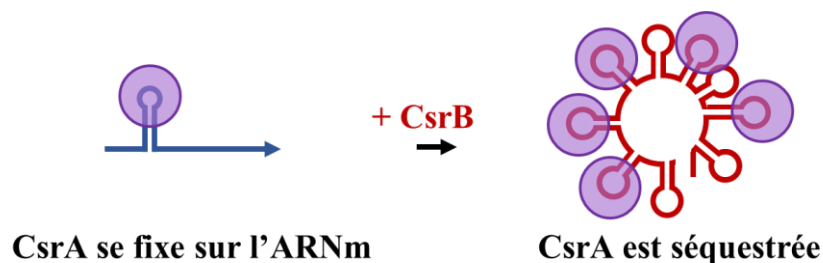


Figure 44 : Mode d'action des petits ARN CsrB et CsrC. La fixation de la protéine CsrA (en violet) sur des motifs GGA présents au sommet de tiges-boucles de certains ARNm altère l'expression de ce transcrit. Quand le niveau du petit ARN CsrB (ou CsrC) augmente, celui séquestre la protéine CsrA limitant son activité régulatrice. D'après (Waters and Storz 2009).

La transcription de CsrB et de CsrC est contrôlée par le système à deux composants BarA-UvrB (Suzuki, Wang et al. 2002). Les signaux conduisant à l'activation de ce système pourraient être

des acides faibles ou encore des intermédiaires du cycle de Krebs (Chavez, Alvarez et al. 2010). Par ailleurs, CsrA pourrait réguler directement ce système à deux composants ce qui constituerait un rétrocontrôle indirect des ARN CsrB et CsrC sur eux-mêmes. Il est intéressant de mentionner que CsrB et CsrC sont ainsi régulés négativement par la protéine membranaire CsrD qui stimule leur dégradation par la RNase E (Suzuki, Babitzke et al. 2006).

1.3.1.3. ARN antisens

Ces ARN sont transcrits sur le brin complémentaire au brin codant les gènes qu'ils régulent. Il y a donc une complémentarité parfaite entre l'ARNm cible et l'ARN régulateur. Les antisens ont d'abord été découverts chez les bactériophages, les plasmides ou dans les transposons où ils sont impliqués dans le maintien du nombre approprié de copies de l'élément mobile considéré. Une autre catégorie d'antisens se retrouve dans les systèmes toxine/antitoxine de classe I. De très nombreux antisens existent mais leur mode d'action est souvent mal caractérisé dans la plupart des cas (Thomason and Storz 2010). Le plus souvent, la régulation est négative, mais il existe des exemples de régulation positive. Classiquement, l'appariement de l'antisens avec l'extrémité 5' non traduite au niveau du site de démarrage de la traduction empêche la fixation du ribosome et donc inhibe la traduction de l'ARNm ciblé. D'autres mécanismes sont aussi possibles : interférence transcriptionnelle ou déstabilisation du messenger cible.

Un exemple original d'antisens est le petit ARN régulateur GadY puisqu'il stabilise son ARNm cible *gadX*, un activateur transcriptionnel impliqué dans la réponse à l'acidité (figure 45). Le gène *gadX* est codé en opéron avec *gadW* qui code une autre protéine impliquée dans la réponse acide. En absence de stress le transcrit *gadXW* est instable et rapidement dégradé. A l'entrée en phase stationnaire, GadY est transcrit sous contrôle de σ^S ou lors d'un stress acide sous celui de GadX. GadY est présent sous trois formes : longue (105 nts) issue du transcrit primaire et deux formes plus courtes (90 et 59 nts) (Opdyke, Kang et al. 2004; Tramonti, De Canio et al. 2008). GadY est codé dans la région intergénique entre *gad X* et *gadW* et va s'apparier parfaitement à la région 3' non traduite de GadX. La fixation de GadY sur le messenger conduit à un clivage du duplex impliquant notamment la RNase III (Opdyke, Fozo et al. 2011). Cette maturation va générer deux ARNm distincts plus stables et permettre la synthèse des protéines GadX et GadW. De façon intéressante, l'antisens GadY nécessite la présence de la chaperonne Hfq essentielle pour de

nombreux ARN codés en *trans* (Opdyke, Kang et al. 2004). La frontière entre les différentes catégories d'ARN régulateurs est parfois floue et GadY en est un bel exemple.

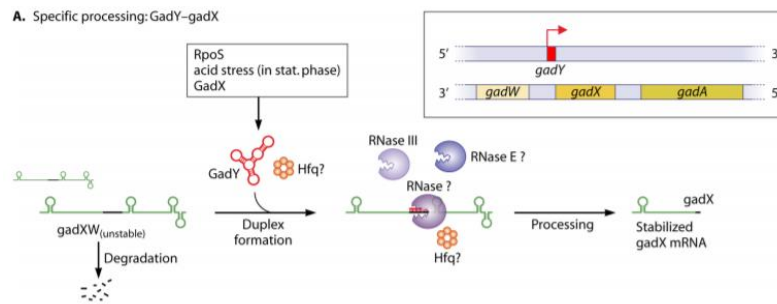


Figure 45 : Mécanisme d'action de GadY. Issue de (Georg and Hess 2011).

1.3.1.4. ARN agissant par complémentarité imparfaite de séquence

Les ARN régulateurs codés en *trans* jouent un rôle majeur dans les circuits de régulation. En effet, plusieurs exemples montrent qu'un petit ARN de ce type peut réguler plusieurs gènes issus de la même voie physiologique.

Ces ARN de 50 à 300 nts souvent très structurés sont généralement localisés dans les régions non codantes et conservées entre espèces bactériennes proches. Contrairement aux antisens, ils ne partagent qu'une complémentarité limitée avec leur ARNm cible ce qui laisse la possibilité à chaque ARN régulateur d'avoir plusieurs cibles et inversement qu'un ARNm puisse être ciblé par différents petits ARN.

En général, la région d'appariement couvre entre 10 et 25 nts mais seuls quelques nucléotides sont critiques pour la régulation. Par exemple, le petit ARN SgrS s'apparie avec l'ARNm *ptsG* permettant la régulation de l'import des sucres dans la cellule (Kawamoto, Koide et al. 2006). La région d'interaction s'étend sur 32 nts avec 23 nts appariés, pourtant seulement quatre mutations uniques dans *sgrS* affectent la régulation qu'il effectue sur *ptsG*.

La fixation de l'ARN régulateur à sa cible ARNm peut avoir diverses conséquences. Cette fixation peut soit inhiber, soit activer la traduction du message. En effet, en fonction de la structure secondaire initiale du message, et donc de l'emplacement du RBS la régulation peut être relativement différente. A la place ou en complément de ces régulations, il peut y avoir stabilisation du message ou à l'inverse dégradation du message en même temps que l'ARN régulateur (figure

46). Il est intéressant de préciser que la dégradation peut provenir uniquement de l'effet de blocage traductionnel.

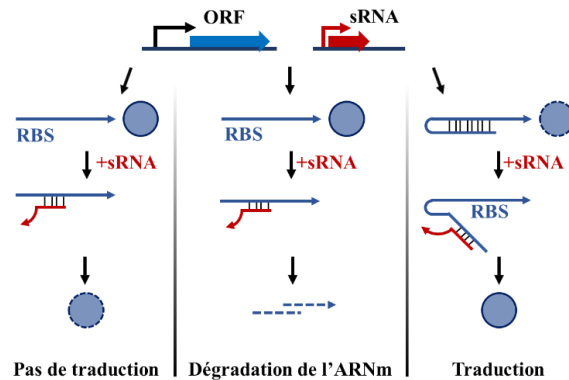


Figure 46 : Mécanismes de régulation des ARN agissant par complémentarité imparfaite de séquence. D'après (Waters and Storz 2009).

1.3.1.4.1. Effet sur la traduction

De manière classique, les ARN régulateurs peuvent se fixer à proximité du site de fixation du ribosome entrant directement en compétition avec celui-ci. Les ARN pouvant ainsi perturber directement la fixation du ribosome sur son site interagissent dans la région -30 à $+20$ (Huttenhofer and Noller 1994). Cependant, il a déjà été rapporté que des ARN se fixant dans la séquence codante ou relativement loin dans la séquence 5' non traduite pouvaient avoir comme conséquence d'inhiber la fixation du ribosome. En effet, chez *B. subtilis*, l'ARN SR1 interagit à de multiples sites dans la phase codante du messager *ahrC* dont le premier se situe entre $+81$ et $+87$ nts après l'ATG (Heidrich, Moll et al. 2007). Il a donc été proposé que la fixation de SR1 induise un changement conformationnel aboutissant à une modification de la fixation du ribosome. Un second exemple montre qu'une fixation dans la séquence codante peut avoir un effet sur la fixation du ribosome mais cette fois en empêchant la formation d'une tige-boucle activatrice (Jagodnik *et al.* en révision). Au laboratoire, il a en effet été montré que le messager *fepA* d'*E. coli* possédait une tige-boucle aidant le ribosome à se fixer au niveau du RBS et permettant ainsi une traduction efficace du messager. En inhibant la formation de cette tige-boucle, les petits ARN OmrA et OmrB sont ainsi capables de réprimer la traduction de *fepA*. Le 5' UTR lointain peut également être ciblé par les petits ARN. En effet, chez *E. coli*, GcvB se fixe relativement loin dans la région 5' non traduite des cibles *argT* et *gltI* respectivement entre -57 et -42 et entre -57 et -45 (Sharma, Darfeuille et al. 2007). GcvB agit selon un mécanisme assez original puisqu'il masque des

séquences riches en C/A s'avérant activatrices en absence de cet ARN. De plus, l'ajout de ces séquences en amont d'un gène non régulé par GcvB en condition sauvage aboutit à un contrôle de ce gène par GcvB. De même, la destruction de ces séquences fait perdre le contrôle par GcvB (Yang, Figueroa-Bossi et al. 2014).

Les ARN régulateurs peuvent également activer la traduction en se fixant dans la région 5' non traduite. Par exemple, les ARN ciblant *rpoS* que sont RprA, DsrA et ArcZ possèdent un site de fixation dans la région -119 à -95 (Mandin and Gottesman 2010). Leur fixation entraîne une modification structurale du messenger et libèrent le site de fixation du ribosome préalablement séquestré dans une tige boucle (figure 46).

L'action des ARN régulateurs sur leur messenger cible entraîne généralement une dégradation de ce dernier et comme nous le verrons dans certains cas il peut s'agir aussi du mécanisme primaire de régulation.

1.3.1.4.2. Effet sur la dégradation

Un des premiers petits ARN régulateurs Hfq-dépendant pour lequel l'effet sur la dégradation d'un messenger a été étudié en détail est RyhB qui participe à la coordination de la réponse de la cellule à une carence en fer. Il a été montré que RyhB induisait une dégradation rapide de ses cibles et était dégradé stochiométriquement avec celles-ci (Massé, Escorcía et al. 2003) ce qui suggère que les ARN régulateurs ne soient pas recyclés. Ainsi ce type de régulation peut présenter l'avantage d'une levée rapide de la régulation des cibles notamment lorsque les conditions environnementales sont modifiées soudainement (ex : sortie de stress). Plus généralement, une analyse par puce à ADN a mis en évidence que RyhB induisait une diminution du niveau d'au moins 18 ARNm codant des protéines utilisant du fer afin de rediriger celui-ci vers les processus cellulaires essentiels (Massé, Vanderpool et al. 2005). Ce mécanisme de dégradation rapide a par la suite été étudié pour un certain nombre de petits ARN et est une propriété qui a souvent été utilisée pour étudier les cibles d'ARN régulateurs.

Bien souvent, la dégradation est sans doute un effet secondaire du blocage traductionnel occasionné par les petits ARN. En effet, les ribosomes étant moins recrutés sur le messenger, ce dernier se trouve « dénudé » et donc potentiellement vulnérable à la dégradation. Néanmoins, la formation du duplex ARN régulateur/ARNm peut, de manière plus directe, entraîner la

dégradation de l'ARNm en recrutant une ou plusieurs RNases en particulier la RNase E et la RNase III (Bandyra, Said et al. 2012). Par exemple, chez *Salmonella*, la fixation de MicC environ 70 nts en aval du codon de démarrage d'*ompD*, n'a aucun effet sur la traduction (Pfeiffer, Papenfort et al. 2009). En revanche, sa fixation promeut un clivage par la RNase E initiant la dégradation du messager.

Une combinaison de ces deux mécanismes a également été suggérée. Il a notamment été montré que la fixation de RyhB sur l'ARNm *sodB* provoquait en plus d'un blocage traductionnel une dégradation par la RNase E indépendante de l'inhibition traductionnelle (Prevost, Desnoyers et al. 2011).

Pour les ARN Hfq-dépendant, un modèle où Hfq aurait pour rôle de recruter la RNase E, facilitant ainsi la dégradation du duplex a été proposé. En effet, Hfq a la capacité d'interagir avec l'extrémité C-terminale de la RNase E (Morita, Maki et al. 2005). Hfq fixerait donc cette région de la RNase E puis, en condition d'induction du petit ARN, fixerait l'ARN régulateur. Ce complexe interagirait avec l'ARNm cible favorisant sa dégradation (figure 47).

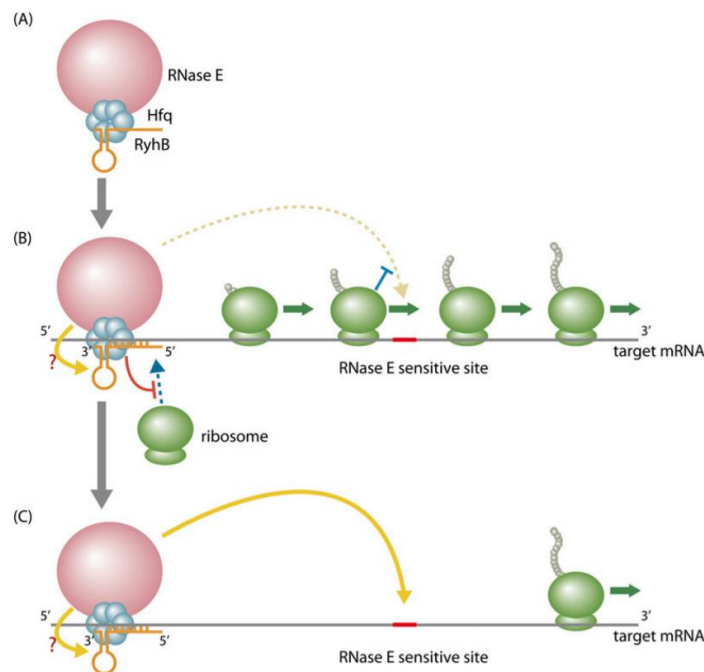


Figure 47 : Modèle de coopération entre petit ARN régulateur, Hfq et la RNase E pour la dégradation d'un messager. Issue de (Morita and Aiba 2011).

De façon intéressante, il a été proposé que la formation du duplex ARNm/ARN régulateur pourrait également stimuler activement la RNase E en plus de favoriser son recrutement (Bandyra, Said et al. 2012). L'interaction entre l'extrémité 5' monophosphate du petit ARN au niveau d'une « poche » en surface de la RNase E stimulerait l'activité catalytique de celle-ci.

Bien que moins fréquent, la fixation d'un ARN régulateur peut également stabiliser un ARNm limitant ainsi sa dégradation. Par exemple, l'isoforme longue de l'ARNm *cfa* de *Salmonella* est stabilisée par la fixation en aval du RBS de l'ARN régulateur RydC associé à Hfq interférant ainsi avec la RNase E qui dégrade normalement ce messenger (Frohlich, Papenfort et al. 2013). Autre exemple, chez *Streptococcus pyogenes*, l'ARN régulateur FasX est capable de stabiliser sa cible ARNm (Ramirez-Pena, Trevino et al. 2010). En effet, en absence de FasX, le messenger *ska* codant un facteur de virulence est rapidement dégradé. FasX en se fixant à l'extrémité 5' de *ska* génère une tige-boucle qui protègera l'ARNm d'une attaque par une RNase.

1.3.2. Focus sur les ARN Hfq-dépendants

Comme je l'ai mentionné précédemment dans l'introduction de cette partie, les premières découvertes d'ARN régulateurs se sont faites pour la plupart de manière fortuite. Par exemple, MicF, le premier petit ARN Hfq-dépendant mis en évidence, a été caractérisé au cours d'une étude portant sur les porines OmpC et OmpF. En surexprimant *ompC* et ses régions de contrôles Mizuno et al, ont observé que l'expression de la porine OmpF était réprimée (Mizuno, Chou et al. 1984; Andersen, Delilhas et al. 1987). Par la suite, ils ont pu montrer que MicF était un ARN de 93 nts situé dans la région intergénique entre *ompC* et *rcsD* et qu'il s'appariait au messenger d'*ompF*, inhibant de ce fait sa traduction. Il a fallu attendre les progrès dans le séquençage du génome puis du transcriptome pour que la plupart des ARN régulateurs dont les Hfq dépendants soient décrits. Au début des années 2000, plusieurs approches incluant des expériences de co-immunoprécipitation avec Hfq suivies d'un séquençage ont pu mettre en évidence des dizaines d'ARN régulateurs ouvrant ainsi un nouveau champ de recherche aussi vaste que passionnant (Argaman, Hershberg et al. 2001; Wassarman, Repoila et al. 2001; Vogel, Bartels et al. 2003; Zhang, Wassarman et al. 2003; Sittka, Lucchini et al. 2008).

1.3.2.1. La protéine Hfq

Hfq a été identifiée pour la première fois comme facteur essentiel à la réplication du brin d'ADN positif du phage Q β chez *E.coli* d'où son nom, acronyme de « Host Factor for phage Q β » (Blumenthal & Carmichael, 1979). Le gène *hfq* est situé au sein du superopéron *amiB-mulL-miaA-hfq-hflX-hflK-hflC* (Tsui, Leung et al. 1994). Sa transcription peut être activée par onze promoteurs dont trois spécifiques à *hfq*. On peut noter que l'expression d'*hfq* est réprimée au niveau post-transcriptionnel via la protéine régulatrice CsrA et Hfq elle-même (Vecerek, Moll et al. 2005; Ziolkowska, Derreumaux et al. 2006; Baker, Eory et al. 2007).

Le gène *hfq* est retrouvé dans la moitié des génomes eubactériens séquencés ce qui inclut de nombreux pathogènes. Il est intéressant de noter que certaines espèces en possèdent deux exemplaires comme *Bacillus anthracis*, l'agent de l'anthrax ou *Burkholderia cenocepacia* (Ramos, Sousa et al. 2011). Des homologues à Hfq sont retrouvés chez certaines archées et eucaryotes (Mura, Randolph et al. 2013). Bien que la séquence du gène *hfq* ne soit pas conservée, il existe néanmoins une conservation de la structure et de l'action de la protéine Hfq. En effet, des expériences de remplacement du gène *hfq* chez *S. enterica* par des homologues eubactériens (*Neisseria meningitidis*, *Aquifex aeolicus*) ou d'archées (*Methanocaldococcus jannaschii*) surproduits à partir de plasmides, ont montré la conservation des propriétés d'Hfq et en particulier sa capacité à fixer les petits ARN régulateurs (Sittka, Sharma et al. 2009). Il existe néanmoins des limites à cette conservation. Tout d'abord parce que les expériences ont été réalisées avec un système de surproduction. Il faudrait vérifier cela dans un contexte plus physiologique en effectuant un remplacement chromosomique. Ensuite, parce qu'il n'a pas été possible de remplacer *hfq* de *S. enterica* par le *hfq* de *S. aureus* (Rochat, Bouloc et al. 2012), ce qui montre qu'il y a probablement des limites au remplacement inter-espèces. De plus, on peut mentionner que le rôle d'Hfq chez les bactéries à Gram positif est à ce jour très discuté, notamment chez *S. aureus*, ce qui peut expliquer l'impossibilité de complémentarité entre le gène d'*hfq* provenant d'une bactérie à Gram positif dans le contexte génétique d'une souche à Gram négatif (Bouloc and Repoila 2016).

Une délétion du gène *hfq* génère des effets pléiotropiques s'expliquant notamment par le rôle crucial de cette protéine dans la régulation par des ARN en *trans*. Il a par exemple été montré chez *Clostridium difficile* qu'une souche délétée pour Hfq est affectée pour la croissance, la morphologie cellulaire mais aussi dans son aptitude à résister à divers stress (Boudry, Gracia et al.

2014). En revanche elle présente une meilleure capacité à sporuler et former des biofilms. De plus, chez certains organismes, Hfq se révèle être un facteur important dans la pathogénie. Il a été montré en 2008 qu'une délétion d'*hfq* provoquait une diminution de la capacité des Salmonelles à infecter des souris ou encore survivre dans des macrophages (Sittka, Pfeiffer et al. 2007). On peut cependant moduler cette observation dans la mesure où les mutants d'*hfq* possèdent de nombreux défauts fonctionnels.

D'un point de vue structural, Hfq est une petite protéine hexamérique de 11,2 kDa composée d'une hélice α N-terminale suivie de cinq feuillets β . Cette protéine cytosolique s'oligomérisse en homohexamères en forme de « doughnut » ce qui est la forme fonctionnelle de cette protéine (figure 48).



Figure 48 : Structure homohexamérique de la protéine Hfq. Représentation réalisée avec Pymol à partir de la structure PDB 2Y90. Chaque monomère est représenté dans une couleur différente.

Les deux faces du « doughnut » sont capables de fixer des molécules d'ARN. La face distale fixe des motifs ARN de type AAYAAAYAA (Y = C ou U) ou des triplets A-R-N (R= A ou G ; N = A/U/G/C) répétés avec une capacité de 18 nts par face distale d'un hexamère (Link, Valentin-Hansen et al. 2009; Lorenz, Gesell et al. 2010). Ce type de motifs est fréquemment retrouvé dans les régions 5' non traduites des ARNm cibles. La face proximale fixe quant à elle des séquences U ou A/U riches et les extrémités 3'OH augmentent fortement l'affinité pour cette face (Sauer and Weichenrieder 2011). Ceci suggère que les ARN Hfq-dépendants se fixent à Hfq préférentiellement par des séquences U riches ancrées dans un terminateur ce qui peut expliquer la protection des extrémité 3' contre les voies de dégradation (Otaka, Ishikawa et al. 2011). De plus, une étude récente montre que les régions 3'UTR pourraient s'avérer une source importante

de petits ARN Hfq-dépendants (Chao, Papenfort et al. 2012). Les auteurs ont notamment caractérisé la régulation de DapZ un petit ARN qui recouvre le 3'UTR de l'ARNm *dapB* chez *Salmonella*. Le deuxième exemple connu d'ARN de ce type est CpxQ (Chao and Vogel 2016). En condition de stress membranaire, le système à deux composants CpxA-CpxR active la transcription de la protéine CpxP. Ce messenger est processé par la RNase E et l'extrémité 3'UTR générée s'avère être le petit ARN CpxQ. L'action coordonnée de CpxP et de CpxQ protège ainsi la bactérie contre les dommages liés à la membrane interne.

Une étude récente a pu montrer que le rôle d'Hfq dans les interactions petit ARN/ARNm serait d'accueillir sur une face l'ARNm et sur l'autre l'ARN régulateur favorisant l'interaction de ces deux éléments (Zhang, Schu et al. 2013). Grâce à l'utilisation de mutants des faces proximales, latérales (Rim) et distales d'Hfq les auteurs ont pu démontrer que les ARN régulateurs pouvaient se fixer d'un côté ou de l'autre d'Hfq mais que chaque ARN avait une face préférentielle. Ainsi il est possible de séparer les ARN en deux classes (figure 49). De façon intéressante, les petits ARN se fixant sur la face proximale et le rim (classe I) sont plus sujet à une dégradation committante avec celle de leur cible alors que ceux se fixant sur les faces proximale et distale (classe II) ne seraient pas ou peu dégradés (Schu, Zhang et al. 2015). Dans les deux classes l'ARN régulateur interagit avec la face proximale car cette région reconnaît des régions riches en U donc le terminateur présent chez tous les petits ARN.

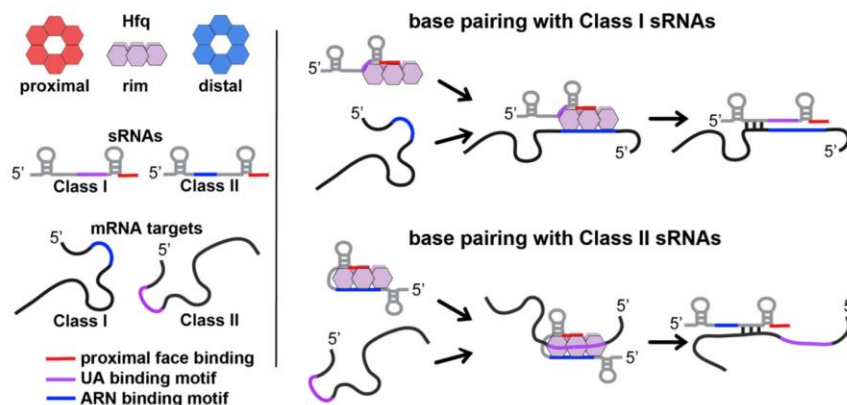


Figure 49 : Mode de fixation des ARN régulateurs sur la protéine Hfq. Les petits ARN de classe I se fixent préférentiellement sur la face proximale d'Hfq et le Rim. Leur messenger cible se fixe quant à lui sur la face distale. A l'inverse, les ARN de classe II se fixent en priorité sur la face proximale d'Hfq et la face distale. Le Rim est fixée par leur messenger cible dans ce cas. Issue de (Schu, Zhang et al. 2015).

Par ailleurs, chez certaines espèces, et en particulier chez les entérobactéries, le duplex formé par l'ARNm-cible et l'ARN codé en *trans* peut être favorisé et stabilisé par l'intervention de la protéine chaperon Hfq. En effet, en l'absence d'Hfq ces ARN sont extrêmement instables et rapidement dégradés. Une des hypothèses pour expliquer ce processus est que les sites de clivage de la RNase E et d'Hfq sont de nature proche (séquences A/U riches) (Moll, Afonyushkin et al. 2003). La fixation d'Hfq sur les petits ARN les protégerait d'une coupure par la RNase E.

Cependant, et bien qu'Hfq soit importante pour de nombreuses régulations, sa présence n'est pas requise chez toutes les bactéries. En effet, chez *Listeria monocytogenes*, même si le génome de cette bactérie possède une copie d'*hfq*, cette chaperonne n'est pas requise pour toutes les régulations en *trans*. En effet, les ARN régulateurs RliB, RliE et RliI ne nécessitent pas l'intervention d'Hfq alors que cette dernière est requise pour la régulation par LhrA/B/C chez cette même bactérie (Mandin, Repoila et al. 2007).

1.3.2.2. Rôles physiologiques des ARN Hfq-dépendants

Chez *E. coli*, seulement quelques ARN régulateurs sont exprimés de façon constitutive. Ainsi, et dans la majorité des cas leur transcription est contrôlée par des facteurs sigma alternatifs ou des régulateurs transcriptionnels, globaux ou spécifiques. On compte par exemple à ce jour trois petits ARN activés par le facteur sigma E (MicA, MicL et RybB) en condition de stress membranaire. Il a été montré par diverses études fonctionnelles qu'une adaptation au stress initial pouvait être observée (figure 50). Par exemple, l'accumulation dans le périplasma de protéines de membrane externe mal repliées va conduire, via sigma E, à l'activation des trois petits ARN précédemment cités. Ces derniers vont rapidement inhiber la synthèse de plus de 20 protéines de membrane externe dont les porines majoritaires ce qui aura donc pour conséquence de diminuer le stress initial en laissant le temps aux protéines déjà synthétisées de s'insérer correctement dans la membrane.

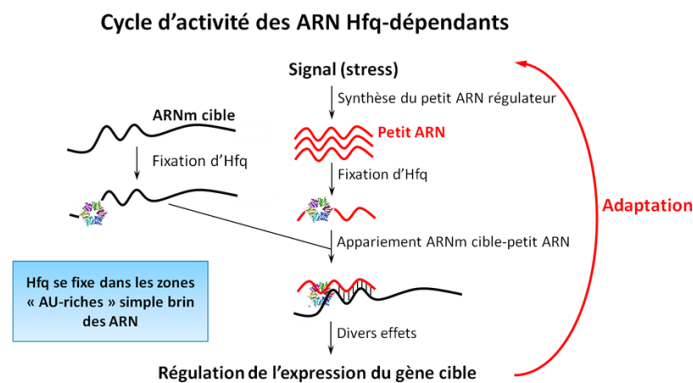


Figure 50 : Cycle d'activité des ARN Hfq-dépendants

Nous savons aujourd'hui que les ARN régulateurs interviennent dans de nombreuses voies de régulation allant du métabolisme du fer (RyhB) à celui des sucres (SgrS) en passant par le contrôle de la composition de la membrane externe (MicA, MicC, MicF, MicL, RybB, OmrA/OmrB). Le tableau 7 rassemble la plupart des ARN Hfq-dépendants connus chez *E. coli* à ce jour. Je détaillerai dans les parties suivantes le signal, les cibles et les mécanismes de régulation connus de trois petits ARN (OmrA, OmrB et RprA) qui ont été au cœur de mon projet de thèse.

Tableau 7 : ARN Hfq-dépendants présents chez *E. coli* utilisés au cours de cette étude

ARN	Signal inducteur	Fonction(s) régulée	Références
ArcZ (RyhA/SraH)	Faible concentration en oxygène, réprimé par ArcA	Active la traduction du facteur sigma S, import de la sérine, chimiotactisme	(Papenfort, Said et al. 2009; Mandin and Gottesman 2010)
ChiX (MicM/RybC/SroB)	Présence de sucres dérivés de la chitine	Répression de la synthèse de chitoporine ChiP, formation des biofilms	(Figuerola-Bossi, Valentini et al. 2009; Bak, Lee et al. 2015)
CpxQ	Stress de membrane	Insertion des protéines de membrane externe	(Chao and Vogel 2016; Grabowicz, Koren et al. 2016)
CyaR (RyeE)	Faible quantité de glucose, activé par CRP	Synthèse des protéines de membrane externe, quorum-sensing, formation des biofilms, motilité (OmpX, LuxS, NadE)	(Johansen, Eriksen et al. 2008; De Lay and Gottesman 2009; Richards and Vanderpool 2011)
DapZ	Condition d'infection	ABC transporteur	(Chao, Papenfort et al. 2012)
DicF	Non rapportée (gène de phage cryptique)	Inhibition de la division cellulaire	(Faubladier, Cam et al. 1990)
DsrA	Réponse à une baisse de la température, stress de membrane	Biosynthèse de la capsule, résistance au stress acide, active le facteur sigma S, réprime H-NS	(Sledjeski and Gottesman 1995; Majdalani, Cunnning et al. 1998; Lease and Belfort 2000; Repoila, Majdalani et al. 2003; Lease, Smith et al. 2004)
FnrS (RydD)	Induit en anaérobie, activé par FNR, ArcA et CRP	Réprime la synthèse des enzymes non nécessaires à la respiration anaérobie (SodB, MaE, GpmA, FolE, FolX)	(Durand and Storz 2010)
GadY (IS183)	Stress acide, régulé par sigma S	Résistance au stress acide	(Opdyke, Kang et al. 2004)
GcvB	Forte concentration en glycine, activé par GcvA et réprimé par GcvR	Répression du transport de peptides (OppA, DppA, GltI, LivK, ArgT, CycA, SsT), régule PhoQ-PhoP	(Urbanowski, Stauffer et al. 2000; Sharma, Darfeuille et al. 2007; Modi, Camacho et al. 2011; Sharma, Papenfort et al. 2011; Coornaert, Chiaruttini et al. 2013; Yang, Figuerola-Bossi et al. 2014)
GlmY (tke1/SroF)	Composés azotés	Biosynthèse des sucres aminés	(Urban, Papenfort et al. 2007)
GlmZ (RyIA/SraJ)	Sucres aminés, réprimé par les composés azotés	Métabolisme des sucres aminés, réparation de l'ADN	(Kalamorz, Reichenbach et al. 2007; Modi, Camacho et al. 2011)
IS118-RyeG	Non rapporté	Motilité, biofilm	(Bak, Lee et al. 2015)
McaS	Carence en source de carbone, activé par CRP	Motilité, biofilm et réparation de l'ADN	(Modi, Camacho et al. 2011; Thomason, Fontaine et al. 2012)
MgrR	Faible concentration en magnésium, activé par le système PhoQ-PhoP	Modification du LPS	(Moon and Gottesman 2009)
MicA (SraD)	Stress d'enveloppe, régulé par sigma E	Répression de la synthèse des porines OmpA/OmpX, modification du LPS, réponse au stress acide, transport du magnésium	(Rasmussen, Eriksen et al. 2005; Udekwa, Darfeuille et al. 2005; Papenfort, Pfeiffer et al. 2006; Udekwa and Wagner 2007; Coornaert, Lu et al. 2010)
MicC (ISO63)	Faible température, milieu pauvre, réprimé par le système EnvZ-OmpR	Répression de la synthèse des porines OmpC/OmpD	(Chen, Zhang et al. 2004; Pfeiffer, Papenfort et al. 2009)
MicF	Nombreux stress extérieurs/stress d'enveloppe, activé par le système EnvZ-OmpR mais aussi par SoxS et MarA, réprimé par Lrp	Répression de la synthèse de la porine OmpF et du régulateur global Lrp	(Delihias and Forst 2001; Holmqvist, Unoson et al. 2012)
MicL	Stress de membrane, régulé par sigma E	Régulation de la lipoprotéine Lpp	(Guo, Updegrove et al. 2014)
OmrA (RygA/SraE)	Haute osmolarité, activé par le système EnvZ-OmpR	Répression de protéines de membrane (CirA, FecA, FepA, OmpT, GntP) et du système EnvZ-OmpR, biofilm, motilité	(Guillier and Gottesman 2006; Guillier and Gottesman 2008; Holmqvist, Reimegard et al. 2010; De Lay and Gottesman 2012)
OmrB (RyGB)	Haute osmolarité, activé par le système EnvZ-OmpR	Répression de protéines de membrane (CirA, FecA, FepA, OmpT, GntP) et du système EnvZ-OmpR, biofilm, motilité	(Guillier and Gottesman 2006; Guillier and Gottesman 2008; Holmqvist, Reimegard et al. 2010; De Lay and Gottesman 2012)
OxyS	Stress oxydatif, activé par OxyR	Métabolisme du formate, réponse globale au stress	(Altuvia, Weinstein-Fischer et al. 1997)
RprA	Stress de membrane, activé par le phosphorelais Rcs	Active le facteur sigma S, Motilité, biofilm	(Majdalani, Chen et al. 2001; Majdalani, Hernandez et al. 2002; Majdalani and Gottesman 2005; Mika, Busse et al. 2012; Papenfort, Espinosa et al. 2015)

RseX	Non rapporté	Répression de la synthèse de porines (OmpC, OmpA)	(Douchin, Bohn et al. 2006)
RybB	Stress d'enveloppe, régulé par sigma E	Répression de la synthèse protéines de membrane externe ou de lipoprotéines (OmpA, OmpC, OmpF, OmpW...)	(Papenfort, Pfeiffer et al. 2006; Thompson, Rhodius et al. 2007; Gogol, Rhodius et al. 2011)
RydC	Non rapporté	Répression d'ABC transporteurs putatifs, motilité	(Bordeau and Felden 2014)
RyhB (SraI)	Carence en fer, réprimé par Fur	Répression des protéines utilisant du fer	(Massé and Gottesman 2002; Salvail and Masse 2012)
SdsN	Induit par divers composés azotés ou la faible température	Régulation des enzymes utilisant des composés azotés, réprime le régulateur NarP	(Vanderpool and Gottesman 2004; Hao, Updegrave et al. 2016)
SdsR ou RyeB	Régulé par sigma S	Chez <i>E. coli</i> réprime MutS ; chez <i>Salmonella</i> régule des protéines membranaires (ex : OmpD), des enzymes et des régulateurs (ex : CRP)	(Frohlich, Papenfort et al. 2012; Gutierrez, Laureti et al. 2013; Frohlich, Haneke et al. 2016)
SgrS (RyaA)	Accumulation de glucose-phosphate	Métabolisme des sucres phosphates	(Vanderpool and Gottesman 2004; Richards and Vanderpool 2011)
Spot42 (Spf)	Forte quantité de glucose, réprimé par CRP	Régulation des sources de carbones alternatives	(Moller, Franch et al. 2002; Beisel and Storz 2011; Richards and Vanderpool 2011)

1.3.2.3. Circuits mixtes de régulation

L'expression des ARN régulateurs est le plus souvent contrôlée au niveau transcriptionnel par des facteurs sigma ou des facteurs de transcription (Mandin and Guillier 2013). En retour, ces mêmes facteurs sont la cible de nombreux petits ARN. Ainsi des circuits mixtes de régulation complexes sont générés au sein de la cellule permettant à celle-ci de s'adapter au mieux aux conditions environnementales (Beisel and Storz 2010). Au cours de ma thèse je me suis particulièrement intéressée aux circuits mixtes entre les ARN Hfq-dépendants et régulateurs de systèmes à deux composants. Quelques exemples ont par le passé montré l'interconnexion entre ces deux modes de régulation.

Par exemple, PhoP active l'expression de l'ARN régulateur MgrR (Moon and Gottesman 2009). Par la suite, il a été montré au laboratoire que le système PhoQ-PhoP était ciblé par les deux petits ARN régulateurs MicA et GcvB (Coornaert, Lu et al. 2010; Coornaert, Chiaruttini et al. 2013). Dans ce cas de figure, la régulation est à sens unique entre les différents opérateurs et on parle en général de signal à simple module ou « single-input module » (figure 51A). Des circuits plus complexes ont également pu être mis en évidence. Par exemple, nous avons vu précédemment que le système à deux composants EnvZ-OmpR régulait la transcription des ARN régulateurs MicC et MicF ainsi que la synthèse des porines OmpC et OmpF. De façon intéressante il a pu être montré que le petit ARN MicC réprimait l'expression d'*ompC* (Chen, Zhang et al. 2004) et que MicF réprimait quant à lui l'expression d'*ompF* (Delilhas and Forst 2001). Nous arrivons donc à un système où une cible se trouve être régulée par un facteur de transcription à la fois de manière

directe, mais également de manière indirecte, via un ARN régulateur agissant au niveau post-transcriptionnel. On parle alors de « feedforward loop » (figure 51B) Dans le cadre de cet exemple, la boucle est dite « logique » ou « cohérente » car les régulations ont la même finalité (activation d'*ompC* et répression d'*ompF*). Dans un certain nombre de cas comme nous l'avons vu avec l'exemple de sigma E et des ARN MicA, MicL et RybB on peut observer une adaptation au stress initial. On parle dans ce cas de boucle de rétrocontrôle ou « feedback loop » (figure 51C). On observe par exemple une boucle de rétrocontrôle directe entre le petit ARN MicF et le régulateur Lrp chez *E. coli* (Holmqvist, Unoson et al. 2012). Elle peut aussi être indirecte comme avec sigma E. Pour finir, les différentes régulations peuvent se croiser et former un régulon relativement dense (figure 51D) que l'on retrouve sous le terme « dense overlapping regulon ».

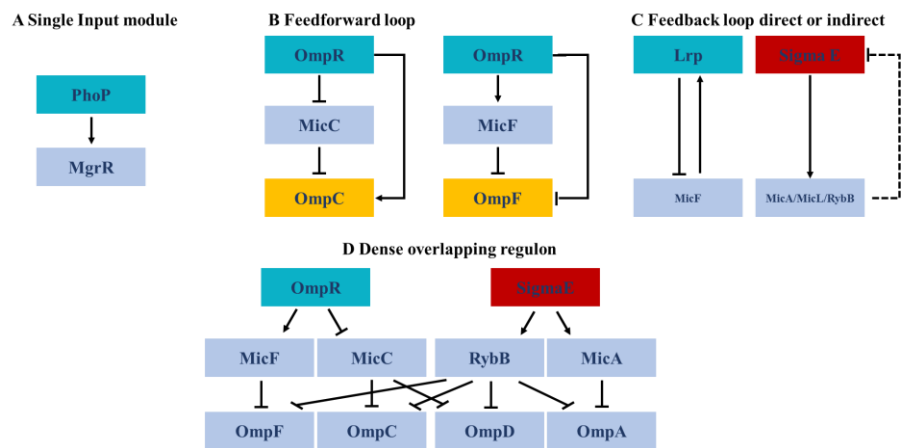


Figure 51 : Différents motifs retrouvés dans les circuits de régulation. A. Dans le cas du « single input motif » la régulation est à sens unique. B. Dans le motif « feedforward loop » un régulateur peut avoir un effet sur sa cible direct mais également indirect en passant par un intermédiaire, ici un petit ARN. C. Dans les « feedback loop » deux éléments régulateurs peuvent se contrôler mutuellement de façon directe comme avec MicF et Lrp où indirectement comme dans le cas des petits ARN Sigma E dépendants. D. Le croisement entre ces différents types de régulation aboutit à un circuit complexe où chacun des éléments peut réguler et être régulé par divers composants (ARN régulateur, facteur de transcription ou facteur sigma).

Ce type de motifs ainsi que leurs propriétés ont par le passé été décrits pour des réseaux entièrement protéiques (Alon 2007) mais les propriétés des circuits mixtes de régulation sont encore mal comprises à ce jour. Néanmoins différentes études mettent en avant les potentiels avantages et les propriétés uniques des systèmes intégrant des ARN régulateurs.

Tout d’abord, la réduction du coût métabolique des régulations médiées par les ARN a souvent été avancée (Altuvia and Wagner 2000; Shimoni, Friedlander et al. 2007). En effet, l’énergie nécessaire pour transcrire des ARN d’environ 100 nts est bien moins importante que celle dépensée

pour générer des régulateurs protéiques de plusieurs kilodaltons. Cet argument peut cependant être modulé selon l'ARN régulateur et la cible considérée (Beisel and Storz 2010). En effet, dans le cas où un seul ARN régulateur contrôle le niveau de nombreuses cibles différentes, alors qu'il aurait fallu plusieurs régulateurs protéiques pour ces mêmes cibles, l'argument énergétique est tout à fait valable. En revanche, si la cible de l'ARN régulateur est un gène très fortement transcrit, il faudra de nombreuses molécules de régulateurs afin de limiter la traduction de la cible. Dans ce cas précis l'arrêt de la transcription par un facteur protéique est une alternative moins coûteuse énergétiquement.

Le second argument avancé en faveur de la régulation par des ARN est que ce type de régulation peut être plus rapide. En effet, comme les ARN régulateurs agissent post-transcriptionnellement, il leur faut bien moins de temps pour moduler le niveau de protéine dans la cellule qu'un facteur de transcription (Beisel and Storz 2010). De plus, les petits ARN sont bien souvent dégradés avec leur cible, le retour à l'état initial est ainsi plus rapide qu'avec une protéine régulatrice possédant un temps de demi-vie plus long (figure 52).

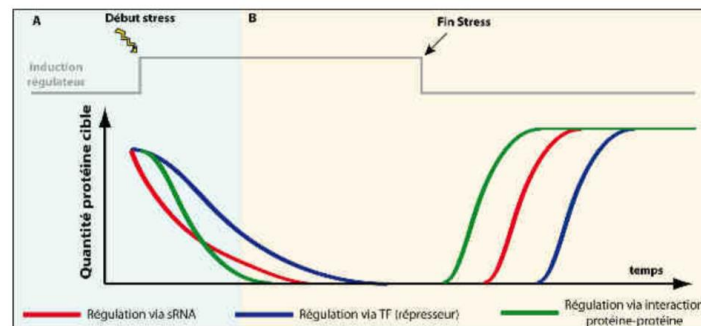


Figure 52 : Simulation de vitesse de réponse au stress. Issue de (Romilly 2012).

En plus d'être très rapide, les régulations ARN sont également très sensibles. Il a été suggéré que lorsque ce type de régulation domine cela puisse limiter les variations entre les cellules du fait du « bruit de fond » limité dans ce type de contrôle (Beisel and Storz 2010). Cette proposition se fonde sur le fait qu'en dessous un certain seuil, où le taux de synthèse de l'ARNm cible est bien plus faible que celui de l'ARN régulateur, la répression est totale. Au-delà de ce seuil, fixé par le taux de transcription de l'ARN régulateur, le niveau de la cible augmente de façon linéaire et le bénéfice de cette régulation s'arrête (figure 53). Ce type de régulation donne la possibilité à la cellule de discriminer entre un signal environnemental transitoire ou un signal plus pérenne dont

le but est d'aboutir à l'expression d'un gène donné. Cette régulation par les petits ARN permet ainsi de suivre plus finement les variations et s'y adapter en conséquence.

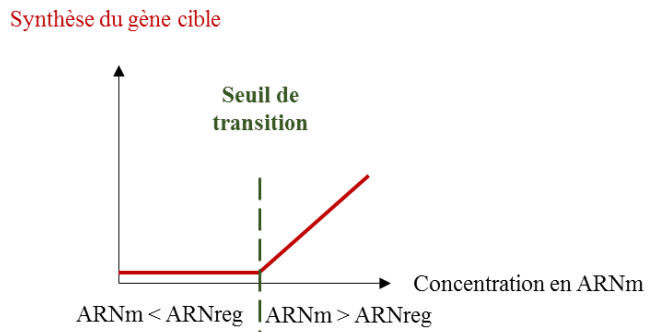


Figure 53 : Modèle de régulation basé sur des ARN. Dans le cas de la répression par des ARN, le contrôle maximal est effectif dès le début du signal. Le niveau de transcription de l'ARNm cible doit dépasser celui de l'ARN régulateur pour outrepasser ce contrôle.

Le troisième argument, mettant en avant l'avantage d'une régulation par des ARN est la capacité à acquérir facilement de nouvelles régulations (Gottesman and Storz 2011; Mandin and Guillier 2013). En effet, les ARN sont des molécules qui évoluent plus rapidement et comme la modification de seulement quelques nucléotides peut suffire pour acquérir un nouvel ARNm-cible c'est donc un avantage non négligeable pour un organisme soumis à des variations constantes et parfois brutales de son environnement. Il est intéressant de constater que bien souvent les régulations par des ARN se limitent à un groupe phylogénétique relativement réduit. Par exemple, la présence des Omr se limite essentiellement au groupe des enterobactériacées.

Jusqu'à présent j'ai exposé les avantages des régulations par des ARN par rapport à la régulation protéique. Cependant l'association de ces deux types est également un apport très important des circuits mixtes de régulation.

Comme nous l'avons vu avec l'exemple des « feedforward loop », un régulateur peut par deux voies (directe ou indirecte) agir sur une de ses cibles. Dans certains cas, le circuit est dit « cohérent » et donc quelle que soit la voie empruntée l'issue est la même. En revanche ce n'est pas toujours le cas ! Par exemple, il a été montré qu'OmpR activait directement la transcription de *csgD*. Cependant OmpR active également la transcription des Omr qui eux inhibent *csgD*. On comprend aisément qu'un aspect cinétique de régulation voit le jour et les propriétés de ces deux voies de régulation sont sans doute différentes pour coexister. Ce type de motifs a déjà été étudié

pour des circuits uniquement protéiques mais il serait intéressant de déterminer si les propriétés sont différentes lorsque l'on inclut des ARN régulateurs.

Par ailleurs, les régulateurs protéiques et les petits ARN d'un même circuit peuvent répondre à des stimuli cellulaires différents. Ainsi intégrer l'ensemble des stimuli afin de coordonner les différents processus cellulaires est probablement un avantage très important des circuits mixtes. Prenons l'exemple du changement de mode de vie entre des bactéries motiles à la recherche de nutriments et d'autres bactéries s'assemblant en biofilm dans une niche écologique riche. Ces deux styles de vie ne nécessitent pas les mêmes protéines de surface, ni les mêmes dépenses énergétiques. On peut donc penser qu'il soit nécessaire que la synthèse des régulateurs majeurs des biofilms et de la motilité qui sont respectivement codés par *csgD* et *flhD* soit finement régulée à la fois par divers facteurs de transcriptions (CRP, OmpR, RcsB ...) issus de systèmes à deux composants ou non mais également par de nombreux petits ARN (OmrA, OmrB, McaS ...). Ces nombreux niveaux de contrôle répondant chacun à des stimuli différents et avec des cinétiques de réponse diverses permettent ainsi à la cellule de choisir le mode de vie le plus adapté aux conditions environnementales dans lesquelles elle évolue.

Pour finir, dans le cadre de la régulation des systèmes à deux composants, par des ARN régulateurs notamment, il est important de préciser que bien que les ARN puissent affecter le niveau du régulateur, le niveau de la forme active de celui-ci peut ne pas être modulé. C'est ce qui est décrit par la robustesse d'un système. Il est donc possible de moduler un régulateur sans affecter toutes ses cibles. J'ai particulièrement étudié cet aspect dans le cadre de l'étude du système EnvZ-OmpR et sa régulation par les petits ARN OmrA/B.

1.3.3. Les petits ARN OmrA et OmrB

1.3.3.1. Découverte des Omr

OmrA et OmrB sont des petits ARN non codants Hfq-dépendants identifiés chez *E. coli* dans trois approches globales basées sur la conservation entre entérobactéries dans des régions intergéniques ainsi que sur la co-précipitation avec Hfq et l'expression dans certaines conditions de croissance (Argaman, Hershberg et al. 2001; Wassarman, Repoila et al. 2001; Vogel, Bartels et al. 2003). Ces ARN avaient été nommés RygA (ou SraE) et RygB dans des études antérieures.

omrA et *omrB* sont adjacents et localisés dans la région intergénique entre les gènes *aas* et *galR* chez *E. coli* codant respectivement une enzyme impliquée dans le recyclage des phospholipides membranaires et le répresseur transcriptionnel associé au transport et catabolisme du D-galactose (figure 54). OmrA est un ARN de 88 nts et OmrB un ARN de 82 nts. Ils sont faiblement exprimés en milieu minimum et en LB. Leur expression maximale est atteinte en phase stationnaire pour OmrA et en phase exponentielle et stationnaire pour OmrB (Argaman, Hershberg et al. 2001; Wassarman, Repoila et al. 2001; Zhang, Wassarman et al. 2003). De plus, le niveau d'OmrB diminue en fin de phase stationnaire (Vogel, Bartels et al. 2003) alors qu'OmrA s'accumule (Argaman, Hershberg et al. 2001).

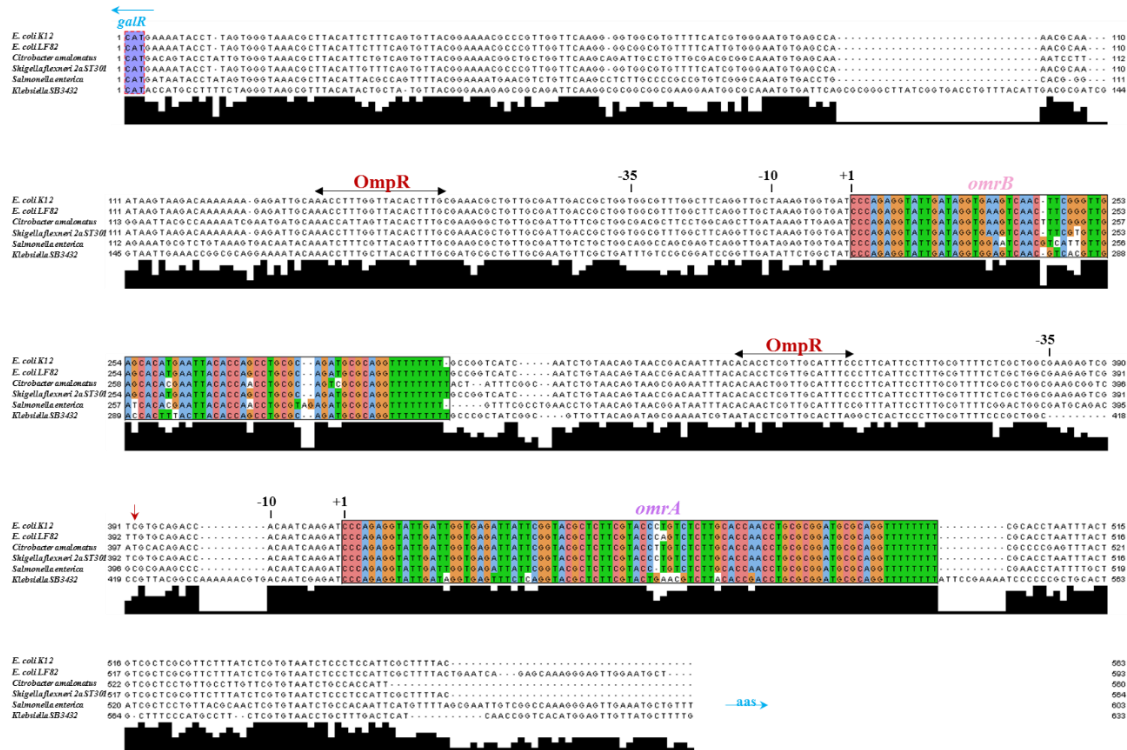


Figure 54 : Organisation et conservation de la région chromosomique *omrA-omrB*. Alignements des gènes *omrA* et *omrB* entre espèces proches effectué avec Clustal W et édité avec Jalview. La conservation des Omr est mise en évidence par le code couleur clustalx. L'ATG du gène *galR* est surligné en mauve. Les sites de fixation d'OmpR prédits sur le promoteur d'*omrA* et d'*omrB* sont indiqués ainsi que les positions -35, -10 et +1 pour chacun des promoteurs. La flèche rouge indique l'unique nucléotide de différence entre la séquence du promoteur d'*omrA* d'*E. coli* K12 et celle d'*E. coli* LF82.

Il est intéressant de préciser que les extrémités 5' et 3' des Omr sont très conservées au sein des entérobactéries mais aussi entre OmrA et OmrB (figure 54). En effet, les 21 premiers nucléotides et les 35 derniers nucléotides des Omr sont pour la plupart identiques, suggérant une importance fonctionnelle de ces régions. La région centrale est quant à elle différente entre les deux Omr.

1.3.3.2. Régulation des Omr par EnvZ-OmpR

A la suite de la découverte de ces petits ARN s'est immédiatement posée la question : quand sont exprimés ces ARN et que font-ils ? Pour comprendre le profil d'expression d'OmrA et OmrB, Guillier et Gottesman (Guillier and Gottesman 2006) ont construit des fusions *PomrA-lacZ* ou *PomrB-lacZ* afin de réaliser un crible génétique. Les souches obtenues sont Lac⁻ (blanches sur milieu Mac ConKey lactose) et elles ont ensuite été transformées par une banque d'ADN génomique d'*E. coli* (banque décrite dans (Ulbrandt, Newitt et al. 1997)). Les transformants présentant un phénotype Lac⁺ (rosés à rouges sur milieu Mac ConKey lactose) ont été sélectionnés avec l'idée que les plasmides qu'ils portent surproduisent un activateur putatif des Omr ou qu'ils titrent un répresseur de ces ARN. Les inserts des plasmides issus des colonies Lac⁺ ont donc été séquencés. Un plasmide (pOmpR) portait la région intergénique *greB-ompR* où se situe la région promotrice d'OmpR, la totalité d'*ompR* mais également les 901 premiers nucléotides d'*envZ* permettant ainsi la traduction d'une protéine de fusion portant les 300 premiers acides aminés d'EnvZ suivis de 43 acides aminés issus de la traduction d'une portion du vecteur. Les autres plasmides isolés lors de ce crible portaient d'autres régions du chromosome d'*E. coli* mais il a été montré que tous agissaient via EnvZ-OmpR.

Comme il était connu que le système à deux composants EnvZ-OmpR était activé par l'osmolarité, les fusions *PomrA-lacZ* ou *PomrB-lacZ* ont été testées en milieu à forte osmolarité, obtenue par addition de sucrose ou mimé par de la procaïne (anesthésique décrit comme activateur d'EnvZ-OmpR) dans le milieu de culture. Par ces expériences, les auteurs ont pu montrer que les fusions *PomrA-lacZ* ou *PomrB-lacZ* étaient activées à haute osmolarité. L'expression de ces fusions était indétectable dans des mutants $\Delta envZ$ ou $\Delta ompR$ à basse ou haute osmolarité confirmant la régulation des Omr par le système à deux composants EnvZ-OmpR (figure 55A et B). De plus, par un Northern blot il a été montré qu'en présence de procaïne, OmrA et OmrB étaient très rapidement induits suggérant une régulation directe par OmpR (figure 55C).

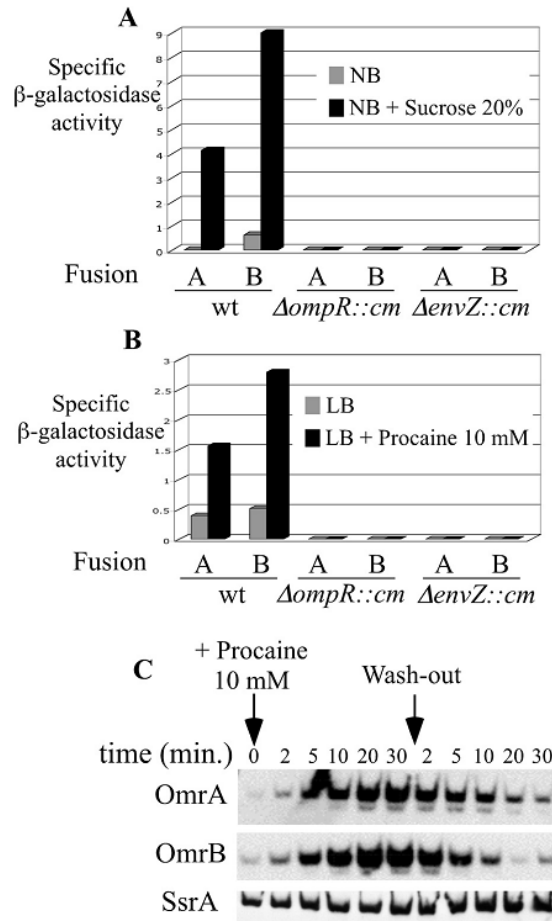


Figure 55 : Le système à deux composants EnvZ-OmpR active la transcription d’*omrA* et *omrB*. (A) Dosage β-galactosidase de fusions *PomrA-lacZ* (fusion A) ou *PomrB-lacZ* (fusion B) dans un contexte sauvage (WT), mutant pour EnvZ ($\Delta envZ$) ou mutant pour OmpR ($\Delta ompR$) avec une culture en NB ou en NB+sucrose. (B) Dosage β-galactosidase de fusions *PomrA-lacZ* ou *PomrB-lacZ* dans un contexte sauvage (WT), mutant pour EnvZ ($\Delta envZ$) ou mutant pour OmpR ($\Delta ompR$) avec une culture en milieu LB ou en LB+procaine. (C) Analyse par Northern Blot d’*OmrA*, *OmrB* et *SsrA*(contrôle de charge) dans une souche sauvage à différents temps après induction avec de la procaine puis lavage des cellules. Issue de (Guillier and Gottesman 2006).

1.3.3.3. Cibles des Omr

Dans l’étude de Guillier & Gottesman en 2006, il a été montré, par une approche transcriptomique dans un premier temps puis plus directement par Northern Blot, que les Omr régulaient négativement l’expression de multiples gènes codant des protéines de la membrane externe telles qu’*ompT*, *cirA*, *fecA* et *fepA* (Guillier and Gottesman 2006). *ompT* code une protéase alors que *cirA*, *fecA* et *fepA* codent des récepteurs TonB dépendants impliqués dans le transport du fer qui sont également des récepteurs pour les phages et les colicines. En 2008, il a également été montré que la région 5’ conservée des Omr se fixait aux ARNm-cible *ompT* et *cirA* dans la région de

démarrage de la traduction (Guillier and Gottesman 2008). Cette région 5' est également requise pour la régulation de toutes les cibles des Omr décrites jusqu'à ce jour. De façon intéressante, il a également été montré que ces ARN étaient capables de réprimer l'expression de l'ARNm *ompR-envZ*, créant ainsi une boucle de rétrocontrôle.

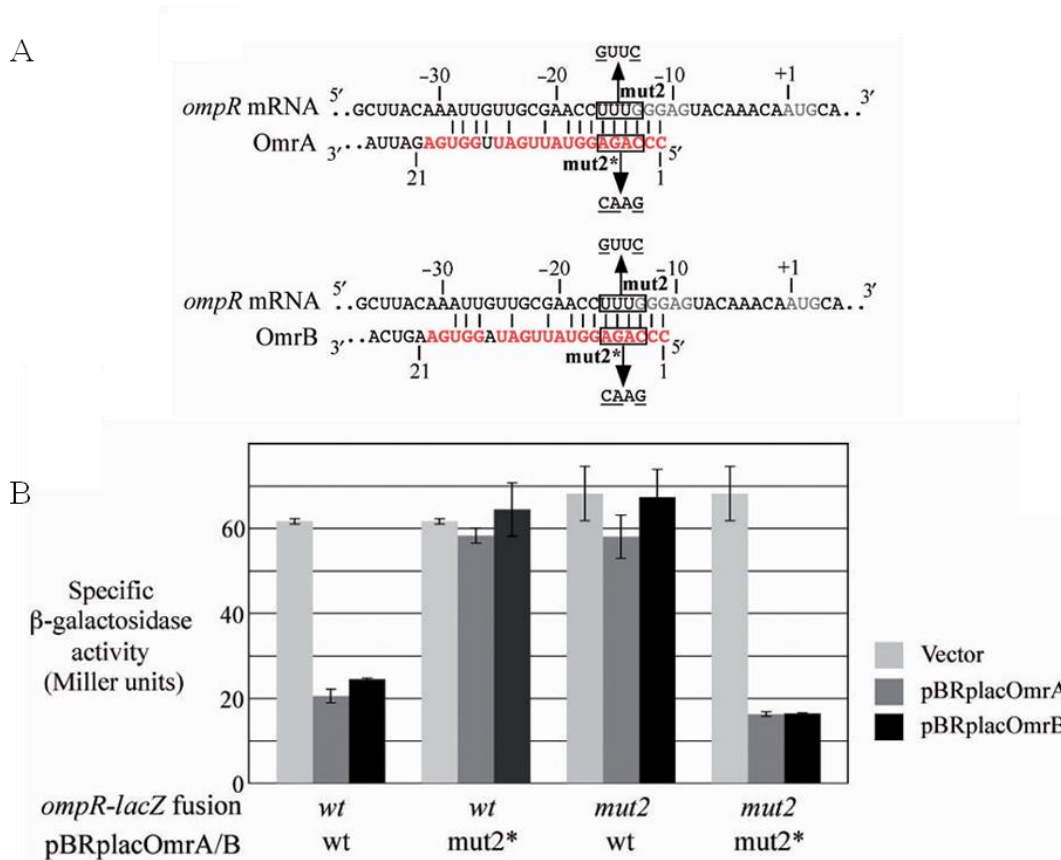


Figure 56 : L'extrémité 5' d'OmrA/B est impliquée dans la régulation de multiples cibles, incluant *ompR*. (A) Appariement prédit entre le messenger *ompR* et OmrA/B. Les nucléotides en rouge sont conservés entre OmrA et OmrB, ceux en gris correspondent au codon de démarrage et au Shine Dalgarno d'*ompR*. (B) Effet des mutations compensatoires dans l'ARNm *ompR-lacZ* et/ou les ARN régulateurs OmrA/B sur l'activité d'une fusion traductionnelle portée par le pBAD-*ompR-lacZ*. Issue de (Guillier and Gottesman 2008).

Ce rétrocontrôle est obtenu par un appariement direct des Omr sur le messenger *ompR-envZ*. En effet, comme cela est présenté sur la figure 56, lorsqu'OmrA ou OmrB sont surexprimés à partir d'un plasmide, l'expression d'une fusion *ompR-lacZ* est réprimée. Lorsqu'un mutant (*mut2*) des zones d'appariement prédit est utilisé, la répression de la fusion est perdue. De plus, les mutants compensatoires, introduits dans la fusion *ompR-lacZ*, sont réprimés par la version mutante des Omr, mais pas par la version sauvage. Cette expérience permet de valider l'appariement prédit entre le messenger *ompR-envZ* et les Omr et confirme donc l'action directe de ces derniers. Notons

que ce résultat était relativement suprenant puisqu'il avait précédemment été décrit que la modulation du niveau d'OmpR et d'EnvZ n'affectait pas le niveau de la forme active de la protéine, OmpR-P (Batchelor and Goulian 2003).

Une autre cible des Omr a été identifiée en 2010 par approche bioinformatique : *csgD*, le facteur de régulation majeur des fibres curli (Holmqvist, Reimegard et al. 2010). Ces fibres permettent aux bactéries de s'agréger entre elles et d'adhérer aux surfaces (verre, plastique, cellules végétales ou animales ...). Elles sont notamment cruciales pour la colonisation primaire des surfaces et dans les étapes de formation du biofilm notamment en augmentant les interactions entre les cellules (Prigent-Combaret, Prensier et al. 2000). Comme dans le cas d'*ompR-envZ*, il s'agit d'un appariement direct entre l'extrémité 5' des Omr et l'ARNm de *csgD* mais l'originalité de cette régulation réside dans la région d'appariement impliquée sur l'ARNm de *csgD* (figure 57). En effet, en s'appariant 70 nts en amont du site de démarrage de la traduction de *csgD*, les Omr modifient la structure de ce dernier (Holmqvist, Reimegard et al. 2010). Le mécanisme précis inhibant le démarrage de la traduction reste cependant encore inconnu. Les Omr peuvent aussi potentiellement activer la synthèse de CsgD par un mécanisme indirect passant par YdaM, une diguanylate cyclase (Mika and Hengge 2014). En effet, la surexpression des Omr diminuerait l'expression d'une fusion *ydaM-gfp* or la fonction d'YdaM est d'activer la transcription de *csgD* via la production de c-diGMP (Weber, Pesavento et al. 2006) De plus, un site d'appariement est prédit entre les Omr et le messenger *ydaM* (figure 57). Ce site recouvre celui de RprA, un petit ARN qui régule directement *ydaM* (Mika, Busse et al. 2012). Une autre voie indirecte d'inhibition de CsgD par les Omr pourrait passer par OmpR. En effet, les Omr répriment la synthèse d'OmpR qui active la transcription de *csgD*.

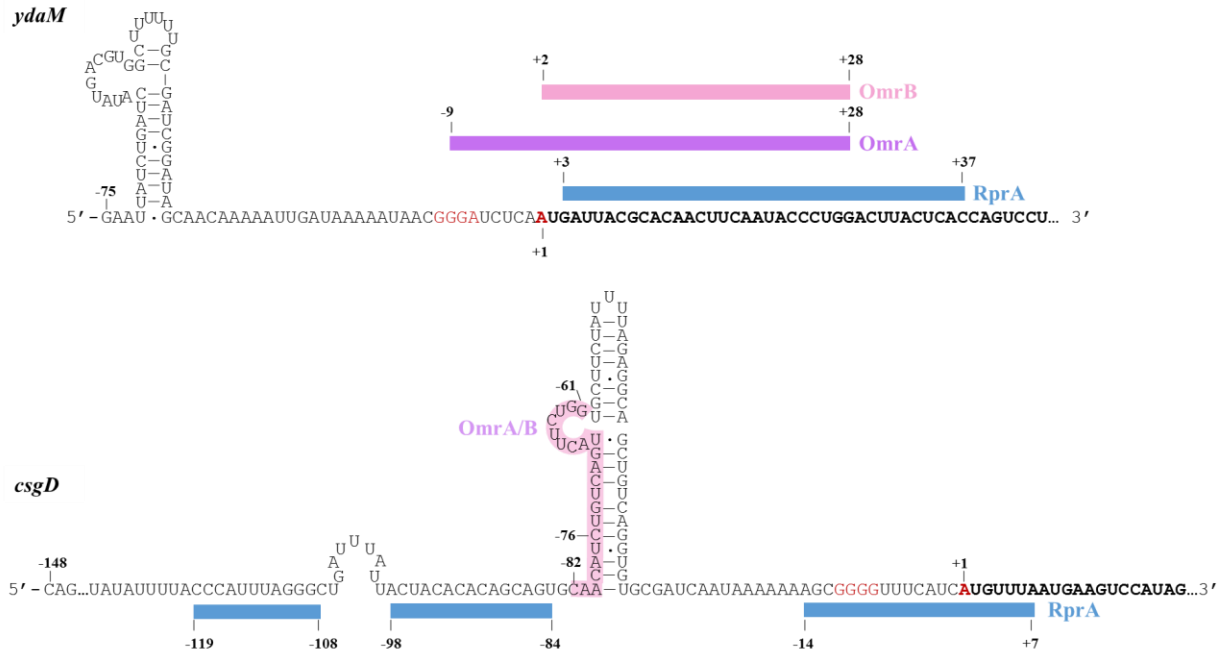


Figure 57 : Appariement des Omr avec le 5'UTR des messagers *csgD* et *ydaM*. Le +1 de traduction et le SD sont indiqués en rouge. Les positions mentionnées sont relatives au message en question. D'après (Mika, Busse et al. 2012).

Enfin, en 2012, il a été montré qu'OmrA et OmrB étaient impliqués dans un réseau complexe contrôlant la motilité. En effet, comme deux autres petits ARN (OxyS et ArcZ) les Omr régulent négativement l'ARNm *flhDC* qui code le régulateur majeur de la synthèse des flagelles (De Lay & Gottesman, 2012). Il est intéressant de mentionner que l'appariement entre les Omr et *flhDC* est relativement étendu puisqu'il s'étend entre les régions -54 à -9 du message et respectivement les régions +50 à +1 ou +49 à +1 d'OmrA et d'OmrB. On peut également noter que pour OmrA l'appariement engage 37 nucléotides alors que seulement 25 nucléotides d'OmrB sont concernés.

En conclusion, nous savons à ce jour que les Omr régulent directement cinq cibles : *ompR-envZ*, *csgD*, *flhDC*, *cirA* et *ompT*. Trois cibles pour lesquelles un appariement prédit avec les Omr sont mentionnés dans la littérature : *fecA*, *fepA* et *ydaM*. On peut cependant préciser que les données de notre laboratoire indiquent que *fepA* est une cible directe des Omr. La figure 58 récapitule le régulon des Omr.

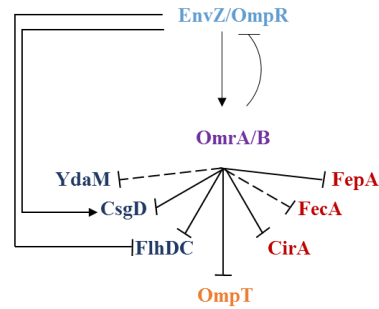


Figure 58 : Régulon des Omr. Les cibles directes sont indiquées en trait plein et les cibles pour lesquelles un appariement est prédit et/ou un contrôle direct est probable sont représentées avec un trait discontinu. Les cibles sont regroupées par fonction : biofilm/motilité (en bleu foncé), capture et import du fer (en rouge), dégradation protéique (en orange). On note la présence d'une feedforward loop entre le système EnvZ-OmpR et les cibles CsgD et FlhDC avec à la fois un contrôle direct ou indirect passant par les Omr. Dans le cas d'FlhDC c'est une boucle cohérente mais ce n'est pas le cas pour CsgD. Il y a également un circuit de feedback entre les Omr et le système EnvZ-OmpR.

1.3.4. RprA

1.3.4.1. Les cibles de RprA

Tout comme les Omr, RprA a été décrit dans les premières approches globales visant à identifier des petits ARN bactériens (Argaman, Hershberg et al. 2001; Wassarman, Repoila et al. 2001) et qui ont pu montrer que cet ARN de 105 nts était exprimé durant la phase stationnaire et fixait Hfq. Sa fonction a été décrite pour la première fois en 2001, au moins partiellement, par Majdalani *et al.* (Majdalani, Chen et al. 2001). Les auteurs avaient réalisé un crible génétique pour rechercher des activateurs de *rpoS* (autre que DsrA). Par la suite, une étude mécanistique (Majdalani, Hernandez et al. 2002) a pu montrer que RprA permettait de libérer aussi le site de fixation du ribosome qui, en l'absence de cet ARN est piégé dans une tige boucle (figure 59).

