

ÉTUDE DES MÉCANISMES DE RÉGULATION DU
RIBORÉGULATEUR LYSINE CHEZ *ESCHERICHIA COLI*

par

Pierre Turcotte

mémoire présenté au département de biologie en vue
de l'obtention du grade de maître ès science (M.Sc.)

FACULTÉ DES SCIENCES
UNIVERSITÉ DE SHERBROOKE

Sherbrooke, Québec, Canada septembre 2019

Le 9 septembre 2019

Le jury a accepté le mémoire de
Pierre Turcotte

Membres du jury

Professeur Daniel Lafontaine
Directeur de recherche
Département de biologie

Professeur Éric Massé
Membre externe
Département de biochimie

Professeur Pierre-Étienne Jacques
Président-rapporteur
Département de biologie

SOMMAIRE

Les riborégulateurs sont un sujet d'étude qui suscite l'intérêt dans la communauté scientifique depuis le début des années 2000. Malgré cet intérêt depuis pratiquement vingt ans, beaucoup de choses sont encore inconnues à propos de riborégulateurs. En effet, certains groupes se sont attardés à découvrir le plus de riborégulateurs possible sans en faire de recherche approfondie à chaque fois. Il est cependant important de comprendre les mécanismes de transcription, le repliement et la régulation des riborégulateurs afin de pouvoir développer des outils de biologie moléculaire ou des antibiotiques par exemple.

Dans ce mémoire il sera principalement question de la régulation du riborégulateur chez *Escherichia coli* ayant pour ligand l'acide aminé lysine (*lysC*). Une étude des modes de régulations a déjà mis en lumière le mécanisme de régulation de la traduction ainsi que l'implication du dégradosome ARN chez ce riborégulateur. Cependant, certains résultats préliminaires nous laissaient croire que la protéine Rho pouvait être impliquée dans la régulation du niveau d'ARN messager de *lysC*. Nous avons donc commencé par étudier brièvement la transcription du riborégulateur ainsi que son changement de conformation en réponse au ligand. Ceci nous a permis d'utiliser les bonnes concentrations de ligand pour étudier la régulation de la protéine Rho. Ces différents modes de régulations présents chez un même riborégulateur nous amènent à penser que ce domaine est beaucoup plus complexe que ce que certains groupes pensaient au départ. C'est un domaine d'étude intéressant dans lequel on a découvert que la pointe de l'iceberg.

Mots clés : Riborégulateur, *lysC*, lysine, régulation, traduction, dégradosome, Rho.

REMERCIEMENTS

J'aimerais d'abord remercier mon directeur de recherche Pr. Daniel Lafontaine pour m'avoir accueilli dans son laboratoire pour ma maîtrise. Merci de m'avoir soutenu et de m'avoir accompagné dans la résolution de problèmes scientifiques.

Je tiens à remercier mes conseillers qui ont contribué à l'avancement de mon projet. Vous m'avez poussé à aller chercher plus loin et étendre mes connaissances lors de chacune des réunions que nous avons eues et ceci me sera utile toute ma vie. Merci pour l'évaluation de mon mémoire.

Merci à Anne-Marie et Mélanie de m'avoir fait confiance lors des entrevues. Merci à mon mentor Adrien Chauvier pour ta patience envers moi, ton partage de connaissance et ta passion contagieuse pour les sciences. Je te souhaite une excellente carrière de chercheur.

J'aimerais remercier les membres présents et passés du laboratoire Lafontaine : merci à Anthony, Laurène, Adrien, Mélanie, Sébastien, Frédéric, Marie-France, Vincent Patrick, Jonathan, Jean-Christophe, Jean-François, Aurélianne, Tithi, Shayma et Anne-Sophie de m'avoir soutenu tout le long de mon parcours.

Pour finir, j'aimerais remercier famille et amis qui m'ont soutenu tout au long de mon parcours. Merci à mon père qui me posait des questions sur mon projet. Merci à ma mère qui a toujours été là pour me soutenir. Un merci spécial à la femme de ma vie qui depuis 4 ans, a su à la fois me soutenir et m'endurer dans les hauts et les bas autant au niveau personnel que professionnel. Merci, Sabrina Coulombe

TABLE DES MATIÈRES

SOMMAIRE	V
REMERCIEMENTS	VI
LISTE DES ABRÉVIATIONS	IX
LISTE DES TABLEAUX	XI
LISTE DES FIGURES	XII
CHAPITRE 1	1
INTRODUCTION GÉNÉRALE	1
1. L'ARN	1
1.1 La transcription	2
1.1.1 Initiation de la transcription	2
1.1.2 Élongation de la transcription	3
1.1.3 La polymérase d'ARN et ses multiples composantes.....	4
1.1.4 Différents facteurs sigma	5
1.2 Les ARN catalytiques	6
1.2.1 Ribonucléase P	7
1.2.2 Ribozyme hammerhead	7
1.2.3 Les ARN de transferts	8
1.3 Les riborégulateurs	10
1.3.1 Leur découverte	10
1.3.2 Anatomie d'un riborégulateur	10
1.3.3 Les classes de riborégulateurs.....	12
1.3.4 Composantes structurales.....	17
1.3.5 Modes de régulation des riborégulateurs	18
1.3.5.1 Régulation traductionnelle.....	18
1.3.5.2 Terminaison intrinsèque.....	18
1.3.5.3 Régulation par la protéine Rho	19
1.4 Dégradosome ARN	26
1.5 Riborégulateur <i>lysC</i>	30
1.5.1. Généralité	30
1.5.2 Structure	31
1.5.3 Liaison de la lysine	33
1.6 Riborégulateurs chez <i>E. coli</i>	34
1.6.1 Gène régulé.....	36
1.6.2 Aptamère, repliement et interactions tertiaires	37
1.6.3 Liaison de la lysine	38
1.6.4 Structure	38
1.6.5 Régulation	38
1.7 Objectifs de maîtrise	39

CHAPITRE 2.....	40
ÉTUDE DE LA TRANSCRIPTION ET DU CHANGEMENT DE CONFORMATION DU RIBORÉGULATEUR <i>IysC</i>.....	40
2.1 INTRODUCTION.....	40
2.2 MATÉRIELS ET MÉTHODES.....	40
2.2.1 Transcription <i>in vitro</i>	40
2.2.2 Identification des sites de pauses.....	41
2.2.3 Transcription des échelles de séquençage à l'aide de nucléotide 3' O-méthyle.....	42
2.2.5 Détermination du K_{switch}	43
2.3 RÉSULTATS.....	44
2.3.1 Transcription <i>in vitro</i>	44
2.3.2 Identification des sites de pauses.....	45
2.3.3 K_{switch}	47
CHAPITRE 3.....	51
ÉTUDE DE LA RÉGULATION DU RIBORÉGULATEUR <i>IysC</i> PAR LA PROTÉINE RHO.....	51
3.1 INTRODUCTION.....	51
3.2 MATÉRIELS ET MÉTHODES.....	52
3.2.1 Essai B-Gal.....	52
3.2.2 Essai de terminaison de la transcription <i>in vitro</i>	52
3.2.3 Prédiction bio-informatique.....	53
3.2.4 Identification du <i>site rut</i> de façon expérimentale.....	53
3.2.4.1 Transcription d'un ARN à l'aide de la polymérase du phage T7.....	53
3.2.4.2 Précipitation et purification des ARN.....	54
3.2.4.3 Déphosphorylation et marquage radioactif en 5' des ARN.....	54
3.2.4.4 Essai de retard sur gel après pontage aux UV.....	55
3.3 RÉSULTATS.....	56
3.3.1 Confirmation de la régulation de Rho par essai Beta-galactosidase.....	56
3.3.2 Localisation de la terminaison de la transcription par la protéine Rho.....	58
3.3.3 Identification du <i>site rut</i> par pontage <i>in vitro</i>	62
4. CONCLUSION.....	64
5. ANNEXE.....	68
6. BIBLIOGRAPHIE.....	70

LISTE DES ABRÉVIATIONS

ADN : Acide désoxyribonucléique

ARN : Acide ribonucléique

ARNm : ARN messenger

ARNp : ARN polymérase

ARNr : ARN ribosomique

ARNt : ARN de transfert

ATP : Adénosine triphosphate

ATPase : Adénosine triphosphatase

B. subtilis : *Bacillus subtilis*

Da : Dalton

E. coli : *Escherichia coli*

EDTA : Éthylènediaminetétraacétique

Facteur σ : Facteur sigma

FMN: Flavine mononucléotide

FRET: *Förster Resonance Energy Transfer*

HDV : *Hepatitis delta virus*

HMP: Hydroxy-méthyl-pyrimidine

Kd : Constante de dissociation

kDa : Kilo Dalton

Mg²⁺ : Ion magnésium

miRNA : Micro-ARN

NTP : Nucléotide triphosphate

Nt: nucléotide

PréQ1: 7-aminométhyl-7-déazaguanine

RBS: *Ribosome Binding Site*

RNase: Ribonucléase

Rut : *Rho Utilization Site*

SAM: S-adénosyl-méthionine

SAH : S-adénosyl-monocystéine

SD: Shine-Dalgarno

SDS : Dodécylsulfate de sodium

TPP : Thiamine pyrophosphate

UTR : *Untranslated Region*

μ : micro

Ψ : Psy

σ : sigma

γ : gamma

α : alpha

Δ : delta

ρ : rho

LISTE DES TABLEAUX

Tableau 1. Tableau récapitulatif des différents facteurs sigma, de leur contexte d'expression et leurs séquences reconnues.....	6
Tableau 2. Les différentes classes de riborégulateurs.....	13
Tableau 3. Tableau des oligonucléotides utilisés.....	68
Tableau 4. Constructions utilisées pour la transcription obtenues par PCR.....	69

LISTE DES FIGURES

Figure 1. Représentation schématique du dogme central de la biologie moléculaire.....	2
Figure 2. Représentation simplifiée de la polymérase ARN chez <i>E. coli</i> en complexe d'élongation.....	5
Figure 3. Représentation de la structure d'un ARN de transfert.....	9
Figure 4. Représentation simplifiée d'un riborégulateur.....	11
Figure 5. Représentation cristallographique des boucles P Q et R de la protéine Rho.....	23
Figure 6. Représentation cristallographique de la protéine Rho en "Open ring complex".....	25
Figure 7. Organisation de la RNase E.....	28
Figure 8. Repliement du riborégulateur <i>lysC</i>	32

Figure 9.	Représentation schématique des riborégulateurs présents chez <i>E. coli</i>	35
Figure 10.	Voie métabolique menant à la synthèse de la lysine.....	36
Figure 11.	Repliement 2D du riborégulateur <i>lysC</i> en présence de ligand chez <i>E. coli</i>	37
Figure 12.	Nucléotides 3' O-méthyle.....	42
Figure 13.	Essai de transcription de l'ADN <i>lysC</i>	45
Figure 14.	Pauses transcriptionnelles du riborégulateur <i>lysC</i>	47
Figure 15.	Détermination du K_{switch}	49
Figure 16.	Validation in vivo de l'effet de la protéine Rho.....	51
Figure 17.	Localisation des sites de terminaison par la protéine Rho.....	59

Figure 18.	Prédiction du site <i>rut</i>	61
Figure 19.	Localisation du <i>site rut</i> à l'aide du pontage aux UV.....	63

CHAPITRE 1

INTRODUCTION GÉNÉRALE

1. L'ARN

Le domaine de l'ARN est en constante évolution. La découverte concernant la structure de l'acide désoxyribonucléique (ADN), en 1953, fut un grand pas pour la biologie moléculaire et probablement indirectement un premier pas vers l'étude de l'ARN (Watson et Crick, 1953). C'est quelques années plus tard que la théorie voulant que l'ARN puisse être un intermédiaire entre l'ADN et la protéine fit son apparition. C'est en 1961 que le premier groupe apporte la preuve de cette théorie. Ce groupe a montré que l'ARNm d'un gène peut interagir avec un ribosome et donc mener à sa traduction (Jacob et Monod, 1961). C'est à ce moment qu'est né le paradigme de la biologie moléculaire. Celui-ci met en avant le principe selon lequel l'information génétique serait encodée dans l'ADN qui serait ensuite transcrite en ARN messager pour ensuite être traduite pour former l'unité active c'est-à-dire, la protéine (Figure 1). Cependant, ce paradigme est assez réducteur de la fonction étendue que peut avoir un ARN. La plupart des organismes ont des ARN qui ne codent pas pour des protéines. Par exemple, chez l'humain, il y a environ 3 % du génome qui est traduit en protéines alors qu'au moins 85 % du génome est transcrit en ARN (Hangauer, Vaughn et McManus, 2013). Il existe donc multitude de fonctions associées à différentes classes d'ARN non codant. Certains de ces ARN seront traités dans ce mémoire afin de nous mener au sujet principal : les riborégulateurs.



Figure 1. Représentation schématique du paradigme de la biologie moléculaire. L'ADN est la matrice contenant toute l'information génétique. Elle est transcrite en une copie temporaire : l'ARN messager. Ce messenger est alors traduit en protéine qui est l'unité effectrice. Les flèches de couleurs représentent les voies classiques alors que les flèches noires représentent des voies alternatives.

1.1 La transcription

La transcription est le processus par lequel la polymérase d'ARN (RNAP, de l'anglais *ribonucleic acid polymerase*) synthétise de l'ARN à partir d'une matrice d'ADN. La transcription nécessite la présence des quatre nucléosides triphosphates (ATP, UTP, CTP et GTP) ainsi que du magnésium divalent (Mg^{2+}) nécessaire à l'activité de l'enzyme. La transcription comporte trois étapes : l'initiation, l'élongation et la terminaison (Boyer, 2006).

1.1.1 Initiation de la transcription

Lors de l'étape d'initiation, la RNAP lie une séquence appelée promoteur qui se trouve au début du gène à amplifier et pour laquelle elle a une grande affinité. Les promoteurs procaryotes possèdent deux régions importantes : la région -35 et la région -10 ayant comme séquence consensus TTGACA et TATAAT. La polymérase peut se lier à l'ADN de façon non spécifique et se déplacer en

direction 5'-3'. La séquence est analysée jusqu'au moment où elle trouve une séquence promotrice (Barrow *et al.*, 1992). Lorsqu'elle la trouve, elle s'associe fortement avec la séquence d'ADN, composant le promoteur, ralentissant du même coup son déplacement sur la matrice. Ce ralentissement permet entre autres d'ouvrir la double hélice afin de pouvoir commencer la transcription (Haugen, Ross et Gourse, 2008). Cette section d'ADN ouverte est appelée bulle de transcription et est longue de 12 à 15 nucléotides (Zuo et Steitz, 2015). La transcription débute avec l'entrée et l'incorporation des trinucleotides complémentaires à la séquence d'ADN à transcrire. La transcription se poursuit de l'extrémité 5' vers le 3' de l'ARN en effectuant à chaque fois une attaque nucléophile du groupement OH sur le phosphate α , ce qui cause le relargage d'un pyrophosphate. Lors de la transcription, il peut arriver qu'il y ait certains transcrits dits abortifs avant l'incorporation du 10^e nucléotide (Liu et Martin, 2009). En effet, dans la littérature, il est décrit que le complexe transcriptionnel est plus stable une fois passé ce cap (Burgess *et al.*, 1969). Le facteur sigma quitte alors la polymérase, qui subit un réarrangement structurel et passe en phase d'élongation.

1.1.2 Élongation de la transcription

Une fois ce réarrangement de la polymérase terminé, la phase d'élongation débute. La polymérase ajoute les ribonucleotides triphosphate du 5' vers le 3' complémentaires au brin matrice de l'ADN qui est en sens inverse. Lorsqu'elle atteint la phase d'élongation, la polymérase devient très processive. Elle peut atteindre une vitesse de 10 à 42 nucléotides incorporés à la seconde, pour une moyenne de 20 nucléotides par seconde à une concentration de 0,5 mM de nucléotides (Tolic *et al.*, 2003). La polymérase continue à transcrire jusqu'à ce qu'elle rencontre un terminateur de la transcription ou jusqu'au moment où elle

se fait décrocher par une autre protéine. L'élongation peut être influencée par plusieurs facteurs tels que la vitesse d'incorporation de nucléotides, les sites de pauses ainsi que la longueur du gène à transcrire. Parfois, il peut aussi y avoir des atténuateurs de la transcription ayant un effet dépendant des conditions cellulaires lors de la transcription (Griffiths *et al.*, 2008). L'atténuateur le plus connu chez *E. coli* est sans doute l'opéron tryptophane. Le principe est le suivant : lorsqu'il y a une concentration élevée de tryptophane dans la cellule bactérienne, il y a la formation d'une tige boucle formant du même coup un terminateur de transcription. Ceci permet d'arrêter la production du tryptophane. Lorsque le niveau de tryptophane redescend, il y a un *stalling* du ribosome ce qui rend possible la formation d'une tige permettant que la transcription suive son cours (Yanofsky, 1981).

1.1.3 La polymérase d'ARN et ses multiples composantes

L'enzyme se compose de cinq sous-unités ($\alpha_2\beta\beta'\omega$) avec une masse d'environ 400 000 Daltons (voir Figure 2). Ces sous-unités sont relativement conservées de la bactérie à l'humain en matière de séquences, structures et fonctions (Ebright, 2000). Les sous-unités β et β' interagissent et forment une structure ressemblant à une pince et on y retrouve le site actif à sa base. Les sous-unités α_1 et α_2 interagissent respectivement avec β et β' et sont impliquées dans les interactions avec l'ADN ou bien encore avec d'autres protéines comme la protéine Rho (décrite dans la section 1.3.5.3).

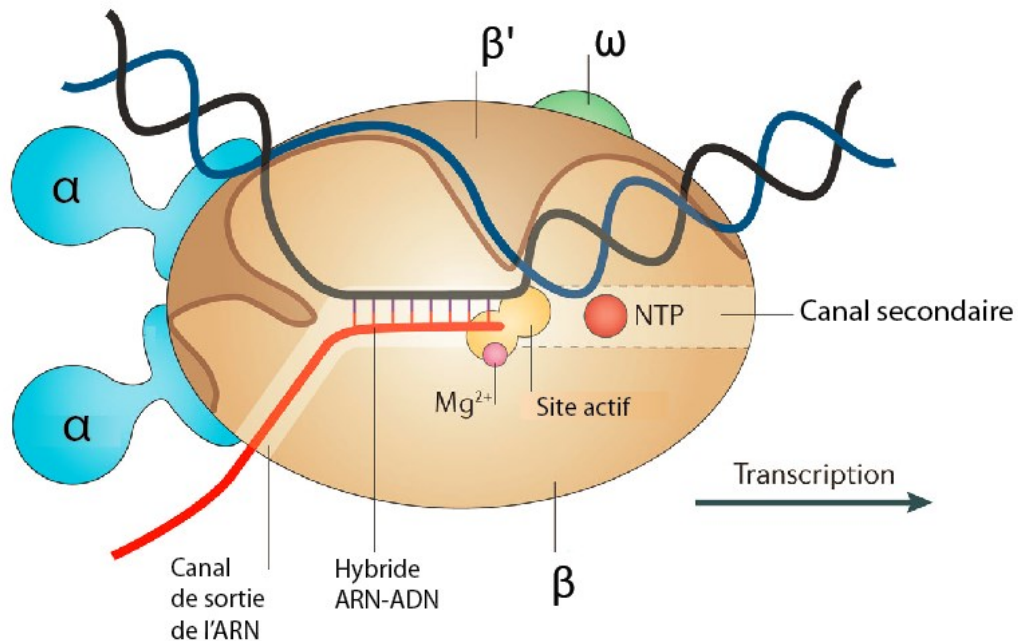


Figure 2. Représentation simplifiée de la polymérase ARN chez *E. coli* en complexe d'élongation. On peut voir ici comment les différentes sous-unités s'assemblent afin de créer une polymérase active avec toutes ses caractéristiques, comme les différents canaux. Adapté de (Santangelo et Artsimovitch, 2011).

1.1.4 Différents facteurs sigma

Chez *E. coli*, il existe 7 facteurs sigma différents (tableau 1). Cette variété de facteurs sigma permet de reconnaître différents promoteurs et d'exprimer différents groupes de gènes. Il existe des séquences consensus qui sont reconnues par les différents facteurs sigma. Plus la séquence du promoteur est similaire au consensus, plus elle sera reconnue fortement. Ceci explique en

partie pourquoi certains gènes sont plus exprimés que d'autres. De plus, il existe une différence d'utilisation des facteurs sigma selon la phase de croissance dans laquelle la bactérie se trouve. La phase de croissance, ainsi que les conditions dans lesquelles la bactérie est cultivée affectent la concentration des facteurs sigmas (Wösten, 1998).

Tableau 1. Tableau récapitulatif des différents facteurs sigma, de leur contexte d'expression et leurs séquences reconnues.

Famille σ^{70}	Protéine chez <i>E. coli</i>	Contexte d'expression	Séquence consensus
			-35 Espaceur -10
σ^{70}	RpoD	Phase exponentielle	TTGACA 16-18 nt TATATT
σ^{70}	RpoS	Phase stationnaire	CTATACT
σ^{70}	FliA	Flagelle et biofilm	TAAA 15 nt GCCGATAA
σ^{70}	RpoE	Choc au froid	GAACCTT 16-17 nt TCTRA
σ^{70}	RpoH	Choc thermique	CTTGAAA 17-16 nt CCCATNT
σ^{70}	Fecl	Carence en fer	Non déterminée
Famille σ^{54}	Protéine chez <i>E. coli</i>	Contexte d'expression	Séquence consensus
			-35 Espaceur -10
σ^{54}	RpoN	Métabolisme de l'azote	TGGCAC 5 TTGCW

1.2 Les ARN catalytiques

Comme mentionné dans la section traitant de l'ARN et du paradigme de la biologie moléculaire, cette molécule était considérée seulement comme une copie temporaire de l'ADN servant de messenger pour produire les protéines. Toutefois, vers la fin des années 1960, Carl Woese propose que l'ARN puisse

posséder une activité catalytique. C'est au tournant des années 80 que l'existence d'autres exemples d'ARN catalytiques ont été mis en lumière par les travaux sur l'épissage et la découverte de la ribonucléase P. Ces travaux ont mis en évidence l'activité autocatalytique de certains ARN. Ces ARN ont été nommés Ribozymes (enzymes ribonucléiques) (Kruger *et al.*, 1982).

1.2.1 Ribonucléase P

La ribonucléase P est une endoribonucléase responsable de la maturation des ARN de transferts. Bien que sa composition et sa structure ne soient pas conservées, on peut cependant observer que l'activité de maturation des ARN est conservée à travers tous les domaines de la vie. Elle a en fait pour rôle de cliver l'extension 5' de l'ARN laissant un 5' monophosphate. Chez les bactéries, la RNase P est composée d'un ARN catalytique assisté par une protéine. La découverte de cette ribonucléase a valu le prix Nobel de chimie à Sidney Altman, en 1989 (Guerrier-Takada *et al.*, 1983) (Altman, 2011).

1.2.2 Ribozyme hammerhead

Le ribozyme tête de marteau (hammerhead) a été identifié en premier chez certains ARN satellites ainsi que chez les viroïdes. C'est par la suite qu'on s'est rendu compte qu'il était très répandu à travers les domaines de la vie (Seehafer *et al.*, 2011). Il doit son nom au fait que sa structure secondaire ressemble à une tête de marteau. Le hammerhead est une sorte d'ARN catalytique ayant comme particularité de pouvoir se cliver lui-même à un endroit très précis. Cette particularité a été mise en lumière en 1986 chez les viroïdes (Buzayan, Gerlach et Bruening, 1986; Hutchins *et al.*, 1986). Cette capacité à s'autocliver permet

d'effectuer des expériences de clivage ciblé de l'ARN. Il est encore aujourd'hui utilisé comme outils de biologie moléculaire. De plus, certaines applications thérapeutiques sont envisagées (Hean et Weinberg 2008).

1.2.3 Les ARN de transferts

Les ARN de transferts ne sont pas des ARN catalytiques à proprement dit. Ils sont quand même dignes de mention. Ils sont l'une des espèces d'ARN non codants les plus répandus dans le domaine du vivant et certainement une des classes les plus étudiés au fil des ans. Leur rôle consiste à apporter les acides aminés aux ribosomes, pour qu'ils puissent produire les protéines. L'ARN de transfert possède plusieurs caractéristiques intéressantes. D'abord, bien que la majorité des nucléotides sont classiques (ATP, UTP, CTP, GTP), il y a un certain nombre de nucléotides qui subissent des modifications post-transcriptionnelles. Ces nucléotides sont modifiés en inosines, pseudouridines, dihydrouridines ou encore subissent une méthylation (Gerber et Keller, 1999; Gray et Gray, 2000). L'ARN de transfert adopte une structure secondaire très précise et très caractéristique. Cette structure est nommée *cloverleaf* (feuille de trèfle), car elle ressemble en effet à un trèfle (Figure 3). Cette structure complexe permet entre autres la reconnaissance de l'ARNt par la ribonucléase P, la reconnaissance du codon ainsi que la reconnaissance par le ribosome (Figure 4).

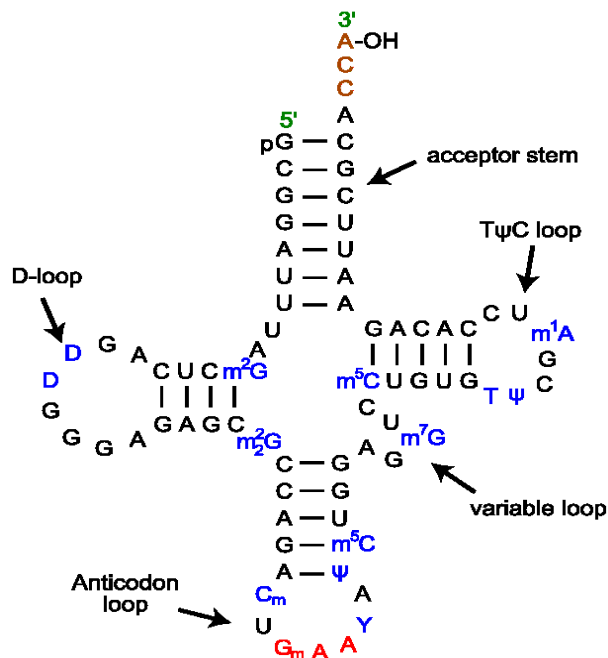


Figure 3. Représentation de la structure d'un ARN de transfert. On peut voir les quatre boucles soit: la boucle D, la boucle anticodon, la boucle variable et la boucle TΨC. Les nucléotides en bleu sont ceux ayant subi une modification alors que ceux en rouge représentent plutôt les nucléotides permettant la reconnaissance de l'acide aminé.

La structure de l'ARN de transfert est tellement importante qu'une seule mutation peut causer une maladie. C'est le cas dans l'ARN de transfert leucine dans les mitochondries. On peut parfois observer une mutation de l'adénine à la position 3243 qui est remplacée par une guanine (El-Hattab *et al.*, 2015). Cette mutation cause le syndrome de MELAS (Mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes). Ce syndrome cause entre autres maux de tête, vomissement et des symptômes semblables à un accident vasculaire. Cet exemple montre encore une fois l'importance de la séquence et la structure d'un ARN même s'il est non codant. C'est aussi le cas des riborégulateurs.

1.3 Les riborégulateurs

1.3.1 Leur découverte

Avant les années 1980-1990, il était inconcevable pour la communauté scientifique qu'un ARN soit capable de lier spécifiquement un métabolite. En effet, on pensait que seules les protéines avaient assez de diversité en acides aminés pour permettre de reconnaître une autre molécule. La découverte de l'opéron tryptophane a permis de commencer à changer les mentalités.

Par la suite, il y a eu la découverte des ribozymes (voir section ARN catalytique). Ces ARN étaient le premier exemple d'une fonction autre que celle de messenger. En effet, ils possèdent une activité catalytique. Puis certains groupes ont tenté de développer un ribozyme répondant à la liaison d'un ligand (Soukup et Breaker, 1999). Une fois cette étape réussie, certains autres groupes ont tenté de trouver dans la nature des ARN qui lient un ligand et qui ont une autre fonction que celle de messenger. En 2002, le groupe du Pr. Ronald Breaker et celui du Pr. Evgeny Nudler ont découvert une nouvelle classe d'ARN régulateur : les riborégulateurs (Mironov *et al.*, 2002 ; Winkler, Nahvi et Breaker, 2002).

1.3.2 Anatomie d'un riborégulateur

Les riborégulateurs sont des séquences d'ARN très structurées majoritairement retrouvées dans le 5' non traduit (5'UTR) des ARNm bactériens. Ces ARN sont composés de deux domaines distincts : l'aptamère et la plateforme d'expression (Figure 4).

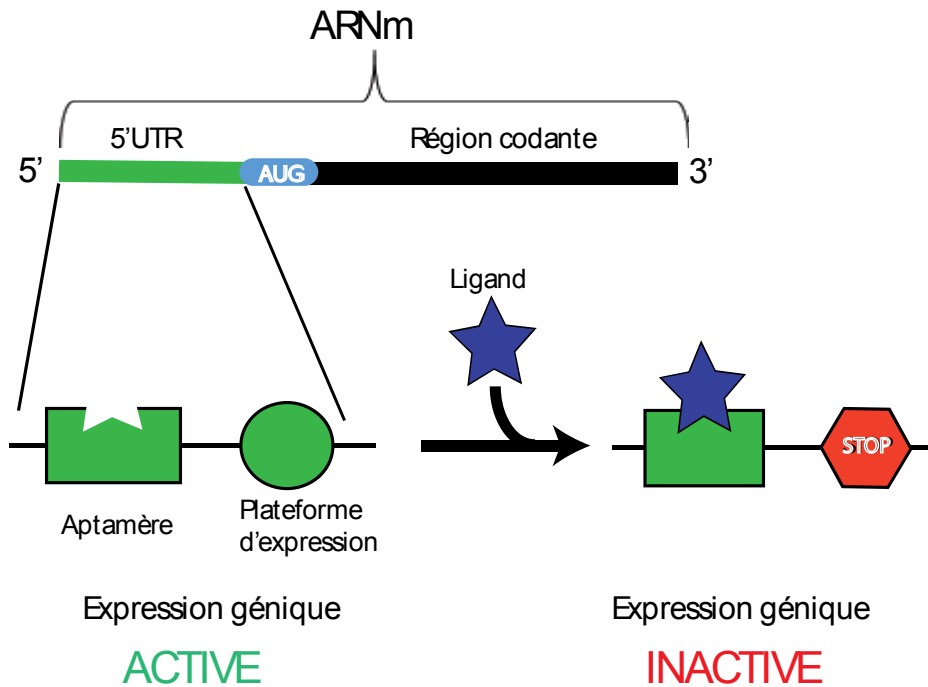


Figure 4. Représentation simplifiée d'un riborégulateur.

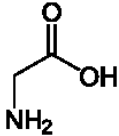
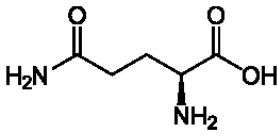
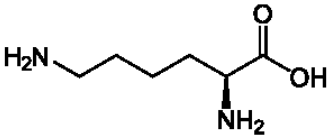
Les riborégulateurs sont des ARN non codants structurés retrouvés dans le 5' UTR. Lorsqu'il n'y a pas de ligand, le riborégulateur est dans une conformation permettant l'expression génique. À la suite de la liaison du ligand, il y a un changement de conformation de la plateforme d'expression ce qui va inactiver l'expression du gène en aval.

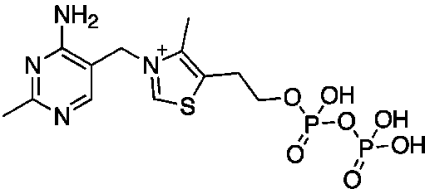
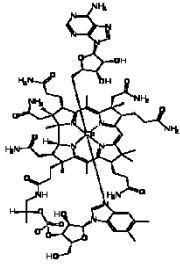
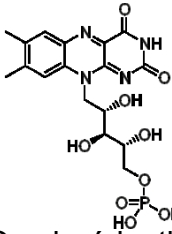
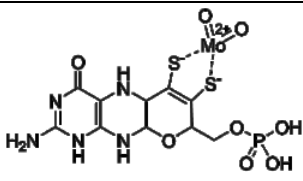
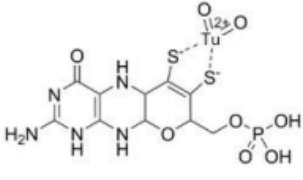
L'aptamère est la partie de l'ARN responsable de reconnaître un métabolite précis qu'on appelle ligand. Une fois cette liaison effectuée, cela provoque un changement de conformation de la plateforme d'expression. Cette dernière a pour rôle de moduler l'expression du gène ou de l'opéron en aval. La plupart des riborégulateurs sont dits "OFF" c'est-à-dire, que la liaison du ligand cause une inhibition de l'expression génique. Il s'agit donc d'un rétrocontrôle négatif (Mandal *et al.*, 2004).

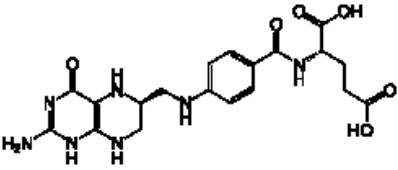
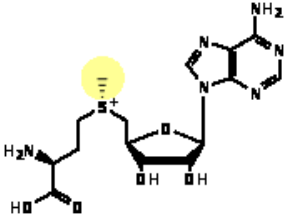
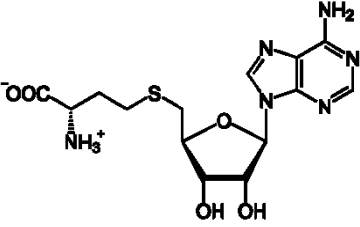
1.3.3 Les classes de riborégulateurs

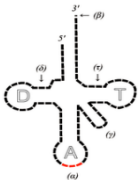
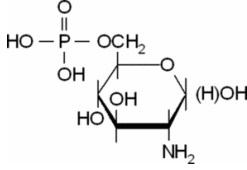
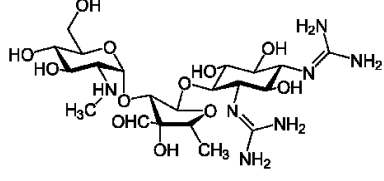
Depuis leur découverte, il y a une grande quantité de riborégulateurs qui ont été identifiés chez une multitude de bactéries allant de commune (*E. coli*) à plus « exotique » (*C. glutamicum*). Il est donc devenu nécessaire d'utiliser un système de classification des riborégulateurs (Tableau 2). À ce jour, il existe 38 classes différentes (McCown *et al.*, 2017). La classe d'un riborégulateur dépend de la nature du ligand qu'il lie. Il peut s'agir par exemple d'acides aminés (glycine, glutamine, lysine), de bases azotées classiques ou des substitués (adénine, guanine, pre-queuosine 1, etc.), de coenzyme (thiamine pyrophosphate, adénosine cobalamine, S-adénosyl-méthionine, etc.) d'ions inorganiques (fluor, magnésium et manganèse) d'autres types de molécules (ARNt, aminoglycoside) et même des paramètres physicochimiques (pH et température). Il existe certains cas où plusieurs classes de riborégulateurs lient le même ligand. C'est le cas chez les riborégulateurs liant le SAM et le pre-Q1.

Tableau 2. Les différentes classes de riborégulateurs.

Type de ligand	Nom du ligand	Structure et fonction	Découverte
Acides Aminés	Glycine	 <p>Composante des protéines</p>	(Mandal <i>et al.</i> , 2004)
	Glutamine	 <p>Composante des protéines</p>	(Ames and Breaker, 2011)
	Lysine	 <p>Composante des protéines</p>	(Sudarsan <i>et al.</i> , 2003)

Coenzymes	Thiamine pyrophosphate	 <p>Décarboxylation</p>	(Winkler, Nahvi and Breaker, 2002)
	Adénosyl Cobalamine	 <p>Transméthylation</p>	(Nahvi <i>et al.</i> , 2002)
	Flavine mononucléotide	 <p>Oxydoréduction</p>	(Mironov <i>et al.</i> , 2002)
Coenzyme (Suite)	Cofacteur molybdène	 <p>Oxydoréduction</p>	(Regulski <i>et al.</i> , 2008)
	Cofacteur tungstène	 <p>Oxydoréduction</p>	(Regulski <i>et al.</i> , 2008)

Coenzyme (suite)	Tétrahydrofolate		THF : (Ames <i>et al.</i> , 2010)
	S-adénosyl-Méthionine	<p>Métabolisme des purines + Acides aminés</p> 	<p>SAM-I : (Winkler <i>et al.</i>, 2003)</p> <p>Sam-II (Corbino <i>et al.</i>, 2005)</p> <p>SAM-III : (Fuchs, Grundy and Henkin, 2006)</p> <p>SAM-IV : (Weinberg <i>et al.</i>, 2008)</p> <p>SAM V : (Poiatea <i>et al.</i>, 2009) <i>et al.</i> 2009</p>
	S-adénosyl-homocystéine	<p>Transfert de groupements méthyle</p>  <p>Intermédiaire dans la synthèse de la cystéine</p>	SAH : (Wang and Breaker, 2008)

Cofacteur métallique	Fluor	F^-	(Baker <i>et al.</i> , 2012)
	Magnésium	Mg^{2+}	(Cromie <i>et al.</i> , 2006)
	Manganèse	Mn^{2+}	(Dambach <i>et al.</i> , 2015)
Paramètres physicochimiques	pH	H^+	(Nechooshtan <i>et al.</i> , 2009)
	Température	Température	(Morita <i>et al.</i> , 1999)
Autres	ARNt	 <p>Synthèse protéique</p>	(Grundy, Winkler and Henkin, 2002)
	Glucosamine 6-phosphate	 <p>Glucide</p>	(Winkler <i>et al.</i> , 2004)
	Aminoglycosides	 <p>Antibiotique</p>	(Jia <i>et al.</i> , 2013)

1.3.4 Composantes structurales

La structure secondaire et tertiaire est très importante dans la fonction des différents ARN. Plusieurs groupes ont étudié plus en détail les structures présentes chez les riborégulateurs. Ces interactions ont été démontrées comme importantes pour la fonction du riborégulateur notamment lors d'interaction entre les tiges de l'aptamère aidant du même coup à la réorganisation du cœur de l'aptamère (Blouin et Lafontaine, 2007 ; Blouin, Chinnappan, et Lafontaine, 2011). Il existe plusieurs types d'interaction. D'abord, il y a les interactions de type boucle-boucle. Ce type d'interaction est caractérisé par l'appariement de type Watson-Crick entre les nucléotides non appariés des boucles se trouvant à l'extrémité de deux tiges distinctes. Ce type d'interaction a été identifié chez plusieurs riborégulateurs (adénine, guanine et 2 -déoxyguanosine) ainsi que le riborégulateur lysine (Garst *et al.*, 2008 ; Serganov, Huang et Patel, 2008). On peut aussi observer des interactions de type hélices-boucles. Cette fois, la boucle interagit non pas avec une autre boucle, mais plutôt avec le sillon mineur de l'hélice d'ARN. On retrouve ce type d'interaction notamment chez les riborégulateurs TPP et lysine et AdoCbl (Serganov *et al.*, 2006 ; Garst *et al.*, 2008 ; Lussier *et al.*, 2015). On peut également observer des interactions de type pseudonœud. Un pseudonœud se forme à partir d'une structure en tige-boucle. Lorsque certains nucléotides de la boucle s'apparient avec une région de l'ARN située à l'extérieur de la tige il y a formation d'une structure avec deux hélices. Ce type d'interaction est retrouvé chez les riboswitch SAM-I, SAM-II SAH et pre-QI (Montange et Batey, 2006 ; Gilbert *et al.*, 2008 ; Lu *et al.*, 2008 ; Wang et Breaker, 2008 ; Jenkins *et al.*, 2011). Certains riborégulateurs comptent un autre type de structure importante : le *Kink-turn*. C'est le cas des riborégulateurs lysine, SAM-I et SAM-IV. Il s'agit d'une structure caractérisée par un coude (bulge) suivi de deux paires de bases A●G. Cet enchaînement spécifique permet à la tige d'atteindre un angle de 60° (Klein *et al.*, 2001). Les riborégulateurs

lysine, SAM-I et SAM-IV ont chacun un motif kink-turn dans leur tige P2 (Blouin and Lafontaine, 2007; Montange and Batey, 2006).

1.3.5 Modes de régulation des riborégulateurs

1.3.5.1 Régulation traductionnelle

Chez certains riborégulateurs, un mécanisme nommé régulation traductionnelle est utilisé. Cette régulation consiste en un changement de conformation de la plateforme d'expression à la suite de la liaison du ligand. Ce changement de conformation mène à la formation d'une tige qui séquestre les RBS et empêche ainsi l'initiation de la traduction (Caron *et al.*, 2012).

1.3.5.2 Terminaison intrinsèque

Chez les bactéries, le processus de terminaison de la transcription le plus retrouvé est celui de la terminaison intrinsèque. Cette terminaison survient lorsque la bactérie transcrit un terminateur canonique de la transcription. Ce terminateur est composé d'une tige boucle stable suivie d'une séquence dites poly-U qui comporte environ neuf résidus uraciles consécutifs (Reynolds, Bermúdez-Cruz and Chamberlin, 1992).