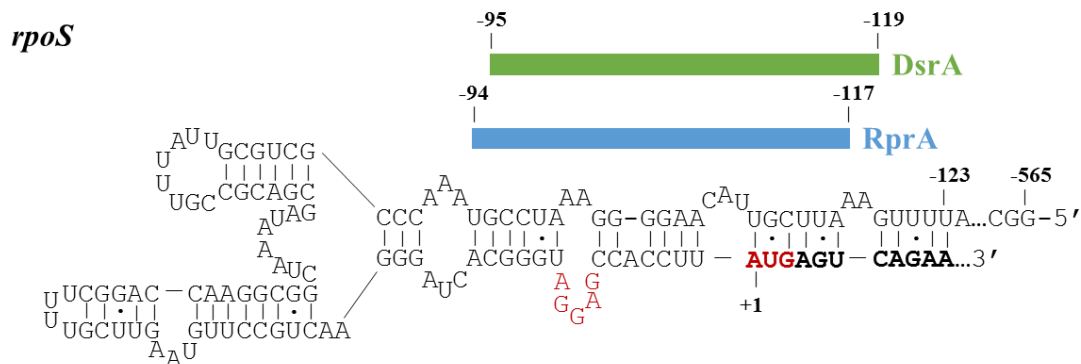


Figure 59 : Sites d'interaction entre *rpoS* et les ARN régulateurs DsrA et RprA. Le +1 de traduction et le SD sont indiqués en rouge. Les positions mentionnées sont relatives au messager *rpoS*. D'après (Mika, Busse et al. 2012).

Outre RpoS, RprA inhibe l'expression de *csgD* et d'*ydaM* (Mika, Busse et al. 2012). Pour ce qui est de *ydaM*, comme on a pu le voir sur la figure 57, l'extrémité 5' de RprA s'apparie tout comme les Omr peu après le codon de démarrage (positions +3 à +37 d'*ydaM*) inhibant de ce fait la fixation du ribosome.

En revanche sa fixation sur le messager *csgD* est bien différente des Omr. En effet, l'interaction entre RprA et *csgD* fait intervenir deux sites. Le site I correspond aux nucléotides -119 à -84 de *csgD* divisé en deux parties Ia (-119 à -108) et Ib (-98 à -84) qui s'apparient avec les nucléotides 60 à 97 de RprA (figure 60). Le site II intervient quant à lui entre les nucléotides -14 à +7 de *csgD* et 28 à 45 de RprA. Des expériences de mutations ont pu montrer que ces deux sites contribuaient à la régulation. L'interaction entre *csgD* et RprA conduit à la diminution du niveau d'ARNm *csgD*. Bien souvent, chez les ARN Hfq-dépendants, la dégradation de leur cible est due à la RNase E généralement associée à Hfq. Dans le cas de *csgD*, il a été montré que ce n'était pas le cas et que la présence de la RNaseE ou de la RNase III n'était pas essentielle à la régulation même si la RNase E et Hfq pouvaient y participer.

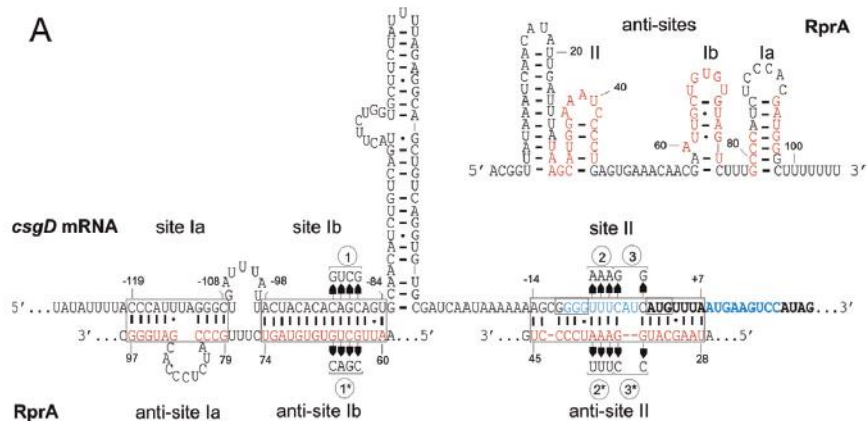


Figure 60 : Sites d'interaction entre RprA et *csgD*. Issue de (Mika, Busse et al. 2012).

Pour finir, il est intéressant de mentionner qu'en 2015, une étude transcriptomique chez *Salmonella* a permis de proposer plusieurs nouvelles cibles de RprA (Papenfort, Espinosa et al. 2015). La cible RicI, impliquée dans la limitation de la conjugaison par les pili, a en particulier été étudiée. En effet, RprA active la traduction par un mécanisme similaire à celui de *rpoS* ; en revanche, alors que la partie 5' de RprA est responsable de l'effet sur *rpoS* c'est la partie 3' de cet ARN qui agit sur *ricI*. D'autres cibles directes ou indirectes ont été suggérées dans cette étude dont le régulateur NarP sur lequel je reviendrai dans ma partie résultat.

1.3.4.2. Régulateurs de RprA

Historiquement, le premier régulateur de RprA identifié est le phosphorelais Rcs (Majdalani, Hernandez et al. 2002). Cette régulation est à priori directe car des mutations au promoteur de *rprA* dans une « boîte RcsB » prédite abolissent ce contrôle. La plupart des systèmes de transduction du signal bactériens fonctionnent comme un phosphorelais simple impliquant seulement deux protéines. Le système RcsCDB est quant à lui un exemple très remarquable et bien plus complexe dans l'utilisation de la phosphorylation pour la transduction d'un signal (pour revue (Majdalani and Gottesman 2005)). Ce système est majoritairement induit en réponse à des dommages liés à la membrane externe et la paroi bactérienne (figure 61). Le senseur de ce système présent dans la membrane interne est la kinase RcsC qui, en réponse au stimulus extérieur, va s'autophosphoryler sur son résidu histidine conservé situé au sein du domaine H1. Contrairement aux systèmes à deux composants classiques, RcsC va d'abord transférer le phosphate à un aspartate situé sur son domaine transmetteur (D1). Après ce transfert intramoléculaire, le phosphate sera transmis au résidu histidine situé sur le domaine H2 de RcsD un transmetteur membranaire intermédiaire. Pour finir le phosphate sera transmis à l'aspartate conservé du régulateur RcsB (domaine D2). Comme les autres régulateurs de réponse de systèmes à deux composants RcsB possède un domaine de fixation à l'ADN et pourra donc se fixer sous forme d'homodimères phosphorylés sur le promoteur de ses gènes cibles tel que le promoteur de *rprA* mais aussi *osmC* et *ftsZ*.

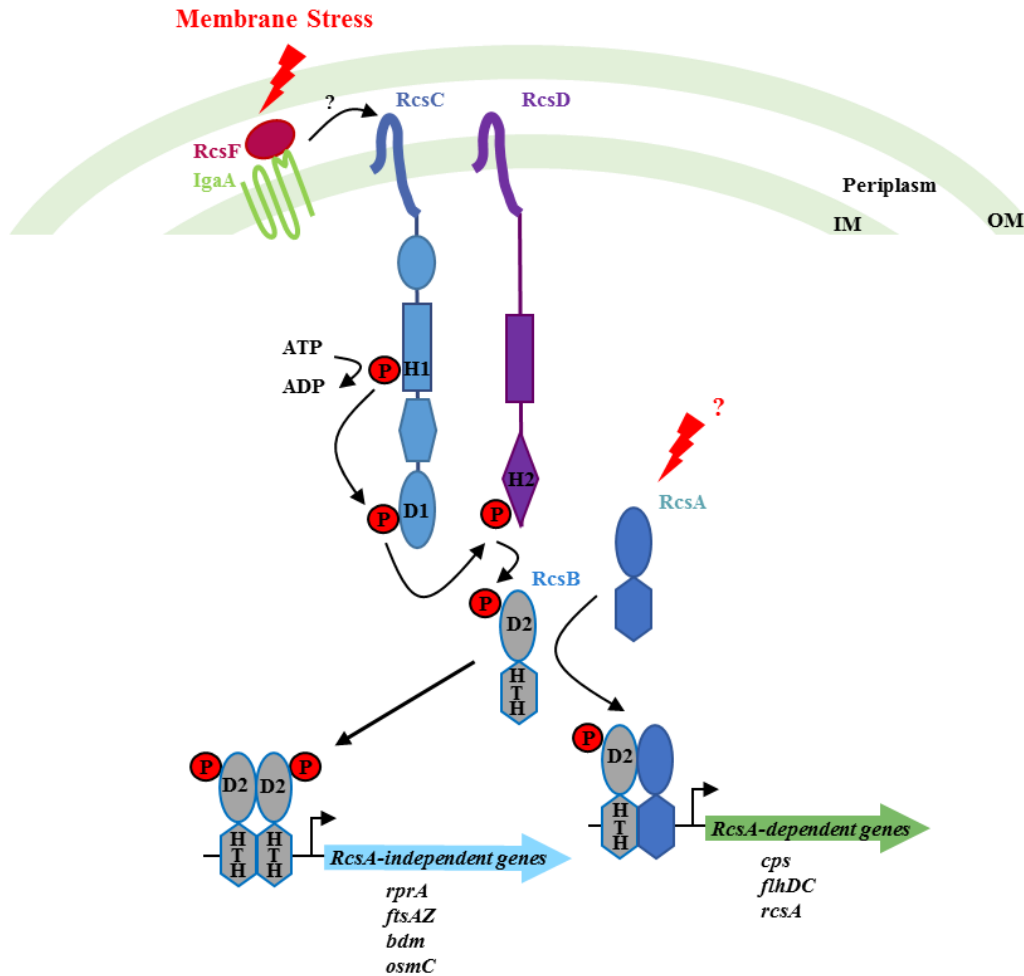


Figure 61 : Le phosphorelais Rcs d'*E. coli*. En présence de stress membranaire RcsF interagit avec IgaA ce qui conduit à l'activation de la cascade du phosphorelais RcsCDB. La première étape est l'autophosphorylation de RcsC au niveau du domaine histidine conservé (H1). Par transfert intramolécule le phosphate se retrouve ensuite sur le domaine donneur (D1) de RcsC puis est transféré au domaine histidine (H2) de RcsD. Le phosphate est par la suite transféré au domaine aspartate conservé du régulateur RcsB (D2). Sous forme d'homodimère (promoteurs RcsA indépendants) ou d'hétérodimères avec la protéine RcsA (gènes RcsA-dépendants) RcsB pourra activer la transcription de nombreux gènes afin de s'adapter aux nouvelles conditions environnementales. Des exemples de gènes pour lesquels des boîtes RcsB ou RcsAB ont pu être définies sont également mentionnés.

La complexité du système ne s'arrête cependant pas là. Le phosphorelais Rcs fait intervenir trois protéines accessoires : RcsA, RcsF et IgaA. En effet, RcsB sous sa forme phosphorylée peut également former des hétérodimères avec la protéine RcsA, un facteur de transcription auxiliaire non phosphorylable. De façon intéressante, les cibles du complexe RcsB-RcsA sont différentes de celles des homodimères de RcsB. C'est pourquoi ces cibles seront nommées RcsA-dépendantes et on compte parmi elles des gènes codants les protéines de capsule *cps* et le régulateur de la motilité *flhDC*. Néanmoins, on peut noter qu'une surexpression de RcsB est suffisante pour activer les

gènes de capsule même en l'absence de RcsA (Brill, Quinlan-Walsh et al. 1988). RcsF, la seconde protéine accessoire, est une lipoprotéine ancrée à la membrane externe qui est impliquée dans la perception du signal et qui agit en amont de RcsC. En effet, la surexpression de RcsF dans une souche délétée pour *rscC* ne suffit pas à induire la synthèse des gènes de capsule (Majdalani, Hernandez et al. 2002) alors que le mutant *rscI37* constitutivement actif suffit pour induire la synthèse de ces mêmes gènes en absence de *rscF* (Majdalani, Heck et al. 2005). Une étude récente a pu mettre en évidence le rôle de RcsF et son interaction avec la troisième protéine accessoire du système : IgaA (Cho, Szewczyk et al. 2014). Cette étude montre dans un premier temps que RcsF interagit avec BamA, la protéine majeure de la machinerie d'assemblage des protéines de la membrane externe à feuillet β (figure 62). En condition « normale », RcsF comme les autres lipoprotéines de membrane externe est convoyée vers la face interne de la membrane externe par la chaperonne LolA. Dans un second temps, RcsF est pris en charge par BamA qui l'assemble en complexe avec la porine OmpA. RcsF se retrouve ainsi répartie à la surface de la cellule et éloignée de la protéine IgaA. En condition de stress de membrane, les molécules de RcsF nouvellement synthétisées éprouvent des difficultés à contacter BamA et se retrouvent dans le périplasme où elles peuvent interagir avec IgaA ce qui a pour conséquence d'activer le système Rcs. Chez *S. enterica* il a été proposé qu'IgaA module le transfert de phosphate entre RcsC et RcsD et maintiendrait le système en mode « off ». Par analogie à Salmonelle on peut imaginer que la liaison de RcsF à IgaA lèverait l'inhibition et activerait le système Rcs. Ce mécanisme complexe permettrait à la cellule de contrôler la disponibilité et la capacité de la machinerie Bam à assembler les protéines de membrane externe et/ou d'exposer les lipoprotéines à la surface de la cellule via ces mêmes protéines de membrane externe.

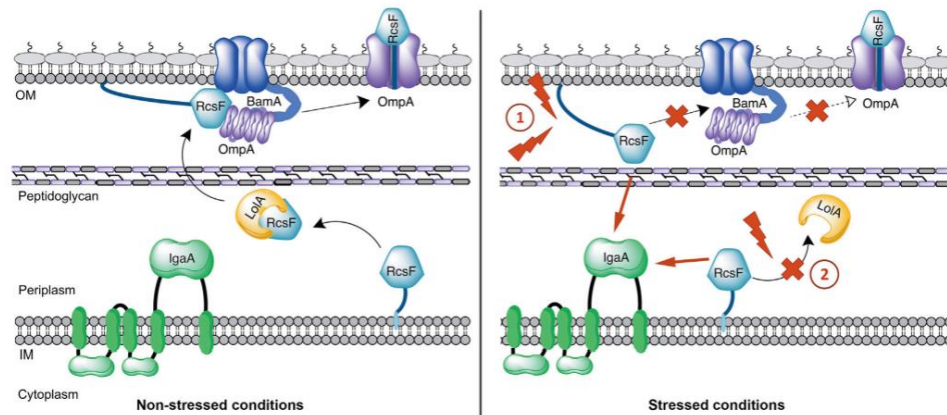


Figure 62 : Mode d'activation du système Rcs : le rôle majeur de RcsF et d'IgaA. Issue de (Cho, Szewczyk et al. 2014)

Le petit ARN RprA est également directement activé par le système à deux composants CpxA-CpxR répondant lui aussi au stress d'enveloppe (Vogt, Evans et al. 2014). De façon intéressante, RprA exerce un rétrocontrôle négatif sur Cpx mais cette fois par un mécanisme indirect encore inconnu.

Pour finir, il a été suggéré que la synthèse de RprA pouvait être réprimée par le régulateur global de la synthèse des flagelles LhrA (Peterson, Carabetta et al. 2006). On peut se demander si cette régulation est directe car seules des expériences de fusion au promoteur de *rprA* en contexte sauvage ou délété pour *lhrA* ont été effectuées.

1.4. La bactérie *Escherichia coli*

Au cours de ma thèse, j'ai eu l'occasion de m'intéresser à l'étude des petits ARN chez la bactérie modèle *E. coli* K12, mais j'ai également eu l'opportunité de travailler avec la bactérie pathogène *E. coli* LF82 associée à la maladie de Crohn. Je me propose donc dans cette dernière partie de présenter, après quelques généralités sur *E. coli*, la souche LF82 ainsi que quelques notions sur la maladie de Crohn.

1.4.1. Généralités

Cet organisme a été décrit pour la première fois en 1885 par un pédiatre allemand du nom de Theodor Escherich qui l'avait isolé de selles de bébés nourris exclusivement au lait maternel. Il observa que le bacille (figure 63), qu'il nomma *Bacterium coli comune*, se trouvait chez tous les individus et qu'il était inoffensif. Par la suite, des études ont montré qu'il existait des variantes pathogènes de cet organisme. Ce n'est qu'à partir de 1919 que Castellani et Chalmers, dans leur troisième édition de leur Manuel of Tropical Diseases, rebaptisèrent le bacille *Escherichia coli* en l'honneur de son découvreur.

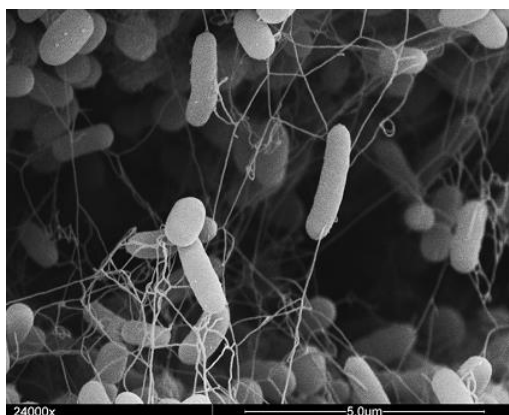


Figure 63 : Image de microscope électronique à balayage d'*Escherichia coli* K12. Issue de (Serra et al., 2013).

Le génome d'*E. coli* K12 a été séquencé en 1997 (Blattner et al., 1997). Il est composé de $4,6 \times 10^6$ paires de bases et compte environ 4000 cadres de lecture ouverts dont 2300 gènes codant des protéines connues ou pour lesquelles le phénotype des mutants est connu. L'ensemble des gènes d'*E. coli* et leurs séquences de régulation couvrent à ce jour 99,5 % de son génome.

1.4.2. Les différentes espèces d'*Escherichia coli*

Bien que la majorité des souches soient commensales et essentiellement retrouvées dans la flore intestinale de nombreux animaux (la plupart des vertébrés, tous les mammifères), des variantes pathogènes existent. Selon leur capacité à traverser des muqueuses lésées, ces souches pathogènes peuvent provoquer des infections urinaires, biliaires, génitales, ou dans de rares cas, une septicémie.

L'espèce *E. coli* est sous-divisée en de nombreux sérotypes. Actuellement 160 types d'antigènes O et 55 types de protéines H sont connus, ce qui fait théoriquement 8800 sérotypes possibles. L'isolat K12 utilisé couramment au laboratoire appartenait au groupe O16 : H48 mais, maintenu en culture depuis 1922, il a perdu par mutation son antigène O et n'est donc plus capable de coloniser l'intestin humain.

L'interaction des souches d'*E. coli* pathogènes avec le système immunitaire de l'hôte est un facteur clé dans le processus de pathogénicité. En effet, les antigènes de type O (lipo-polysaccharides) présents à la surface de la bactérie, protègent celle-ci notamment de la fixation des anticorps et de la phagocytose mais aussi de l'action lytique du complément. D'autre part, les antigènes K (capsules polysaccharidiques) sécrétées à la surface de certaines souches pathogènes permettent à la bactérie d'échapper au système de défense de l'hôte. Le système immunitaire peut également être évité grâce à la variation antigénique des protéines de surface telles que les protéines H (protéines majeures des flagelles).

Outre les *E. coli* induisant des diarrhées sur lesquelles je reviendrais, il existe le pathovar des *E. coli* pathogènes extraintestinales (ExPEC). Ces dernières sont impliquées dans des affections non-intestinales telles que les infections urinaires, les méningites ou les septicémies. En effet, 90% des infections urinaires sont causées par les souches d'*E. coli* uropathogènes (UPEC). Ces bactéries utilisent les fimbriae pour se lier aux cellules endothéliales et coloniser ensuite la vessie. Les hémolysines produites entraînent la lyse des cellules des voies urinaires.

Par ailleurs, les souches d'*E. coli* inductrices de diarrhées peuvent être classées en six pathovars sur la base des modes d'interaction hôte-pathogène et des signes cliniques liés à l'infection : *E. coli* entérotoxigéniques (ETEC), *E. coli* entéroinvasives (EIEC), *E. coli* entéropathogènes (EPEC), *E. coli* entérohémorragiques (EHEC), *E. coli* entéroaggrégatives (EAaggEC) et enfin *E. coli* à

adhésion diffuse (DAEC) (Kaper, Nataro et al. 2004). On peut noter qu'un nouveau groupe a fait son apparition plus récemment : les AIEC pour *E. coli* adhérentes et invasives. C'est ce groupe qui m'a intéressé au cours de cette thèse.

Les souches de type ETEC et EPEC infectent essentiellement les jeunes enfants. On retrouve également ce type de souche dans les cas de « diarrhée du voyageur » ou « turista » bien que d'autres espèces bactériennes soient impliquées dans ce type de pathologie.

Le pouvoir pathogène de ces différentes souches d'*E. coli* repose essentiellement sur la présence de fimbriae (sorte de pili) qui permettent l'adhésion des bactéries aux cellules intestinales mais aussi la sécrétion de toxines thermostables et/ou thermolabiles.

On note un certain nombre de similitudes entre les mécanismes d'invasion de ces souches et ceux observés chez des bactéries de type *Shigella*. Par exemple, certaines d'entre elles, les AIEC notamment, sont capables de lyser leur vacuole d'internalisation et de se déplacer à l'intérieur de la cellule infectée puis pénétrer dans les cellules adjacentes.

1.4.3. *Escherichia coli* et la flore intestinale : entre commensalisme et pathogénie

1.4.3.1. Le microbiote intestinal

Le terme « microbiote intestinal » définit l'ensemble des micro-organismes (archées, bactéries, champignons, protistes et virus) peuplant le tractus digestif d'un individu de l'intestin grêle à l'anus. Par l'analyse de l'ARNr 16S mais aussi de l'ensemble des gènes codés par le microbiote (microbiome), une étude a pu montrer que celui-ci était composé de plus d'une centaine d'espèces bactériennes différentes aérobies mais surtout anaérobies (Qin, Li et al. 2010). Chez un adulte sain, on retrouve environ 1000 espèces moléculaires dont 60% à 70% ne sont pas cultivables en laboratoire. Chaque individu possède sa propre flore intestinale de composition relativement stable à partir d'environ 2 ans. Pour autant, la structure globale est relativement la même d'un individu à l'autre. En effet, deux grands phyla constituent près de 90% de la population totale : les Firmicutes et les Bacteroidetes (figure 64). Les Clostridies constituent la majeure partie du phylum des Firmicutes. Deux autres phyla moins abondants sont également rapportés : les Protéobactéries et les Actinobactéries.

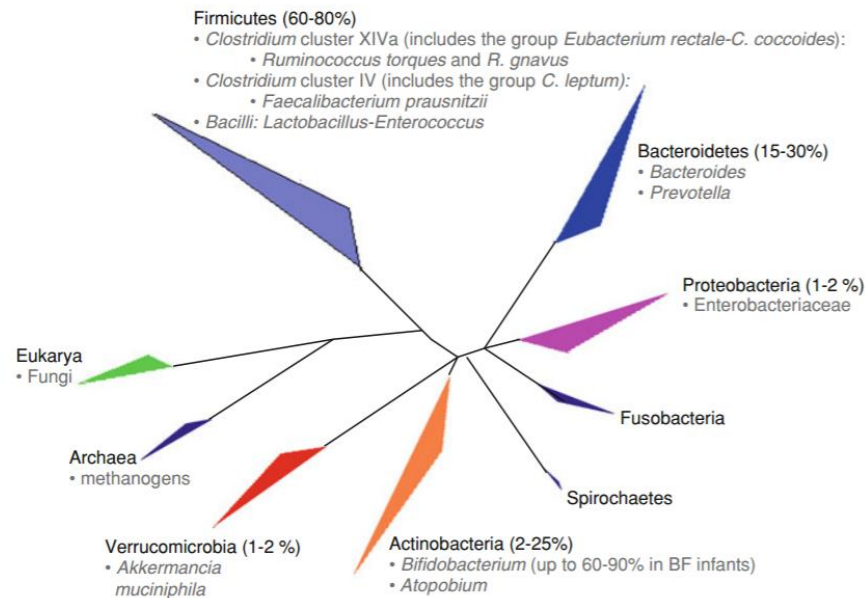


Figure 64 : Représentation schématique de l'arbre phylogénétique des phyla majoritaires retrouvés dans le microbiote intestinal humain. Entre parenthèse se trouve la proportion du phylum chez un adulte. La nomenclature BF (pour Breast-Feed) correspond à l'allaitement durant la petite enfance (Cheng, Palva et al. 2013)

On peut préciser que la flore est répartie de façon plus ou moins dense le long du tractus digestif. En effet, la majeure partie de la flore est située au niveau de l'iléon et du colon avec une densité par gramme de contenu intestinal d'environ 10^7 à 10^{12} bactéries alors que la densité de la flore stomacale n'est que de 0 à 10^2 (Sartor 2008).

E. coli fait partie des premières bactéries à coloniser l'intestin des nouveau-nés. A l'âge adulte, elle représenterait moins de 1% des bactéries présentes dans le microbiote.

1.4.3.2. La maladie de Crohn

La maladie de Crohn est une pathologie humaine caractérisée par une inflammation chronique de l'appareil digestif (Shanahan 2002). En général, elle touche l'iléon situé à l'extrémité de l'intestin grêle à la jonction avec le colon. Néanmoins, d'autres portions du tube digestif peuvent être touchées. L'inflammation s'étend relativement profondément dans la muqueuse intestinale, provoquant des gonflements et générant des douleurs. Dans un certain nombre de cas, le tissu cicatriciel formé à l'intérieur de l'intestin peut créer une sténose, c'est-à-dire le rétrécissement du conduit digestif dû à un épaississement de la paroi. Cette sténose peut ralentir la progression des aliments dans le système digestif aboutissant à de vives douleurs ou des crampes. Une autre

complication de cette maladie est la formation d'ulcères aux alentours de la zone inflammée. Ces ulcères peuvent engendrer des fistules traitables par voie médicamenteuse ou par chirurgie. Dans les cas extrêmes certains patients sont contraints de subir des interventions chirurgicales visant à retirer des portions plus ou moins importantes de l'intestin inflammé. Globalement les patients atteints de la maladie de Crohn souffrent d'anémie et de carences en protéines et en vitamines du fait de la mauvaise absorption des composés par la portion atteinte de l'intestin et doivent donc suivre un traitement de supplémentation.

La cause exacte de cette pathologie est aujourd'hui inconnue mais il s'agirait d'une réaction anormale du système immunitaire. De nombreuses études ont proposé que la flore intestinale et plus précisément la présence de certains pathogènes contribuent à la survenue de la maladie de Crohn (Tawfik, Flanagan et al. 2014). Par des approches globales, il a été montré que la dysbiose générale du microbiote intestinal constaté chez les patients atteints de la maladie de Crohn avait pour conséquence l'augmentation du groupe des entérobactériacées au détriment des firmicutes (Kang, Denman et al. 2010; Mondot, Kang et al. 2011). Il a donc été proposé que cette dysbiose ait pour conséquence de diminuer les fonctions de barrière du microbiote et favoriserait la colonisation de l'épithélium par des AIEC. En effet, il a été observé dans les lésions précoces ou chroniques de l'intestin une surreprésentation des souches d'*E. coli* de 50 % à 100 % par rapport au nombre total de bactéries aérobies et anaérobies (Darfeuille-Michaud, Neut et al. 1998) dont la plupart sont adhérentes aux cellules épithéliales et capables de les envahir d'où leur nom d'AIEC. De plus, la présence d'AIEC induit la sécrétion de TNF- α par les macrophages (Glasser, Boudeau et al. 2001). Cette cytokine pro-inflammatoire est notamment responsable du « gonflement » des intestins observé chez les patients atteints de la maladie de Crohn. On peut cependant noter que les études restent prudentes et suggèrent que la présence des AIEC pourrait aussi être une conséquence de la maladie et non la cause directe et serait d'une certaine façon un facteur aggravant.

Outre la dysbiose, deux autres théories tentent d'expliquer les origines de la maladie de Crohn. Différents auteurs suggèrent que la pathologie serait due à une translocation bactérienne excessive. Ce processus est défini comme le passage de bactéries viables d'origine digestive à travers la barrière de la muqueuse intestinale en direction des ganglions mésentériques (ganglions lymphatiques entourant les vaisseaux irrigant le tube digestif) puis de là vers d'autres organes internes (foie et rate principalement). Chez les patients atteints de la maladie de Crohn plusieurs

études ont montré que certaines bactéries intestinales (*E. coli*, *Clostridium spp.*, *Enterococcus spp.*, *Proteus spp.*, et *Bacteroides fragilis*) étaient retrouvées au niveau des ganglions mésentériques suggérant qu'elles aient passé la barrière intestinale (Ambrose, Johnson et al. 1984; Takesue, Ohge et al. 2002).

Pour finir, il a été suggéré que la maladie de Crohn et son évolution impliqueraient divers agents infectieux tels que des virus (virus Epstein Barr notamment), des levures (*Candida albicans*) ou encore des bactéries pathogènes. Outre les souches d'*E. coli* adhérentes invasives mentionnées précédemment, il a été montré que des souches de *Mycobacterium avium paratuberculosis*, d'*Helicobacter* mais aussi de *Yersinia* et de *Listeria* pourraient être responsables du développement de la maladie de Crohn (Pineton de Chambrun, Colombel et al. 2008).

Cette pathologie semble toucher de manière égale les femmes et les hommes et s'établit généralement entre 13 et 30 ans. Néanmoins, différentes études montrent que différents facteurs environnementaux et le régime alimentaire favoriseraient la survenue de la dysbiose. Le tabac (Mahid, Minor et al. 2007), un régime riche en gras ou en sucre et pauvre en fibre (Chapman-Kiddell, Davies et al. 2010) sont considérés comme des facteurs de risque. Pour finir, une étude a pu montrer que la maladie de Crohn pouvait avoir des origines génétiques (Jostins, Ripke et al. 2012). 30 polymorphismes génétiques semblent être spécifiquement associés à cette maladie et principalement reliés aux gènes impliqués dans la reconnaissance des pathogènes (notamment une mutation du récepteur TLR4 limitant la réponse au LPS) et la destruction des bactéries au sein des macrophages (notamment via l'autophagosome).

A ce jour, différentes stratégies thérapeutiques seules ou en association sont disponibles mais ne permettent pas encore d'empêcher totalement les récives. La prise d'anti-inflammatoires permet de contrôler les poussées et de limiter les récives mais est souvent associée à des effets secondaires indésirables. Le recours à la chirurgie est bien souvent incontournable en cas de complication de type ulcère, sténose ou fistule. Malheureusement l'excision des segments atteints n'exclut pas la survenue de nouvelles poussées. Ces dernières années des thérapeutiques nouvelles ont émergé et sont essentiellement basées sur la manipulation de la flore intestinale allant de la prise d'antibiotiques à l'ajout de prébiotiques ou probiotiques dans l'alimentation jusqu'à des cas de transplantation fécale. Une étude métagénomique récente montre que la prise d'antibiotiques

chez les patients atteints de la maladie de Crohn amplifierait la dysbiose associée à cette pathologie suggérant que cette piste thérapeutique soit peu prometteuse (Gevers, Kugathasan et al. 2014).

1.4.4. La souche d'*Escherichia coli* LF82

Parmi les différentes souches associées à la maladie de Crohn, la souche d'*E. coli* LF82 a tout particulièrement été étudiée et est considérée comme la souche de référence. Le génome de cette souche séquencé en 2010 montre qu'elle appartient au groupe phylogénétique B2 et qu'elle a sans doute évolué par acquisition de gènes de *Salmonella* ou de *Yersinia* à partir d'une souche ExPEC (Miquel, Peyretailade et al. 2010).

Outre sa capacité à adhérer aux cellules épithéliales intestinales (Darfeuille-Michaud, Neut et al. 1998), cette souche est capable d'envahir l'épithélium (Boudeau, Glasser et al. 1999). De façon intéressante, le mécanisme d'internalisation de LF82 dans les cellules épithéliales se révèle bien différent des autres souches d'*E. coli* et s'apparente plus à celui de certaines *Salmonelles* ou *Shigelles* voire des souches de *Listeria*. En effet, le contact entre LF82 et la cellule épithéliale induit la formation d'extensions des microvillosités permettant ainsi d'entourer la bactérie puis de l'internaliser (figure 65). On retrouve ce processus chez les *Salmonelles* ou les *Shigelles*. La bactérie se retrouve ainsi dans des vacuoles d'endocytose qu'elle peut par la suite lyser pour se retrouver libre dans le cytoplasme. C'est donc un processus actif qui permet à LF82 d'envahir les cellules épithéliales.

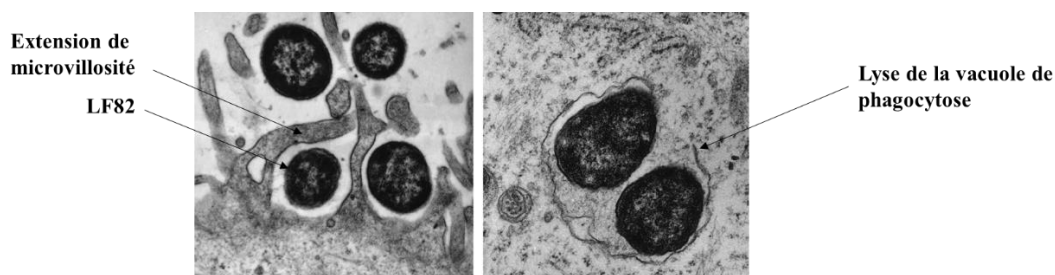


Figure 65 : Cliché de microscopie électronique à transmission de cellules épithéliales Hep-2 infectées par la souche LF82. A gauche, bactérie englobée par l'extension d'une microvillosité issue de la cellule infectée (x 21600). A droite, la membrane de la vacuole en cours de lyse par la souche LF82, illustrant sa capacité à s'échapper des vacuoles d'endocytose (x 28800). D'après (Boudeau, Glasser et al. 1999).

Par la suite, la même équipe de recherche a pu montrer, en réalisant un crible génétique, que plusieurs gènes étaient impliqués dans le processus d'adhésion et d'internalisation dans les cellules épithéliales. Les gènes qui ressortent de ce crible codent pour les pili de type I, les flagelles et des

lipoprotéines de surface. Des mutants des gènes *fimI* et *fimF* codant des constituants des pili conservent leur capacité d'adhésion aux cellules épithéliales mais perdent celle d'induire l'internalisation. L'expression de l'opéron *fim* de LF82 chez *E. coli* K12 ne confère pas à cette dernière la capacité d'invasion des cellules épithéliales suggérant que les pili ne sont cependant pas la seule structure impliquée dans ce processus (Boudeau, Barnich et al. 2001). En 2003, il a en effet été montré que les flagelles jouaient aussi un rôle dans la virulence de LF82. La délétion du gène *fliC* codant la flagelline (protéine structurale majoritaire du flagelle), bloque le processus d'invasion de LF82 et réduit sa capacité à adhérer aux cellules épithéliales intestines (Barnich, Boudeau et al. 2003). Deux lipoprotéines, NlpI et YfgL, sont également importantes dans le processus de virulence de LF82. Un mutant de la protéine de surface NlpI réduit fortement la capacité d'adhésion et d'invasion de LF82 (Barnich, Bringer et al. 2004). Ce mutant est non motile et exprime très peu de pili de type I pourtant NlpI ne semble pas agir directement sur l'opéron *fim* ou le régulateur des flagelles FlhDC. Au regard de l'effet pléiotropique du mutant *nlpI*, il se pourrait qu'il soit relié à un système de régulation plus large de type système à deux composants. En effet, chez *E. coli* K12 la lipoprotéine NlpE est requise pour induire l'adhésion à des surfaces hydrophobes via le système à deux composants CpxA-CpxR. On peut imaginer qu'un mécanisme semblable ait lieu chez LF82.

La seconde lipoprotéine importante dans la virulence est YfgL. Cette protéine est normalement impliquée dans la synthèse et la dégradation du peptidoglycane. La délétion du gène *yfgL* a pour conséquence une diminution très forte (34 fois moins) des capacités d'adhésion et d'invasion de LF82 (Rolhion, Barnich et al. 2005). Dans le mutant d'*yfgL* il est également observé une diminution des vésicules de membrane externe (OMV). Ces vésicules sont formées par bourgeonnement de la membrane externe chez les bactéries à Gram négatif (Kuehn and Kesty 2005). Il a été suggéré que ces vésicules soient impliquées dans les communications entre bactéries de la même espèce ou inter-espèces et plus récemment dans les interactions hôtes-pathogènes en délivrant par exemple des facteurs de virulence (toxines, adhésines, protéases...). Les OMV pourraient également jouer un rôle de protection chez les bactéries non pathogènes afin de se débarrasser de composés toxiques ou des phages. Dans le cas de LF82, il semblerait que la délétion d'*yfgL* diminue la production de peptidoglycane limitant *in fine* la production des OMV.

LF82, comme les autres AIEC associées à la maladie de Crohn, possède par ailleurs la capacité de survivre et se répliquer dans les macrophages. Les macrophages sont les cellules immunitaires capables d'éliminer les agents pathogènes de l'organisme par phagocytose. Une fois « l'intrus » capturé dans la phagosome (vésicule intracellulaire) ce dernier va fusionner avec le lysosome, une vésicule dont le pH est très acide (entre 3.5 et 5) et contenant de nombreuses enzymes et radicaux libres de l'oxygène. Dans cet environnement peu favorable la plupart des bactéries sont tuées mais ce n'est pas le cas de LF82. Une étude montre que le nombre de bactéries LF82 retrouvées dans les macrophages humain est 2 fois supérieur à 4h qu'à 1h après infection (Glasser, Boudeau et al. 2001). Dans ce même test, avec une souche d'*E. coli* K12, entre 1h et 4h la moitié des bactéries ont été tuées. D'autres tests avec des lignées de macrophages murins J774-A1 montrent que LF82 est capable de résister à l'activité bactéricide des macrophages pendant au moins 5 jours. Par ailleurs, un pic à 48h post-infection est observé où les bactéries sont jusqu'à 72 fois plus nombreuses qu'une heure après infection. LF82 induit la formation d'une large vacuole dans laquelle elle va persister et se répliquer. Contrairement à ce qui se passe dans les cellules épithéliales LF82 ne semble pas lyser la vacuole car aucune bactérie n'est retrouvée dans le cytoplasme des macrophages. Par ailleurs, LF82 n'induit pas la mort cellulaire des macrophages infectés comme font d'autres pathogènes intracellulaires tels que les Salmonelles ou les Shigelles. En revanche, les macrophages infectés sécrètent des taux élevés de TNF- α (Glasser, Boudeau et al. 2001). Cette cytokine, importante dans la médiation de l'inflammation aiguë, est normalement produite par des macrophages infectés en réponse au lipopolysaccharide (LPS) des bactéries à Gram négatif. La sécrétion de TNF- α permet notamment le recrutement d'autres macrophages au site de l'infection. Il est intéressant de rapprocher cette information de l'observation de granulomes associés aux lésions de la maladie de Crohn. Les granulomes associés à cette pathologie sont des agrégats formés de bactéries (on retrouve *E. coli* dans 80% des granulomes de la maladie de Crohn), de macrophages, de lymphocytes et de cellules géantes multinucléées (Ryan, Kelly et al. 2004). On peut noter que la présence de ces granulomes est une des caractéristiques histologiques étudiées dans le processus de diagnostic de la maladie de Crohn. Différentes études ont pu montrer que le TNF- α augmentait la formation des granulomes. Par ailleurs, *in vitro* LF82 induit l'agrégation de macrophages humains infectés puis le recrutement de lymphocytes circulants et de cellules géantes multinucléées formant une structure ressemblant fortement aux granulomes associés à la maladie de Crohn (Meconi, Vercellone et al. 2007). On peut donc penser que LF82,

en stimulant la production de TNF- α par les macrophages, conduirait à la formation des granulomes. De plus, il a été montré que la forte sécrétion de TNF- α était nécessaire pour que la réplication de LF82 dans les macrophages soit maximale (Bringer, Billard et al. 2012).

Un autre rôle du TNF- α a été suggéré dans une étude en 2007 (Barnich, Carvalho et al. 2007). En effet, il a été montré que LF82 adhérerait préférentiellement à la bordure des entérocytes iléaux (un des quatre type de cellules retrouvé dans l'épithélium intestinal) des patients atteints de la maladie de Crohn et plus précisément au niveau du récepteur CEACAM6 (CarcinoEmbryonic Antigen-related Cell Adhesion Molecule 6), surexprimé chez ce type de patient. Les auteurs précisent que la stimulation des cellules épithéliales par du TNF- α et des cytokines pro-inflammatoires (INF- γ) aboutit à la surexpression de ce récepteur de même que l'infection de ces même cellules par une AIEC. Ainsi les AIEC telles que LF82 pourraient stimuler la production de TNF- α , induisant la surproduction du récepteur à la surface des cellules épithéliales et donc favorisant leur colonisation (figure 66).

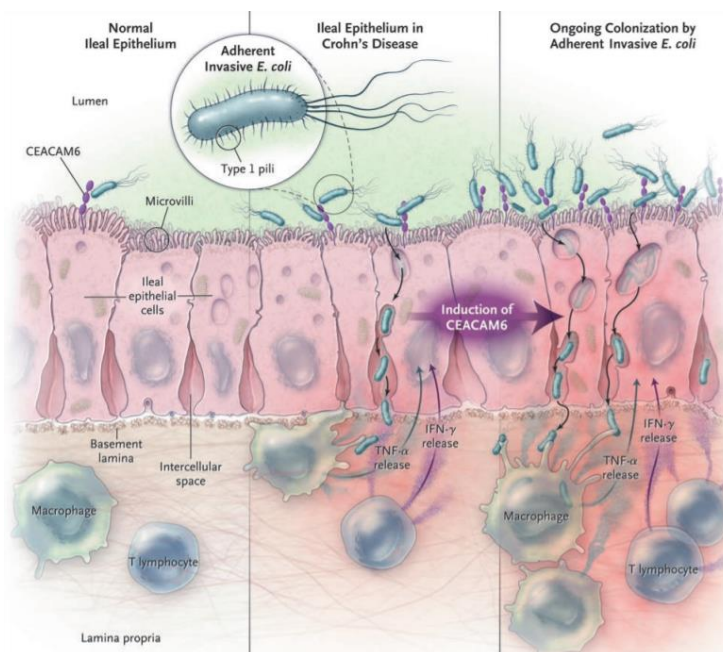


Figure 66 : Modèle de colonisation des cellules épithéliales par une bactérie de type AIEC telle que LF82. Les bactéries de type AIEC pourraient via leurs pili de type I se fixer au récepteur CEACAM6 ce qui leur permettrait d'adhérer et d'envahir l'épithélium intestinal. Une fois cette barrière passée, elles pourraient être internalisées au sein des macrophages qui en retour sécrèteraient du TNF- α . Le rôle de cette molécule serait non seulement de recruter d'autres macrophages et lymphocytes (formation des granulomes) mais aussi de stimuler l'expression du récepteur CEACAM6 facilitant ainsi l'entrée de nouvelles AIEC. Issue de (Abraham and Cho 2007).

De façon intéressante, une fois dans les phagolysosomes, LF82 n'interfère pas avec la maturation des macrophages et est donc capable de résister à l'acidité et à l'activité hydrolytique de la cathepsine D active dans ce compartiment (Bringer, Glasser et al. 2006). Au contraire l'acidité semble être un signal important pour la régulation des gènes de virulence nécessaires à la survie et la multiplication de LF82 au sein des macrophages. En effet, la neutralisation du pH des phagolysosomes contenant des bactéries de type LF82 par des agents tels que la chloroquine ou du chlorure d'ammonium inhibe leur réplication.

Une multitude de gènes sont impliqués dans la réponse des bactéries au stress généré par le processus infectieux mais je me suis particulièrement intéressée au rôle du système à deux composants EnvZ-OmpR dans la pathogenèse de LF82. Une forte osmolarité améliore de façon significative l'adhésion entre LF82 et des cellules épithéliales intestinales (Rolhion, Carvalho et al. 2007). L'analyse du profil des porines OmpA, OmpC et OmpF dans ces conditions révèle une augmentation du niveau d'OmpC et une diminution du niveau d'OmpF. Les auteurs ont donc logiquement pensé que cela pouvait impliquer le système EnvZ-OmpR. Comme attendu, la délétion d'*ompR* dans une souche LF82 diminue très fortement ses capacités à adhérer et envahir les cellules épithéliales (figure 67A).

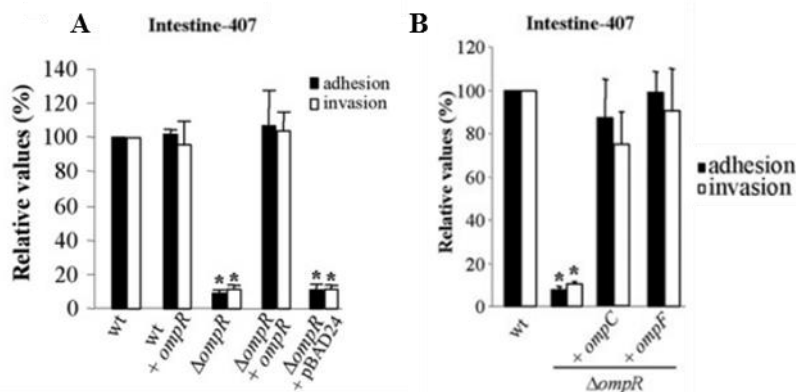


Figure 67 : Effet de la délétion d'*ompR* sur les capacités d'adhésion et d'invasion de LF82. L'adhésion et l'invasion de souches sauvages (wt) ou délétées pour OmpR ($\Delta ompR$) transformées par le plasmide surproduisant OmpR (+*ompR*) [A], OmpC (+*ompC*) [B], OmpF (+*ompF*) [B] ou le vecteur vide (pBAD24) [A]. Issue de (Rolhion, Carvalho et al. 2007).

Les auteurs de cette publication précisent que c'est la porine OmpC qui est directement impliquée dans l'adhésion et l'invasion. En effet, à faible osmolarité ou dans une souche délétée pour *ompC*, ces capacités sont fortement diminuées. Ils précisent également que la délétion d'*ompC* n'affecte

pas la formation des pili de type I et des flagelles requis pour l'adhésion et l'invasion. En revanche, dans un mutant d'*ompR*, la surexpression d'*ompC* ou d'*ompF* restaure l'adhésion et l'invasion de cellules épithéliales (figure 67B).

Les auteurs proposent également un modèle dans lequel la forte osmolarité du tractus intestinal va induire EnvZ-OmpR conduisant à l'augmentation de l'expression d'*ompC*. Dans ce cas, l'accumulation d'OmpC dans le périplasme activerait sigma E aboutissant à la surexpression des pili de type I et des flagelles à la surface de la bactérie (figure 68). Ces données montrent donc que les protéines de membrane externe jouent un rôle prépondérant dans la pathogénie de LF82. Il est intéressant de préciser que chez 55% des patients atteints de la maladie de Crohn on retrouve des anticorps dirigés contre la porine OmpC, renforçant l'implication de cette protéine dans la pathogénèse (Landers, Cohavy et al. 2002).

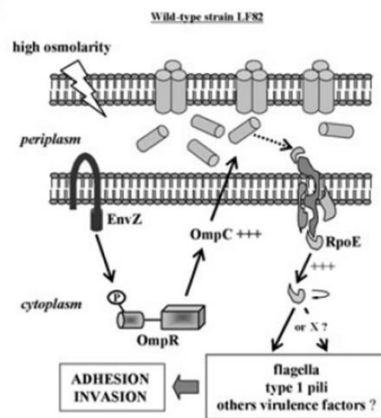


Figure 68 : Rôle potentiel du système à deux composants EnvZ-OmpR et d'OmpC dans la virulence de LF82. Issue de (Rolhion, Carvalho et al. 2007).

On peut néanmoins préciser que la régulation via le système EnvZ-OmpR pourrait être un peu différente de la souche d'*E. coli* non pathogène K12. En effet, la délétion d'*ompR* conduit chez LF82 à une forte diminution des flagelles alors que chez K12 la motilité est augmentée dans un mutant d'*ompR* (figure 69) suggérant une augmentation du nombre de flagelles. De même chez *E. coli* K12 l'induction de sigma E conduit plutôt à une diminution des pili et flagelles alors qu'il s'agirait d'une augmentation chez LF82.

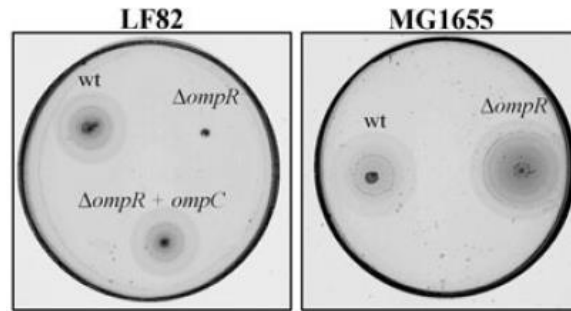


Figure 69 : Test de motilité sur agar mou des souches LF82 ou MG1655 (*E. coli* K12). Plusieurs contextes génétiques sont testés : sauvage (wt), mutant pour *ompR* ($\Delta ompR$) sans/avec surproduction d'*ompC*. Issue de (Rolhion, Carvalho et al. 2007)

Pour finir on peut imaginer que l'implication du système EnvZ-OmpR soit également liée au fait que ce système réponde au stress acide comme nous l'avons vu chez *E. coli* K12 ce qui pourrait être aussi le cas chez LF82. De manière globale les systèmes à deux composants sont importants dans la physiologie bactérienne et notamment au cours du processus d'infection comme on peut le voir avec LF82 et EnvZ-OmpR. Aussi comprendre les mécanismes mis en jeu via ces systèmes au cours de l'infection peut nous renseigner sur la virulence des souches pathogènes.

1.5. Objectifs de la thèse

Nous avons pu voir au cours de cette introduction que l'expression génétique chez les bactéries est un processus finement régulé impliquant de nombreux mécanismes de contrôle, médiés notamment par les systèmes à deux composants et les petits ARN régulateurs. L'objectif de mon travail de thèse a été de comprendre de quelle manière peuvent s'interconnecter les réseaux de régulation impliquant ces deux types d'acteurs agissant souvent à des niveaux de contrôle différents : au niveau transcriptionnel pour les systèmes à deux composants et post-transcriptionnel pour les petits ARN régulateurs.

Je me suis d'abord attachée à caractériser la boucle de rétrocontrôle existant entre les petits ARN OmrA/B et le système à deux composants EnvZ-OmpR et à déterminer quelles propriétés cette boucle procurait à l'expression de chacun d'eux. Je me suis intéressée à cette régulation non seulement chez *E. coli* K12 mais aussi dans le contexte de l'infection de macrophages par une souche pathogène d'*E. coli*, LF82.

Par la suite, j'ai commencé à étudier de manière plus générale les interactions entre systèmes à deux composants et les petits ARN régulateurs, notamment en me penchant sur la régulation de la synthèse du système NarQ-NarP par des petits ARN.

2. Matériels et méthodes

La partie « matériels et méthodes » qui suit concerne les expériences mentionnées en dehors de mon article de recherche. Pour toutes les expériences ayant attiré à ce dernier, se reporter à sa propre partie « matériels et méthodes ».

2.1. Construction des souches d'*E. coli* non pathogènes

Les tableaux 8, 9 et 10 récapitulent les souches, plasmides et oligonucléotides que j'ai utilisés et/ou construits. Les PCR préparatives des produits de construction ont été réalisées avec l'ADN polymérase Phusion (NEB). Les PCR de vérification ont été effectuées sur colonie avec la Purple Taq (Ozyme) ou la Taq classique (NEB). Dans tous les cas, j'ai suivi les recommandations des fournisseurs pour le protocole. Pour combiner les différents allèles mutants et les fusions, des transductions avec le phage P1 ont été généralement utilisées. J'ai par ailleurs utilisé des allèles mutants de la collection KEIO pour construire mes mutants du système NarX-NarL ; NarQ-NarP (Baba, Ara et al. 2006).

Pour préparer nos plasmides nous utilisons le protocole classique décrit dans les kits NucleoSpin Plasmid (mini-préparations) et NucleoBond PC100 (maxi-préparation) commercialisés par la société Macherey-Nagel. Pour construire le pRprAmut2 je me suis servi comme matrice du pRprA et ai utilisé le kit de mutagenèse dirigée QuikChange II Site-directed Mutagenesis (Agilent) avec les oligonucléotides listés dans le tableau 10. Les différentes transformations des souches d'*E. coli* par les plasmides ont été réalisées par une méthode classique dite « Transformation and Storage Solution » ou « TSS » (Chung, Niemela et al. 1989) ou par électroporation.

Tableau 8 : Souches et plasmides utilisés au cours du projet « Autres régulateurs des Omr »

Nom de la souche	Génotype	Références
MG1655	<i>E.coli</i> K-12	F. Blattner
DJ624	DJ480 <i>mal ::lacIq</i>	D. Jin
NM1100	MG1655 mini lambda-Tet	N. Majdalani
MG1004	DJ480λRS-104+8-OmrA (fusion PomrA- <i>lacZ</i>)	(Guillier, Gottesman et al. 2006)
MG1005	DJ480λRS-104+8-OmrB (fusion PomrB- <i>lacZ</i>)	(Guillier, Gottesman et al. 2006)
MG1811	MG1004 <i>ompR ::cat</i>	Souche du laboratoire
MG1812	MG1005 <i>ompR ::cat</i>	Souche du laboratoire
MG2192	MG1004 <i>fadR ::cat</i>	Cette étude
MG2193	MG1005 <i>fadR ::cat</i>	Cette étude
MG2194	MG1004 <i>gcvA ::cat</i>	Cette étude
MG2195	MG1005 <i>gcvA ::cat</i>	Cette étude
MG2196	MG1004 <i>rstAB ::spc</i>	Cette étude
MG2197	MG1005 <i>rstAB ::spc</i>	Cette étude
Plasmide	Caractéristiques	Usage
pHDB3	Ampicilline, (Ulbrandt, Newitt et al. 1997)	Vecteur contrôle
pOmpR	Ampicilline, dérivé de pHDB3 issu d'un crible génétique, (Guillier and Gottesman 2006)	Surproduction d'OmpR
pRstAB	Ampicilline, dérivé de pHDB3 issu d'un crible génétique décrit dans cette étude	Plasmide activateur d'OmrB

Tableau 9 : Souches et plasmides utilisés au cours du projet « Le système à deux composants NarQ-NarP »

Nom de la souche	Génotype	Références
MG1655	<i>E.coli</i> K-12	F. Blattner
DJ624	DJ480 <i>mal ::lacIq</i>	D. Jin
NM1100	MG1655 mini lambda-Tet	N. Majdalani
DH300	<i>rprA142-lacZ</i>	(Majdalani, Hernandez et al. 2002)
DH311	<i>ΔrcsB</i>	(Majdalani, Hernandez et al. 2002)
MG1497	NM1100 <i>crp ::cat</i>	Souche du laboratoire
MG1505	MG1497 <i>mal ::lacIq</i>	Souche du laboratoire
MG1508	MG1505 Ptet- <i>cat-sacB-lacZ</i>	Souche du laboratoire
MG2114	MG1505 <i>rrnBt2-cat-sacB-lacZ-17</i>	Souche du laboratoire
AB1000	MG1508 Ptet- <i>baeR-lacZ</i>	Cette étude
AB1001	MG1508 Ptet- <i>basR-lacZ</i>	Cette étude
AB1002	MG1508 Ptet- <i>evgA-lacZ</i>	Cette étude
AB1003	MG1508 Ptet- <i>kdpE-lacZ</i>	Cette étude
AB1004	MG1508 Ptet- <i>narP-lacZ</i>	Cette étude
AB1005	MG1508 Ptet- <i>phoB-lacZ</i>	Cette étude
AB1006	MG1508 Ptet- <i>rstA-lacZ</i>	Cette étude
AB1007	MG1508 Ptet- <i>uvrY-lacZ</i>	Cette étude
AB1008	MG1508 Ptet- <i>yedW-lacZ</i>	Cette étude
AB1009	AB1000 <i>hfq ::cat</i>	Cette étude
AB1010	AB1001 <i>hfq ::cat</i>	Cette étude
AB1011	AB1002 <i>hfq ::cat</i>	Cette étude
AB1012	AB1003 <i>hfq ::cat</i>	Cette étude
AB1013	AB1004 <i>hfq ::cat</i>	Cette étude
AB1014	AB1005 <i>hfq ::cat</i>	Cette étude
AB1015	AB1006 <i>hfq ::cat</i>	Cette étude
AB1016	AB1007 <i>hfq ::cat</i>	Cette étude
AB1017	AB1008 <i>hfq ::cat</i>	Cette étude
AB1018	AB1004 <i>rprA ::tet</i>	Cette étude
AB1022	MG1508 Ptet- <i>narP-30-lacZ rprA ::tet</i>	Cette étude
AB1029	AB1018 <i>rprA ::tet</i>	Cette étude
AB1033	MG1508 Ptet- <i>narP-105-lacZ rprA ::tet</i>	Cette étude
AB1037	MG1508 Ptet- <i>narPmut2-lacZ rprA ::tet</i>	Cette étude
AB1042	MG2114 PnapF-85+20- <i>lacZ</i>	Cette étude
AB1043	MG2114 PnirB-130+16- <i>lacZ</i>	Cette étude
AB1044	AB1042 <i>narP ::kan</i>	Cette étude
AB1045	AB1043 <i>narP ::kan</i>	Cette étude
AB1048	AB1044 <i>narL ::cat</i>	Cette étude
AB1049	AB1045 <i>narL ::cat</i>	Cette étude
AB1052	AB1042 <i>narL ::cat</i>	Cette étude
AB1053	AB1043 <i>narL ::cat</i>	Cette étude
Plasmide	Caractéristiques	Usage
pBRplac	Ampicilline, tétracycline, P _{LacO-1} cloné dans pBR322 (Guillier, Gottesman et al. 2006)	Vecteur contrôle
pRprA	Ampicilline, tétracycline, P _{LacO-1} cloné dans pBR322 (Mandin and Gottesman 2010)*	Surproduction de RprA
pRprAmut2	Construit à partir du pRprA par mutagenèse dirigée	Surproduction de RprAmut2

* Les autres plasmides utilisés pour la banque de petits ARN sont également du même type et décrit dans cette publication

Tableau 10 : Oligonucléotides utilisés dans le projet « Le système à deux composants NarQ-NarP »

Nom	Séquence	Utilisation
Construction de souches		
3'BaeRlacZ	TAACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGACGATGCGGCAGGCGTCGGCTTCC	<i>baeR-lacZ</i>
5'PtetBaeR	TAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACACCGCTGGAACGGGATTACAGAGAG	<i>baeR-lacZ</i>
3'BasRlacZ	TAACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGACGTTTCCTCATTCGCGACCAGCATATAGC	<i>basR-lacZ</i>
5'PtetBasR	TAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACTTGCAGGAGAGTGAGTGAATGAAAATTC	<i>basR-lacZ</i>
3'EvgAlacZ	TAACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGCGGATTTTGTACGTTGTGCGAATGTG	<i>evgA-lacZ</i>
5'PtetEvgA	TAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACATACTGTGCGAATTATCTTAAAGGAAG	<i>evgA-lacZ</i>
3'KdpElacZ	TAACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGACGATAAACCAGTAGCCAATACC	<i>kdpE-lacZ</i>
5'PtetKdpE	TAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACGCAAACTGCCCTGAACTTGAAG	<i>kdpE-lacZ</i>
3'NarPlacZ	TAACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGACTTGTGCCCCGCGTTGTTGCAGGAAC	<i>narP-lacZ</i>
5'PtetNarP	TAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACTTTATAATGAAAATGATGCCAAAGC	<i>narP-lacZ</i>
3'PhoBlacZ	TAACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGACAAAGCGGGTTGAAAAACGATATCCTGTACC	<i>phoB-lacZ</i>
5'PtetPhoB	TAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACGATTAATGATCGCAACCTATTATTAC	<i>phoB-lacZ</i>
3'RstAlacZ	TAACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGACTTCCCATGCATGAGGCGAAAAAGATAGC	<i>rstA-lacZ</i>
5'PtetRstA	TAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACTTTTTATATCTACCGTGAATGTTATGAAC	<i>rstA-lacZ</i>
3'UvrYlacZ	TAACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGCTGACTTGATAATGTCTCCGCATTAC	<i>uvrY-lacZ</i>
5'PtetUvrY	TAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACGACTAATATCAGTAGCGTTATCC	<i>uvrY-lacZ</i>
3'YedWlacZ	TAACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGACTTTTTTACCCTACGAATGAATAGC	<i>yedW-lacZ</i>
5'PtetYedW	TAGAGATTGACATCCCTATCAGTGATAGAGATACTGAGCACATGAAGATTCTACTTATTGAAG	<i>yedW-lacZ</i>
3'NarPlacZshort	TAACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGACTTACCTGAAAAGGTTGCTTCAGG	<i>narP30-lacZ</i>
narPmut2for	GAT TCT AAT AAA ACG AGT GGA GAC TAC TAT GCC TGA AGC AAC	<i>narPmut2-lacZ</i>
narPmut2rev	GTT GCT TCA GGC ATA GTA GTC TCC ACT CGT TTT ATT AGA ATC	<i>narPmut2-lacZ</i>
narLCmR_for	CCT TTA TTC CCG AAA AAA CTT TCA CAG ACG TCC AAG GAG ATA CCC ATG GAG AAA AAA ATC ACT GG	<i>narL ::cat</i>
narLCmR_rev	GAT GCA TTG TCA AAC GAC GAA CTG CGC TGG GAA CCG TAA TTA CGC CCC GCC CTG CCA CTC ATC	<i>narL ::cat</i>
rrnBt2-napF-for	GGCGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTTTCCCGCCACTCTTTTGATC	<i>napF-lacZ</i>
lacZ-17-napF-rev	CAGTGAATCCGTAATCATGGTCATAGCTGTTTCTGTGTGAATTAATAAGCCATTTTTATAGC	<i>napF-lacZ</i>
rrnBt2-nirB-for	GGCGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCATGAAATATCAGACAATTC	<i>nirB-lacZ</i>
lacZ-17-nirB-rev	CAGTGAATCCGTAATCATGGTCATAGCTGTTTCTGTGTGAGATTTCTTTTCTATTACCGC	<i>nirB-lacZ</i>
Construction de plasmides		
rprAmut2for	GATTTATAAGCATGGAATCCCCACTCTGAAACAACGAATTGCTG	pRprAmut2
rprAmut2rev	CAGCAATTCGTTGTTTCAGAGTGGGGATTTCATGCTTATAAATC	pRprAmut2
Séquençage		
fadR-164for	CTTAACGGTCAGGCAGGAGTGAGGCAAG	<i>fadR ::cat</i>
NarXLfor	TAC CCA TAG TGA GTA CAG TGA CTT C	<i>narXL::kan</i>
NarQfor	GCA ATG TTC ATG TGC CCA GAA TAA G	<i>narQ::kan</i>
NarPforseq	GAT GCC AAA GCG AAC GAC AAG GTT G	<i>narP ::kan</i>
KanR_rev	GCA TTG CAT CAG CCA TGA TGG ATA C	<i>Kan^R (KEIO)</i>
narX1594for	CTG ACC GTC CAG GAT AAC GGC TGC	<i>narL::cat</i>
narLseq_rev	GCT GGA CGC TGA TCA TGC CTG ACT GG	<i>narL::cat</i>

2.2. Conditions de culture des souches d'*E. coli* non pathogènes

Le milieu utilisé pour la culture des différentes souches est le milieu Luria-Bertani ou LB. Il existe diverses compositions de LB. De façon générale, ce milieu est composé de peptides et peptones de caséine, de vitamines, d'oligo-éléments et de minéraux.

Composition de notre milieu LB : tryptone (10 g/L), extrait de levure (5g/L) et NaCl (10 g/L).

J'ai également utilisé du milieu minimum glucose dont la composition est la suivante : milieu minimum A additionné de 1 mM de MgSO₄ et 0.2% de glucose. Pour le crible génétique 0.004% de Xgal a été ajouté sur boîte.

Enfin pour les tests en anaérobiose je me suis servie de MOPS glucose 1X additionné ou non de 40 mM de NaNO₃ (tableau 11).

Tableau 11 : Composition du milieu MOPS pour les cultures en anaérobiose

Composants MOPS 10X	Concentration finale
Tampon MOPS pH 7.4	400 mM
Tricine pH 7.4	40 mM
FeSO ₄	20 mM
MgSO ₄	5 mM
NH ₄ Cl	100 mM
Composition MOPS Glucose 1X	Concentration finale
MOPS 10X	1X
Glucose	0.8%
Tampon phosphate pH 7.4	2.6 mM
Cas amino-acides	1%

Les différents antibiotiques utilisés ont été l'ampicilline à 150 µg/mL (Amp) en milieu liquide et 100 µg/mL en milieu solide, la kanamycine à 25 µg/mL (Kan), la tétracycline à 10 µg/mL (Tet) et le chloramphénicol à 5 ou 10 µg/mL (Cm).

2.3. Souches de LF82 et infection des macrophages

2.3.1. Construction des souches

Une souche de LF82 (Glasser, Boudeau et al. 2001) délétée pour le gène *bla* (LF82 Δbla – don du laboratoire de Nicolas Barnich à Clermont-Ferrand) et transformée par le plasmide pKOBEGA (Chaveroche, Ghigo et al. 2000) a servi de base pour la construction du mutant $\Delta ompR$. Le plasmide porte les gènes de recombinaison du phage Lambda Red et permet donc de réaliser une recombinaison homologue classique. La délétion a été introduite par recombinaison avec un fragment PCR portant les régions d'homologie du gène *ompR* interrompu par une cassette de résistance au chloramphénicol. Les bornes de la délétion sont identiques au mutant d'*E. coli* K12 décrit précédemment.

2.3.2. Infection des macrophages

La lignée de cellulaire utilisée est une lignée de monocytes humains THP-1. Ces monocytes sont différenciés en macrophages grâce à un traitement à la PMA (phorbol 12-myristate 13 acétate). Dans un premier temps les monocytes sont décongelés et cultivés à 37°C en suspension dans du milieu de culture RPMI 1640 (GE Healthcare) supplémenté avec 10% de sérum de veau fœtal (Dutscher) et de la glutamine 2mM final (GE Healthcare). Le milieu de culture est changé tous les deux ou trois jours et les cellules réparties à chaque passage dans des flasques de 75 cm² de façon à ce qu'elles puissent se multiplier. Les monocytes sont dénombrés en cellule de Malassez puis le nombre de cellules nécessaire à l'expérience est prélevé.

Afin d'induire la différenciation des monocytes en macrophages de la PMA à 20 ng/mL final est ajoutée à la suspension cellulaire avant d'ensemencer les plaques 24 puits (1mL/puit). Après environ 18h d'incubation à 37°C les cellules doivent former des amas et adhérer au fond des puits. Le milieu est retiré par retournement et les cellules lavées 2 fois au PBS 1X pour retirer les éventuels monocytes non différenciés. Du milieu frais est ajouté, les cellules sont ainsi prêtes pour l'infection.

Les suspensions de bactéries (MOI 100 bactéries par cellules) sont réparties dans les différents puits. La suspension de bactérie est une préculture nocturne diluée à une DO=0.5 dans du PBS 1X. Afin de synchroniser l'entrée des bactéries dans les macrophages, les plaques sont centrifugées 10 min à 900 rpm à température ambiante puis incubées 10 min à 37°C. Le début de la centrifugation

constitue le T0 de l'infection. A l'issue des 10 min d'incubation, les cellules sont lavées deux fois au PBS 1X afin d'ôter la majeure partie des bactéries non internalisées dans les macrophages. 1 ml de milieu frais contenant de la gentamicine à 50 µg/ml final est ajouté après ces deux lavages. La gentamicine permet d'éviter la prolifération des éventuelles bactéries restantes après lavages mais aussi des bactéries relarguées dans le milieu suite à la mort de certains macrophages. Les plaques sont ensuite remises dans l'incubateur 40 min (pour le point 1h post-infection), 2h40 min (pour le point 3h post-infection), 5h40 min (pour le point 6h post-infection) ou 23h40 min (pour le point 24h post-infection).

2.3.3. Extraction des ARN au Trizol®

Afin d'extraire les ARN des bactéries et des macrophages après l'infection, une méthode d'extraction au Trizol® est utilisée. Le milieu de culture est retiré (les macrophages étant des cellules adhérentes) et les cellules lavées 2 fois avec 1ml de PBS 1X froid. 1ml de Trizol® froid est ensuite ajouté dans chacun des puits et les cellules sont décollées et mélangées avec le Trizol® à l'aide d'un grattoir. La suspension est transférée dans un Eppendorf et 200µl de chloroforme sont ajoutés. Le tube est vortexé vigoureusement pendant 15 sec et incubé 3 min à température ambiante. Cette étape est suivie d'une centrifugation de 5 min à 6500 rpm à 4°C afin de séparer les phases aqueuses et phénoliques. Avec cette méthode la phase phénolique apparaît rose et la phase aqueuse translucide. La phase aqueuse est transférée dans un Eppendorf et on y ajoute 500µl d'isopropanol pour précipiter les ARN. Après 10 min à température ambiante, les échantillons sont centrifugés à 13500 rpm pendant 10 min à 4°C. On retire la totalité du surnageant et le culot est lavé deux fois avec de l'éthanol à 75% froid. Après séchage le culot est repris dans 25 µl d'eau DEPC.

2.4. Northern blot

La procédure d'extraction d'ARN au phénol et l'analyse par Northern blot que nous utilisons est décrite dans (Brosse, Korobeinikova et al. 2016). Les séquences des sondes biotinylées utilisées pour révéler les ARN sont listées dans le tableau 12.

Tableau 12 : Sondes biotinylées pour les Northern

Sonde	Séquence
RprA-bio	Bio-CAGGGGATTTCCATGCTTATAAATCAATATGTTGATTTATAACc
SsrA-bio	Bio-CGCCACTAACAACTAGCCTGATTAAGTTTTAACGCTTCA
5S-bio	Bio-CTACCATCGGCGCTACGGCGTTTCACTTCTGAGTTCG
OmrA-bio	Bio-CAGGTTGGTGCAAGAGACAGGGTACGAAGAGCGTACCG
OmrB-bio	Bio-CGCAGGCTGGTGTAATTCATGTGCTCAACCCGAAGTTGA
Spot42-bio	Bio-GTAAAAGGTCTGAAAGATAGAACATCTTACCTCTGTACCC

2.5. Western blot « Phostag »

J'ai optimisé cette technique au laboratoire afin de pouvoir discriminer sur Western blot la forme phosphorylée de la forme non phosphorylée de la protéine OmpR. L'originalité de la technique réside dans l'utilisation d'un réactif appelé Phos-tag. Ce dernier se fixe au groupement phosphate et ralentit ainsi la migration sur gel d'acrylamide des protéines phosphorylées. La liaison phosphate sur les protéines des systèmes à deux composants étant très labile les extraits cellulaires doivent être préparés de façon plus « douce » que pour un Western blot classique. Nous utilisons donc un mélange commercial de détergent, le Bugbuster (Novagen), afin de lyser les cellules sans chauffer les extraits. Pour la plupart des étapes, les extraits doivent être conservés dans la glace pour limiter toute dégradation de la liaison phosphate.

5mL de culture sont prélevés, refroidis dans la glace puis centrifugés à 8000 rpm à 4°C pendant 10 min. La majeure partie du surnageant est ôtée. Le culot est repris dans le reliquat de surnageant puis transféré dans un Eppendorf de 1,5 mL et centrifugé à 8000 rpm à 4°C pendant 10 min. La totalité du surnageant est enlevée et le culot repris dans du tampon Bugbuster 1X à une concentration de 40 DO/mL. Après 30 sec de vortex à température ambiante, est ajouté ¼ de volume de tampon de dépôt 5X-SDS (Tris-HCl 250 mM pH 6,8 ; SDS 8 % ; glycérol 25 % ; β -mercaptoéthanol 0,572 M ; bleu de bromophénol). Ce mélange est centrifugé à 13200 rpm à 4°C

pendant 10 min. Le lysat clarifié est prélevé et transféré dans un Eppendorf de 1,5 mL et conservé dans la glace. 20 μ L de ce lysat clarifié est déposé sur gel SDS-PAGE (composition détaillée en annexe 2) contenant le réactif Phostag (25 μ M) et du MnCl₂ (50 μ M) dans le gel de séparation. La migration s'effectue dans du tampon Tris-Glycine SDS 1X. Comme pour le gel classique elle se fait en deux étapes : 30 min à 80 V puis 2h-2h30 à 130 V. Le gel est démoulé et équilibré pendant 10 min dans une solution d'EDTA 1 mM afin de chélater les ions manganèse du gel qui pourraient perturber le transfert. Le gel est ensuite immergé pendant 10 min dans le tampon de transfert. Les protéines sont transférées sur la membrane de nitrocellulose (Hybond C Super – GE Healthcare) par transfert électrique de 2h à 100 V dans un tampon Tris-Glycine-SDS. Après transfert la membrane est incubée 30 min dans une solution de blocage (PBS 1 X ; Tween 0,05 % ; Lait demi-écrémé en poudre 5 %). Elle est ensuite incubée avec l'anticorps primaire anti-OmpR (lapin) dilué au 1/1500ème dans de la solution de blocage pendant une nuit à 4°C sous agitation lente. Le lendemain trois lavages au PBS-T (PBS 1 X ; Tween 0,05 %) sont réalisés avant d'incuber la membrane avec l'anticorps secondaire (anti-lapin couplé à la peroxydase HRP, Sigma A6154) dilué au 1/2000ème dans de la solution de blocage. Cette incubation d'une heure est suivie de trois lavages au PBS-T. La révélation se fait ensuite grâce au kit Covalight (Covalab) et à l'appareil d'acquisition de chimiluminescence Chemidoc (Biorad).

3. Résultats

3.1. Circuit mixte de régulation entre les petits ARN OmrA et OmrB et le système à deux composants EnvZ-OmpR chez *Escherichia coli*

3.1.1. Résultats publiés

Dans l'article qui suit, nous avons montré qu'OmpR activait directement la transcription d'*omrA* et d'*omrB* en se fixant à leur promoteur. Cette activation permet la production des petits ARN OmrA et OmrB qui, via leur extrémité 5' conservée, ciblent à leur tour plusieurs ARN messagers-cibles et notamment le messager *ompR-envZ*.

En accord avec la robustesse du système à deux composants EnvZ-OmpR, nous avons montré que ce contrôle d'*ompR-envZ* par les Omr n'affectait pas le niveau de la forme phosphorylée d'OmpR, posant ainsi la question de l'intérêt d'une telle régulation.

Au moins une partie de la réponse à cette question est venue de l'étude plus poussée de la régulation d'*omrA* et d'*omrB* par OmpR. En effet, nos résultats ont aussi démontré que certaines cibles d'OmpR et en particulier OmrA/B, ne sont pas robustes et répondent à la fois à des variations du niveau d'OmpR-P, mais également de la forme non phosphorylée d'OmpR. Ce phénomène permet donc à ces ARN régulateurs, via le contrôle des niveaux d'OmpR, de limiter leur propre synthèse en ayant un effet limité sur l'expression des autres cibles d'OmpR comme les porines OmpC et OmpF.

Ce travail met en lumière l'originalité de la régulation existant entre le système à deux composants EnvZ-OmpR et les petits ARN OmrA et OmrB. Il soulève également diverses interrogations et notamment celle de savoir si ce type de régulation avec une robustesse différentielle des cibles peut être retrouvé chez d'autres systèmes à deux composants chez *E. coli*.

Unexpected properties of sRNA promoters allow feedback control via regulation of a two-component system

Anaïs Brosse¹, Anna Korobeinikova¹, Susan Gottesman² and Maude Guillier^{1,*}

¹CNRS UMR8261, Associated with University Paris Diderot, Sorbonne Paris Cité, Institut de Biologie Physico-Chimique, 75005 Paris, France and ²Laboratory of Molecular Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

Received July 19, 2015; Revised June 21, 2016; Accepted July 07, 2016

ABSTRACT

Two-component systems (TCS) and small regulatory RNAs (sRNAs) are both widespread regulators of gene expression in bacteria. TCS are in most cases transcriptional regulators. A large class of sRNAs act as post-transcriptional regulators of gene expression that modulate the translation and/or stability of target-mRNAs. Many connections have been recently unraveled between these two types of regulators, resulting in mixed regulatory circuits with poorly characterized properties. This study focuses on the negative feedback circuit that exists between the EnvZ-OmpR TCS and the OmrA/B sRNAs. We have shown that OmpR directly activates transcription from the *omrA* and *omrB* promoters, allowing production of OmrA/B sRNAs that target multiple mRNAs, including the *ompR-envZ* mRNA. This control of *ompR-envZ* by the Omr sRNAs does not affect the amount of phosphorylated OmpR, i.e. the presumably active form of the regulator. Accordingly, expression of robust OmpR targets, such as the *ompC* or *ompF* porin genes, is not affected by OmrA/B. However, we find that several OmpR targets, including OmrA/B themselves, are sensitive to changing total OmpR levels. As a result, OmrA/B limit their own synthesis. These findings unravel an additional layer of control in the expression of some OmpR targets and suggest the existence of differential regulation within the OmpR regulon.

INTRODUCTION

Regulation of gene expression plays a crucial role in the rapid adaptation of bacteria to their environment, a process that allows them to colonize and survive in various media.

This control can take place at all stages of gene expression and via a great diversity of precise molecular mechanisms.

For instance, examples of transcriptional control have been reported at the level of transcription initiation, elongation or termination, and the regulators can be DNA- or RNA-binding proteins, but also RNA molecules. Among the most widely used regulators of transcription in bacteria are the two-component systems (TCS), which are typically composed of a sensor protein and its cognate response regulator (1). In response to specific stimuli, the sensor controls the phosphorylation status of the regulator by acting as a kinase and/or as a phosphatase. In the phosphorylated state, the regulator typically controls transcription of multiple genes. More than 30 TCS have been identified in the *Escherichia coli* genome and allow both sensing and adaptation to the environment, in response to extremely diverse input signals, such as phosphate limitation, low magnesium, pH or osmolarity.

EnvZ-OmpR is one of the most extensively studied TCS (2). EnvZ autophosphorylates upon sensing cognate stimuli and controls the level of phosphorylated OmpR (OmpR-P). Even though the precise signal detected by EnvZ is still unclear, it is known that signals such as increasing osmolarity, temperature or acid pH (among others) induce the activity of the EnvZ-OmpR TCS (3). OmpR directly controls the transcription of several genes, in particular those encoding the major porins in *E. coli*, OmpC and OmpF. While *ompC* expression is activated in conditions of high osmolarity (high OmpR-P), OmpR represses or activates *ompF* expression in high and low osmolarity, respectively (4 and references therein). In addition to these two genes, OmpR also directly controls transcription of other genes such as *dtpA* (a.k.a. *tppB*) (5), *flhDC* (6), *csgD* (7) and *bolA* (8), that respectively encode an oligopeptide transporter, the regulator for flagella biogenesis, the activator of curli production and a transcriptional regulator also involved in controlling the switch from a planktonic to a sessile lifestyle (9). Note that regulation of *dtpA* by EnvZ-OmpR was investigated in

*To whom correspondence should be addressed. Tel: +33 1 58 41 52 38; Fax: +33 1 58 41 50 25; Email: maude.guillier@ibpc.fr
Present address: Anna Korobeinikova, Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region 142290, Russia.

detail; interestingly, it is different from regulation of both *ompC* and *ompF* in that *dtpA* transcription does not seem to respond to changes in osmolarity (5). Furthermore, OmpR was found to bind to several dozen sites in both *E. coli* and *Salmonella* in a recent ChIP-on-chip analysis, indicating the existence of more direct targets (10). Although the OmpR regulons were found to differ in the two species, with only a few common targets, OmpR controls genes involved in pH homeostasis and acid stress response in both *E. coli* and *Salmonella*. In agreement with this, OmpR appears to play an important role in acid resistance (11). At least 4 small RNAs (sRNAs) are also known to be controlled by EnvZ-OmpR: MicC, MicF, OmrA and OmrB (12–14) (see below).

Regulatory RNAs have emerged as major regulators of gene expression in recent years, in virtually all organisms. In bacteria, they are typically rather short (a few hundred nts or less) and are therefore referred to as small RNAs (sRNAs). So far, the best characterized class corresponds to sRNAs that can base-pair with target-mRNA(s) encoded by another locus of the genome and, as a result, regulate either positively or negatively, their translation and/or stability (15). This base-pairing is in most cases mediated by short and imperfect sRNA-mRNA duplexes and in several bacterial species, such as *E. coli* for instance, requires the action of a hexameric RNA chaperone called Hfq (16,17).

Base-pairing sRNAs are involved in the regulation of diverse cellular functions as crucial as iron homeostasis or quorum-sensing. Interestingly, many sRNAs target genes that encode outer membrane proteins (OMP) or are involved in synthesis of membrane appendages such as curli or flagella (14,18–22); this is true in particular for the various OmpR-dependent sRNAs. MicC and MicF were, for instance, shown to repress expression of *ompC* and *ompF* porin genes, respectively (13,23). In addition, MicC also represses the synthesis of OmpD, another abundant porin in *Salmonella* (24). OmrA/B are involved in membrane re-modeling as well, since they negatively regulate the synthesis of several OMPs, such as the OmpT protease and the CirA, FecA and FepA receptors for iron-siderophore complexes (25). They also repress the synthesis of several transcriptional regulators such as CsgD (18) and FlhD₂C₂ (21), the key regulators for curli formation and flagella biogenesis respectively, as well as the EnvZ-OmpR TCS, their own transcriptional activator (26).

sRNAs play major roles in shaping gene expression in bacteria. They are involved in different regulatory circuits, and they most often participate in mixed regulatory networks that combine transcriptional and post-transcriptional controls, mediated by proteins and sRNAs respectively (27,28). While the occurrence of sRNAs in these circuits is clearly recognized, the properties they can provide are still poorly understood in most cases. sRNAs were in particular reported in feedforward motifs (FFM), where a regulator controls expression of a target gene both directly and indirectly via control of another regulator. The Spot 42 sRNA participates, for instance, in a multioutput coherent FFM with the global regulator CRP to repress expression of multiple genes involved in the use of diverse non-preferred carbon sources. Interestingly, this mixed FFM allowed an accelerated response following circuit deactivation, a dynamic property that had not been observed be-

fore with purely transcriptional FFM, suggesting a unique role for sRNAs in these motifs (29). FFM can also involve positively acting sRNAs: RprA sRNA is an activator of the synthesis of RicI, a membrane protein that inhibits conjugation, both by direct base-pairing with the *ricI* mRNA and via positive control of RpoS. Because the resulting mixed coherent FFM operates with an AND-logic (i.e. both RprA and RpoS are required for *ricI* activation), it ensures that synthesis of RicI, and hereby conjugation inhibition, occurs under sustained activation of RprA only, with a delay in comparison to the synthesis of RpoS (30). Another recurrent motif involving sRNAs and regulatory proteins is the feedback circuit where one regulator controls and is controlled by a second regulator. This study focuses on one example of such a feedback: the direct repression of *ompR-envZ* expression by OmrA/B, whose transcription is activated by this TCS. Our findings reveal an exquisite elaboration in the way in which the well-studied EnvZ-OmpR TCS works, which allows for autocontrol of OmrA/B.

MATERIALS AND METHODS

Strains and plasmids, general microbiology techniques

Strains used in this study are listed in Supplementary Table S1 and their construction is described in supplementary material. Briefly, mutant alleles were either obtained from existing sources or engineered by lambda Red recombination, and moved by transduction when required. Strains were grown in LB, and when necessary, antibiotics were used at the following concentrations: 150 mg/ml ampicillin, 25 mg/ml kanamycin, 10 mg/ml chloramphenicol and 10 mg/ml tetracycline. When required, expression of various sRNAs from pBRplac derivatives was induced using 100 mM IPTG. The Phusion DNA polymerase was used for cloning and amplification of DNA fragments for recombineering.

OmpR-His purification and *in vitro* phosphorylation

Protein OmpR was expressed and purified from strain BL21 (DE3) transformed with a pET15b plasmid carrying the *ompR* gene with a 6His tag at its N-terminus under control of a T7 promoter (31). Purification was mostly as described previously. Briefly, cells were diluted 250-fold from an overnight culture into fresh LB-Ampicillin and grown to exponential phase. When the optical density at 600 nm reached 0.6, synthesis of OmpR-His was induced with addition of 1 mM IPTG for 3 h. Cells were then centrifuged and, after freezing, pellets were resuspended in A buffer (20 mM sodium phosphate buffer pH 7.44, 0.5 M NaCl, 10 mM imidazole), and protease inhibitors (Complete mix, Roche) and 1 mg/ml lysozyme were added. After sonication and clarification by centrifugation, the supernatant was loaded onto a column of Ni-NTA agarose resin (Macherey-Nagel) equilibrated in A buffer. After washing, OmpR-His was eluted with increasing concentrations of imidazole. Fractions containing sufficiently pure OmpR-His were then pooled and dialyzed against TEGD buffer (20 mM Tris-HCl pH 7.6, 0, EDTA 0.1 mM, glycerol 5%, NaCl 0.4M, DTT 0.1mM).

In vitro phosphorylation was performed prior to each transcription reaction. For this purpose, 10 mM OmpR-

His was incubated in phosphorylation buffer (50 mM Tris-HCl pH 7.6, 20 mM MgCl₂, 50 mM KCl) in presence of 25 mM acetylphosphate (Sigma reference 01409) at 37 °C for 2 h with moderate shaking. Phosphorylation was systematically checked by Phos-Tag electrophoresis and found to be reproducibly around 50%.

***In vitro* transcription**

In vitro transcription reactions were performed according to (32) and (33) using plasmid DNA as templates. 50 ng template DNA was incubated at 37 °C for 10 min in the reaction buffer (40 mM HEPES pH 7.4, 60 mM potassium glutamate, 2 mM MgCl₂, 0.05% Nonidet P-40) in presence of NTPs (final concentration of ATP, CTP and GTP is 0.2 mM, 10 nM UTP, 2.5 mCi [*a*-³²P] UTP) and 10 mg/ml bovine serum albumin (BSA). A constant volume of phosphorylated or unphosphorylated OmpR-His (or buffer for the control) was added at the required concentration and incubation was continued for 10 min. 1 ml of *E. coli* RNA polymerase (Epicentre) at 0.5 units/ml was then added and transcription was allowed to proceed for 20 min at 37 °C; it was then stopped by addition of 1 volume of stop solution (95% formamide, 25 mM EDTA, 0.05% xylene cyanol and bromophenol blue). Transcription products were analyzed on a 6% sequencing gel next to 5 labeled pBR322/MspI ladder. Quantification was performed with ImageQuant software and used RNAI transcript as a reference to normalize the different reactions.

Phos-Tag electrophoresis

Protein extraction and gel electrophoresis for separation of phosphorylated and unphosphorylated OmpR was mostly as described previously (34) with the following modifications. Briefly, cells from 5 ml of culture were collected by centrifugation and resuspended in Bugbuster detergent mix (Merck) at a concentration equivalent to 40 OD₆₀₀/ml. 1/4 volume of SDS loading buffer 5× (250 mM Tris-HCl pH 6.8, 8% sodium dodecyl sulfate (SDS), 25% glycerol, 572 mM β-mercaptoethanol, 0.10% bromophenol blue) was then added. After a new centrifugation, supernatant was collected and 20 ml were loaded onto a Phos-Tag containing gel. Gel composition and electrophoresis was as in (34). After migration, gel was equilibrated in 1 mM ethylenedi-aminetetraacetic acid (EDTA) to chelate Mn²⁺ ions, and transfer was then as for standard protein gels (see below).

Western blot

After a 500-fold dilution from an overnight culture into fresh medium, cells were grown to exponential phase, collected by centrifugation and pellets were resuspended in SDS-loading buffer containing DTT (Biolabs). A volume equivalent to 0.15 A₆₀₀ was then loaded on a 10% SDS-PAGE gel and after separation, proteins were transferred onto an Hybond C-super membrane (GE Healthcare). Membrane was blocked for 1 h prior to overnight incubation with the anti-OmpR antibody diluted 1500-fold in phosphate-buffered saline (PBS)-0.05% Tween 20. After

washing and incubation with the secondary antibody, detection was performed using the Covalight chemiluminescent reagent kit (Covalab).

-Galactosidase activity

Cells were diluted 500-fold into fresh medium from an overnight culture and grown to an optical density of 0.4 at 600 nm. The β-galactosidase activity was then measured following Miller's protocol (35) and results presented here correspond to the average of at least two independent experiments (except for Figure 4C where a representative experiment is shown, with triplicates presented in Figure S6). For the results that are shown as ratios, activities are given in Supplementary Table S3.

RNA extraction and northern-blot analysis

Total RNA was extracted from cells grown to exponential phase after a 500-fold dilution from overnight culture into fresh medium, using hot phenol as previously described (14). A constant amount of total RNA was then separated on an 8% acrylamide gel in TBE 1× (for northern for OmrA/B, *bolA* and *SsrA*) or on a 1% agarose gel in MOPS 1× (for *ompC*, *ompF* and *SsrA*) prior to electric transfer onto an Hybond-N+ membrane (GE Healthcare). Detection was then performed using 5 end biotinylated oligonucleotides (sequence in Supplementary Table S2) and the Brightstar biodetect kit (Ambion) following manufacturer's instructions.

RESULTS

The OmpR response regulator directly activates transcription of *omrA* and *omrB* genes

Our previous results identified OmpR as an activator of *omrA/B* transcription; both OmrA and OmrB levels increased with high osmolarity or upon procaine addition, in an EnvZ- and OmpR-dependent process (25). Further-more, recent OmpR ChIP experiments suggest direct binding of OmpR to these promoters, indicating that activation by OmpR is direct regulation (10). To further investigate this, *in vitro* transcription reactions were performed using templates driven from the *omrA* or *omrB* promoter in presence of increasing amounts of OmpR, phosphorylated or not; the *ompC* promoter was included as a positive control for activation by OmpR (Figure 1). All three promoters, (starting 114 nts upstream of the TSS for *omrA/B* or 437 nts upstream for *ompC*), were cloned in the pRLG770 vector (36), allowing the production of ~200 nt-transcripts carrying the first 50 nts of each RNA fused to 5S rRNA. An N-terminal His₆-tagged version of OmpR (31) was purified and, when indicated, phosphorylated *in vitro* in the presence of acetylphosphate. The efficiency of the phosphorylation reaction was systemically assessed for each experiment and was reproducibly found to be around 50% (Supplementary Figure S1C). In these experiments, OmpR or OmpR-P was used at 25-, 50- and 125-fold excess over the template DNA (in molar ratio), which corresponds to a concentration of respectively 15, 30, 75 nM; an additional concentration of 250-fold (or 150 nM) was also included for

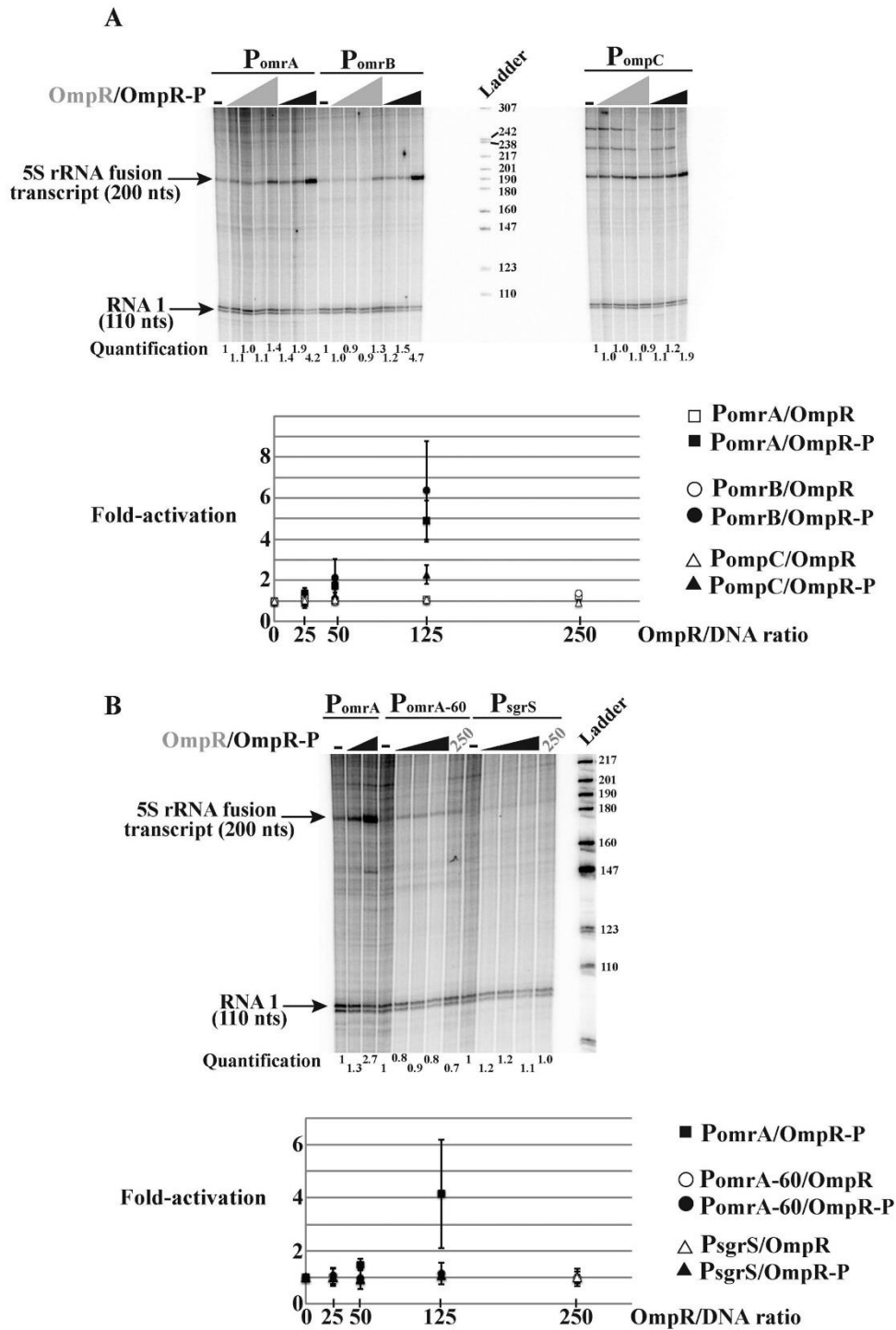


Figure 1. OmpR-P activates *omrA* and *omrB* transcription *in vitro*. (A) *In vitro* transcription reactions were performed using plasmid templates carrying regions -114 to $+50$ of either the *OmrA* or *OmrB* promoter, or -437 to $+50$ of *ompC* promoter (relative to P1 transcription start site), fused to 5S to give a 200 nt transcript. When indicated, OmpR was added at a molar ratio of 25, 50, 125 and 250 with the template DNA, and OmpR-P at a molar ratio of 25, 50 and 125. Products were separated on a sequencing gel and a portion of a representative gel is shown where transcripts driven from *PomrA*, *PomrB* or *PompC* promoter and the internal control RNA1 are visible. The ladder corresponds to radiolabeled pBR322/MspI digestion products and indicated sizes refer thus to double-stranded DNA fragments. Quantification of 5S rRNA fusion transcripts was performed after normalization to RNA1 and the level is arbitrarily set at 1 for control without OmpR for each promoter. The graph below represents the average activation of transcription from *PomrA*, *PomrB* or *PompC* in presence of increasing OmpR or OmpR-P concentrations, as calculated from two independent experiments. Error bars indicate standard deviations and, when not visible, are smaller than the symbols. (B) Similar *in vitro* transcription reactions were performed using templates carrying a short version of *omrA* promoter (*PomrA-60*, carries nts -60 to $+50$ relative to TSS) or the *sgrS* promoter (from nts -70 to $+50$). OmpR-P was used at a molar ratio of 25-, 50- and 125-fold with template DNA, and OmpR at 250-fold. *PomrA* was used here as a positive control for activation by 50- or 125-fold OmpR-P. Full gels are shown in Supplementary Figure S1.

unphosphorylated OmpR. As a comparison, total OmpR cellular concentration was estimated at 3500 OmpR molecules per cell (ca. 3.5 mM) (37) with ~25% being phosphorylated in the presence of procaine or sucrose (34). The transcription products were separated on a sequencing gel and normalized to the RNA1 transcript (Figure 1A; a full representative gel is shown in Supplementary Figure S1). Transcription of *ompC* was, as expected, dependent on OmpR-P, but only at the highest concentration; in this condition, *ompC* transcription was activated by 1.9-fold relative to the control without OmpR. OmpR-P had a stronger effect on the *omr* promoters: a 50-fold excess of OmpR-P allowed activation of transcription from both the *omrA* and *omrB* promoters (by 1.9- and 1.5-fold, respectively, relative to the controls without OmpR) while a 125-fold excess further increased *omr* transcription (by 4.2- and 4.7-fold respectively). In contrast, the unphosphorylated form of OmpR did not significantly activate transcription from these three promoters. As an additional control, transcription driven by two non-OmpR-regulated promoters was studied as well. The first control promoter is a truncated version of P_{omrA} , starting only 60 nts upstream of the TSS, which is not sufficient to allow activation by OmpR *in vivo* (14). The second control is the promoter of *SgrS*, another Hfq-dependent sRNA, whose transcription responds to phosphosugar stress via the *SgrR* activator and is independent of OmpR (38 and see below). Addition of OmpR, phosphorylated or not, had no effect on transcription originating from either control promoter, showing that the *in vitro* activation of *omrA*, *omrB* or *ompC* transcription is specific (Figure 1B). Altogether, these results strongly suggest that the previously reported control of *OmrA* or *OmrB* by EnvZ-OmpR is direct.

OmrA and OmrB are highly sensitive to various mutants of EnvZ-OmpR TCS

While OmpR regulates many genes, the details of regulation are different for different promoters, allowing differential expression of members of this regulon. For instance, *ompC* transcription is induced at high osmolarity, whereas *ompF* transcription is repressed at high osmolarity, but activated at low osmolarity (2). Furthermore, the analysis of *dtpA* expression in different mutants of *envZ* and *ompR* showed that it behaved differently from both *ompC* and *ompF* (5). We were interested in comparing *omrA* and *omrB* regulation by EnvZ and OmpR to that for *ompC* and *ompF*, the best-studied OmpR-regulated genes, with the expectation that this should provide a better sense of when expression of these sRNAs is physiologically important. For this purpose, we chose to analyze expression of *lacZ* fusions to the promoters of *omrA/B* in strains carrying different mutant alleles of *envZ* or *ompR*. As controls, expression of *ompC*- or *ompF*-*lacZ* promoter fusions was measured in the same mutants.

The *envZ* mutant alleles used were *envZ390* and *envZ473*. *EnvZ390* was isolated in a screen for EnvZ mutants that were deficient in kinase activity but that retained phosphatase activity (39). With such a mutant, the level of phosphorylated OmpR is therefore expected to be low. In contrast, the *EnvZ473* mutant retained kinase activity, but

lost its phosphatase activity; the level of OmpR-P should therefore be high (40,41). This was confirmed *in vivo* through protein electrophoresis in acrylamide gels containing Phos-Tag, a phosphate chelator that allows separation of phosphorylated proteins (42), followed by immunoblotting with an OmpR specific antibody (a kind gift from Dr Ann Stock). In the wt strain, OmpR-P was detected after treatment with procaine, an anesthetic known to activate the EnvZ-OmpR TCS. In the *envZ390* and *envZ473* strains, our results confirmed the respectively low and high levels of OmpR-P, with or without procaine (Figure 2A; an independent experiment is shown in Supplementary Figure S2A). In the *envZ473* mutant strain, note that the vast majority of total OmpR was phosphorylated, which differs from what was observed by Barbieri *et al.* (34). These differences were found to be due to the different strains used in the two studies, an MC4100 derivative in (34) and an MG1655 derivative in this study (Supplementary Figure S2B). Another striking result from Figure 2A is that the level of total OmpR was significantly increased in the *envZ473* allele. This result was confirmed when the same extracts were separated on a classical sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE) gel (Supplementary Figure S2C). This increase is most likely due to a transcriptional up-regulation as expression of a promoter fusion between *Pompr* and *lacZ* was increased by more than 10-fold in the same mutant (Supplementary Figure S2D).

Two mutants of the OmpR response regulator were also used. These were isolated in a screen for OmpR mutants yielding either an $OmpC^-$, $OmpF^+$ phenotype (*ompR2* class) or, in contrast, an $OmpC^+$, $OmpF^-$ phenotype (*ompR3* class) (43). The mutant representatives of these two classes used in this study are, for *ompR2*, *ompR472*, where Val203 is changed to Met and, for *ompR3*, *ompR107*, where Arg15 is changed to Cys. These two mutants were reported to be defective in dephosphorylation (44,45). In previous *in vitro* studies, *OmpR107* bound to *ompC* and *ompF* promoters similarly to the wt protein, whereas *OmpR472* displayed a higher affinity for the *ompF* promoter region and a weaker affinity for *ompC* (44,46). Again, Phos-Tag electrophoresis confirmed the increase in the OmpR-P/OmpR ratio in these two mutant strains *in vivo* (Figure 2A and Supplementary Figure S2A).

We then compared the effect of the various mutants on *omrA*, *omrB*, *ompC* and *ompF* transcription. For this purpose, we constructed *lacZ* fusions to the promoters of each of these genes, carrying all upstream sequences known to be required for regulation by OmpR (see Supplementary Figure S3 for a schematic representation of the different fusions used in this work, and experimental procedures for details). As expected, expression of the four resulting fusions was strongly dependent on *ompR*: the β -galactosidase activities were decreased from 7- to 78- fold in an *ompR^-* strain in the absence of procaine (Figure 2B). A similar pattern was observed in the presence of procaine (data not shown).

The effect of the *envZ390* allele was similar to that of the *ompR* deletion for *omrA*, *omrB* and *ompC* fusions (compare fold activation in Figure 2B to fold repression in Figure 2C), suggesting that OmpR-P is strictly required for activation at these promoters. While *ompF* transcription was significantly decreased as well in the *envZ390* context (by 3-fold), the effect of this mutant allele was in this case not nearly as

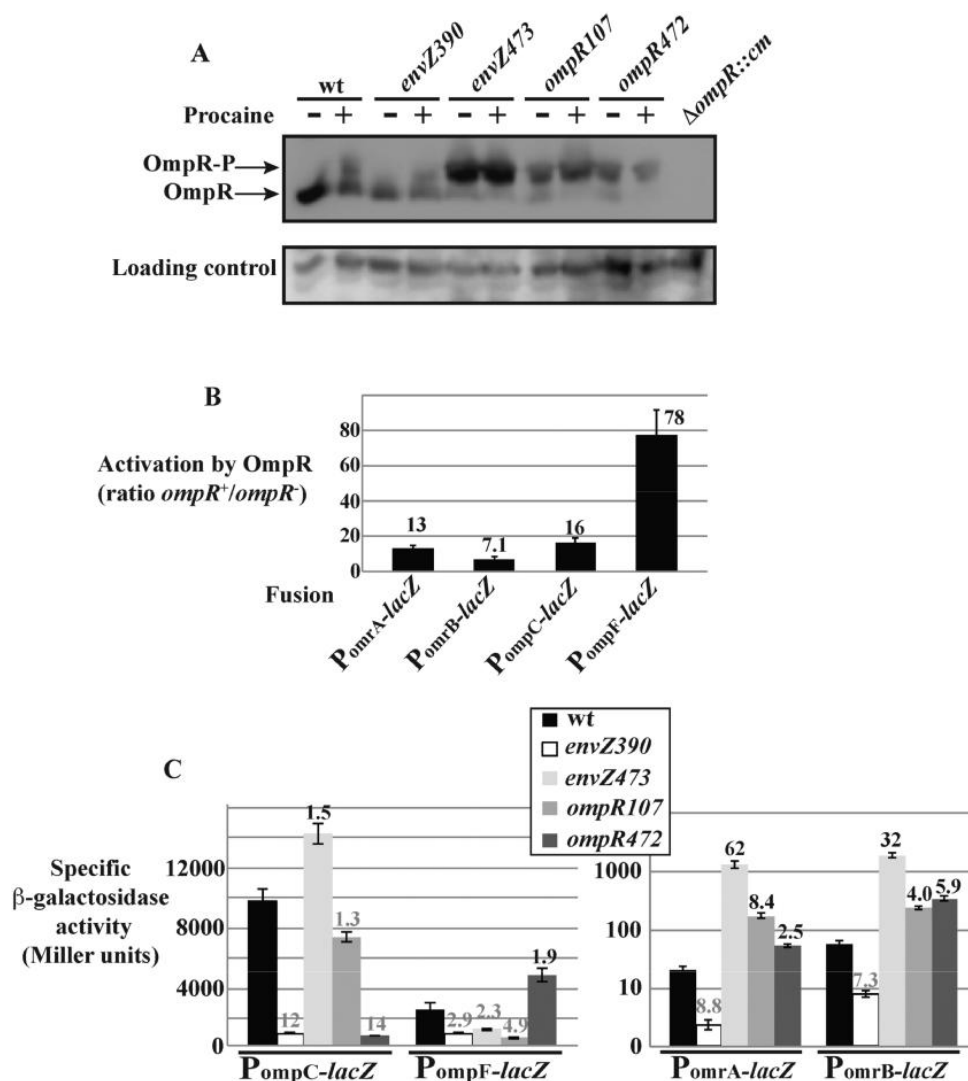


Figure 2. Transcription of *OmrA* and *OmrB* sRNAs is strongly dependent on EnvZ-OmpR TCS. (A) The effect of various mutants in *envZ* or *ompR* on OmpR phosphorylation *in vivo* was analyzed by western blot following protein separation in presence of Phos-Tag. Samples were taken from strains DJ480 (wt), MG1670 (*envZ390*), MG1671 (*envZ473*), MG1672 (*ompR107*), MG1673 (*ompR472*) and MG1035 (*ompR::cm*) grown in LB to exponential phase, before or after addition of 10 mM procaine for 10 min. A non-specific band obtained from cross-hybridization with an anti-EF-Tu antibody is shown as a loading control. (B) The activation of promoter fusions to *omrA*, *omrB*, *ompC* or *ompF* was calculated as the ratio between the activity of those fusions in $ompR^+$ and $ompR^-$ strains. The average and the standard deviations from at least two independent experiments are shown. Strains used here are, for $ompR^+$ and $ompR^-$ respectively, MG1004 and MG1811 (P_{omrA} fusion), MG1005 and MG1812 (P_{omrB} fusion), MG1863 and MG1891 (P_{ompC} fusion) and MG1690 and MG1810 (P_{ompF} fusion). In this experiment, the β -galactosidase activities in the $ompR^+$ strain were, in Miller units, 19.1 (P_{omrA} fusion), 49.4 (P_{omrB}), 6010 (P_{ompC}) and 3270 (P_{ompF}). (C) The β -galactosidase activity of the same promoter fusions in wt strains or in the same set of *envZ-ompR* mutants used in (A) was measured in LB medium in exponential phase. wt, *envZ390*, *envZ473*, *ompR107* and *ompR472* strains are respectively MG1892, MG1893, MG1894, MG1895 and MG1896 (P_{ompC} fusion), MG1690, MG1692, MG1694, MG1696, MG1698 (P_{ompF} fusion), MG1004, MG1299, MG1300, MG1301, MG1302 (P_{omrA} fusion) and MG1005, MG1303, MG1304, MG1305 and MG1306 (P_{omrB} fusion). Numbers above the bars indicate repression- or activation-fold (in gray or black respectively) in expression relative to the wt strain; note that right-hand panel uses log scale.

strong as the *ompR* deletion strain, that decreased expression of the *ompF* fusion by 78-fold (Figure 2B and C). This is possibly due to the presence of residual OmpR-P in this mutant that may be sufficient to activate *ompF*. In line with this, OmpR-P is still detectable by Phos-Tag in the *envZ390* background in presence of procaine (Figure 2A). The three other mutants, each of which has increased levels of OmpR-P, activated transcription from both the *omrA* and *omrB* promoters, with *envZ473* being by far the most efficient, since it increased activity of the *omrA* and *omrB* promoter

fusions by 62- and 32-fold respectively (Figure 2C, note log-arithmetic scale). These effects differed from those observed with the fusions to *ompC* and *ompF* promoters. For instance, activation of *ompC* or *ompF* fusion in the *envZ473*, *ompR107* or *ompR472* mutants was never more than 2-fold (Figure 2C), and as expected from previous work, *envZ473* and *ompR107* repressed the *ompF* fusion, while *ompR472* repressed the *ompC* fusion (41,43). The basal level of transcription of *OmrA/B* is strikingly lower than *ompC/F*. To rule out that those *envZ/ompR* mutants affect *OmrA/B* via

non-specific regulation of poorly transcribed genes, the effect of these mutants was analyzed on control promoter fusions. These controls were *lacZ* fusions to the short version of OmrA promoter or to the SgrS promoter, i.e. the same control promoters as used for the *in vitro* transcription re-actions. None of these fusions was significantly affected by the different *envZ/ompR* mutants, even though their expression in wt cells is lower than the expression of the Omr promoter fusions (Figure 2 and Supplementary Figure S4A). This confirms that the effect of EnvZ-OmpR on OmrA/B transcription is specific and not due to a low transcription level.

These effects on the *omr* fusions were also different from what was reported with the *dtpA* gene, whose transcription was unaffected in *envZ473* and repressed in *ompR107* and *ompR472* strains (5). Therefore, *omrA* and *omrB* are differentially regulated by EnvZ-OmpR in comparison to the other well-studied members of the regulon and they display a stronger dependence on changes in total OmpR and/or OmpR-P levels than *ompC* and *ompF*.

ompF and *ompC* have long non-coding leaders and are subject to post-transcriptional regulation by the MicF and MicC sRNAs, themselves OmpR-dependent. While the *ompF* and *ompC* promoter fusions used above did not include the leaders and sites of sRNA interaction, the effects of the different *envZ* and *ompR* alleles were very similar to that shown previously with 'long' transcriptional fusions isolated by Silhavy and coworkers and carrying the leaders together with the beginning of the *ompC/F* ORF (41,47) (fusions *ompC-lacZ* and *ompF-lacZ*, Supplementary Figure

S4, panels B and C, compare to P_{ompC}-*lacZ* and P_{ompF}-*lacZ* fusions used in Figure 2). This suggests that the effect of the mutants used here on *ompC/F* is primarily at the level of transcription initiation. As a further confirmation, fusions were created in which P_{tet} replaced the *ompF* or *ompC* promoter, but the leader of these genes and the translation start site was retained. These fusions were much less affected by loss of OmpR (Supplementary Figure S4B, P_{tet}-*ompC/F-lacZ*). Post-transcriptional regulation of *ompF* expression is further examined later in this study.

OmrA/B-mediated repression of *ompR* does not change OmpR-P levels

OmrA/B sRNAs regulate the expression of multiple targets through their 5 conserved region; one target is the bi-cistronic operon encoding their own transcriptional regulator, the EnvZ-OmpR TCS (Figure 3A). We showed previously that, when overproduced, OmrA/B directly pair to the *ompR* translation initiation region and repress *ompR* expression as well as the levels of *ompR-envZ* mRNA (26). This control of *ompR-envZ* by OmrA/B also occurs upon induction of expression of OmrA/B from their natural locus. This was assayed by comparing expression of a PBAD-*ompR-lacZ* translational fusion in wt and *omrAB* cells when the EnvZ-OmpR TCS was activated using 20 mM procaine. Expression of the wt fusion was repressed 4-fold upon addition of procaine. This repression is specific to *ompR* as it was not observed on another PBAD-driven translational fusion (PBAD-*phoP-lacZ*). Furthermore, it is clearly due to the induction of the OmrA/B sRNAs, since procaine did not affect the wt fusion in a *omrAB* strain, and

it did not affect a fusion carrying the *ompRmut2* change that abolished the control by OmrA/B (Figure 3B).

Interestingly, a mathematical model of OmpR-P levels performed by Batchelor and Goulian predicted that OmpR-P was robust (i.e. insensitive) to changes in *ompR* or *envZ* expression, as long as total OmpR levels remained much greater than EnvZ. This prediction was validated when transcription of *ompC* and *ompF* was followed over a wide range of OmpR or EnvZ levels (48,49 and personal communication). Therefore, one would predict that modulating expression of *ompR-envZ* by OmrA/B should not affect the level of OmpR-P and regulation of *ompC* and *ompF*. To test whether levels of OmpR phosphorylation changed with changing OmpR levels, proteins were extracted from cells overproducing either wt Omr or a derivative mutated in the 5 end (*mut2*) that no longer repressed *ompR* (Figure 3A and 26). These extracts were then subjected to electrophoresis on a Phos-Tag containing gel followed by a western blot analysis to monitor both unphosphorylated OmpR and OmpR-P levels. The results were in complete agreement with the previously reported robustness: over-production of wt OmrA/B, but not of the mutant versions, decreased the total amount of OmpR as expected, but the level of OmpR-P remained unchanged (Figure 3C; an independent repeat is shown in Supplementary Figure S5). This robustness, however, leaves unexplained what the physiological role of regulating *ompR-envZ* expression by the OmrA/B sRNAs might be. Is this truly a negative feedback circuit, if the levels of OmpR-P, the presumably active form of the protein, do not change?

Not all OmpR targets are as robust as *ompC/F* to changes in OmpR levels

A first hint toward resolving this question came from the fact that in earlier work we were able to isolate a plasmid overexpressing OmpR from its native promoter as a multi-copy activator of OmrA/B (hereafter referred to as pOmpR (14)). This indicates that at least some OmpR targets are sensitive to changes in *ompR* expression. Indeed, expression of *omr* fusions was strongly activated by this plasmid (roughly 12- and 7-fold for OmrA and OmrB, respectively), while expression of the *ompC/F* promoter fusions was only modestly changed (1.3-fold activation or 1.4-fold repression respectively), in agreement with the robustness of these two latter targets (Figure 4A). This difference in behavior suggested a qualitative difference in the response of the *omrA/B* promoters, compared to the *ompC/F* promoters.

However, the pOmpR plasmid carries a truncated *ompR-envZ* operon with only two-thirds of *envZ* ORF and it is not clear whether the resulting EnvZ protein (if produced) is functional or not. Another construct was therefore engineered on the chromosome to modulate transcription of the whole operon: the promoter region of *ompR* was replaced by a P_{LacO-1} (hereafter Plac) promoter inducible with IPTG, preceded by a tetracycline resistance gene so that this allele can be easily moved between strains. As visible in Figure 4B, this construct allowed us to increase total OmpR by varying concentrations of IPTG, with the level of OmpR at 10⁻² M IPTG being similar to the level of OmpR produced

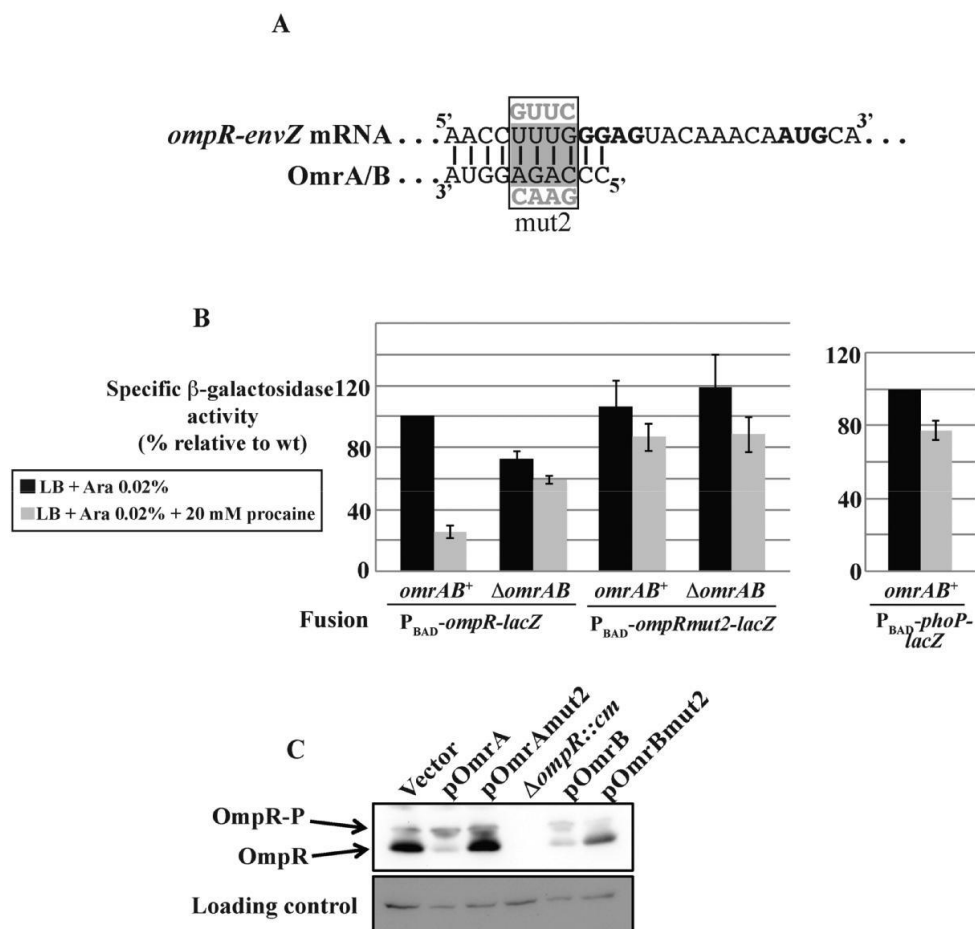


Figure 3. Control of *ompR* expression by OmrA/B changes OmpR, but not OmpR-P levels. (A) Seed-pairing interaction between OmrA/B 5' end and *ompR-envZ* mRNA (26). Shine-Dalgarno sequence and start codon of *ompR* are in bold, and mut2 changes in OmrA/B or *ompR* mRNA are shown in gray. (B) The β -galactosidase activity of an *ompR-lacZ* translational fusion, wt or with the mut2 change, expressed from a P_{BAD} inducible promoter was measured in *omrAB*⁺ or *omrAB*⁻ strains in a medium supplemented or not with 20 mM procaine. Strains used in this experiment are MG2174, MG2175, MG2176 and MG2177. As a control, the activity of a P_{BAD} -*phoP-lacZ* translational fusion (strain MG1425) was measured in the same media. Activities are expressed as % of the wt fusions in *omrAB*⁺ cells in the absence of procaine; these reference activities are, in Miller units, 2238 (*ompR* fusion) and 542 (*phoP* fusion). (C) *In vivo* phosphorylation of OmpR was analyzed upon overproduction of OmrA/B or of their mut2 derivatives in strain DJ480. An extract from a *ompR::cm* strain (MG1035) was loaded on the same gel. A non-specific band revealed with the anti-OmpR antibody was used as a loading control.

from the pOmpR plasmid mentioned above. Further-more, even in the absence of IPTG, this construct synthesizes significantly more OmpR than from the chromosomal wt promoter. When combined with the different OmpR target fusions, changing OmpR levels from this construct using a range of IPTG from 0 to 10^{-3} M, we observed that transcription of *omrA/B* fusions increased significantly (between 7- and 8-fold), while again, expression of *ompC/F* fusions varied much more modestly (1.3-fold) (Figure 4C and Supplementary Figure S6).

To rule out the possibility that these differences between OmrA/B and *ompC/F* expression were inherent to the fusions, the levels of the OmrA/B sRNAs and of the *ompC* mRNA were analyzed by northern blot in similar experiments to those above, where OmpR levels were changed using either the pOmpR plasmid or by induction of the Tet- P_{lac} -*ompR* construct with IPTG. As previously, while OmrA/B increased greatly upon OmpR overproduction, *ompC* mRNA was not significantly affected (Figure 4D).

The levels of *ompF* mRNA were also followed and results are discussed below (see Figure 7 and accompanying text).

Altogether, our data confirm that the levels of OmpR-P are robust to changes in *ompR-envZ* expression and that this translates into robust transcriptional control of some targets such as *ompC* and *ompF*. However, this is not true for all OmpR targets. In particular, *omrA/B* transcription is not robust and expression of these two sRNAs varies greatly with changing levels of the OmpR protein. **OmrA/B limit their own synthesis through feedback control**

One prediction that can be made in the light of these results is that, by regulating *ompR* translation, OmrA/B sRNAs should not modify the expression from robust promoter targets such as *ompC/F* but should modulate that of non-robust targets such as the *omrA* and *omrB* promoters. To test this, the activity of the OmrA/B- and *ompC/F*-promoter fusions was monitored upon overexpression of

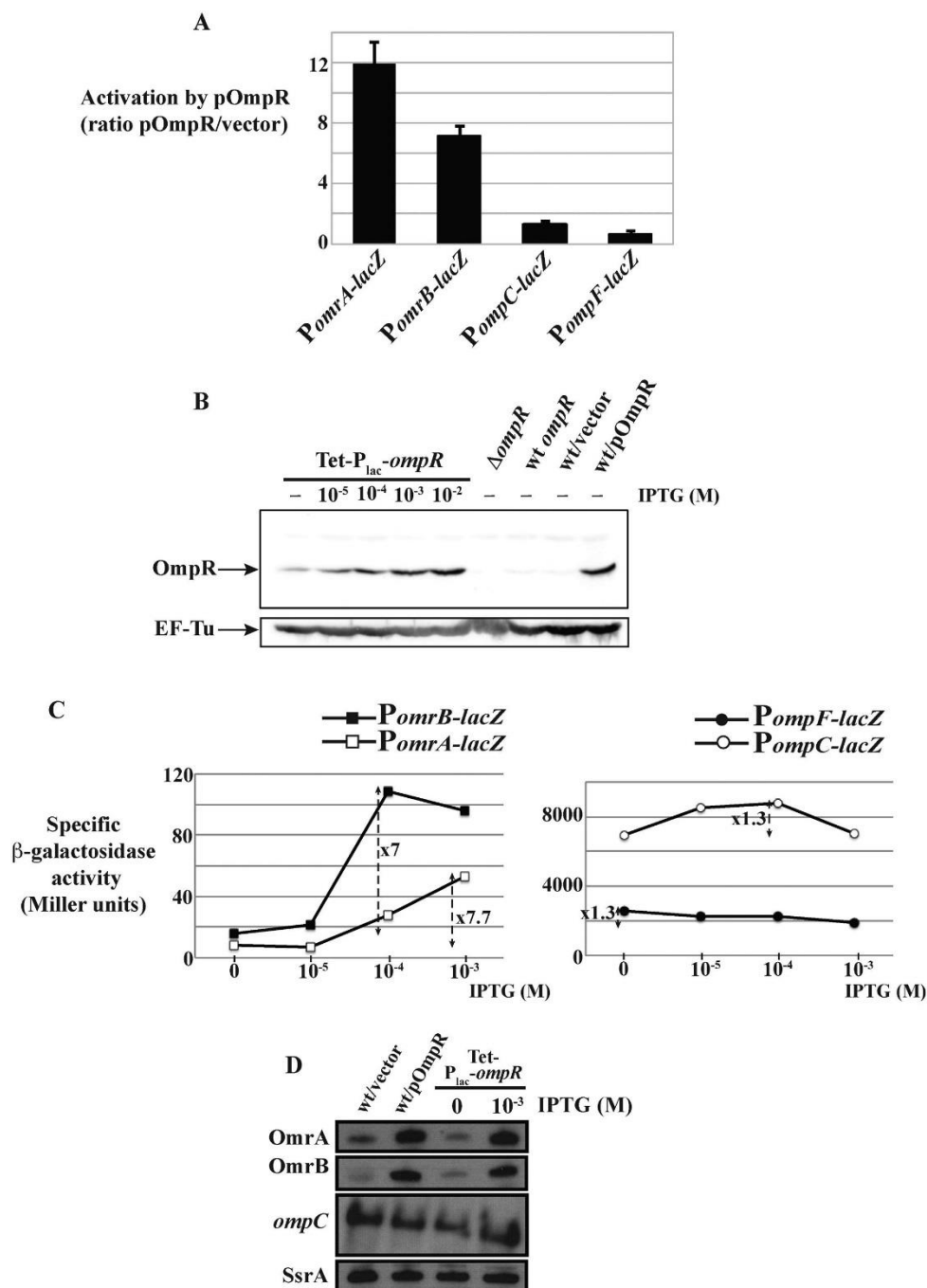


Figure 4. Different patterns of robustness in the OmpR regulon. (A) The activation of OmrA, OmrB, *ompC* or *ompF* promoter fusions (strains MG1004, MG1005, MG1892 and MG1690 respectively) was calculated as the ratio between the β -galactosidase activities of strains transformed by pOmpR and the pHDB3 empty vector respectively. Activities in presence of pHDB3 were, in Miller units, 33 (P_{omrA}), 87 (P_{omrB}), 7313 (P_{ompC}) and 3510 (P_{ompF}). (B) Western blot analysis of OmpR levels from cells carrying *ompR-envZ* operon expressed from its own promoter (wt *ompR*, strain MG1988) or from the

Tet-P_{lac}-*ompR* construct (strain MG2000). Cells were grown to exponential phase in LB supplemented or not with IPTG at the indicated concentrations. As controls, the level of OmpR was also analyzed from strain MG1004 transformed with pHDB3 or pOmpR, and from the *ompR* strain MG1811. Immunoblot detection of EF-Tu is shown as a loading control. (C) The β -galactosidase activity of the same promoter fusions as in panel (A) was followed when *ompR-envZ* expression was modulated from the Tet-P_{lac}-*ompR* allele by increasing IPTG concentrations. Results of a representative experiment are shown here and two additional independent repeats are shown in Supplementary Figure S6. Strains used in this experiment are MG2000, MG2002, MG2004 and MG2006. (D) Levels of OmrA, OmrB sRNAs and of *ompC* mRNA were followed by northern blot when changing OmpR levels. SsrA is used as a loading control. Strains are MG1004 (wt) and MG2000 (Tet-P_{lac}-*ompR*).

either OmrA/B or their mut2 derivative that no longer represses *ompR-envZ* (Figure 5A). The overproduction of the different Omr variants was achieved using plasmids carrying the corresponding *omr* gene under control of an IPTG-inducible promoter in strains deleted for the *omrAB* chromosomal genes. As a control, the same experiment was carried out in an isogenic strain where the chromosomal copy of *ompR-envZ* carries the compensatory change to mut2, i.e. its expression is no longer controlled by wt OmrA/B but is controlled by the mutant OmrA/Bmut2. This *om-pRmut2* mutation consists of a change in nts -13 and -17 (relative to the *ompR* ATG start codon). Therefore, it is not expected to change the amino acid sequence of OmpR, and it does not significantly affect *ompR* expression, at least when expressed from a P_{BAD} promoter (26).

The presence of plasmid pOmrA decreased expression of both P_{omrA}- and P_{omrB}-*lacZ* fusions, by 3- and 2.4-fold, respectively, in the wt strain, but had no effect in the *ompRmut2* background (Figure 5B, top panel). In contrast, overexpression of OmrAmut2 did not significantly change the activity of either fusion in the wt context, but decreased it in presence of the *ompRmut2* variant, clearly showing that the observed repression is due to the down-regulation of *ompR-envZ* by OmrA. Similar results were obtained when OmrB (or OmrBmut2) was overproduced, but with some-what lower repression factors. Interestingly, repression of either fusion was always higher in the mut2/mut2 situation than in the wt/wt. There are at least two hypotheses that could explain this observation; these models are not mutually exclusive. First, it is possible that the sRNA-mRNA duplex formation and/or stability was increased by the compensatory changes. Alternatively, since all targets of OmrA/B known so far are regulated through the 5 end region of these sRNAs (18,21,26), it is plausible that, in the context of the compensatory changes, *ompRmut2* is the only target (or at least one among very few) of OmrA/Bmut2. The end result is that, in the absence of other targets, the level of OmrA/Bmut2 available to base-pair with *om-pRmut2* is increased compared to the wt/wt situation.

As expected based on the results of Figure 4, this ability of OmrA/B to repress their own synthesis through *ompR* regulation was not paralleled when we looked at the expression of *ompC* and *ompF*. Instead, the activity of both P_{ompC}- and P_{ompF}-*lacZ* fusions was unaffected, or only very weakly affected, by overexpression of OmrA/B wt or mut2 (Figure 5B, lower panels).

The experiments shown in Figure 5B were performed in LB medium, i.e. under experimental conditions where the activity of EnvZ/OmpR TCS is not expected to be strongly induced and where overproduction of OmrA/B is driven by an inducible promoter, whose activity is completely independent of EnvZ-OmpR. To get a better sense of the effect of OmrA/B when expressed from the chromosome, the activity of the same set of promoter fusions was compared in *omrAB* wt or deleted strains (Figure 5C). This experiment was performed in an *ompR107* background where OmrA/B are induced (Figure 2). In line with results obtained with OmrA/B overproduction, deletion of *omrAB* reproducibly increased the activity of the fusions to the promoters of OmrA/B by ~40 to 50%, but did not significantly affect *ompC* or *ompF* expression (Figure 5C). Therefore, the difference in the responses of the *omrA/B*- and *ompC/F*- promoters to changing OmpR levels allows autocontrol of OmrA/B sRNAs through the negative feedback circuit that exists between EnvZ-OmpR TCS

and OmrA/B; at the same time, *ompC/F* transcription remains resistant to the effect of OmrA/B.

A role for non-phosphorylated OmpR in activating *omrA* and *omrB* transcription?

What could explain the sensitivity of non-robust promoters such as P_{omrA} or P_{omrB} to changing OmpR levels given that the levels of OmpR-P are themselves robust? An obvious hypothesis is that, in contrast to *ompC/F*, *omrA/B* promoters might be activated not only by OmpR-P, but also by the non-phosphorylated form of OmpR. To address this question, we first worked with mutants of the phosphorylation site of OmpR (D55A or D55N) but found that they were unable to activate *omr* transcription, even when over-produced (data not shown). Because it is difficult to rule out the possibility that these variant OmpR proteins might be inactive, we then decided instead to follow the activity of the different promoter fusions when using strains with a wt OmpR protein but very low level of OmpR-P. This was achieved by working in strains lacking the OmpR kinase EnvZ. However, it is known that in the absence of EnvZ, some OmpR-P is still present and is generated in large part through phosphorylation of OmpR by acetylphosphate (see e.g. 50). The *envZ* deletion was therefore used in conjunction with a deletion of the *ackA-pta* genes, required for synthesis of acetylphosphate. The resulting strains display a low expression of the four OmpR targets, almost as low as in the *ompR* null mutant (Figure 6A; note that each wt level is set to 100%, to allow easier comparison between the promoters, although total levels of expression are very different, see wt bars in Figure 6B). OmpR-P is thus required for expression of all these genes including *omrA/B*, which is consistent with the results obtained in the *envZ390* mutant (Figure 2) and in the *in vitro* transcription reactions (Figure 1). In this *envZ ackA-pta* background, overproduction of OmpR from the Tet-P_{lac-ompR} allele slightly increased expression of both the *ompF* and *ompC* promoter fusions, compared to the *envZ ackA-pta* strain with *ompR* expressed from its own promoter, suggesting that OmpR-P levels may be slightly higher (Figure 6A, compare dark gray and black bars). The stronger activation of *ompF* is consistent with the fact that *ompF* is activated at lower OmpR-P levels compared to *ompC* (2). The presence of OmpR-P in these strains is most likely due to cross-phosphorylation by a histidine kinase other than EnvZ or to the presence of acetylphosphate in the LB growth medium. This cross-phosphorylation is relatively inefficient, as judged by the low level of P_{ompC}-*lacZ* transcription, compared to a wt strain (Figure 6A, compare black and white bars). Strikingly, overexpression of OmpR in these strains leads to strong activation of both *omrA* and *omrB* transcription relative to the wt strain (by 19- and 10-fold respectively), despite the lower level of OmpR-P (Figure 6A). This shows that, even though expression of these two genes strictly requires OmpR-P, it can nonetheless be modulated by changes in the levels of unphosphorylated OmpR. While a detailed

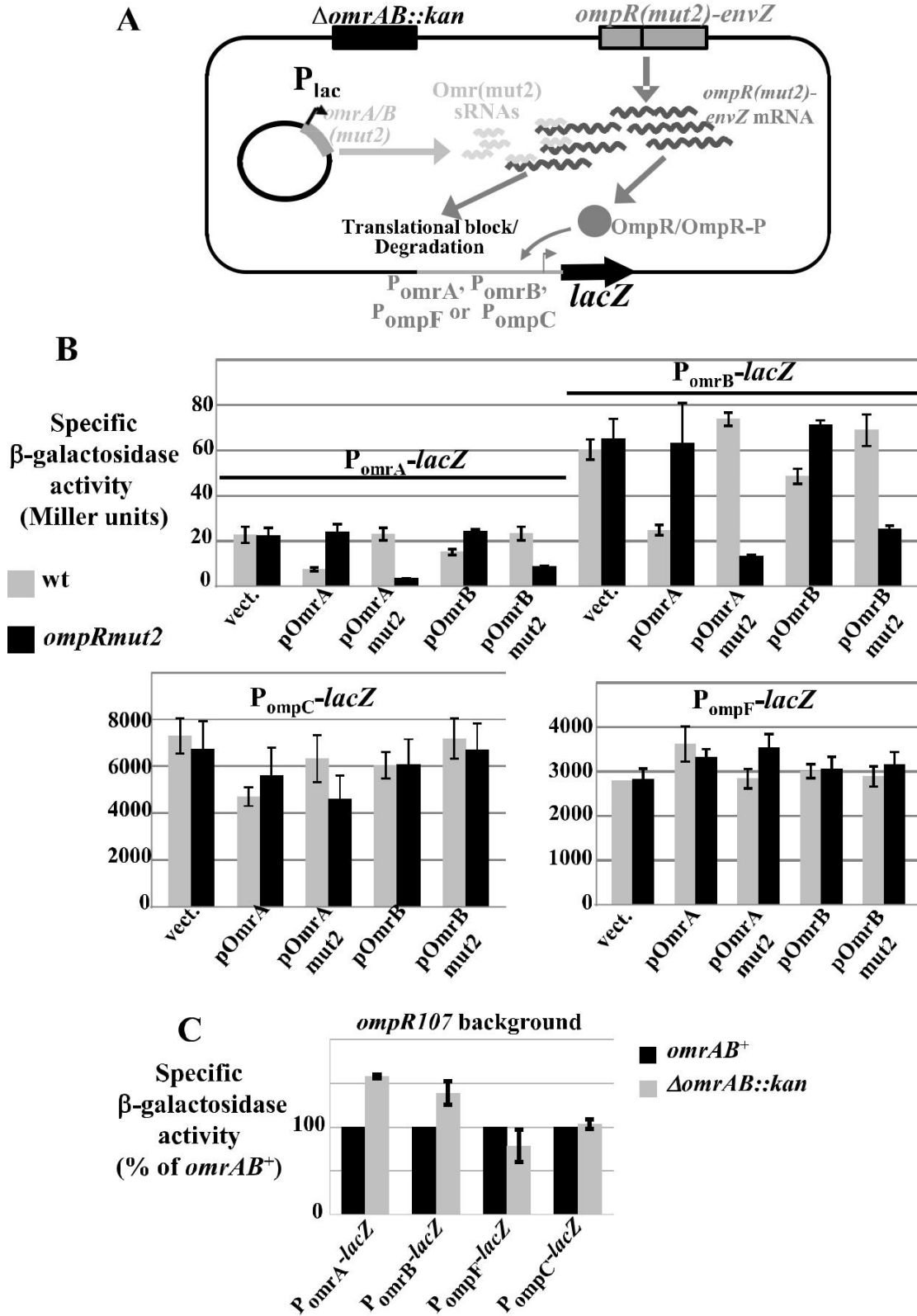


Figure 5. Feedback control in OmrA/B expression. (A) Scheme depicting the experiment performed in (B): the β -galactosidase activity of promoter fusions to OmrA, OmrB, *ompC* and *ompF* was measured upon overproduction of OmrA, OmrB or of their *mut2* derivatives, in strains with a wt copy of *ompR* (gray bars) or with a mutant version carrying the *ompRmut2* compensatory change to OmrA/Bmut2 (black bars). Strains used here are MG1014, MG1666, MG1017, MG1667, MG1901, MG1902, MG1686, MG1688 (panel B, see strain table for details). (C) Activity of the same set of promoter fusions was compared in *omrAB*⁺ and *omrAB*⁻ strains in the *ompR107* background. Values are represented as % of wt strain for each fusion; the activities of the wt strains were, in Miller units, 135 (P_{omrA} fusion), 173 (P_{omrB}), 547 (P_{ompF}) and 5042 (P_{ompC}). Strains used are MG1301, MG1908, MG1305, MG1909, MG1696, MG1910, MG1895 and MG1911.

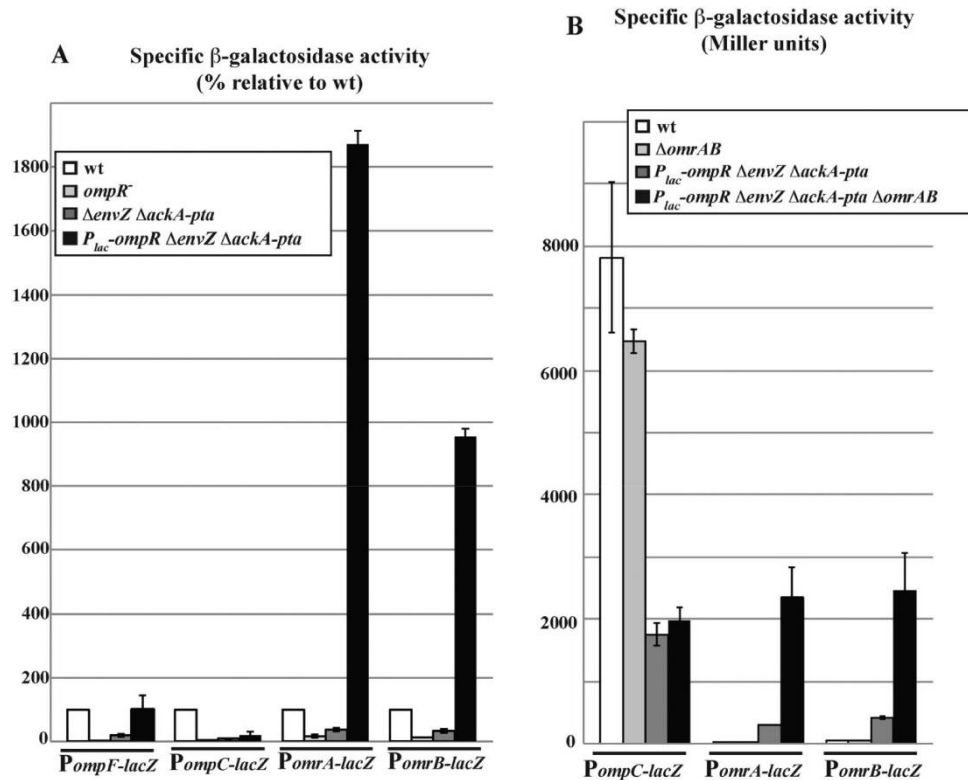


Figure 6. Expression of *ompA* and *ompB* can be activated by non-phosphorylated OmpR. (A) The β -galactosidase activity of promoter fusions to *ompA*, *ompB*, *ompC* or *ompF* was measured in LB in exponentially growing cells with varying OmpR levels and/or phosphorylation. Activities are shown relative to that of wt cells. Strains used in this experiment are MG1690, MG1810, MG2043, MG2044 (P_{ompF}), MG1892, MG2018, MG2045, MG2046 (P_{ompC}), MG1004, MG1811, MG2039, MG2040 (P_{ompA}), MG1005, MG1812, MG2041 and MG2042 (P_{ompB}). The β -galactosidase activities of the different fusions in the wt strains were, in Miller units, 22 (P_{ompA} fusion), 63 (P_{ompB}), 11292 (P_{ompC}) and 3618 (P_{ompF}). (B) Activity of the same *ompA*, *ompB* and *ompC* transcriptional fusions was measured in wt cells or upon OmpR overproduction in *envZ* and *ackA-pta* background, both in *ompRAB*⁺ or *ompRAB*⁻ strains. Strains are MG1892, MG2203, MG2046 and MG2206 (P_{ompC} fusion), MG1004, MG2201, MG2040 and MG2204 (P_{ompA}), and MG1005, MG2202, MG2042 and MG2205 (P_{ompB}).

understanding of the mechanism involved is still lacking, this would set up the basis for the different robustness to OmpR levels that were observed for OmpA/B and *ompC/F*.

Interestingly, activation of the OmpA and OmpB fusions upon OmpR overproduction in this low OmpR-P background was even more pronounced in *ompRAB* cells (Figure 6B). In this context, *ompA* or *ompB* transcription was increased 123- and 49-fold respectively, i.e. roughly 8- and 6-times more than in the *ompR*⁺ cells. In contrast, deletion of *ompRAB* had no effect on the decrease in *ompC* transcription due to the lower OmpR-P levels. This is a further demonstration that OmpA/B limit *ompR* overexpression and that this feedback results in lowering their own transcription.

ompF mRNA levels are not robust to OmpR

When looking at *ompC* mRNA by northern blot with changing OmpR levels, we found that, similarly to what was observed with the transcriptional fusion, *ompC* mRNA was barely affected (Figure 4D). In striking contrast however, while transcription initiation of *ompF* was also robust to OmpR (Figures 4 and 5), *ompF* mRNA levels were clearly reduced when *ompR* expression increased (Figure 7A). This difference should be due to an OmpR-dependent process controlling *ompF* expression downstream of transcription

initiation. In agreement with this possibility, a translational *ompF-lacZ* fusion expressed from a P_{tet} constitutive promoter still displayed some dependency on OmpR: expression of this fusion was lower in an *ompR*⁺ than an *ompR*⁻ strain (ratio of 0.64, see Supplementary Figure S4B), and was 2-fold repressed when OmpR synthesis was induced with IPTG from the Tet- $P_{lac-ompR}$ allele (Figure 7B). Most of this post-transcriptional regulation by OmpR appears to be due to MicF under the conditions used here, since deletion of *micF* abolishes the effect of overproducing OmpR (Figure 7B). This suggests that MicF is itself a non-robust OmpR target even though, given the complexity of MicF regulation (51), further experiments are required to determine whether this is due to the direct non-robust activation of *micF* transcription by OmpR. These experiments indicate that even though transcription of *ompF* is clearly robust, synthesis of the OmpF porin might nonetheless follow a non-robust behavior under some conditions because of post-transcriptional control. Furthermore, the involvement of MicF in this process does not rule out that yet other actors might play a role as well in the non-robust response of *ompF* mRNA levels.