La transcription du terminateur cause une pause transcriptionnelle de la polymérase. Par la suite, la polymérase finit par décrocher. À ce jour, le mécanisme par lequel la polymérase termine la transcription est encore flou. Il existe pour le moment deux modèles proposés : le modèle de la translocation et le modèle allostérique.

Le premier modèle propose que la formation de la tige boucle appliquerait une force qui tire sur l'ARN. Cette force ferait en sorte que la polymérase serait déplacée en aval sans pouvoir continuer la transcription (Santangelo *et* Roberts, 2004).

Le deuxième modèle quant à lui propose que la formation de la tige mène à une modification de la conformation de la polymérase ce qui déstabilise la polymérase et la fait décrocher (Epshtein *et al.*, 2007).

1.3.5.3 Régulation par la protéine Rho

La protéine Rho est une protéine hexamérique en forme d'anneau de 267 kDa liant l'ARN qui a été découverte par J.W Roberts en 1969 (Roberts, 1969). Comme bien d'autre découverte en science, elle fut identifiée par erreur. En effet, les auteurs cherchaient à identifier des facteurs protéiques pouvant augmenter le niveau de transcription des ARNm dans la bactérie. Leurs expériences leur ont permis d'isoler une protéine qui régulait négativement la transcription de certains gènes. Par la suite, J.P Richardson a continué à étudier cette protéine. Il a notamment mis en évidence l'activité ATPase et hélicase de la protéine Rho. En effet, une fois que la protéine lie l'ARN elle peut se déplacer sur celle-ci grâce à son activité ATPase et défaire les structures en amonts.

1.3.5.3.1 Anatomie des domaines fonctionnels

Avant l'élucidation du cristal de la protéine complexée à des acides nucléiques, certains groupes ont tenté d'élucider le rôle des différents domaines de la protéine Rho (Bear *et al.*, 1985; Dolan, Marshall et Richardson, 1990; Wei et Richardson, 2001a, 2001b). Les informations de différentes études, le cristal et l'analyse de différents mutants ont permis d'identifier les différents domaines fonctionnels de la protéine. La principale région liant l'ARN se nomme domaine de liaison primaire. Elle est située entre les acides aminés 22 et 211 de la protéine (Modrak et Richardson, 1994). La boucle P qui est associée à la liaison de l'ATP et à son activité ATPase est située entre les acides aminés 179 et 183. Elle fait partie des ATPase de la famille RecA et est hautement conservée (Opperman et Richardson, 1994 ; Wei et Richardson, 2001a). La boucle Q se retrouve entre les acides aminés 278 et 290. Cette région comprend aussi le site de liaison secondaire à l'ARN (Wei et Richardson, 2001a, 2001 b ; Skordalakes et Berger, 2003). La boucle R quant à elle est une région un peu méconnue de la protéine. Elle se trouve entre les acides aminés 322-326, chez *E. coli* (Miwa, Horiguchi et Shigesada, 1995 ; Burgess et Richardson, 2001 ; Skordalakes et Berger, 2003).

1.3.5.3.2 Composition de la protéine Rho

On peut retrouver Rho dans plusieurs états d'assemblage en solution. La population majoritaire demeure cependant la forme totalement active : l'hexamère. Plus exactement, elle consiste en un trimère de dimère. Lorsque le transcrit est de 22 nucléotides ou moins, Rho s'assemble en dodécamère (Geiselman *et al.*, 1992).

L'assemblage de la protéine dépend principalement de l'environnement ionique de la solution, la concentration de la protéine et la présence ou non de cofacteurs (J Geiselman *et al.*, 1992).

Bien que ce soit encore sujet à débat, il est mentionné dans la littérature que la protéine Rho active peut exister sous deux formes : la forme "anneau ouvert" (*open ring complex*) ou "anneau fermé" (*closed ring complex*). Cependant, certains groupes tendent à montrer que les deux conformations existent. En effet, la conformation dite "anneau ouvert" serait transitoire alors que la conformation fermée serait adoptée une fois que le site de liaison secondaire a reconnu l'ARN (Figure 5) (Rabhi *et al.*, 2011).

1.3.5.3.3 Site primaire de liaison à l'ARN

Le domaine N-terminal de la protéine comporte trois hélices alpha et cinq feuillets bêta. Cet enchainement comprend le domaine de liaison des oligonucléotides ou oligosaccharides. Il y a en fait trois régions d'interaction distinctes comprises dans le N-terminal. Les résidus Leu58 et Phe62 créent une poche hydrophobe pour la liaison du ribose (Bogden *et al.*, 1999). Les acides aminés Glu108, Arg109, et Try110 forment une enceinte hydrophobe pouvant effectuer une interaction Van der Waals avec une base cytidine. Cette petite pochette de liaison est assez grande pour accommoder les pyrimidines, mais trop étroite pour les purines, ce qui explique que les sites *rut* sont riches en cytidine et pauvre en guanine. La troisième région est composée des acides aminés Arg66, Glu 56 et Asp78 qui ont pour rôle de discriminer encore plus les pyrimidines et les purines. La disposition de cette pochette et la capacité à faire des interactions de type Watson-Crick permet une préférence pour les cytidines par rapport aux uraciles. Le cristal de la protéine aura permis de démontrer que le site de liaison à l'ARN n'est pas en périphérie de la protéine telle que suggéré

auparavant. En effet, l'ARN suit plutôt un trajet de zigzag sur chacun des monomères de la protéine (Figure 5B) (Skordalakes and Berger, 2003).

1.3.5.3.4 Site secondaire de liaison à l'ARN

Le site de liaison de l'ARN secondaire de Rho est caractérisé par la boucle Q et R. Alors que la boucle R a des implications dans la liaison à la fois de l'ATP et de l'ARN (Xu, Kohn et Widger, 2002), la boucle Q s'étend quant à elle vers le centre de l'anneau. Elle est strictement destinée à la liaison de l'ARN. La proximité de la boucle R par rapport au domaine de liaison ou à la boucle P de l'ATP soulève une explication possible du couplage de la liaison de l'ARN à l'hydrolyse de l'ATP (Figure 5). Une fois que le site de liaison secondaire est lié à l'ARN la protéine passerait en conformation anneau fermé (Figure 6 A et B).

1.3.5.3.5 Liaison de l'ATP

Le domaine de liaison à l'ATP est situé à l'interface de la portion C-terminal entre les monomères adjacents. Une chose assez mystérieuse a été soulevée par certains groupes. Les six monomères composant la protéine ont tous une séquence identique et possèdent une densité d'électron égale au niveau de leur poche de liaison à l'ATP. Cependant, des études antérieures ont démontré que les six sous-unités de Rho sont fonctionnellement différentes en ce qui concerne la liaison de l'ATP (Stitt, 1988 ; Geiselmann *et al.*, 1992 ; Kim, Shigesada et Patel, 1999). Il existe trois sites de liaison forts et trois sites de liaison faibles pour l'ATP et l'ARN (Geiselmann *et al.*, 1992). Les sites de liaison forts effectuent une mauvaise hydrolyse de l'ATP alors que les sites de liaison faibles sont efficaces pour l'hydrolyse.

L'hydrolyse de l'ATP peut être ciblée pour inhiber l'action de Rho. En effet, il existe un inhibiteur non compétitif nommé bicyclomycine. Cette molécule lie le résidu Glu112 de la protéine Rho et empêche la liaison de la molécule d'eau nécessaire à l'hydrolyse de l'ATP (Skordalakes *et al.*, 2005; Furukawa *et al.*, 2015).

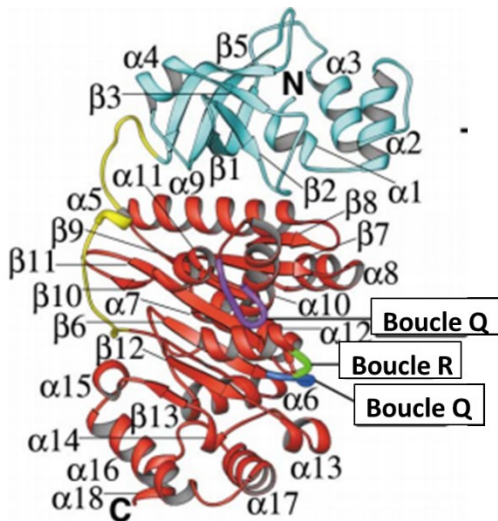


Figure 5. Représentation cristallographique des boucles P Q et R de la protéine Rho.

1.3.5.3.6 Site *rut* (Rho ut^{il}ization site) et sites de terminaison

Le site de liaison de la protéine Rho à l'ARN (*site rut*) est une séquence hautement dégénérée. À ce jour, aucune séquence consensus ou règle précise par rapport au *site rut* n'a été trouvée. Ce qu'on sait de ces sites, c'est qu'ils seraient long d'environ 80 nucléotides, riche en cytosine, faible en guanine et généralement non structuré (Banerjee *et al.*, 2006). Il est donc très difficile de prédire l'emplacement d'un *site rut* dans le génome d'*E. coli*. Récemment, un

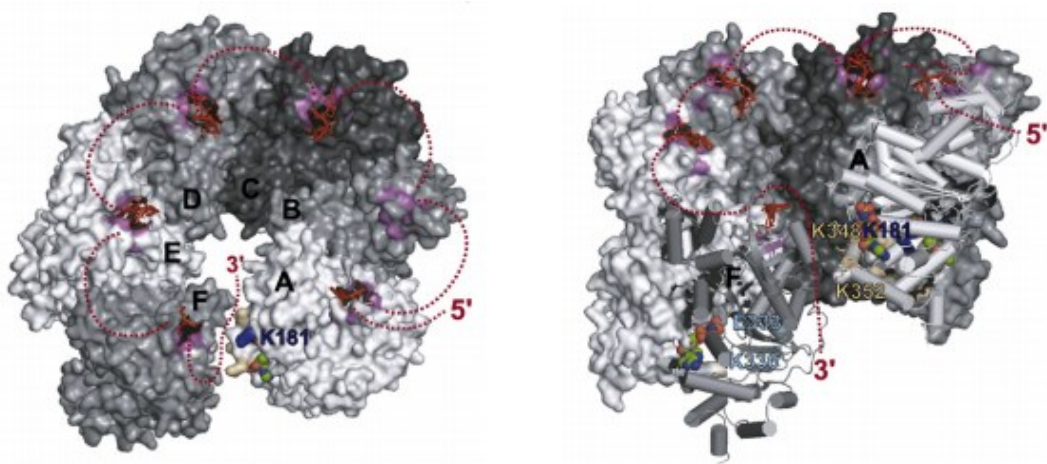
groupe a tenté de le faire en se focalisant sur les bulles de 80 nucléotides riches en C et faible en G dans le génome d'*E. coli*. Bien que cet algorithme soit un excellent point de départ, il n'est pas encore assez précis pour trouver de façon assez convaincante les sites *rut* (Nadiras *et al.*, 2018).

Pour ce qui est de la terminaison, il est important de savoir qu'on parle le plus souvent de plusieurs sites de terminaison plus ou moins bien définis plutôt qu'un seul site franc. Le premier site de terminaison est généralement situé à environ 60 à 90 nucléotides en aval du site de liaison de Rho. La protéine Rho régulerait un peu plus de 30 % des gènes chez *E. coli* (Dar et Sorek, 2018).

1.3.5.3.7 Étapes de la terminaison dépendante de la protéine Rho

La protéine Rho se lie à l'ARN non traduit sur un site riche en C (*site rut*). La liaison du *site rut* par la protéine Rho est réalisée par le domaine de liaison N-terminal aussi appelé primaire. Puis, il y a liaison de l'ARN au site de liaison secondaire (Figure 6B). Les sites secondaires liés à l'ARN activent l'hydrolyse de l'ATP. L'énergie dérivée de l'hydrolyse de l'ATP alimente l'activité translocase/hélicase de Rho pour dérouler le duplex d'ARN/ADN (Banerjee *et al.*, 2006). Par la suite, la transcription est terminée par un processus qui demeure inconnu, malgré les 49 dernières années consacrées à la compréhension de cette protéine.

A)



B)

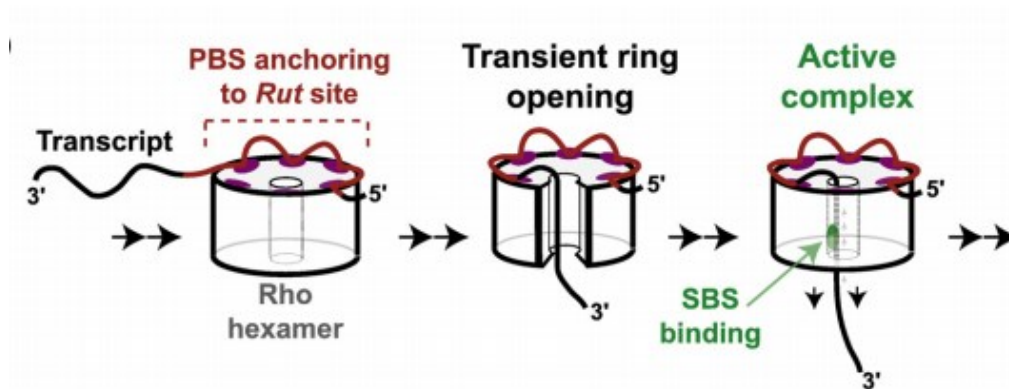


Figure 6. Représentation cristallographique de la protéine Rho en « Open ring complex».

A) Sur l'image de gauche, on voit la protéine avec une vision de haut. On peut voir les 6 sous-unités assemblées avec une légère ouverture (open ring). Le tracé en pointillé rouge représente les endroits où l'ARN contacte la protéine. L'image de droite est une rotation de 90 degrés par rapport à l'image de gauche. On peut ainsi mieux voir l'ARN qui passe par le centre de la protéine après avoir contacté les sites primaires de liaison des sous-unités. B) Représentation schématique des étapes de liaison entre l'ARN et la protéine Rho.

1.4 Dégradosome ARN

Une fois qu'il y a arrêt de l'expression d'un gène, la cellule se retrouve avec des ARN qui ne servent plus à rien dans l'immédiat. Il y a donc un processus de dégradation de l'ARNm qui est enclenché. Ces mécanismes sont étudiés depuis plusieurs années chez *E. coli*. En 1979, un groupe a trouvé l'identité de l'enzyme clé dans l'initiation de cette dégradation (Ono and Kuwano, 1979). Cette enzyme est la RNase E. Bien qu'elle soit capable de dégrader l'ARNm par ses propres moyens, elle forme un complexe avec d'autres protéines afin de faciliter la dégradation complète de l'ARNm (Khemici *et al.*, 2005). Elle peut s'associer à la Polynucléotide phosphorylase, afin de mieux dégrader les ARNm ainsi que pour effectuer la maturation de l'ARN ribosomal 5S (Carpousis *et al.*, 1994 ; Causton *et al.*, 1994). Lorsque l'hélicase RhIB et l'é nolase s'associent à leur tour, on appelle ce complexe le dégradosome (Miczak *et al.*, 1996 ; Py *et al.*, 1996). Le dégradosome est responsable de la majorité de la dégradation des ARNm chez *E. coli* (Bernstein *et al.*, 2002).

1.4.1 Endoribonucléase RNase E

La RNase E est une protéine de 118,22 kDa (1061 acide aminés) divisée en deux parties. Le domaine N-terminal possède l'activité catalytique qui permet l'endoclivage de l'ARN. La RNase E est en fait un homotétramère (Callaghan *et al.*, 2005) et c'est cet assemblage qui forme le site catalytique de la protéine (Figure 7A). Il est cependant possible de la trouver sous formes de dimère ou de trimère. Le domaine N terminal est lui-même sous-divisé en deux domaines : le grand domaine et le petit domaine. L'assemblage du tétramère est possible grâce à l'interaction des petits domaines les uns avec les autres. Le grand

domaine, quant à lui, a pour rôle la liaison et le positionnement de l'ARN. Il est divisé en sous-domaine. Le sous-domaine S1 est essentiel à l'activité de la RNase E. Il est impliqué dans la liaison et dans la dimérisation la protéine. Le sous-domaine RNase H a une grande ressemblance avec l'endoribonucléase H, mais ne possède pas d'activité catalytique. Il jouerait un rôle dans la liaison de la partie 5' préférentiellement monophosphorylé. Il est important de noter que le niveau de phosphorylation de l'ARN n'affecte pas la liaison de ce dernier (Jiang and Belasco, 2004), mais influence plutôt la vitesse de clivage (Mackie, 1998 ; Jiang et Belasco, 2004). En effet, la RNase E dégrade de 20 à 30 fois plus rapidement les ARN monophosphorylés par rapport à ceux triphosphorylés. Le domaine catalytique est retrouvé dans le sous-domaine DNase I, nommé ainsi à cause de sa similarité avec la DNase I. Lorsqu'un ARNm est en contact avec la RNase E, les sous-domaines S1 et DNase I se rapprochent. Ce rapprochement permet le clivage par l'attaque nucléophile sur le squelette de l'ARNm d'un ion magnésium, lié au résidu Asp303 et Asp346 du domaine DNase I (Koslover *et al.*, 2008).

La partie C-terminale de la RNase E est une longue région non structurée, où les autres protéines composant le dégradosome s'assemblent. Un monomère d'hélicase RhlB se lie aux acides aminés 698 à 762, un dimère d'énolase se lie aux acides aminés 823 à 850 et d'un trimère de PNPase entre la position 1039 et 1061 (Vanzo *et al.*, 1998; Coburn *et al.*, 1999; Symmons *et al.*, 2002; Callaghan *et al.*, 2004; Chandran and Luisi, 2006; Marcaida *et al.*, 2006)(Figure 7B).

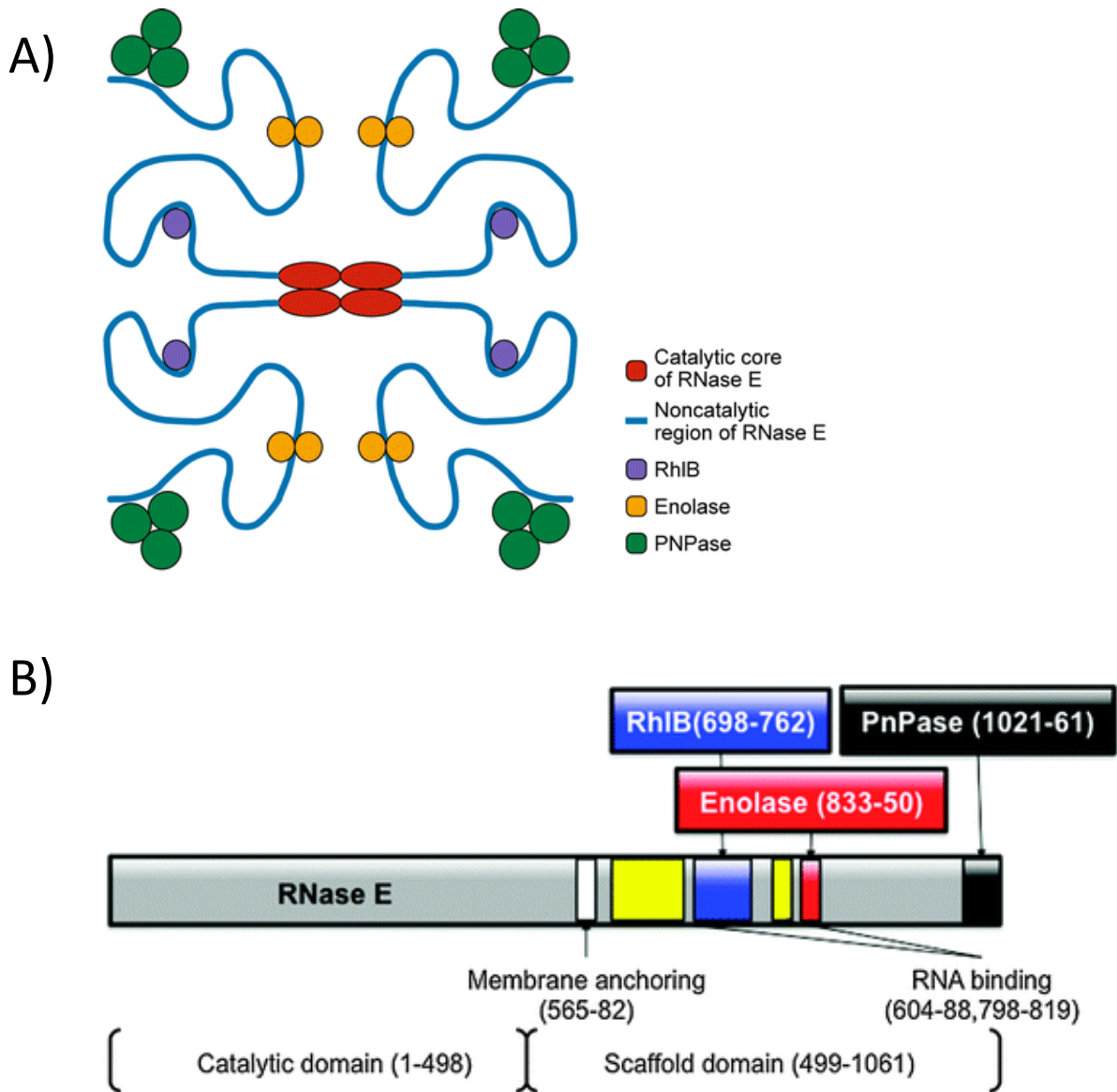


Figure 7. Organisation de la RNase E. A) Représentation schématique de la RNase E. Chaque RNase E est composée de quatre sous-unités identiques orientées tel que démontré ci-dessus. B) Représentation linéaire de la RNase E montrant ainsi les sites de liaisons des différentes protéines. Tiré de (Carpousis, 2007).