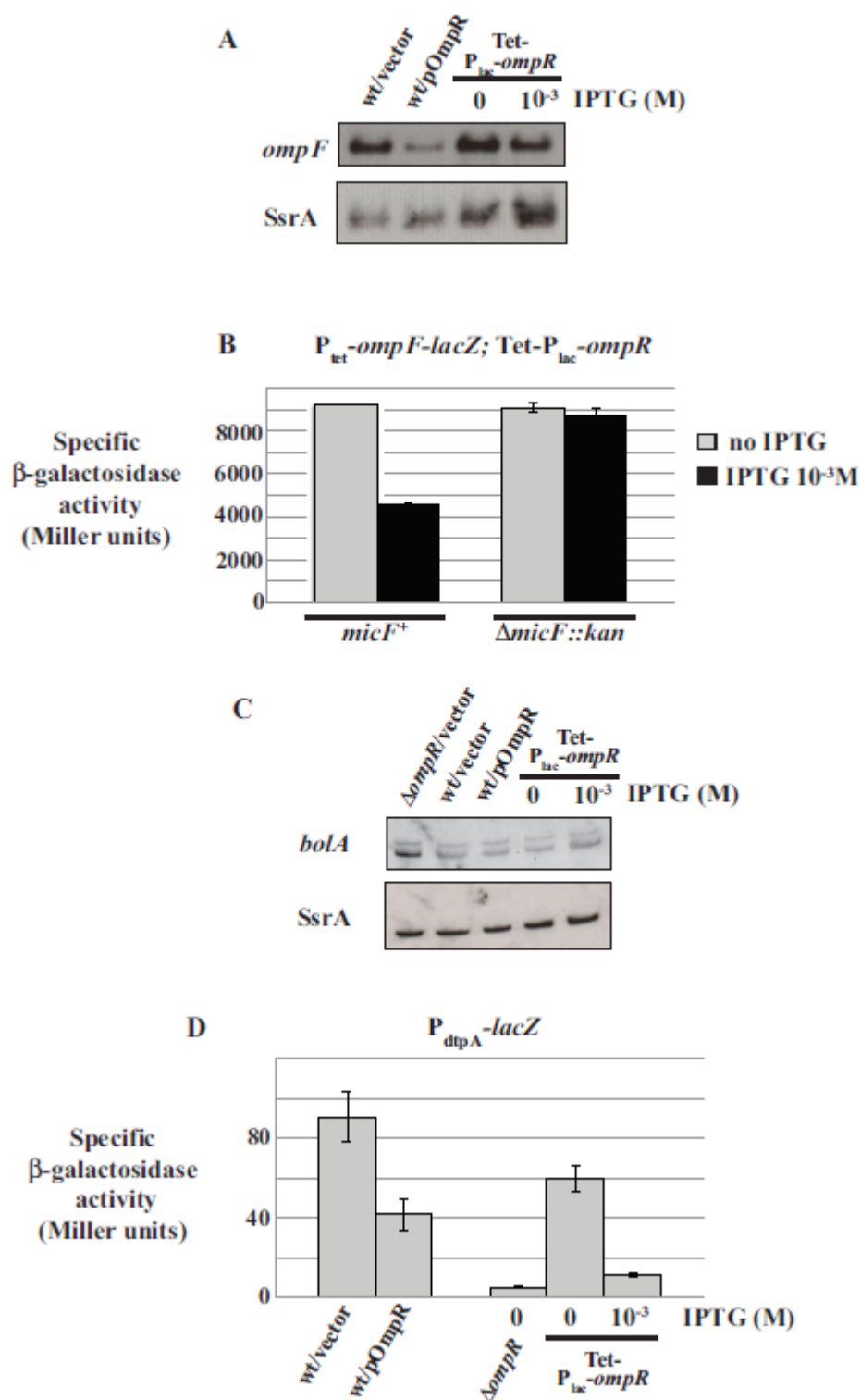


Figure 7. Robustness and non-robustness of other OmpR targets. **(A)** The same RNA samples used in Figure 4D were analyzed by northern blot probed for *ompF* and for SsrA as a loading control. **(B)** MicF sRNA is involved in the post-transcriptional control of *ompF* upon OmpR overproduction. The β -galactosidase activity of $P_{tet-ompF-lacZ}$ translational fusion was measured with and without IPTG induction of *ompR* overexpression in *micF*⁺ and *micF*⁻ cells (strains MG2127 and MG2167 respectively). **(C)** *bolA* mRNA levels were analyzed by northern blot in strains MG1004 (wt) transformed with pHDB3 or pOmpR, MG1811 (*ompR*) transformed with pHDB3 and in MG2000 ($Tet-P_{lac-ompR}$) with or without IPTG. SsrA levels were probed as well and used as a loading control. **(D)** Expression of a *dtpA-lacZ* promoter fusion was followed in a set of strains with various OmpR levels (MG2138 (wt), MG2169 (*ompR*) and MG2170 ($Tet-P_{lac-ompR}$)).

Robustness and other OmpR targets

EnvZ-OmpR directly controls transcription of other genes in addition to MicF, OmrA/B and *ompC/F*. One can therefore wonder whether these other OmpR targets are sensitive to changes in OmpR levels. We have tested this for two of them: *bolA* and *ntpA*. Transcription of *bolA* can originate from two promoters and OmpR was reported to repress transcription from the proximal promoter (8). Accordingly, two distinct *bolA* mRNAs were visible by northern blot and only the levels of the shorter transcript increased in an *ompR* deleted strain. When OmpR was overproduced, either from pOmpR or from Tet-Plac-*ompR* using IPTG, levels of both *bolA* transcripts remained unchanged, indicating that *bolA* transcription, like that of *ompC/F*, responds only to OmpR-P and is thus robust to changes in OmpR levels (Figure 7C). Regarding *ntpA*, transcription was followed using a transcriptional fusion carrying nts -201 to +15 of *ntpA* (relative to TSS) upstream of *lacZ* (from nt -17). Consistent with the known regulation of *ntpA* by OmpR (5), expression of

this P_{ntpA}-*lacZ* fusion was strongly reduced in the absence of OmpR. Interestingly, activity of the fusion was also decreased upon OmpR overproduction, by 2-fold in presence of pOmpR and by 5-fold when *ompR* expression was induced from Tet-Plac-*ompR* (Figure 7D). It seems therefore that *ntpA* transcription is activated by OmpR (or OmpR-P) at low levels but repressed at higher levels, reminiscent of the regulation of *ompF* by OmpR-P. More importantly for the scope of this study, this result shows that *ntpA* is another non-robust OmpR target whose expression is modulated in response to changes in OmpR levels. Note at last that both OmpR and OmpR-P were shown to bind the *fimB* promoter of an uropathogenic *E. coli* strain, indicating the existence of yet other targets sensitive to changes in OmpR levels (52).

DISCUSSION

The results reported here confirm and extend our understanding of the EnvZ/OmpR two-component system, how it is itself regulated and how it regulates targets. We summarize our findings in Figure 8.

A dual OmpR regulon?

OmpR-P levels were previously predicted to be robust when *ompR* or *envZ* expression varied, and accordingly *ompC* and *ompF* transcription remained constant over a wide range of EnvZ or OmpR levels, from 10-fold less to 10-fold more relative to wt levels (48). However, this robustness did not apply upon stronger variations of either EnvZ or OmpR levels (48) as well as upon overproduction of the cytoplasmic domain of EnvZ (53,54).

The use of Phos-Tag allowed us to directly confirm robustness of OmpR-P, at least upon overproduction of OmrA or OmrB sRNA. Our data furthermore show that this translates into robust transcriptional regulation of some targets such as *ompC* and *ompF*. However, this is not the case for all OmpR targets as the promoters of *omrA/B* and *ntpA* were found, for instance, to be regulated by changes in OmpR levels as well.

One prediction that can be made from these findings is that the OmpR regulon is actually composed of at least

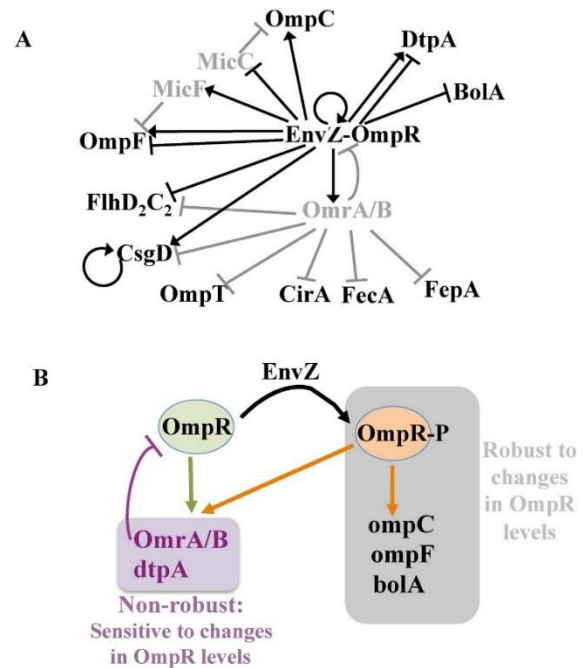


Figure 8. Robustness and differential regulation within the OmpR regulon. (A) Regulatory circuit centered on the EnvZ-OmpR-OmrA/B feed-back motif. Transcriptional and post-transcriptional regulations are in black and gray respectively, with arrows indicating activations and perpendicular bars indicating repressions. sRNAs are in gray. For simplicity, only gene products relevant to this study are indicated. Note that direct control has not been demonstrated in some aspects of this circuit (for instance, MicC control by OmpR, or OmpR autoregulation). (B) Differential regulation of robust or non-robust OmpR targets by the non-phosphorylated form of OmpR. See text for more details.

two subsets of genes. The first group, corresponding to robust targets (gray square in Figure 8B), should be sensitive only to signals affecting the level of OmpR-P, e.g. osmolarity. The second group is composed of the non-robust targets whose expression should depend on both the phosphorylation status of OmpR and its expression levels (purple square in Figure 8B). This is for instance what is observed for OmrA/B, which autoregulate their own synthesis by repressing *ompR* translation (Figure 5). Interestingly, control of *ompR* expression is not limited to the action of the Omr sRNAs, as it was reported that *ompR-envZ* transcription is also regulated by cAMP-CRP (via a complex pattern depending on the transcription start site), inhibited by IHF (55,56) and is possibly positively autoregulated by OmpR (10 and Supplementary Figure S2). The non-robust control of some OmpR targets is therefore expected to allow integration of multiple signals in controlling only a subset of genes of the OmpR regulon.

Robustness has been proposed to minimize output variability between cells that are subject to the same input but can present stochastic variations in the concentrations of some cellular components (57). Given the importance of the *ompC/F* structural genes that encode the major porins in *E. coli* and whose ratio will impact the entry of various beneficial or

harmful compounds, one can understand why keeping their transcription independent of stochastic variations might matter. In contrast, regarding OmrA/B, it seems logical that these OmpR targets involved in feedback control of *ompR* expression are non-robust, which is expected to allow them to keep *ompR* expression levels within a given range and underlies their observed autoregulation.

Furthermore, our data indicate that *bolA* transcription is robust to OmpR while *ntpA* is not. In the future, it will be interesting to determine the nature of the other OmpR targets in order to understand a possible link between signals affecting *ompR* expression and the function of non-robust targets.

Finally, it is worth noting that, even if transcription of some targets is robust per se, their expression can nonetheless be modulated in response to changes in OmpR levels by a non-robust post-transcriptional step. This is what we have found for *ompF* for instance (Figure 7). Similarly, one can expect that OmrA/B control of *flhDC* and *csgD* will provide these targets a non-robust response to OmpR.

Mechanism of regulation of Omr sRNAs transcription by EnvZ-OmpR

An important question that is raised by the results of this study is the nature of the promoter determinants that underly the robust or non-robust regulation by OmpR, together with the molecular mechanism involved. Our data strongly suggest that unphosphorylated OmpR is involved in transcription regulation at the Omr promoters, but also possibly at the promoters of other non-robust targets. In some cases, the unphosphorylated form of a response regulator has been shown to specifically activate gene expression at some target-promoters (e.g. 58,59). However, in the case studied here, our results suggest that at least some OmpR-P is required for transcription activation at the Omr promoters. Indeed, *envZ* or *ackA-pta* mutants that lower OmpR-P reduce OmrA/B transcription as much as an *ompR* deletion allele (Figures 2C and 6). Furthermore, mutants of the phosphorylation site of OmpR (D55A or D55N) failed to activate Omr transcription, even when overproduced (our unpublished results), which also suggests that OmpR-P is essential for transcription activation, even at the Omr promoters.

A first hypothesis that could explain both the requirement for OmpR-P and the sensitivity to unphosphorylated OmpR is that Omr promoters could simultaneously bind both forms of OmpR. As OmpR phosphorylation has been found to promote dimerization and subsequently DNA-binding (34), OmpR-P presumably binds as a dimer and this binding is likely to be essential for control. Unphosphorylated OmpR may bind as well and promote OmpR-P mediated transcription regulation, for instance by activating OmpR-P binding. It is not clear in what form unphosphorylated OmpR would bind, perhaps as a monomer or in complex with another protein, and what promoter elements would allow this binding. An alternative possibility is that OmpR and OmpR-P could form heterodimers able to control gene expression specifically at the Omr promoters.

It is also possible that the observed regulation of Omr synthesis by unphosphorylated OmpR is indirect. Instead, OmpR could control the level of another regulator of Omr

transcription; both the nature of this regulator and how its levels would be controlled by OmpR remains to be identified to support this hypothesis. Additional studies are required to discriminate between these possibilities, and identify the promoter elements that make Omr sensitive to both unphosphorylated OmpR and OmpR-P.

Regulating regulators: a recurrent theme in sRNA biology

Synthesis of base-pairing sRNAs is often controlled by transcriptional regulators and, in turn, base-pairing sRNAs commonly target genes encoding transcriptional regulators (60,61 for reviews).

In some cases, sRNAs control expression of their own regulator, thereby creating a feedback circuit, such as the one described here. This has been shown for instance for MicF that negatively controls *lrp* translation in a double negative feedback circuit (62,63). Another striking example is that of the *Vibrio* Qrr sRNAs involved in quorum-sensing and that lie at the core of multiple negative feedback circuits (64), which participate in modulating the cell density at which the switch occurs as well as the kinetics of the switch (65–67).

The results presented in this study show that the feedback circuit mediated by OmrA/B on *ompR-envZ* expression can limit the expression of OmrA/B in steady-state experiments (Figures 5 and 6B). However, this circuit could provide other properties to OmrA/B expression. It could, for instance, restrict Omr induction to a shorter time frame following OmpR activation, allow a surge in Omr levels or provide a faster recovery once the inducing signal(s) for OmpR has disappeared. Another possibility is that this feedback could be involved in ensuring that Omr synthesis occurs only in response to sustained stimulation of OmpR and would filter transient signals, or in limiting cell-to-cell variations among the bacterial population. Interestingly, a recent RNAseq study showed that OmrB, and to a lesser extent OmrA, was strongly induced when *Salmonella* enters eukaryotic host cells (68). Our findings may thus also be relevant for host-pathogen interactions.

Regulatory network motifs and their properties were first defined by studies of protein transcriptional regulators. However, while many transcriptional regulators are subject to direct autoregulation, examples of protein-based negative feedback circuits reminiscent of the one described in this work are extremely rare in purely transcriptional networks (69). Understanding its properties could thus provide important information regarding not only OmrA/B, but also the role of this specific regulatory circuit in gene expression.

In addition to EnvZ-OmpR TCS, other examples of robustness have been reported such as the phosphorylation of IDH (Isocitrate Dehydrogenase) enzyme by the IDHKP kinase/phosphatase in the glyoxylate bypass pathway (70). Given that prediction of robustness relies on the fact that EnvZ is a bifunctional kinase/phosphatase and that OmpR levels are much greater than those of EnvZ, it is likely that robustness will apply to other TCS as well. In this regard, results obtained with the PhoQ/PhoP TCS suggest that PhoP-P levels might indeed be robust to some changes in

total PhoP (71). Our results can thus have implications for these other cases and it is therefore of interest to determine whether these systems also have non-robust targets as well.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to D. Walthers and L. Kenney for providing the OmpR-His overproducing plasmid and the pRLG770 vector, and to L. Morgan for advice regarding the purification procedure. We also thank A. Stock for the anti-OmpR antibody and help with the Phos-Tag technique, T. Silhavy, A. Stock, E. Hobbes and N. Majdalani for providing strains or plasmids, and W. Ross for advice with *in vitro* transcription. Discussions with T. Silhavy, M. Goulian, L. Kenney, A. Stock, T. Msadek and F. Repoila throughout this work were greatly appreciated. Thanks to H. Kemble and C. Chiaruttini for the *ompR* transcriptional fusion and the *ara::lacI^q-Kan^R* allele respectively. We are grateful to M. Springer, J. Jagodnik, C. Chiaruttini and C. Condon for comments on the manuscript.

FUNDING

CNRS, ANR [ANR-2010-JCJC-130801 and ANR-14-CE10-0004-01 to M.G.]; ‘Initiative d’Excellence’ program from the French State [‘DYNAMO’, ANR-11-LABX-0011]; Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. Funding for open access charge: ANR (french national research agency) [ANR-14-CE10-0004-01].

Conflict of interest statement. None declared.

REFERENCES

- Goulian, M. (2010) Two-component signaling circuit structure and properties. *Curr. Opin. Microbiol.*, **13**, 184–189.
- Pratt, L. and Silhavy, T. (1995) Porin regulon of *Escherichia coli*. In: Hoch, J. and Silhavy, T. (eds). *Two-Component Signal Transduction*. American Society for Microbiology, pp. 105–127.
- Pratt, L.A., Hsing, W., Gibson, K.E. and Silhavy, T.J. (1996) From acids to *osmZ*: multiple factors influence synthesis of the OmpF and OmpC porins in *Escherichia coli*. *Mol. Microbiol.*, **20**, 911–917.
- Kenney, L.J. (2002) Structure/function relationships in OmpR and other winged-helix transcription factors. *Curr. Opin. Microbiol.*, **5**, 135–141.
- Goh, E.B., Siino, D.F. and Igo, M.M. (2004) The *Escherichia coli tppB* (*ydgR*) gene represents a new class of OmpR-regulated genes. *J. Bacteriol.*, **186**, 4019–4024.
- Shin, S. and Park, C. (1995) Modulation of flagellar expression in *Escherichia coli* by acetyl phosphate and the osmoregulator OmpR. *J. Bacteriol.*, **177**, 4696–4702.
- 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.
- Yamamoto, K., Nagura, R., Tanabe, H., Fujita, N., Ishihama, A. and Utsumi, R. (2000) Negative regulation of the *bolA*1p of *Escherichia coli* K-12 by the transcription factor OmpR for osmolarity response genes. *FEMS Microbiol. Lett.*, **186**, 257–262.
- Dressaire, C., Moreira, R.N., Barahona, S., Alves de Matos, A.P. and Arraiano, C.M. (2015) BolA is a transcriptional switch that turns off motility and turns on biofilm development. *mBio*, **6**, doi:10.1128/mBio.02352-14.
- Quinn, H.J., Cameron, A.D. and Dorman, C.J. (2014) Bacterial regulon evolution: distinct responses and roles for the identical OmpR proteins of *Salmonella Typhimurium* and *Escherichia coli* in the acid stress response. *PLoS Genet.*, **10**, e1004215.
- Stinccone, A., Daudi, N., Rahman, A.S., Antczak, P., Henderson, I., Cole, J., Johnson, M.D., Lund, P. and Falciani, F. (2011) A systems biology approach sheds new light on *Escherichia coli* acid resistance. *Nucleic Acids Res.*, **39**, 7512–7528.
- Takayanagi, K., Maeda, S. and Mizuno, T. (1991) Expression of *micF* involved in porin synthesis in *Escherichia coli*: two distinct cis-acting elements respectively regulate *micF* expression positively and negatively. *FEMS Microbiol. Lett.*, **67**, 39–44.
- Chen, S., Zhang, A., Blyn, L.B. and Storz, G. (2004) MicC, a second small-RNA regulator of Omp protein expression in *Escherichia coli*. *J. Bacteriol.*, **186**, 6689–6697.
- Guillier, M. and Gottesman, S. (2006) Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. *Mol. Microbiol.*, **59**, 231–247.
- Storz, G., Vogel, J. and Wassarman, K.M. (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Mol. Cell*, **43**, 880–891.
- De Lay, N., Schu, D.J. and Gottesman, S. (2013) Bacterial small RNA-based negative regulation: Hfq and its accomplices. *J. Biol. Chem.*, **288**, 7996–8003.
- Vogel, J. and Luisi, B.F. (2011) Hfq and its constellation of RNA. *Nat. Rev. Microbiol.*, **9**, 578–589.
- 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. *EMBO J.*, **29**, 1840–1850.
- 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*. *Mol. Microbiol.*, **84**, 36–50.
- 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.
- De Lay, N. and Gottesman, S. (2012) A complex network of small non-coding RNAs regulate motility in *Escherichia coli*. *Mol. Microbiol.*, **86**, 524–538.
- Mika, F., Busse, S., Possling, A., Berkholz, 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*. *Mol. Microbiol.*, **84**, 51–65.
- Mizuno, T., Chou, M.-Y. and Inouye, M. (1984) A unique mechanism regulating gene expression: Translational inhibition by a complementary RNA transcript (*micRNA*). *Proc. Natl. Acad. Sci. U.S.A.*, **81**, 1966–1970.
- Pfeiffer, V., Papenfort, K., Lucchini, S., Hinton, J.C. and Vogel, J. (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat. Struct. Mol. Biol.*, **16**, 840–846.
- Guillier, M. and Gottesman, S. (2006) Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. *Mol. Microbiol.*, **59**, 231–247.
- Guillier, M. and Gottesman, S. (2008) The 5' end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. *Nucleic Acids Res.*, **36**, 6781–6794.
- Beisel, C.L. and Storz, G. (2010) Base pairing small RNAs and their roles in global regulatory networks. *FEMS Microbiol. Rev.*, **34**, 866–882.
- Jagodnik, J., Thieffry, D. and Guillier, M. (2016) Bacterial small RNAs in mixed regulatory circuits. In: de Bruijn, F.J. (ed). *Stress and environmental regulation of gene expression and adaptation in bacteria*. Wiley-Blackwell Publishers, Hoboken, pp. 371–382.
- Beisel, C.L. and Storz, G. (2011) The base-pairing RNA Spot 42 participates in a multioutput feedforward loop to help enact catabolite repression in *Escherichia coli*. *Mol. Cell*, **41**, 286–297.
- Papenfort, K., Espinosa, E., Casades, J. and Vogel, J. (2015) Small RNA-based feedforward loop with AND-gate logic regulates extrachromosomal DNA transfer in *Salmonella*. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E4772–E4781.

31. Rhee, J.E., Sheng, W., Morgan, L.K., Nolet, R., Liao, X. and Kenney, L.J. (2008) Amino acids important for DNA recognition by the response regulator OmpR. *J. Biol. Chem.*, **283**, 8664–8677.
32. Walther, D., Li, Y., Liu, Y., Anand, G., Yan, J. and Kenney, L.J. (2011) *Salmonella enterica* response regulator SsrB relieves H-NS silencing by displacing H-NS bound in polymerization mode and directly activates transcription. *J. Biol. Chem.*, **286**, 1895–1902.
33. Ross, W. and Gourse, R.L. (2009) Analysis of RNA polymerase-promoter complex formation. *Methods*, **47**, 13–24.
34. Barbieri, C.M., Wu, T. and Stock, A.M. (2013) Comprehensive analysis of OmpR phosphorylation, dimerization, and DNA binding supports a canonical model for activation. *J. Mol. Biol.*, **425**, 1612–1626.
35. Miller, J.H. (1992) *A Short Course in Bacterial Genetics*, Cold Spring Harbor Laboratory Press, NY.
36. Ross, W., Thompson, J.F., Newlands, J.T. and Gourse, R.L. (1990) *E. coli* Fis protein activates ribosomal RNA transcription *in vitro* and *in vivo*. *EMBO J.*, **9**, 3733–3742.
37. Cai, S.J. and Inouye, M. (2002) EnvZ-OmpR interaction and osmoregulation in *Escherichia coli*. *J. Biol. Chem.*, **277**, 24155–24161.
38. Vanderpool, C.K. and Gottesman, S. (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol. Microbiol.*, **54**, 1076–1089.
39. Hsing, W., Russo, F.D., Bernd, K.K. and Silhavy, T.J. (1998) Mutations that alter the kinase and phosphatase activities of the two-component sensor EnvZ. *J. Bacteriol.*, **180**, 4538–4546.
40. Verhoef, C., Lugtenberg, B., van Boxtel, R., de Graaff, P. and Verheij, H. (1979) Genetics and biochemistry of the peptidoglycan-associated proteins b and c of *Escherichia coli* K12. *Mol. Gen. Genet.*, **169**, 137–146.
41. Hall, M.N. and Silhavy, T.J. (1981) The *ompB* locus and the regulation of the major outer membrane porin proteins of *Escherichia coli* K12. *J. Mol. Biol.*, **146**, 23–43.
42. Kinoshita-Kikuta, E., Aoki, Y., Kinoshita, E. and Koike, T. (2007) Label-free kinase profiling using phosphate affinity polyacrylamide gel electrophoresis. *Mol. Cell Proteomics*, **6**, 356–366.
43. Slauch, J.M. and Silhavy, T.J. (1989) Genetic analysis of the switch that controls porin gene expression in *Escherichia coli* K-12. *J. Mol. Biol.*, **210**, 281–292.
44. Tran, V.K., Oropeza, R. and Kenney, L.J. (2000) A single amino acid substitution in the C terminus of OmpR alters DNA recognition and phosphorylation. *J. Mol. Biol.*, **299**, 1257–1270.
45. Aiba, H., Nakasai, F., Mizushima, S. and Mizuno, T. (1989) Evidence for the physiological importance of the phosphotransfer between the two regulatory components, EnvZ and OmpR, in osmoregulation in *Escherichia coli*. *J. Biol. Chem.*, **264**, 14090–14094.
46. Mizuno, T., Kato, M., Jo, Y.L. and Mizushima, S. (1988) Interaction of OmpR, a positive regulator, with the osmoregulated *ompC* and *ompF* genes of *Escherichia coli*. Studies with wild-type and mutant OmpR proteins. *J. Biol. Chem.*, **263**, 1008–1012.
47. Hall, M.N. and Silhavy, T.J. (1979) Transcriptional regulation of *Escherichia coli* K-12 major outer membrane protein 1b. *J. Bacteriol.*, **140**, 342–350.
48. Batchelor, E. and Goulian, M. (2003) Robustness and the cycle of phosphorylation and dephosphorylation in a two-component regulatory system. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 691–696.
49. Siryaporn, A., Perchuk, B.S., Laub, M.T. and Goulian, M. (2010) Evolving a robust signal transduction pathway from weak cross-talk. *Mol. Syst. Biol.*, **6**, 452.
50. Hsing, W. and Silhavy, T.J. (1997) Function of conserved histidine-243 in phosphatase activity of EnvZ, the sensor for porin osmoregulation in *Escherichia coli*. *J. Bacteriol.*, **179**, 3729–3735.
51. Delias, N. and Forst, S. (2001) *MicF*: An antisense RNA gene involved in response of *Escherichia coli* to global stress factors. *J. Mol. Biol.*, **313**, 1–12.
52. Rentschler, A.E., Lovrich, S.D., Fitton, R., Enos-Berlage, J. and Schwan, W.R. (2013) OmpR regulation of the uropathogenic *Escherichia coliformB* gene in an acidic/high osmolality environment. *Microbiology*, **159**, 316–327.
53. Wang, L.C., Morgan, L.K., Godakumbura, P., Kenney, L.J. and Anand, G.S. (2012) The inner membrane histidine kinase EnvZ senses osmolality via helix-coil transitions in the cytoplasm. *EMBO J.*, **31**, 2648–2659.
54. Foo, Y.H., Gao, Y., Zhang, H. and Kenney, L.J. (2015) Cytoplasmic sensing by the inner membrane histidine kinase EnvZ. *Progr. Biophys. Mol. Biol.*, **118**, 119–129.
55. Tsui, P., Huang, L. and Freundlich, M. (1991) Integration host factor binds specifically to multiple sites in the *ompB* promoter of *Escherichia coli* and inhibits transcription. *J. Bacteriol.*, **173**, 5800–5807.
56. Huang, L., Tsui, P. and Freundlich, M. (1992) Positive and negative control of *ompB* transcription in *Escherichia coli* by cyclic AMP and the cyclic AMP receptor protein. *J. Bacteriol.*, **174**, 664–670.
57. Shinar, G., Milo, R., Martinez, M.R. and Alon, U. (2007) Input output robustness in simple bacterial signaling systems. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 19931–19935.
58. Dahl, M.K., Msadek, T., Kunst, F. and Rapoport, G. (1992) The phosphorylation state of the DegU response regulator acts as a molecular switch allowing either degradative enzyme synthesis or expression of genetic competence in *Bacillus subtilis*. *J. Biol. Chem.*, **267**, 14509–14514.
59. Latasa, C., Garcia, B., Echeverez, M., Toledo-Arana, A., Valle, J., Campoy, S., Garcia-del Portillo, F., Solano, C. and Lasa, I. (2012) *Salmonella* biofilm development depends on the phosphorylation status of RcsB. *J. Bacteriol.*, **194**, 3708–3722.
60. Gopel, Y. and Gorke, B. (2012) Rewiring two-component signal transduction with small RNAs. *Curr. Opin. Microbiol.*, **15**, 132–139.
61. Mandin, P. and Guillier, M. (2013) Expanding control in bacteria: interplay between small RNAs and transcriptional regulators to control gene expression. *Curr. Opin. Microbiol.*, **16**, 125–132.
62. Corcoran, C.P., Podkaminski, D., Papenfort, K., Urban, J.H., Hinton, J.C. and Vogel, J. (2012) Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Mol. Microbiol.*, **84**, 428–445.
63. Holmqvist, E., Unoson, C., Reimegard, J. and Wagner, E.G. (2012) A mixed double negative feedback loop between the sRNA MicF and the global regulator Lrp. *Mol. Microbiol.*, **84**, 414–427.
64. Rutherford, S.T., van Kessel, J.C., Shao, Y. and Bassler, B.L. (2011) AphA and LuxR/HapR reciprocally control quorum sensing in vibrios. *Genes Dev.*, **25**, 397–408.
65. Svenningsen, S.L., Waters, C.M. and Bassler, B.L. (2008) A negative feedback loop involving small RNAs accelerates *Vibrio cholerae*'s transition out of quorum-sensing mode. *Genes Dev.*, **22**, 226–238.
66. Tu, K.C., Waters, C.M., Svenningsen, S.L. and Bassler, B.L. (2008) A small-RNA-mediated negative feedback loop controls quorum-sensing dynamics in *Vibrio harveyi*. *Mol. Microbiol.*, **70**, 896–907.
67. Tu, K.C., Long, T., Svenningsen, S.L., Wingreen, N.S. and Bassler, B.L. (2010) Negative feedback loops involving small regulatory RNAs precisely control the *Vibrio harveyi* quorum-sensing response. *Mol. Cell*, **37**, 567–579.
68. Westermann, A.J., Forstner, K.U., Amman, F., Barquist, L., Chao, Y., Schulte, L.N., Muller, L., Reinhardt, R., Stadler, P.F. and Vogel, J. (2016) Dual RNA-seq unveils noncoding RNA functions in host-pathogen interactions. *Nature*, **529**, 496–501.
69. Alon, U. (2006) *Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall/CRC, Boca Raton.
70. LaPorte, D.C., Thorsness, P.E. and Koshland, D.E. Jr (1985) Compensatory phosphorylation of isocitrate dehydrogenase. A mechanism for adaptation to the intracellular environment. *J. Biol. Chem.*, **260**, 10563–10568.
71. Shin, D. and Groisman, E.A. (2005) Signal-dependent binding of the response regulators PhoP and PmrA to their target promoters *in vivo*. *J. Biol. Chem.*, **280**, 4089–4094.

Supplementary data:

Strains

Strains used in this study are derivatives of MG1655 and are listed in Table S1. Mutations *ΔmalT::kan*, *ΔenvZ::kan*, *ΔackA-pta::cat*, and *ΔomrAB::amp* were engineered by homologous recombination of PCR fragments carrying the antibiotic resistance gene surrounded by homology regions in a strain expressing the recombination genes from a mini-λ (Yu, Ellis et al. 2000; Court, Swaminathan et al. 2003). Sequence of primers used for those PCR is given in Table S2. The *ΔmicF::kan* allele (strain GSO159 in (Hobbs, Astarita et al. 2010)) was obtained from E. Hobbs. These mutations, as well as the *ara::lacI^q-Kan^R* allele, where a *lacI^q-Kan^R* construct replaces the *araBAD* genes (constructed by C. Chiaruttini), were then moved by P1 transduction when necessary.

Construction of *Tet-P_{lac}-ompR* allele was also by recombineering. The TetR gene was first amplified from plasmid DNA with primers TetplacompRfor/rev1, and the product was then subjected to a second PCR with primers TetplacompRfor/rev2. The resulting DNA fragment was then recombined in an *envZ* wt or *DenvZ::kan* strain. In this *Tet-P_{lac}-ompR* construct, transcription of *ompR* is expected to start 55 nts upstream of *ompR* start codon.

Mutations *envZ390*, *envZ473*, *ompR107* and *ompR472* were obtained from T. Silhavy (Table S1); they were co-transduced with *malT* into *ΔmalT::kan* recipient strains, selecting for growth on maltose. The presence of mutations was checked by sequencing.

Strains carrying *lacZ* fusions to *omrA* or *omrB* promoters were described previously; these fusions encompass nts -104 to +8 of PomrA or PomrB (relative to TSS) upstream of *lacZ* (Guillier and Gottesman 2006). A first set of *ompC*- and *ompF-lacZ* fusions, referred to as "long" fusions in Table S1, was obtained from T. Silhavy (strains MH225 (Hall and Silhavy 1979) and MH513 (Hall and Silhavy 1981) respectively) and moved by P1 transduction. Because these fusions are subject to post-transcriptional control by MicC or MicF for instance (our unpublished data), leaderless promoter fusions were engineered by recombination to focus on transcription initiation. They carry respectively nts -356 to +15 of *ompC* (relative to transcription start site from P1, itself 81 nts upstream of *ompC* start codon) or nts -410 to +5 of *ompF* (relative to transcription start site), upstream of the *lacZ* gene, starting 17 nts upstream

of its start codon. In other words, these fusions should contain all known OmpR binding sites and extend 14 nts (for *ompC*) or 4 nts (for *ompF*) beyond the transcription start site. PCR fragments suitable for recombineering were generated by PCR as follows: (i) *ompC* or *ompF* promoter region was amplified by PCR with primers pairs lacIupompCfor2/+15ompC-lacZrev or lacIupompFfor/ompF-lacZrev respectively, (ii) fragment -17 to +120 of lacZ gene was PCR amplified with primers pairs +15ompC-lacZfor/lacZ96-120 (for *ompC*) or ompF-lacZfor/lacZ96-120 (for *ompF*) and (iii) products of PCR (i) and (ii) were purified and mixed to allow PCR amplification of the full-length fragment using primers pairs lacIupompCfor2/lacZ96-120 or lacIupompFfor/lacZ96-120. These fragments were then used to replace by recombineering the P_{tet}-*cat-sacB* construct in front of *lacZ* in strain MG1508. Presence of fusions was confirmed by sequencing.

Translational fusions between *ompC* or *ompF* and *lacZ* under the control of the Ptet promoter were also constructed by recombineering in strain MG1508. For this purpose, PCR products were amplified with primers pairs 5'Ptet-ompC/3'ompC-lacZ and 5'Ptet-ompF/3'ompF-lacZ before recombineering. The resulting fusions were checked by sequencing and carry the 81-nts long 5'UTR of *ompC* or the 110-nts 5'UTR of *ompF* followed by the 30 first nts of the ORF between Ptet and *lacZ* (starting at its 10th aa).

ompR-lacZ translational fusions, wt or carrying the mut2 change, under the control of P_{BAD} promoter, were also made. They carry nts -35 (which corresponds to TSS-1 in (Quinn, Cameron et al. 2014)) to +30 of *ompR* or *ompRmut2* (relative to AUG) fused to *lacZ* 10th aa. To construct these fusions, the required *ompR* (*mut2*) regions were amplified with primers PBADompR-35 and lacZ37-73rev using DNA from strains carrying *ompR* (*mut2*)-*lacZ* fusions (but with a longer 5'UTR, see strains MG1398 and MG1403 in (Guillier and Gottesman 2008) for instance) as templates. The resulting PCR products were then recombined in strain PM1205 (Mandin and Gottesman 2009) that carries a P_{BAD}-*cat-sacB-lacZ* allele, and fusions were checked by sequencing.

Transcriptional P_{ompR}-*lacZ* fusion, where nts -243 to -35 of *ompR* (relative to start codon) are placed in front of *lacZ*, starting at nt-17, was constructed as follows. Promoter region of *ompR* and *lacZ* early ORF were PCR amplified with primers pairs 5'lacIompR/ompR-lacZrev and ompR-lacZfor/lacZ37-73rev respectively. PCR products were then mixed and re-amplified

with primers 5'lacIPompR and lacZ37-73rev. The resulting DNA fragment was then recombined into strain MG1508 and fusion was checked by sequencing. Similarly, for construction of P_{dpA}-lacZ, *dpA* promoter region (nts -200 to +14) was amplified with primers 5'rrnB-tppB and 3'tppB-lacZ and used to replace by recombineering the *cat-sacB* cassette in a strain carrying a *cat-sacB-lacZ* construct downstream of the *rrnB* t2 transcriptional terminator (at the *lacZ* locus).

Plasmids

Templates for *in vitro* transcription were prepared as follows. DNA fragments encompassing nts -104 to +50 (relative to transcription start site) of PomrA or PomrB, nts -437 to +50 of PompC, nts -70 to +50 of PsgrS, or nts -60 to +50 of PomrA (PomrA-60) were amplified by PCR using primers allowing insertion of EcoRI and HindIII restriction sites (Table S2). These fragments were then cloned in the pRLG770 vector (Ross, Thompson et al. 1990) cut by the same enzymes to give plasmids pRLG-omrA, pRLG-omrB, pRLG-ompC, pRLG-sgrS and pRLG-omrA-60 respectively, where the cloned promoter regions drive transcription of a product containing the first 50 nts of each RNA fused to 5S rRNA.

1. Yu, D. G., Ellis, H. M., Lee, E. C., Jenkins, N. A., Copeland, N. G., and Court, D. L. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 97, 5978-5983
2. Court, D. L., Swaminathan, S., Yu, D., Wilson, H., Baker, T., Bubunenko, M., Sawitzke, J., and Sharan, S. K. (2003) Mini-lambda: a tractable system for chromosome and BAC engineering. *Gene* 315, 63-69
3. Hobbs, E. C., Astarita, J. L., and Storz, G. (2010) Small RNAs and small proteins involved in resistance to cell envelope stress and acid shock in *Escherichia coli*: analysis of a bar-coded mutant collection. *J. Bact.* 192, 59-67
4. Guillier, M., and Gottesman, S. (2006) Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. *Mol. Microbiol.* 59, 231-247
5. Hall, M. N., and Silhavy, T. J. (1979) Transcriptional regulation of *Escherichia coli* K-12 major outer membrane protein 1b. *J. Bacteriol.* 140, 342-350
6. Hall, M. N., and Silhavy, T. J. (1981) The *ompB* locus and the regulation of the major outer membrane porin proteins of *Escherichia coli* K12. *J. Mol. Biol.* 146, 23-43
7. Quinn, H. J., Cameron, A. D., and Dorman, C. J. (2014) Bacterial regulon evolution: distinct responses and roles for the identical OmpR proteins of *Salmonella Typhimurium* and *Escherichia coli* in the acid stress response. *PLoS Genet.* 10, e1004215
8. Guillier, M., and Gottesman, S. (2008) The 5' end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. *Nucleic acids research* 36, 6781-6794
9. Mandin, P., and Gottesman, S. (2009) A genetic approach for finding small RNAs regulators of genes of interest identifies RybC as regulating the DpiA/DpiB two-component system. *Mol. Microbiol.* 72, 551-565
10. Ross, w., Thompson, J. F., Newlands, J. T., and Gourse, R. L. (1990) *E. coli* Fis protein activates ribosomal RNA transcription *in vitro* and *in vivo*. *EMBO J.* 9, 3733-3742

Table S1. Strains and plasmids used in this study.

Strain	Characteristics	Source/Construction
MG1655	wt strain	
DJ480	MG1655 <i>AlacX74</i>	D. Jin, NIH
JMS97	<i>ompR107</i> mutant strain	T. Silhavy's lab
MH225	<i>ompC-lacZ</i> "long" fusion	(Hall and Silhavy 1979)
MH513	<i>ompF-lacZ</i> "long" fusion	(Hall and Silhavy 1981)
MH760	<i>ompR472</i> mutant strain	(Hall and Silhavy 1981)
MH1471	<i>envZ473</i> mutant strain	(Hall and Silhavy 1981)
NM300	DJ480 mini- λ tet ^R	N. Majdalani, NIH
WH68	<i>envZ390</i> mutant strain	T. Silhavy's lab
CV5200	DJ480 P _{sgrS} - <i>lacZ</i> (nts -70 to +4 of P _{sgrS} fused to <i>lacZ</i>)	(Vanderpool and Gottesman 2004)
MG1003	DJ480 Δ <i>omrAB::kan</i>	(Guillier and Gottesman 2006)
MG1004	DJ480 P _{omrA} - <i>lacZ</i> (nts -104 to +8 of P _{omrA} fused to <i>lacZ</i>)	(Guillier and Gottesman 2006)
MG1005	DJ480 P _{omrB} - <i>lacZ</i> (nts -104 to +8 of P _{omrB} fused to <i>lacZ</i>)	(Guillier and Gottesman 2006)
MG1014	DJ480 P _{omrA} - <i>lacZ</i> Δ <i>omrAB::kan</i>	MG1004+P1(MG1003)
MG1017	DJ480 P _{omrB} - <i>lacZ</i> Δ <i>omrAB::kan</i>	MG1005+P1(MG1003)
MG1022	DJ480 P _{omrA-60} - <i>lacZ</i> (nts -60 to +8 of P _{omrA} fused to <i>lacZ</i>)	(Guillier and Gottesman 2006)
MG1035	DJ480 Δ <i>ompR::cm</i>	(Guillier and Gottesman 2006)
MG1284	NM300 Δ <i>malT::kan</i>	Recombineering
MG1285	DJ480 Δ <i>malT::kan</i>	DJ480+P1(MG1284)
MG1286	MG1004 Δ <i>malT::kan</i>	MG1004+P1(MG1284)
MG1287	MG1005 Δ <i>malT::kan</i>	MG1005+P1(MG1284)
MG1295	DJ480 <i>ompC-lacZ</i> ("long" fusion)	DJ480+P1(MH225)
MG1296	DJ480 <i>ompF-lacZ</i> ("long" fusion)	DJ480+P1(MH513)
MG1297	DJ480 Δ <i>malT::kan ompC-lacZ</i> ("long" fusion)	MG1285+P1(MH225)
MG1298	DJ480 Δ <i>malT::kan ompF-lacZ</i> ("long" fusion)	MG1285+P1(MH513)
MG1299	DJ480 P _{omrA} - <i>lacZ envZ390</i>	MG1286+P1(WH68)
MG1300	DJ480 P _{omrA} - <i>lacZ envZ473</i>	MG1286+P1(MH1471)
MG1301	DJ480 P _{omrA} - <i>lacZ ompR107</i>	MG1286+P1(JMS97)
MG1302	DJ480 P _{omrA} - <i>lacZ ompR472</i>	MG1286+P1(MH760)
MG1303	DJ480 P _{omrB} - <i>lacZ envZ390</i>	MG1287+P1(WH68)
MG1304	DJ480 P _{omrB} - <i>lacZ envZ473</i>	MG1287+P1(MH1471)
MG1305	DJ480 P _{omrB} - <i>lacZ ompR107</i>	MG1287+P1(JMS97)
MG1306	DJ480 P _{omrB} - <i>lacZ ompR472</i>	MG1287+P1(MH760)
MG1310	DJ480 <i>ompC-lacZ</i> ("long" fusion) <i>envZ390</i>	MG1297+P1(WH68)
MG1311	DJ480 <i>ompC-lacZ</i> ("long" fusion) <i>envZ473</i>	MG1297+P1(MH1471)
MG1312	DJ480 <i>ompC-lacZ</i> ("long" fusion) <i>ompR107</i>	MG1297+P1(JMS97)
MG1313	DJ480 <i>ompC-lacZ</i> ("long" fusion) <i>ompR472</i>	MG1297+P1(MH760)
MG1314	DJ480 <i>ompF-lacZ</i> ("long" fusion) <i>envZ390</i>	MG1298+P1(WH68)
MG1315	DJ480 <i>ompF-lacZ</i> ("long" fusion) <i>envZ473</i>	MG1298+P1(MH1471)
MG1316	DJ480 <i>ompF-lacZ</i> ("long" fusion) <i>ompR107</i>	MG1298+P1(JMS97)
MG1317	DJ480 <i>ompF-lacZ</i> ("long" fusion) <i>ompR472</i>	MG1298+P1(MH760)
MG1346	NM300 <i>ompR::cat-sacB</i>	Recombineering
MG1351	DJ480 <i>ompRmut2</i>	Recombineering
MG1413	DJ480 P _{omrA} - <i>lacZ ompRmut2</i>	Lysogenization of MG1351 with λ RS-104+8OmrA, see (Guillier and Gottesman 2006)
MG1414	DJ480 P _{omrB} - <i>lacZ ompRmut2</i>	Lysogenization of MG1351 with λ RS-104+8OmrB, see (Guillier and Gottesman 2006)
MG1508	MG1655 <i>mal::lacI^q P_{LtetO-1}-cat-sacB-lacZ</i> mini- λ tet ^R	(Coornaert, Chiaruttini et al. 2013)
MG1665	DJ480 <i>ompRmut2</i> Δ <i>omrAB::kan</i>	MG1351+P1(MG1003)
MG1666	DJ480 P _{omrA} - <i>lacZ ompRmut2</i> Δ <i>omrAB::kan</i>	MG1413+P1(MG1003)

MG1667	DJ480 P _{omrB} -lacZ ompRmut2 ΔomrAB::kan	MG1414+P1(MG1003)
MG1670	DJ480 envZ390	MG1285+P1(WH68)
MG1671	DJ480 envZ473	MG1285+P1(MH1471)
MG1672	DJ480 ompR107	MG1285+P1(JMS97)
MG1673	DJ480 ompR472	MG1285+P1(MH760)
MG1676	MG1508 P _{ompF} -lacZ (nts -410 to +5 of P _{ompF} fused to lacZ)	Recombineering
MG1686	DJ480 P _{ompF} -lacZ ΔomrAB::kan	MG1003+P1(MG1676)
MG1688	DJ480 P _{ompF} -lacZ ompRmut2 ΔomrAB::kan	MG1665+P1(MG1676)
MG1690	DJ480 P _{ompF} -lacZ	DJ480+P1(MG1676)
MG1692	DJ480 P _{ompF} -lacZ envZ390	MG1670+P1(MG1676)
MG1694	DJ480 P _{ompF} -lacZ envZ473	MG1671+P1(MG1676)
MG1696	DJ480 P _{ompF} -lacZ ompR107	MG1672+P1(MG1676)
MG1698	DJ480 P _{ompF} -lacZ ompR472	MG1673 + P1(MG1676)
MG1807	DJ480 ompC-lacZ ("long" fusion) ΔompR::cm	MG1295+P1(MG1035)
MG1808	DJ480 ompF-lacZ ("long" fusion) ΔompR::cm	MG1296+P1(MG1035)
MG1810	DJ480 P _{ompF} -lacZ ΔompR::cm	MG1690+P1(MG1035)
MG1811	DJ480 P _{omrA} -lacZ ΔompR::cm	MG1004+P1(MG1035)
MG1812	DJ480 P _{omrB} -lacZ ΔompR::cm	MG1005+P1(MG1035)
MG1831	MG1508 P _{LietO-1} -ompC-lacZ (nts -81 to +30 of ompC (relative to AUG) fused to lacZ 10 th aa)	Recombineering
MG1833	MG1508 P _{LietO-1} -ompF-lacZ (nts -110 to +30 of ompF (relative to AUG) fused to lacZ 10 th aa)	Recombineering
MG1860	MG1508 P _{LietO-1} -ompC-lacZ ΔompR::cm	MG1831+P1(MG1035)
MG1861	MG1508 P _{LietO-1} -ompF-lacZ ΔompR::cm	MG1833+P1(MG1035)
MG1863	MG1508 P _{ompC} -lacZ (nts -356 to +15 of P _{ompC} fused to lacZ)	Recombineering
MG1891	MG1508 P _{ompC} -lacZ ΔompR::cm	MG1863+P1(MG1035)
MG1892	DJ480 P _{ompC} -lacZ	DJ480+P1(MG1863)
MG1893	DJ480 P _{ompC} -lacZ envZ390	MG1670+P1(MG1863)
MG1894	DJ480 P _{ompC} -lacZ envZ473	MG1671+P1(MG1863)
MG1895	DJ480 P _{ompC} -lacZ ompR107	MG1672+P1(MG1863)
MG1896	DJ480 P _{ompC} -lacZ ompR472	MG1673+P1(MG1863)
MG1901	DJ480 P _{ompC} -lacZ ΔomrAB::kan	MG1003+P1(MG1863)
MG1902	DJ480 P _{ompC} -lacZ ompRmut2 ΔomrAB::kan	MG1665+P1(MG1863)
MG1908	DJ480 P _{omrA} -lacZ ompR107 ΔomrAB::kan	MG1301+P1(MG1003)
MG1909	DJ480 P _{omrB} -lacZ ompR107 ΔomrAB::kan	MG1305+P1(MG1003)
MG1910	DJ480 P _{ompF} -lacZ ompR107 ΔomrAB::kan	MG1696+P1(MG1003)
MG1911	DJ480 P _{ompC} -lacZ ompR107 ΔomrAB::kan	MG1895+P1(MG1003)
MG1969	MG1508 P _{ompR} -lacZ (nts -208 to -1 of P _{ompR} (relative to TSS-1 in (Quinn, Cameron et al. 2014))fused to lacZ)	Recombineering
MG1972	MG1508 P _{ompR} -lacZ malt::kan	MG1969+P1(MG1284)
MG1974	MG1508 P _{ompR} -lacZ ΔompR::cm	MG1969+P1(MG1035)
MG1979	MG1508 P _{ompR} -lacZ envZ390	MG1972+P1(WH68)
MG1980	MG1508 P _{ompR} -lacZ envZ473	MG1972+P1(MH1471)
MG1981	MG1508 P _{ompR} -lacZ ompR107	MG1972+P1(JMS97)
MG1982	MG1508 P _{ompR} -lacZ ompR472	MG1972+P1(MH760)
MG1985	Tet-P _{lac} -ompR strain	Recombineering
MG1988	DJ480 P _{omrA} -lacZ ara::lacI ^q -Kan ^R	MG1004+P1(ara::lacI ^q -Kan ^R)
MG2000	DJ480 P _{omrA} -lacZ Tet-P _{lac} -ompR ara::lacI ^q -Kan ^R	MG1004+P1(MG1985), then +P1(ara::lacI ^q -Kan ^R)
MG2002	DJ480 P _{omrB} -lacZ Tet-P _{lac} -ompR ara::lacI ^q -Kan ^R	MG1005+P1(MG1985), then +P1(ara::lacI ^q -Kan ^R)
MG2004	DJ480 P _{ompF} -lacZ Tet-P _{lac} -ompR ara::lacI ^q -Kan ^R	MG1690+P1(MG1985), then +P1(ara::lacI ^q -Kan ^R)
MG2006	DJ480 P _{ompC} -lacZ Tet-P _{lac} -ompR ara::lacI ^q -Kan ^R	MG1892+P1(MG1985), then +P1(ara::lacI ^q -Kan ^R)
MG2018	DJ480 P _{ompC} -lacZ ΔompR::cm	MG1892 + P1(MG1035)
MG2023	ΔenvZ::kan strain	Recombineering

MG2024	Tet-Plac- <i>ompR</i> $\Delta envZ::kan$ strain	Recombineering
MG2034	NM300 $\Delta ackA-pta::cat$	Recombineering
MG2039	DJ480 P _{omrA} - <i>lacZ</i> $\Delta envZ::kan \Delta ackA-pta::cat$	MG1004+P1(MG2023), then +P1(MG2034)
MG2040	DJ480 P _{omrA} - <i>lacZ</i> Tet-P _{lac} - <i>ompR</i> $\Delta envZ::kan \Delta ackA-pta::cat$	MG1004+P1(MG2024), then +P1(MG2034)
MG2041	DJ480 P _{omrB} - <i>lacZ</i> $\Delta envZ::kan \Delta ackA-pta::cat$	MG1005+P1(MG2023), then +P1(MG2034)
MG2042	DJ480 P _{omrB} - <i>lacZ</i> Tet-P _{lac} - <i>ompR</i> $\Delta envZ::kan \Delta ackA-pta::cat$	MG1005+P1(MG2024), then +P1(MG2034)
MG2043	DJ480 P _{ompF} - <i>lacZ</i> $\Delta envZ::kan \Delta ackA-pta::cat$	MG1690+P1(MG2023), then +P1(MG2034)
MG2044	DJ480 P _{ompF} - <i>lacZ</i> Tet-P _{lac} - <i>ompR</i> $\Delta envZ::kan \Delta ackA-pta::cat$	MG1690+P1(MG2024), then +P1(MG2034)
MG2045	DJ480 P _{ompC} - <i>lacZ</i> $\Delta envZ::kan \Delta ackA-pta::cat$	MG1892+P1(MG2023), then +P1(MG2034)
MG2046	DJ480 P _{ompC} - <i>lacZ</i> Tet-P _{lac} - <i>ompR</i> $\Delta envZ::kan \Delta ackA-pta::cat$	MG1892+P1(MG2024), then +P1(MG2034)
MG2127	MG1508 P _{LetO-1} - <i>ompF-lacZ</i> Tet-P _{lac} - <i>ompR</i>	MG1833+P1(MG1985)
MG2138	MG1655 <i>mal::lacI^q</i> P _{dtpA} - <i>lacZ</i> (nts -200 to +14 of P _{dtpA} fused to <i>lacZ</i>)	Recombineering
MG2167	MG1508 P _{LetO-1} - <i>ompF-lacZ</i> Tet-P _{lac} - <i>ompR</i> $\Delta micF::kan$	MG2127+P1($\square micF::kan$)
MG2169	MG1655 <i>mal::lacI^q</i> P _{dtpA} - <i>lacZ</i> $\Delta ompR::cm$	MG2138+P1(MG1035)
MG2170	MG1655 <i>mal::lacI^q</i> P _{dtpA} - <i>lacZ</i> Tet-P _{lac} - <i>ompR</i>	MG2138+P1(MG1985)
MG2174	MG1655 <i>mal::lacI^q</i> $\Delta araBAD$ <i>araC⁺</i> P _{BAD} - <i>ompR-lacZ</i> (nts -35 to +30 of <i>ompR</i> (relative to AUG) fused to <i>lacZ</i> 10 th aa)	Recombineering
MG2175	MG1655 <i>mal::lacI^q</i> $\Delta araBAD$ <i>araC⁺</i> P _{BAD} - <i>ompRmut2-lacZ</i> (nts -35 to +30 of <i>ompR</i> (relative to AUG) fused to <i>lacZ</i> 10 th aa)	Recombineering
MG2176	MG1655 <i>mal::lacI^q</i> $\Delta araBAD$ <i>araC⁺</i> P _{BAD} - <i>ompR-lacZ</i> $\Delta omrAB::kan$	MG2174+P1(MG1003)
MG2177	MG1655 <i>mal::lacI^q</i> $\Delta araBAD$ <i>araC⁺</i> P _{BAD} - <i>ompRmut2-lacZ</i> $\Delta omrAB::kan$	MG2175+P1(MG1003)
MG2178	DJ480 P _{omrA-60} - <i>lacZ</i> $\Delta ompR::cm$	MG1022+P1(MG1035)
MG2179	DJ480 P _{sgrS} - <i>lacZ</i> $\Delta ompR::cm$	CV5200+P1(MG1035)
MG2180	DJ480 P _{omrA-60} - <i>lacZ</i> $\Delta malT::kan$	MG1022+P1(MG1284)
MG2181	DJ480 P _{sgrS} - <i>lacZ</i> $\Delta malT::kan$	CV5200+P1(MG1284)
MG2182	DJ480 P _{omrA-60} - <i>lacZ</i> <i>envZ390</i>	MG2180+P1(WH68)
MG2183	DJ480 P _{omrA-60} - <i>lacZ</i> <i>envZ473</i>	MG2180+P1(MH1471)
MG2184	DJ480 P _{omrA-60} - <i>lacZ</i> <i>ompR107</i>	MG2180+P1(JMS97)
MG2185	DJ480 P _{omrA-60} - <i>lacZ</i> <i>ompR472</i>	MG2180+P1(MH760)
MG2186	DJ480 P _{sgrS} - <i>lacZ</i> <i>envZ390</i>	MG2181+P1(WH68)
MG2187	DJ480 P _{sgrS} - <i>lacZ</i> <i>envZ473</i>	MG2181+P1(MH1471)
MG2188	DJ480 P _{sgrS} - <i>lacZ</i> <i>ompR107</i>	MG2181+P1(JMS97)
MG2189	DJ480 P _{sgrS} - <i>lacZ</i> <i>ompR472</i>	MG2181+P1(MH760)
MG2198	$\Delta omrAB::amp$ strain	Recombineering
MG2201	DJ480 P _{omrA} - <i>lacZ</i> $\Delta omrAB::amp$	MG1004+P1(MG2198)
MG2202	DJ480 P _{omrB} - <i>lacZ</i> $\Delta omrAB::amp$	MG1005+P1(MG2198)
MG2203	DJ480 P _{ompC} - <i>lacZ</i> $\Delta omrAB::amp$	MG1892+P1(MG2198)
MG2204	DJ480 P _{omrA} - <i>lacZ</i> Tet-P _{lac} - <i>ompR</i> $\Delta envZ::kan \Delta ackA-pta::cat$ $\Delta omrAB::amp$	MG2040+P1(MG2198)
MG2205	DJ480 P _{omrB} - <i>lacZ</i> Tet-P _{lac} - <i>ompR</i> $\Delta envZ::kan \Delta ackA-pta::cat$ $\Delta omrAB::amp$	MG2042+P1(MG2198)
MG2206	DJ480 P _{ompC} - <i>lacZ</i> Tet-P _{lac} - <i>ompR</i> $\Delta envZ::kan \Delta ackA-pta::cat$ $\Delta omrAB::amp$	MG2046+P1(MG2198)
Plasmids	Characteristics	Source
pRLG770	ColEI, AmpR, used for cloning <i>omrA</i> , <i>omrB</i> or <i>ompC</i> promoter in front of 5S RNA gene	(Ross, Thompson et al. 1990)
pRLG-OmrA	pRLG770 derivative, fragment -114 to +50 of <i>PomrA</i> (relative to TSS) cloned upstream of 5S	This study
pRLG-OmrB	pRLG770 derivative, fragment -114 to +50 of <i>PomrB</i> (relative to TSS) cloned upstream of 5S	This study

pRLG-ompC	pRLG770 derivative, fragment -437 to +50 of PompC (relative to TSS) cloned upstream of 5S	This study
pRLG-OmrA-60	pRLG770 derivative, fragment -60 to +50 of PomrA (relative to TSS) cloned upstream of 5S	This study
pRLG-SgrS	pRLG770 derivative, fragment -70 to +50 of PsgrS (relative to TSS) cloned upstream of 5S	This study
pBRplac	empty vector, pBR322 derivative carrying a modified P _{LacO-1} promoter, Amp ^R , Tet ^R	(Guillier and Gottesman 2006)
pBRplacOmrA	<i>omrA</i> under control of P _{LacO-1}	(Guillier and Gottesman 2006)
pBRplacOmrB	<i>omrB</i> under control of P _{LacO-1}	(Guillier and Gottesman 2006)
pBRplacOmrA mut2	<i>omrA</i> mut2 under control of P _{LacO-1}	(Guillier and Gottesman 2008)
pBRplacOmrB mut2	<i>omrB</i> mut2 under control of P _{LacO-1}	(Guillier and Gottesman 2008)

1. Hall, M. N., and Silhavy, T. J. (1979) J. Bacteriol. 140, 342-350
2. Hall, M. N., and Silhavy, T. J. (1981) J. Mol. Biol. 146, 23-43
3. Vanderpool, C. K., and Gottesman, S. (2004) Mol. Microbiol. 54, 1076-1089
4. Guillier, M., and Gottesman, S. (2006) Mol. Microbiol. 59, 231-247
5. Coornaert, A., Chiaruttini, C., Springer, M., and Guillier, M. (2013) PLoS Genet. 9, e1003156
6. Quinn, H. J., Cameron, A. D., and Dorman, C. J. (2014) PLoS Genet. 10, e1004215
7. Ross, w., Thompson, J. F., Newlands, J. T., and Gourse, R. L. (1990) EMBO J. 9, 3733-3742
8. Guillier, M., and Gottesman, S. (2008) Nucleic acids research 36, 6781-6794

Table S2. Oligonucleotides used in this study.

Oligonucleotide	5'-> 3' sequence	Comment
Strain construction		
lacZ96-120	gctattaccgccagctggcgaaaggg	
lacZ37-73rev	gttgggtaacgccagggtttcccagtcacgacgttg	
malT::kanfor	ttcctcattttccacagtgaagtgattaactatgctgattccgAAAGCCACGTTGTGTCTCA A	Construction <i>malT::kan</i> allele
malT::kanrev	gagatgtaagccggatctggcgcgttatccggctaaacttacacTTAGAAAACTCATCGA GCA	
Δ envZ::kanfor	gtctactgtttgtaccggacggctctaagcatgaggcattgAAAGCCACGTTGTGTCTC AA	Construction Δ envZ::kan allele
Δ envZ::kanrev	aagcaaacgggaggcaccttcgctcccgtttatttaccTTAGAAAACTCATCGAGC A	
ackAcatfor	cctgacgttttttagccacgtatcaattataggtactccCCTGTGACGGAAGATCACTTC GC	Construction Δ ackA-pta::cat allele
ptacatrev	atccgcagcgcgaaagctcgggatgatgacgagattactgTTATCACTTATTCAGGCGTAG C	
Δ omr-ampfor	cgctgttgccgattgaccgctggtgcgcttggctcaggttgcCATTCAAATATGTATCCG CTC	Construction Δ omrAB::amp allele
Δ omr-amprev	cgacagtaaattaggtgcgaaaaaacctcgcgatccgcgaggttGTAAACTTGGTCTG ACAGTTA	
TetplacompRfor	ggctgaaattcataccagatttagctggtgacgaacgtgaTCCTAATTTTTGTTGACACT CTA	Construction <i>Tet-P_{lac}-ompR</i> allele
TetplacompRrev1	gtatctgttatccgctcacaatgtcaatgttatccgctcacatttatCTTCTTTATCGTAAAAA ATGCC	
TetplacompRrev2	aggttcgcaacaattgtaagcgtgtattcttaaaaaagcgtctcaGTATCTTGTATCCGCT CACAATG	
+15ompC-lacZfor	ggagaatggactgccgactgattaatcacacaggaacagctatgac	P _{ompC} -lacZ fusion
+15ompC-lacZrev	gtcatagctgtttcctgtgtgattaatcagtcggcaagtccattctcc	
lacIupompCfor2	gcgaagcggcatgcatttacgttgacaccatcgaatggcgcAAAGAAGGGTAAAAAAA ACCGAATGCG	
ompF-lacZfor	tggaaagatgcctgcagacatcacacaggaacagctatgac	P _{ompF} -lacZ fusion
ompF-lacZrev	gtcatagctgtttcctgtgtgatgctcgcaggcatctttcca	
lacIupompFfor	gcgaagcggcatgcatttacgttgacaccatcgaatggcgcTATTTGCAACCCCGCCAT AAC	
5'Ptet-ompC	tagagattgacatccctatcagtgatagagatactgagcacTTGCCGACTGATTAATGAG GG	P _{tet} -ompC-lacZ fusion
3'ompC-lacZ	taacgccagggtttcccagtcacgacgttgtaaacgacGACCAGGAGGGACAGTACT TTAAC	
5'Ptet-ompF	tagagattgacatccctatcagtgatagagatactgagcacAGACACATAAAGACACCA AACTC	P _{tet} -ompF-lacZ fusion
3'ompF-lacZ	taacgccagggtttcccagtcacgacgttgtaaacgacGATCACTGCCAGAATATTG CGC	
ompR-lacZfor	gtgagcttttaagaatacacgctcacacaggaacagctatgac	P _{ompR} -lacZ fusion
ompR-lacZrev	gtgtgagcgtgtattcttaaaaaagctcac	
5'lacIPompR	cgaagcggcatgcatttacgttgacaccatcgaatggcgcCCAGGGGCGTTTTTCATCTC G	

5'rrnB-tppB	ggcgcagaagccatcctgacggatggccttttgcgtttGTAAAAACGCTAAATAGCA GAACATATCG	P _{dtpA} -lacZ fusion
3'tppB-lacZ	cagtgaatccgtaatcatggtcatagctgtttcctgtgtgaGGGAGGGGAAACGGCAGGT AATGTT	
PBADompR-35	acctgacgcttttatcgcaactctactgtttccatTTACAAATTGTTGCGAACC	P _{BAD} -ompR(<i>mut2</i>)- lacZ fusion (with lacZ37-73rev)
Plasmids construction		
Eco-114OmrA	GCAGAATTCGCCGGTCATCAATCTGTAACAG	Construction pRLG-omrA
OmrA+50Hind	CGAAAGCTTACAGGGTACGAAGAGCGTACCG	
Eco-114OmrB	GCAGAATTCAAACGCAAATAAGTAAGACAAAAAAG	Construction pRLG-omrB
OmrB+50Hind	CGAAAGCTTTAATTCATGTGCTCAACCCGAAG	
Eco-437ompC	GCAGAATTCAAAGAAGGGTAAAAAAAACCGAATGC	Construction pRLG-ompC
ompC+50Hind	CGAAAGCTTGCTTTTTTATGCCACTGCATACTG	
Eco-70sgrS	GCAGAATTCATAAAAAGGGGAACCTCTGTGCAAAAAGACAG	Construction pRLG-sgrS
SgrS+50Hind	CGAAAGCTTGTAATAATAGTGCTGATAAAACTGACGCATG	
Eco-60OmrA	GCAGAATTCCTTCATTCTTTGCGTTTTCTCGCTGGCGAAG	Construction pRLG-omrA-60 (with OmrA+50Hind)
Biotinylated probes for Northern- Blot		
OmrA-probe	Bio-caggttggtgcaagagacaggtacgaagagcgtaccg	
OmrB-probe	Bio-cgcaggtggtgtaattcatgtgctcaaccgaagtga	
ompC-probe	Bio-cttcagcagcgtttgctgcccctgctaccagcagagctgg	
ompF-probe	Bio-gctactaacagagcagggacgatcactgccagaatattgc	
bolA-probe	Bio-ggttggaacgcccccttaattttctctatccgctcacgtac	
SsrA-probe	Bio-cgccactaacaactagcctgattaagtttaacgctca	
23S-probe	Bio-tttagccccgttacatcttccgcgagccg	

Table S3. Summary of b-galactosidase activities for results that are given only as ratios in the main text. Shown are the average values of two independent experiments, expressed either as the average of activities or as the average of the ratios with the reference strain (wt or carrying the vector control).

Fig.	Strain	Growth medium	Average b-galactosidase activity (Miller units)	Standard deviation	Average ratio +/- st. dev
2B	MG1004	LB	19.1	8.58	13.3 +/- 1.34
2B	MG1811	LB	1.42	0.60	
2B	MG1005	LB	49.4	22.8	7.1 +/- 1.49
2B	MG1812	LB	7.6	4.52	
2B	MG1863	LB	6010	2753	16.3 +/- 2.66
2B	MG1891	LB	356	119	
2B	MG1690	LB	3270	496	77.6 +/- 13.7
2B	MG1810	LB	42.1	0.23	
4A	MG1004/pHDB3	LB-Amp	32.6	10.3	11.8 +/- 1.50
4A	MG1004/pOmpR	LB-Amp	377	84.3	
4A	MG1005/pHDB3	LB-Amp	87.1	26.3	7.1 +/- 0.71
4A	MG1005/pOmpR	LB-Amp	607	136	
4A	MG1892/pHDB3	LB-Amp	7313	2711	1.32 +/- 0.22
4A	MG1892/pOmpR	LB-Amp	9534	3328	
4A	MG1690/pHDB3	LB-Amp	3510	1412	0.71 +/- 0.15
4A	MG1690/pOmpR	LB-Amp	2341	546	
5C	MG1301	LB	135	6.2	1.58 +/- 0.02
5C	MG1908	LB	213	6.6	
5C	MG1305	LB	173	11.6	1.39 +/- 0.13
5C	MG1909	LB	239	7.3	
5C	MG1696	LB	547	132	0.79 +/- 0.17
5C	MG1910	LB	418	4.5	
5C	MG1895	LB	5042	392	1.03 +/- 0.05
5C	MG1911	LB	5195	139	
6A	MG1690	LB	3618	352	0.011 +/- 0.001
6A	MG1810	LB	40.1	7.69	
6A	MG2043	LB	691	111	0.19 +/- 0.05
6A	MG2044	LB	3496	1308	
6A	MG1892	LB	11292	1002	0.04 +/- 0.005
6A	MG2018	LB	462	97.8	
6A	MG2045	LB	923	195	0.08 +/- 0.01
6A	MG2046	LB	2430	1395	
6A	MG1004	LB	22.1	2.53	0.17 +/- 0.05
6A	MG1811	LB	3.75	1.62	
6A	MG2039	LB	8.25	2.04	0.37 +/- 0.05
6A	MG2040	LB	414	56.8	
6A	MG1005	LB	62.8	19.6	18.7 +/- 0.42
6A	MG1812	LB	7.94	2.36	
6A	MG2041	LB	19.9	2.22	0.13 +/- 0.002
6A	MG2042	LB	603	202	
S4B	MG1295	LB	640	101	37.9 +/- 6.40
S4B	MG1807	LB	17.4	4.99	
S4B	MG1863	LB	6010	2753	16.3 +/- 2.66
S4B	MG1891	LB	356	119	
S4B	MG1831	LB	10485	1543	0.93 +/- 0.01
S4B	MG1860	LB	11296	1572	
S4B	MG1296	LB	243	186	91.5 +/- 8.7
S4B	MG1808	LB	2.65	1.97	
S4B	MG1690	LB	3270	496	77.6 +/- 13.7
S4B	MG1810	LB	42.1	0.23	
S4B	MG1833	LB	3340	808	0.64 +/- 0.00
S4B	MG1861	LB	5187	1261	

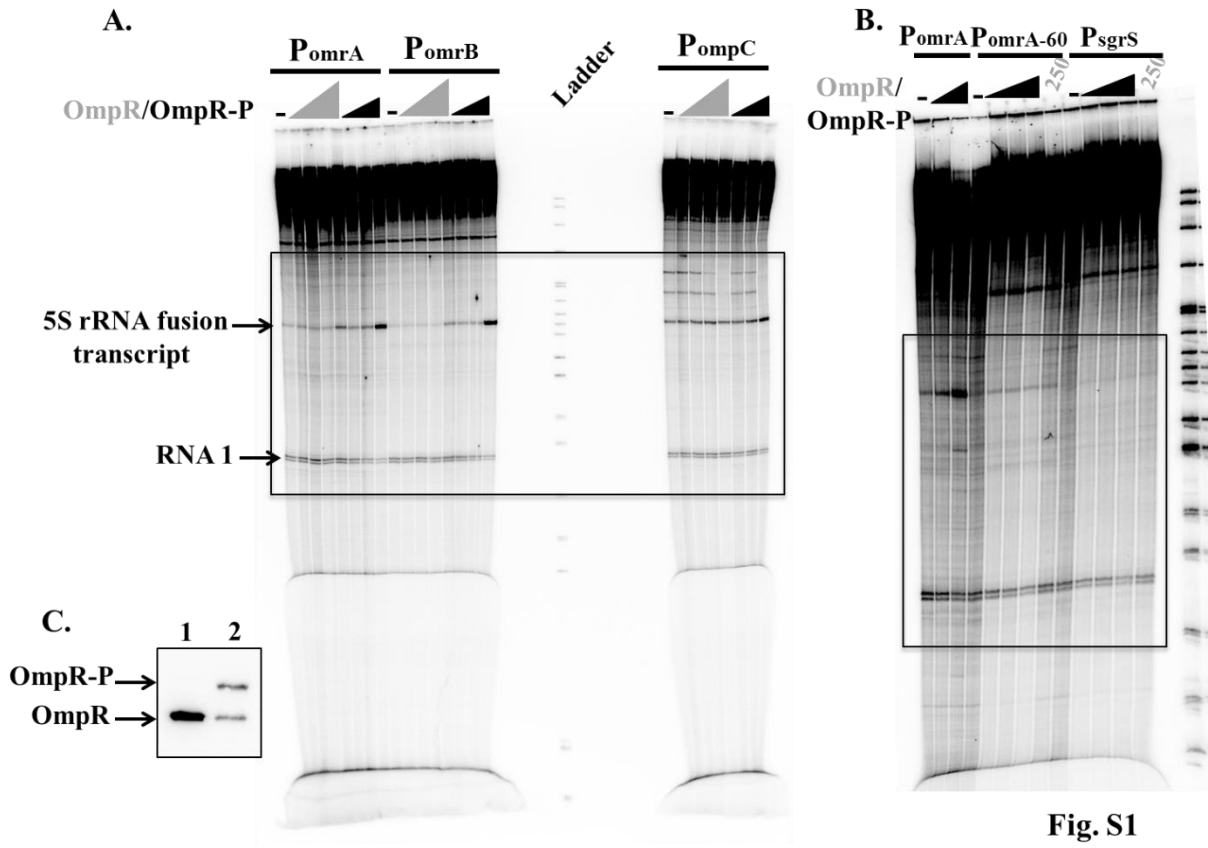


Fig. S1

Figure S1. Direct activation of *omrA* and *omrB* transcription by OmpR-P. (A, B) Shown are the entire sequencing gels whose a portion is presented in Fig. 1 (indicated by the squares). See Fig. 1 for details. (C) A representative *in vitro* phosphorylation reaction of OmpR with acetylphosphate. OmpR and OmpR-P levels were monitored by Western-blot following separation on Phos-Tag containing protein gel before (lane 1) and after (lane 2) incubation of purified OmpR-His with acetylphosphate.

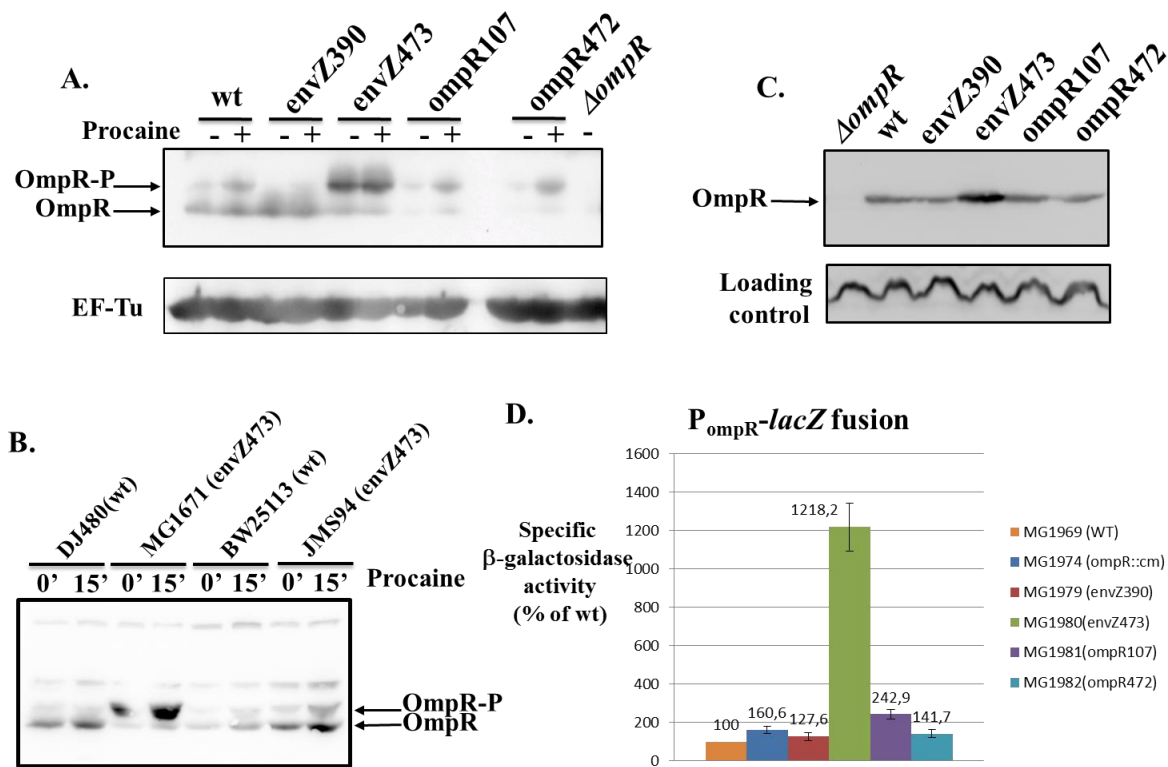
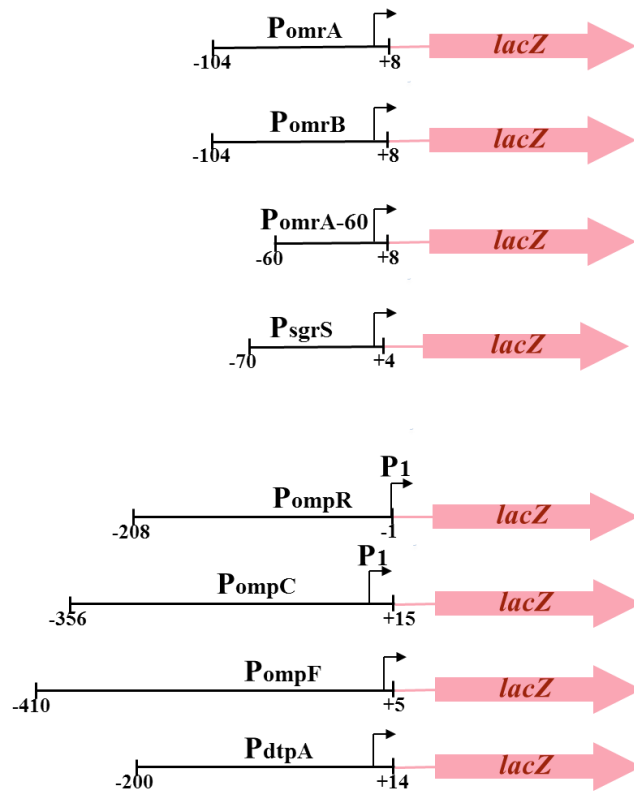


Fig.S2

Figure S2. Effect of *ompR-envZ* mutants on OmpR and OmpR-P levels and *ompR* transcription. (A) An independent repeat of the experiment shown in Fig. 2A is presented. (B) Side-by-side comparison of OmpR phosphorylation in different *envZ473* mutant strains. Samples for Phos-Tag analysis were taken from cultures grown to mid-exponential phase (0.3 OD₆₀₀) and treated or not with 10mM procaine for 15 minutes. Strains used in this experiment are DJ480 and MG1671 (wt and *envZ473* respectively, in our MG1655 background) and BW25113 and JMS94 (as in (Barbieri, Wu et al. 2013), wt and *envZ473* respectively) (C) Same protein extracts as those used in Fig. 2A (in the absence of procaine only) were loaded on a classical SDS-PAGE gel prior to immunodetection with an anti-OmpR antibody. (D) The β-galactosidase activity of an *ompR* transcriptional fusion was measured in LB in strains carrying various mutants of *ompR* or *envZ*. This fusion carries nts -243 to -35 of *ompR* (relative to the ATG start codon) upstream of *lacZ* gene (starting at nt-17).

A. Transcriptional fusions



B. Translational fusions

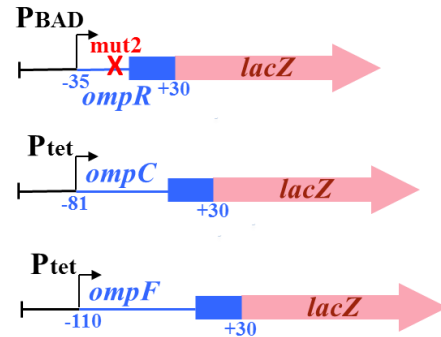


Fig.S3

Figure S3. Schematic depiction of the different fusions used in this study. (A) Transcriptional fusions. Numbering refers to the TSS of the gene whose promoter is fused to *lacZ*. When multiple TSS have been described, numbering is relative to the TSS from the proximal promoter. Pink thin line indicates *lacZ* sequence upstream of *lacZ* ORF, and *lacZ* ORF is shown as a pink thick line. (B) Translational fusions. Numbering now refers to the AUG start codon and the P_{BAD} or P_{tet} promoters drive transcription of an mRNA whose TSS is the same than from *ompR* P1, *ompC* P1 and *PompF* respectively. The red cross indicates position of the *mut2* change in *ompR-lacZ* fusion. Note that fusions were inserted at the *lacZ* locus, except for *PomrA*-, *PomrB*-, *PomrA-60*- and *PsgrS-lacZ* that are at the lambda attachment site (Vanderpool and Gottesman 2004; Guillier and Gottesman 2006).

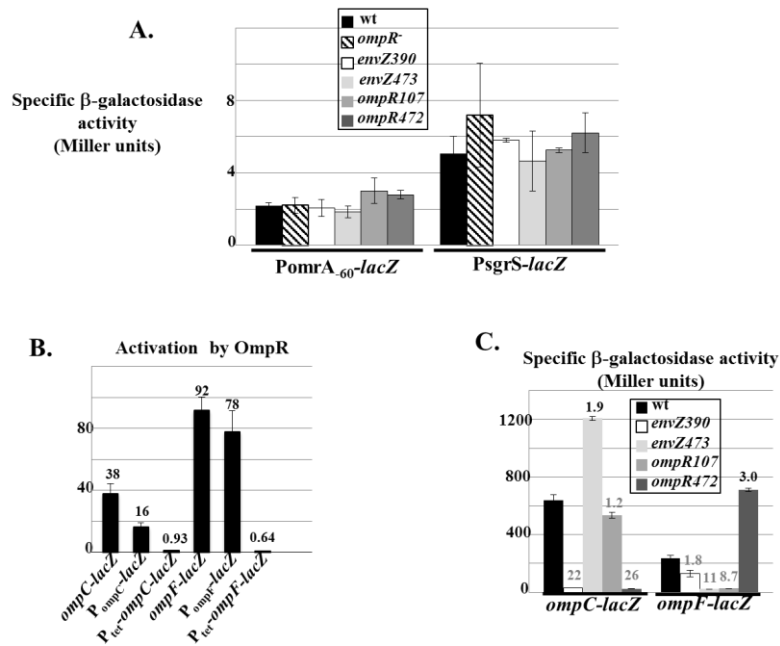


Fig.S4

Figure S4. *ompC* and *ompF* expression is regulated by OmpR primarily at the transcriptional level. (A) The short version of OmpA promoter and the SgrS promoter are not regulated by OmpR. The β -galactosidase activity of transcriptional fusions to either promoter was measured in various *ompR/envZ* mutants. Strains are MG1022, MG2178, MG2182, MG2183, MG2184 and MG2185 (*PomrA*₋₆₀) and CV5200, MG2179, MG2186, MG2187, MG2188 and MG2189 (*PsgrS*). (B) Ratio in β -galactosidase activities in *ompR*⁺ and *ompR*⁻ cells is shown for various *ompC* and *ompF* fusions. *ompC*- and *ompF*-*lacZ* are the fusions that were previously isolated by the lab of T. Silhavy and were transferred to the same strain background as the other fusions; they carry some of *ompC/F* ORF and are sensitive to MicC and MicF respectively (data not shown). *P_{ompC}-* and *P_{ompF}-lacZ* are the leaderless promoter fusions used throughout this study. *P_{tet}-ompC-* and *P_{tet}-ompF-lacZ* are translational fusions expressed from a *P_{tet}* promoter (constitutively expressed in these strains) and allowing production of fusions proteins carrying the first 10 aminoacids of OmpC or OmpF respectively. The activity of the different fusions in the *ompR*⁺ strains was, in Miller units, 640 (*ompC-lacZ*), 6010 (*P_{ompC}-lacZ*), 10485 (*P_{tet}-ompC-lacZ*), 243 (*ompF-lacZ*), 3270 (*P_{ompF}-lacZ*) and 3340 (*P_{tet}-ompF-lacZ*). Strains are MG1295, MG1807, MG1863, MG1891, MG1831, MG1860, MG1296, MG1808, MG1690, MG1810, MG1833 and MG1861. (C) The effect of *ompR-envZ* mutants was assayed on the *ompC-* and *ompF-lacZ* fusions. Strains are MG1295, MG1310, MG1311, MG1312, MG1313, MG1296, MG1314, MG1315, MG1316 and MG1317.

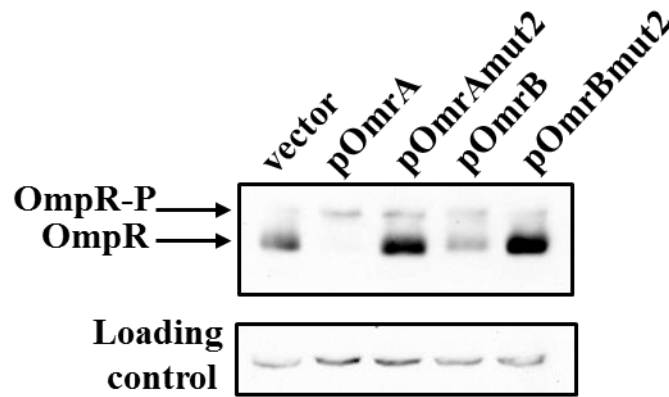


Figure S5. OmrA/B repress *ompR*, but do not affect OmpR-P levels. An independent repeat from the experiment shown in Fig. 3C is presented. See Fig. 3C for details.

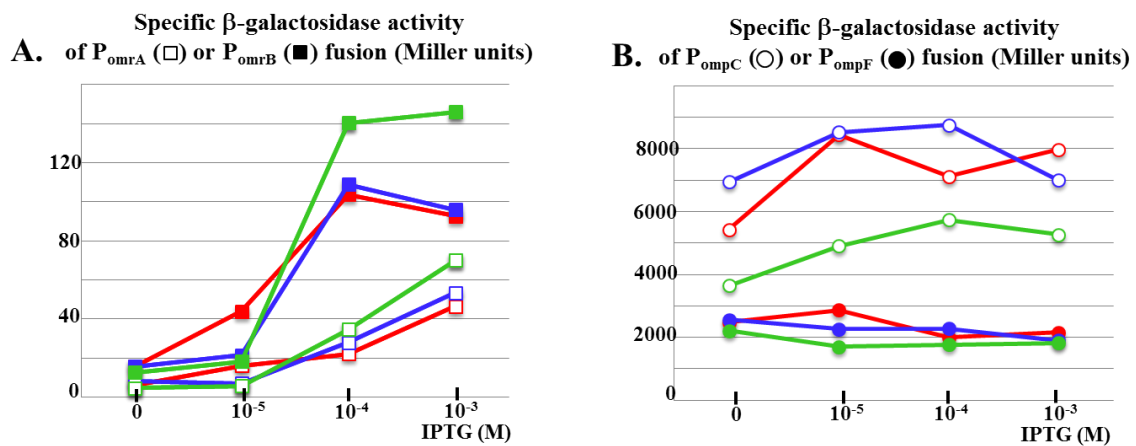


Fig. S6

Figure S6. OmrA/B transcription is much more sensitive to changes in OmpR levels than *ompC/F* transcription. Shown are the b-galactosidase activities measured in independent triplicate repeats of Fig. 4C, with each experiment indicated by a different color. (A) Data obtained with the OmrA or OmrB promoter fusion are indicated by open and closed squares, respectively (panel A) and those obtained with *ompC* or *ompF* promoter fusion by open and closed circles (panel B). Data shown in Fig. 4C are here in blue. See Fig. 4C for details.

3.1.2. Autres régulateurs des Omr

Nos travaux nous laissent penser que les Omr puissent être régulés par d'autres facteurs dépendants ou indépendants d'OmpR. Aussi pour la suite de ce projet j'ai poursuivi ces deux pistes.

3.1.2.1. Partenaires d'OmpR

Nous avons pu montrer dans nos travaux publiés qu'OmpR activait les Omr non seulement via sa forme phosphorylée mais aussi via un mécanisme peu commun faisant intervenir la forme non phosphorylée d'OmpR. En effet, nous avons pu mettre en évidence qu'en présence de faible quantité d'OmpR-P dans la cellule la forme non phosphorylée d'OmpR pouvait parfaitement activer les Omr (figure 6 de l'article). Nous avons évoqué deux hypothèses. La première serait qu'il y ait formation d'hétérodimères entre les quelques molécules d'OmpR-P et les monomères d'OmpR non phosphorylés. La fixation de ces hétérodimères activerait la transcription des Omr. On peut aussi penser que la présence de monomères non phosphorylés favoriserait le recrutement des molécules d'OmpR-P qui activeraient ensuite la transcription des Omr. Dans les deux cas il s'agirait de jouer sur les deux formes d'OmpR pour activer les Omr. Pour tester cette hypothèse j'ai réalisé une expérience de transcription *in vitro* en mélangeant à une quantité fixe d'OmpR-P (non suffisante pour activer fortement la transcription dans les transcription *in vitro* précédentes) des quantités croissantes d'OmpR non phosphorylé (figure 70).

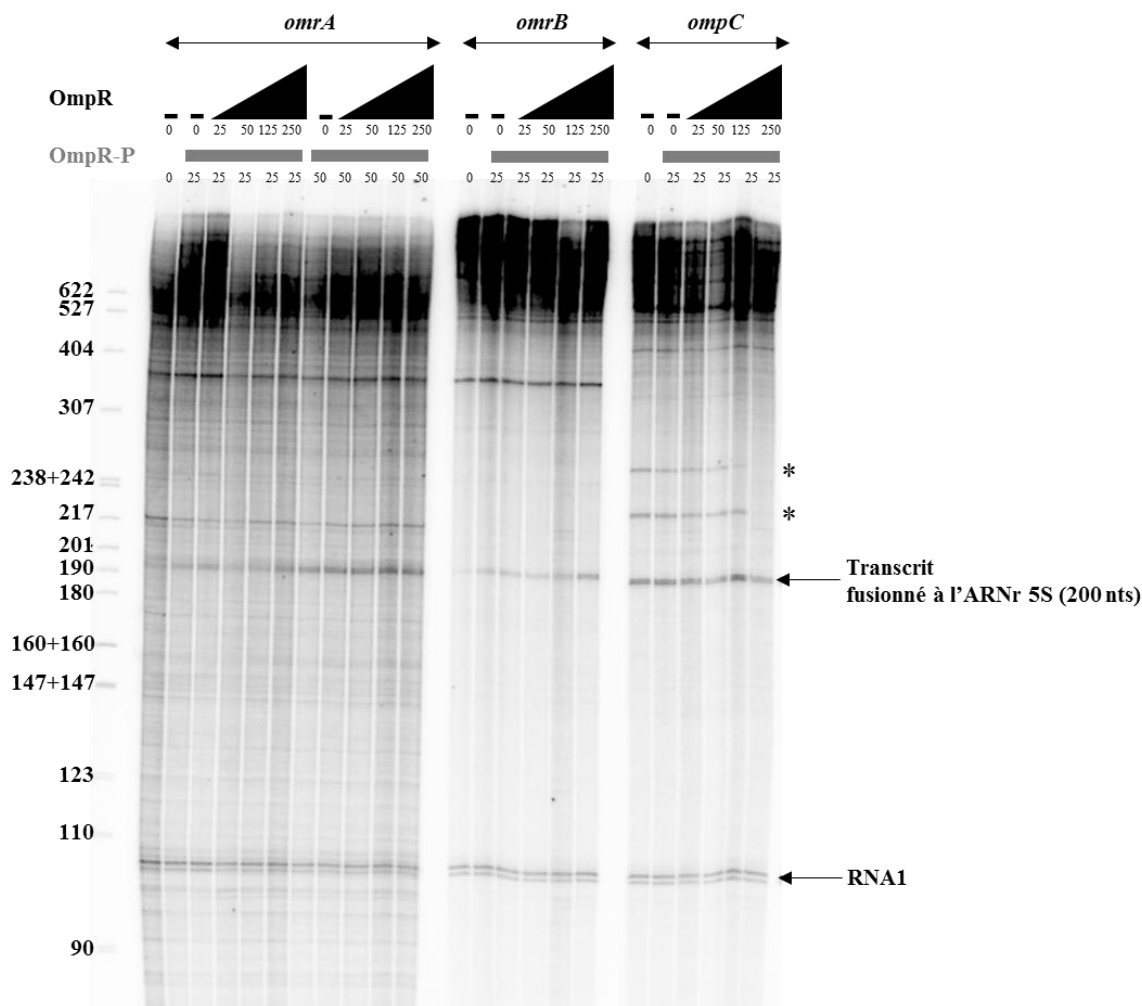


Figure 70 : Activation des *omrA/B* et d'*ompC* par des mélanges OmpR/OmpR-P. Les réactions de transcription *in vitro* ont été réalisées en utilisant les plasmides matrices portant les régions -114+50 soit du promoteur d'*omrA* soit celui d'*omrB* ou la région -437 à +50 d'*ompC* (numérotation relative au +1 de transcription) fusionnées au 5S afin de donner un transcrit de 200 nts. OmpR a été ajoutée dans des ratio molaires de 25, 50, 125 ou 250 par rapport à la matrice ADN et OmpR-P dans un ratio molaire de 25 ou 50. Les produits des réactions ont été séparés sur un gel de séquençage. La portion du gel présentée permet de visualiser les transcrits issus de *PomrA*, *PomrB* et *PompC* ainsi que le contrôle interne RNA1. Le marqueur de taille correspond aux produits de digestion du pBR322 par MspI. Les * correspondent aux produits de transcription issus des deux autres promoteurs d'*ompC* OmpR-indépendants.

Avec cette approche, il est impossible de reconstituer *in vitro* l'activation d'*omrA/B* par la forme non phosphorylée d'OmpR, même en présence d'une faible quantité d'OmpR-P. Je me suis donc penchée sur la seconde hypothèse où nous imaginons que ce serait l'intervention d'un partenaire qui permettrait à la forme non phosphorylée d'OmpR d'activer les Omr. Si nous gardons à l'esprit qu'un minimum de forme phosphorylée d'OmpR est peut-être nécessaire pour activer les Omr, ces hétérodimères pourraient être un moyen de positionner les molécules d'OmpR-P uniquement dans les positions stratégiques pour assurer l'activation

transcriptionnelle, les autres sites seraient requis plutôt pour favoriser l'interaction entre OmpR-P et la polymérase.

Dans l'article de Guillier *et al.* en 2006, il est fait mention de trois autres plasmides activateurs des Omr dépendants d'OmpR (Guillier and Gottesman 2006), qui sont donc des candidats potentiels pour être des partenaires d'OmpR (figure 71).

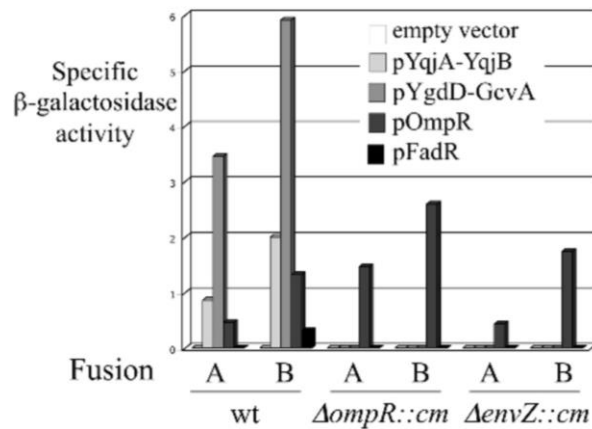


Figure 71 : Différents activateurs en multicopie des Omr agissant via EnvZ-OmpR. Issue de (Guillier and Gottesman 2006).

Le plasmide pYqjA-YqjB porte deux protéines de membrane interne dont YqjB renommé ultérieurement MzrA et décrit pour activer le système EnvZ-OmpR via EnvZ uniquement (Gerken, Charlson et al. 2009). Je ne me suis donc pas intéressée à ce plasmide car l'effet est indirect. Le second plasmide porte le régulateur FadR impliqué dans la régulation du métabolisme des acides gras. Enfin, le troisième plasmide porte les gènes YgdD et GcvA codant respectivement une protéine de membrane interne et un activateur transcriptionnel impliqué dans la régulation du métabolisme de la glycine. Une étude de délétion effectuée sur ce dernier plasmide a pu montrer qu'il s'agissait de la portion contenant YgdD qui était sans doute responsable de l'effet.

L'effet sur les Omr de ces deux derniers plasmides suggérait qu'YgdD et FadR pouvaient être des partenaires directs d'OmpR. Afin de déterminer si ces gènes sont requis pour la réponse des Omr à la forme non phosphorylée d'OmpR, nous avons mesuré l'activation par le pOmpR des fusions P_{omrA} -lacZ et P_{omrB} -lacZ dans un contexte délété pour *fadR* ou pour *gcvA-ygdD* (figure 72).

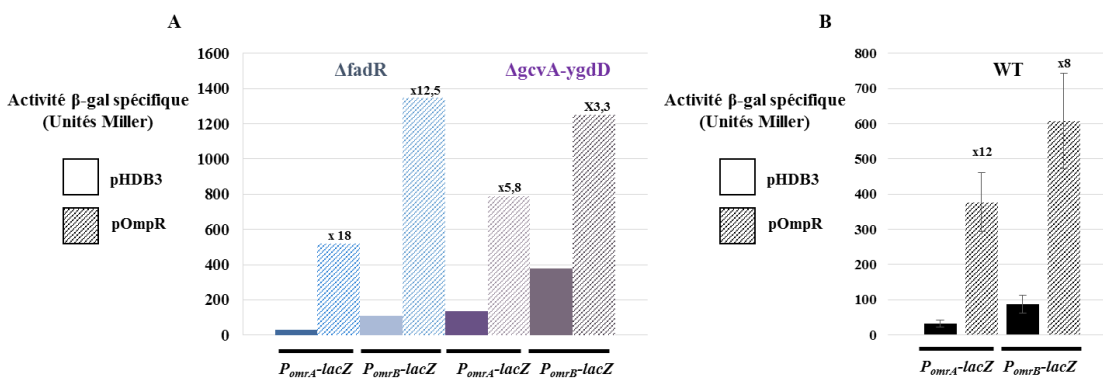


Figure 72 : Activation des Omr par OmpR dans un contexte Δ *fadR* ou Δ *gcvA-ygdD* (A) en comparaison du contexte sauvage (B). A. Les activités sont mesurées en phase exponentielle en milieu LB additionné d'ampicilline à 150 μ g/ml à 37°C. Les souches utilisées sont MG2192 (P_{omrA} -lacZ ; Δ *fadR*), MG2193 (P_{omrB} -lacZ ; Δ *fadR*), MG2194 (P_{omrA} -lacZ ; Δ *gcvA-ygdD*), et MG2195 (P_{omrB} -lacZ ; Δ *gcvA-ygdD*) transformées par les plasmides pHDB3 (vecteur vide) ou pOmpR. B. Les activités des fusions P_{omrA} -lacZ (MG1004) et P_{omrB} -lacZ (MG1005) dans un contexte sauvage et transformées par le pHDB3 ou le pOmpR correspondent à celles de la figure 4A de l'article que nous avons publié (Brosse, Korobeinikova et al. 2016). Les facteurs d'activation entre le pHDB3 et le pOmpR sont mentionnés systématiquement au dessus de la souche portant le pOmpR.

Premièrement, on remarque qu'il n'est pas nécessaire que FadR ou qu'YgdD soit présent pour avoir un effet du pOmpR sur les fusions au promoteur des *omr*. Cependant, on peut noter que bien que la délétion de *fadR* ne module pas le niveau de base des fusions elle semble en revanche augmenter l'activation des Omr par OmpR. La délétion *gcvA-ygdD* augmente quant à elle le niveau de base des fusions au promoteur des *omr*. Cette modification a un impact direct sur le niveau d'activation qui dans ce cas diminue. En effet, alors que l'activation par OmpR dans un contexte sauvage est autour de 12 pour la fusion P_{omrA} -lacZ et 7.5 pour la fusion P_{omrB} -lacZ, dans ce mutant ce facteur est réduit à 5.7 pour la fusion P_{omrA} -lacZ et à 3.3 pour la fusion P_{omrB} -lacZ. Ces premiers résultats sont à manier avec prudence car ils devront être reproduits en intégrant également les fusions dans le contexte sauvage afin de s'assurer que les facteurs d'activation sont bien significativement différents. Cependant, on peut d'ores et déjà suggérer que FadR n'est a priori pas le partenaire recherché car il n'agit pas directement sur le niveau des fusions au promoteur des *omr*. Il pourrait cependant contribuer de manière indirecte à la régulation par OmpR. D'autre part, YgdD n'étant pas connu pour être un régulateur transcriptionnel, l'effet pourrait également être indirect. On peut imaginer qu'YgdD contrôle le niveau d'un répresseur se fixant sur le promoteur des *omr* remplacé par OmpR lors de l'induction de celui-ci conduisant à l'activation importante des Omr. Dans le mutant d'*ygdD* le répresseur ne serait plus présent et donc laisserait la possibilité aux molécules d'OmpR (non phosphorylé notamment) présentes dans la cellule à un niveau basal d'activer le promoteur des

omr, ces derniers étant très sensibles au niveau d'OmpR. Pour tester cette hypothèse et vérifier que l'effet est dépendant d'OmpR nous pourrions refaire ce même type d'expérience dans un contexte $\Delta ompR$.

3.1.2.2. Régulateurs indépendants d'OmpR

Nous avons vu qu'OmpR jouait un rôle prépondérant dans le contrôle des Omr mais il est possible que d'autres régulateurs ciblent le promoteur des *omr*. En particulier, nous avons observé de manière reproductible que dans un mutant d'OmpR, il restait une légère expression des Omr en milieu minimum que nous n'observons pas en milieu riche de type LB (figure 73). Ceci suggère l'existence d'un activateur de ces petits ARN en milieu minimum ou d'un répresseur en milieu riche.

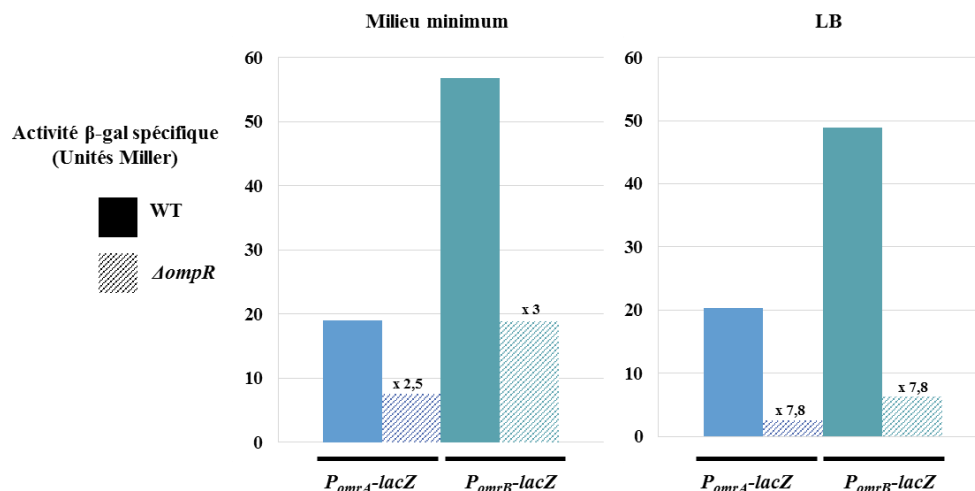


Figure 73 : Expression des Omr en milieu minimum ou en LB en l'absence d'*ompR*. Les activités sont mesurées en phase exponentielle en milieu LB ou en milieu minimum glucose à 37°C. Les souches utilisées sont MG1004 ($P_{omrA-lacZ}$), MG1811 ($P_{omrA-lacZ}$; $\Delta ompR$), MG1005 ($P_{omrB-lacZ}$) et MG1812 ($P_{omrB-lacZ}$; $\Delta ompR$). Les facteurs de répression sont systématiquement mentionnés au dessus de la souche $\Delta ompR$.

Afin de trouver un activateur des Omr indépendant d'OmpR nous avons alors choisi de réaliser un crible génétique avec comme rapporteur la fusion $P_{omrB-lacZ}$ dans un contexte $\Delta ompR$. Cette souche a été transformée avec une banque de plasmides portant des fragments de l'ordre de 1 à 5 kb et couvrant l'ensemble du génome d'*E. coli* (Ulbrandt, Newitt et al. 1997). Le crible a été effectué sur milieu minimum additionné de Xgal afin de récupérer des clones activateurs (plus bleus). Sur environ 7700 clones obtenus après transformation, bleu très clair pour la plupart, nous avons isolé 48 clones bleu foncé. Parmi ceux-ci, 14 le sont restés après purification et retransformation, leur insert a alors été séquencé (tableau 13).

Tableau 13 : Résultats du crible génétique.

Clone	Commentaires
K2/K3/M1/N2/H2 *	' <i>pck-envZ-ompR-greB</i>
H1	' <i>envZ-ompR</i>
N1	' <i>envZ-ompR</i>
G1/O4 *	' <i>envZompR-greB-yhgF</i> '
D1	' <i>mdtH-rimJ-yceH</i> '
C3/M2	' <i>insH1-slp-dctR-yhiD</i> '
J2	' <i>folM-ydgC-rstA-rstB-tus</i> '
O1	? ' <i>rstB-tus</i> ' **

* la séquence de ces clones est strictement identique

** la séquence complète de l'insert n'a pu être obtenue suite à un réarrangement du plasmide en revanche une des bornes est identique au plasmide J2

9 plasmides portaient la totalité du gène *ompR*. Sur ces 9 clones, 5 étaient identiques suggérant la saturation du crible. Pour les plasmides restants nous avons choisi de nous focaliser sur celui ayant le phénotype le plus stable que l'on nommera par la suite pRstAB car il porte notamment les gènes du système à deux composants RstB-RstA. Ce plasmide porte précisément et dans cet ordre une portion de *folM* (codant une réductase), *ydgC* (codant une hypothétique protéine), *rstA* (régulateur transcriptionnel), *rstB* (senseur du système RstB-RstA) et enfin une portion de *tus* (codant une protéine impliquée dans l'inhibition de la réplication). Le système à deux composants RstB-RstA est impliqué dans la réponse à l'acidité. En effet, il active notamment la transcription du gène *asr* (acid shock protein) (Ogasawara, Hasegawa et al. 2007). RstA réprime également la transcription de *csgD*. La transcription du gène *rstA* est également sous le contrôle du système à deux composants PhoQ-PhoP (Minagawa, Ogasawara et al. 2003) qui de son côté perçoit les variations de concentration en Mg^{2+} (Garcia Vescovi, Soncini et al. 1996).

Après transformation dans les souches portant les fusions P_{omrA} -lacZ ou P_{omrB} -lacZ dans un contexte sauvage ou $\Delta ompR$, nous avons pu déterminer que le pRstAB activait de manière spécifique *OmrB* (figure 74). De plus, l'activation est bien indépendante d'OmpR puisque l'activation par le pRstAB de la fusion au promoteur d'*omrB* dans une souche délétée pour *OmpR* est même supérieure à celle détectée dans une souche où *OmpR* est présent dans les conditions testées. On peut imaginer que l'effet soit amplifié dans la souche $\Delta ompR$ du fait de l'absence de ce régulateur majeur des *Omr*. De plus, l'activation par le pRstAB est bien plus modeste que celle d'OmpR suggérant que ce régulateur d'*OmrB* serait sans doute secondaire. Néanmoins, ce résultat est vraiment très intéressant et prometteur car à ce jour, il s'agirait d'une autre condition où les *Omr* pourraient avoir une expression différentielle. En effet, il a été

proposé que RpoS régule spécifiquement OmrA (Levi-Meyrueis, Monteil et al. 2014; Peano, Wolf et al. 2015). Ici nous montrons à l'inverse que RstB-RstA, pourrait réguler spécifiquement OmrB en milieu minimum. Des données préliminaires nous montraient en effet, que le plasmide pRstAB n'avait aucun effet sur la fusion au promoteur d'*omrB* en LB. Ces données pourraient expliquer que les deux Omr aient été conservés au cours de l'évolution chez *E. coli* notamment.

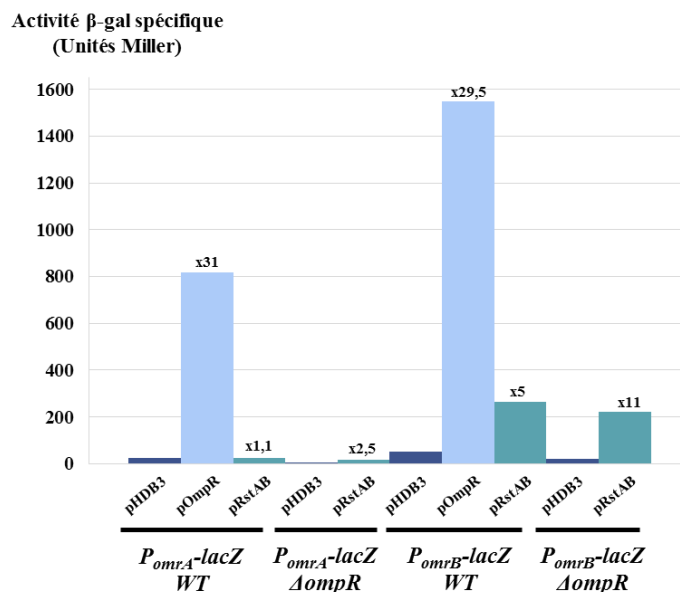


Figure 74 : Effets du plasmide pRstAB sur le promoteur d'*omrA* ou d'*omrB* dans un contexte sauvage ou mutant pour *ompR*. Les activités sont mesurées en phase exponentielle en milieu minimum glucose additionné d'ampicilline à 150µg/ml à 37°C. Les souches utilisées sont MG1004 (*P_{omrA}-lacZ*), MG1811 (*P_{omrA}-lacZ* ; Δ*ompR*), MG1005 (*P_{omrB}-lacZ*) et MG1812 (*P_{omrB}-lacZ* ; Δ*ompR*) transformées par les plasmides pHDB3 (vecteur vide), pOmpR ou pRstAB. Les facteurs d'activation des plasmides pOmpR et pRstAB sont indiqués systématiquement au dessus de la souche transformée par le plasmide en question.

Par ailleurs, cette expérience nous montre également que l'effet du pOmpR est bien plus important en milieu minimum qu'en LB. En effet, les facteurs d'activation en milieu minimum sont autour de 30 alors qu'en LB, il s'agit de facteurs d'activation entre 8 et 12 (figure 72). De plus, le niveau de base des fusions est plus faible. On peut donc imaginer qu'en milieu minimum sans stress il y ait moins d'OmpR-P dans la cellule ce qui conduirait à une activité de base des Omr plus faible.

Afin de tester si l'effet de la forme non phosphorylée d'OmpR serait dépendant de RstB-RstA nous avons comparé le niveau des fusions *P_{omrA}-lacZ* et *P_{omrB}-lacZ* dans un contexte sauvage ou mutant pour *rstAB* en présence ou en absence du plasmide surproduisant OmpR (figure 75). En

effet, le fait d'avoir trouvé ce régulateur dans une souche $\Delta ompR$ n'exclut pas que qu'OmpR puisse agir en partenariat avec RstB-RstA.

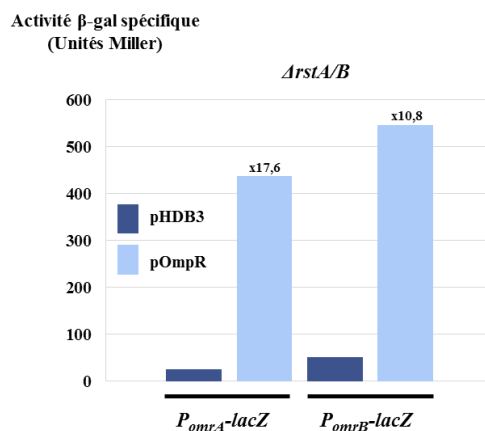


Figure 75 : Effets du plasmide pOmpR sur le promoteur d'*omrA* ou d'*omrB* dans un contexte sauvage ou mutant pour *rstAB*. Les activités sont mesurées en phase exponentielle en milieu LB additionné d'ampicilline à 150µg/ml à 37°C. Les souches utilisées sont MG2196 ($P_{omrA-lacZ}$; $\Delta rstAB$), et MG2197 ($P_{omrB-lacZ}$; $\Delta rstAB$) transformées par les plasmides pHDB3 (vecteur vide) ou le pOmpR. Les facteurs d'activation par le pOmpR sont indiqués au dessus de la souche portant ce plasmide.

La fusion $P_{omrA-lacZ}$ est activée d'un facteur 17.6 par le plasmide pOmpR dans le contexte $\Delta rstAB$ ce qui est un peu supérieur à celui obtenu dans le contexte sauvage (figure 72). Par ailleurs, la fusion $P_{omrB-lacZ}$ est activée quant à elle d'un facteur 10.8 par le plasmide pOmpR dans le contexte $\Delta rstAB$ alors qu'elle l'était d'un facteur 7.5 sans la délétion de *rstAB*. Ces résultats montrent que l'effet du pOmpR est conservé dans une souche $\Delta rstAB$ et donc que la régulation qu'il exerce est sans doute indépendante de RstB/A. Par ailleurs, on peut imaginer que l'augmentation des facteurs d'activation soit due au fait que l'absence de ce second régulateur laisse la possibilité à plus de molécules d'OmpR de se fixer sur le promoteur des *omr*.

3.2. La régulation des Omr par OmpR chez la bactérie pathogène *Escherichia coli* LF82

3.2.1. Analyse des données de RNAseq

Dans le but d'étudier le profil d'expression de l'ensemble des gènes de la souche LF82 au cours de l'infection des macrophages, Gaëlle Demarre dans le groupe d'Olivier Espeli au Collège de France a réalisé un séquençage haut débit des ARN totaux (RNAseq) de bactéries internalisées dans des macrophages humains, 1h, ou 6h après infection. Dans ces extraits ont retrouvé ainsi l'ARN des bactéries et des macrophages ce qui est une différence majeure avec de nombreuses autres analyses de ce type. Les ARN totaux ont aussi été extraits de bactéries cultivées pendant 1h ou 6h dans le milieu de culture des macrophages afin de servir de contrôle.

Les données collectées sont exprimées comme un nombre de lectures par gène et permettent d'estimer l'abondance des ARN dans les différentes conditions. Dans cette étude, le nombre de lectures a été normalisé en fonction du nombre de lectures d'un gène par rapport au nombre total de lectures s'alignant sur la totalité du génome bactérien. Après normalisation, les ratios entre le nombre de lecture d'un gène après 1h ou 6h d'infection et le nombre de lecture de ce même gène dans les bactéries cultivées pendant 1h seules ont été analysés et représentent un facteur d'induction. Ce choix a été fait car les bactéries à 1h et à 6h dans les macrophages se trouvent plutôt en phase exponentielle de croissance tout comme à 1h dans le milieu de culture. Au contraire, au bout de 6h dans le milieu de culture les bactéries sont en phase stationnaire.

De façon intéressante en réalisant cette analyse, Gaëlle Demarre et Olivier Espeli ont observé que nombre de petits ARN dont OmrA et OmrB étaient induits lors de la croissance dans les macrophages (figure 76). Les Omr font d'ailleurs partie des ARN régulateurs les plus induits. Un certain nombre d'ARN régulateurs sont également réprimés et quelques uns tels que CsrC (titrateur de CsrA) ont un comportement différent entre 1h et 6h dans les macrophages.

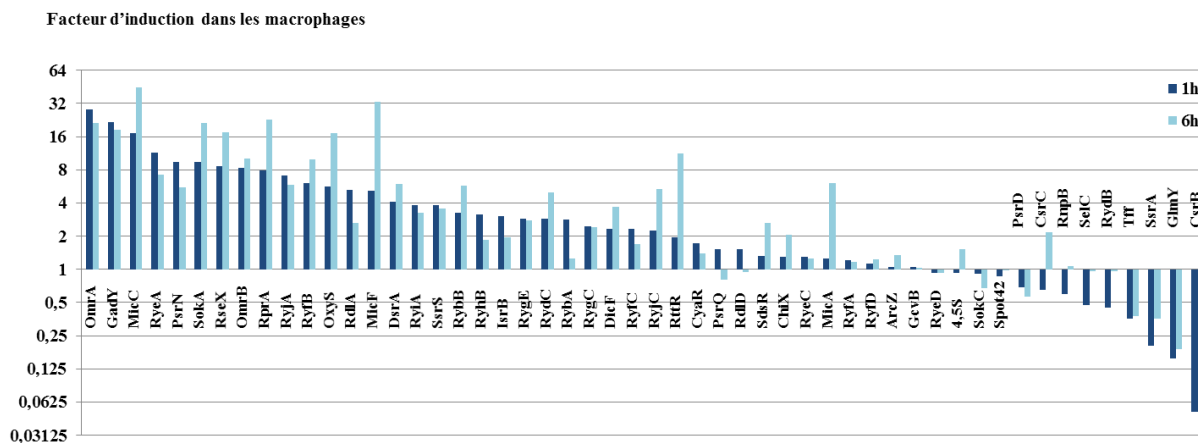


Figure 76 : Profil d'expression des petits ARN régulateurs de LF82 au cours de l'infection. Les facteurs d'induction sont exprimés pour 1h ou 6h dans des macrophages. (A partir des données de Gaëlle Demarre et Olivier Espeli ; Communication personnelle)

Les facteurs d'induction pour les petits ARN peuvent être relativement élevés atteignant un facteur 44 pour MicC et 32 pour MicF. Dans un certain nombre de cas, l'induction est plus importante à 6h qu'à 1h suggérant que des événements secondaires participent à l'induction. Pour OmrA et OmrB, les ARN sur lesquels je me suis focalisée, ils sont induits respectivement 28 fois et 8 fois à 1h post-infection et 21 fois et 10 fois à 6h post-infection. L'induction d'OmrA, OmrB et MicF suggère l'implication du système à deux composants EnvZ-OmpR ce que j'ai vérifié par la suite. Néanmoins, de façon étonnante MicC est également induit alors qu'on s'attendrait à une répression médiée par OmpR. Les petits ARN connus pour être induits en réponse à divers stress tels que GadY, RprA ou RybB qui répondent respectivement à sigma S, au système Rcs ou à sigma E sont également induits ici ce qui est cohérent avec l'internalisation de la bactérie dans la vacuole des macrophages. Cependant, il doit y avoir des subtilités car tous les petits ARN qui répondent à sigma E ou d'autres facteurs de stress ne sont pas systématiquement induits.

Les facteurs d'induction observés sont relativement importants pour des petits ARN mais il est intéressant de mettre ces facteurs en regard de ceux obtenu pour des messagers. En effet, les facteurs peuvent atteindre plus de 10 000 fois par exemple pour le messager d'*ydeO*. La figure 77 présente les facteurs d'induction parmi les plus importants obtenus pour certains des messagers dans cette expérience (*ydeO*, *emrK* et *frc* codant respectivement un régulateur transcriptionnel impliqué dans la survie en condition de croissance anaérobie ou acide, une pompe à efflux et enfin une enzyme requise lors de l'adaptation au stress acide) ainsi que ceux

de certains membres du régulon d'OmpR (les messagers *ompC*, *ompF*, *csgD*, *flhD*, *bolA*, *tpxB* et les petits ARN OmrA et OmrB). On peut noter que les plus forts facteurs correspondent aux messagers de gènes activés par le système à deux composants EvgS-EvgA répondant au stress acide, ce qui était aussi attendu du fait que LF82 se localise dans la vacuole des macrophages.

Facteur d'induction dans les macrophages

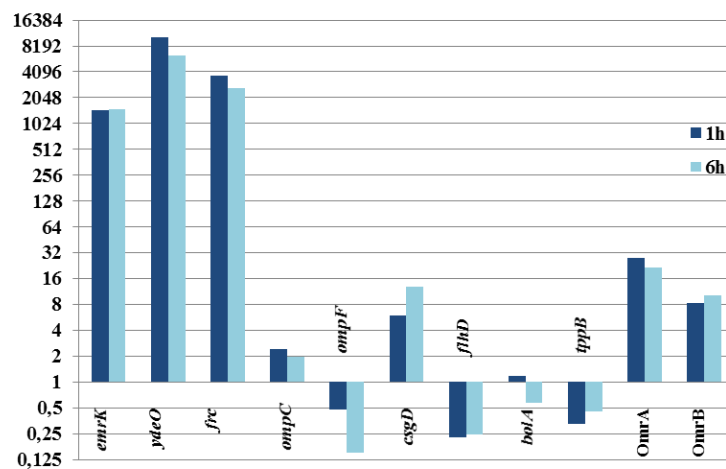


Figure 77 : Profil d'expression des messagers les plus exprimés dans le séquençage et de certains gènes issus du régulon OmpR chez la bactérie LF82 au cours de l'infection. Les facteurs d'induction sont exprimés pour 1h ou 6h dans des macrophages. (D'après les données de Gaëlle Demarre et Olivier Espeli ; Communication personnelle).

Au vu de ces données de séquençage, ma contribution à ce projet a alors été de confirmer les données globales sur les Omr en particulier par une méthode plus directe qu'est le Northern-blot et de s'intéresser au rôle d'OmpR dans l'induction des Omr.

3.2.2. Contrôle des Omr par OmpR chez LF82 au cours de l'infection

Après avoir construit par recombinaison une souche de LF82 délétée pour *ompR* ($\Delta ompR$), nous avons pu réaliser une expérience d'infection des macrophages par LF82 sauvage ou $\Delta ompR$ suivie d'une extraction des ARN totaux à 1h/3h/6h/24h post-infection. En contrôle nous avons utilisé une culture de 3h dans le milieu de culture des macrophages. Etant donné que la quantité d'ARN bactérien total obtenu pour les échantillons post-infection est très faible par rapport à une culture bactérienne seule nous avons choisi de réaliser une gamme de dilution des ARN bactériens seuls et utilisé une sonde pour l'ARNr 5S comme contrôle pour l'analyse par Northern Blot (figure 78). Cette stratégie nous permet de quantifier l'induction des Omr au cours de l'infection.

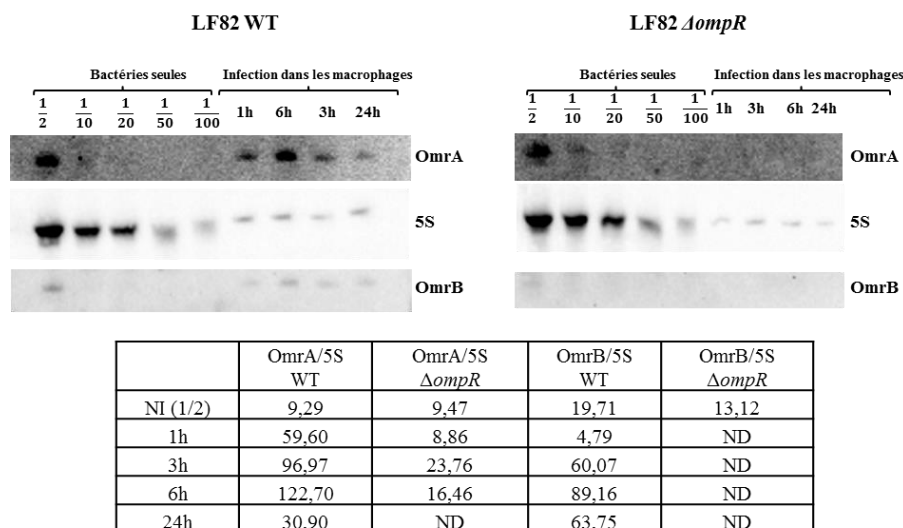


Figure 78 : Induction des Omr au cours de l'infection chez LF82. Les Northern Blot (partie haute de la figure) montrent clairement que la délétion d'OmpR inhibe fortement l'expression des Omr au cours de l'infection. Une quantification de ces résultats a été réalisée avec le logiciel Image Quant et est présentée dans le tableau. Le choix a été fait de prendre comme contrôle non induit la dilution au 1/2 des bactéries seules. ND : l'intensité de la bande est trop faible pour être quantifiée.

Ces résultats confirment tout d'abord qu'au cours de l'infection, OmrA et OmrB sont bien induits et présentent un pic à 6h. Tout comme chez *E. coli* K12 OmrA est mieux exprimé qu'OmrB. De plus ces données nous permettent de voir que l'induction des Omr est totalement dépendante de la présence du régulateur OmpR ce qui suggère non seulement que la régulation des Omr par OmpR puisse être similaire à *E. coli* K12 en condition de stress mais aussi qu'OmpR est activé lors de l'infection.

En revanche, la délétion d'*ompR* dans la souche LF82 ayant poussé dans le milieu de culture des macrophages n'affecte que très peu le niveau d'OmrA mais plus fortement le niveau d'OmrB. Au contraire, chez *E. coli* K12, la délétion d'*ompR* limite très fortement aussi bien le niveau d'OmrA que d'OmrB (voir partie 1.3.3.2.). Cette observation est très intéressante et reproductible en culture LB (figure 79). Elle suggère qu'il existe quelques différences de régulation d'OmrA par OmpR entre ces deux souches d'*E. coli* et notamment qu'il puisse y avoir un contrôle différent dans un milieu de culture synthétique de laboratoire. On peut imaginer la présence d'un autre régulateur chez l'une ou l'autre des souches qui serait activateur chez LF82 ou répresseur chez K12. Le promoteur d'OmrA des deux souches ne diffère que d'un seul nucléotide (C pour K12 et T pour LF82 en position -11 par rapport au +1 d'OmrA-

voir figure 54) et il serait intéressant de vérifier si cette mutation a un effet direct sur la régulation.

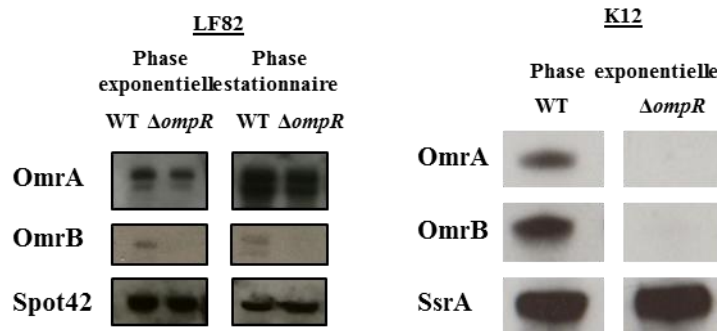


Figure 79 : Expression d'OmrA et d'OmrB en phase exponentielle ou en phase stationnaire chez LF82 ou K12. Les souches sauvages (WT) et délétée pour *ompR* ($\Delta ompR$) ont été cultivées en LB à 37°C et des extraits ARN ont été préparés et analysés par Northern Blot. Panel droit : résultat antérieur du laboratoire.

Par ailleurs, des expériences de survie dans les macrophages réalisées par Gaëlle Demarre au Collège de France montrent que le mutant $\Delta ompR$, est très affecté dans la survie à l'intérieur des macrophages suggérant qu'OmpR ai un rôle prépondérant au cours de l'infection. Il serait donc intéressant de savoir si les Omr ont eux même un rôle direct dans la virulence. Un article de 2011 démontre qu'un mutant de la chaperone Hfq est très affecté dans ses capacités d'invasion de l'épithélium intestinal du nématode *C. elegans* (Simonsen, Nielsen et al. 2011). Cette donnée suggère que les petits ARN, et potentiellement les Omr, pourraient avoir un rôle direct dans le processus infectieux.

3.3. NarQ-NarP, un autre système régulé par des petits ARN Hfq-dépendants

J'ai choisi de présenter dans ce manuscrit une autre partie de mon projet de thèse qui est encore en cours. Certaines expériences n'ont été réalisées qu'une fois d'où l'absence de barres d'erreurs sur certains graphiques.

3.3.1. Identification des petits ARN régulant l'expression de *narP*

Afin d'avoir une vue générale de la régulation des systèmes à deux composants par des petits ARN Hfq-dépendants, nous avons choisi de nous focaliser sur neuf systèmes à deux composants (tableau 14). Ces systèmes ont tout d'abord été sélectionnés pour la co-précipitation de leur messenger avec Hfq (Zhang, Schu et al. 2013). En effet, Hfq est une chaperone connue pour stabiliser les interactions ARN régulateur/ARNm. Ainsi, si le messenger co-précipite avec Hfq, il y a de fortes chances pour qu'il soit régulé par un petit ARN. Cependant, nous savons que les ARNm codant des régulateurs de systèmes à deux composants peuvent être exprimés à des niveaux faibles en absence de stress rendant leur détection dans une approche globale délicate. Ce critère n'est donc pas suffisant pour estimer l'étendue de la régulation des systèmes à deux composants par des petits ARN.

Nous avons également décidé d'ajouter comme critère supplémentaire l'impact de leur délétion dans une analyse transcriptomique (Oshima, Aiba et al. 2002). En effet, une étude de 2012 a pu montrer que la délétion de ce type de système pouvait moduler l'expression de nombreux gènes. L'idée de ce critère est de sélectionner des systèmes clé pour la physiologie cellulaire notamment en condition de stress. Ainsi, en condition de stress agir directement et rapidement sur l'expression du régulateur via un petit ARN pourrait permettre de contrôler simultanément plusieurs voies ce qui *in fine* permettrait à l'organisme de s'adapter au mieux aux nouvelles conditions environnementales. Pour finir, nous avons concentré nos efforts sur des systèmes encore peu connus afin d'élargir le champ d'étude des circuits mixtes de régulation.

Tableau 14 : Systèmes à deux composants choisis pour cette étude. A titre de comparaison nous avons mentionné les facteurs obtenus avec le système EnvZ-OmpR.

Système à deux composants	Nombre de gènes surexprimés (Oshima <i>et al.</i> 2002)	Nombre de gènes réprimés (Oshima <i>et al.</i> 2002)	Facteur d'enrichissement avec Hfq (Zhang 2013)		Fonction
			2008	2009	
BaeS-BaeR	14	8	0.7	2.4	Réponse au stress de membrane
BasS-BasR	20	12	-	-	Réponse au fer
EvgS-EvgA	2	3	0.2	5/0.3	Résistance aux antibiotiques et stress acide
KdpD-KdpE	10	10	0.3	0.5	Transport du potassium
NarQ-NarP	NarQ 20 NarP 3	NarQ 25 NarP 9	38	32.9	Métabolisme des nitrates
PhoR-PhoB	1	8	-	-	Métabolisme du phosphate
RstB-RstA	12	22	-	-	Réponse stress acide
BarA-UvrY	BarA 2 UvrY 86	BarA 24 UvrY 37	-	-	Métabolisme global
YedV-YedW	6	7	6.7	-	Réponse au stress oxydatif
EnvZ-OmpR	71	54	4	2.7	Réponse au stress osmotique et stress acide

Le système BaeS-BaeR est connu pour répondre au stress de membrane et de façon intéressante il a été montré que des interconnexions existaient entre ce système et les voies de signalisation Cpx, Psp, Rcs et sigma E afin de maintenir l'homéostasie de l'enveloppe dans des conditions de stress diverses (Bury-Mone, Nomane et al. 2009).

Le système BasS-BasR est quant à lui impliqué dans la réponse au fer et au zinc notamment en régulant différents gènes dont les fonctions sont associées à la structure et à la fonction de la membrane (Ogasawara, Shinohara et al. 2012).

Peu de données sont connues sur le système EvgS-EvgA mais on sait qu'il est impliqué dans la résistance au stress acide et aux antibiotiques. De plus, via la protéine b1500, qui interagit avec la kinase PhoQ, le système à deux composants PhoQ-PhoP est activé en réponse à la forme phosphorylée d'EvgA (Eguchi, Itou et al. 2007).

Ensuite, le système KdpE-KdpD est associé au transport du potassium et de façon intéressante il a été montré qu'en l'absence de la kinase KdpD, KdpE pouvait être phosphorylé par PhoB un autre des régulateurs que nous avons choisi impliqué pour sa part dans le métabolisme du phosphate (Schramke, Laermann et al. 2017).

En ce qui concerne RstB-RstA, son expression est régulée par le système à deux composants PhoQ-PhoP et cible différents gènes nécessaires à l'adaptation au stress acide notamment le gène *asr* (Ogasawara, Hasegawa et al. 2007).

BarA-UvrY est connu pour activer la transcription des petits ARN CsrB et CsrC, qui eux même régulent l'activité du facteur global CsrA (Chavez, Alvarez et al. 2010). Il active également la transcription de *rpoS*, le gène codant le facteur alternatif sigma S (Mukhopadhyay, Audia et al. 2000). Il semblerait donc que ce système à deux composants ait un effet global allant de la modulation du métabolisme carboné à celle de la motilité et de la formation des biofilm en passant par la virulence chez les souches d'*E. coli* uropathogènes (Teplitski, Goodier et al. 2003). La délétion de ce système affecte d'ailleurs de nombreuses cibles. On peut aussi remarquer que ce système à deux composants n'est pas codé en opéron et que la délétion d'*uvrY* affecte un plus grand nombre de cibles que celle de *barA*. On peut donc imaginer qu'une autre kinase puisse également phosphoryler UvrY expliquant l'importance de ce régulateur.

Pour finir YedV-YedW est un système peu connu qui semble être associé au stress oxydatif.

Pour le système NarX-NarP associé au métabolisme des nitrates nous verrons qu'il nous a beaucoup intéressé ; aussi pour plus de détails se reporter au chapitre 1.2.6. de l'introduction.

Les fusions des régulateurs de système à deux composants avec *lacZ* ont été construites en choisissant le +1 de transcription décrit dans la littérature ou en analysant les données de « RNAseq » disponibles (Thomason, Bischler et al. 2015). Le tableau de la figure 80 récapitule les différents systèmes sélectionnés et la position du +1 retenue par rapport à l'ATG. La région 5'UTR est suivie de toute la phase ouverte excepté le codon stop et mise en phase avec le gène rapporteur *lacZ* (figure 80). De plus, le promoteur originel a été remplacé par le promoteur P_{tet} afin de s'affranchir des effets transcriptionnels et avoir une meilleure expression car en général l'expression des régulateurs des systèmes à deux composants est faible.



Système à deux composants	Gène du régulateur	Position du +1 par rapport à l'ATG (*GTG ** TTG)	Région présente sur la fusion
BaeS-BaeR	<i>baeR</i>	-30	-30 à +720
BasS-BasR	<i>basR</i>	-18	-18 à +666
EvgS-EvgA	<i>evgA</i>	-125	-125 à +612
KdpD-KdpE	<i>kdpE</i>	-39 *	-39 à + 675
NarQ-NarP	<i>narP</i>	-150	-150 à +645
PhoR-PhoB	<i>phoB</i>	-41	-41 à +687
RstB-RstA	<i>rstA</i>	-13 *	-13 à +726
BarA-UvrY	<i>uvrY</i>	-44 **	-44 à +654
YedV-YedW	<i>yedW</i>	0 (pas de leader décrit)	+1 à +669

Figure 80 : Construction des différentes fusions traductionnelles. Le schéma représente la structure générique des fusions traductionnelles. Le tableau précise quant à lui les régions présentes sur la fusion pour chacun des systèmes étudiés ainsi que la position du +1 par rapport à l'ATG.

Ces fusions sont intégrées au locus *lacZ* d'une souche d'*E. coli*, puis nous avons transduit l'allèle *hfq::cat* dans ces souches. Ceci permet d'avoir une indication possible de la régulation par des petits ARN Hfq-dépendants. Bien évidemment une régulation n'est pas exclue même sans effet du mutant d'Hfq. La figure 81 présente les résultats du dosage de ces différentes fusions.

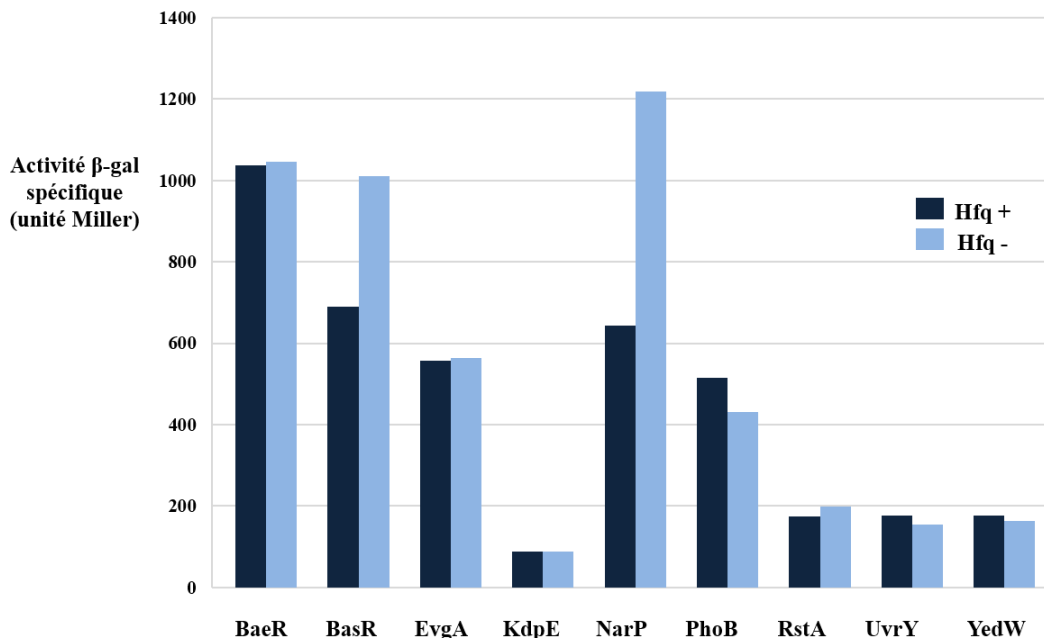


Figure 81 : Expression des différentes fusions des régulateurs des systèmes à deux composants dans un contexte sauvage ou muté pour *hfq*. Les souches utilisées sont les suivantes avec respectivement la fusion sauvage et son dérivé muté : P_{tet} -*baeR-lacZ* (AB1000-AB1009) ; P_{tet} -*basR-lacZ* (AB1001-AB1010) ; P_{tet} -*evgA-lacZ* (AB1002-AB1011) ; P_{tet} -*kdpE-lacZ* (AB1003-AB1012) ; P_{tet} -*narP-lacZ* (AB1004-AB1013) ; P_{tet} -*phoB-lacZ* (AB1005-AB1014) ; P_{tet} -*rstA-lacZ* (AB1006-AB1015) ; P_{tet} -*uvrY-lacZ* (AB1007-AB1016) ; P_{tet} -*yedW-lacZ* (AB1008-AB1017) . Les activités sont mesurées en phase exponentielle en milieu en milieu LB à 37°C.

L'effet d'Hfq est le plus prononcé pour les fusions P_{tet} -*basR-lacZ* et P_{tet} -*narP-lacZ* avec respectivement des facteurs d'activation de 1.5 et de 1.9 dans les mutants *d'hfq*. Ceci suggère que des petits ARN Hfq-dépendants réprimant l'expression de ces deux fusions ne peuvent agir ou sont moins présents dans le mutant *hfq*.

Nous avons décidé pour la suite de l'étude de nous concentrer sur la fusion P_{tet} -*narP-lacZ* car l'effet est le plus important. La souche portant cette fusion a été transformée par la banque de plasmides surproduisant la plupart des petits ARN Hfq-dépendants d'*E. coli* décrits dans la littérature (figure 82). Cette banque disponible au laboratoire à été construite par Pierre Mandin (Mandin and Gottesman 2010) et a été complétée au cours du temps avec le plasmide pMicL que nous avons récupéré (Guo, Updegrave et al. 2014) mais également les plasmides portant DapZ et CpxQ que nous avons construits au laboratoire. Tous ces petits ARN sont sous le contrôle du promoteur P_{lac} donc leur expression est inductible par de l'IPTG.

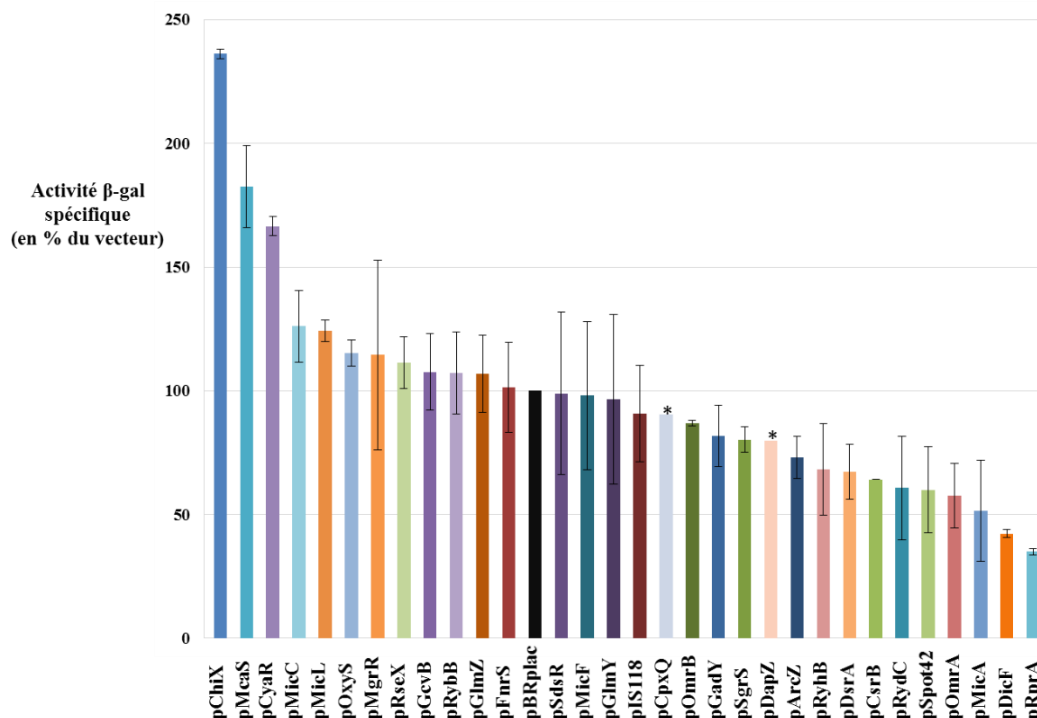


Figure 82 : Activité spécifique de la souche AB1004 portant la fusion P_{tet} -*narP*-*lacZ* et transformée avec la banque de petits ARN Hfq-dépendants. L'activité de la souche transformée par le vecteur pBRplac a été arbitrairement fixée à 100%. Pour les petits ARN DapZ et CpxQ mentionnés par une étoile le dosage n'a été effectuée qu'une seule fois car nous avons construit ces plasmides après les autres. Les activités sont mesurées en phase exponentielle en milieu LB Amp150 μ g/ml IPTG 10^{-4} M à 37°C.

Nous avons normalisé les résultats en plaçant le vecteur à 100% et exprimé les autres données en fonction de ce dernier. On peut noter que certains petits ARN (ChiX, McaS ...) activent la fusion, néanmoins ces derniers sont connus pour titrer Hfq (Moon and Gottesman 2011) donc l'effet de ces plasmides n'a pas été étudié plus en détails.

Par ailleurs, et de façon intéressante, plusieurs petits ARN répriment l'expression de la fusion. Le facteur le plus important à savoir 2.8 est attribué au plasmide surproducteur de de RprA. Ce petit ARN est connu pour réguler positivement *rpoS* (Majdalani, Chen et al. 2001) et *ricI* (Papenfort, Espinosa et al. 2015) mais aussi négativement *csgD* et *ydaM* (Mika, Busse et al. 2012). Sa synthèse est contrôlée quant à elle par le phosphorelais Rcs (Majdalani, Hernandez et al. 2002).