1.4.2 Polynucléotide phosphorylase (PNPase)

La polynucléotide phosphorylase (PNPase) est une exonucléase 3'-5'. C'est une enzyme très processive, mais elle est inhibée lorsqu'elle rencontre des structures d'ARN double brin (Blum et al., 1999 ; Spickler et Mackie, 2000). La PNPase nécessite une extrémité saillante de 7 à 10 nucléotides pour qu'elle puisse dégrader un ARN structuré (Coburn et al., 1999 ; Py et al., 1996). La PNPase requiert donc la coopération de protéines pouvant contourner cette limitation. C'est le cas de la polyA polymérase (PAP1) qui peut ajouter un poly A qui permet d'atteindre les 7 à 10 nucléotides nécessaires ou encore certaines hélicases permettant de défaire les structures (Blum et al., 1999 ; Carpousis et al., 1999 ; Coburn et al., 1999). Contrairement aux autres exonucléase 3-5' qui hydrolysent l'ARN, la PNPase est une enzyme phospholytique. Elle clive les liens phosphodiester reliant deux bases de l'ARN à l'aide d'un phosphate inorganique. Cette réaction est très près de l'équilibre sur le plan énergétique (Deutscher, 2006 ; Singer, 1958). La PNPase n'est pas essentielle chez *E. coli*. Cependant, lorsqu'une autre exonucléase 3'-5' est délétée, elle est capable de compenser afin de permettre la viabilité cellulaire (Deutscher, 2006). En plus de son rôle dans la dégradation des ARN, elle joue un rôle mineur dans la maturation des ARN ribosomiaux (Deutscher et al., 1988 ; Li et Deutscher, 1994).

1.4.3 Hélicase RhIB

Membre de la famille des DEAD-Box helicase, RhIB possède un domaine ATPase, un domaine hélicase et un domaine de liaison non spécifique à l'ARN en C-terminal. Son rôle dans le dégradosome consiste à défaire les structures secondaires de l'ARNm activant l'activité de la PNPase qui ne peut dégrader que les ARN simple brin. La PNPase peut interagir avec la partie C-terminale et ce

même sans la RNase E. Suite à cet assemblage, il a été observé que l'activité ATPase de RhlB est de 15 fois supérieur (Liou *et al.*, 2002).

1.4.4 Énolase

L'énolase est une enzyme dite glycolytique chez *E. coli*. Elle catalyse la conversion du 2-phosphoglycérate en phosphoénolpyruvate dans la néoglucogenèse. Elle est produite en grande quantité, mais seulement une faible proportion est liée au dégradosome (Py *et al.* 1996). Le rôle joué par l'énolase dans la dégradation des ARNm est encore peu connu. Cependant, il a été observé que l'énolase peut affecter la stabilité de plusieurs ARNm impliqués dans la génération d'énergie de la cellule (Lee, Bernstein et Cohen, 2002). Elle participe aussi à la réponse au stress des sucres phosphorylés (Morita, Maki et Aiba, 2005). Il a donc été proposé que l'énolase servirait à détecter l'état des réserves d'énergie de la cellule.

1.5 Riborégulateur *lysC*

1.5.1. Généralité

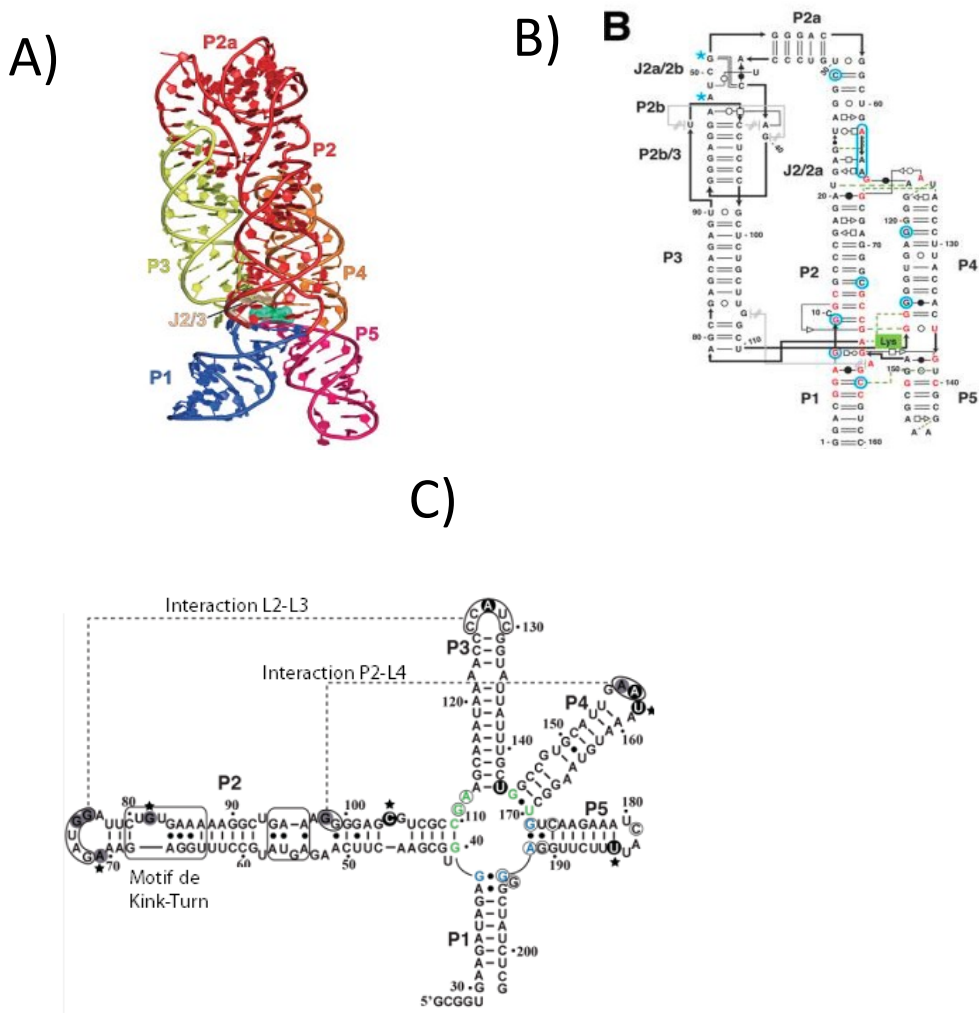
Le riborégulateur lysine, aussi appelé L-box ou LYS-element, est un motif ARN relativement abondant chez les bactéries. Dans une étude menée par le groupe de Breaker, il est démontré que, dans les 700 génomes bactériens analysés, on retrouve environ 600 riborégulateurs lysine (Breaker, 2012). Ces riborégulateurs sont présents principalement chez les firmicutes et gammaproteobacteria (Serganov, Huang et Patel, 2008). La découverte du riborégulateur lysine a été faite par Rodionov en 2003 (Rodionov *et al.*, 2003). Il a été découvert grâce à

des analyses bio-informatiques. La première étude sur ce riborégulateur est parue la même année. Ce groupe a démontré *in vitro* que, chez *B. subtilis*, la liaison de la lysine à l'aptamère induit un changement de conformation de la structure du riborégulateur (Sudarsan *et al.*, 2003). Chez *B. subtilis*, ce changement de conformation permet de stabiliser la tige P1 et de former un terminateur intrinsèque de la transcription au niveau de la plateforme d'expression.

1.5.2 Structure

Lorsqu'on regarde la structure bidimensionnelle du riborégulateur lysine, on peut remarquer plusieurs caractéristiques. D'abord, elle est composée de 5 tiges. Ces tiges sont nommées de P1 à P5 et ont en leur centre une portion qu'on appelle le cœur de l'aptamère c'est-à-dire, l'endroit où va se lier le ligand. Ces tiges sont capables d'interagir entre elles afin de former des interactions dites tertiaires. En 2008, le cristal du riborégulateur a été résolu. Il a été possible de voir qu'il y a un empilement (stacking) des tiges P1 et P2 ainsi que P4 et P5. Cette géométrie de l'aptamère ressemble en quelque sorte à un X asymétrique. Le cristal a aussi permis de visualiser l'interaction entre la boucle de la tige P2 et celle de la tige P3 (Garst *et al.*, 2008 ; Serganov, Huang et Patel, 2008). L'importance de cette interaction pour le bon repliement ainsi que la fonction du riborégulateur avait été observée en 2007, par le groupe du professeur Lafontaine (Blouin et Lafontaine, 2007). Le motif kink-turn dans la tige P2 serait quant à lui essentiel pour l'interaction avec la boucle de la tige P3. Toujours au niveau de l'aptamère, on remarque qu'il y a aussi une interaction entre la tige P2 et la boucle de la tige P4. Les bases adénines présentes dans la boucle font des liens avec le sillon mineur de l'hélice P2 via leur face Watson-Crick (Garst *et al.*, 2008 ; Serganov, Huang et Patel, 2008). Cette interaction est encore une fois

importante pour le repliement et la fonction et a été démontrée par le groupe du professeur Lafontaine, en 2011 (Blouin, Chinnappan et Lafontaine, 2011). De plus, il est également démontré que la longueur des tiges P1, P4 et P5 dans l'aptamère ainsi que la tige P6 présente dans la plateforme sont importantes afin de répondre de façon optimale à la présence de lysine.



En plus du repliement en 2D cette figure démontre les interactions L2-L3 et P2-L4 ainsi que les nucléotides y participant. Ces interactions sont cruciales pour la liaison du ligand. Adapté de (Blouin, Chinnappan et Lafontaine, 2011)

1.5.3 Liaison de la lysine

Le repliement tertiaire mentionné à la section précédente permet une organisation du cœur très précise (Figure 8). Ceci permet la création d'une pochette de liaison de la lysine (Serganov, Huang et Patel, 2008). L'identité des nucléotides composant ce cœur est très importante. En effet, une étude bio-informatique du groupe du Pr. Breaker en 2007 a démontré que les nucléotides du cœur sont très conservés à travers la phylogénie. La résolution du cristal chez *T. maritima* a permis de mieux comprendre comment s'effectuait la liaison. Le groupement aminé de la lysine fait deux ponts hydrogène avec le groupement hydroxyle et la base azotée de la base G140. Du côté du groupement carboxylé de la lysine, on peut voir avec le cristal qu'il y a une interaction avec la paire de bases wobble G114-U140, la paire de bases Watson-Crick G12-C79 et le groupement 2'-hydroxyle de la base G11 (Serganov, Huang et Patel, 2008). Toujours dans ce même cristal, on peut voir qu'un ion potassium est nécessaire à la liaison du ligand. Il interagit de façon stable avec le groupement carboxylé de la lysine ainsi que les bases G12 et G14, situées à la jonction de P1 et P2. Il est donc primordial d'utiliser des conditions avec la présence de potassium lors de l'étude du riborégulateur. On peut aussi noter que la portion N-terminale de la lysine interagit avec le cœur de l'aptamère via, une combinaison de diverses interactions électrostatiques.

En 2003, Sudarsan *et al.* ont démontré que le riborégulateur discriminait la L-lysine de la D-lysine par un facteur 50. C'était assez contrintuitif à l'époque, car les deux molécules possèdent la même longueur de chaîne ainsi que les mêmes

groupements chimiques (Sudarsan *et al.*, 2003). En fait, ces deux molécules ne diffèrent que par l'orientation de ces différents groupements. Pourtant, le riborégulateur lie la D-lysine avec peu d'affinité. Grâce au travail de quelques groupes, il a été possible de comprendre les caractéristiques nécessaires pour que l'aptamère lie le ligand. Tout d'abord, l'arrangement stéréochimique est très important comme en témoigne la discrimination L et D-lysine. Ensuite, la longueur de la molécule est importante étant donné que les groupements carboxylés et aminés participent à des interactions avec le cœur (Garst *et al.*, 2008). Le groupement C4 et N6 sont plus permissifs au changement. En effet, on peut mettre un groupement avec une taille plus importante en position N6 à condition qu'il soit de charge positive à pH neutre et qu'il soit en mesure de faire des ponts hydrogène (Blount *et al.*, 2007).

1.6 Riborégulateurs chez *E. coli*

Chez *E. coli*, il y a présentement 8 riborégulateurs qui ont été identifiés et confirmés (Figure 9). Ils lient différentes classes de ligands comme des acides aminés (lysine), des cofacteurs enzymatiques (thiamine pyrophosphate, adénosine cobalamine, flavine mononucléotide) ou encore des ions métalliques (magnésium et manganèse). Dans cette étude, il sera question du riborégulateur *lysC* chez *E. coli*.

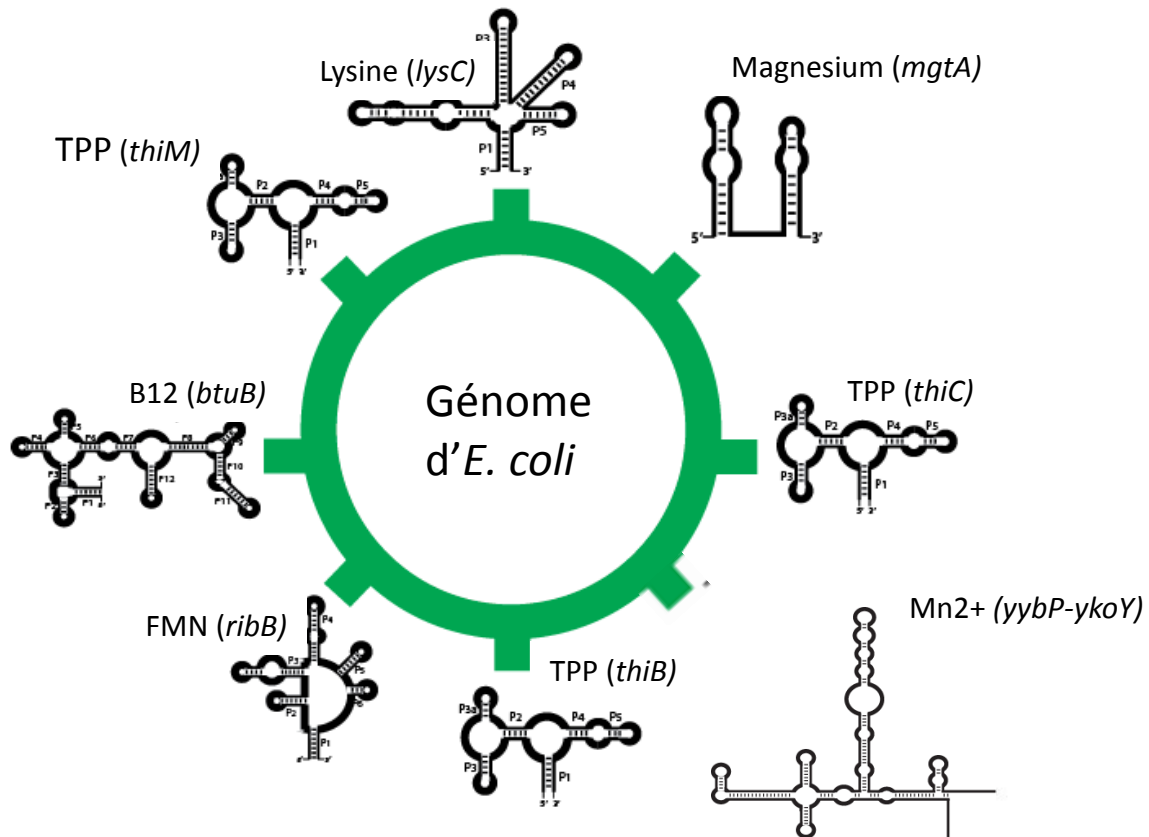


Figure 9. Représentation schématique des riborégulateurs présents chez *E. coli*. Chez *E. coli* il y a actuellement 8 riborégulateurs identifiés et confirmés expérimentalement.

1.6.1 Gène régulé

Le riborégulateur *lysC* est dans le 5' non traduit du gène du même nom. Le gène *lysC* code pour l'aspartokinase III. Cette kinase catalyse l'ajout d'un phosphate en position 4 du L-aspartate, la première étape dans la voie métabolique de la synthèse de la lysine (Figure10) (Margarita, Cohen, Borne, et Patte, 1974). Ce gène mène à l'expression de l'aspartokinase, une protéine de 48,5 kDa. Pour obtenir son effet catalytique, cette enzyme doit se dimériser. L'activité catalytique de cette protéine se fait à l'aide de sa portion N-terminale (résidus 3 à 291) alors que sa portion responsable de la régulation se trouve en C-terminal (300-349) (Kotaka, Ren, Lockyer, Hawkins, et Stammers, 2006). Lorsqu'il y a trop de lysine, la production de l'aspartokinase est stoppée et les dimères s'assemblent en tétrade : la forme inactive de l'enzyme (Richaud *et al.*, 1973) (Niles et Westhead, 1973).

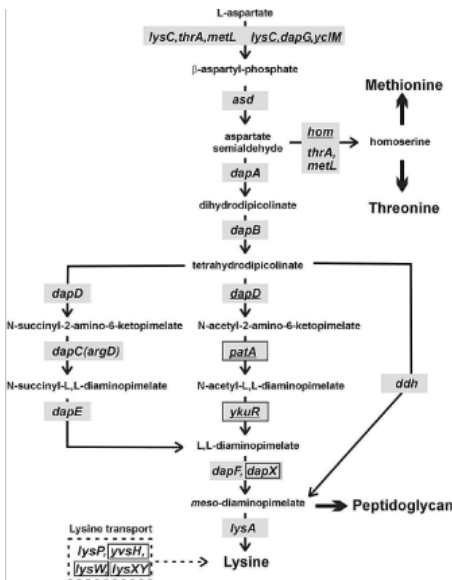


Figure 10. Voie métabolique menant à la synthèse de la lysine. L'aspartokinase III régulée par le riborégulateur *lysC* est la première enzyme dans la voie métabolique et ceci permet de limiter la perte d'énergie en cas de surplus de lysine.

1.6.2 Aptamère, repliement et interactions tertiaires

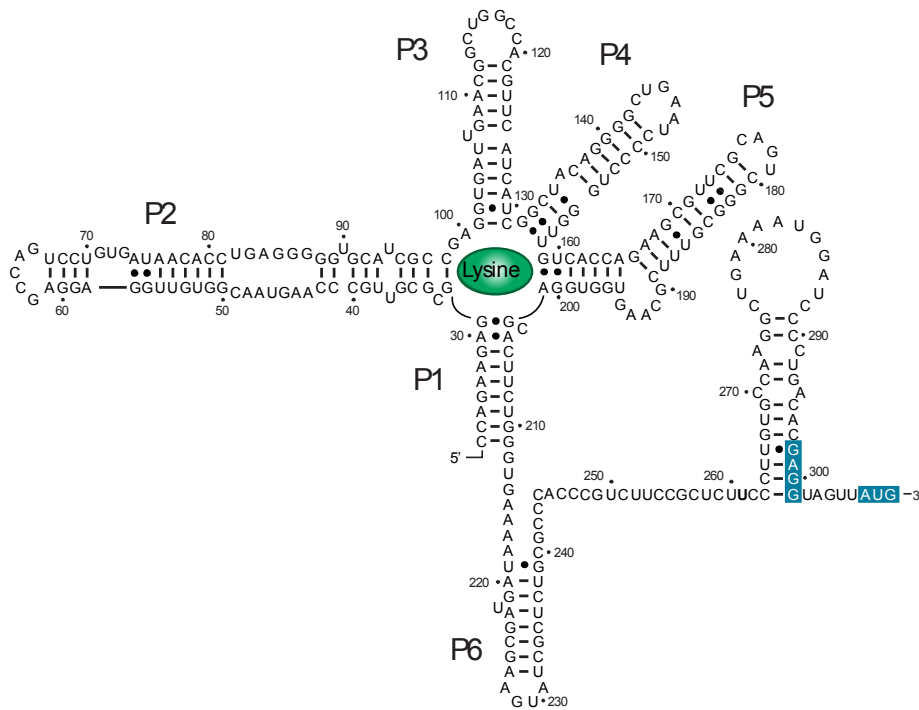


Figure 11. Repliement 2D du riborégulateur *lysC* en présence de ligand chez *E.coli*. Modélisation du repliement du riborégulateur lié à la lysine basée sur les cristaux chez *T. maritima* et les données chez *B. subtilis*. Comme chez *B. subtilis*, l'aptamère est composé de cinq régions paires (P1 à P5). On peut observer la complémentarité de séquence entre L2-L3 et P2-L4 comme décrite chez le riborégulateur *lysC* chez *B. subtilis*. Cependant, contrairement à ce dernier, le riborégulateur *lysC* régule au niveau de la traduction plutôt que de compter sur la formation d'un terminateur intrinsèque. La version du riborégulateur *lysC* présente chez *E. coli* possède donc une tige séquestratrice du RBS ainsi qu'une tige P6 permettant de compétitionner avec cette séquestration lorsqu'il n'y a pas de liaison du ligand (Adapté de Caron 2012).

1.6.3 Liaison de la lysine

Bien qu'il n'y ait pas d'étude exhaustive sur la liaison de la lysine chez le riborégulateur *lysC* chez *E. coli*, Caron et al. ont mis en lumière un mutant permettant d'abolir la liaison de la lysine. La mutation est en position 31 du riborégulateur (G31C) (Caron et al., 2012).

1.6.4 Structure

Il n'y a pas à ce jour d'étude de la structure du riborégulateur lysine chez *E. coli*. Cependant lorsqu'on regarde au niveau de l'interaction de la boucle de la tige P2 et celle de la tige P3, on peut voir que la complémentarité de séquence est présente chez *E. coli* ce qui laisse présager que l'interaction est conservée (Figure 11). Pour ce qui est de l'interaction de la boucle de la tige 2 avec la tige de la P2 on peut aussi voir une complémentarité au niveau des nucléotides comme chez *B. subtilis* (Fig 11). On peut donc penser que cette interaction existe chez *E. coli*.

1.6.5 Régulation

Tout d'abord, le promoteur du gène *lysC* est sensible à la présence de lysine. Le riborégulateur quant à lui utilise différents mécanismes afin d'accomplir la régulation du gène en aval. Il est régulé au niveau de la traduction via la séquestration du RBS lors de la liaison de la lysine à l'aptamère. Le riborégulateur *lysC* régule aussi le niveau d'ARNm du gène via le dégradosome ARN. Le riborégulateur compte deux sites ciblés par le dégradosome, nommés

sites 1 et 2. Ces sites sont séquestrés dans des structures lorsqu'il n'y a pas de ligand alors qu'ils sont exposés lors de la liaison de la lysine à l'aptamère.

1.7 Objectifs de maitrise

À ce jour, le riborégulateur *lysC* chez *Escherichia coli* est l'un des riborégulateurs possédant la plus grande complexité au niveau de sa régulation. Cependant, depuis la publication en 2012 de Kerry Hollands ou plus récemment, Laurene Bastet et Adrien Chauvier, on peut voir que la protéine Rho est souvent impliquée dans la régulation des riborégulateurs présents chez *E. coli* (Hollands *et al.*, 2012; Bastet *et al.*, 2017; Chauvier *et al.*, 2017).

Le premier chapitre abordera la caractérisation *in vitro* du riborégulateur *lysC*. Cette étude permettra de mieux comprendre le comportement du riborégulateur au niveau de sa transcription ainsi que de son changement de conformation à la suite de la liaison du ligand

Le deuxième chapitre a pour but de vérifier si le riborégulateur *lysC*, en plus de la régulation traductionnelle et la dégradation nucléolitique, a recourt à la protéine Rho pour réguler son niveau d'ARNm.

CHAPITRE 2

ÉTUDE DE LA TRANSCRIPTION ET DU CHANGEMENT DE CONFORMATION DU RIBORÉGULATEUR *lysC*.

2.1 INTRODUCTION

Comme mentionné dans l'introduction de ce mémoire, le riborégulateur *lysC* a été identifié en 2003 (Sudarsan *et al.*, 2003). Depuis, ce riborégulateur a été identifié ou prédit chez plusieurs organismes, mais les études du riborégulateur en soi sont assez rares. À ce jour, les seules informations disponibles sur le riborégulateur *lysC* entier chez *E. coli* (aptamère+plateforme) se retrouvent dans l'étude réalisée par Marie-Pier Caron (Caron *et al.*, 2012) Cet article démontre les mécanismes de régulation *in vivo* du riborégulateur. Cependant, au début de cette maîtrise, on en connaissait bien peu sur la structure, les pauses et le changement de conformation de ce riborégulateur. Ce chapitre portera donc sur une étude des caractéristiques du riborégulateur *lysC*.

2.2 MATÉRIELS ET MÉTHODES

2.2.1 Transcription *in vitro*

Les matrices d'ADN pour effectuer la transcription *in vitro* à cycle unique ont été préparées par réaction de polymérisation en chaîne avec les couples d'amorces décrite dans le tableau décrit dans le tableau 4 (Annexe). Les transcriptions *in vitro* ont été réalisées dans un tampon contenant 20 mM Tris-

HCl pH 8.0, 20 mM MgCl₂, 20 mM KCl, 14 mM 2-mercaptoethanol et 0.1 mM EDTA. La matrice d'ADN a été incubée à 37 °C pendant 5 minutes avec la polymérase et le facteur sigma. Par la suite, le trinuécléotide initiateur GUA, les nucléotides (NTP) CTP/GTP ainsi que [α -³²P] UTP ont été ajoutés. La solution a été incubée 10 minutes à 37 °C afin d'obtenir un complexe d'élongation stable (EC) long de 11 nucléotides. La transcription a été complétée par l'addition des quatre nucléotides (concentration dépendante de l'expérience subséquente) ainsi que l'ajout d'héparine pour empêcher la ré-initiation de la transcription par la polymérase (solution FL). La lysine a été ajoutée à une concentration de 100 μ M lorsqu'indiqué.

2.2.2 Identification des sites de pauses

Les sites auxquels la polymérase d'ARN effectue ses pauses durant la transcription (pauses transcriptionnelles) ont été déterminés tels que décrits précédemment (Chauvier *et al.*, 2017). En résumé, les transcriptions à cycle unique ont été réalisées telles que décrites au paragraphe 2.2.1 avec les modifications suivantes. Le EC-11 est transcrit. Ensuite, de l'eau a été ajoutée pour obtenir un volume final de 80 μ L. Puis, le même volume de solution FL a été ajouté (rNTP 50 μ M final). Le tout a été incubé à 37°C. Un aliquot de 10 μ L a été prélevé à chacun de temps prédéterminé. Cet aliquot a été mis dans 30 μ L de solution d'arrêt (95 % formamide, 20 mM EDTA et 0.4 % SDS). Les réactions ont été déposées sur gel d'acrylamide dénaturant (acrylamide : bis-acrylamide 19 : 1, 8 M urée) avec des échelles de séquençage (section 2.2.3).

2.2.3 Transcription des échelles de séquençage à l'aide de nucléotide 3' O-méthyle

Cette technique est encore une fois dérivée d'une transcription à cycle unique classique avec une finalité différente. On tente d'obtenir des arrêts de transcriptions à tous les nucléotides afin de s'en servir comme échelle de séquençage. Le EC-11 a été transcrit tel que décrit précédemment. Des réactions contenant un nucléotide 3'-O-méthyle (Figure 12) et les trois autres nucléotides classiques ont été ajoutés au EC-11 pour poursuivre la transcription. Les concentrations de chacun des nucléotides modifiés nécessaires pour obtenir une échelle claire dépendent de la nature de la séquence nucléotidique de la matrice à transcrire. Pour le riborégulateur *lysC*, il faut respectivement 50, 50, 25 et 17 μM de nucléotides modifiés (A, U, C et G). Les quatre réactions ont été migrées sur gel d'acrylamide dénaturant (acrylamide : bis-acrylamide 19 : 1 18M urée) avec les expériences nécessitant des échelles de séquençage.

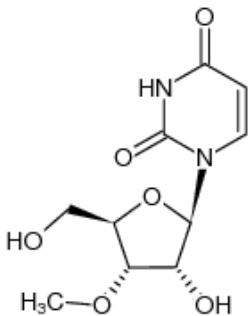


Figure 12. Nucléotide 3' O-méthyle. Représentation d'un nucléotide modifié. On peut voir l'ajout d'un groupement méthyle lié à un oxygène en position 3 du sucre. L'incorporation de ce nucléotide cause l'arrêt de transcription, car le 3' du brin en synthèse devient 3' O-CH₃.

2.2.4 Essai de cartographie à la ribonucléase H

Afin de réaliser un essai de cartographie à la RNase H, une matrice d'ARN est réalisée préalablement telle que décrite dans la section transcription *in vitro*. Une réaction de transcription avec ligand et une réaction en absence de ligand ont été transcrites. Une fois complétés, 8 μL de l'ARN nouvellement transcrit ont été prélevés puis mis en contact avec 2 μL de sonde à 37 °C pendant 5 minutes. Pendant l'incubation de la sonde, la RNase H a été diluée dans son tampon à raison de 0,6 μL d'enzyme pour 49,4 μL de tampon (5 mM Tris-HCl, pH 8.0, 20 mM MgCl₂, 100mM KCl, 50 μM EDTA et 10 mM 2-mercaptoethanol) pour une activité finale de 0.12 U/ μL . Une fois l'incubation avec la sonde terminée, 10 μL de la dilution de l'enzyme a été ajoutée puis incubée pendant cinq minutes. Une fois les cinq minutes écoulées, les réactions ont été arrêtées par l'ajout de 20 μL de solution d'arrêt et mise 2 minutes à 95 °C. Les réactions ont été migré sur gel d'acrylamide dénaturant (acrylamide : bis-acrylamide 19 : 18M urée).

2.2.5 Détermination du K_{switch}

Le K_{switch} est le nom donné à la concentration de ligand nécessaire pour faire basculer un riborégulateur d'un état vers un autre (actif vers inactif ou vice et versa).

Cette valeur a été déterminée avec la technique de la cartographie à la RNase H avec quelques modifications. D'abord, 20 réactions ont été transcrites avec une concentration de ligand croissante. Les valeurs qui ont été utilisées pour le

riborégulateur *lysC* varient entre 0 et 50 mM. Une fois les transcriptions terminées, les réactions ont été incubées avec la sonde d'ADN 1205 AC ou 1251 AC (Tableau 3) puis avec la RNase H comme mentionné précédemment. Les réactions ont été déposées sur gel d'acrylamide pour visualiser la coupure. Le pourcentage de coupure a été quantifié grâce au logiciel Quantity One. Le pourcentage de clivage par rapport à la concentration de ligand a été modélisé puis le logiciel a calculé la concentration à laquelle 50% de la population a changé de conformation ce qui correspond au K_{switch} .

2.3 RÉSULTATS

2.3.1 Transcription *in vitro*

Le riborégulateur a été transcrit du +1 de transcription jusqu'à 100 nucléotides de la séquence codante. Cette longueur de matrice est un choix arbitraire. Des tests préliminaires *in vitro* semblaient démontrer que la séquence de 20 codons utilisés par Marie-Pier Caron ne nous permettait pas de voir tous les événements de régulation du riborégulateur alors que cette longueur était suffisante *in vivo*. Par la suite, les promoteurs naturels du riborégulateur qui sont sensibles à la présence de ligand ont été remplacés par le promoteur *PlacUV5* qui lui, ne devrait pas être sensible à la lysine. Comme le démontre la Figure 13, l'ajout de ligand n'affecte pas le niveau d'ARN produit par transcription *in vitro* à cycle unique. Ce résultat permet de confirmer la variation qu'on observe dans les expériences subséquentes sont un effet du ligand et non un effet du promoteur.

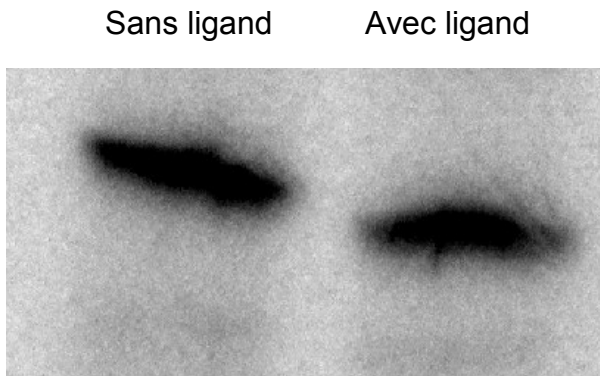


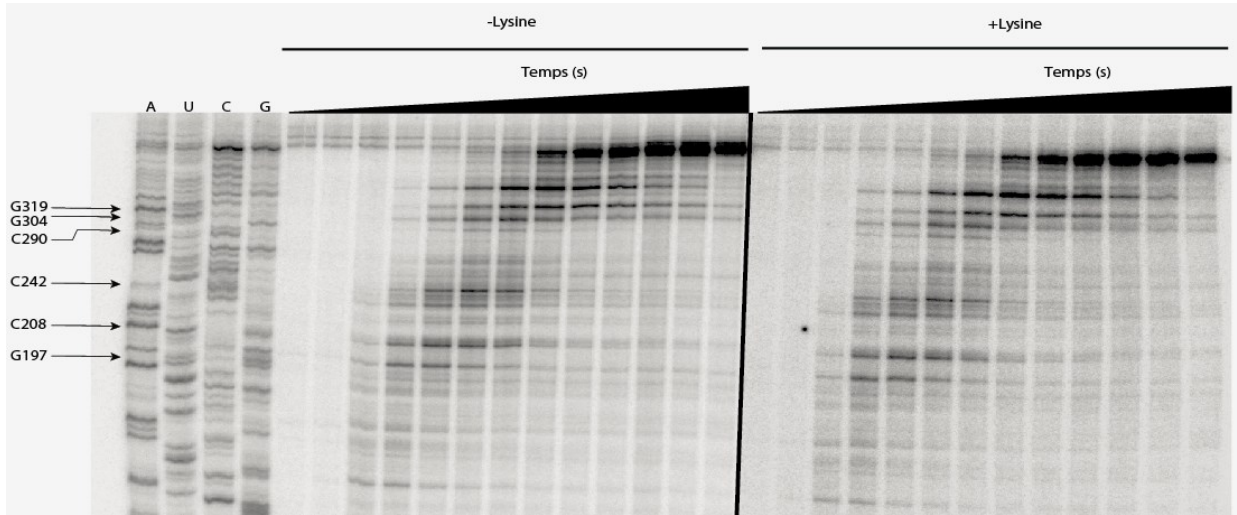
Figure 13. Essai de transcription de la matrice d'ADN *lysC*. Transcription de la matrice *lysC P/acUV5 + 100nt* de la séquence codante en cycle unique.

2.3.2 Identification des sites de pauses

Depuis quelques années, des études sur les sites de pauses tendent à démontrer leurs importances dans divers processus biologiques comme le repliement de l'ARN et la régulation génétique. Il était donc intéressant de vérifier s'il y avait de sites de pauses transcriptionnels chez le riborégulateur *lysC* et ensuite savoir leur emplacement dans la séquence. La matrice *P/acUV5+lysC+100nt* de la séquence codante a donc été transcrite. La transcription est effectuée avec une concentration finale de nucléotide de 50 μ M. Ceci permet de ralentir la transcription afin de visualiser les différents sites de pauses. Il y a plusieurs sites de pauses comme le démontre la Figure 14A. Cependant, les efforts ont été mis à identifier la position des pauses les plus fortes. Les deux premières pauses identifiées, soit en position 197 et 208 sont intéressantes d'un point de vue repliement de l'aptamère. En effet, elles sont positionnées dans la tige P5 et au dernier nucléotide de la partie 3' de la tige P1.

Ce qu'il faut savoir c'est que lorsque la polymérase est arrêtée à une pause, il y a de 8 à 14 nucléotides qui sont protégés. Lorsqu'on regarde des positions de pause, on doit garder en tête ce qui est dans la polymérase et par conséquent ne peut pas faire de paires de bases. Ce qui suggère que la pause en position 208 pourrait aider au repliement de l'aptamère, mais aussi laisser plus de temps au ligand de s'y lier. Afin de vérifier cette hypothèse, il serait intéressant de faire une étude plus exhaustive comme dans l'article de Chauvier et al. 2017. Les pauses près de l'AUG pourraient être impliquées dans la modulation de la régulation du niveau d'ARN messager. Il est aussi intéressant de noter qu'on peut voir une pause forte à l'AUG qui respecte le consensus décrit par Larson (Larson *et al.*, 2014). Les sites de pauses ont par la suite été quantifiés selon la méthode mentionnée dans la section 2.2.2. On peut remarquer qu'il ne semble pas y avoir d'effet du ligand sur la durée de la pause, sauf pour la pause à la position 309. Elle semble être légèrement moins longue en présence de ligand. Il faudrait faire plus d'expérience pour confirmer cette hypothèse, car la différence est faible et pourrait être due à une erreur du manipulateur ou de quantification. Il faudrait entre autres faire varier la concentration de nucléotides lors de la transcription afin d'accélérer la transcription et voir si la variation avec et sans ligand s'agrandit telle qu'observée pour la pause PsC du riborégulateur *thiC* (Chauvier *et al.*, 2017).

A)



Demie-vie des trois principales pauses du riborégulateur *lysC*

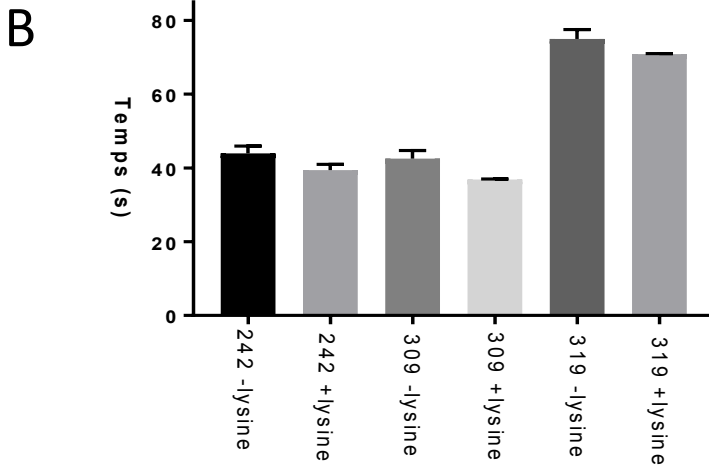


Figure 14. Pauses transcriptionnelles du riborégulateur *lysC*. A) Expérience de *time pausing* déposé sur gel d'acrylamide 5 % avec des échelles 3'-O-méthyle afin de déterminer les sites de pauses. B) Quantification des demi-vies des principales pauses identifiées en A). Les demi-vies ont été déterminées tel que décrit par le groupe de Landick (Landick, Wang et Chan, 1996).

2.3.3 K_{switch}

Depuis quelques années, une technique permettant de quantifier la concentration de ligand causant un changement de conformation de l'ARN est disponible pour l'étude des riborégulateurs. Cette mesure est appelée le K_{switch} . Il ne faut cependant pas mélanger le K_d et le K_{switch} . Le K_d est une mesure qui reflète l'affinité qu'à l'ARN pour un ligand. La valeur de K_d représente donc la concentration où 50 % des ARN ont lié le ligand. Les K_d des riborégulateurs ont pour la plupart été déterminés par cartographie chimique au magnésium (*in line probing*) et avec seulement l'aptamère et non le riborégulateur complet. Le K_{switch} quant à lui représente la concentration de ligand à laquelle 50 % des ARN (riborégulateurs complets) passent d'une conformation à une autre (ON à OFF ou OFF à ON). On peut donc déjà postuler que la valeur de K_{switch} sera égale ou supérieure à la valeur du K_d puisqu'il faut que le ligand se lie avant de provoquer un changement de conformation. Dans le cas du riborégulateur *lysC*, le K_d est de 1 μM (Sudarsan *et al.*, 2003). On s'attendait donc à avoir une valeur de K_{switch} plus élevée. La construction *P/acUV5+ lysC 100 nt ORF* a été transcrite afin de trouver des sondes pour la RNase H afin de pouvoir passer à l'étape du K_{switch} .

On peut voir à la Figure 15A les endroits ciblés par les différentes sondes. Le choix de sondes a été fait selon deux critères. Le premier critère est de cibler une zone qui devrait changer de structure lors du passage de l'état actif à inactif. C'est pour cette raison que la région qui compose la tige P1 du riborégulateur a été ciblée. Pour la deuxième sonde, la région ciblée est celle qui compose la tige séquestratrice du RBS, car on sait grâce aux travaux de Marie-Pier Caron, que le riborégulateur *lysC* est traductionnel et donc que cette région change de conformation. Une fois les régions intéressantes déterminées, plusieurs sondes de 10 nucléotides en se décalant de 1 nucléotide en 3' à chaque fois ont été testées. Dans le cas de *lysC*, on observe une forte coupure de l'ARN en absence de ligand et une faible coupure lorsque le riborégulateur lie le ligand.