De plus, il est possible de prédire un appariement entre le messenger de NarP et le petit ARN (figure 83). Dans la suite du projet, je me suis donc attachée à caractériser la régulation de *narP* par RprA et notamment déterminer si cette interaction était directe.

En ce qui concerne les autres petits ARN identifiés dans cette approche, je développerai les quelques pistes que nous avons poursuivies dans le paragraphe 3.3.5.

3.3.2. Le messenger *narP*, une cible directe de RprA

Nous avons vu dans le paragraphe précédent qu'un appariement pouvait être prédit entre le messenger de NarP et RprA. Aussi nous avons réalisé les mutations dans la zone d'appariement prédite à la fois sur l'ARNm et sur RprA présentées dans la figure 83 afin de valider *in vivo* cette interaction. On peut mentionner qu'une première série de mutations a été effectuée (décalage d'un nucléotide coté 3') mais celle-ci affectait trop le Shine Dalgarno de *narP* et avait donc pour conséquence une expression très faible de la fusion.

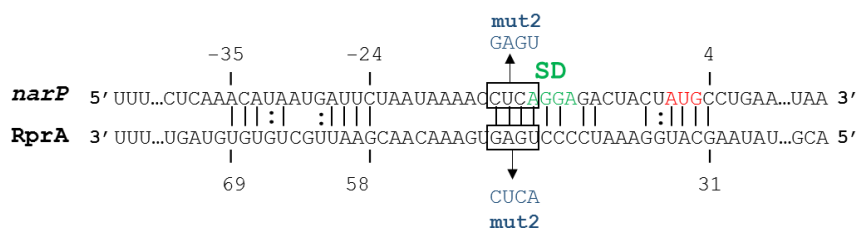


Figure 83 : Mutations compensatoires dans RprA et dans la fusion P_{tet} -*narP*-*lacZ*. Prédiction effectuée avec le logiciel IntaRNA (Mann, Wright et al. 2017).

Les mutations ont été introduites par PCR avec des oligonucléotides portant la mutation pour obtenir le plasmide pRprAmut2. La fusion mutée a quant à elle été construite par PCR puis recombriquée. Après transduction de l'allèle *rprA* ::*tet* dans les souches portant les fusions sauvages et mutantes afin de s'affranchir de la copie de RprA chromosomique, celles-ci ont été transformées par les plasmides pBRplac (vecteur), pRprA et pRprAmut2. La figure 84 présente les résultats des dosages d'activité β -galactosidase dans ces différentes souches portant les fusions.

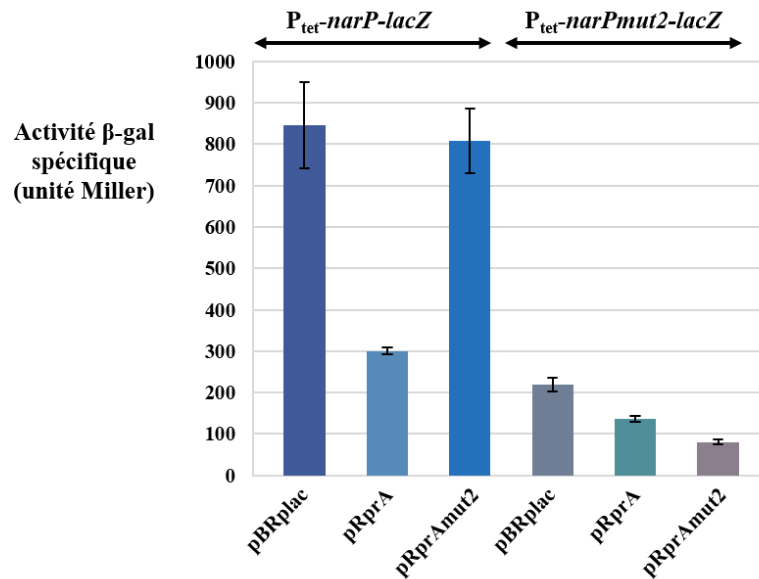


Figure 84 : *narP* est une cible directe de RprA. Activité spécifique des souches AB1018 et AB3037 portant respectivement la fusion $P_{tet}\text{-narP-lacZ}$ et la fusion $P_{tet}\text{-narPmut2-lacZ}$ transformées par les plasmides pBRplac, pRprA et pRprAmut2. Les activités sont mesurées en phase exponentielle en milieu LB Amp150 $\mu\text{g/ml}$ IPTG 10^{-4} M à 37°C.

L'expression de la fusion sauvage est réprimée d'un facteur 3 lorsque RprA est surexprimé. De la même manière la fusion mutante est réprimée 2.7 fois par le plasmide mutant. En revanche la mutation 2 supprime la répression que l'on observait en condition sauvage. Dans le contexte de la fusion mutante on peut noter que le plasmide portant RprA conserve un léger effet, ce qui indique qu'il y a sans doute d'autres régions de RprA pouvant participer à l'effet notamment celles plus en amont que nous pouvions prédire. Cet effet peut également être dû à un appariement alternatif entre le messager mutant et le petit ARN sauvage.

Afin de vérifier que l'effet du plasmide pRprAmut2 est bien dû à une perte d'interaction et non à une surproduction moindre de la version mutante de RprA les niveaux d'ARN ont été vérifiés par Northern blot (figure 85).

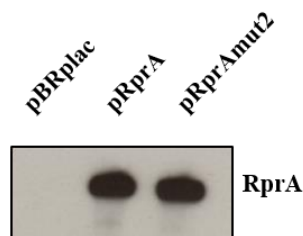


Figure 85 : Niveaux de surproduction de RprA et RprAmut2. La souche AB1018 ($P_{tet}\text{-narP-lacZ}$; $rprA ::tet$) transformée par les plasmides pBRplac, pRprA ou pRprAmut2 a été cultivée en LB Amp 150 $\mu\text{g/ml}$ IPTG 10^{-4} M à 37°C. Les ARN ont été extraits en phase exponentielle de croissance et analysés par Northern Blot grâce à une sonde spécifique de RprA.

Au vu des résultats des fusions et du Northern nous pouvons conclure que *narP* est bien une cible directe de RprA. Le mécanisme est probablement canonique à savoir une inhibition de la fixation de la sous-unité 30S du ribosome étant donné que RprA se fixe dans la région du SD.

Nous avons également cherché à comprendre quelle portion du messenger était requise pour avoir une régulation par RprA. Pour cela des fusions de taille différente ont été construites : une fusion dite « fusion courte » $P_{tet}\text{-narP}_s\text{-lacZ}$ possédant seulement 30 nucléotides après l'AUG et une fusion dite « fusion 105 » $P_{tet}\text{-narP}_{105}\text{-lacZ}$ possédant 105 nucléotides après l'AUG. Nous avons choisi cette position car les résultats de RNAseq après cross-link avec Hfq (Holmqvist, Wright et al. 2016), suggéraient la présence d'un site de fixation d'Hfq entre les 24^{ème} et 35^{ème} nucléotides de l'ORF de *narP*. Après transformation par le plasmide RprA nous avons comparé l'expression de ces deux fusions plus courtes à celle de la fusion $P_{tet}\text{-narP-lacZ}$. Les résultats sont présentés dans la figure 86.

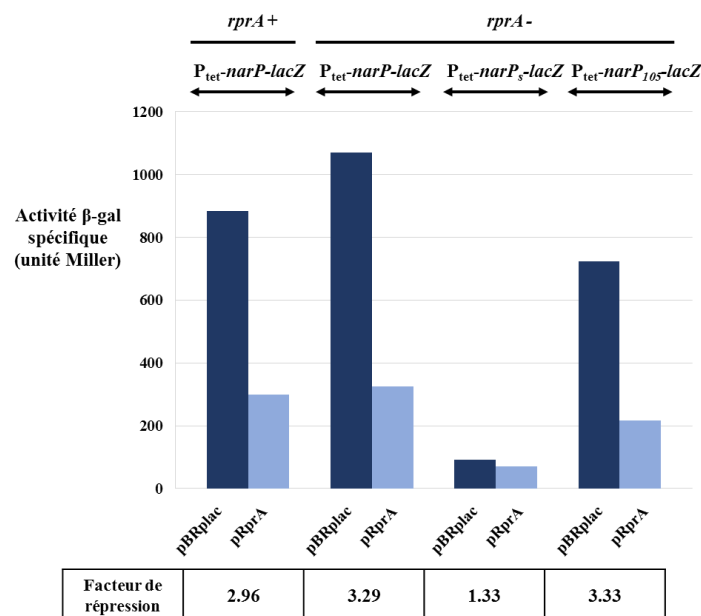


Figure 86 : Les 105 premiers nucléotides de la phase codante de *narP* sont requis pour la régulation par RprA. Activité spécifique des souches AB1004, AB1018, AB1022 et AB3033 portant respectivement la fusion $P_{tet-narP-lacZ}$ et les fusions $P_{tet-narP-lacZ}$, $P_{tet-narP_s-lacZ}$ et $P_{tet-narP_{105-lacZ}}$ transformées par les plasmides pBRplac et pRprA. Les activités sont mesurées en phase exponentielle en milieu LB Amp150 $\mu\text{g/ml}$ IPTG 10^{-4} M à 37°C .

Ces résultats montrent que les 105 premiers nucléotides de la phase codante de *narP* sont suffisants pour observer le contrôle par RprA. Ceci suggère que seule cette portion de *narP* soit requise pour la régulation. On peut néanmoins noter que l’expression de la fusion $P_{tet-narP_{105-lacZ}}$ est plus faible que la version non tronquée ($P_{tet-narP-lacZ}$) suggérant que la phase codante puisse participer à une stabilité du messager *narP*. De plus, les résultats obtenus avec la fusion portant seulement 30 nucléotides après l’AUG montrent que l’on a une activité près de 10 fois plus faible que pour la fusion $P_{tet-narP-lacZ}$ suggérant qu’il pourrait aussi y avoir d’autres déterminants pour la stabilité du messager dans la portion 30 à 105 après l’AUG. Par ailleurs, dans cette configuration on perd la plupart de l’effet dû à RprA. Ceci va dans le sens de la présence du site de fixation d’Hfq identifié sur les données de RNAseq. Il serait donc très intéressant de réitérer cette expérience dans un contexte de délétion d’*hfq*. Nous pourrions à la fois savoir si Hfq stabilise le messager *narP* en comparant les fusions de tailles différentes mais aussi si l’absence d’Hfq abolit la régulation de *narP* par RprA, à condition toutefois que les niveaux de RprA soient suffisants en l’absence d’Hfq.

Par ailleurs, j’ai tenté de confirmer la régulation de *narP* par RprA en travaillant avec la copie chromosomique de cet ARN. L’approche retenue était de se placer dans un contexte activant

RprA à savoir un mutant du système Rcs et de voir l'effet de la délétion de *rprA* sur le niveau de messenger *narP*. A ce jour, je n'ai pas encore réussi à obtenir ces souches. Le mutant de Rcs étant très mucoïde, toutes les transductions de la délétion *rprA* ::*tet* testées à ce jour ont échoué.

3.3.3. Effet de la modulation du niveau de *narP* sur les cibles du système NarQ-NarP

Comme nous l'avons vu dans le chapitre consacré à la régulation d'OmpR par les Omr, moduler un régulateur au niveau de son messenger peut ne pas affecter l'expression de ses cibles s'il existe une robustesse du système. Il nous a donc semblé intéressant de voir si sur ce nouveau système étudié, la modulation de *narP* par RprA pouvait affecter les cibles de NarP. Nous avons choisi deux opérons-cibles de NarP à savoir *napFDAGHBC* et *nirBDC* afin d'étudier cet aspect. Ces cibles étant aussi des cibles de NarL nous avons décidé de travailler dans un contexte mutant pour *narL* pour s'affranchir des potentiels effets de compensation. Des fusions transcriptionnelles P_{napF} -*lacZ* et P_{nirB} -*lacZ* ont été construites puis combinées à une délétion *narL* ::*cat* et/ou *narP* ::*kan*.

La figure 87 présente l'activité des fusions P_{napF} -*lacZ* et P_{nirB} -*lacZ* dans les contextes sauvages, délété pour *narL*, délété pour *narP* ou un double mutant.

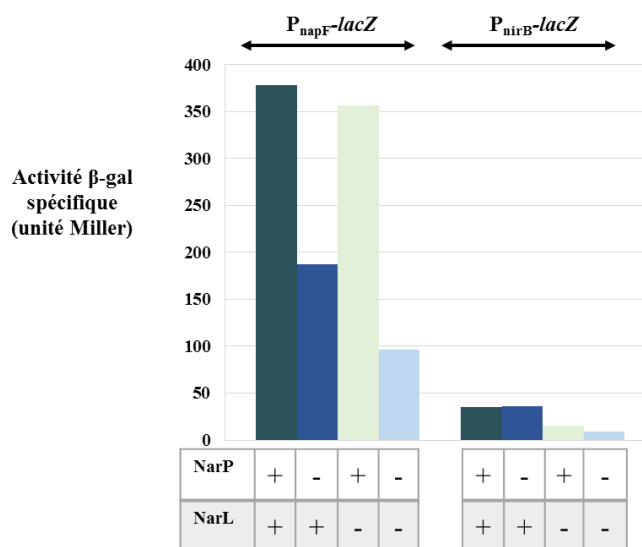


Figure 87 : Contrôle des opérons *napFDAGHBC* et *nirBDC* par les régulateurs NarL et NarP. Activité spécifique de la fusion P_{napF} -*lacZ* et P_{nirB} -*lacZ* dans un contexte sauvage (AB1042 – AB1043), mutant pour *narP* (AB1044 – AB1045), mutant pour *narL* (AB1052 – AB1053) ou double mutant pour *narP* et *narL* (AB1048 – AB1049). Les activités sont mesurées en phase exponentielle en milieu LB à 37°C.

La délétion de *narP* a un effet plus important sur la fusion $P_{\text{napF}}\text{-lacZ}$ que la délétion de *narL*. Dans le double mutant l'expression de la fusion est encore diminuée par rapport au simple mutant *narP*. Ces résultats sont tout à fait en accord avec ceux obtenus par Stewart *et al.* (Stewart, Bledsoe et al. 2003) lorsqu'ils ont caractérisé la régulation de *napF* par NarP et NarL. De la même façon, la fusion $P_{\text{nirB}}\text{-lacZ}$ se comporte comme attendu si l'on se réfère à l'article de Tyson *et al.* (Tyson, Cole et al. 1994). En effet, sur cette fusion la délétion de *narL* a un effet plus important que celle de *narP* et le double mutant est le contexte où la fusion est la moins exprimée. Pour la suite de l'étude, nous avons choisi pour le moment de ne travailler qu'avec la fusion $P_{\text{napF}}\text{-lacZ}$ car cette fusion est mieux exprimée que la fusion $P_{\text{nirB}}\text{-lacZ}$ et est plus sensible aux diminutions du niveau de NarP.

Les précédentes mesures étant effectuées en LB nous avons également voulu vérifier que la fusion $P_{\text{napF}}\text{-lacZ}$ répondait comme attendu en condition aérobie mais surtout dans les conditions d'induction du système NarQ-NarP à savoir en anaérobiose avec ou sans la présence de nitrates. Pour ces tests les cultures ont été faites non plus en LB mais en MOPS supplémenté en casamino-acides afin d'avoir un milieu plus contrôlé. Ne disposant pas de chambre de Freiter, nous avons utilisé des tubes fermés et sans agitation pour créer une micro-anaérobiose. Plus précisément, les cellules ont été cultivées à 37°C en aérobie jusqu'à une densité optique de 0.05 puis transférées avec ou sans nitrates dans des tubes Eppendorfs fermés par un parafilm et placés dans un bain à 37°C. Nous avons d'ailleurs vérifié que ces conditions permettaient de simuler l'anaérobiose en vérifiant par Northern blot l'induction du petit ARN FnrS (données préliminaires). En effet, celui-ci est induit en anaérobiose alors que son niveau en aérobie est très faible comme cela avait été décrit précédemment (Durand and Storz 2010).

Les conditions utilisées induisant bien l'anaérobiose, l'activité de la fusion $P_{\text{napF}}\text{-lacZ}$ dans les contextes génétiques différents et dans les conditions de culture différentes a pu être étudiée. La figure 88 présente les différents résultats obtenus.

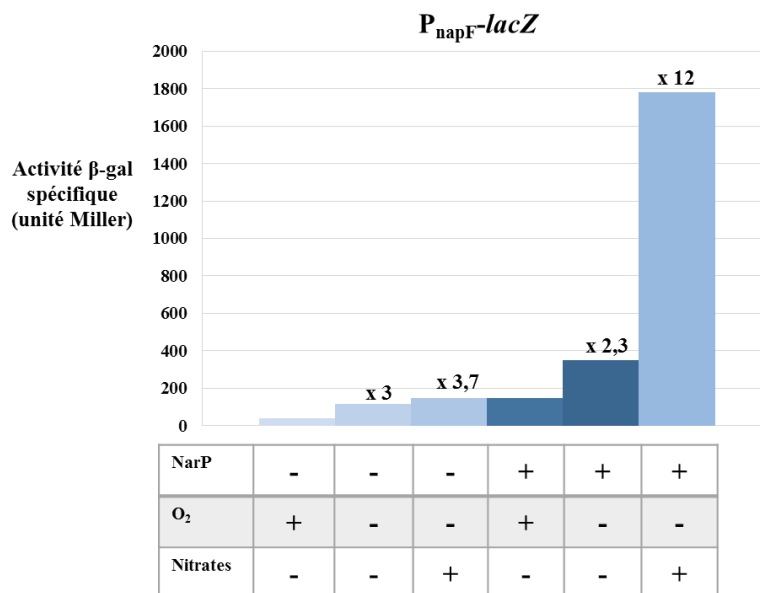


Figure 88 : L'opéron *napFDAGHBC* est bien spécifiquement induit par NarP en micro-anaérobie et en présence de nitrates. Activité spécifique de la fusion P_{napF}-lacZ dans un contexte sauvage pour *narP* (AB1052) ou mutant pour *narP* (AB1048) ; le gène *narL* est toujours délété. Les activités sont mesurées en phase exponentielle en milieu MOPS avec ou sans oxygène et additionné ou non de 40 mM de NaNO₃ à 37°C.

Comme on peut le voir la fusion P_{napF}-lacZ est bien activée spécifiquement par NarP lorsque l'on cultive ces souches sans oxygène et avec des nitrates. En effet, dans le double mutant $\Delta narL \Delta narP$, l'activation de la fusion est plus faible, même en condition d'induction par les nitrates. Ces résultats sont en accord avec ceux décrits par Stewart *et al.* (Stewart, Bledsoe et al. 2003) qui démontraient que l'expression de *napF* était induite par NarP en anaérobiose et en présence de nitrates.

Comme notre système fonctionne, les souches portant la fusion P_{napF}-lacZ ont ensuite été transformées par les plasmides pBRplac, pRprA et pRprAmut2 afin de savoir si la modulation de *narP* par RprA avait un effet sur sa cible NapF. Les activités des fusions ont été mesurées à la fois en condition aérobie mais aussi en anaérobiose avec ou sans la présence de nitrates. La figure 89 présente ces différents résultats.

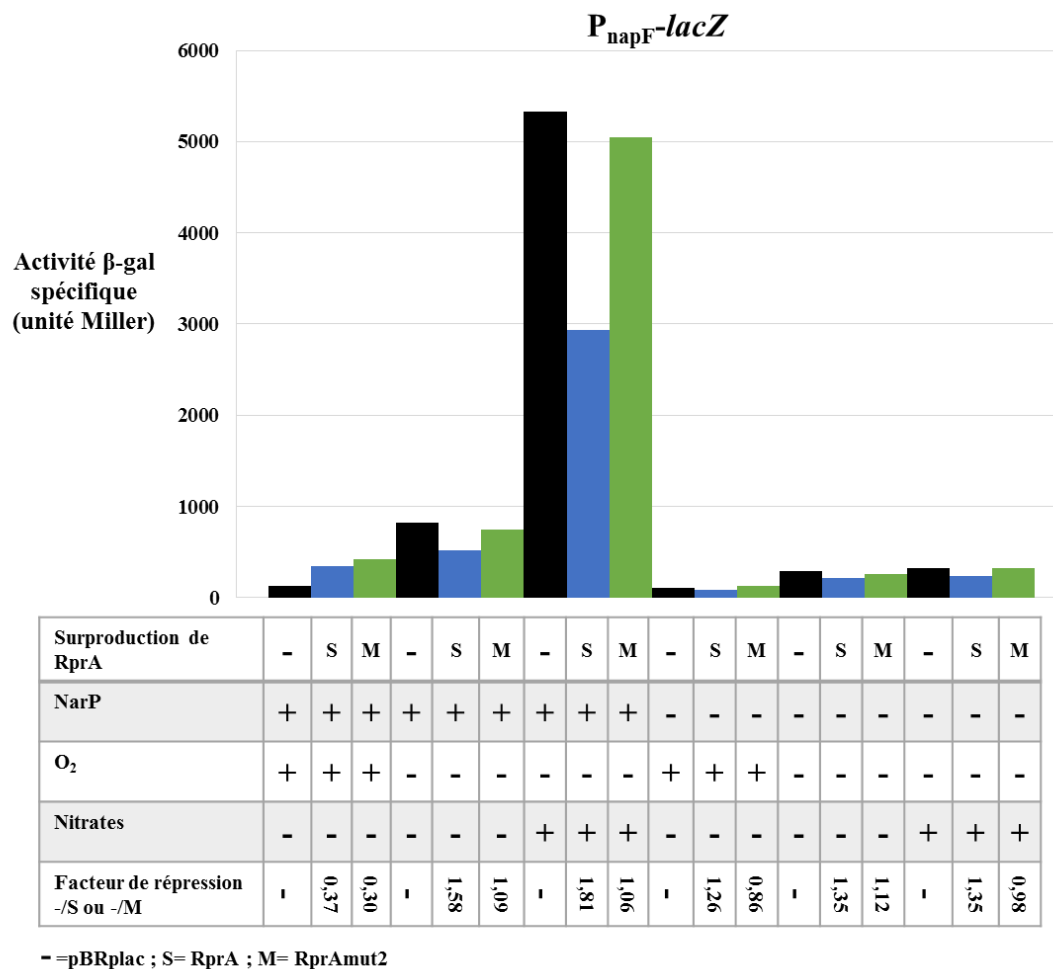


Figure 89 : Le contrôle de *narP* par RprA diminue le niveau de l'opéron *napFDAGHBC*. Activité spécifique des souches portant la fusion P_{napF}-lacZ dans un contexte sauvage pour *narP* (AB1052) ou mutant pour *narP* (AB1048) et transformées soit par le plasmide pBRplac, pRprA ou pRprAmut2. Les activités sont mesurées en phase exponentielle en milieu MOPS Amp150µg/ml IPTG 10⁻⁴ M avec ou sans oxygène et additionné ou non de 40 mM de NaNO₃ à 37°C.

Ces résultats montrent qu'en condition d'induction du système NarQ-NarP (présence de nitrates), la surproduction de RprA diminue l'activation de la transcription de *napF* par NarP. Cet effet est bien spécifique à RprA et lié à son appariement au messenger de *narP* puisque l'on perd cet effet en surproduisant la version mut2 de RprA. Cet effet est également visible en anaérobie seule même s'il est plus modeste que lorsque celle-ci s'accompagne de l'ajout de nitrates. Si l'on se place en aérobie la surproduction de RprA n'affecte pas le niveau de la fusion P_{napF}-lacZ bien qu'on sache que cela modifie le niveau du messenger de *narP*. Ceci permet de supposer que la modulation par RprA de *narP* tend à moduler la quantité de protéine NarP et de sa forme active, logiquement NarP phosphorylée, induisant ainsi une diminution de

l'activation de la transcription de *napF*. L'effet est également spécifique de NarP puisque la majeure partie de l'effet est perdu dans le mutant de *narP*. Comme il n'est pas totalement aboli on peut imaginer que RprA agisse sur un autre régulateur de *napF* autre que NarL ou NarP.

La suite logique de ce travail sera de vérifier les niveaux d'ARN de *narP*, *napF* et RprA dans les mêmes conditions expérimentales que ces dosages β -galactosidase. Nous pourrions également étudier le niveau de protéine NarP et NarP-P lorsque RprA est surproduit. A ce jour, différentes constructions sont en cours pour ajouter une étiquette à la protéine NarP car nous ne disposons pas d'anticorps dirigés contre ce régulateur.

3.3.4. Réponse de RprA aux nitrates

Au cours des expériences précédentes, nous avons extrait les ARN pour contrôler les conditions d'anaérobiose en absence ou en présence de nitrates notamment en vérifiant la détection de FnrS. Nous en avons alors profité pour suivre les niveaux de RprA dans ces conditions. Nous avons ainsi pu remarquer que l'ajout de nitrates avait pour effet d'augmenter le niveau de RprA (figure 90).

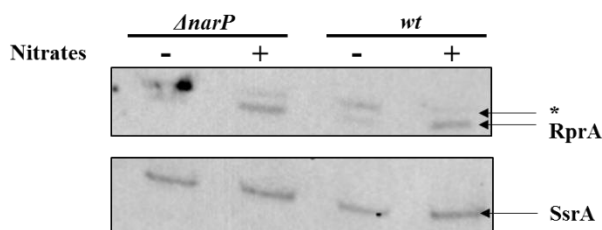


Figure 90 : RprA est activé par la présence de nitrates indépendamment de NarP. Les souches AB1048 (P_{napF} -*lacZ*- ; $\Delta narP$) et AB1052 (P_{napF} -*lacZ*- ; *narP* sauvage) ont été cultivées en micro-anaérobiose en milieu MOPS additionné ou non de nitrates (40 mM de $NaNO_3$) à 37°C. Ces souches sont par ailleurs $\Delta narL$. Les ARN ont été extraits en phase exponentielle de croissance et analysés par Northern Blot grâce à des sondes spécifiques de RprA ou du contrôle SsrA. * Seule la bande inférieure de la piste RprA correspond à cet ARN, la bande supérieure provient d'un sondage précédent pour FnrS.

De manière intéressante, l'activation de RprA par les nitrates est indépendante de la présence de NarP ou de NarL. Aussi, comme RprA est un petit ARN activé par le système de phosphorelais Rcs nous avons voulu savoir si le changement des niveaux de RprA par les nitrates passait par le système Rcs. Pour cela nous avons utilisé un mutant de RcsB, le facteur de transcription qui active directement RprA. L'activité d'une fusion P_{rprA} -*lacZ* a été mesurée dans un contexte sauvage ou $\Delta rcsB$ (figure 91).

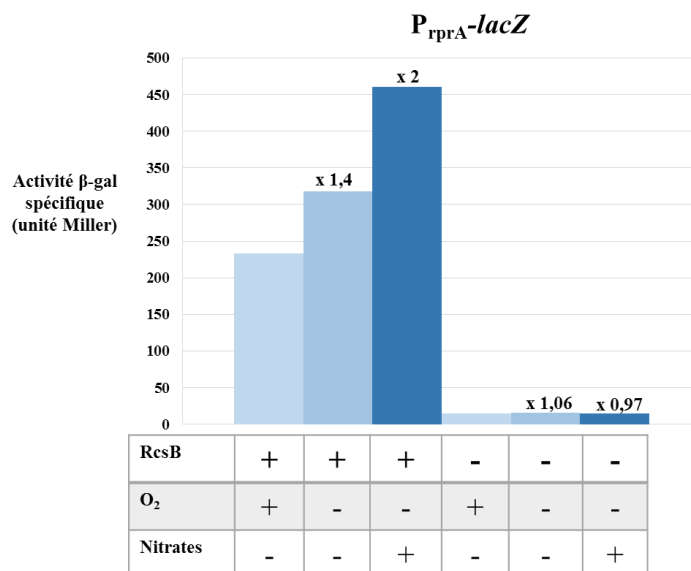


Figure 91 : L'activation de RprA par les nitrates passe par le système Rcs. Activité spécifique des souches portant la fusion P_{rprA}-lacZ dans un contexte sauvage pour *rcsB* (DH300) ou mutant pour *rcsB* (DH311). Les activités sont mesurées en phase exponentielle en milieu MOPS avec ou sans oxygène et additionné ou non de 40 mM de NaNO₃ à 37°C.

Ces données seront à confirmer car le dosage n'a été réalisé qu'une seule fois. Cependant, il semblerait que l'activation de RprA par les nitrates passe par le système Rcs. Ce résultat est très prometteur car il pourrait s'agir d'un nouveau signal capté par le système Rcs. Comme attendu, l'activité de la fusion P_{rprA}-lacZ dans le mutant $\Delta rcsB$ est relativement faible (autour de 15 unités Miller). Il serait donc intéressant d'étudier à nouveau l'induction de RprA par les nitrates dans un mutant constitutivement actif de RcsB ou dans une souche $\Delta rcsC$ où les niveaux de RprA sont augmentés d'un facteur 10 environ (Majdalani, Hernandez et al. 2002) ce qui nous permettrait également de suivre les niveaux de RprA par Northern-blot.

3.3.5. Dépendance à RprA des autres petits ARN régulant *narP*

Comme nous l'avons vu dans le chapitre 3.3.1. d'autres petits ARN avaient pour effet de réprimer l'activité de la fusion Ptet-*narP*-lacZ. Nous avons sélectionné OmrA, Spot42, MicA et DicF car leurs effets étaient les plus forts et nous avons voulu savoir si leur effet était direct ou passait par RprA. La souche portant la fusion Ptet-*narP*-lacZ et l'allèle *rprA* ::*tet* a été transformée par les plasmides portant les différents petits ARN. La figure 92 présente les niveaux d'activité de la fusion Ptet-*narP*-lacZ dans ces conditions.

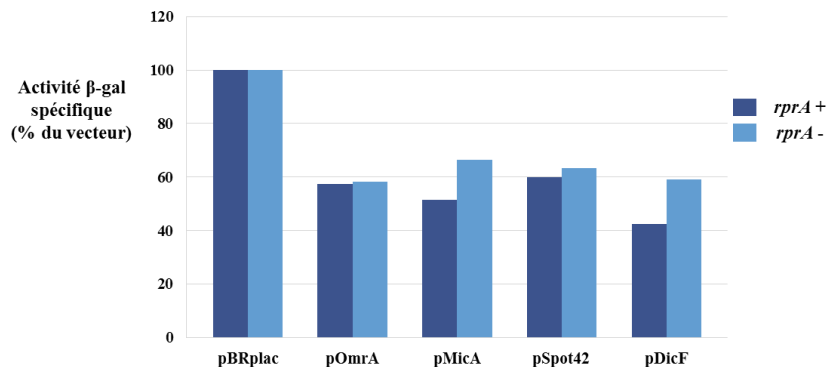


Figure 92 : L'effet d'OmrA, MicA, Spot42 et DicF est indépendant de RprA. Activité spécifique de la fusion P_{tet} -*narP-lacZ* dans la souche sauvage pour *rprA* (AB1004) ou délété pour *rprA* (AB1018) et transformée par les plasmides pBRplac, pOmrA, pMicA, pSpot42 et pDicF. Les activités sont mesurées en phase exponentielle en milieu en LB Amp150 μ g/ml IPTG 10^{-4} M à 37°C.

Ces résultats indiquent que l'effet d'OmrA, MicA, Spot42 et DicF est indépendant de celui de RprA. Au cours de ce projet le laboratoire de Gisela Storz a publié un article indiquant que *narP* pouvait être régulé par un nouveau petit ARN appelé SdsN (Hao, Updegrove et al. 2016). Aussi, nous avons voulu savoir si l'effet de nos plasmides passait par SdsN. La figure 93 présente les niveaux d'activité de la fusion P_{tet} -*narP-lacZ* transformée par les plasmides pBRplac, pRprA, pOmrA, pMicA, pSpot42 et pDicF dans un contexte de délétion de *sdsN*.

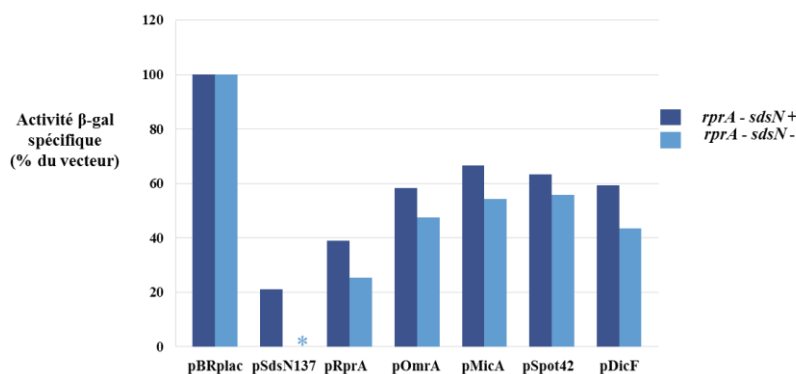


Figure 93 : L'effet d'OmrA, MicA, Spot42 et DicF est indépendant de SdsN. Activité spécifique de la fusion P_{tet} -*narP-lacZ* dans la souche délétée de *rprA* et sauvage pour *sdsN* (AB1018) ou également délétée de *sdsN* (AB1029) et transformée par les plasmides pBRplac, pRprA, pSdsN137, pOmrA, pMicA, pSpot42 et pDicF. Les activités sont mesurées en phase exponentielle en milieu en LB Amp150 μ g/ml IPTG 10^{-4} M à 37°C. * indique que le dosage n'a pas été réalisé pour cette souche.

Au regard de ces données on peut dire que l'effet des différents plasmides ne dépend pas de SdsN ni de RprA. Cependant, l'absence de ces deux régulateurs de *narP* pourrait augmenter artificiellement l'effet des autres plasmides, d'où l'observation d'une répression plus importante avec ces derniers.

4. Conclusion et perspectives

Mes travaux de thèse sur les circuits mixtes de régulation entre les systèmes à deux composants et les ARN régulateurs ont permis d'ajouter un niveau de complexité supplémentaire à la régulation réciproque des OmpR par le système à deux composants EnvZ-OmpR. Nous avons pu constater qu'il existe une hiérarchie entre les diverses cibles d'OmpR (figure 94). En effet, nous avons décrit deux classes de promoteurs sensibles à OmpR : les promoteurs d'*ompC* et d'*ompF* peu sensibles aux variations de niveau d'OmpR et qualifiés de robustes, et ceux des Omr, qui sont à l'inverse sensibles à ces variations et donc ne sont pas robustes, ce qui leur permet de s'autoréguler.

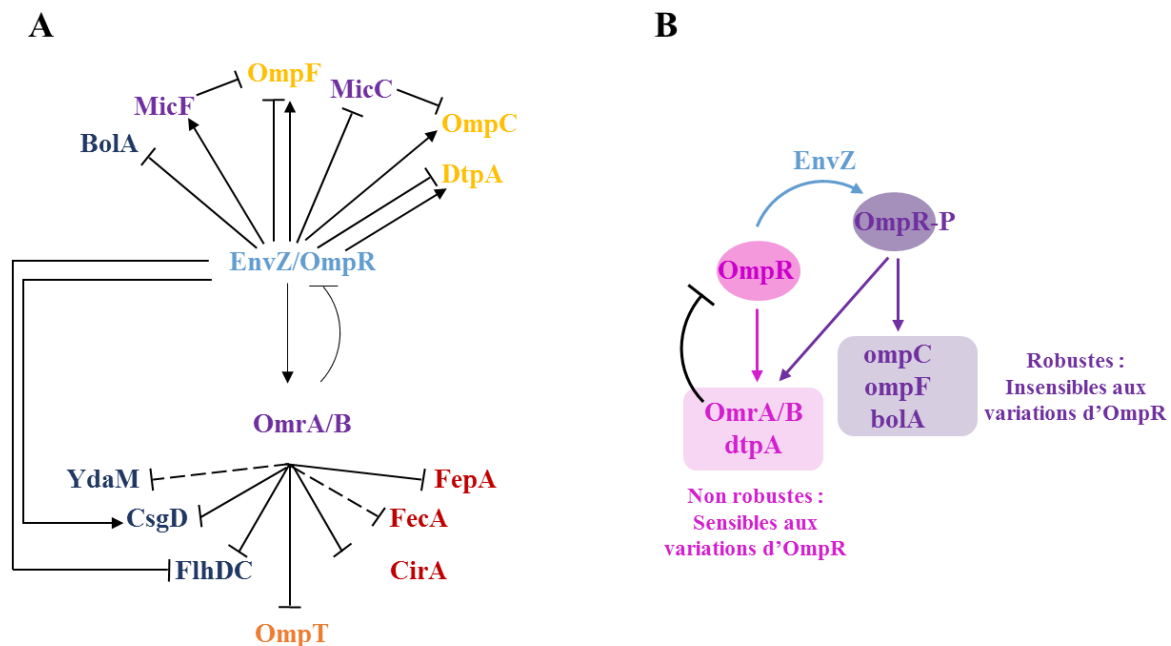


Figure 94 : Résumé du travail sur le système EnvZ-OmpR. A. Interconnexion entre le régulon d'OmpR et celui des Omr. B. Robustesse des cibles du régulateur OmpR.

Mes travaux sur le système à deux composants NarQ-NarP ont pu de leur côté mettre en évidence une régulation de ce système par des petits ARN régulateurs. De façon intéressante, RprA, qui cible le messenger *narP*, répond au même signal cellulaire que la kinase NarQ créant un lien physiologique entre le petit ARN et sa cible. Enfin, et contrairement au système EnvZ-OmpR, la modulation du niveau de NarP par RprA ne semble pas robuste puisqu'elle affecte au moins une cible de NarP.

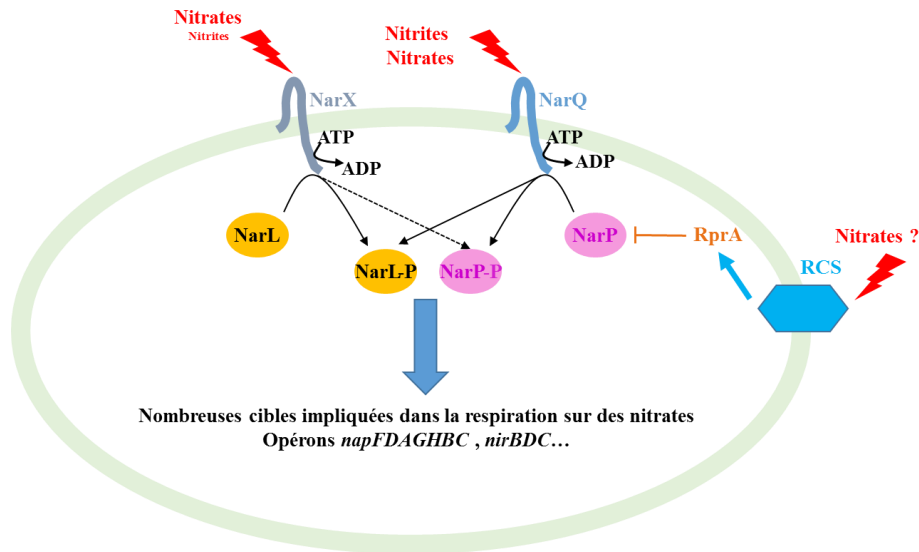


Figure 95 : Résumé du travail sur le système NarQ-NarP et son lien avec le petit ARN RprA.

Mes travaux sur la bactérie pathogène LF82 montrent une forte induction d'OmrA/B OmpR-dépendante au cours de l'infection de macrophages posant ainsi la question du rôle des régulations mixtes entre systèmes à deux composants et petits ARN dans la pathogénie bactérienne. L'ensemble de ces travaux soulève encore de nombreuses interrogations que je me propose d'évoquer ici.

4.1. Réseau de régulation entre les Omr et EnvZ-OmpR

4.1.1. Quel mécanisme pour expliquer la régulation des Omr par OmpR ?

Nous avons pu montrer que les Omr étaient sensibles à la forme non phosphorylée d'OmpR. Cependant, seul, OmpR non phosphorylé ne semble pas être suffisant pour activer la transcription d'*omrA* ou d'*omrB*. Notre première approche pour la recherche des partenaires d'OmpR, en testant divers gènes candidats, n'ayant pas été fructueuse, nous envisageons une approche globale en réalisant un crible génétique afin d'identifier un mutant où l'activation des Omr par la surproduction d'OmpR (via l'allèle Tet-P_{lac-ompR}) serait abolie. Pour ce faire, il serait intéressant de construire une souche où Tet-P_{lac-ompR} serait fusionné à un rapporteur X d'une part et d'autre part où les promoteurs d'*omrA* et d'*omrB* seraient fusionnés à des rapporteurs Y et Z. Une fois les rapporteurs les plus adéquats choisis, nous pourrions réaliser une mutagenèse sur cette souche. L'idée est de chercher un mutant où le niveau des rapporteurs Y et Z diminue sans affecter le niveau du rapporteur X. En effet, une variation du rapporteur X,

relié au Tet- $P_{lac-ompR}$, serait signe d'une mutation affectant la surproduction d'OmpR ce que nous ne recherchons pas. De plus, nous rechercherons en priorité des mutations affectant à la fois les rapporteurs Y et Z reliés à l'activation du promoteur d'*omrA* et d'*omrB* pour augmenter la probabilité qu'il s'agisse d'un partenaire d'OmpR. Dans cette approche il est crucial de bien choisir les rapporteurs pour faciliter la sélection des mutants d'intérêt. Nous pourrions par exemple nous servir de rapporteurs fluorescents afin de sélectionner les mutants en utilisant un microscope ou un cytomètre de flux. Nous pourrions également utiliser un rapporteur de type *lacZ* et faire notre crible sur un milieu additionné de Xgal ou de type « Mac Conkey ». Le choix du type de mutagenèse est également très important. En effet, avec une approche de type « transposon », nous pourrions échouer à détecter le partenaire si ce dernier est codé par un gène essentiel bien que cela soit peu probable s'il s'agit d'un régulateur.

4.1.2. Quels sont les autres régulateurs des Omr ?

Nos premiers résultats ont pu montrer que le système à deux composants RstB-RstA, au moins quand il est surproduit, activait spécifiquement *omrB* en milieu minimum. De nombreuses questions restent en suspens suite à ces données et notamment est ce que RstA est un régulateur direct des Omr ? Pour répondre à cette question, une des premières expériences à envisager serait de confirmer que l'effet est bien dû à la surproduction de RstB-RstA, par exemple en construisant un plasmide ne portant que ces gènes. Nous pourrions ensuite nous intéresser à l'effet de la délétion de *rstAB* sur la fusion *omrB-lacZ* dans un contexte où *ompR* est absent. Il se pourrait en effet que la présence d'OmpR masque l'effet de la délétion de ce second régulateur. L'étape suivante sera, si l'effet est confirmé, de vérifier le niveau du petit ARN OmrB dans un mutant $\Delta rstAB$. Nous devons aussi démontrer que l'effet est direct. Pour cela nous pourrions vérifier l'interaction entre RstA et le promoteur d'*omrB*, ou utiliser une approche de type « transcription *in vitro* » pour valider l'activation d'*omrB* par RstA. Nous pourrions également choisir de nous tourner vers des approches plus globales de type « DNA sampling » (Butala, Busby et al. 2009). Le but de cette méthode est de récupérer les protéines se fixant à une séquence d'ADN d'intérêt tel qu'un promoteur. Non seulement cela nous donnerait une indication supplémentaire de l'interaction entre le promoteur d'*omrB* et RstA mais cela pourrait nous donner accès à d'autres candidats protéiques capables de fixer le promoteur d'*omrB*, ou d'*omrA* si nous l'utilisons aussi comme cible.

4.1.3. Y a-t-il d'autres cibles d'OmpR non robustes et quel en est l'intérêt ?

Nous avons pu montrer que les Omr étaient des cibles d'OmpR non robustes et les premiers résultats avec la cible *dtpA* suggèrent fortement que cette cible soit également sensible aux variations d'expression d'*ompR-envZ*. A contrario, l'expression d'*ompC*, d'*ompF* et de *bolA*, les autres cibles testées dans le cadre de cet article, n'est pas modifiée quand la concentration d'OmpR varie ; ces cibles sont donc robustes. Une suite logique de cette étude serait de s'intéresser aux autres cibles d'OmpR décrites et notamment *csgD*. Cette cible est particulièrement intéressante car elle est à la fois activée au niveau transcriptionnel par OmpR-P mais aussi réprimée au niveau post-transcriptionnel par les Omr. Pour déterminer si sa transcription répond également à des variations d'OmpR nous pourrions regarder l'effet du Tet- $P_{lac-ompR}$ ou du plasmide pOmpR sur une fusion transcriptionnelle *csgD-lacZ* où la zone de fixation des Omr serait absente. Se pose alors immédiatement la question de l'intérêt d'une telle régulation à première vue incohérente. On peut à ce sujet noter que *csgD* présente des différences d'expression au sein des populations (Grantcharova, Peters et al. 2010). On peut donc se demander si cette double régulation est à l'origine de ce phénotype.

L'intérêt majeur d'avoir des cibles comme OmrA/B qui sont moins robustes qu'OmpC/F est qu'ils peuvent limiter leur propre synthèse via la boucle de rétrocontrôle et, plus généralement, répondre à des variations d'EnvZ et d'OmpR. Elles seront donc sensibles non seulement aux signaux qui modifient la phosphorylation d'OmpR, mais aussi à ceux qui modifient son expression via CRP, IHF ou HNS, ou même OmrA/B (Pratt, Hsing et al. 1996). Dans ce cadre il serait particulièrement intéressant de rechercher des régulateurs de l'expression d'*ompR-envZ* qui pourraient nous donner des indications sur les signaux induisant les Omr.

La limitation du niveau des Omr pourrait avoir un avantage cinétique. En effet, la perception du signal inducteur d'EnvZ-OmpR active la production des Omr. Ces derniers vont cibler différents ARNm entraînant très probablement la dégradation du complexe ARNm/Omr (par analogie avec les autres systèmes Hfq-dépendants). Lorsque le signal inducteur cesse, la présence de la boucle peut permettre de restaurer plus rapidement un niveau basal des Omr car ils auront au préalable limité leur propre synthèse. Cette boucle pourrait permettre également un retour à la normale plus rapide du niveau d'ARNm *ompR-envZ*, qui augmente lors de l'activation d'EnvZ-OmpR via un autocontrôle (données préliminaires).

Par ailleurs, on sait que les transcriptions ne sont pas homogènes au sein d'une population bactérienne. Ainsi on peut envisager qu'en cas de transcription plus importante d'OmpR ou des Omr, la boucle de rétrocontrôle garantira un niveau relativement stable de ces molécules et donc l'homéostasie du système. Certaines de mes données préliminaires suggèrent que le niveau d'OmpR est relativement homogène au sein d'une population cellulaire en absence de stress. Il serait donc intéressant de savoir si cette homogénéité est perturbée en l'absence de la boucle de rétrocontrôle des Omr. De la même manière, on peut imaginer que si d'autres cibles d'OmpR sont non robustes, elles pourraient, via la boucle, être exprimées de manière plus homogène dans la population. Cela peut présenter un avantage si l'ensemble de la population est dirigée par exemple vers la formation d'un biofilm ou au contraire une stabilisation dans un état planctonique.

4.1.4. Les Omr ont-ils un rôle direct dans la virulence de LF82 ?

Nous avons pu établir que les petits ARN OmrA et OmrB étaient très fortement induits au cours de l'infection de macrophages par LF82 et ce de manière OmpR-dépendante. LF82 étant un pathogène, nous pouvons donc nous poser la question de l'implication d'OmrA et d'OmrB dans la virulence de cette souche. Sont-ils directement impliqués dans le processus infectieux par exemple en régulant des facteurs de virulence ou participent-ils seulement à la survie des bactéries au sein des macrophages ?

Le mutant $\Delta ompR$ est très affecté dans ses capacités d'infection aussi nous pourrions rechercher le(s) cible(s) responsable(s) de cet effet. Il est peu probable que les Omr soient seuls responsables de cet effet mais on peut cependant noter que l'induction de ces ARN est la plus forte parmi toutes les cibles du régulon OmpR.

Les gènes *omrA/B* n'étant pas situés dans un îlot de pathogénicité et si l'on tient compte de leur fonction chez la bactérie modèle *E. coli* K12 non pathogène, il semblerait qu'ils puissent participer à la persistance et éventuellement à la multiplication de LF82 au sein des macrophages. Ils pourraient notamment être impliqués dans la résistance au stress acide supporté par la bactérie au moment de son internalisation dans la vacuole du macrophage. Une approche intéressante serait d'étudier ces étapes dans une souche de LF82 sauvage ou délétée pour *ompR* (en surproduisant OmpC si besoin pour limiter le défaut d'infection, Rolhion,

Carvalho et al. 2007) ou pour les Omr. La multiplication et la persistance peuvent être suivies en utilisant la méthode de dilution de fluorescence développée par Sophie Helaine et David Holden, disponible au Collège de France (Helaine, Thompson et al. 2010). Cette technique est basée sur la production de deux fluorophores (GFP et mcherry) par les bactéries au cours du temps. L'expression du premier fluorophore est sous le contrôle d'un promoteur constitutif alors que le second est sous contrôle d'un promoteur inductible. Une population de bactéries « pré-induites » exprimant les deux fluorophores est utilisée pour infecter des macrophages (sans l'inducteur). Les bactéries se multipliant dans les macrophages (sans inducteur donc) verront leur fluorescence se diluer par deux à chaque génération. Les cellules en revanche qui ne prolifèrent pas conserveront leur fluorescence initiale. Il est ainsi possible d'évaluer par microscopie le ratio entre les différents types de cellules et de comparer celui-ci en fonction de la souche considérée.

Par ailleurs, l'induction importante des Omr dépend du système EnvZ-OmpR au cours de l'infection. Le promoteur des Omr étant très sensible à la présence d'OmpR-P mais également à celle d'OmpR non phosphorylé, il serait donc intéressant de comparer les niveaux d'OmpR et d'OmpR-P avant ou après infection dans une souche LF82 par exemple en réalisant des extraits à analyser sur gel Phos-tag. Ceci permettrait de déterminer si l'induction d'OmrA/B au cours de l'infection est due à l'activation d'OmpR par phosphorylation et/ou à une forte augmentation de l'expression d'*ompR*. La difficulté de cette procédure résidera dans le fait de récupérer suffisamment de matériel pour une détection par Western blot.

4.2. Réseau de régulation entre NarP et RprA

4.2.1. Quel est le mécanisme de contrôle de *narP* par RprA ?

Nous avons pu voir que RprA régula le niveau de NarP en interagissant directement avec le messenger *narP*. La suite de cette étude sera de caractériser plus en détails le mécanisme d'action mis en jeu dans cette régulation.

Premièrement nous pourrions suivre directement le niveau de messenger *narP* suite à l'induction de RprA. Cela permettrait de confirmer par une approche plus directe nos résultats de fusion traductionnelle. Il serait également intéressant de savoir si une RNase, telle que la RNase E, pourrait intervenir et dégrader *narP* suite à l'interaction avec RprA et si c'est le cas, si l'effet est dépendant ou non d'un blocage traductionnel. Bien que fort probable au regard de la

localisation du site de fixation de RprA, nous devons également confirmer que RprA empêche bien la fixation de la sous-unité 30S du ribosome en réalisant une expérience de toeprint.

En parallèle, il serait intéressant d'investiguer plus en détails le rôle de la phase codante de *narP* dans sa régulation par RprA. En effet, nous avons vu que les 105 premiers nucléotides de l'ORF de *narP* étaient nécessaires à la régulation. Une de nos hypothèses serait que cette portion de l'ARNm de *narP* contienne un site de fixation d'Hfq. En se fixant, Hfq pourrait stabiliser l'interaction entre *narP* et RprA. Nous souhaiterions donc valider le site de fixation d'Hfq prédit. Cette étude serait d'autant plus intéressante que la position du site de fixation d'Hfq est peu classique car relativement éloignée du site d'appariement. Nous pourrions par la suite étudier la structure du messenger *narP* afin de comprendre si un repliement spécifique de la phase codante est requis pour la stabilité du messenger. Ce genre d'étude peut se faire en sondant chimiquement la structure du messenger avec des portions variables de la phase codante.

Enfin, nous souhaiterions suivre le niveau de protéine NarP totale mais aussi le niveau de la forme phosphorylée de NarP en présence ou en absence de la régulation par RprA. Des constructions sont actuellement en cours pour marquer la protéine NarP car nous ne disposons pas d'anticorps dirigé contre ce régulateur. Avec la même idée que l'étude sur EnvZ-OmpR nous pourrions suivre les niveaux de NarP et NarP-P en utilisant la technique de Phos-tag.

4.2.2. Comment la régulation de NarP par RprA se met-elle en place en réponse aux nitrates ?

De premières indications suggèrent que l'induction de RprA par les nitrates passe par le système Rcs. Si elles se valident, il nous faudra par la suite déterminer si l'induction de Rcs par les nitrates peut se faire via RcsF ou s'il s'agit d'un autre mécanisme. La suite sera de comprendre la logique de la régulation de *narP* par RprA. En effet, s'ils sont induits par le même signal il se pourrait que RprA ajoute un niveau supplémentaire dans la complexité de la cross-régulation entre les systèmes NarX-NarL et NarQ-NarP. Il a été montré que NarX était 100 fois plus sensible à la présence de nitrates que de nitrites alors que NarQ ne faisait pas de différence entre ces deux ligands. On peut imaginer qu'en présence de nitrates, RprA pourrait limiter le nombre de molécules de NarP et donc limiter l'induction de ses cibles qui par ailleurs sont déjà activées par NarL. Il s'agirait donc d'une forme d'économie d'énergie pour la cellule.

Plus généralement, il sera important de comprendre le lien physiologique entre les signaux induisant RprA, via Rcs par exemple, et ceux induisant NarP. A priori le lien entre les défauts d'enveloppe sentis par Rcs et les nitrates n'est pas évident. Comprendre également le lien entre *narP* et les autres cibles de RprA (*rpoS*, *csgD*, *ydaM*, *ricI*) pourrait permettre de faire ce lien physiologique.

4.2.3. Comment fonctionne la régulation de NarP par les autres petits ARN ?

Nous avons vu que l'effet d'OmrA, Spot42, DicF et MicA était à la fois indépendant de RprA mais aussi de SdsN. La suite de cette étude serait de déterminer si leur effet est direct ou passe par un ou plusieurs intermédiaires. Des prédictions d'appariement entre OmrA et *narP* suggèrent une interaction directe, cependant les diverses constructions testées à ce jour ne m'ont pas permis de valider cette hypothèse. Pour les trois autres ARN, les prédictions sont moins claires et je m'oriente plutôt vers l'hypothèse d'un effet indirect.

Plus généralement, il serait intéressant de comprendre comment se met en place la régulation de NarP par les différents petits ARN qui le ciblent. On peut remarquer d'ores et déjà que RprA a la capacité d'inhiber *narP* directement mais également indirectement en passant par le petit ARN SdsN. En effet, RprA active RpoS qui à son tour active SdsN, l'autre ARN répresseur de *narP* (figure 97).

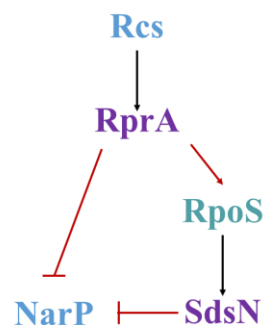


Figure 97 : Régulation directe et indirecte de NarP par RprA. Les régulations transcriptionnelles sont mentionnées par des flèches noires et les régulations post-transcriptionnelles par des flèches rouges.

Nous savons déjà que l'effet de RprA, au moins quand il est surproduit, est indépendant de SdsN mais il serait intéressant de déterminer si cela est vrai aussi pour l'effet d'une délétion de RprA sur l'expression de *narP*. Nous pourrions aussi voir si l'effet de SdsN est modulé en absence de RprA. De plus, nous pourrions suivre les niveaux de SdsN en présence ou en absence

d'induction de RprA. D'une manière générale, les propriétés de ce type de réseaux qui combinent régulations transcriptionnelles et post-transcriptionnelles sont encore mal comprises. On peut cependant imaginer que la régulation directe par RprA est plus rapide, mais que la régulation indirecte via RpoS et SdsN permet ensuite une répression plus importante de *narP*. L'intervention de différents régulateurs permet ainsi d'intégrer divers signaux qui pourront ensuite moduler la synthèse de NarP.

4.3. Les circuits mixtes de régulation

4.3.1. Quels déterminants pour expliquer la robustesse de certains systèmes à deux composants ?

J'ai pu au cours de ma thèse m'intéresser en détails au fonctionnement de deux systèmes à deux composants. Alors que EnvZ-OmpR est un système robuste où le niveau de forme phosphorylée est indépendant des variations d'EnvZ ou d'OmpR, cela ne semble pas être le cas du système NarQ-NarP. Il serait intéressant d'étudier la robustesse de tous les systèmes à deux composants afin de savoir quels en sont les déterminants.

Les critères posés pour le modèle de robustesse d'EnvZ-OmpR, à savoir une kinase bifonctionnelle et un niveau de régulateur très supérieur à celui du senseur, sont assez courants pour les systèmes à deux composants. En effet, plusieurs kinases sont bifonctionnelles et quant au niveau relatif des régulateurs de réponse et des histines kinases, une estimation d'après les données du laboratoire de J. Weissman sur les taux de synthèse déterminés pour les protéines appartenant aux systèmes à deux composants (figure 98) suggère que les régulateurs de réponse sont souvent en excès (Li, Burkhardt et al. 2014). Sur cette base, on peut ainsi proposer que les systèmes ArcB-ArcA, Rcs, CpxA-CpxR, NarX-NarL, PhoQ-PhoP et BarA-UvrY soient robustes.

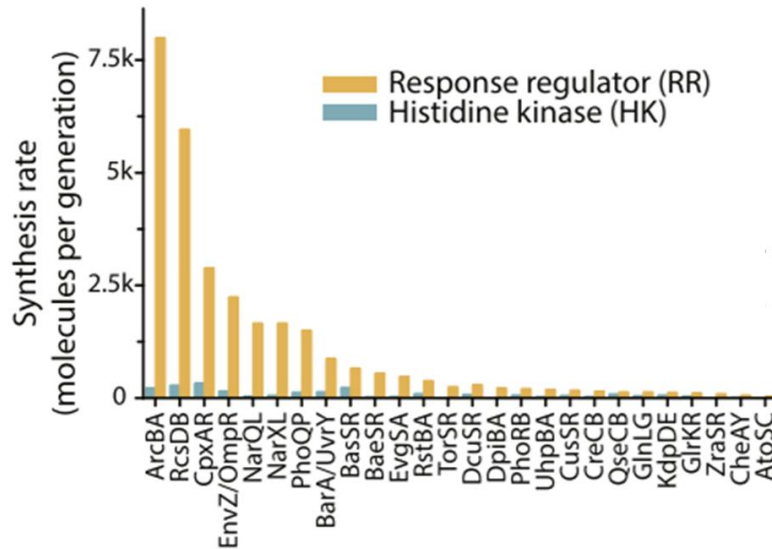


Figure 98 : Taux de synthèse des molécules appartenant aux systèmes à deux composants d'*E. coli*. Sur les 30 systèmes, 26 sont ici représentés. Données issues de (Li, Burkhardt et al. 2014).

Cette propriété semble indépendante de l'organisation des gènes du système à deux composants en opéron puisque, parmi tous les systèmes que je viens d'évoquer, seuls *cpxRA* et *phoPQ* forment un opéron avec le gène du régulateur en amont de celui de la kinase. Plus généralement, prédire la robustesse d'un système à deux composants est encore assez délicat et nécessite donc des études plus expérimentales.

4.3.2. Quelles autres propriétés pour les circuits mixtes ?

Mes travaux ont pu mettre en évidence des propriétés tout à fait uniques des circuits de régulation intégrant des systèmes à deux composants et des petits ARN. Non seulement ces circuits peuvent faire appel à des mécanismes d'action peu communs tels que la réponse des Omr à la forme non phosphorylée d'OmpR, mais également intégrer des signaux divers. En effet, la réponse de RprA aux nitrates via le système Rcs est tout à fait inattendue compte tenu des connaissances actuelles sur ce système.

Les systèmes à deux composants et petits ARN sont des éléments de régulation très répandus chez les bactéries. Chez *E. coli* on compte une trentaine de système à deux composants et au moins autant de petits ARN Hfq-dépendants, aussi l'étude des régulations mixtes offre de nombreuses perspectives et pourrait mettre à jour de nouveaux mécanismes d'action et des réponses physiologiques tout à fait inédites.

Plus généralement, nous savons déjà qu'un certain nombre de petits ARN sont régulés par des systèmes à deux composants mais également par des facteurs sigma ou d'autres facteurs de transcription globaux. De plus, dans un certain nombre de cas, les petits ARN sont les cibles les plus fortement régulées par ces facteurs. Les petits ARN sigma E dépendants en sont un très bel exemple mais cela se vérifie également pour les petits ARN ciblés par des systèmes à deux composants. Nous avons en effet vu que la régulation des Omr par OmpR était très forte et très sensible.

Inversement, de nombreux régulateurs globaux sont ciblés par des petits ARN. Ce type de régulation est très intéressant car elle peut affecter de nombreuses cibles simultanément. Cependant, toutes les cibles du régulateur ne sont pas nécessairement affectées comme on peut le voir avec les systèmes à deux composants. En effet, contrairement aux facteurs de transcription classiques et aux facteurs sigma, les régulateurs de systèmes à deux composants sont modifiés post-traductionnellement et la régulation de l'expression de ces systèmes ne va pas forcément moduler les niveaux de la forme modifiée, qui est a priori la forme active. On peut aussi imaginer qu'en absence de signal inducteur, d'un système à deux composants par exemple, il soit utile pour la cellule de ne pas traduire les ARNm de ces systèmes, ce qui constituerait une économie d'énergie pour la cellule. Cette réflexion est également valable pour les autres facteurs de transcription contrôlés par des petits ARN. Par ailleurs, nombre de systèmes à deux composants semblent également être autorégulés, ainsi les petits ARN pourraient permettre de limiter cet autocontrôle.

Il est également remarquable de voir qu'un certain nombre de petits ARN répondent à des signaux comparables au système à deux composants qu'ils régulent. Il semblerait que ce soit le cas entre RprA et le système NarQ-NarP mais d'autres exemples ont par le passé été décrits dans la littérature. Nous savons que l'induction des Omr passe par OmpR et donc répond aux mêmes signaux, mais il a également été proposé que MicF et RprA, tous deux activés par le stress d'enveloppe, affectent l'expression ou l'activité du système CpxA-CpxR également impliqué dans le même type de stress (Vogt, Evans et al. 2014). On peut ainsi se demander si ce lien physiologique ne se retrouverait pas systématiquement dans les autres régulations de systèmes à deux composants par des petits ARN.

En conclusion, mes travaux ne lèvent le voile que sur quelques propriétés fascinantes des régulations mixtes. Continuer dans cette voie d'étude nous permettrait d'en apprendre beaucoup sur la physiologie bactérienne et la façon qu'ont les bactéries de sentir leur environnement. Ces organismes exceptionnels ont développé et développent encore de formidables outils leur permettant de coloniser de nouvelles niches ou s'adapter à des environnements changeants et parfois hostiles comme nous avons pu le voir avec la survie de la souche LF82 au sein de la vacuole des macrophages. Ainsi, tous les secrets de la bactérie modèle *E. coli* ne sont pas encore percés et réservent sans doute encore de belles découvertes.

5. Annexes

Annexe 1 : Revue sur les méthodes utilisées pour l'étude des interactions entre ARN régulateurs et messagers.

Dans les trois domaines du vivant, l'ARN n'est pas seulement impliqué dans l'expression de l'information génétique, il exerce aussi des fonctions cellulaires extrêmement diverses. Cette polyvalence est principalement due au fait que les molécules d'ARN peuvent par appariement former une grande variété de structures secondaires. Ils peuvent ainsi jouer divers rôles. Ces molécules sont aussi capables de cibler sélectivement d'autres acides nucléiques. Ceci est particulièrement vrai pour les petits ARN régulateurs bactériens qui agissent par appariement de base imparfait avec des ARNm cibles, et contrôlent ainsi leur expression par différents mécanismes. Dans cette revue nous donnons un aperçu des approches *in vivo* et *in vitro* qui sont actuellement utilisées pour comprendre les mécanismes mis en place par les ARN régulateurs bactériens afin de contrôler l'expression des gènes bactériens.



Mechanistic study of base-pairing small regulatory RNAs in bacteria



Jonathan Jagodnik, Anaïs Brosse, Thao Nguyen Le Lam, Claude Chiaruttini, Maude Guillier*

CNRS UMR8261, Associated with University Paris Diderot, Sorbonne Paris Cité, Institut de Biologie Physico-Chimique, 75005 Paris, France

ARTICLE INFO

Article history:

Received 3 August 2016

Accepted 22 September 2016

Available online 28 September 2016

Keywords:

Gene expression regulation

Post-transcriptional control

RNA stability

RNA processing sRNA-

mRNA base pairing RNA

structure probing

SHAPE

Toeprint

Hfq

ABSTRACT

In all three kingdoms of life, RNA is not only involved in the expression of genetic information, but also carries out extremely diverse cellular functions. This versatility is essentially due to the fact that RNA molecules can exploit the power of base pairing to allow them to fold into a wide variety of structures through which they can perform diverse roles, but also to selectively target and bind to other nucleic acids. This is true in particular for bacterial small regulatory RNAs that act by imperfect base-pairing with target mRNAs, and thereby control their expression through different mechanisms. Here we outline an overview of *in vivo* and *in vitro* approaches that are currently used to gain mechanistic insights into how these sRNAs control gene expression in bacteria.

© 2016 Elsevier Inc. All rights reserved.

Contents

1. Introduction	67
2. <i>In vivo</i> approaches	69
2.1. Studying the effect of an sRNA on a target using reporter gene fusions	69
2.2. Effect of the sRNA on mRNA levels and stability	70
2.3. A role for other actors in addition to the sRNA-mRNA duplex?	71
3. <i>In vitro</i> approaches	71
3.1. Synthesis, labeling and detection of RNA molecules for <i>in vitro</i> studies	71
3.2. <i>In vitro</i> characterization of the sRNA-mRNA interaction	71
3.3. <i>In vitro</i> study of translation initiation: the toeprinting assay	73
3.4. <i>In vitro</i> translation experiments	74
3.5. Studying the degradation or processing of sRNA-mRNA duplexes <i>in vitro</i>	74
4. Conclusion	74
Funding	75
Acknowledgements	75
References	75

Abbreviations: CMCT, 1-cyclohexyl-32-(morpholinoethyl)carbodiimide metho-p-toluene sulfonate; DMS, dimethyl sulfate; FACS, fluorescence-activated cell sorting; ketoxal, -ethoxy- α -ketobutyraldehyde; IM7, 1-methyl-7-nitroisatoic anhydride; NMIA, N-methylisatoic anhydride; nt, nucleotide; ORF, open reading frame; PAGE, polyacrylamide gel electrophoresis; RBS, ribosome binding site; RT, reverse-transcription; SHAPE, selective 2° -hydroxyl acylation analyzed by primer extension; UTR, untranslated region.

* Corresponding author.

E-mail address: maude.guillier@ibpc.fr (M. Guillier).

<http://dx.doi.org/10.1016/j.ymeth.2016.09.012>

1046-2023/ 2016 Elsevier Inc. All rights reserved.

1. Introduction

It is now recognized that the transcribed fraction of genomes is much higher than previously anticipated and that RNA molecules play versatile roles in processes other than the translation of genetic expression into proteins [1]. The characterization of several of these RNAs highlighted their role in a wide variety of different biological functions, which include enzymatic activity,

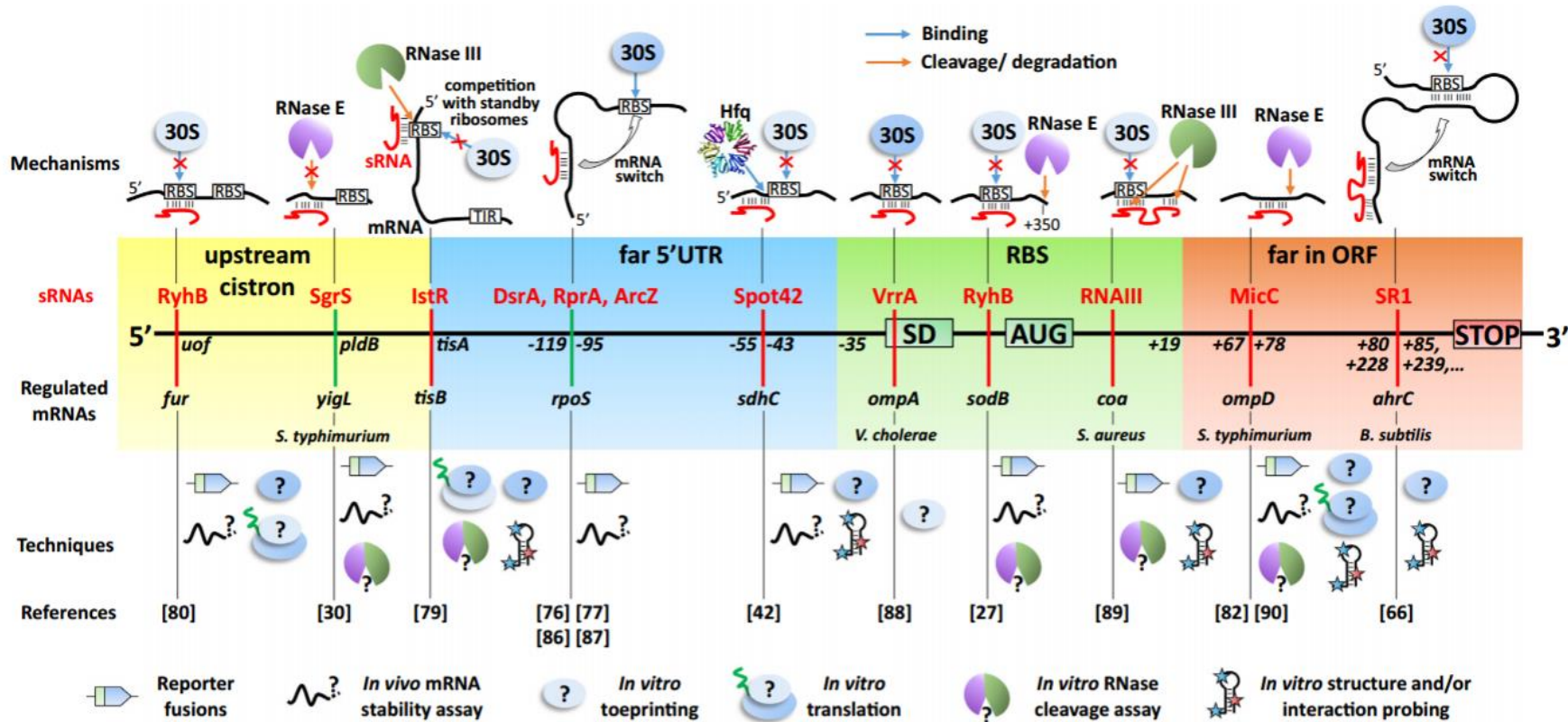


Fig. 1. Diversity of sRNA-based regulatory mechanisms and the techniques used to dissect them. Selected examples of sRNAs that activate (green vertical bars) or repress (red vertical bars) expression of target mRNAs by pairing at various positions of the transcript. The regulatory mechanism and the experimental approaches that were used are schematically depicted above and below, respectively. Four distinct regions for sRNA-binding are shown on the mRNA, with the RBS corresponding to the nts 35 to +19 as determined in [85]. Positions of the mRNA (or the cistron when relevant) where the sRNA binds are shown when pairing occurs outside of the RBS. When different from *E. coli*, the organism in which the study was done is indicated. See text for details. TIR: translation initiation region. (See above-mentioned references for further information.)