Pour cette construction, on obtient une valeur de K_{switch} de 23.65 μM . Comme prévu, cette valeur est plus haute que le K_d . On peut aussi voir qu'en plus de posséder un K_d élevé et une affinité stricte envers la L-lysine, le riborégulateur *lysC* nécessite une forte concentration de lysine afin de changer de conformation. En effet, la majorité des riborégulateurs chez *E. coli* ont un K_{switch} dans l'ordre du nanomolaire sauf *lysC* et *tbpA* qui sont dans l'ordre du bas micro molaire. Cette donnée est importante pour toutes les expériences subséquentes, car la concentration de ligand utilisé doit être plus élevée que le K_{switch} afin de s'assurer que le changement de conformation s'effectue.

CHAPITRE 3

ÉTUDE DE LA RÉGULATION DU RIBORÉGULATEUR *lysC* PAR LA PROTÉINE RHO.

3.1 INTRODUCTION

Lors de mon arrivée au laboratoire, les modes de régulation du riborégulateur *lysC* avaient déjà publiés (Caron *et al.*, 2012). On pouvait y voir que le riborégulateur lysine était régulé au niveau de la traduction, car lors de la liaison du ligand il y avait formation d'une tige qui séquestrait le RBS afin que le ribosome ne puisse pas initier la traduction. On pouvait aussi y apprendre que le niveau d'ARN était régulé par l'action du dégradosome ARN. De plus, des résultats de Northern Blot semblaient démontrer que la protéine Rho ne régulait pas le niveau d'ARN. Cependant, on peut voir dans certains articles que le riborégulateur lysine semble être régulé par cette protéine (Sedlyarova *et al.*, 2016b ; Bastet *et al.*, 2017). Ces résultats et le Northern Blot ne sont pas en total désaccord. Pour bien comprendre il faut penser que dans l'expérience de Northern Blot on regarde sur la disparition de l'ARN entier donc s'il y a une RNase quelconque qui agit sur l'ARN on ne verra pas l'effet de l'inhibition de la protéine Rho par la bicyclomycine en Northern Blot. Dans ce chapitre il sera question de confirmer que la régulation du niveau d'ARNm est affectée par le facteur de terminaison de la transcription Rho. Il sera aussi question de caractériser cette régulation notamment déterminer les sites de terminaison et le site de liaison de la protéine sur l'ARN.

3.2 MATÉRIELS ET MÉTHODES

3.2.1 Essai B-Gal

L'essai Beta-galactosidase a été réalisé tel que décrit par Bastet et *al.* (Bastet et *al.*, 2017). En résumé, une souche de *E. coli* MG1655 contenant le plasmide pFR Δ avec le riborégulateur *lysC* et 6 codons clonés à l'intérieur a été mise en pré-culture O/N à 37 °C dans un milieu minimum M63 glycérol. Le matin suivant la culture est diluée dans du milieu frais puis elle a été incubée à 37 °C jusqu'à l'obtention d'une densité optique de 0,1. Une fois la densité optique de 0,1 atteinte, l'arabinose a été ajoutée. Par la suite, la bicyclomicine a été ajoutée là où indiquée.

3.2.2 Essai de terminaison de la transcription *in vitro*

Cette technique a été réalisée telle que décrite par Bastet et *al.* Brièvement, la construction P/*lacUV5+lysC+100nt* de la séquence codante a été utilisée pour cet essai. L'ARN a été transcrit dans le tampon de terminaison (40 mM Tris-HCl, pH 8.0, 50 mM KCl, 5 mM MgCl₂ et 1,5 mM DTT) avec le tri-nucléotide GUA les nucléotides ATP et CTP et le [α -³²P] UTP. Une fois le EC-11 formé, la réaction est incubée 5 minutes avec 50 nm de la protéine Rho. Par la suite, le mélange FL a été ajouté comme décrit dans la section transcription *in vitro* avec les modifications suivantes. L'héparine a été remplacée par de la rifampicine à une concentration finale de 1,5 μ g/mL et on ajoute la protéine NusG à la même concentration que la protéine Rho dans les réactions indiquées. Une étape d'extraction des protéines par phénol/chloroforme a été réalisée afin d'enlever Rho et NusG. Cette étape est réalisée afin d'éviter que la réaction reste prise

dans les puits du gel. Par la suite, la solution d'arrêt (95 % formamide, 20 mM EDTA et 0,4 % SDS) a été ajoutée dans un ratio 1 : 1.

3.2.3 Prédiction bio-informatique

En partant du principe que les sites *rut* sont des régions riches en cytidine il est possible de prédire des régions qui sont susceptibles d'être reconnu par Rho. En effet, un graphique qui représentant le pourcentage de cytidine par rapport au pourcentage de guanine dans la séquence d'intérêt a été réalisé. Pour ce faire, il s'agit de prendre des fenêtres de 10 nucléotides et on décale par la suite la fenêtre de 1 nucléotide. Chaque nucléotide est donc couvert par 10 fenêtres différentes. Par la suite, on peut calculer le pourcentage de cytidine par rapport au pourcentage de guanine pour chaque nucléotide.

3.2.4 Identification du *site rut* de façon expérimentale

3.2.4.1 Transcription d'un ARN à l'aide de la polymérase du phage T7

La transcription des ARN a été réalisée à partir d'une matrice d'ADN double brin (ADNdb), en utilisant la polymérase d'ARN du phage T7 (Milligan *et al.*, 1987). La matrice d'ADNdb a été obtenue à partir d'une PCR, auquel un un promoteur T7 en amont de la séquence à transcrire a été ajouté. Afin d'augmenter l'efficacité de la transcription, la séquence GCG est ajoutée au début de la matrice à transcrire (Pleiss, Derrick et Uhlenbeck, 1998). La matrice d'ADN a été incubée à 37 °C pendant 3 heures en présence de tampon de transcription FRANK 1X (40 mM Tris-HCl (pH 8,0), 0,01 % Triton X-100 (V/V), 20 mM MgCl₂ et 2 mM

spermidine), 5 mM rNTP, 10mM DTT et 1 µl de la polymérase d'ARN du phage T7.

3.2.4.2 Précipitation et purification des ARN

Une fois la transcription réalisée, les ARN ont été précipités dans une solution composée de 75 % d'éthanol, 67 mM d'acétate d'ammonium et 15 µg/ml de glycogène à -80 °C pendant 30 min. Une fois cette étape réalisée, les tubes ont été centrifugés à 4 °C pendant 20 minutes, à 13 000 rpm. Les surnageants ont été enlevés, les culots ont été lavés avec de l'éthanol 75 % froid (-20°C) et centrifugés à 4°C pendant 5 minutes, à 13 000 rpm. L'excédent de liquide a été enlevé à l'aide d'une micropipette afin qu'il n'y ait plus d'éthanol visible avant de laisser sécher. Les culots ont ensuite été brièvement séchés et resuspendus dans de l'eau. Un volume de formamide 100 % a été ajouté et les ARN ont été purifiés sur gel d'acrylamide (acrylamide:bis-acrylamide 19:1, TBE 1X et 8 M urée). Une fois la migration effectuée, les bandes contenant les ARN sont détectées par ombrage aux UV puis coupées. Les morceaux de bandes sont mis dans des microtubes avec 500 µl d'eau puis ont été élués pendant O/N à 4°C et récupérés à l'aide d'une précipitation telle que décrite précédemment.

3.2.4.3 Déphosphorylation et marquage radioactif en 5' des ARN

Le marquage radioactif en 5' des ARN nécessite d'abord une étape de déphosphorylation. Cette étape a été réalisée en utilisant la phosphatase alcaline de la crevette (rSAP) en incubant 50 pmoles d'ARN pendant 1 heure à 37°C en présence de tampon SAP 1X (50 mM Tris-HCl pH 8,5). L'enzyme a ensuite été désactivée par dénaturation thermique à 65 °C pendant 15 minutes. Le marquage radioactif en 5' a été effectué avec la polynucléotide kinase du

bactériophage T4 (T4 PNK) en incubant 5 pmoles d'ARN déphosphorylé pendant 1 heure à 37 °C en présence de tampon PNK 1X (70 mM Tris-HCl (pH 7,6), 10 mM MgCl₂, 5 mM DTT) et 10 mCi [γ -³²P] -ATP. L'ARN radiomarqué a ensuite été purifié par une électrophorèse sur gel de polyacrylamide (acrylamide:bis-acrylamide 19:1, TBE 1X et 8 M urée), élués pendant 16 heures dans 500 μ L l'eau à 4°C et récupérés à l'aide d'une précipitation telle que décrite précédemment. Chaque marquage a ensuite été resuspendu dans un volume de 10 μ L. Le volume peut être ajusté en fonction de l'efficacité du marquage.

3.2.4.4 Essai de retard sur gel après pontage aux UV

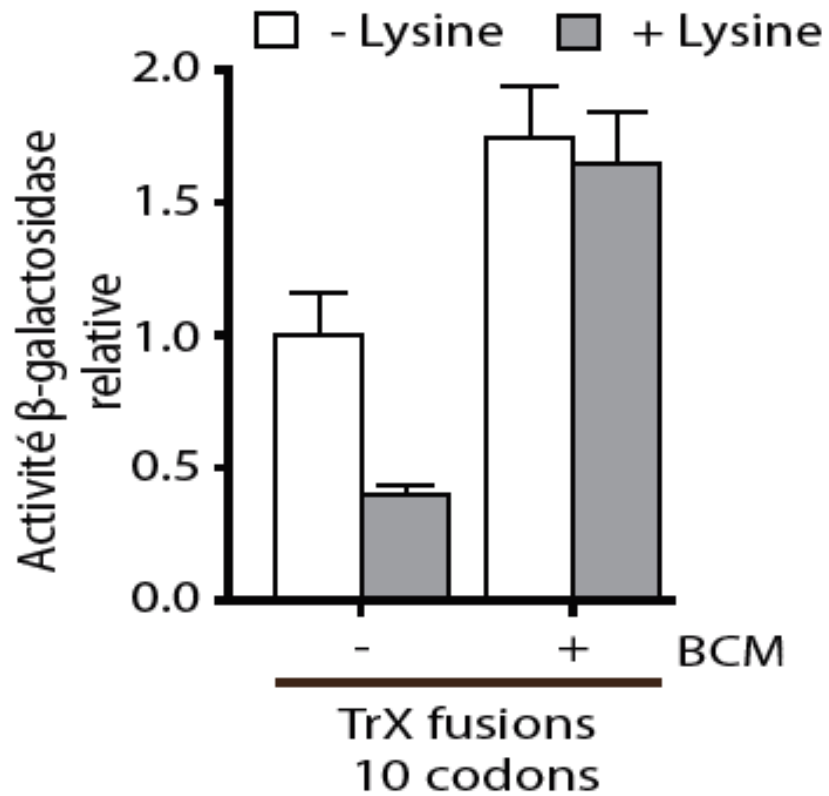
Une fois l'ARN T7 transcrit élué et marqué, la radioactivité présente dans le marquage a été quantifiée à l'aide d'un compteur à scintillation. Une fois le nombre de comptes par minute (cpm) déterminé, une dilution a été réalisée afin d'atteindre dans chaque échantillon une activité de 3000 cpm/ μ L. Pour effectuer le pontage, 1 μ L de la dilution, 1 μ L de tampon TKM 5x (20 mM tris 50mM KCl 20 mM MgCl₂) 20 picomoles de Rho ont été mis dans un microtube. Le volume a été ajusté à 5 μ L puis la réaction a été incubée à 37°C pendant 5 minutes afin de permettre la liaison de la protéine à l'ARN. Par la suite, le mélange a été exposé aux ultraviolets à une longueur d'onde de 254 nm à une énergie totale de 500 mJ/cm². Une fois ce pontage réalisé, la réaction a été sur gel de polyacrylamide (acrylamide:bis-acrylamide 19:1, TBE 1X et 8 M urée) afin d'observer un retard sur gel.

3.3 RÉSULTATS

3.3.1 Confirmation de la régulation de Rho par essai Beta-galactosidase

Dans la littérature, il était question que Rho eût probablement un impact sur l'ARN du gène *lysC* depuis la parution des publications de Bastet et al et du groupe de Nudler (Sedlyarova *et al.*, 2016 a ; Bastet *et al.*, 2017). Cependant, dans ce dernier on ne faisait que suggérer que la protéine Rho était impliquée. Afin de confirmer l'implication de la protéine Rho, l'essai de gène rapporteur de la Beta-galactosidase. Tout d'abord, une fusion transcriptionnelle a été construite (Figure 15B) afin de suivre le niveau de l'ARN dans différentes conditions. Une fois cette construction obtenue, des cultures a été réalisées en triplicata pour les différentes conditions (avec/sans ligand, avec/sans bicyclomycine). Les résultats sont présentés en Figure 16A. Tout d'abord, on peut voir dans les deux premières colonnes qu'il y a une régulation du niveau d'ARN lorsqu'on ajoute de la lysine dans le milieu de culture. Ce résultat était plutôt attendu, car on avait observé la même chose dans précédemment (Caron *et al.*, 2012). Cependant, lorsqu'on regarde la deuxième partie du graphique on peut voir que l'ajout de bicyclomycine a pour effet d'augmenter l'activité beta-galactosidase ce qui rapporte une augmentation du niveau d'ARN dans la cellule. Avec ce résultat et en prenant compte les articles précédents, on peut affirmer que le riborégulateur *lysC* est bel et bien régulé par la protéine Rho. Un autre détail ressort des résultats. En effet, on peut remarquer une augmentation du niveau d'ARN même dans la condition sans ligand. Ceci implique que la protéine Rho peut probablement réguler de façon plus légère le niveau d'ARN même si le riborégulateur est dans sa conformation active.

A)



B)



Figure 16. Validation in vivo de l'effet de la protéine Rho. A) Suivi du niveau d'ARN via le système de gène rapporteur β -galactosidase. On peut voir l'effet du ligand sur le niveau d'ARN ainsi que l'abolition de la régulation avec l'ajout de bicyclomycine. B) Schéma simplifié de la construction introduite dans la bactérie.

3.3.2 Localisation de la terminaison de la transcription par la protéine Rho

À la suite des résultats obtenus avec la technique d'essai beta-galactosidase, on pouvait affirmer que le gène *lysC* était régulé par la protéine Rho. Cependant, les sites de terminaison de la transcription étaient toujours inconnus. Afin de déterminer ces sites, une terminaison de la transcription par la protéine rho a été effectuée avec l'ARN de *lysC* (*PlacUV5+lysC+ 100nt*). On peut remarquer plusieurs choses d'intéressantes à la Figure 17. D'abord, on voit *in vitro* ce qu'on voyait *in vivo* c'est-à-dire, que la protéine Rho régule. On peut cependant voir que la protéine seule n'a pas un grand impact sur la terminaison de la transcription sur cet ARN. Par contre, si on regarde les pistes avec la protéine Rho en combinaison avec la protéine NusG, on peut voir une grande amélioration dans la terminaison. On peut aussi voir que c'est l'interaction entre les protéines NusG et Rho qui provoque cette augmentation de la terminaison et non une accumulation d'une terminaison des deux protéines. En effet, on peut voir que la protéine NusG n'est pas capable de causer une terminaison de la transcription lorsque Rho est absente. Il est intéressant de noter dans ce résultat que la terminaison s'effectue en tout temps (avec et sans ligand). Cependant, la terminaison est environ deux fois plus efficace lorsqu'il y a du ligand lié au riborégulateur.

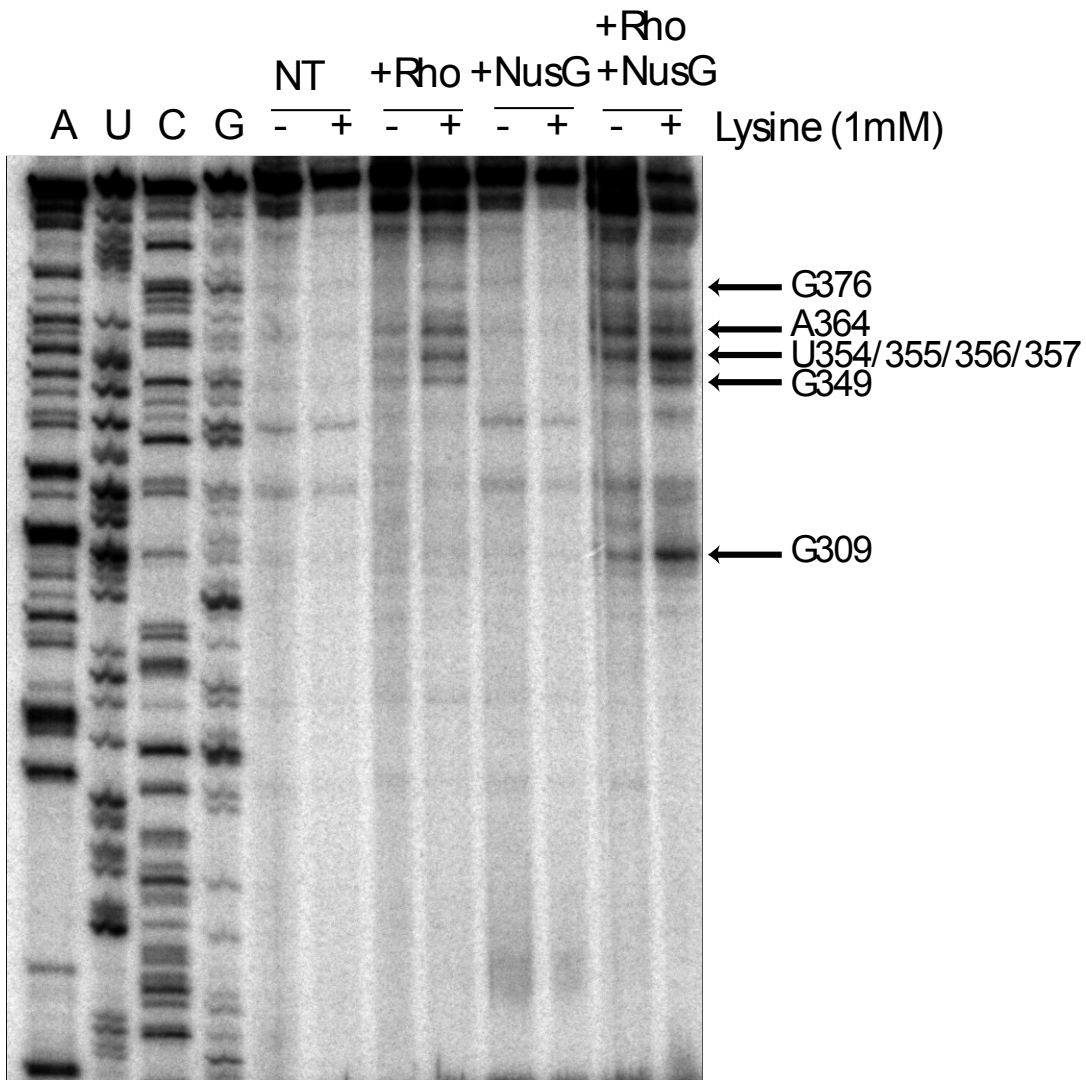


Figure 17. Localisation des sites de terminaison par la protéine Rho.

Les différentes conditions de réaction ont été migrées sur gel avec la présence d'échelle de séquençage. Dans les premières pistes, on voit que la transcription est propre avec ou sans ligand. Dans les pistes suivantes on peut noter que la terminaison est à son maximum avec Rho, NusG et avec ligand.