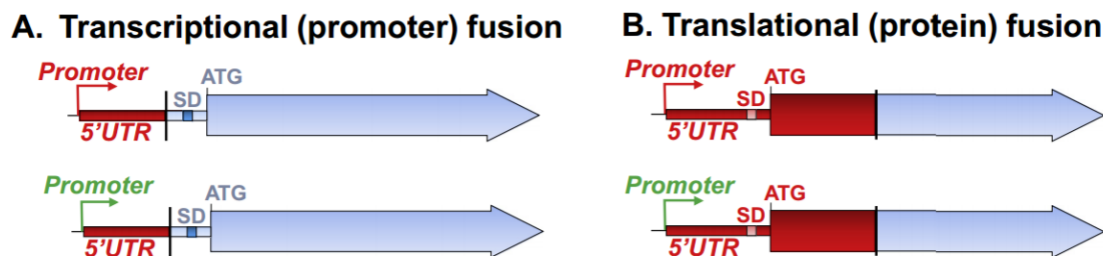


Fig. 2. Schematics of the transcriptional and translational fusions. Sequences of the gene of interest and of the reporter gene are shown in red and blue, respectively. Black vertical bars indicate the fusion points. Green arrows represent heterologous promoters, either constitutive or inducible, that replace the promoter of the gene of interest. Note that transcriptional fusions can carry some or all of the coding sequence of the target gene as well.

maintenance of genome integrity by protection against foreign DNA, processing of precursor RNAs, methylation or pseudouridylation of rRNA, and modification of protein activity [2].

RNA molecules have also been found to play important regulatory roles at each stage of gene expression via extremely diverse mechanisms. This is particularly true in bacteria where, in a few model organisms at least, detailed understanding of their modes of action has elucidated novel modes of RNA-based control [3].

Bacterial regulatory RNAs are extremely diverse from different standpoints. If we consider their synthesis for instance, they can be *bona fide* primary transcripts made from free-standing genes, but also originate from promoters internal to coding sequences (e.g. MicL [4]), or 3' untranslated regions (e.g. DapZ [5]). Regulatory RNAs can also be generated by the processing of longer mRNAs (e.g. CpxQ [6]), tRNAs precursors [7], or even from longer isoforms of regulatory RNAs (e.g. SdsN [8]). Lastly, regulatory RNAs can constitute an integral part of an mRNA: this is the case for riboswitches for instance [9]. Another source of diversity within regulatory RNAs comes from their various modes of action. While riboswitches affect expression of *cis* encoded genes in response to ligand binding and conformational rearrangement, other regulatory RNAs modify the activity of a protein or more often act *via* base-pairing with mRNA. Base-pairing RNAs fall into two main classes: (i) the antisense RNAs (asRNAs), transcribed from the DNA strand opposite to their target gene, and that consequently have perfect and extended complementarity to the mRNA and (ii) the trans-encoded regulatory RNAs that target mRNA(s) encoded by a different *locus via* imperfect complementarity and therefore can regulate multiple genes in most cases [10].

The aim of this review is to provide an overview of the different approaches that can be used to study mechanisms by which bacterial regulatory RNAs control gene expression. We have chosen to focus more specifically on RNAs acting by imperfect base-pairing because this represents a major class of regulatory RNAs. Hereafter we will refer to these as small RNAs (sRNAs), as their typical size is in the 70–200 nts range. Furthermore, we will not cover the techniques aiming at the identification of direct mRNA targets of sRNAs (reviewed in [3], see also [11] for a more recent technique); we will thus focus on sRNAs and targets that were previously shown to interact directly. Most approaches mentioned below have nonetheless much broader applications and can be used in particular to study the action of other regulators, including RNAs, that do not act via imperfect base-pairing.

The best way to get insights into the detailed mode of action of an sRNA is to combine *in vivo* and *in vitro* approaches, that will bring complementary information. The techniques to be used will depend on the sRNA/target pair being studied and the different questions that it raises. The most common techniques that have previously successfully been used in the study of sRNAs are described below, starting with the *in vivo* approaches. The various

mechanisms and techniques that are discussed in this review are summarized in Fig. 1.

2. *In vivo* approaches

Imperfectly base-pairing sRNAs typically affect the translation and/or the stability of target mRNAs either positively or negatively [11] and can also directly impact transcription of a target gene [12,13]. It is thus useful, once an sRNA has been shown to directly affect the levels of a given target mRNA or protein to study which step of gene expression is affected. This is generally performed *in vivo* by assessing whether the expression of the sRNA affects that of its target gene, as measured with transcriptional or translational fusions to a reporter gene, and by looking at the mRNA levels and/ or stability.

2.1. Studying the effect of an sRNA on a target using reporter gene fusions

The general idea underlying the use of fusions is to monitor the expression of a given target gene with a read-out that is easily and precisely quantifiable. Two types of fusions exist depending on the region of the gene of interest attached to the reporter gene (Fig. 2). In the case of transcriptional fusions (or promoter fusions), the transcription signals are those of the target gene while the translation signals are from the reporter gene. Changes in the expression of a promoter fusion thus indicate changes in the transcription levels of the target gene or in the stability of the reporter mRNA. The situation is different with translational fusions (or protein fusions) where transcription and translation are driven by sequences from the target gene resulting in a chimeric protein (Fig. 2B). It is also common to use translational fusions expressed from a heterologous promoter to focus on regulatory events occurring downstream of transcription initiation. Depending on the goal, this promoter can be constitutive or inducible. The same promoter swapping approach can be performed with promoter fusions to discriminate effects at the promoter level from effects on transcription elongation or stability of the fusion mRNA.

Translational fusions have been extensively used to look at the effect of mutations in the regulatory regions located on both the trans-encoded sRNA and its target mRNA or to validate base-pairing interactions between sRNA and target mRNA via compensatory changes, but fusions can provide much more information regarding the effect of sRNAs. They can for instance allow deciphering the level of control by comparing the action of sRNAs on transcriptional and translational fusions. They can also be useful to narrow down the regions of a given gene that are required for sRNA control by constructing fusions carrying various regions of the target. In this regard, it is possible to change the promoter as already mentioned above, or to introduce deletions in the 5'UTR

or in the ORF of the target gene. Another powerful application of gene fusions is that they can be exploited to identify different factors involved in sRNA control of a given gene through using genetic screens or selections. A genetic selection aiming at identifying regulators of the ChiP chitoporin in *Salmonella* revealed that ChiP synthesis is regulated by Hfq and by the ChiX sRNA for instance, while a genetic screen led to further characterization of the system and ultimately demonstrated that levels of the ChiX sRNA are controlled by an inducible “bait target” of this sRNA [14]. In the same study, mutants were isolated in *chiP* or ChiX in cis, highlighting that genetics and gene fusions can also identify bases important for the control in the sRNA or the target.

Different reporter genes have been successfully used, with the most common ones being *lacZ*, the gene encoding β -galactosidase, and genes coding for fluorescent proteins such as *gfp*. These reporters have different advantages: *lacZ* provides a highly sensitive assay for gene expression and, because it gives rise to easily discernable phenotypes on plates, it has been widely used in genetic screens or selections [see [15] for another example], while *gfp* (or other fluorescent reporters) can be used to detect heterogeneity of gene expression in a clonal population, by FACS or microscopy, for instance [16]. Fusions have been engineered on the chromosome of several bacterial species at different loci, but they have also been expressed from plasmids in several cases [17,18].

Fusion of a reporter gene to a target gene results in the synthesis of a chimeric mRNA that might no longer be controlled by its cognate sRNA. This could be due to specific features in the structure of the chimeric mRNA which might prevent sRNA binding, or affect its translation and/or its stability. In cases where a gene fusion does not accurately report control by an sRNA, it is possible to use smaller tags instead (e.g. Flag or 6xHis tag) with the same idea that it is easy to follow gene expression [19]. Even these smaller tags can interfere with “natural” gene regulation and in this case, it is still possible to study the sRNA control by directly following the target protein or mRNA levels in response to changes in the sRNA expression, using western- or northern-blotting, respectively. In some cases, it is also possible to easily follow the levels of a protein based on its function, if it gives rise to a discernable phenotype or a quantifiable measurement for instance.

One last point that is worth mentioning here is that fusions are not restricted to the target mRNAs. Fusions to sRNA promoters have been widely used to study their transcriptional control. More importantly for the scope of this review, chimeras between different domains of sRNAs have been used to show that a limited region of an sRNA is necessary and sufficient for control [20,21].

2.2. Effect of the sRNA on mRNA levels and stability

If the expression of a gene is regulated by an sRNA, it might be interesting to determine whether the levels of mRNA transcribed from this gene are correlated with the expression of the regulatory sRNA, but also whether the stability of this mRNA is affected. The most straightforward way to assess the stability of an mRNA is to measure its decay rate after blocking its synthesis, which is commonly achieved via the addition of rifampicin to cultures. To test whether an sRNA controls the stability of a target mRNA, isogenic strains expressing or lacking this sRNA are thus typically grown up to a specific point in the growth-phase before addition of rifampicin. Total RNA is then extracted at several time-points just before and following rifampicin addition and the amount of target mRNA remaining is determined for each of these time-points, by northern-blotting for instance. The half-life of the mRNA, which is defined as the time required to decrease the amount of the mRNA 2-fold, is calculated using linear fits to a semilog plot of the hybridization data versus time.

If, in such an experiment, an sRNA affects the stability of its target, this raises several questions, and in particular (i) whether this is a primary effect of the sRNA-mRNA interaction, (ii) whether this change in stability is required for the control, and (iii) what are the RNases involved and what is their role?

Regarding the first point, it is well established that translating ribosomes protect the mRNA from ribonucleases [22–24] and decreasing the translation of an mRNA with an sRNA is thus expected to decrease its stability. Nonetheless, sRNAs can also directly target mRNA for degradation, independently of translational control. This can be assessed by determining whether degradation of the target mRNA is still sRNA dependent in an *in vitro* degradation assay, i.e. in the absence of the translation machinery (see Section 3.5). It is also possible to study *in vivo* whether the sRNA affects the stability of the target mRNA in the absence of translating ribosomes. It has been shown for instance that the SgrS sRNA induced a strong decrease in the *ptsG* mRNA levels, when the translation efficiency was reduced by a mutation in the Shine-Dalgarno sequence, or by the addition of kasugamycin, a translation initiation inhibitor [25,26]. Similarly, RyhB sRNA was found to repress *sodB* mRNA levels even when *sodB* translation was blocked by the introduction of a premature stop codon [27].

Recently, a ribosome profiling approach following the induction of the RyhB sRNA has proven useful to discriminate between targets that are regulated at the level of translation and/or at the level of stability [28]. While this does not address the question of whether changes in stability of targets are a consequence of changes in translation (for targets for which both stability and translation are affected), this study identified novel RyhB targets, some of which are regulated only at the level of translation, and this approach is likely to be applied to other sRNAs in the near future.

Importantly, even though an sRNA might change the stability and thus the levels of a target mRNA, this does not necessarily mean that these changes are required for the control of the protein levels. If we consider again the regulation of *ptsG* by SgrS for instance, the synthesis of the *ptsG*-encoded IICB^{Glc} protein is still controlled in the absence of *ptsG* degradation [25]. In contrast, the control of the surface protein A by the regulatory RNAIII in *Staphylococcus aureus* requires that the RNase III endonuclease cleaves the resulting RNAIII-mRNA duplex [29]. In other words, it is important to determine whether degradation of the mRNA is required for sRNA control of the protein level on a “case-by-case” basis.

Note that sRNAs can also increase target mRNAs stability: in several cases, sRNAs have been reported to induce changes in the processing of their targets, leading to the production of a more stable form of the transcript. Such processing events can be followed *in vivo* by analyzing the 5' or 3' ends of the target, by primer extension (for the 5' end) or RACE experiments [30,31].

Another important question to address the mechanism of control is to determine which RNase(s) are responsible for the observed change in stability or processing of the target mRNA. This can be done through a candidate approach by testing whether these changes are still observed in strains mutated for known RNases. Obvious candidates that can be tested are RNase E (in Gram-negative bacteria), RNases Y and J1 (in Gram-positive bacteria) and RNase III. These enzymes have major impact in cellular RNA turnover and have already been shown to play important roles in control mediated by sRNAs [32,33]. Note however that even if these RNases are involved in the control *in vivo*, this can be indirect, and again an *in vitro* assay is required to discriminate direct and indirect effects (see Section 3.5).

If no single RNase is identified in such an assay, one can combine mutations for various RNases as this issue could be due to (partial) functional redundancy. Alternatively, one can look for other factors involved in the control, e.g. by genetic methods,

which could reveal previously uncharacterized RNase(s) as well as other genes whose function in sRNA-mediated regulation will have to be deciphered in further detail. The role of PNPase in mediating the stability of several sRNAs was discovered by a combined genetic screen and selection for example [15].

2.3. A role for other actors in addition to the sRNA-mRNA duplex?

Other factors can be required for regulation of gene expression by sRNAs acting by base-pairing with mRNAs. For instance, in Gram-negative bacteria, the Hfq RNA chaperone has been shown to be involved in both the stabilization of the sRNAs and sRNA-mRNA duplex formation [34]. More generally, the precise understanding of the regulatory mechanism requires that these factors be identified and their action further characterized. A good example comes from the genetic study showing that the product of the *rapZ* gene was involved in the control of *glmS* by the GlmZ sRNA [35], while further work later showed that RapZ specifically directs the RNase E endonuclease to this sRNA or to the GlmY “mimic”, thus protecting GlmZ from cleavage in this latter case [36]. If we want to identify factors required for the action of an sRNA on a given mRNA, it is *a priori* possible to take advantage of genetic screens or selections to look for factors that affect control when mutated (or overproduced). A more in-depth characterization of these factors and of their mode of action is then required, possibly through *in vitro* experiments. Because sRNAs may only have a subtle effect on expression of some targets, getting such genetic screens to work can be difficult and all factors may not be identified by this approach. One can thus envision to search for partners interacting with a given sRNA by a biochemical approach in parallel [37,38]. Immunoprecipitation of RNA-binding proteins and subsequent analysis of associated RNAs has also been used to identify protein partners for sRNAs (and mRNAs) [e.g. 39–41]. How these factors intervene in the regulation then requires further analysis, and *in vitro* approaches might be particularly helpful for this purpose.

3. *In vitro* approaches

While *in vivo* results are essential and highly informative in regard to the key aspects of sRNA regulation of gene expression, the use of *in vitro* techniques is necessary to further elucidate the molecular mechanisms and to define the minimal number of actors required for control and to decipher how they act. For instance, the Hfq chaperone is required *in vivo* for the action of many sRNAs, but it is often unclear whether it is because it stabilizes them, facilitates duplex formation and/or has a more direct role in the control. This question can be addressed *in vitro*: it has been shown for instance that, in the control of *sdhC* mRNA by Spot42 sRNA, binding of the sRNA to the target mRNA recruits Hfq to the vicinity of the RBS and that Hfq is the factor that competes with ribosome binding [42]. In contrast, in many other cases, the sRNA is sufficient to compete with ribosome binding *in vitro* without addition of Hfq [e.g. 43].

It is important to emphasize that *in vitro* experiments are subject to artefacts, as the experimental conditions necessarily vary from the cellular context. Even if concentrations and ratios of the different molecules present in the *in vitro* assay are carefully chosen to reflect those encountered *in vivo* as much as possible, this change in experimental conditions, including the absence of all other cellular components, can lead to non-specific effects. *In vitro* approaches are thus usually used to follow up an *in vivo* study and must necessarily include specificity controls; in the case of sRNAs, such a control can be a mutant derivative that no longer regulates target gene expression *in vivo*. *In vitro* approaches involve working with purified and/or *in vitro* synthesized components. Proteins are most often purified from strains in which they are overexpressed, and general considerations for the preparation of RNA components are described below.

3.1. Synthesis, labeling and detection of RNA molecules for *in vitro* studies

In most cases, RNA molecules for *in vitro* studies are obtained by *in vitro* transcription using RNA polymerase from the T7 bacteriophage [44–46]. Depending on the experiment, it might be necessary to label the RNA, which is most often performed either during transcription, by including a labeled NTP in the transcription mixture, or post-transcriptionally, e.g. at the 5' end using T4 polynucleotide kinase and [³²P-γ]ATP, or at the 3' end using T4 RNA Ligase and [³²P]pCp. This labeling provides an easy detection of the full-length transcripts and cleavage products in RNA structure probing experiments or processing/cleavage assays (see below). The different RNA fragments are separated by denaturing PAGE and their sizes (as well as cleavage sites) inferred from ladders generated by partial digestion of RNA by T1 RNase, usually coupled with partial alkaline digestion. Note however that it is essential that the RNA molecules that are to be labeled are homogeneous in size, which requires a PAGE purification; for this reason direct labeling can only be used for relatively short RNAs (<200 nucleotides). Alternatively, tags have been developed that allow purification of longer transcripts with homogeneous ends [47].

Another approach that is widely used to analyze cleavages of *in vitro* (or *in vivo*) transcripts is their detection as stops to cDNA extension from a DNA primer labeled at its 5' end either with ³²P or with a fluorescent dye. The resulting cDNA fragments are subsequently separated by denaturing PAGE alongside with DNA fragments generated by sequencing reactions with the same primer. Sequencing reactions can be performed either on the *in vitro* transcripts or on the DNA template used for transcription. The sites of cleavage are inferred from the length of the resulting cDNA fragments compared to that of the DNA fragments generated by sequencing. Even if it involves an extra step prior to analysis, there are several advantages to this method. The first is that it is not restricted to the analysis of cleavage products but it also applies to the analysis of chemical modification of RNA or formation of translation initiation complex (see below). The second advantage is that RNA of any size can be probed because labeled primers can be annealed anywhere along the RNA molecule, and multiple primers can be used in separate extension reactions to detect modifications along the entire molecule. The third advantage is that there is no need to gel-purify the RNA or to label it in advance. This method will not work however if the region of interest is located very close to the 3' end of the RNA under study since cDNA extension requires annealing of a primer downstream of the concerned region.

Before starting *in vitro* experiments, which are often time-consuming and labor-intensive, it must be realized that the probability of misfolding increases dramatically for larger RNAs where several potential isoforms can exist. It is thus especially important to restrict, if possible, the region of the mRNA that will be used in the *in vitro* experiments, for instance when there are *in vivo* indications that only a given region is sufficient for control (see above). Regarding the sRNAs, the probability of misfolding is not considered to be a major issue, given their short sizes in general. Nonetheless, sRNAs have sometimes been replaced by short DNA or RNA oligonucleotides mimicking the seed-pairing region in several *in vitro* approaches, which turned out to be both easier from an experimental point of view and highly informative [21,48].

3.2. *In vitro* characterization of the sRNA-mRNA interaction

Even when one sRNA-mRNA interaction site has been validated *in vivo*, usually *via* introduction of compensatory changes that re-establish sRNA-mRNA base pairing, it might be informative to

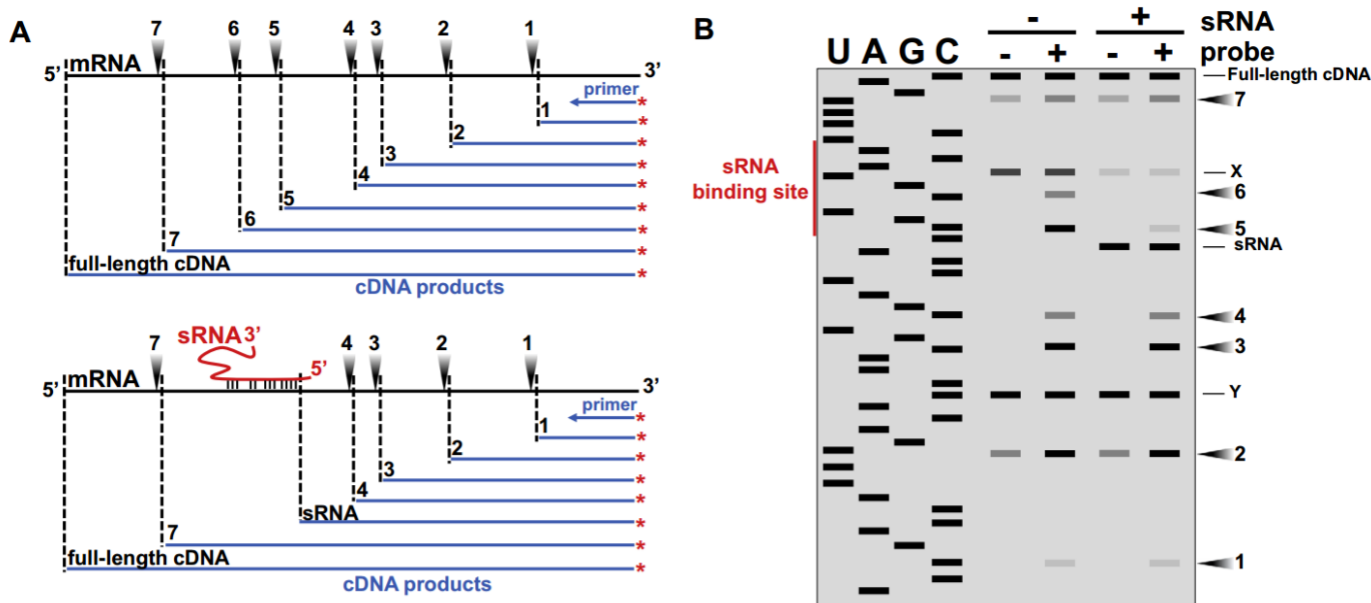


Fig. 3. Principle of the RNA structural probing experiments and analysis by RT. (A) Nucleotides reacting with the probe are indicated by black arrowheads and blue lines represent the cDNA species produced by RT of the labeled DNA primer. RT is blocked one nt downstream of a reactive residue and/or RNA structure. (B) Scheme of the fractionation of the RT products after sequencing PAGE. Nucleotides reactive to the probe present different reactivities (here, moderate for positions 1 and 7; strong for positions 2 and 6). Reactivities of positions 5 and 6 to the probe are highly dependent on binding of the sRNA, and other RT stops – potentially caused by mRNA secondary structures or cleavages – are depicted as either being sensitive to sRNA binding (X) or unchanged upon sRNA binding (Y).

determine (i) whether the interaction site is restricted to the base pairs that have been validated or is longer, and (ii) whether additional interaction sites exist. While *in vivo* data can provide partial answers to these questions, for instance by narrowing down the regions required for the control, it is also possible to precisely analyze the sRNA-mRNA duplex *in vitro*. Possibly the simplest experiment for this purpose is the analysis of reverse-transcription (RT) stops or pauses induced by the binding of the sRNA to its cognate mRNA. This analysis may provide an idea of the position of the 3' end of the duplex on the mRNA side and may also indicate whether several pairing-regions exist. However, not all sRNA-mRNA duplexes are sufficiently stable to give rise to such signals and the absence of RT-stops upon sRNA binding is thus not sufficient to rule out a potential pairing. Alternatively, a detailed view of the *in vitro* secondary structure of the sRNA and mRNA, alone or in complex, can be provided by RNA secondary structure probing, whose principle is shown in Fig. 3.

In most cases, intermolecular RNA-RNA interactions rely on nucleotides that are located in single-stranded regions, such as loops or unstructured regions. Determination of the secondary structure of either the sRNA or the mRNA is thus expected to identify the nucleotides which are potentially available for base pairing, but also those directly involved in the interaction, if probing is performed after duplex formation. RNA secondary structure probing in solution consists of monitoring the reactivity of RNA nucleotides towards different probes, such as chemical reagents or RNases that display different specificities. RNases such as T1 (cleaves after unpaired guanines), T2 (cleaves after unpaired nts, with a preference for adenines), or V1 and RNase III (both of which cleave in double-stranded regions) have been extensively used for probing structure of sRNA, mRNA or sRNA-mRNA complex [49–51]. However, nucleotides buried inside the RNA molecule are difficult for RNases to access (because of their large size) and thus interpretation of probing data with RNases has to be restricted to the positive detection of cleavages rather than the absence of cleavages. Chemical probes can also be used for RNA structure probing. Because they are much smaller than RNases, they can reach almost every part of the RNA molecule and they are thus ideally suited for

the study of RNA structure. A first group of chemical reagents is sequence-dependent and forms stable covalent adducts at accessible and reactive Watson-Crick positions with one or two of the four RNA bases. The most popular base-specific probes are DMS, which reacts with adenosine and cytosine residues [52], kethoxal, which modifies guanosine residues [53–55] and CMCT, which probes guanosine and uridine residues [56]. A second group of reagents, including lead (II) [57] or SHAPE reagents such as NMIA or 1M7 [58], targets the ribose-phosphate backbone with the advantage of being thus sequence-independent. They ultimately measure local nucleotide flexibility or dynamics with the assumption that more “flexible” regions (such as single-stranded regions) are modified more efficiently than more “rigid” regions (such as double-stranded regions). However it should be noted that in contrast to SHAPE reagents which form adducts at the 2' hydroxyl group of the ribose moiety, lead (II) cleavage of the RNA backbone has strong biases towards sites of metal coordination that may make the interpretation of secondary structure measurements ambiguous [59,60]. Finally, *in-line* probing [61] uses the structure-dependent autocleavage capability of RNA and single-stranded regions are expected to degrade over time as they are more flexible and hence unstable.

Probing methods that interrogate all nucleotides are particularly concise and highly informative because they present two advantages compared to the use of base-specific probes: (i) the reactivities of the RNA nucleotides can be measured for almost every residue in a single experiment while base-specific probes are typically limited in their range of selectivity and should be used in parallel and (ii) the reactivities can be quantitatively compared. For more detailed information about the chemistries of the various modifying reactions, the reader should refer to some excellent reviews or articles describing the use of chemical RNA structure probing techniques [58,61–65].

In addition to precisely mapping the interaction regions for a given RNA-RNA duplex, *in vitro* RNA probing can provide key insights about possible structural rearrangement of the mRNA following sRNA binding. In some cases, the existence of such alternative structures is essential for the action of the sRNA, possibly

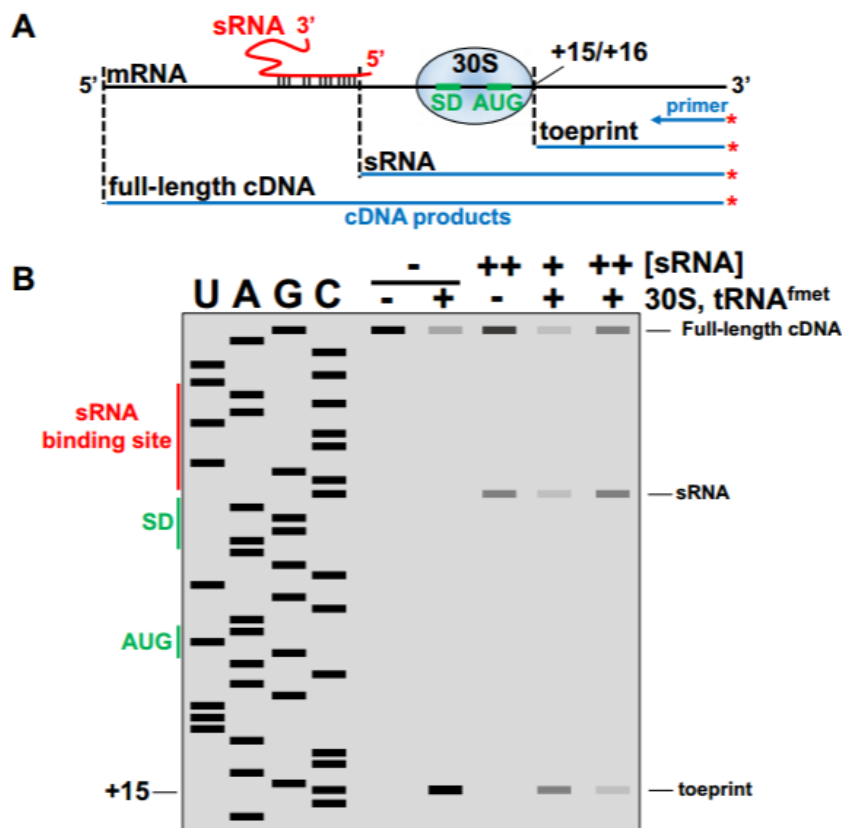


Fig. 4. Principle of the toeprinting experiment. (A) Translation signals are indicated in green, and positions of the toeprint signal are given relative to the translation start. Different cDNA produced by RT from a labeled DNA primer are shown as blue lines. (B) Scheme of the fractionation of the RT products after sequencing PAGE. In this example, sRNA binding to the mRNA inhibits the toeprint signal in a concentration-dependent manner.

because they will ultimately affect the accessibility of the RBS [66]. Similar studies have also been performed upon addition of Hfq to determine its role in mRNA or sRNA structure [67,68].

Interestingly, some of the chemical probes described above, such as DMS, lead (II) or SHAPE reagents, can enter living cells and thus be used for characterization of RNA structures *in vivo* [54,65,69,70]. RNA structure characterization has undergone critical advances recently by combining chemical probing with high-throughput sequencing, providing comprehensive insights into the RNA structure of most transcripts of a given organism in a single experiment [69,71]. The “RNA structurome” of some bacteria was for instance determined using RNase probing [72,73] and this is likely to be repeated soon in more bacterial species and/or using different probing reagents, possibly *in vivo*. It is thus advisable to take advantage of these recent data sets (if applicable) before starting RNA structure probing experiments, which can be tedious and time-consuming.

3.3. *In vitro* study of translation initiation: the toeprinting assay

Many sRNAs act at the level of translation initiation by affecting binding of the 30S ribosomal subunit to the target mRNA. Therefore, one of the first mechanisms that should be investigated, when studying sRNAs that modify translation of targets, is inhibition of ribosome binding to the mRNA. For this purpose, a straightforward experiment to perform is toeprinting. Basically, toeprinting is a primer extension inhibition assay which is used to examine the formation of the translation initiation ternary complex formed between the mRNA, the 30S ribosomal subunit and the initiator tRNA^{fMet} [74]. This assay relies on the fact that reverse transcriptase is “stopped” (usually at nts 15/16 relative to the first nt of the ORF) when it encounters a ribosome sitting at the RBS, thereby allowing

to determine the precise position of the ribosome on mRNA (Fig. 4). When a toeprint assay is performed in presence of sRNAs, those have to be incubated with the mRNA prior to the addition of the 30S ribosomal subunit and the initiator tRNA, as they are generally unable to displace the ternary initiation complex. They are thus usually added after denaturation and refolding of the mRNA [75], but can also be denatured and renatured together with the mRNA, which should favor sRNA-mRNA duplex formation. Note also that other factors, e.g. Hfq, might be required for toeprinting inhibition by sRNAs. This is either because it is directly involved in the competition with ribosome binding [42], or because it can increase the formation or the stability of the sRNA-mRNA duplex to more efficiently compete with ternary complex (our unpublished results).

Some essential controls must be included in the toeprinting assay: (i) a control performed in the absence of both the ribosome and the sRNA will determine the premature primer extension arrests that are due to stable mRNA secondary structure or cleavages; (ii) a control lacking ribosomes but containing the sRNA at the highest concentration used in the assay will determine the primer extension arrests that are due to base pairing of the sRNA with mRNA whenever detectable; (iii) a control performed using a mutant version of the sRNA, which has lost its regulatory effect *in vivo* and is expected not to affect toeprinting *in vitro*, will ensure that the observed effect is specific.

In most instances, sRNAs decrease the toeprint signal. In such cases, if the sRNA interacts with the mRNA at or close to the RBS, signal reduction is a strong indication that it acts *via* competition with the ribosome to bind the mRNA. If it interacts instead with a region located outside of the RBS, a possible scenario is that the sRNA-mRNA interaction triggers a rearrangement of the mRNA secondary structure, which will ultimately impede access of the ribosome to the RBS [66]. In this regard, probing the structure of

the mRNA following sRNA binding (see above) can be highly informative regarding the occurrence of such a structural switch. If, in contrast, the outcome of sRNA binding is an increase in the toeprint signal, it is likely that the sRNA base pairs with an mRNA region located outside of the RBS and a plausible mechanism is disruption of an inhibitory secondary structure that sequesters the RBS of the target mRNA by the sRNA. This is a classical and well-documented mechanism that is used by several sRNAs to activate gene expression [76–78]. In this scenario also, probing mRNA structure in the context of an sRNA-mRNA duplex should be informative. Finally, if the sRNA has an effect on the toeprint signal without binding at or in the vicinity of the RBS and without indication of a structural switch, alternative mechanisms have to be envisioned and further elucidated. For instance, the sRNA can instead compete with ribosomes for binding to a “ribosome standby” site located far upstream of the RBS of the target mRNA [79], or control expression of the target mRNA by translational coupling after acting on the expression of an upstream ORF [80].

3.4. *In vitro* translation experiments

Cell-free protein synthesis using bacterial extracts can be used to recapitulate control by an sRNA, as well as to study the effect of specific factors and to determine the level at which control takes place [50]. This kind of approach has become more popular with the development of an *in vitro* protein synthesis assay including T7 RNA polymerase, ribosomes and the other purified components of the *E. coli* translation machinery [81]. Although this assay is best suited to the study of the sRNA-mRNA pairs existing in *E. coli*, it might also be adaptable to the study of sRNA-mRNA pairs from other bacteria as long as the mRNA under study is translated using the *E. coli* machinery.

Using this assay, one can follow the effect of a regulatory sRNA on the expression of a given gene by analyzing the synthesis of the encoded protein either by protein gel staining, western-blotting, or ³⁵S labeling, depending on translation efficiency. As previously, an essential control that must be performed is looking at translation in the presence of a mutated sRNA that is unable to regulate the translation of the target mRNA *in vivo*.

This *in vitro* assay allows mRNA transcription and protein synthesis from a template DNA carrying a T7 promoter. If an sRNA effect on protein synthesis is detected with such a template, the use of a heterologous promoter strongly suggests that the sRNA does not act at the promoter level. It is also possible to infer whether the effect is transcriptional or translational from *in vivo* data provided by the use of fusions with reporter genes (see above). It must be realized that transcription by the T7 RNA polymerase is much faster than transcription by the *E. coli* enzyme [22], which will result in an uncoupling of transcription and translation. This might impair regulatory mechanisms that require transcription and translation to be coupled, and lead to formation of mRNA structures that differ from the native ones, possibly preventing the action of sRNAs. Alternatively, an mRNA template can be used instead. Even if it is also synthesised by the T7 RNA polymerase, it can be denatured and refolded prior to be added to this assay, which will make more unlikely the possibility of unwanted RNA structures. Another advantage of using an mRNA template is that it can help to precisely monitor mRNA levels and the sRNA-mRNA ratios used in the experiment.

If an sRNA is shown to affect the translation of a given gene in this assay, it could be due to an effect either on translation initiation or at a later step. To discriminate between these possibilities, data obtained from toeprinting experiments could be highly informative in supporting or dismissing translation initiation as the key-step of control. In the event where it seems that the sRNA controls translation elongation or termination (even though no

such case has been reported so far), it would be interesting to look for a truncated polypeptide resulting from premature translation termination or elongation pausing triggered by sRNA.

Because no ribonucleases are present in the commercialized *in vitro* translation assay, another important question this test can help address is the relative contribution of translation and mRNA degradation in the control of gene expression by an sRNA, which, as mentioned above, is extremely difficult to determine by *in vivo* analyses. Alternatively, this can also be deciphered by the reverse approach of studying how an sRNA affects the target mRNA degradation or processing *in vitro*, in the absence of translation.

3.5. Studying the degradation or processing of sRNA-mRNA duplexes *in vitro*

Whenever an sRNA is suspected to affect the stability and/or the processing of a target mRNA, it is worth undertaking an *in vitro* study of the mRNA fate, as well as that of the sRNA, in order (i) to determine whether this change in stability is a consequence of a change in translation or, in contrast, occurs in the absence of translation as well, (ii) to identify the nuclease(s) involved and possibly (iii) to uncover other mechanistic details. One can for instance follow the cleavage of a labeled target mRNA by a purified RNase, in the presence or absence of the regulatory sRNA. In this way, it was shown that the MicC sRNA induced a specific cleavage of *ompD* mRNA by the catalytic domain of RNase E after pairing within the *ompD* ORF [82]. Furthermore, the addition of MicC variants with different 5' ends in these experiments showed that the 5' monophosphorylated MicC was more efficient at inducing cleavage of *ompD* than the triphosphorylated form, suggesting that, in addition to driving the RNase E cleavage close to the site of interaction, MicC also activated the enzyme by providing a 5' monophosphorylated end. This *in vitro* assay thus confirmed that degradation of an sRNA-targeted mRNA was independent of translation and provided mechanistic insights into how the degradation takes place.

In vitro cleavage assays have also proven useful in the study of positively acting sRNAs: for instance, the activation of the sugar phosphatase encoding *yigL* gene by the SgrS sRNA is also an RNase E-dependent process [30]. In this case, SgrS binding in the upstream ORF of the *pldB-yigL* bicistronic operon was shown to stabilize the *yigL* mRNA by preventing cleavage by RNase E at a critical site. *In vitro* cleavage reactions aiming at understanding the role of RNA processing in the action of sRNAs have also been performed with purified RNA degradosomes [27], with the RNase III endonuclease [79], or with the PNPase 3'–5' exonuclease [83], and are likely to be developed for other exo- or endonucleases. As for all *in vitro* approaches, these analyses have to be interpreted with caution and results must be supported by *in vivo* data. An *in vitro* cleavage of an sRNA-mRNA duplex by RNase III could for example simply be due to the fact that this endonuclease recognizes and cleaves double-stranded structures (which is the property that is used for RNA structural probing). It is thus essential in this case to show that a deletion of the *mc* gene also affects control, or at least the levels of the target mRNA *in vivo* [see e.g. 79,84].

4. Conclusion

The different approaches discussed in this review can be generally applied to the study of the regulation of gene expression by bacterial base pairing sRNAs. So far, they have provided an in-depth understanding of the molecular mechanism underlying the action of sRNAs for a few sRNA-mRNA pairs in only a small number of bacteria. In the future, there is no doubt that their application, certainly with some adjustments, to other sRNAs from a larger

set of bacteria, will lead to uncover yet other mechanisms. It is also likely that the use of these techniques in conjunction with high-throughput sequencing, aiming for instance to the determination of RNA structures, will become more and more widespread. This will be highly interesting as the availability of more experimental data will give us a better sense of how reliable these approaches are.

The techniques described here can also be useful to study how other regulatory RNAs, including the antisense RNAs, that are widely present in all genomes, regulate gene expression. Because the perfect complementarity between an antisense RNA and its target might render the analysis more difficult, the mode of action of antisense RNAs is still extremely poorly characterized in most cases. Overall, mechanistic studies are likely to provide exciting results in the years to come.

Funding

Work in our lab is supported by the CNRS, ANR grant (ANR-14-CE10-0004-01) to MG and by the “Initiative d’Excellence” program from the French State (Grant “DYNAMO”, ANR-11-LABX-0011). JJ is the recipient of a “Labex Dynamo” Ph.D. fellowship.

Acknowledgements

We are grateful to Ciaran Condon, Sylvain Durand and Mathias Springer for critical reading of the manuscript.

References

- J.T. Wade, D.C. Grainger, Pervasive transcription: illuminating the dark matter of bacterial transcriptomes, *Nat. Rev. Microbiol.* 12 (2014) 647–653.
- T.R. Cech, J.A. Steitz, The noncoding RNA revolution—trashing old rules to forge new ones, *Cell* 157 (2014) 77–94.
- E.G. Wagner, P. Romby, Small RNAs in bacteria and archaea: who they are, what they do, and how they do it, *Adv. Genet.* 90 (2015) 133–208.
- M.S. Guo, T.B. Updegrove, E.B. Gogol, S.A. Shabalina, C.A. Gross, et al., MicL, a new sigmaE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein, *Genes Dev.* 28 (2014) 1620–1634.
- Y. Chao, K. Papenfort, R. Reinhardt, C.M. Sharma, J. Vogel, An atlas of Hfq-bound transcripts reveals 3′ UTRs as a genomic reservoir of regulatory small RNAs, *EMBO J.* 31 (2012) 4005–4019.
- Y. Chao, J. Vogel, A 3′ UTR-derived small RNA provides the regulatory noncoding arm of the inner membrane stress response, *Mol. Cell* 61 (2016) 352–363.
- D. Lalaouna, M.C. Carrier, S. Semsey, J.S. Brouard, J. Wang, et al., A 3′ external transcribed spacer in a tRNA transcript acts as a sponge for small RNAs to prevent transcriptional noise, *Mol. Cell* 58 (2015) 393–405.
- Y. Hao, T.B. Updegrove, N.N. Livingston, G. Storz, Protection against deleterious nitrogen compounds: role of sigmaS-dependent small RNAs encoded adjacent to *sdhA*, *Nucleic Acids Res.* 44 (2016) 6935–6948.
- B.J. Tucker, R.R. Breaker, Riboswitches as versatile gene control elements, *Curr. Opin. Struct. Biol.* 15 (2005) 342–348.
- L.S. Waters, G. Storz, Regulatory RNAs in bacteria, *Cell* 136 (2009) 615–628.
- S. Melamed, A. Peer, R. Faigenbaum-Romm, Y.E. Gatt, N. Reiss, et al., Global Mapping of Small RNA-Target Interactions in Bacteria, *Mol. Cell* 63 (2016) 884–897.
- X. Wang, S.C. Ji, H.J. Jeon, Y. Lee, H.M. Lim, Two-level inhibition of galK expression by Spot42: degradation of mRNA mK2 and enhanced transcription termination before the galK gene, *Proc. Natl. Acad. Sci. U.S.A.* 112 (2015) 7581–7586.
- N. Sedlyarova, I. Shamovsky, B.K. Bharati, V. Epshtein, J. Chen, sRNA-Mediated Control of Transcription Termination in *E. coli*, *Cell* 167 (2016) 111–121 e13.
- N. Figueroa-Bossi, M. Valentini, L. Malleret, F. Fiorini, L. Bossi, Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target, *Genes Dev.* 23 (2009) 2004–2015.
- N. De Lay, S. Gottesman, Role of polynucleotide phosphorylase in sRNA function in *Escherichia coli*, *RNA* 17 (2011) 1172–1189.
- N. Grantcharova, V. Peters, C. Monteiro, K. Zakikhany, U. Romling, Bistable expression of CsgD in biofilm development of *Salmonella enterica* serovar typhimurium, *J. Bacteriol.* 192 (2010) 456–466.
- J.H. Urban, J. Vogel, Translational control and target recognition by *Escherichia coli* small RNAs in vivo, *Nucleic Acids Res.* 35 (2007) 1018–1037.
- C.P. Corcoran, D. Podkaminski, K. Papenfort, J.H. Urban, J.C. Hinton, et al., Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA, *Mol. Microbiol.* 84 (2012) 428–445.
- M. Guillier, S. Gottesman, The 5′ end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator, *Nucleic Acids Res.* 36 (2008) 6781–6794.
- K. Papenfort, M. Bouvier, F. Mika, C.M. Sharma, J. Vogel, Evidence for an autonomous 5′ target recognition domain in an Hfq-associated small RNA, *Proc. Natl. Acad. Sci. U.S.A.* 107 (2010) 20435–20440.
- M. Bouvier, C.M. Sharma, F. Mika, K.H. Nierhaus, J. Vogel, Small RNA binding to 5′ mRNA coding region inhibits translational initiation, *Mol. Cell* 32 (2008) 827–837.
- O.V. Makarova, E.M. Makarov, R. Sousa, M. Dreyfus, Transcribing of *Escherichia coli* genes with mutant T7 RNA polymerases: stability of lacZ mRNA inversely correlates with polymerase speed, *Proc. Natl. Acad. Sci. U.S.A.* 92 (1995) 12250–12254.
- A. Deana, J.G. Belasco, Lost in translation: the influence of ribosomes on bacterial mRNA decay, *Genes Dev.* 19 (2005) 2526–2533.
- M. Dreyfus, Killer and protective ribosomes, *Prog. Mol. Biol. Transl. Sci.* 85 (2009) 423–466.
- T. Morita, Y. Mochizuki, H. Aiba, Translational repression is sufficient for gene silencing by bacterial small noncoding RNAs in the absence of mRNA destruction, *Proc. Natl. Acad. Sci. U.S.A.* 103 (2006) 4858–4863.
- H. Kawamoto, T. Morita, A. Shimizu, T. Inada, H. Aiba, Implications of membrane localization of target mRNA in the action of a small RNA: mechanism of post-transcriptional regulation of glucose transporter in *Escherichia coli*, *Genes Dev.* 19 (2005) 328–338.
- K. Prevost, G. Desnoyers, J.F. Jacques, F. Lavoie, E. Masse, Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage, *Genes Dev.* 25 (2011) 385–396.
- J. Wang, W. Rennie, C. Liu, C.S. Carmack, K. Prevost, et al., Identification of bacterial sRNA regulatory targets using ribosome profiling, *Nucleic Acids Res.* 43 (2015) 10308–10320.
- E. Huntzinger, S. Boisset, C. Saveanu, Y. Benito, T. Geissmann, et al., *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate spa gene expression, *EMBO J.* 24 (2005) 824–835.
- K. Papenfort, Y. Sun, M. Miyakoshi, C.K. Vanderpool, J. Vogel, Small RNA-mediated activation of sugar phosphatase mRNA regulates glucose homeostasis, *Cell* 153 (2013) 426–437.
- N. Obana, Y. Shirahama, K. Abe, K. Nakamura, Stabilization of *Clostridium perfringens* collagenase mRNA by VR-RNA-dependent cleavage in 5′ leader sequence, *Mol. Microbiol.* 77 (2010) 1416–1428.
- D. Lalaouna, M. Simoneau-Roy, D. Lafontaine, E. Masse, Regulatory RNAs and target mRNA decay in prokaryotes, *Biochim. Biophys. Acta* 2013 (1829) 742–747.
- S. Durand, A. Tomasini, F. Braun, C. Condon, P. Romby, SRNA and mRNA turnover in Gram-positive bacteria, *FEMS Microbiol. Rev.* 39 (2015) 316–330.
- J. Vogel, B.F. Luisi, Hfq and its constellation of RNA, *Nat. Rev. Microbiol.* 9 (2011) 578–589.
- F. Kalamorz, B. Reichenbach, W. Marz, B. Rak, B. Gorke, Feedback control of glucosamine-6-phosphate synthase GlmS expression depends on the small RNA GlmZ and involves the novel protein YhbJ in *Escherichia coli*, *Mol. Microbiol.* 65 (2007) 1518–1533.
- Y. Gopel, K. Papenfort, B. Reichenbach, J. Vogel, B. Gorke, Targeted decay of a regulatory small RNA by an adaptor protein for RNase E and counteraction by an anti-adaptor RNA, *Genes Dev.* 27 (2013) 552–564.
- N. Said, R. Rieder, R. Hurwitz, J. Deckert, H. Urlaub, et al., In vivo expression and purification of aptamer-tagged small RNA regulators, *Nucleic Acids Res.* 37 (2009) e133.
- A. Smirnov, K.U. Forstner, E. Holmqvist, A. Otto, R. Günster, et al., Grad-seq guides the discovery of ProQ as a major small RNA binding protein, *Proc. Natl. Acad. Sci. U.S.A.* (2016), in press. <http://dx.doi.org/10.1073/pnas.1609981113>.
- A. Zhang, D.J. Schu, B.C. Tjaden, G. Storz, S. Gottesman, Mutations in interaction surfaces differentially impact *E. coli* Hfq association with small RNAs and their mRNA targets, *J. Mol. Biol.* 425 (2013) 3678–3697.
- E. Holmqvist, P.R. Wright, L. Li, T. Bischler, L. Barquist, et al., Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo, *EMBO J.* 35 (2016) 991–1011.
- L. Attaiach, A. Boughammoura, C. Brochier-Armanet, O. Allatif, F. Peillard-Fiorente, et al., Silencing of natural transformation by an RNA chaperone and a multitarget small RNA, *Proc. Natl. Acad. Sci. U.S.A.* 113 (2016) 8813–8818.
- G. Desnoyers, E. Masse, Noncanonical repression of translation initiation through small RNA recruitment of the RNA chaperone Hfq, *Genes Dev.* 26 (2012) 726–739.
- A. Coornaert, C. Chiaruttini, M. Springer, M. Guillier, Post-transcriptional control of the *Escherichia coli* PhoQ-PhoP two-component system by multiple sRNAs involves a novel pairing region of GcvB, *PLoS Genet.* 9 (2013) e1003156.
- P. Davanloo, A.H. Rosenberg, J.J. Dunn, F.W. Studier, Cloning and expression of the gene for bacteriophage T7 RNA polymerase, *Proc. Natl. Acad. Sci. U.S.A.* 81 (1984) 2035–2039.
- S. Tabor, C.C. Richardson, A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes, *Proc. Natl. Acad. Sci. U.S.A.* 82 (1985) 1074–1078.
- C. Cazenave, O.C. Uhlenbeck, RNA template-directed RNA synthesis by T7 RNA polymerase, *Proc. Natl. Acad. Sci. U.S.A.* 91 (1994) 6972–6976.
- G. Di Tomasso, A. Salvail-Lacoste, J. Bouvette, J.G. Omichinski, P. Legault, Affinity purification of in vitro transcribed RNA with homogeneous ends using a 3′-ARiBo tag, *Methods Enzymol.* 549 (2014) 49–84.
- S. Chen, A. Zhang, L.B. Blyn, G. Storz, MicC, a second small-RNA regulator of *Omp* protein expression in *Escherichia coli*, *J. Bacteriol.* 186 (2004) 6689–6697.

- [49] C. Romilly, C. Lays, A. Tomasini, I. Caldelari, Y. Benito, et al., A non-coding RNA promotes bacterial persistence and decreases virulence by regulating a regulator in *Staphylococcus aureus*, *PLoS Pathog.* 10 (2014) e1003979.
- [50] E. Holmqvist, J. Reimegard, M. Sterk, N. Grantcharova, U. Romling, et al., Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis, *EMBO J.* 29 (2010) 1840–1850.
- [51] A. Eyraud, P. Tattevin, S. Chabelskaya, B. Felden, A small RNA controls a protein regulator involved in antibiotic resistance in *Staphylococcus aureus*, *Nucleic Acids Res.* 42 (2014) 4892–4905.
- [52] D.A. Peattie, W. Gilbert, Chemical probes for higher-order structure in RNA, *Proc. Natl. Acad. Sci. U.S.A.* 77 (1980) 4679–4682.
- [53] R. Shapiro, J. Hachmann, The reaction of guanine derivatives with 1,2-dicarbonyl compounds, *Biochemistry* 5 (1966) 2799–2807.
- [54] R. Shapiro, B.I. Cohen, S.J. Shiuey, H. Maurer, On the reaction of guanine with glyoxal, pyruvaldehyde, and kethoxal, and the structure of the acylguanines. A new synthesis of N2-alkylguanines, *Biochemistry* 8 (1969) 238–245.
- [55] H.F. Noller, J.B. Chaires, Functional modification of 16S ribosomal RNA by kethoxal, *Proc. Natl. Acad. Sci. U.S.A.* 69 (1972) 3115–3118.
- [56] N.W. Ho, P.T. Gilham, The reversible chemical modification of uracil, thymine, and guanine nucleotides and the modification of the action of ribonuclease on ribonucleic acid, *Biochemistry* 6 (1967) 3632–3639.
- [57] M. Lindell, P. Romby, E.G. Wagner, Lead(II) as a probe for investigating RNA structure in vivo, *RNA* 8 (2002) 534–541.
- [58] E.J. Merino, K.A. Wilkinson, J.L. Coughlan, K.M. Weeks, RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE), *J. Am. Chem. Soc.* 127 (2005) 4223–4231.
- [59] M. Forconi, D. Herschlag, Metal ion-based RNA cleavage as a structural probe, *Methods Enzymol.* 468 (2009) 91–106.
- [60] R.C. Spitale, P. Crisalli, R.A. Flynn, E.A. Torre, E.T. Kool, et al., RNA SHAPE analysis in living cells, *Nat. Chem. Biol.* 9 (2013) 18–20.
- [61] E.E. Regulski, R.R. Breaker, In-line probing analysis of riboswitches, *Methods Mol. Biol.* 419 (2008) 53–67.
- [62] C. Brunel, P. Romby, Probing RNA structure and RNA-ligand complexes with chemical probes, *Methods Enzymol.* 318 (2000) 3–21.
- [63] C. Chiaruttini, F. Allemand, M. Springer, Structural probing of RNA thermosensors, *Methods Mol. Biol.* 540 (2009) 233–245.
- [64] C. Ehresmann, F. Baudin, M. Mougél, P. Romby, J.-P. Ebel, et al., Probing the structure of RNAs in solution, *Nucleic Acids Res.* 15 (1987) 9109–9128.
- [65] K.M. Weeks, Advances in RNA structure analysis by chemical probing, *Curr. Opin. Struct. Biol.* 20 (2010) 295–304.
- [66] N. Heidrich, I. Moll, S. Brantl, In vitro analysis of the interaction between the small RNA SR1 and its primary target *ahrC* mRNA, *Nucleic Acids Res.* 35 (2007) 4331–4346.
- [67] Y. Peng, J.E. Curtis, X. Fang, S.A. Woodson, Structural model of an mRNA in complex with the bacterial chaperone Hfq, *Proc. Natl. Acad. Sci. U.S.A.* 111 (2014) 17134–17139.
- [68] V. Bordeau, B. Felden, 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 Res.* 42 (2014) 4682–4696.
- [69] P. Ge, S. Zhang, Computational analysis of RNA structures with chemical probing data, *Methods* 79–80 (2015) 60–66.
- [70] K.M. Weeks, D.M. Mauger, Exploring RNA structural codes with SHAPE chemistry, *Acc. Chem. Res.* 44 (2011) 1280–1291.
- [71] K.E. Watters, A.M. Yu, E.J. Strobel, A.H. Settle, J.B. Lucks, Characterizing RNA structures in vitro and in vivo with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq), *Methods* 103 (2016) 34–48.
- [72] C. Del Campo, A. Bartholomäus, I. Fedyunin, Z. Ignatova, Secondary structure across the bacterial transcriptome reveals versatile roles in mRNA regulation and function, *PLoS Genet.* 11 (2015) e1005613.
- [73] F. Righetti, A.M. Nuss, C. Twittenhoff, S. Beele, K. Urban, et al., Temperature-responsive in vitro RNA structure of *Yersinia pseudotuberculosis*, *Proc. Natl. Acad. Sci. U.S.A.* 113 (2016) 7237–7242.
- [74] D. Hartz, D.S. McPheeters, R. Traut, L. Gold, Extension inhibition analysis of translation initiation complexes, *Methods Enzymol.* 164 (1988) 419–425.
- [75] P. Fechter, C. Chevalier, G. Yusupova, M. Yusupov, P. Romby, et al., Ribosomal initiation complexes probed by toeprinting and effect of trans-acting translational regulators in bacteria, *Methods Mol. Biol.* 540 (2009) 247–263.
- [76] N. Majdalani, C. Cuning, D. Sledjeski, T. Elliott, S. Gottesman, DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription, *Proc. Natl. Acad. Sci. U.S.A.* 95 (1998) 12462–12467.
- [77] N. Majdalani, D. Hernandez, S. Gottesman, Regulation and mode of action of the second small RNA activator of RpoS translation, RprA, *Mol. Microbiol.* 46 (2002) 813–826.
- [78] L. Archambault, J. Linscott, N. Swerdlow, K. Boyland, E. Riley, et al., Translational efficiency of rpoS mRNA from *Borrelia burgdorferi*: effects of the length and sequence of the mRNA leader region, *Biochem. Biophys. Res. Commun.* 433 (2013) 73–78.
- [79] F. Darfeuille, C. Unoson, J. Vogel, E.G. Wagner, An antisense RNA inhibits translation by competing with standby ribosomes, *Mol. Cell* 26 (2007) 381–392.
- [80] B. Vecerek, I. Moll, U. Blasi, Control of Fur synthesis by the non-coding RNA RyhB and iron-responsive decoding, *EMBO J.* 26 (2007) 965–975.
- [81] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, et al., Cell-free translation reconstituted with purified components, *Nat. Biotechnol.* 19 (2001) 751–755.
- [82] K.J. Bandyra, N. Said, V. Pfeiffer, M.W. Gorna, J. Vogel, et al., The seed region of a small RNA drives the controlled destruction of the target mRNA by the endoribonuclease RNase E, *Mol. Cell* 47 (2012) 943–953.
- [83] K.J. Bandyra, D. Sinha, J. Syrjänen, B.F. Luisi, N.R. De Lay, The ribonuclease polynucleotide phosphorylase can interact with small regulatory RNAs in both protective and degradative modes, *RNA* 22 (2016) 360–372.
- [84] J. Vogel, L. Argaman, E.G. Wagner, S. Altuvia, The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide, *Curr. Biol.* 14 (2004) 2271–2276.
- [85] A. Huttenhofer, H.F. Noller, Footprinting mRNA-ribosome complexes with chemical probes, *EMBO J.* 13 (1994) 3892–3901.
- [86] P. Mandin, S. Gottesman, Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA, *EMBO J.* 29 (2010) 3094–3107.
- [87] C.A. McCullen, J.N. Benhammou, N. Majdalani, S. Gottesman, Mechanism of positive regulation by DsrA and RprA small noncoding RNAs: pairing increases translation and protects rpoS mRNA from degradation, *J. Bacteriol.* 192 (2010) 5559–5571.
- [88] T. Song, F. Mika, B. Lindmark, Z. Liu, S. Schild, et al., A new *Vibrio cholerae* sRNA modulates colonization and affects release of outer membrane vesicles, *Mol. Microbiol.* 70 (2008) 100–111.
- [89] C. Chevalier, S. Boisset, C. Romilly, B. Masquida, P. Fechter, et al., *Staphylococcus aureus* RNAIII binds to two distant regions of *coa* mRNA to arrest translation and promote mRNA degradation, *PLoS Pathog.* 6 (2010) e1000809.
- [90] V. Pfeiffer, K. Papenfort, S. Lucchini, J.C. Hinton, J. Vogel, Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation, *Nat. Struct. Mol. Biol.* 16 (2009) 840–846.

Annexe 2 : Composition détaillée des gels SDS-PAGE « Phos-tag »

Gel de « stacking »

Tris-HCl 1M pH 6,8	312.5 µL
Acrylamide-bisacrylamide 40 % (19 : 1)	250 µL
SDS 20 %	12.5 µL
APS 10 %	25 µL
TEMED	2.5 µL
Eau stérile	1.9 mL

➔ Polymérisation en 30 min à température ambiante

Gel de séparation

Tris-HCl 1M pH 8,8	1.5 mL
Acrylamide-bisacrylamide 40 % (19 : 1)	1 mL
Solution Phos-tag	20 µL
MnCl ₂ 10 mM	20 µL
SDS 20 %	20 µL
APS 10 %	40 µL
TEMED	4 µL
Eau stérile	1.4 mL

➔ Polymérisation en 30 min à température ambiante

6. Bibliographie

- Abraham, C. and J. H. Cho (2007). "Bugging of the intestinal mucosa." *N Engl J Med* **357**(7): 708-710.
- Abraham, J. M., C. S. Freitag, et al. (1985). "An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*." *Proc Natl Acad Sci U S A* **82**(17): 5724-5727.
- Abraham, S. N., J. P. Babu, et al. (1985). "Protection against *Escherichia coli*-induced urinary tract infections with hybridoma antibodies directed against type 1 fimbriae or complementary D-mannose receptors." *Infect Immun* **48**(3): 625-628.
- Al Ibrahim, N. and J. Green (2017). "Regulation of the *Escherichia coli* ydhY-T operon in the presence of alternative electron acceptors." *Microbiology*.
- Allemand, F., J. Haentjens, et al. (2007). "*Escherichia coli* ribosomal protein L20 binds as a single monomer to its own mRNA bearing two potential binding sites." *Nucleic Acids Res* **35**(9): 3016-3031.
- Alon, U. (2007). "Network motifs: theory and experimental approaches." *Nat Rev Genet* **8**(6): 450-461.
- Altuvia, S. and E. G. Wagner (2000). "Switching on and off with RNA." *Proc Natl Acad Sci U S A* **97**(18): 9824-9826.
- Altuvia, S., D. Weinstein-Fischer, et al. (1997). "A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator." *Cell* **90**(1): 43-53.
- Altuvia, S., A. Zhang, et al. (1998). "The *Escherichia coli* OxyS regulatory RNA represses fhlA translation by blocking ribosome binding." *EMBO J* **17**(20): 6069-6075.
- Ambrose, N. S., M. Johnson, et al. (1984). "Incidence of pathogenic bacteria from mesenteric lymph nodes and ileal serosa during Crohn's disease surgery." *Br J Surg* **71**(8): 623-625.
- Andersen, J., N. Delihias, et al. (1987). "The isolation and characterization of RNA coded by the micF gene in *Escherichia coli*." *Nucleic Acids Res* **15**(5): 2089-2101.
- Argaman, L., M. Elgrably-Weiss, et al. (2012). "RelA protein stimulates the activity of RyhB small RNA by acting on RNA-binding protein Hfq." *Proc Natl Acad Sci U S A* **109**(12): 4621-4626.
- Argaman, L., R. Hershberg, et al. (2001). "Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*." *Curr Biol* **11**(12): 941-950.
- Arraiano, C. M., J. M. Andrade, et al. (2010). "The critical role of RNA processing and degradation in the control of gene expression." *FEMS Microbiol Rev* **34**(5): 883-923.
- Artsimovitch, I. and R. Landick (2002). "The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand." *Cell* **109**(2): 193-203.
- Baba, T., T. Ara, et al. (2006). "Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection." *Mol Syst Biol* **2**: 2006 0008.
- Babitzke, P. and S. R. Kushner (1991). "The Ams (altered mRNA stability) protein and ribonuclease E are encoded by the same structural gene of *Escherichia coli*." *Proc Natl Acad Sci U S A* **88**(1): 1-5.
- Bak, G., J. Lee, et al. (2015). "Identification of novel sRNAs involved in biofilm formation, motility, and fimbriae formation in *Escherichia coli*." *Sci Rep* **5**: 15287.
- Baker, C. S., L. A. Eory, et al. (2007). "CsrA inhibits translation initiation of *Escherichia coli* hfq by binding to a single site overlapping the Shine-Dalgarno sequence." *J Bacteriol* **189**(15): 5472-5481.
- Baker, C. S., I. Morozov, et al. (2002). "CsrA regulates glycogen biosynthesis by preventing translation of glgC in *Escherichia coli*." *Mol Microbiol* **44**(6): 1599-1610.
- Ban, N., P. Nissen, et al. (2000). "The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution." *Science* **289**(5481): 905-920.
- Bandyra, K. J., N. Said, et al. (2012). "The seed region of a small RNA drives the controlled destruction of the target mRNA by the endoribonuclease RNase E." *Mol Cell* **47**(6): 943-953.
- Barbieri, C. M., T. Wu, et al. (2013). "Comprehensive analysis of OmpR phosphorylation, dimerization, and DNA binding supports a canonical model for activation." *J Mol Biol* **425**(10): 1612-1626.
- Bardwell, J. C., P. Regnier, et al. (1989). "Autoregulation of RNase III operon by mRNA processing." *EMBO J* **8**(11): 3401-3407.

- Barnich, N., J. Boudeau, et al. (2003). "Regulatory and functional co-operation of flagella and type 1 pili in adhesive and invasive abilities of AIEC strain LF82 isolated from a patient with Crohn's disease." *Mol Microbiol* **48**(3): 781-794.
- Barnich, N., M. A. Bringer, et al. (2004). "Involvement of lipoprotein NlpI in the virulence of adherent invasive Escherichia coli strain LF82 isolated from a patient with Crohn's disease." *Infect Immun* **72**(5): 2484-2493.
- Barnich, N., F. A. Carvalho, et al. (2007). "CEACAM6 acts as a receptor for adherent-invasive E. coli, supporting ileal mucosa colonization in Crohn disease." *J Clin Invest* **117**(6): 1566-1574.
- Barrett, J. F. and J. A. Hoch (1998). "Two-component signal transduction as a target for microbial anti-infective therapy." *Antimicrob Agents Chemother* **42**(7): 1529-1536.
- Barrick, D., K. Villanueva, et al. (1994). "Quantitative analysis of ribosome binding sites in E.coli." *Nucleic Acids Res* **22**(7): 1287-1295.
- Batchelor, E. and M. Goulian (2003). "Robustness and the cycle of phosphorylation and dephosphorylation in a two-component regulatory system." *Proc Natl Acad Sci U S A* **100**(2): 691-696.
- Beier, D. and R. Frank (2000). "Molecular characterization of two-component systems of Helicobacter pylori." *J Bacteriol* **182**(8): 2068-2076.
- Beisel, C. L. and G. Storz (2010). "Base pairing small RNAs and their roles in global regulatory networks." *FEMS Microbiol Rev* **34**(5): 866-882.
- Beisel, C. L. and G. Storz (2011). "The base-pairing RNA spot 42 participates in a multioutput feedforward loop to help enact catabolite repression in Escherichia coli." *Mol Cell* **41**(3): 286-297.
- Bell, L. C., M. D. Page, et al. (1993). "Insertion of transposon Tn5 into a structural gene of the membrane-bound nitrate reductase of Thiophaera pantotropha results in anaerobic overexpression of periplasmic nitrate reductase activity." *J Gen Microbiol* **139**(12): 3205-3214.
- Berg, B. L. and V. Stewart (1990). "Structural genes for nitrate-inducible formate dehydrogenase in Escherichia coli K-12." *Genetics* **125**(4): 691-702.
- Bernardini, M. L., A. Fontaine, et al. (1990). "The two-component regulatory system ompR-envZ controls the virulence of Shigella flexneri." *J Bacteriol* **172**(11): 6274-6281.
- Blanc-Potard, A. B. and E. A. Groisman (1997). "The Salmonella selC locus contains a pathogenicity island mediating intramacrophage survival." *EMBO J* **16**(17): 5376-5385.
- Bongaerts, J., S. Zoske, et al. (1995). "Transcriptional regulation of the proton translocating NADH dehydrogenase genes (nuoA-N) of Escherichia coli by electron acceptors, electron donors and gene regulators." *Mol Microbiol* **16**(3): 521-534.
- Bonnefoy, V. and J. A. DeMoss (1992). "Identification of functional cis-acting sequences involved in regulation of narK gene expression in Escherichia coli." *Mol Microbiol* **6**(23): 3595-3602.
- Bordeau, V. and B. Felden (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 Res* **42**(7): 4682-4696.
- Boudeau, J., N. Barnich, et al. (2001). "Type 1 pili-mediated adherence of Escherichia coli strain LF82 isolated from Crohn's disease is involved in bacterial invasion of intestinal epithelial cells." *Mol Microbiol* **39**(5): 1272-1284.
- Boudeau, J., A. L. Glasser, et al. (1999). "Invasive ability of an Escherichia coli strain isolated from the ileal mucosa of a patient with Crohn's disease." *Infect Immun* **67**(9): 4499-4509.
- Boudry, P., C. Gracia, et al. (2014). "Pleiotropic role of the RNA chaperone protein Hfq in the human pathogen Clostridium difficile." *J Bacteriol* **196**(18): 3234-3248.
- Boudvillain, M., N. Figueroa-Bossi, et al. (2013). "Terminator still moving forward: expanding roles for Rho factor." *Curr Opin Microbiol* **16**(2): 118-124.
- Bouloc, P. and F. Repoila (2016). "Fresh layers of RNA-mediated regulation in Gram-positive bacteria." *Curr Opin Microbiol* **30**: 30-35.
- Bouvier, M. and A. J. Carpousis (2011). "A tale of two mRNA degradation pathways mediated by RNase E." *Mol Microbiol* **82**(6): 1305-1310.
- Bremer, H. and P. P. Dennis (1996). *Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates*. Washington D.C., ASM Press.

- Brendel, V., G. H. Hamm, et al. (1986). "Terminators of transcription with RNA polymerase from *Escherichia coli*: what they look like and how to find them." *J Biomol Struct Dyn* **3**(4): 705-723.
- Brill, J. A., C. Quinlan-Walsh, et al. (1988). "Fine-structure mapping and identification of two regulators of capsule synthesis in *Escherichia coli* K-12." *J Bacteriol* **170**(6): 2599-2611.
- Bringer, M. A., E. Billard, et al. (2012). "Replication of Crohn's disease-associated AIEC within macrophages is dependent on TNF-alpha secretion." *Lab Invest* **92**(3): 411-419.
- Bringer, M. A., A. L. Glasser, et al. (2006). "The Crohn's disease-associated adherent-invasive *Escherichia coli* strain LF82 replicates in mature phagolysosomes within J774 macrophages." *Cell Microbiol* **8**(3): 471-484.
- Brosse, A., A. Korobeinikova, et al. (2016). "Unexpected properties of sRNA promoters allow feedback control via regulation of a two-component system." *Nucleic Acids Res* **44**(20): 9650-9666.
- Browning, D. F. and S. J. Busby (2016). "Local and global regulation of transcription initiation in bacteria." *Nat Rev Microbiol* **14**(10): 638-650.
- Browning, D. F., J. A. Cole, et al. (2000). "Suppression of FNR-dependent transcription activation at the *Escherichia coli* nir promoter by Fis, IHF and H-NS: modulation of transcription initiation by a complex nucleo-protein assembly." *Mol Microbiol* **37**(5): 1258-1269.
- Browning, D. F., J. A. Cole, et al. (2004). "Transcription activation by remodelling of a nucleoprotein assembly: the role of NarL at the FNR-dependent *Escherichia coli* nir promoter." *Mol Microbiol* **53**(1): 203-215.
- Bury-Mone, S., Y. Nomane, et al. (2009). "Global analysis of extracytoplasmic stress signaling in *Escherichia coli*." *PLoS Genet* **5**(9): e1000651.
- Butala, M., S. J. Busby, et al. (2009). "DNA sampling: a method for probing protein binding at specific loci on bacterial chromosomes." *Nucleic Acids Res* **37**(5): e37.
- Cabrera, J. E. and D. J. Jin (2003). "The distribution of RNA polymerase in *Escherichia coli* is dynamic and sensitive to environmental cues." *Mol Microbiol* **50**(5): 1493-1505.
- Cai, S. J. and M. Inouye (2002). "EnvZ-OmpR interaction and osmoregulation in *Escherichia coli*." *J Biol Chem* **277**(27): 24155-24161.
- Cai, S. J. and M. Inouye (2003). "Spontaneous subunit exchange and biochemical evidence for trans-autophosphorylation in a dimer of *Escherichia coli* histidine kinase (EnvZ)." *J Mol Biol* **329**(3): 495-503.
- Capra, E. J. and M. T. Laub (2012). "Evolution of two-component signal transduction systems." *Annu Rev Microbiol* **66**: 325-347.
- Cavicchioli, R., I. Schroder, et al. (1995). "The NarX and NarQ sensor-transmitter proteins of *Escherichia coli* each require two conserved histidines for nitrate-dependent signal transduction to NarL." *J Bacteriol* **177**(9): 2416-2424.
- Chao, Y., K. Papenfort, et al. (2012). "An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs." *EMBO J* **31**(20): 4005-4019.
- Chao, Y. and J. Vogel (2016). "A 3' UTR-Derived Small RNA Provides the Regulatory Noncoding Arm of the Inner Membrane Stress Response." *Mol Cell* **61**(3): 352-363.
- Chapman-Kiddell, C. A., P. S. Davies, et al. (2010). "Role of diet in the development of inflammatory bowel disease." *Inflamm Bowel Dis* **16**(1): 137-151.
- Chaverroche, M. K., J. M. Ghigo, et al. (2000). "A rapid method for efficient gene replacement in the filamentous fungus *Aspergillus nidulans*." *Nucleic Acids Res* **28**(22): E97.
- Chavez, R. G., A. F. Alvarez, et al. (2010). "The physiological stimulus for the BarA sensor kinase." *J Bacteriol* **192**(7): 2009-2012.
- Chen, H., M. Bjerknes, et al. (1994). "Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs." *Nucleic Acids Res* **22**(23): 4953-4957.
- Chen, S., A. Zhang, et al. (2004). "MicC, a second small-RNA regulator of Omp protein expression in *Escherichia coli*." *J Bacteriol* **186**(20): 6689-6697.
- Cheng, J., A. M. Palva, et al. (2013). "Contribution of the intestinal microbiota to human health: from birth to 100 years of age." *Curr Top Microbiol Immunol* **358**: 323-346.