Au niveau de la position des terminaisons dans l'ARN, il y a plusieurs détails intéressants. Malgré le fait qu'il est déjà connu dans la littérature que les pauses donnent plus de temps à Rho de rejoindre la polymérase, il était intéressant de voir que les sites de terminaison les plus intenses se situent à la position exacte ou à quelques nucléotides des pauses les plus fortes du riborégulateur. Il est aussi intéressant de noter que le premier site de terminaison se situe à la position 309 soit, le G du codon d'initiation de la traduction. Ceci nous donne déjà un indice sur la position du *site rut*. On sait maintenant que le site de liaison de la protéine Rho est à l'intérieur de la séquence du riborégulateur puisque la protéine doit nécessairement lier l'ARN avant de pouvoir terminer la transcription. De plus, lorsqu'on construit le graphique représentant le pourcentage de C par rapport au pourcentage de G comme expliqué à la section 3.2.3 on peut voir une bulle cytidine riche guanine pauvre (Fig 18A). La région grise dans le graphique de la Figure 17A représente à la fois l'endroit où se trouvent les sites utilisés par le dégradosome et la région dans laquelle l'ARN n'est pas structuré. Donc, sans avoir encore de résultats probants il est possible d'imaginer que la région en grise puisse être au cœur du *site rut*. En effet, elle correspond aux critères d'un *site rut* et elle est située à 70 nucléotides du premier site de terminaison (Figure18B)

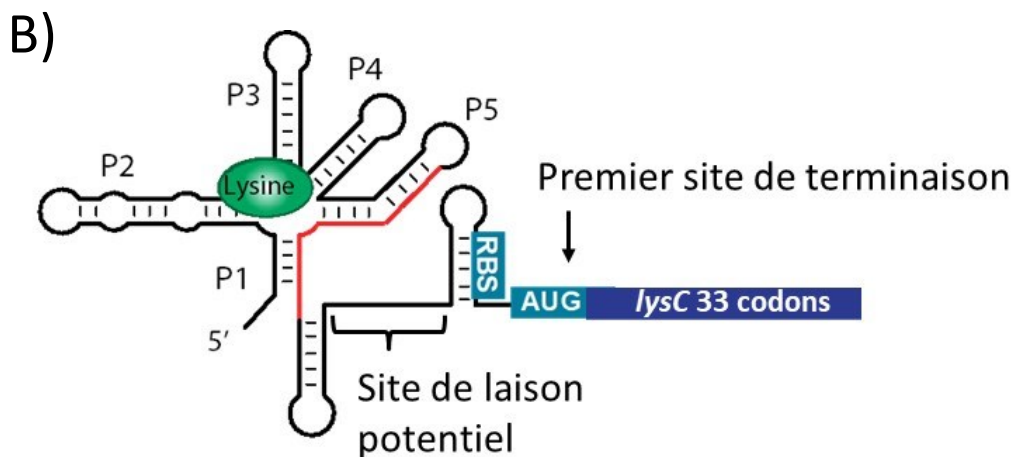
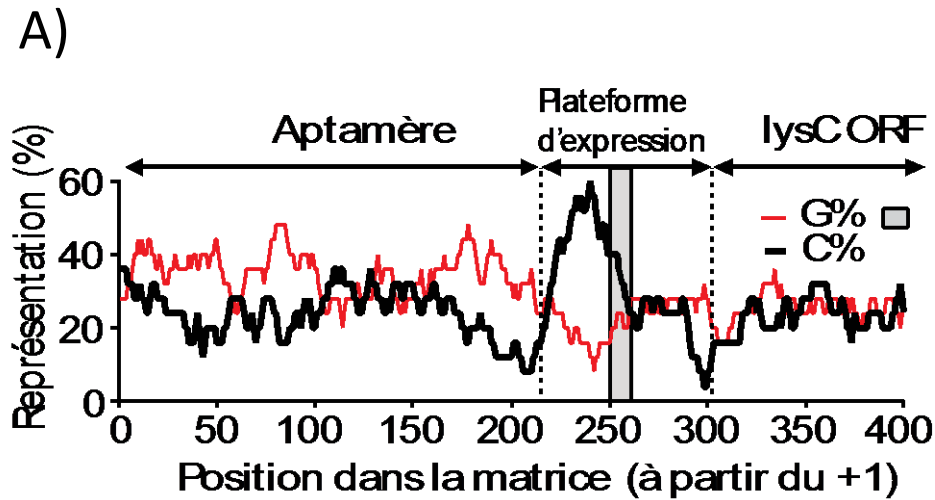


Figure 18. Prédiction du *site rut*. Représentation du pourcentage de cytidine par rapport au pourcentage de guanine afin de trouver une région pauvre en G et riche en C. La région en grise représente l'endroit ciblé par le dégradosome ARN et on peut voir que cela se situe dans une région riche en cytidine et pauvre en guanine. B) Représentation graphique du site de liaison potentiel de la protéine. Elle tient compte de la prédiction bio-informatique, la position de la première terminaison par la protéine rho ainsi que les informations dans la littérature à propos de la distance entre un *site rut* et le premier site de terminaison.

3.3.3 Identification du *site rut* par pontage *in vitro*.

Les pontages *in vitro* ont été réalisés tels que décrits à la section 3.2.4 (Figure 19A). Afin de vérifier l'énergie nécessaire permettant d'obtenir un bon pontage, de multiples énergies ont été utilisées. À la Figure 19B on peut voir plusieurs éléments. D'abord, on peut voir que l'ARN non traité ainsi que l'ARN traité aux UV sans l'ajout de Rho migrent sensiblement à la même hauteur sur le gel. Ceci implique donc que le pontage aux UV ne lie pas plusieurs ARN les uns avec les autres. Par contre, il est difficile de dire seulement avec ce résultat si l'ARN se pontage sur lui-même. Ensuite on peut voir différents pontages effectués avec une énergie croissante (500 mJ/cm^2 à 999 mJ/cm^2). On peut voir que même à la plus faible énergie utilisée il y a environ 100 % de pontage. Cette énergie a donc été utilisée afin de réaliser les expériences subséquentes. Afin de déterminer la séquence minimale du riborégulateur nécessaire à avoir un pontage de l'ARN et de Rho, des transcrits tronqués ont été réalisés. L'ARN du riborégulateur *lysC* a été artificiellement tronqué grâce à différentes constructions par PCR (voir Tableau 4). Ainsi, dans cette expérience l'ARN dit non tronqué correspond à *lysC* transcrit jusqu'au nucléotide 308 (1 nucléotide avant le premier site de terminaison). L'ARN a été par la suite raccourci pour obtenir un tronqué 290, 280 260 et un tronqué 250. Comme on peut voir à la Figure 18C, on observe un pontage efficace peu importe la longueur du transcrit. Ce résultat semble contredire notre prédiction bio-informatique et nous ramène à la case départ. Cependant, le résultat à la Figure 18C a été obtenu dans la dernière semaine de cette maîtrise et n'a pas pu être tenté à nouveau. Il serait intéressant de réaliser un triplicata avant de conclure trop vite. Aussi, il faut noter que 500 mJ/cm^2 représente une énorme quantité d'énergie donc il faudrait peut-être essayer de diminuer l'énergie envoyée à l'ARN.

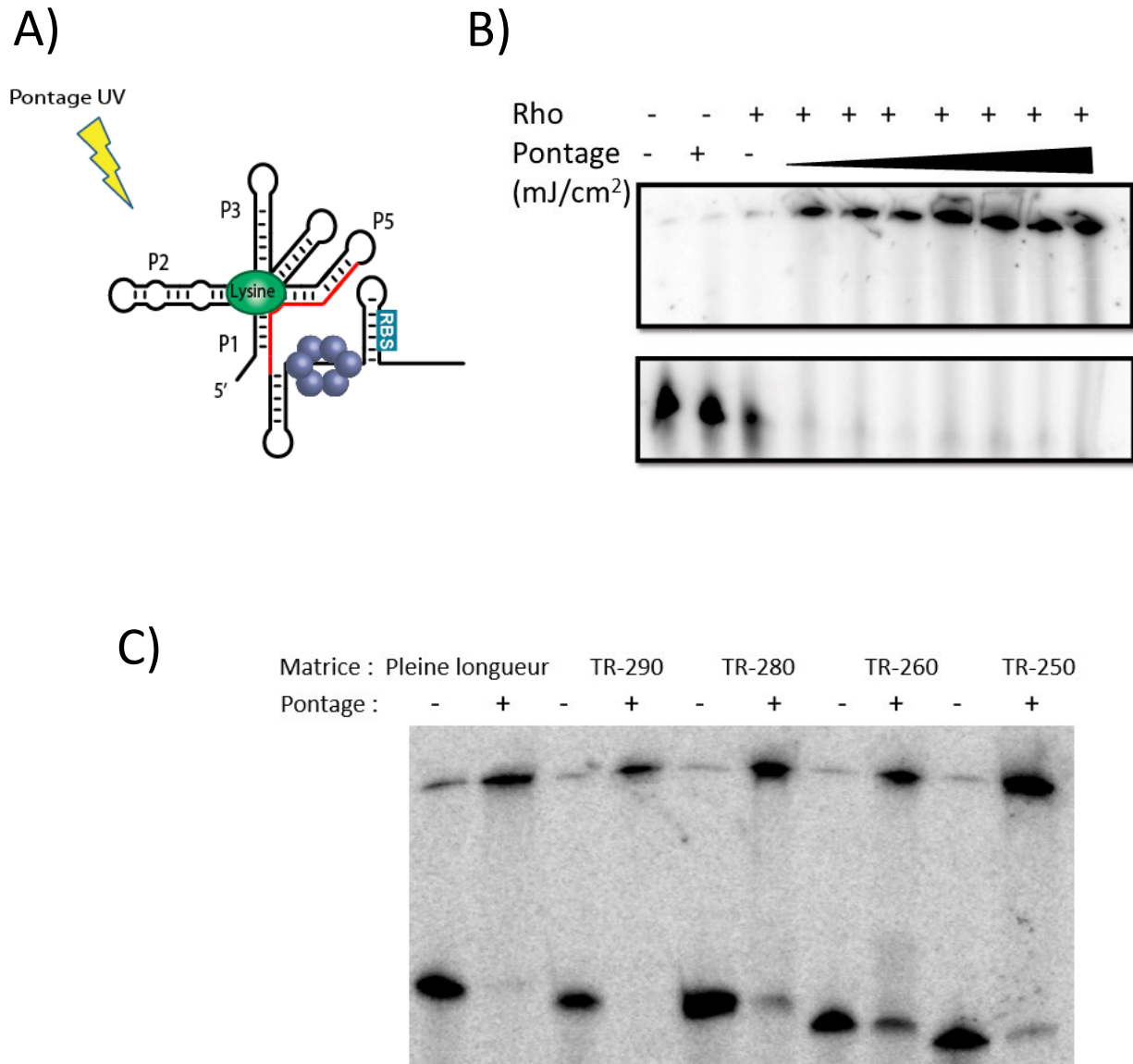


Figure 19. Localisation du *site rut* à l'aide du pontage aux UV. A) Schéma simplifié de l'expérience. Le riborégulateur est transcrit avec la polymérase T7 puis incubé avec la protéine Rho puis soumis aux UV. B) Essai préliminaire de pontage avec l'ARN pleine longueur (+1 de transcription à l'AUG) avec différentes concentrations de Rho. C) Essai de pontage avec des transcrits tronqués. Les chiffres situés après les lettres TR représentent la position à laquelle le transcrit est arrêté.

4. CONCLUSION

Dans ce mémoire, le riborégulateur *lysC* a été étudié. D'abord, cette étude a porté sur une brève caractérisation *in vitro* du riborégulateur (pauses, K_{switch}). Puis nous avons démontré que le riborégulateur *lysC* chez *E. coli* possède un troisième mécanisme permettant de réguler l'expression du gène en aval. En effet nous avons mis en lumière la régulation du niveau d'ARN par la protéine Rho puis nous avons tenté de caractériser cette régulation. Les sites de terminaisons pour la construction possédant 100 nt de la séquence codante ont été identifiés. Malgré tous les efforts et les multiples techniques utilisées, il n'est pas encore à ce jour possible de localiser avec certitude l'endroit où se lie la protéine Rho. Vers la fin de cette maîtrise, nous avons été en mesure d'identifier que notre réserve de protéine Rho avait une activité compromise. Nous avons donc recommencé la purification de la protéine ce qui a donné les résultats de pontage aux UV présenté en Figure 17. Cependant, il est impossible de dire si cette diminution d'activité peut expliquer qu'aucune technique classique de détermination d'un site de liaison d'une protéine liant l'ARN n'ait permis d'identifier le *site rut*. Une autre avenue possible pour expliquer l'impossibilité le trouver serait que le *site rut* de *lysC* serait un *site rut* faible. Dans la littérature, on retrouve certains articles qui discutent de la force d'un *site rut* (Nadiras *et al.*, 2018). Les sites *rut* qui sont dit faibles nécessitent la présence de la protéine NusG afin de pouvoir lier l'ARN. Il faudrait donc recommencer tous les essais de cartographie chimique ou enzymatique avec la présence de Rho et NusG dans la réaction. Cependant, il est important de noter qu'une enzyme a généralement besoin de certaines conditions (tampon spécifique, ions spécifiques). Il faut donc garder en tête que l'ajout de protéines comme Rho et NusG qui sont dans leur propre tampon peuvent causer des complications comme une concentration trop élevée en glycérol qui peut inhiber l'action de l'enzyme. Une autre voie à explorer

serait de modifier les conditions de pontage aux UV. En effet, le pontage a été réalisé à la base avec certaines données de la littérature, mais il n'y avait pas beaucoup d'information pertinente pour le pontage de la protéine Rho. L'énergie envoyée était basée sur une étude *in vivo* dans laquelle les auteurs utilisaient 800 mJ/cm². Considérant que dans l'expérience *in vivo* l'énergie doit passer au travers du milieu de culture et de la paroi bactérienne une énergie moindre (500 mJ/cm²) a été utilisée pour les expériences *in vitro*. Cependant, rien n'indique que dans cette condition nous n'allons pas obtenir des faux positifs. Il serait donc intéressant de vérifier qu'elle est l'énergie minimale requise pour observer un pontage qui s'approche du 100 %. Il serait important de vérifier la concentration minimale requise de la protéine Rho pour observer un pontage qui lui aussi s'approche de 100 % afin de limiter la possibilité que trop de protéine Rho cause un pontage non spécifique. Il sera aussi important de trouver un ARN auquel la protéine ne pourra pas ponter dans les conditions déterminées ultérieurement.

Utiliser trois modes de régulation en plus d'avoir un promoteur sensible au ligand nous amène à nous poser certaines questions. Pourquoi avoir autant de mécanismes différents de régulation ? Est-ce favorable énergétiquement parlant pour la cellule ? Pourquoi ces différents modes de régulation ont-ils été conservés par l'évolution ? Existe-t-il des conditions où un certain mode de régulation est favorisé alors que dans d'autres conditions est défavorisé ? Malheureusement, il est impossible de répondre à ces questions avec les connaissances actuelles. En fait, pour répondre certaine question comme celle sur l'évolution, il faudrait effectuer un travail de phylogénie puis tester chacun des ancêtres d'*E. coli* afin de vérifier quels mécanismes ils possèdent. Cependant, on peut commencer la réflexion en se demandant quand agissent les différents mécanismes. La reconnaissance du promoteur est nécessaire pour commencer la transcription donc on pourrait qualifier de régulation du promoteur par le ligand de pré-transcriptionnel. Ceci laisse présager que la régulation plus "grossière"

est effectuée par le promoteur. Par la suite viennent les mécanismes de régulation co-transcriptionnels c'est-à-dire ceux qui agissent alors que la transcription est toujours en cours (Rho et séquestration du RBS). Ces mécanismes sont probablement en place pour s'assurer de ne pas dépenser d'énergie à transcrire tout le gène si le niveau de lysine est assez élevé. Pour ce qui est du dégradosome, le moment où il agit est un peu nébuleux. Cependant, il semble jouer un rôle de recycleur des ressources c'est-à-dire qu'il dégraderait les transcrits abortifs ou terminés prématurément pour recycler les nucléotides. Il serait donc possiblement post-transcriptionnel. Bien qu'il n'y ait pas encore de preuve tangible de ce modèle, il y a certains événements étranges observés dans les résultats de Caron et *al.*, ainsi que dans ce mémoire qui pourrait nous donner des informations précieuses.

Au courant de cette maîtrise, nous avons aussi remarqué que certains résultats ne pouvaient être expliqués que par l'inhibition d'un seul mode de régulation. En effet, si on regarde la Figure 4 dans l'article de Caron et *al.*, on peut voir que lorsqu'on effectue le retrait de la séquence site 1 on n'observe aucune régulation par le ligand. Dans l'article on attribue ce résultat au fait qu'on enlève la séquence ciblée par le dégradosome ARN. Par contre, dans ce mémoire les résultats obtenus montrent clairement que la protéine Rho est impliquée dans la régulation du niveau d'ARN. On devrait alors voir un effet de Rho même si on inhibe le dégradosome. Un autre résultat qui démontre ce phénomène est la Figure 15 dans ce mémoire. Lorsqu'on inhibe Rho avec la bicyclomycine, on perd à nouveau toute la régulation alors que nous devrions voir un effet du dégradosome ARN. On pourrait penser que la bicyclomycine inhibe aussi le dégradosome, mais des essais de Northern blot ont déjà été réalisés dans le laboratoire et on ne voit aucun effet sur l'activité de dégradation du complexe. Il serait donc intéressant dans le futur de poursuivre les travaux qui ont été commencés vers la fin de cette maîtrise. Il serait intéressant de tenter de

découpler les modes de régulation en performant des essais Beta-galactosidase avec différentes souches mutantes (dégradosome et rho) en combinaison avec différentes constructions de *lysC* (longue et courte). Ceci permettra de mieux comprendre la dynamique de régulation complexe du riborégulateur *lysC*.

“Toute science crée une nouvelle ignorance”

- Henri Michaux

5. ANNEXE

Tableau 3. Liste des oligonucléotides.

Oligonucleotides	Séquences 5'-3'
1168 AC	GGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAAT GTGTGGGTACTACCTGCGCTAGCGCA
1170 AC	GGCGTCAAATCAGCTACGCTG
1175 AC	CCTCTTCTGG
1205 AC	CAGCCTTGGC
1251 AC	TTTCAGCCTT
2170 PT	TGTAATACGACTCACTATAGGCCAGAAGAGGCGCGTTGC
2315 PT	ATAACTACCTCGTGTCAGGGGATCCA
2783 PT	AGAGCGGAAGACGGGTGGGCGCAGA
2784 PT	ACGGGTGGGCGCAGAGCGATACTTC
2785 PT	ACTTCGCTACTATTTTCACCCAGAA
3059 PT	GGGATCCATTTTCAGCCTTGGCAC
3236 PT	GGCACAAGGGAAGAGCGGAAGACGG
3237 PT	TTCAGCCTTGGCACAAGGGAAGAGC

Tableau 4. Constructions utilisées pour la transcription obtenues par PCR

Oligonucléotides	Constructions
<hr/>	
Transcription In vitro	
<hr/>	
<i>PlacUV5+ lysC 100 nt ORF</i>	1168 AC-1170 AC (ADN génomique)
<i>lysC T7-Tronqué 230</i>	2170 PT-2785 PT (ADN Génomique)
<i>lysC T7-Tronqué 250</i>	2170 PT-2784 PT (ADN génomique)
<i>lysC T7 Tronqué 260</i>	2170 PT-2783 PT (ADN génomique)
<i>lysC T7 Tronqué 270</i>	2170 PT-3236 PT (ADN génomique)
<i>lysC T7-Tronqué 280</i>	2170 PT-3237 PT (ADN génomique)
<i>lysC T7-Tronqué 290</i>	2170 PT-3059 PT (ADN génomique)
<i>lysC T7 FL 308 nt</i>	2170 PT-2315 PT (ADN génomique)

6. BIBLIOGRAPHIE

Altman, S. (2011) 'Ribonuclease P.', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. The Royal Society, 366(1580), pp. 2936–41. doi: 10.1098/rstb.2011.0142.

Ames, T. D. *et al.* (2010) 'A eubacterial riboswitch class that senses the coenzyme tetrahydrofolate.', *Chemistry & biology*, 17(7), pp. 681–5. doi: 10.1016/j.chembiol.2010.05.020.

Ames, T. D. and Breaker, R. R. (2011) 'Bacterial aptamers that selectively bind glutamine.', *RNA biology*, 8(1), pp. 82–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21282981> (Accessed: 3 September 2018).

Baker, J. L. *et al.* (2012) 'Widespread Genetic Switches and Toxicity Resistance Proteins for Fluoride', *Science*, 335(6065), pp. 233–235. doi: 10.1126/science.1215063.

Banerjee, S. *et al.* (2006) 'Rho-dependent transcription termination: more questions than answers.', *Journal of microbiology (Seoul, Korea)*, 44(1), pp. 11–22. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1838574&tool=pmcentrez&rendertype=abstract> (Accessed: 14 December 2015).

Barrow, C. J. *et al.* (1992) *A Role for Interaction of the RNA Polymerase Flap Domain with the Subunit in Promoter Recognition*, 21. J. E. Hueser, *J. Muscle Res. Cell Motil.* Available at: www.sciencemag.org (Accessed: 3 September 2018).

Bastet, L. *et al.* (2017) *Translational control and Rho-dependent transcription termination are intimately linked in riboswitch regulation.*, *Nucleic acids research*. doi: 10.1093/nar/gkx434.

Bear, D. G. *et al.* (1985) 'Escherichia coli transcription termination factor rho has a two-domain structure in its activated form.', *Proceedings of the National Academy of Sciences of the United States of America*, 82(7), pp. 1911–5. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2580303> (Accessed: 10 September 2018).

Bernstein, J. A. *et al.* (2002) 'Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays.', *Proceedings of the National Academy of Sciences of the United States of America*, 99(15), pp. 9697–702. doi: 10.1073/pnas.112318199.