- Cheng, Z. F. and M. P. Deutscher (2005). "An important role for RNase R in mRNA decay." Mol Cell **17**(2): 313-318.
- Cheng, Z. F., Y. Zuo, et al. (1998). "The vacB gene required for virulence in *Shigella flexneri* and *Escherichia coli* encodes the exoribonuclease RNase R." J Biol Chem **273**(23): 14077-14080.
- Cheung, J. and W. A. Hendrickson (2009). "Structural analysis of ligand stimulation of the histidine kinase NarX." Structure **17**(2): 190-201.
- Chiang, R. C., R. Cavicchioli, et al. (1992). "Identification and characterization of narQ, a second nitrate sensor for nitrate-dependent gene regulation in *Escherichia coli*." Mol Microbiol **6**(14): 1913-1923.
- Chiang, R. C., R. Cavicchioli, et al. (1997). "'Locked-on' and 'locked-off' signal transduction mutations in the periplasmic domain of the *Escherichia coli* NarQ and NarX sensors affect nitrate- and nitrite-dependent regulation by NarL and NarP." Mol Microbiol **24**(5): 1049-1060.
- Cho, B. K., C. L. Barrett, et al. (2008). "Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*." Proc Natl Acad Sci U S A **105**(49): 19462-19467.
- Cho, B. K., E. M. Knight, et al. (2008). "Genome-wide analysis of Fis binding in *Escherichia coli* indicates a causative role for A-/AT-tracts." Genome Res **18**(6): 900-910.
- Cho, S. H., J. Szewczyk, et al. (2014). "Detecting envelope stress by monitoring beta-barrel assembly." Cell **159**(7): 1652-1664.
- Choe, M. and W. S. Reznikoff (1993). "Identification of the regulatory sequence of anaerobically expressed locus aeg-46.5." J Bacteriol **175**(4): 1165-1172.
- Chung, C. T., S. L. Niemela, et al. (1989). "One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution." Proc Natl Acad Sci U S A **86**(7): 2172-2175.
- Collins, J. A., I. Irnov, et al. (2007). "Mechanism of mRNA destabilization by the glmS ribozyme." Genes Dev **21**(24): 3356-3368.
- Condon, C. and H. Putzer (2002). "The phylogenetic distribution of bacterial ribonucleases." Nucleic Acids Res **30**(24): 5339-5346.
- Constantinidou, C., J. L. Hobman, et al. (2006). "A reassessment of the FNR regulon and transcriptomic analysis of the effects of nitrate, nitrite, NarXL, and NarQP as *Escherichia coli* K12 adapts from aerobic to anaerobic growth." J Biol Chem **281**(8): 4802-4815.
- Cook, G. M., J. Membrillo-Hernandez, et al. (1997). "Transcriptional regulation of the cydDC operon, encoding a heterodimeric ABC transporter required for assembly of cytochromes c and bd in *Escherichia coli* K-12: regulation by oxygen and alternative electron acceptors." J Bacteriol **179**(20): 6525-6530.
- Coornaert, A., C. Chiaruttini, et al. (2013). "Post-transcriptional control of the *Escherichia coli* PhoQ-PhoP two-component system by multiple sRNAs involves a novel pairing region of GcvB." PLoS Genet **9**(1): e1003156.
- Coornaert, A., C. Chiaruttini, et al. (2013). "Post-transcriptional control of the *Escherichia coli* PhoQ-PhoP two-component system by multiple sRNAs involves a novel pairing region of GcvB." PLoS Genet **9**(1): e1003156.
- Coornaert, A., A. Lu, et al. (2010). "MicA sRNA links the PhoP regulon to cell envelope stress." Mol Microbiol **76**(2): 467-479.
- Cotter, P. A. and R. P. Gunsalus (1989). "Oxygen, nitrate, and molybdenum regulation of dmsABC gene expression in *Escherichia coli*." J Bacteriol **171**(7): 3817-3823.
- Court, D. L., J. Gan, et al. (2013). "RNase III: Genetics and function; structure and mechanism." Annu Rev Genet **47**: 405-431.
- Court, D. L., S. Swaminathan, et al. (2003). "Mini-lambda: a tractable system for chromosome and BAC engineering." Gene **315**: 63-69.
- da Costa, P. N., M. Teixeira, et al. (2003). "Regulation of the flavorubredoxin nitric oxide reductase gene in *Escherichia coli*: nitrate repression, nitrite induction, and possible post-transcription control." FEMS Microbiol Lett **218**(2): 385-393.
- Dahl, M. K., T. Msadek, et al. (1992). "The phosphorylation state of the DegU response regulator acts as a molecular switch allowing either degradative enzyme synthesis or expression of genetic competence in *Bacillus subtilis*." J Biol Chem **267**(20): 14509-14514.

- Darfeuille-Michaud, A., C. Neut, et al. (1998). "Presence of adherent *Escherichia coli* strains in ileal mucosa of patients with Crohn's disease." *Gastroenterology* **115**(6): 1405-1413.
- Darwin, A. J. and V. Stewart (1995). "Nitrate and nitrite regulation of the Fnr-dependent aeg-46.5 promoter of *Escherichia coli* K-12 is mediated by competition between homologous response regulators (NarL and NarP) for a common DNA-binding site." *J Mol Biol* **251**(1): 15-29.
- Darwin, A. J., K. L. Tyson, et al. (1997). "Differential regulation by the homologous response regulators NarL and NarP of *Escherichia coli* K-12 depends on DNA binding site arrangement." *Mol Microbiol* **25**(3): 583-595.
- De Las Penas, A., L. Connolly, et al. (1997). "SigmaE is an essential sigma factor in *Escherichia coli*." *J Bacteriol* **179**(21): 6862-6864.
- De Lay, N. and S. Gottesman (2009). "The Crp-activated small noncoding regulatory RNA CyaR (RyeE) links nutritional status to group behavior." *J Bacteriol* **191**(2): 461-476.
- De Lay, N. and S. Gottesman (2012). "A complex network of small non-coding RNAs regulate motility in *Escherichia coli*." *Mol Microbiol* **86**(3): 524-538.
- Deana, A., H. Celesnik, et al. (2008). "The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal." *Nature* **451**(7176): 355-358.
- Delgado, J., S. Forst, et al. (1993). "Identification of a phosphorylation site and functional analysis of conserved aspartic acid residues of OmpR, a transcriptional activator for ompF and ompC in *Escherichia coli*." *Mol Microbiol* **10**(5): 1037-1047.
- Delihans, N. and S. Forst (2001). "MicF: an antisense RNA gene involved in response of *Escherichia coli* to global stress factors." *J Mol Biol* **313**(1): 1-12.
- Desai, S. K. and L. J. Kenney (2017). "To approximately P or Not to approximately P? Non-canonical activation by two-component response regulators." *Mol Microbiol* **103**(2): 203-213.
- Desai, S. K., R. S. Winardhi, et al. (2016). "The horizontally-acquired response regulator SsrB drives a *Salmonella* lifestyle switch by relieving biofilm silencing." *Elife* **5**.
- Dorman, C. J., S. Chatfield, et al. (1989). "Characterization of porin and ompR mutants of a virulent strain of *Salmonella typhimurium*: ompR mutants are attenuated in vivo." *Infect Immun* **57**(7): 2136-2140.
- Douchin, V., C. Bohn, et al. (2006). "Down-regulation of porins by a small RNA bypasses the essentiality of the regulated intramembrane proteolysis protease RseP in *Escherichia coli*." *J Biol Chem* **281**(18): 12253-12259.
- Dressaire, C., R. N. Moreira, et al. (2015). "BolA is a transcriptional switch that turns off motility and turns on biofilm development." *MBio* **6**(1): e02352-02314.
- Dreyfus, M. (2009). "Killer and protective ribosomes." *Prog Mol Biol Transl Sci* **85**: 423-466.
- Durand, S. and G. Storz (2010). "Reprogramming of anaerobic metabolism by the FnrS small RNA." *Mol Microbiol* **75**(5): 1215-1231.
- Dutta, R., L. Qin, et al. (1999). "Histidine kinases: diversity of domain organization." *Mol Microbiol* **34**(4): 633-640.
- Eguchi, Y., J. Itou, et al. (2007). "B1500, a small membrane protein, connects the two-component systems EvgS/EvgA and PhoQ/PhoP in *Escherichia coli*." *Proc Natl Acad Sci U S A* **104**(47): 18712-18717.
- Eichler, K., A. Buchet, et al. (1996). "Identification and characterization of the caiF gene encoding a potential transcriptional activator of carnitine metabolism in *Escherichia coli*." *J Bacteriol* **178**(5): 1248-1257.
- Enz, S., S. Mahren, et al. (2003). "Analysis of the ferric citrate transport gene promoter of *Escherichia coli*." *J Bacteriol* **185**(7): 2387-2391.
- Faubladier, M., K. Cam, et al. (1990). "*Escherichia coli* cell division inhibitor DicF-RNA of the dicB operon. Evidence for its generation in vivo by transcription termination and by RNase III and RNase E-dependent processing." *J Mol Biol* **212**(3): 461-471.
- Figueroa-Bossi, N., M. Valentini, et al. (2009). "Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target." *Genes Dev* **23**(17): 2004-2015.
- Filenko, N., S. Spiro, et al. (2007). "The NsrR regulon of *Escherichia coli* K-12 includes genes encoding the hybrid cluster protein and the periplasmic, respiratory nitrite reductase." *J Bacteriol* **189**(12): 4410-4417.

- Forst, S., J. Delgado, et al. (1989). "Phosphorylation of OmpR by the osmosensor EnvZ modulates expression of the ompF and ompC genes in Escherichia coli." Proc Natl Acad Sci U S A **86**(16): 6052-6056.
- Frohlich, K. S., K. Haneke, et al. (2016). "The target spectrum of SdsR small RNA in Salmonella." Nucleic Acids Res **44**(21): 10406-10422.
- Frohlich, K. S., K. Papenfort, et al. (2012). "A conserved RpoS-dependent small RNA controls the synthesis of major porin OmpD." Nucleic Acids Res **40**(8): 3623-3640.
- Frohlich, K. S., K. Papenfort, et al. (2013). "A small RNA activates CFA synthase by isoform-specific mRNA stabilization." EMBO J **32**(22): 2963-2979.
- Garcia Vescovi, E., F. C. Soncini, et al. (1996). "Mg²⁺ as an extracellular signal: environmental regulation of Salmonella virulence." Cell **84**(1): 165-174.
- Garmendia, J., C. R. Beuzon, et al. (2003). "The roles of SsrA-SsrB and OmpR-EnvZ in the regulation of genes encoding the Salmonella typhimurium SPI-2 type III secretion system." Microbiology **149**(Pt 9): 2385-2396.
- Georg, J. and W. R. Hess (2011). "cis-antisense RNA, another level of gene regulation in bacteria." Microbiol Mol Biol Rev **75**(2): 286-300.
- Georgellis, D., O. Kwon, et al. (1999). "Amplification of signaling activity of the arc two-component system of Escherichia coli by anaerobic metabolites. An in vitro study with different protein modules." J Biol Chem **274**(50): 35950-35954.
- Gerasimova, A. V., M. S. Gelfand, et al. (2003). "ArcA regulator of Gamma-Proteobacteria: Identification of the Binding Signal and Description of the Regulon." Biophysics **48**(Suppl 1): S21-S25.
- Gerken, H., E. S. Charlson, et al. (2009). "MzrA: a novel modulator of the EnvZ/OmpR two-component regulon." Mol Microbiol **72**(6): 1408-1422.
- Gerstel, U., C. Park, et al. (2003). "Complex regulation of csgD promoter activity by global regulatory proteins." Mol Microbiol **49**(3): 639-654.
- Gevers, D., S. Kugathasan, et al. (2014). "The treatment-naive microbiome in new-onset Crohn's disease." Cell Host Microbe **15**(3): 382-392.
- Ghosh, S. and M. P. Deutscher (1999). "Oligoribonuclease is an essential component of the mRNA decay pathway." Proc Natl Acad Sci U S A **96**(8): 4372-4377.
- Glasser, A. L., J. Boudeau, et al. (2001). "Adherent invasive Escherichia coli strains from patients with Crohn's disease survive and replicate within macrophages without inducing host cell death." Infect Immun **69**(9): 5529-5537.
- Gogol, E. B., V. A. Rhodius, et al. (2011). "Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon." Proc Natl Acad Sci U S A **108**(31): 12875-12880.
- Goh, E. B., P. J. Bledsoe, et al. (2005). "Hierarchical control of anaerobic gene expression in Escherichia coli K-12: the nitrate-responsive NarX-NarL regulatory system represses synthesis of the fumarate-responsive DcuS-DcuR regulatory system." J Bacteriol **187**(14): 4890-4899.
- Goh, E. B., D. F. Siino, et al. (2004). "The Escherichia coli tppB (ydgR) gene represents a new class of OmpR-regulated genes." J Bacteriol **186**(12): 4019-4024.
- Golby, P., D. J. Kelly, et al. (1998). "Transcriptional regulation and organization of the dcuA and dcuB genes, encoding homologous anaerobic C4-dicarboxylate transporters in Escherichia coli." J Bacteriol **180**(24): 6586-6596.
- Gottesman, S. and G. Storz (2011). "Bacterial small RNA regulators: versatile roles and rapidly evolving variations." Cold Spring Harb Perspect Biol **3**(12).
- Grabowicz, M., D. Koren, et al. (2016). "The CpxQ sRNA Negatively Regulates Skp To Prevent Mistargeting of beta-Barrel Outer Membrane Proteins into the Cytoplasmic Membrane." MBio **7**(2): e00312-00316.
- Grantcharova, N., V. Peters, et al. (2010). "Bistable expression of CsgD in biofilm development of Salmonella enterica serovar typhimurium." J Bacteriol **192**(2): 456-466.
- Green, R. and H. F. Noller (1997). "Ribosomes and translation." Annu Rev Biochem **66**: 679-716.
- Greive, S. J. and P. H. von Hippel (2005). "Thinking quantitatively about transcriptional regulation." Nat Rev Mol Cell Biol **6**(3): 221-232.

- Guillier, M. and S. Gottesman (2006). "Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs." Mol. Microbiol. **59**(1): 231-247.
- Guillier, M. and S. Gottesman (2006). "Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs." Mol Microbiol **59**(1): 231-247.
- Guillier, M. and S. Gottesman (2008). "The 5' end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator." Nucleic Acids Res **36**(21): 6781-6794.
- Guillier, M. and S. Gottesman (2008). "The 5' end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator." Nucleic Acids Res. **36**(21): 6781-6794.
- Guillier, M., S. Gottesman, et al. (2006). "Modulating the outer membrane with small RNAs." Genes Dev **20**(17): 2338-2348.
- Guo, M. S., T. B. Updegrave, et al. (2014). "MicL, a new sigmaE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein." Genes Dev **28**(14): 1620-1634.
- Gushchin, I., I. Melnikov, et al. (2017). "Mechanism of transmembrane signaling by sensor histidine kinases." Science **356**(6342).
- Gutierrez, A., L. Laureti, et al. (2013). "beta-Lactam antibiotics promote bacterial mutagenesis via an RpoS-mediated reduction in replication fidelity." Nat Commun **4**: 1610.
- Haentjens-Sitri, J., F. Allemand, et al. (2008). "A competition mechanism regulates the translation of the *Escherichia coli* operon encoding ribosomal proteins L35 and L20." J Mol Biol **375**(3): 612-625.
- Hall, M. N. and T. J. Silhavy (1979). "Transcriptional regulation of *Escherichia coli* K-12 major outer membrane protein 1b." J. Bacteriol. **140**(2): 342-350.
- Hall, M. N. and T. J. Silhavy (1981). "The *ompB* locus and the regulation of the major outer membrane porin proteins of *Escherichia coli* K12." J. Mol. Biol. **146**(1): 23-43.
- Hall, M. N. and T. J. Silhavy (1981). "The *ompB* locus and the regulation of the major outer membrane porin proteins of *Escherichia coli* K12." J Mol Biol **146**(1): 23-43.
- Hansen, U. M. and W. R. McClure (1980). "Role of the sigma subunit of *Escherichia coli* RNA polymerase in initiation. II. Release of sigma from ternary complexes." J Biol Chem **255**(20): 9564-9570.
- Hao, Y., T. B. Updegrave, et al. (2016). "Protection against deleterious nitrogen compounds: role of sigmaS-dependent small RNAs encoded adjacent to *sdiA*." Nucleic Acids Res **44**(14): 6935-6948.
- Harley, C. B. and R. P. Reynolds (1987). "Analysis of *E. coli* promoter sequences." Nucleic Acids Res **15**(5): 2343-2361.
- Harlocker, S. L., L. Bergstrom, et al. (1995). "Tandem binding of six OmpR proteins to the *ompF* upstream regulatory sequence of *Escherichia coli*." J Biol Chem **270**(45): 26849-26856.
- Hasona, A., W. T. Self, et al. (2001). "Transcriptional regulation of the *moe* (molybdate metabolism) operon of *Escherichia coli*." Arch Microbiol **175**(3): 178-188.
- Hayden, J. D. and S. E. Ades (2008). "The extracytoplasmic stress factor, sigmaE, is required to maintain cell envelope integrity in *Escherichia coli*." PLoS One **3**(2): e1573.
- Head, C. G., A. Tardy, et al. (1998). "Relative binding affinities of OmpR and OmpR-phosphate at the *ompF* and *ompC* regulatory sites." J Mol Biol **281**(5): 857-870.
- Heidrich, N., I. Moll, et al. (2007). "In vitro analysis of the interaction between the small RNA SR1 and its primary target *ahrC* mRNA." Nucleic Acids Res **35**(13): 4331-4346.
- Helaine, S., J. A. Thompson, et al. (2010). "Dynamics of intracellular bacterial replication at the single cell level." Proc Natl Acad Sci U S A **107**(8): 3746-3751.
- Higashitani, A., Y. Nishimura, et al. (1993). "Osmoregulation of the fatty acid receptor gene *fadL* in *Escherichia coli*." Mol Gen Genet **240**(3): 339-347.
- Hobbs, E. C., J. L. Astarita, et al. (2010). "Small RNAs and small proteins involved in resistance to cell envelope stress and acid shock in *Escherichia coli*: analysis of a bar-coded mutant collection." J. Bact. **192**(1): 59-67.
- Holmqvist, E., J. Reimegard, et al. (2010). "Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis." EMBO J **29**(11): 1840-1850.

- Holmqvist, E., C. Unoson, et al. (2012). "A mixed double negative feedback loop between the sRNA MicF and the global regulator Lrp." *Mol Microbiol* **84**(3): 414-427.
- Holmqvist, E., P. R. Wright, et al. (2016). "Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo." *EMBO J* **35**(9): 991-1011.
- Hommais, F., E. Krin, et al. (2001). "Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS." *Mol Microbiol* **40**(1): 20-36.
- Hong, E., H. M. Lee, et al. (2007). "Structure of an atypical orphan response regulator protein supports a new phosphorylation-independent regulatory mechanism." *J Biol Chem* **282**(28): 20667-20675.
- Hsing, W., F. D. Russo, et al. (1998). "Mutations that alter the kinase and phosphatase activities of the two-component sensor EnvZ." *J Bacteriol* **180**(17): 4538-4546.
- Hsing, W. and T. J. Silhavy (1997). "Function of conserved histidine-243 in phosphatase activity of EnvZ, the sensor for porin osmoregulation in Escherichia coli." *J Bacteriol* **179**(11): 3729-3735.
- Huang, K. J. and M. M. Igo (1996). "Identification of the bases in the ompF regulatory region, which interact with the transcription factor OmpR." *J Mol Biol* **262**(5): 615-628.
- Huang, K. J., J. L. Schieberl, et al. (1994). "A distant upstream site involved in the negative regulation of the Escherichia coli ompF gene." *J Bacteriol* **176**(5): 1309-1315.
- Huang, L., P. Tsui, et al. (1992). "Positive and negative control of ompB transcription in Escherichia coli by cyclic AMP and the cyclic AMP receptor protein." *J Bacteriol* **174**(3): 664-670.
- Hui, M. P., P. L. Foley, et al. (2014). "Messenger RNA degradation in bacterial cells." *Annu Rev Genet* **48**: 537-559.
- Huttenhofer, A. and H. F. Noller (1994). "Footprinting mRNA-ribosome complexes with chemical probes." *EMBO J* **13**(16): 3892-3901.
- Huynh, T. N., L. L. Chen, et al. (2015). "Sensor-response regulator interactions in a cross-regulated signal transduction network." *Microbiology* **161**(7): 1504-1515.
- Igo, M. M. and T. J. Silhavy (1988). "EnvZ, a transmembrane environmental sensor of Escherichia coli K-12, is phosphorylated in vitro." *J Bacteriol* **170**(12): 5971-5973.
- Ishihama, A., A. Kori, et al. (2014). "Intracellular concentrations of 65 species of transcription factors with known regulatory functions in Escherichia coli." *J Bacteriol* **196**(15): 2718-2727.
- Iuchi, S. and E. C. Lin (1987). "The narL gene product activates the nitrate reductase operon and represses the fumarate reductase and trimethylamine N-oxide reductase operons in Escherichia coli." *Proc Natl Acad Sci U S A* **84**(11): 3901-3905.
- Iuchi, S. and E. C. Lin (1988). "arcA (dye), a global regulatory gene in Escherichia coli mediating repression of enzymes in aerobic pathways." *Proc Natl Acad Sci U S A* **85**(6): 1888-1892.
- Iuchi, S. and E. C. Lin (1993). "Adaptation of Escherichia coli to redox environments by gene expression." *Mol Microbiol* **9**(1): 9-15.
- Jacob, F. and J. Monod (1961). "Genetic regulatory mechanisms in the synthesis of proteins." *J Mol Biol* **3**: 318-356.
- Jagodnik, J., A. Brosse, et al. (2017). "Mechanistic study of base-pairing small regulatory RNAs in bacteria." *Methods* **117**: 67-76.
- Jain, C. and J. G. Belasco (1995). "RNase E autoregulates its synthesis by controlling the degradation rate of its own mRNA in Escherichia coli: unusual sensitivity of the rne transcript to RNase E activity." *Genes Dev* **9**(1): 84-96.
- Jin, T. and M. Inouye (1993). "Ligand binding to the receptor domain regulates the ratio of kinase to phosphatase activities of the signaling domain of the hybrid Escherichia coli transmembrane receptor, Taz1." *J Mol Biol* **232**(2): 484-492.
- Johansen, J., M. Eriksen, et al. (2008). "Down-regulation of outer membrane proteins by noncoding RNAs: unraveling the cAMP-CRP- and sigmaE-dependent CyaR-ompX regulatory case." *J Mol Biol* **383**(1): 1-9.
- Jostins, L., S. Ripke, et al. (2012). "Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease." *Nature* **491**(7422): 119-124.
- Jung, K., K. Hamann, et al. (2001). "K⁺ stimulates specifically the autokinase activity of purified and reconstituted EnvZ of Escherichia coli." *J Biol Chem* **276**(44): 40896-40902.

- Kaiser, M. and G. Sawers (1995). "Nitrate repression of the *Escherichia coli* pfl operon is mediated by the dual sensors NarQ and NarX and the dual regulators NarL and NarP." *J Bacteriol* **177**(13): 3647-3655.
- Kalamorz, F., B. Reichenbach, et al. (2007). "Feedback control of glucosamine-6-phosphate synthase GlmS expression depends on the small RNA GlmZ and involves the novel protein YhbJ in *Escherichia coli*." *Mol Microbiol* **65**(6): 1518-1533.
- Kalman, L. V. and R. P. Gunsalus (1989). "Identification of a second gene involved in global regulation of fumarate reductase and other nitrate-controlled genes for anaerobic respiration in *Escherichia coli*." *J Bacteriol* **171**(7): 3810-3816.
- Kang, S., S. E. Denman, et al. (2010). "Dysbiosis of fecal microbiota in Crohn's disease patients as revealed by a custom phylogenetic microarray." *Inflamm Bowel Dis* **16**(12): 2034-2042.
- Kaper, J. B., J. P. Nataro, et al. (2004). "Pathogenic *Escherichia coli*." *Nat Rev Microbiol* **2**(2): 123-140.
- Kara, A., M. Vickers, et al. (2015). "Genome-wide prediction of prokaryotic two-component system networks using a sequence-based meta-predictor." *BMC Bioinformatics* **16**: 297.
- Kato, N., M. Tsuzuki, et al. (1995). "Gene activation by the *Escherichia coli* positive regulator OmpR: a mutational study of the DNA-binding domain of OmpR." *Mol Gen Genet* **248**(4): 399-406.
- Kawamoto, H., Y. Koide, et al. (2006). "Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq." *Mol Microbiol* **61**(4): 1013-1022.
- Keiler, K. C. (2015). "Mechanisms of ribosome rescue in bacteria." *Nat Rev Microbiol* **13**(5): 285-297.
- Kenney, L. J. (2002). "Structure/function relationships in OmpR and other winged-helix transcription factors." *Curr Opin Microbiol* **5**(2): 135-141.
- Keseler, I. M., A. Mackie, et al. (2013). "EcoCyc: fusing model organism databases with systems biology." *Nucleic Acids Res* **41**(Database issue): D605-612.
- Khemici, V., L. Poljak, et al. (2008). "The RNase E of *Escherichia coli* is a membrane-binding protein." *Mol Microbiol* **70**(4): 799-813.
- Kiley, P. J. and H. Beinert (1998). "Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster." *FEMS Microbiol Rev* **22**(5): 341-352.
- Kirby, J. R. (2009). "Chemotaxis-like regulatory systems: unique roles in diverse bacteria." *Annu Rev Microbiol* **63**: 45-59.
- Kolb, A., S. Busby, et al. (1993). "Transcriptional regulation by cAMP and its receptor protein." *Annu Rev Biochem* **62**: 749-795.
- Kondo, H., A. Nakagawa, et al. (1997). "*Escherichia coli* positive regulator OmpR has a large loop structure at the putative RNA polymerase interaction site." *Nat Struct Biol* **4**(1): 28-31.
- Koo, B. M., V. A. Rhodius, et al. (2009). "Mutational analysis of *Escherichia coli* sigma28 and its target promoters reveals recognition of a composite -10 region, comprised of an 'extended -10' motif and a core -10 element." *Mol Microbiol* **72**(4): 830-843.
- Kuehn, M. J. and N. C. Kesty (2005). "Bacterial outer membrane vesicles and the host-pathogen interaction." *Genes Dev* **19**(22): 2645-2655.
- Kushner, S. R. (1996). mRNA Decay. *Escherichia coli and Salmonella*. F. C. Neidhardt. Washington, D. C., ASM Press. **1**: 849-860.
- Lacour, S., A. Kolb, et al. (2003). "Nucleotides from -16 to -12 determine specific promoter recognition by bacterial sigmaS-RNA polymerase." *J Biol Chem* **278**(39): 37160-37168.
- Landers, C. J., O. Cohavy, et al. (2002). "Selected loss of tolerance evidenced by Crohn's disease-associated immune responses to auto- and microbial antigens." *Gastroenterology* **123**(3): 689-699.
- Latasa, C., B. Garcia, et al. (2012). "*Salmonella* biofilm development depends on the phosphorylation status of RcsB." *J Bacteriol* **194**(14): 3708-3722.
- Laub, M. T. and M. Goulian (2007). "Specificity in two-component signal transduction pathways." *Annu Rev Genet* **41**: 121-145.
- Lease, R. A. and M. Belfort (2000). "Riboregulation by DsrA RNA: trans-actions for global economy." *Mol Microbiol* **38**(4): 667-672.
- Lease, R. A., D. Smith, et al. (2004). "The small noncoding DsrA RNA is an acid resistance regulator in *Escherichia coli*." *J Bacteriol* **186**(18): 6179-6185.

- Lee, A. I., A. Delgado, et al. (1999). "Signal-dependent phosphorylation of the membrane-bound NarX two-component sensor-transmitter protein of *Escherichia coli*: nitrate elicits a superior anion ligand response compared to nitrite." *J Bacteriol* **181**(17): 5309-5316.
- Lee, A. K., C. S. Detweiler, et al. (2000). "OmpR regulates the two-component system SsrA-ssrB in *Salmonella* pathogenicity island 2." *J Bacteriol* **182**(3): 771-781.
- Leonardo, M. R. and S. Forst (1996). "Re-examination of the role of the periplasmic domain of EnvZ in sensing of osmolarity signals in *Escherichia coli*." *Mol Microbiol* **22**(3): 405-413.
- Lesnik, E. A., R. Sampath, et al. (2001). "Prediction of rho-independent transcriptional terminators in *Escherichia coli*." *Nucleic Acids Res* **29**(17): 3583-3594.
- Levi-Meyrueis, C., V. Monteil, et al. (2014). "Expanding the RpoS/sigmaS-network by RNA sequencing and identification of sigmaS-controlled small RNAs in *Salmonella*." *PLoS One* **9**(5): e96918.
- Li, G. W., D. Burkhardt, et al. (2014). "Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources." *Cell* **157**(3): 624-635.
- Lin, H. Y., P. J. Bledsoe, et al. (2007). "Activation of *yeaR*-*yoaG* operon transcription by the nitrate-responsive regulator NarL is independent of oxygen-responsive regulator Fnr in *Escherichia coli* K-12." *J Bacteriol* **189**(21): 7539-7548.
- Link, T. M., P. Valentin-Hansen, et al. (2009). "Structure of *Escherichia coli* Hfq bound to polyribadenylate RNA." *Proc Natl Acad Sci U S A* **106**(46): 19292-19297.
- Lorenz, C., T. Gesell, et al. (2010). "Genomic SELEX for Hfq-binding RNAs identifies genomic aptamers predominantly in antisense transcripts." *Nucleic Acids Res* **38**(11): 3794-3808.
- Lynch, A. S. and E. C. Lin (1996). "Transcriptional control mediated by the ArcA two-component response regulator protein of *Escherichia coli*: characterization of DNA binding at target promoters." *J Bacteriol* **178**(21): 6238-6249.
- Mackie, G. A. (1998). "Ribonuclease E is a 5'-end-dependent endonuclease." *Nature* **395**(6703): 720-723.
- Mackie, G. A. (2013). "RNase E: at the interface of bacterial RNA processing and decay." *Nat Rev Microbiol* **11**(1): 45-57.
- Maeda, S. and T. Mizuno (1990). "Evidence for multiple OmpR-binding sites in the upstream activation sequence of the *ompC* promoter in *Escherichia coli*: a single OmpR-binding site is capable of activating the promoter." *J Bacteriol* **172**(1): 501-503.
- Maeda, S., K. Takayanagi, et al. (1991). "Activation of the osmoregulated *ompC* gene by the OmpR protein in *Escherichia coli*: a study involving synthetic OmpR-binding sequences." *J Biochem* **110**(3): 324-327.
- Mahid, S. S., K. S. Minor, et al. (2007). "The role of smoking in Crohn's disease as defined by clinical variables." *Dig Dis Sci* **52**(11): 2897-2903.
- Majdalani, N., S. Chen, et al. (2001). "Regulation of RpoS by a novel small RNA: the characterization of RprA." *Mol Microbiol* **39**(5): 1382-1394.
- Majdalani, N., C. Cuning, et al. (1998). "DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription." *Proc Natl Acad Sci U S A* **95**(21): 12462-12467.
- Majdalani, N. and S. Gottesman (2005). "The Rcs phosphorelay: a complex signal transduction system." *Annu Rev Microbiol* **59**: 379-405.
- Majdalani, N., M. Heck, et al. (2005). "Role of RcsF in signaling to the Rcs phosphorelay pathway in *Escherichia coli*." *J Bacteriol* **187**(19): 6770-6778.
- Majdalani, N., D. Hernandez, et al. (2002). "Regulation and mode of action of the second small RNA activator of RpoS translation, RprA." *Mol Microbiol* **46**(3): 813-826.
- Malpica, R., G. R. Sandoval, et al. (2006). "Signaling by the *arc* two-component system provides a link between the redox state of the quinone pool and gene expression." *Antioxid Redox Signal* **8**(5-6): 781-795.
- Mandin, P. and S. Gottesman (2009). "A genetic approach for finding small RNAs regulators of genes of interest identifies RybC as regulating the DpiA/DpiB two-component system." *Mol. Microbiol.* **72**(3): 551-565.
- Mandin, P. and S. Gottesman (2010). "Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA." *EMBO J* **29**(18): 3094-3107.

- Mandin, P. and M. Guillier (2013). "Expanding control in bacteria: interplay between small RNAs and transcriptional regulators to control gene expression." *Curr Opin Microbiol* **16**(2): 125-132.
- Mandin, P., F. Repoila, et al. (2007). "Identification of new noncoding RNAs in *Listeria monocytogenes* and prediction of mRNA targets." *Nucleic Acids Res* **35**(3): 962-974.
- Mann, M., P. R. Wright, et al. (2017). "IntaRNA 2.0: enhanced and customizable prediction of RNA-RNA interactions." *Nucleic Acids Res*.
- Mansilla, M. C., L. E. Cybulski, et al. (2004). "Control of membrane lipid fluidity by molecular thermosensors." *J Bacteriol* **186**(20): 6681-6688.
- Maris, A. E., M. Kaczor-Grzeskowiak, et al. (2005). "Primary and secondary modes of DNA recognition by the NarL two-component response regulator." *Biochemistry* **44**(44): 14538-14552.
- Martinez-Antonio, A. and J. Collado-Vides (2003). "Identifying global regulators in transcriptional regulatory networks in bacteria." *Curr Opin Microbiol* **6**(5): 482-489.
- Martinez-Hackert, E. and A. M. Stock (1997). "The DNA-binding domain of OmpR: crystal structures of a winged helix transcription factor." *Structure* **5**(1): 109-124.
- Mascher, T., J. D. Helmann, et al. (2006). "Stimulus perception in bacterial signal-transducing histidine kinases." *Microbiol Mol Biol Rev* **70**(4): 910-938.
- Massé, E., F. E. Escorcia, et al. (2003). "Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*." *Genes Dev* **17**(19): 2374-2383.
- Massé, E. and S. Gottesman (2002). "A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*." *Proc Natl Acad Sci U S A* **99**(7): 4620-4625.
- Massé, E., C. K. Vanderpool, et al. (2005). "Effect of RyhB small RNA on global iron use in *Escherichia coli*." *J Bacteriol* **187**(20): 6962-6971.
- Mathy, N., L. Benard, et al. (2007). "5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA." *Cell* **129**(4): 681-692.
- Matsubara, M., S. I. Kitaoka, et al. (2000). "Tuning of the porin expression under anaerobic growth conditions by his-to-Asp cross-phosphorelay through both the EnvZ-osmosensor and ArcB-anaerosensor in *Escherichia coli*." *Genes Cells* **5**(7): 555-569.
- Matsubara, M. and T. Mizuno (1999). "EnvZ-independent phosphotransfer signaling pathway of the OmpR-mediated osmoregulatory expression of OmpC and OmpF in *Escherichia coli*." *Biosci Biotechnol Biochem* **63**(2): 408-414.
- Matsunaga, J., E. L. Simons, et al. (1997). "*Escherichia coli* RNase III (*rnc*) autoregulation occurs independently of *rnc* gene translation." *Mol Microbiol* **26**(5): 1125-1135.
- Mattison, K., R. Oropeza, et al. (2002). "The linker region plays an important role in the interdomain communication of the response regulator OmpR." *J Biol Chem* **277**(36): 32714-32721.
- McLaren, R. S., S. F. Newbury, et al. (1991). "mRNA degradation by processive 3'-5' exoribonucleases in vitro and the implications for prokaryotic mRNA decay in vivo." *J Mol Biol* **221**(1): 81-95.
- Meconi, S., A. Vercellone, et al. (2007). "Adherent-invasive *Escherichia coli* isolated from Crohn's disease patients induce granulomas in vitro." *Cell Microbiol* **9**(5): 1252-1261.
- Membrillo-Hernandez, J. and E. C. Lin (1999). "Regulation of expression of the *adhE* gene, encoding ethanol oxidoreductase in *Escherichia coli*: transcription from a downstream promoter and regulation by *fnr* and *RpoS*." *J Bacteriol* **181**(24): 7571-7579.
- Merrick, M. J. (1993). "In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N)." *Mol Microbiol* **10**(5): 903-909.
- Mika, F., S. Busse, et al. (2012). "Targeting of *csgD* by the small regulatory RNA RprA links stationary phase, biofilm formation and cell envelope stress in *Escherichia coli*." *Mol Microbiol* **84**(1): 51-65.
- Mika, F. and R. Hengge (2014). "Small RNAs in the control of RpoS, CsgD, and biofilm architecture of *Escherichia coli*." *RNA Biol* **11**(5): 494-507.
- Minagawa, S., H. Ogasawara, et al. (2003). "Identification and molecular characterization of the Mg²⁺ stimulon of *Escherichia coli*." *J Bacteriol* **185**(13): 3696-3702.
- Miquel, S., E. Peyretailade, et al. (2010). "Complete genome sequence of Crohn's disease-associated adherent-invasive *E. coli* strain LF82." *PLoS One* **5**(9).
- Mizuno, T. (1997). "Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of *Escherichia coli*." *DNA Res* **4**(2): 161-168.

- Mizuno, T., M. Y. Chou, et al. (1984). "A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA)." Proc Natl Acad Sci U S A **81**(7): 1966-1970.
- Modi, S. R., D. M. Camacho, et al. (2011). "Functional characterization of bacterial sRNAs using a network biology approach." Proc Natl Acad Sci U S A **108**(37): 15522-15527.
- Moll, I., T. Afonyushkin, et al. (2003). "Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs." RNA **9**(11): 1308-1314.
- Moller, T., T. Franch, et al. (2002). "Spot 42 RNA mediates discoordinate expression of the E. coli galactose operon." Genes Dev **16**(13): 1696-1706.
- Mondot, S., S. Kang, et al. (2011). "Highlighting new phylogenetic specificities of Crohn's disease microbiota." Inflamm Bowel Dis **17**(1): 185-192.
- Moon, K. and S. Gottesman (2009). "A PhoQ/P-regulated small RNA regulates sensitivity of Escherichia coli to antimicrobial peptides." Mol Microbiol **74**(6): 1314-1330.
- Moon, K. and S. Gottesman (2011). "Competition among Hfq-binding small RNAs in Escherichia coli." Mol Microbiol **82**(6): 1545-1562.
- Moreno, F., J. E. Gonzalez-Pastor, et al. (2002). "The regulation of microcin B, C and J operons." Biochimie **84**(5-6): 521-529.
- Morita, T. and H. Aiba (2011). "RNase E action at a distance: degradation of target mRNAs mediated by an Hfq-binding small RNA in bacteria." Genes Dev **25**(4): 294-298.
- Morita, T., K. Maki, et al. (2005). "RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs." Genes Dev **19**(18): 2176-2186.
- Mukhopadhyay, S., J. P. Audia, et al. (2000). "Transcriptional induction of the conserved alternative sigma factor RpoS in Escherichia coli is dependent on BarA, a probable two-component regulator." Mol Microbiol **37**(2): 371-381.
- Munch, R., K. Hiller, et al. (2005). "Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes." Bioinformatics **21**(22): 4187-4189.
- Mura, C., P. S. Randolph, et al. (2013). "Archaeal and eukaryotic homologs of Hfq: A structural and evolutionary perspective on Sm function." RNA Biol **10**(4): 636-651.
- Neidhardt, F. C. and H. E. Umberger (1996). Chemical composition of *Escherichia coli*. Escherichia coli and Salmonella ; Cellular And Molecular Biology. F. C. Neidhardt. Washington, D. C. , ASM Press. **2**: 13-16.
- Nissen, P., J. Hansen, et al. (2000). "The structural basis of ribosome activity in peptide bond synthesis." Science **289**(5481): 920-930.
- Nixon, B. T., C. W. Ronson, et al. (1986). "Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes ntrB and ntrC." Proc Natl Acad Sci U S A **83**(20): 7850-7854.
- Noeske, J., M. R. Wasserman, et al. (2015). "High-resolution structure of the Escherichia coli ribosome." Nat Struct Mol Biol **22**(4): 336-341.
- Noriega, C. E., H. Y. Lin, et al. (2010). "Asymmetric cross-regulation between the nitrate-responsive NarX-NarL and NarQ-NarP two-component regulatory systems from Escherichia coli K-12." Mol Microbiol **75**(2): 394-412.
- Norioka, S., G. Ramakrishnan, et al. (1986). "Interaction of a transcriptional activator, OmpR, with reciprocally osmoregulated genes, ompF and ompC, of Escherichia coli." J Biol Chem **261**(36): 17113-17119.
- Novick, R. P. and E. Geisinger (2008). "Quorum sensing in staphylococci." Annu Rev Genet **42**: 541-564.
- Nudler, E. and M. E. Gottesman (2002). "Transcription termination and anti-termination in E. coli." Genes Cells **7**(8): 755-768.
- Ogasawara, H., A. Hasegawa, et al. (2007). "Genomic SELEX search for target promoters under the control of the PhoQP-RstBA signal relay cascade." J Bacteriol **189**(13): 4791-4799.
- Ogasawara, H., S. Shinohara, et al. (2012). "Novel regulation targets of the metal-response BasS-BasR two-component system of Escherichia coli." Microbiology **158**(Pt 6): 1482-1492.
- Opdyke, J. A., E. M. Fozo, et al. (2011). "RNase III participates in GadY-dependent cleavage of the gadX-gadW mRNA." J Mol Biol **406**(1): 29-43.

- Opdyke, J. A., J. G. Kang, et al. (2004). "GadY, a small-RNA regulator of acid response genes in *Escherichia coli*." *J Bacteriol* **186**(20): 6698-6705.
- Oshima, T., H. Aiba, et al. (2002). "Transcriptome analysis of all two-component regulatory system mutants of *Escherichia coli* K-12." *Mol Microbiol* **46**(1): 281-291.
- Osterberg, S., T. del Peso-Santos, et al. (2011). "Regulation of alternative sigma factor use." *Annu Rev Microbiol* **65**: 37-55.
- Otaka, H., H. Ishikawa, et al. (2011). "PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action." *Proc Natl Acad Sci U S A* **108**(32): 13059-13064.
- Page, L., L. Griffiths, et al. (1990). "Different physiological roles of two independent pathways for nitrite reduction to ammonia by enteric bacteria." *Arch Microbiol* **154**(4): 349-354.
- Papenfort, K., E. Espinosa, et al. (2015). "Small RNA-based feedforward loop with AND-gate logic regulates extrachromosomal DNA transfer in *Salmonella*." *Proc Natl Acad Sci U S A* **112**(34): E4772-4781.
- Papenfort, K., V. Pfeiffer, et al. (2006). "SigmaE-dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global omp mRNA decay." *Mol Microbiol* **62**(6): 1674-1688.
- Papenfort, K., N. Said, et al. (2009). "Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA." *Mol Microbiol* **74**(1): 139-158.
- Park, H. and M. Inouye (1997). "Mutational analysis of the linker region of EnvZ, an osmosensor in *Escherichia coli*." *J Bacteriol* **179**(13): 4382-4390.
- Peakman, T., J. Crouzet, et al. (1990). "Nucleotide sequence, organisation and structural analysis of the products of genes in the nirB-cysG region of the *Escherichia coli* K-12 chromosome." *Eur J Biochem* **191**(2): 315-323.
- Peano, C., J. Wolf, et al. (2015). "Characterization of the *Escherichia coli* sigma(S) core regulon by Chromatin Immunoprecipitation-sequencing (ChIP-seq) analysis." *Sci Rep* **5**: 10469.
- Pena-Sandoval, G. R. and D. Georgellis (2010). "The ArcB sensor kinase of *Escherichia coli* autophosphorylates by an intramolecular reaction." *J Bacteriol* **192**(6): 1735-1739.
- Peterson, C. N., V. J. Carabetta, et al. (2006). "LrhA regulates rpoS translation in response to the Rcs phosphorelay system in *Escherichia coli*." *J Bacteriol* **188**(9): 3175-3181.
- Pfeiffer, V., K. Papenfort, et al. (2009). "Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation." *Nat Struct Mol Biol* **16**(8): 840-846.
- Pineton de Chambrun, G., J. F. Colombel, et al. (2008). "Pathogenic agents in inflammatory bowel diseases." *Curr Opin Gastroenterol* **24**(4): 440-447.
- Portier, C., L. Dondon, et al. (1990). "Translational autocontrol of the *Escherichia coli* ribosomal protein S15." *J Mol Biol* **211**(2): 407-414.
- Potter, L., H. Angove, et al. (2001). "Nitrate reduction in the periplasm of gram-negative bacteria." *Adv Microb Physiol* **45**: 51-112.
- Pratt, L. A., W. Hsing, et al. (1996). "From acids to osmZ: multiple factors influence synthesis of the OmpF and OmpC porins in *Escherichia coli*." *Mol Microbiol* **20**(5): 911-917.
- Pratt, L. A. and T. J. Silhavy (1995). "Identification of base pairs important for OmpR-DNA interaction." *Mol Microbiol* **17**(3): 565-573.
- Prevost, K., G. Desnoyers, et al. (2011). "Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage." *Genes Dev* **25**(4): 385-396.
- Prigent-Combaret, C., E. Brombacher, et al. (2001). "Complex regulatory network controls initial adhesion and biofilm formation in *Escherichia coli* via regulation of the csgD gene." *J Bacteriol* **183**(24): 7213-7223.
- Prigent-Combaret, C., G. Prensier, et al. (2000). "Developmental pathway for biofilm formation in curli-producing *Escherichia coli* strains: role of flagella, curli and colanic acid." *Environ Microbiol* **2**(4): 450-464.
- Purcell, E. B. and S. Crosson (2008). "Photoregulation in prokaryotes." *Curr Opin Microbiol* **11**(2): 168-178.
- Qin, J., R. Li, et al. (2010). "A human gut microbial gene catalogue established by metagenomic sequencing." *Nature* **464**(7285): 59-65.

- Quinn, H. J., A. D. Cameron, et al. (2014). "Bacterial regulon evolution: distinct responses and roles for the identical OmpR proteins of *Salmonella Typhimurium* and *Escherichia coli* in the acid stress response." PLoS Genet. **10**(3): e1004215.
- Quinn, H. J., A. D. Cameron, et al. (2014). "Bacterial regulon evolution: distinct responses and roles for the identical OmpR proteins of *Salmonella Typhimurium* and *Escherichia coli* in the acid stress response." PLoS Genet **10**(3): e1004215.
- Rabin, R. S. and V. Stewart (1992). "Either of two functionally redundant sensor proteins, NarX and NarQ, is sufficient for nitrate regulation in *Escherichia coli* K-12." Proc Natl Acad Sci U S A **89**(18): 8419-8423.
- Rabin, R. S. and V. Stewart (1993). "Dual response regulators (NarL and NarP) interact with dual sensors (NarX and NarQ) to control nitrate- and nitrite-regulated gene expression in *Escherichia coli* K-12." J Bacteriol **175**(11): 3259-3268.
- Ralston, A. (2008). "Simultaneous Gene Transcription and Translation in Bacteria." Nature Education **1**(1).
- Ramirez-Pena, E., J. Trevino, et al. (2010). "The group A Streptococcus small regulatory RNA FasX enhances streptokinase activity by increasing the stability of the ska mRNA transcript." Mol Microbiol **78**(6): 1332-1347.
- Ramos, C. G., S. A. Sousa, et al. (2011). "The second RNA chaperone, Hfq2, is also required for survival under stress and full virulence of *Burkholderia cenocepacia* J2315." J Bacteriol **193**(7): 1515-1526.
- Rasmussen, A. A., M. Eriksen, et al. (2005). "Regulation of ompA mRNA stability: the role of a small regulatory RNA in growth phase-dependent control." Mol Microbiol **58**(5): 1421-1429.
- Raynal, L. C. and A. J. Carpousis (1999). "Poly(A) polymerase I of *Escherichia coli*: characterization of the catalytic domain, an RNA binding site and regions for the interaction with proteins involved in mRNA degradation." Mol Microbiol **32**(4): 765-775.
- Rentschler, A. E., S. D. Lovrich, et al. (2013). "OmpR regulation of the uropathogenic *Escherichia coli* fimB gene in an acidic/high osmolality environment." Microbiology **159**(Pt 2): 316-327.
- Repoila, F., N. Majdalani, et al. (2003). "Small non-coding RNAs, co-ordinators of adaptation processes in *Escherichia coli*: the RpoS paradigm." Mol Microbiol **48**(4): 855-861.
- Rezuchova, B., H. Miticka, et al. (2003). "New members of the *Escherichia coli* sigmaE regulon identified by a two-plasmid system." FEMS Microbiol Lett **225**(1): 1-7.
- Richard, D. J., G. Sawers, et al. (1999). "Transcriptional regulation in response to oxygen and nitrate of the operons encoding the [NiFe] hydrogenases 1 and 2 of *Escherichia coli*." Microbiology **145** (Pt 10): 2903-2912.
- Richards, G. R. and C. K. Vanderpool (2011). "Molecular call and response: the physiology of bacterial small RNAs." Biochim Biophys Acta **1809**(10): 525-531.
- Righetti, F. and F. Narberhaus (2014). "How to find RNA thermometers." Front Cell Infect Microbiol **4**: 132.
- Rivas, E., R. J. Klein, et al. (2001). "Computational identification of noncoding RNAs in *E. coli* by comparative genomics." Curr Biol **11**(17): 1369-1373.
- Rochat, T., P. Bouloc, et al. (2013). "Gene expression control by selective RNA processing and stabilization in bacteria." FEMS Microbiol Lett **344**(2): 104-113.
- Rochat, T., P. Bouloc, et al. (2012). "Lack of interchangeability of Hfq-like proteins." Biochimie **94**(7): 1554-1559.
- Rolhion, N., N. Barnich, et al. (2005). "Strong decrease in invasive ability and outer membrane vesicle release in Crohn's disease-associated adherent-invasive *Escherichia coli* strain LF82 with the yfgL gene deleted." J Bacteriol **187**(7): 2286-2296.
- Rolhion, N., F. A. Carvalho, et al. (2007). "OmpC and the sigma(E) regulatory pathway are involved in adhesion and invasion of the Crohn's disease-associated *Escherichia coli* strain LF82." Mol Microbiol **63**(6): 1684-1700.
- Romby, P. and M. Springer (2007). Translational Control in Prokaryotes. Translational Control in Biology and Medicine. M. B. Matthews, N. Sonenberg and J. W. B. Hershey, Cold Spring Harbor Laboratory Press 803-827.
- Romilly, C. (2012). Fonction de nouveaux ARN non codant dans la régulation de l'expression des gènes de *Staphylococcus aureus* : adaptation à l'environnement et virulence, Université de Strasbourg.

- Ross, w., J. F. Thompson, et al. (1990). "*E. coli* Fis protein activates ribosomal RNA transcription *in vitro* and *in vivo*." *EMBO J* **9**: 3733-3742.
- Ross, W., J. F. Thompson, et al. (1990). "E.coli Fis protein activates ribosomal RNA transcription *in vitro* and *in vivo*." *EMBO J* **9**(11): 3733-3742.
- Rowe, J. L., G. L. Starnes, et al. (2005). "Complex transcriptional control links NikABCDE-dependent nickel transport with hydrogenase expression in *Escherichia coli*." *J Bacteriol* **187**(18): 6317-6323.
- Russo, F. D. and T. J. Silhavy (1991). "EnvZ controls the concentration of phosphorylated OmpR to mediate osmoregulation of the porin genes." *J Mol Biol* **222**(3): 567-580.
- Russo, F. D., J. M. Slauch, et al. (1993). "Mutations that affect separate functions of OmpR the phosphorylated regulator of porin transcription in *Escherichia coli*." *J Mol Biol* **231**(2): 261-273.
- Ryan, P., R. G. Kelly, et al. (2004). "Bacterial DNA within granulomas of patients with Crohn's disease-detection by laser capture microdissection and PCR." *Am J Gastroenterol* **99**(8): 1539-1543.
- Salmon, K., S. P. Hung, et al. (2003). "Global gene expression profiling in *Escherichia coli* K12. The effects of oxygen availability and FNR." *J Biol Chem* **278**(32): 29837-29855.
- Salvail, H. and E. Masse (2012). "Regulating iron storage and metabolism with RNA: an overview of posttranscriptional controls of intracellular iron homeostasis." *Wiley Interdiscip Rev RNA* **3**(1): 26-36.
- Sarma, V. and P. Reeves (1977). "Genetic locus (*ompB*) affecting a major outer-membrane protein in *Escherichia coli* K-12." *J Bacteriol* **132**(1): 23-27.
- Sartor, R. B. (2008). "Microbial influences in inflammatory bowel diseases." *Gastroenterology* **134**(2): 577-594.
- Sauer, E. and O. Weichenrieder (2011). "Structural basis for RNA 3'-end recognition by Hfq." *Proc Natl Acad Sci U S A* **108**(32): 13065-13070.
- Schramke, H., V. Laermann, et al. (2017). "Revisiting regulation of potassium homeostasis in *Escherichia coli*: the connection to phosphate limitation." *Microbiologyopen*.
- Schu, D. J., A. Zhang, et al. (2015). "Alternative Hfq-sRNA interaction modes dictate alternative mRNA recognition." *EMBO J* **34**(20): 2557-2573.
- Shanahan, F. (2002). "Crohn's disease." *Lancet* **359**(9300): 62-69.
- Sharma, C. M., F. Darfeuille, et al. (2007). "A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites." *Genes Dev* **21**(21): 2804-2817.
- Sharma, C. M., K. Papenfort, et al. (2011). "Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA." *Mol Microbiol* **81**(5): 1144-1165.
- Shimoni, Y., G. Friedlander, et al. (2007). "Regulation of gene expression by small non-coding RNAs: a quantitative view." *Mol Syst Biol* **3**: 138.
- Shin, S. and C. Park (1995). "Modulation of flagellar expression in *Escherichia coli* by acetyl phosphate and the osmoregulator OmpR." *J Bacteriol* **177**(16): 4696-4702.
- Shine, J. and L. Dalgarno (1974). "The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites." *Proc Natl Acad Sci U S A* **71**(4): 1342-1346.
- Simonsen, K. T., G. Nielsen, et al. (2011). "A role for the RNA chaperone Hfq in controlling adherent-invasive *Escherichia coli* colonization and virulence." *PLoS One* **6**(1): e16387.
- Siryaporn, A. and M. Goulian (2008). "Cross-talk suppression between the CpxA-CpxR and EnvZ-OmpR two-component systems in *E. coli*." *Mol Microbiol* **70**(2): 494-506.
- Siryaporn, A., B. S. Perchuk, et al. (2010). "Evolving a robust signal transduction pathway from weak cross-talk." *Mol Syst Biol* **6**: 452.
- Sittka, A., S. Lucchini, et al. (2008). "Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq." *PLoS Genet* **4**(8): e1000163.
- Sittka, A., V. Pfeiffer, et al. (2007). "The RNA chaperone Hfq is essential for the virulence of *Salmonella typhimurium*." *Mol Microbiol* **63**(1): 193-217.

- Sittka, A., C. M. Sharma, et al. (2009). "Deep sequencing of Salmonella RNA associated with heterologous Hfq proteins in vivo reveals small RNAs as a major target class and identifies RNA processing phenotypes." *RNA Biol* **6**(3): 266-275.
- Skerker, J. M., B. S. Perchuk, et al. (2008). "Rewiring the specificity of two-component signal transduction systems." *Cell* **133**(6): 1043-1054.
- Skerker, J. M., M. S. Prasol, et al. (2005). "Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis." *PLoS Biol* **3**(10): e334.
- Slauch, J. M. and T. J. Silhavy (1989). "Genetic analysis of the switch that controls porin gene expression in Escherichia coli K-12." *J Mol Biol* **210**(2): 281-292.
- Sledjeski, D. and S. Gottesman (1995). "A small RNA acts as an antisilencer of the H-NS-silenced rcsA gene of Escherichia coli." *Proc Natl Acad Sci U S A* **92**(6): 2003-2007.
- Smith, A. M., R. T. Fuchs, et al. (2010). "Riboswitch RNAs: regulation of gene expression by direct monitoring of a physiological signal." *RNA Biol* **7**(1): 104-110.
- Sorek, R., C. M. Lawrence, et al. (2013). "CRISPR-mediated adaptive immune systems in bacteria and archaea." *Annu Rev Biochem* **82**: 237-266.
- Spiro, S. and J. R. Guest (1990). "FNR and its role in oxygen-regulated gene expression in Escherichia coli." *FEMS Microbiol Rev* **6**(4): 399-428.
- Squire, D. J., M. Xu, et al. (2009). "Competition between NarL-dependent activation and Fis-dependent repression controls expression from the Escherichia coli yeaR and ogt promoters." *Biochem J* **420**(2): 249-257.
- Stewart, V. (1988). "Nitrate respiration in relation to facultative metabolism in enterobacteria." *Microbiol Rev* **52**(2): 190-232.
- Stewart, V. (2003). "Biochemical Society Special Lecture. Nitrate- and nitrite-responsive sensors NarX and NarQ of proteobacteria." *Biochem Soc Trans* **31**(Pt 1): 1-10.
- Stewart, V., P. J. Bledsoe, et al. (2003). "Dual overlapping promoters control napF (periplasmic nitrate reductase) operon expression in Escherichia coli K-12." *J Bacteriol* **185**(19): 5862-5870.
- Stewart, V. and J. Parales, Jr. (1988). "Identification and expression of genes narL and narX of the nar (nitrate reductase) locus in Escherichia coli K-12." *J Bacteriol* **170**(4): 1589-1597.
- Stincone, A., N. Daudi, et al. (2011). "A systems biology approach sheds new light on Escherichia coli acid resistance." *Nucleic Acids Res* **39**(17): 7512-7528.
- Stock, A. M., V. L. Robinson, et al. (2000). "Two-component signal transduction." *Annu Rev Biochem* **69**: 183-215.
- Suzuki, K., P. Babitzke, et al. (2006). "Identification of a novel regulatory protein (CsrD) that targets the global regulatory RNAs CsrB and CsrC for degradation by RNase E." *Genes Dev* **20**(18): 2605-2617.
- Suzuki, K., X. Wang, et al. (2002). "Regulatory circuitry of the CsrA/CsrB and BarA/UvrY systems of Escherichia coli." *J Bacteriol* **184**(18): 5130-5140.
- Takahashi, K., T. Hattori, et al. (1994). "Repression of in vitro transcription of the Escherichia coli fnr and nar X genes by FNR protein." *FEBS Lett* **340**(1-2): 59-64.
- Takesue, Y., H. Ohge, et al. (2002). "Bacterial translocation in patients with Crohn's disease undergoing surgery." *Dis Colon Rectum* **45**(12): 1665-1671.
- Tanaka, T., S. K. Saha, et al. (1998). "NMR structure of the histidine kinase domain of the E. coli osmosensor EnvZ." *Nature* **396**(6706): 88-92.
- Tawfik, A., P. K. Flanagan, et al. (2014). "Escherichia coli-host macrophage interactions in the pathogenesis of inflammatory bowel disease." *World J Gastroenterol* **20**(27): 8751-8763.
- Taylor, R. K., M. N. Hall, et al. (1983). "Isolation and characterization of mutations altering expression of the major outer membrane porin proteins using the local anaesthetic procaine." *J Mol Biol* **166**(3): 273-282.
- Teplitski, M., R. I. Goodier, et al. (2003). "Pathways leading from BarA/SirA to motility and virulence gene expression in Salmonella." *J Bacteriol* **185**(24): 7257-7265.
- Thomason, M. K., T. Bischler, et al. (2015). "Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in Escherichia coli." *J Bacteriol* **197**(1): 18-28.
- Thomason, M. K., F. Fontaine, et al. (2012). "A small RNA that regulates motility and biofilm formation in response to changes in nutrient availability in Escherichia coli." *Mol Microbiol* **84**(1): 17-35.

- Thomason, M. K. and G. Storz (2010). "Bacterial antisense RNAs: how many are there, and what are they doing?" *Annu Rev Genet* **44**: 167-188.
- Thompson, K. M., V. A. Rhodius, et al. (2007). "SigmaE regulates and is regulated by a small RNA in *Escherichia coli*." *J Bacteriol* **189**(11): 4243-4256.
- Tokishita, S., A. Kojima, et al. (1991). "Transmembrane signal transduction and osmoregulation in *Escherichia coli*. Functional importance of the periplasmic domain of the membrane-located protein kinase, EnvZ." *J Biol Chem* **266**(11): 6780-6785.
- Tomizawa, J., T. Itoh, et al. (1981). "Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA." *Proc Natl Acad Sci U S A* **78**(3): 1421-1425.
- Tomomori, C., T. Tanaka, et al. (1999). "Solution structure of the homodimeric core domain of *Escherichia coli* histidine kinase EnvZ." *Nat Struct Biol* **6**(8): 729-734.
- Tramonti, A., M. De Canio, et al. (2008). "GadX/GadW-dependent regulation of the *Escherichia coli* acid fitness island: transcriptional control at the *gadY-gadW* divergent promoters and identification of four novel 42 bp GadX/GadW-specific binding sites." *Mol Microbiol* **70**(4): 965-982.
- Tsui, H. C., H. C. Leung, et al. (1994). "Characterization of broadly pleiotropic phenotypes caused by an *hfq* insertion mutation in *Escherichia coli* K-12." *Mol Microbiol* **13**(1): 35-49.
- Tsui, P., L. Huang, et al. (1991). "Integration host factor binds specifically to multiple sites in the *ompB* promoter of *Escherichia coli* and inhibits transcription." *J Bacteriol* **173**(18): 5800-5807.
- Tyson, K. L., J. A. Cole, et al. (1994). "Nitrite and nitrate regulation at the promoters of two *Escherichia coli* operons encoding nitrite reductase: identification of common target heptamers for both NarP- and NarL-dependent regulation." *Mol Microbiol* **13**(6): 1045-1055.
- Udekwu, K. I., F. Darfeuille, et al. (2005). "Hfq-dependent regulation of *OmpA* synthesis is mediated by an antisense RNA." *Genes Dev* **19**(19): 2355-2366.
- Udekwu, K. I. and E. G. Wagner (2007). "Sigma E controls biogenesis of the antisense RNA *MicA*." *Nucleic Acids Res* **35**(4): 1279-1288.
- Ulbrandt, N. D., J. A. Newitt, et al. (1997). "The *E. coli* signal recognition particle is required for the insertion of a subset of inner membrane proteins." *Cell* **88**(2): 187-196.
- Ulrich, L. E. and I. B. Zhulin (2010). "The MiST2 database: a comprehensive genomics resource on microbial signal transduction." *Nucleic Acids Res* **38**(Database issue): D401-407.
- Uden, G. and J. Bongaerts (1997). "Alternative respiratory pathways of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors." *Biochim Biophys Acta* **1320**(3): 217-234.
- Urban, J. H., K. Papenfort, et al. (2007). "A conserved small RNA promotes discoordinate expression of the *glmUS* operon mRNA to activate *GlmS* synthesis." *J Mol Biol* **373**(3): 521-528.
- Urbanowski, M. L., L. T. Stauffer, et al. (2000). "The *gcvB* gene encodes a small untranslated RNA involved in expression of the dipeptide and oligopeptide transport systems in *Escherichia coli*." *Mol Microbiol* **37**(4): 856-868.
- Utsumi, R., R. E. Brissette, et al. (1989). "Activation of bacterial porin gene expression by a chimeric signal transducer in response to aspartate." *Science* **245**(4923): 1246-1249.
- Vanderpool, C. K. and S. Gottesman (2004). "Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system." *Mol Microbiol* **54**(4): 1076-1089.
- Vanderpool, C. K. and S. Gottesman (2004). "Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system." *Mol. Microbiol.* **54**: 1076-1089.
- Vecerek, B., I. Moll, et al. (2005). "Translational autocontrol of the *Escherichia coli* *hfq* RNA chaperone gene." *RNA* **11**(6): 976-984.
- Vieu, E. (2004). Dissection du mécanisme de terminaison/antiterminaison au niveau du terminateur du phage lambda : application à l'étude des complexes ARN-protéines *in vivo*, UNIVERSITE D'ORLEANS. DOCTEUR DE L'UNIVERSITE D'ORLEANS.
- Villa, E., J. Sengupta, et al. (2009). "Ribosome-induced changes in elongation factor Tu conformation control GTP hydrolysis." *Proc Natl Acad Sci U S A* **106**(4): 1063-1068.

- Vo, N. V., L. M. Hsu, et al. (2003). "In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 3. Influences of individual DNA elements within the promoter recognition region on abortive initiation and promoter escape." *Biochemistry* **42**(13): 3798-3811.
- Vogel, J., L. Argaman, et al. (2004). "The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide." *Curr Biol* **14**(24): 2271-2276.
- Vogel, J., V. Bartels, et al. (2003). "RNomics in Escherichia coli detects new sRNA species and indicates parallel transcriptional output in bacteria." *Nucleic Acids Res* **31**(22): 6435-6443.
- Vogt, S. L., A. D. Evans, et al. (2014). "The Cpx envelope stress response regulates and is regulated by small noncoding RNAs." *J Bacteriol* **196**(24): 4229-4238.
- Vvedenskaya, I. O., J. S. Sharp, et al. (2012). "Growth phase-dependent control of transcription start site selection and gene expression by nanoRNAs." *Genes Dev* **26**(13): 1498-1507.
- Walthers, D., V. K. Tran, et al. (2003). "Interdomain linkers of homologous response regulators determine their mechanism of action." *J Bacteriol* **185**(1): 317-324.
- Wang, H. and R. P. Gunsalus (2000). "The nrfA and nirB nitrite reductase operons in Escherichia coli are expressed differently in response to nitrate than to nitrite." *J Bacteriol* **182**(20): 5813-5822.
- Wang, H. and R. P. Gunsalus (2003). "Coordinate regulation of the Escherichia coli formate dehydrogenase fdnGHI and fdhF genes in response to nitrate, nitrite, and formate: roles for NarL and NarP." *J Bacteriol* **185**(17): 5076-5085.
- Wang, H., C. P. Tseng, et al. (1999). "The napF and narG nitrate reductase operons in Escherichia coli are differentially expressed in response to submicromolar concentrations of nitrate but not nitrite." *J Bacteriol* **181**(17): 5303-5308.
- Wang, L. C., L. K. Morgan, et al. (2012). "The inner membrane histidine kinase EnvZ senses osmolality via helix-coil transitions in the cytoplasm." *EMBO J* **31**(11): 2648-2659.
- Wang, X., A. K. Dubey, et al. (2005). "CsrA post-transcriptionally represses pgaABCD, responsible for synthesis of a biofilm polysaccharide adhesin of Escherichia coli." *Mol Microbiol* **56**(6): 1648-1663.
- Wang, Y. and P. L. deHaseh (2003). "Sigma 32-dependent promoter activity in vivo: sequence determinants of the groE promoter." *J Bacteriol* **185**(19): 5800-5806.
- Wassarman, K. M., F. Repoila, et al. (2001). "Identification of novel small RNAs using comparative genomics and microarrays." *Genes Dev* **15**(13): 1637-1651.
- Waters, L. S. and G. Storz (2009). "Regulatory RNAs in bacteria." *Cell* **136**(4): 615-628.
- Weber, H., C. Pesavento, et al. (2006). "Cyclic-di-GMP-mediated signalling within the sigma network of Escherichia coli." *Mol Microbiol* **62**(4): 1014-1034.
- Wei, B. L., A. M. Brun-Zinkernagel, et al. (2001). "Positive regulation of motility and flhDC expression by the RNA-binding protein CsrA of Escherichia coli." *Mol Microbiol* **40**(1): 245-256.
- Williams, S. B. and V. Stewart (1997). "Discrimination between structurally related ligands nitrate and nitrite controls autokinase activity of the NarX transmembrane signal transducer of Escherichia coli K-12." *Mol Microbiol* **26**(5): 911-925.
- Wilson, C. J., H. Zhan, et al. (2007). "The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding." *Cell Mol Life Sci* **64**(1): 3-16.
- Winkler, W. C., A. Nahvi, et al. (2004). "Control of gene expression by a natural metabolite-responsive ribozyme." *Nature* **428**(6980): 281-286.
- Wright, P. R., A. S. Richter, et al. (2013). "Comparative genomics boosts target prediction for bacterial small RNAs." *Proc Natl Acad Sci U S A* **110**(37): E3487-3496.
- Wu, H., K. L. Tyson, et al. (1998). "Regulation of transcription initiation at the Escherichia coli nir operon promoter: a new mechanism to account for co-dependence on two transcription factors." *Mol Microbiol* **27**(2): 493-505.
- Wurtzel, E. T., M. Y. Chou, et al. (1982). "Osmoregulation of gene expression. I. DNA sequence of the ompR gene of the ompB operon of Escherichia coli and characterization of its gene product." *J Biol Chem* **257**(22): 13685-13691.
- Xiao, M., Y. Lai, et al. (2016). "Transcriptional Regulation of the Outer Membrane Porin Gene ompW Reveals its Physiological Role during the Transition from the Aerobic to the Anaerobic Lifestyle of Escherichia coli." *Front Microbiol* **7**: 799.
- Xu, M., S. J. Busby, et al. (2009). "Activation and repression at the Escherichia coli ynfEFGHI operon promoter." *J Bacteriol* **191**(9): 3172-3176.

- Yamamoto, K., K. Hirao, et al. (2005). "Functional characterization in vitro of all two-component signal transduction systems from *Escherichia coli*." *J Biol Chem* **280**(2): 1448-1456.
- Yamamoto, K., R. Nagura, et al. (2000). "Negative regulation of the *bolA1p* of *Escherichia coli* K-12 by the transcription factor *OmpR* for osmolarity response genes." *FEMS Microbiol Lett* **186**(2): 257-262.
- Yang, Q., N. Figueroa-Bossi, et al. (2014). "Translation enhancing ACA motifs and their silencing by a bacterial small regulatory RNA." *PLoS Genet* **10**(1): e1004026.
- Yanofsky, C. (1981). "Attenuation in the control of expression of bacterial operons." *Nature* **289**(5800): 751-758.
- Yoshida, M., A. Ishihama, et al. (2015). "Cross talk in promoter recognition between six NarL-family response regulators of *Escherichia coli* two-component system." *Genes Cells* **20**(7): 601-612.
- Yu, D. G., H. M. Ellis, et al. (2000). "An efficient recombination system for chromosome engineering in *Escherichia coli*." *Proc. Natl. Acad. Sci. USA* **97**(11): 5978-5983.
- Zengel, J. M. and L. Lindahl (1994). "Diverse mechanisms for regulating ribosomal protein synthesis in *Escherichia coli*." *Prog Nucleic Acid Res Mol Biol* **47**: 331-370.
- Zhang, A., D. J. Schu, et al. (2013). "Mutations in interaction surfaces differentially impact *E. coli* Hfq association with small RNAs and their mRNA targets." *J Mol Biol* **425**(19): 3678-3697.
- Zhang, A., K. M. Wassarman, et al. (2003). "Global analysis of small RNA and mRNA targets of Hfq." *Mol Microbiol* **50**(4): 1111-1124.
- Zhu, Y., L. Qin, et al. (2000). "Phosphatase activity of histidine kinase EnvZ without kinase catalytic domain." *Proc Natl Acad Sci U S A* **97**(14): 7808-7813.
- Ziolkowska, K., P. Derreumaux, et al. (2006). "Hfq variant with altered RNA binding functions." *Nucleic Acids Res* **34**(2): 709-720.
- Zubay, G., D. Schwartz, et al. (1970). "Mechanism of activation of catabolite-sensitive genes: a positive control system." *Proc Natl Acad Sci U S A* **66**(1): 104-110.
- Zusman, D. R., A. E. Scott, et al. (2007). "Chemosensory pathways, motility and development in *Myxococcus xanthus*." *Nat Rev Microbiol* **5**(11): 862-872.