Blouin, S., Chinnappan, R. and Lafontaine, D. A. (2011) 'Folding of the lysine riboswitch: importance of peripheral elements for transcriptional regulation.', *Nucleic acids research*, 39(8), pp. 3373–87. doi: 10.1093/nar/gkq1247.

Blouin, S. and Lafontaine, D. A. (2007) 'A loop loop interaction and a K-turn motif located in the lysine aptamer domain are important for the riboswitch gene regulation control', *RNA*, 13(8), pp. 1256–1267. doi: 10.1261/rna.560307.

Blount, K. F. *et al.* (2007) 'Antibacterial lysine analogs that target lysine riboswitches.', *Nature chemical biology*, 3(1), pp. 44–9. doi: 10.1038/nchembio842.

Bogden, C. E. *et al.* (1999) 'The structural basis for terminator recognition by the Rho transcription termination factor.', *Molecular cell*, 3(4), pp. 487–93. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10230401> (Accessed: 10 September 2018).

Boyer, R. F. (2006) 'Chapter 11: DNA Replication and Transcription: Biosynthesis of DNA and RNA', in *Concepts in Biochemistry*, pp. 315–351.

Breaker, R. R. (2012) 'Riboswitches and the RNA World', *Cold Spring Harbor Perspectives in Biology*, 4(2), pp. a003566–a003566. doi: 10.1101/cshperspect.a003566.

Burgess, B. R. and Richardson, J. P. (2001) 'Transcription Factor Rho Does Not Require a Free End to Act as an RNA-DNA Helicase on an RNA', *Journal of Biological Chemistry*, 276(20), pp. 17106–17110. doi: 10.1074/jbc.M101856200.

Burgess, R. R. *et al.* (1969) 'Factor stimulating transcription by RNA polymerase.', *Nature*, 221(5175), pp. 43–6. doi: 10.1038/221043a0.

Buzayan, J. M., Gerlach, W. L. and Bruening, G. (1986) 'Satellite tobacco ringspot virus RNA: A subset of the RNA sequence is sufficient for autolytic processing.', *Proceedings of the National Academy of Sciences of the United States of America*, 83(23), pp. 8859–62. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=387032&tool=pmcentrez&rendertype=abstract> (Accessed: 3 June 2016).

Callaghan, A. J. *et al.* (2004) 'Studies of the RNA Degradosome-organizing Domain of the Escherichia coli Ribonuclease RNase E', *Journal of Molecular Biology*, 340(5), pp. 965–979. doi: 10.1016/j.jmb.2004.05.046.

Callaghan, A. J. *et al.* (2005) 'Structure of Escherichia coli RNase E catalytic domain and implications for RNA turnover', *Nature*, 437(7062), pp. 1187–1191. doi: 10.1038/nature04084.

Caron, M.-P. *et al.* (2012) 'Dual-acting riboswitch control of translation initiation and mRNA decay.', *Proceedings of the National Academy of Sciences of the United States of America*, 109(50), pp. E3444-53. doi: 10.1073/pnas.1214024109.

Caron, M. P. *et al.* (2012) 'Dual-acting riboswitch control of translation initiation and mRNA decay', *Proc Natl Acad Sci U S A*. 2012/11/22, 109(50), pp. E3444-53. doi: 10.1073/pnas.12140241091214024109 [pii].

- Carpousis, A. J. *et al.* (1994) 'Copurification of E. coli RNAase E and PNPase: evidence for a specific association between two enzymes important in RNA processing and degradation.', *Cell*, 76(5), pp. 889–900. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7510217> (Accessed: 14 January 2019).
- Carpousis, A. J. (2007) 'The RNA degradosome of Escherichia coli: an mRNA-degrading machine assembled on RNase E.', *Annual review of microbiology*. Annual Reviews, 61, pp. 71–87. doi: 10.1146/annurev.micro.61.080706.093440.
- Causton, H. *et al.* (1994) 'mRNA degradation in Escherichia coli: a novel factor which impedes the exoribonucleolytic activity of PNPase at stem-loop structures.', *Molecular microbiology*, 14(4), pp. 731–41. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7534370> (Accessed: 14 January 2019).
- Chandran, V. and Luisi, B. F. (2006) 'Recognition of Enolase in the Escherichia coli RNA Degradosome', *Journal of Molecular Biology*, 358(1), pp. 8–15. doi: 10.1016/j.jmb.2006.02.012.
- Chauvier, A. *et al.* (2017) 'Transcriptional pausing at the translation start site operates as a critical checkpoint for riboswitch regulation', *Nature Communications*, 8, p. 13892. doi: 10.1038/ncomms13892.
- Coburn, G. A. *et al.* (1999) 'Reconstitution of a minimal RNA degradosome demonstrates functional coordination between a 3' exonuclease and a DEAD-box RNA helicase.', *Genes & development*, 13(19), pp. 2594–603. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10521403> (Accessed: 11 January 2019).
- Corbino, K. A. *et al.* (2005) 'Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in alpha-proteobacteria.', *Genome biology*, 6(8), p. R70. doi: 10.1186/gb-2005-6-8-r70.
- Cromie, M. J. *et al.* (2006) 'An RNA Sensor for Intracellular Mg²⁺', *Cell*, 125(1), pp. 71–84. doi: 10.1016/j.cell.2006.01.043.

- Dambach, M. *et al.* (2015) 'The ubiquitous yybP-ykoY riboswitch is a manganese-responsive regulatory element.', *Molecular cell*. Elsevier Inc., 57(6), pp. 1099–109. doi: 10.1016/j.molcel.2015.01.035.
- Dar, D. and Sorek, R. (2018) 'High-resolution RNA 3'-ends mapping of bacterial Rho-dependent transcripts', *Nucleic Acids Research*, 46(13), pp. 6797–6805. doi: 10.1093/nar/gky274.
- Dolan, J. W., Marshall, N. F. and Richardson, J. P. (1990) 'Transcription termination factor rho has three distinct structural domains.', *The Journal of biological chemistry*, 265(10), pp. 5747–54. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2318834> (Accessed: 10 September 2018).
- Ebright, R. H. (2000) 'RNA Polymerase: Structural Similarities Between Bacterial RNA Polymerase and Eukaryotic RNA Polymerase II', *Journal of Molecular Biology*, 304(5), pp. 687–698. doi: 10.1006/jmbi.2000.4309.
- El-Hattab, A. W. *et al.* (2015) 'MELAS syndrome: Clinical manifestations, pathogenesis, and treatment options', *Molecular Genetics and Metabolism*, 116(1–2), pp. 4–12. doi: 10.1016/j.ymgme.2015.06.004.
- Epshtein, V. *et al.* (2007) 'An Allosteric Path to Transcription Termination', *Molecular Cell*, 28(6), pp. 991–1001. doi: 10.1016/j.molcel.2007.10.011.
- Fuchs, R. T., Grundy, F. J. and Henkin, T. M. (2006) 'The S(MK) box is a new SAM-binding RNA for translational regulation of SAM synthetase.', *Nature structural & molecular biology*, 13(3), pp. 226–33. doi: 10.1038/nsmb1059.
- Furukawa, K. *et al.* (2015) 'Bacterial Riboswitches Cooperatively Bind Ni²⁺ or Co²⁺ Ions and Control Expression of Heavy Metal Transporters', *Molecular Cell*. Elsevier Inc., 57(6), pp. 1088–1098. doi: 10.1016/j.molcel.2015.02.009.
- Garst, A. D. *et al.* (2008) 'Crystal structure of the lysine riboswitch regulatory mRNA element.', *The Journal of biological chemistry*, 283(33), pp. 22347–51.

doi: 10.1074/jbc.C800120200.

Geiselman, J. *et al.* (1992) *Physical Properties of the Escherichia coli Transcription Termination Factor Rho. 1. Association States and Geometry of the Rho Flamer-*’, *Biochemistry*. UTC. Available at: <https://pubs.acs.org/sharingguidelines> (Accessed: 7 August 2018).

Geiselman, J. *et al.* (1992) ‘Physical properties of the Escherichia coli transcription termination factor rho. 1. Association states and geometry of the rho hexamer.’, *Biochemistry*, 31(1), pp. 111–21. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1370623> (Accessed: 10 September 2018).

Gerber, A. P. and Keller, W. (1999) ‘An adenosine deaminase that generates inosine at the wobble position of tRNAs.’, *Science (New York, N.Y.)*, 286(5442), pp. 1146–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10550050> (Accessed: 29 August 2018).

Gilbert, S. D. *et al.* (2008) ‘Structure of the SAM-II riboswitch bound to S-adenosylmethionine’, *Nature Structural & Molecular Biology*, 15(2), pp. 177–182. doi: 10.1038/nsmb.1371.

Gray, M. C. M. W. and Gray, M. W. (2000) ‘Pseudouridine in RNA: What, Where, How, and Why’, *IUBMB Life (International Union of Biochemistry and Molecular Biology: Life)*, 49(5), pp. 341–351. doi: 10.1080/152165400410182.

Griffiths, A. J. F. *et al.* (2008) *Introduction to Genetic Analysis*. W H Freeman & Co.

Grundy, F. J., Winkler, W. C. and Henkin, T. M. (2002) ‘tRNA-mediated transcription antitermination in vitro: Codon-anticodon pairing independent of the ribosome’, *Proceedings of the National Academy of Sciences*, 99(17), pp. 11121–11126. doi: 10.1073/pnas.162366799.

Guerrier-Takada, C. *et al.* (1983) ‘The RNA moiety of ribonuclease P is the

catalytic subunit of the enzyme', *Cell*, 35(3), pp. 849–857. doi: 10.1016/0092-8674(83)90117-4.

Hangauer, M. J., Vaughn, I. W. and McManus, M. T. (2013) 'Pervasive Transcription of the Human Genome Produces Thousands of Previously Unidentified Long Intergenic Noncoding RNAs', *PLoS Genetics*. doi: 10.1371/journal.pgen.1003569.

Haugen, S. P., Ross, W. and Gourse, R. L. (2008) 'Advances in bacterial promoter recognition and its control by factors that do not bind DNA.', *Nature reviews. Microbiology*. NIH Public Access, 6(7), pp. 507–19. doi: 10.1038/nrmicro1912.

Hollands, K. *et al.* (2012) 'Riboswitch control of Rho-dependent transcription termination.', *Proceedings of the National Academy of Sciences of the United States of America*, 109(14), pp. 5376–81. doi: 10.1073/pnas.1112211109.

JACOB, F. and MONOD, J. (1961) 'Genetic regulatory mechanisms in the synthesis of proteins.', *Journal of molecular biology*, 3, pp. 318–56. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/13718526> (Accessed: 22 May 2015).

Jenkins, J. L. *et al.* (2011) 'Comparison of a PreQ₁ Riboswitch Aptamer in Metabolite-bound and Free States with Implications for Gene Regulation', *Journal of Biological Chemistry*, 286(28), pp. 24626–24637. doi: 10.1074/jbc.M111.230375.

Jia, X. *et al.* (2013) 'Riboswitch Regulation of Aminoglycoside Resistance Acetyl and Adenyl Transferases', *Cell*, 153(7), pp. 1419–1420. doi: 10.1016/j.cell.2013.05.050.

Jiang, X. and Belasco, J. G. (2004) 'Catalytic activation of multimeric RNase E and RNase G by 5'-monophosphorylated RNA.', *Proceedings of the National Academy of Sciences of the United States of America*, 101(25), pp. 9211–6. doi: 10.1073/pnas.0401382101.

- Khemici, V. *et al.* (2005) 'Evidence in vivo that the DEAD-box RNA helicase RhlB facilitates the degradation of ribosome-free mRNA by RNase E.', *Proceedings of the National Academy of Sciences of the United States of America*, 102(19), pp. 6913–8. doi: 10.1073/pnas.0501129102.
- Kim, D. E., Shigesada, K. and Patel, S. S. (1999) 'Transcription termination factor Rho contains three noncatalytic nucleotide binding sites.', *The Journal of biological chemistry*, 274(17), pp. 11623–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10206972> (Accessed: 14 January 2019).
- Klein, D. J. *et al.* (2001) 'The kink-turn: a new RNA secondary structure motif', *The EMBO Journal*, 20(15), pp. 4214–4221. doi: 10.1093/emboj/20.15.4214.
- Koslover, D. J. *et al.* (2008) 'The crystal structure of the Escherichia coli RNase E apoprotein and a mechanism for RNA degradation.', *Structure (London, England : 1993)*, 16(8), pp. 1238–44. doi: 10.1016/j.str.2008.04.017.
- Kotaka, M. *et al.* (2006) 'Structures of R- and T-state *Escherichia coli* Aspartokinase III', *Journal of Biological Chemistry*, 281(42), pp. 31544–31552. doi: 10.1074/jbc.M605886200.
- Kruger, K. *et al.* (1982) 'Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena', *Cell*, 31(1), pp. 147–157. doi: 10.1016/0092-8674(82)90414-7.
- Landick, R., Wang, D. and Chan, C. L. (1996) 'Quantitative analysis of transcriptional pausing by Escherichia coli RNA polymerase: his leader pause site as paradigm.', *Methods in enzymology*, 274, pp. 334–53. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8902817> (Accessed: 18 December 2015).
- Larson, M. H. *et al.* (2014) 'A pause sequence enriched at translation start sites drives transcription dynamics in vivo.', *Science (New York, N.Y.)*, 344(6187), pp. 1042–7. doi: 10.1126/science.1251871.

Lee, K., Bernstein, J. A. and Cohen, S. N. (2002) 'RNase G complementation of rne null mutation identifies functional interrelationships with RNase E in Escherichia coli.', *Molecular microbiology*, 43(6), pp. 1445–56. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11952897> (Accessed: 11 January 2019).

Liou, G.-G. *et al.* (2002) 'DEAD box RhlB RNA helicase physically associates with exoribonuclease PNPase to degrade double-stranded RNA independent of the degradosome-assembling region of RNase E.', *The Journal of biological chemistry*, 277(43), pp. 41157–62. doi: 10.1074/jbc.M206618200.

Liu, X. and Martin, C. T. (2009) 'Transcription elongation complex stability: the topological lock.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 284(52), pp. 36262–70. doi: 10.1074/jbc.M109.056820.

Lu, C. *et al.* (2008) 'Crystal structures of the SAM-III/S(MK) riboswitch reveal the SAM-dependent translation inhibition mechanism.', *Nature structural & molecular biology*, 15(10), pp. 1076–83. doi: 10.1038/nsmb.1494.

Lussier, A. *et al.* (2015) 'A Kissing Loop Is Important for btuB Riboswitch Ligand Sensing and Regulatory Control.', *The Journal of biological chemistry*, 290(44), pp. 26739–51. doi: 10.1074/jbc.M115.684134.

Mackie, G. A. (1998) 'Ribonuclease E is a 5'-end-dependent endonuclease', *Nature*, 395(6703), pp. 720–724. doi: 10.1038/27246.

Mandal, M. *et al.* (2004) 'A Glycine-Dependent Riboswitch That Uses Cooperative Binding to Control Gene Expression', *Science*, 306(5694), pp. 275–279. doi: 10.1126/science.1100829.

Marcaida, M. J. *et al.* (2006) 'The RNA degradosome: life in the fast lane of adaptive molecular evolution.', *Trends in biochemical sciences*, 31(7), pp. 359–65. doi: 10.1016/j.tibs.2006.05.005.

Mccown, P. J. *et al.* (2017) 'Riboswitch diversity and distribution', *RNA*. doi: 10.1261/rna.061234.117.

Miczak, A. *et al.* (1996) 'Proteins associated with RNase E in a multicomponent ribonucleolytic complex.', *Proceedings of the National Academy of Sciences of the United States of America*, 93(9), pp. 3865–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8632981> (Accessed: 14 January 2019).

Milligan, J. F. *et al.* (1987) 'Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates.', *Nucleic acids research*, 15(21), pp. 8783–98. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/3684574> (Accessed: 14 January 2019).

Mironov, A. S. *et al.* (2002) 'Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria.', *Cell*, 111(5), pp. 747–56. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12464185> (Accessed: 28 April 2016).

Miwa, Y., Horiguchi, T. and Shigesada, K. (1995) 'Structural and functional dissections of transcription termination factor rho by random mutagenesis.', *Journal of molecular biology*, 254(5), pp. 815–37. doi: 10.1006/jmbi.1995.0658.

Modrak, D. and Richardson, J. P. (1994) 'The RNA-binding domain of transcription termination factor rho: isolation, characterization, and determination of sequence limits.', *Biochemistry*, 33(27), pp. 8292–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7518246> (Accessed: 10 September 2018).

Montange, R. K. and Batey, R. T. (2006) 'Structure of the S-adenosylmethionine riboswitch regulatory mRNA element', *Nature*, 441(7097), pp. 1172–1175. doi: 10.1038/nature04819.

Morita, M. T. *et al.* (1999) 'Translational induction of heat shock transcription factor sigma32: evidence for a built-in RNA thermosensor.', *Genes & development*, 13(6), pp. 655–65. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/10090722> (Accessed: 4 January 2019).

Morita, T., Maki, K. and Aiba, H. (2005) 'RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs.', *Genes & development*, 19(18), pp. 2176–86. doi: 10.1101/gad.1330405.

Nadiras, C. *et al.* (2018) 'A multivariate prediction model for Rho-dependent termination of transcription.', *Nucleic acids research*. Oxford University Press, 46(16), pp. 8245–8260. doi: 10.1093/nar/gky563.

Nahvi, A. *et al.* (2002) 'Genetic control by a metabolite binding mRNA.', *Chemistry & biology*, 9(9), p. 1043. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12323379> (Accessed: 16 December 2015).

Nechooshtan, G. *et al.* (2009) 'A pH-responsive riboregulator.', *Genes & development*, 23(22), pp. 2650–62. doi: 10.1101/gad.552209.

Niles, E. G. and Westhead, E. W. (1973) 'The variable subunit structure of lysine-sensitive aspartylkinase from *Escherichia coli* TIR-8.', *Biochemistry*, 12(9), pp. 1715–22. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/4572992> (Accessed: 21 August 2018).

Ono, M. and Kuwano, M. (1979) 'A conditional lethal mutation in an *Escherichia coli* strain with a longer chemical lifetime of messenger RNA', *Journal of Molecular Biology*. Academic Press, 129(3), pp. 343–357. doi: 10.1016/0022-2836(79)90500-X.

Opperman, T. and Richardson, J. P. (1994) 'Phylogenetic analysis of sequences from diverse bacteria with homology to the *Escherichia coli* rho gene.', *Journal of bacteriology*, 176(16), pp. 5033–43. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8051015> (Accessed: 10 September 2018).

Pleiss, J. A., Derrick, M. L. and Uhlenbeck, O. C. (1998) 'T7 RNA polymerase

produces 5' end heterogeneity during in vitro transcription from certain templates.', *RNA (New York, N.Y.)*, 4(10), pp. 1313–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9769105> (Accessed: 14 January 2019).

Poiata, E. *et al.* (2009) 'A variant riboswitch aptamer class for S-adenosylmethionine common in marine bacteria.', *RNA (New York, N.Y.)*, 15(11), pp. 2046–56. doi: 10.1261/rna.1824209.

Py, B. *et al.* (1996) 'A DEAD-box RNA helicase in the Escherichia coli RNA degradosome', *Nature*, 381(6578), pp. 169–172. doi: 10.1038/381169a0.

Rabhi, M. *et al.* (2011) 'Mutagenesis-Based Evidence for an Asymmetric Configuration of the Ring-Shaped Transcription Termination Factor Rho', *Journal of Molecular Biology*, 405(2), pp. 497–518. doi: 10.1016/j.jmb.2010.11.006.

Regulski, E. E. *et al.* (2008) 'A widespread riboswitch candidate that controls bacterial genes involved in molybdenum cofactor and tungsten cofactor metabolism', *Molecular Microbiology*, 68(4), pp. 918–932. doi: 10.1111/j.1365-2958.2008.06208.x.

Reynolds, R., Bermúdez-Cruz, R. M. and Chamberlin, M. J. (1992) 'Parameters affecting transcription termination by Escherichia coli RNA polymerase. I. Analysis of 13 rho-independent terminators.', *Journal of molecular biology*, 224(1), pp. 31–51. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1372365> (Accessed: 10 September 2018).

Richaud, C. *et al.* (1973) 'Subunit structure of aspartokinase 3 of Escherichia coli K12.', *European journal of biochemistry*, 40(2), pp. 619–29. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/4360904> (Accessed: 21 August 2018).

Roberts, J. W. (1969) 'Termination Factor for RNA Synthesis', *Nature*. Available at: <http://www.nature.com/nature/journal/v224/n5225/pdf/2241168a0.pdf> (Accessed: 31 March 2017).

Rodionov, D. A. *et al.* (2003) 'Regulation of lysine biosynthesis and transport genes in bacteria: yet another RNA riboswitch?', *Nucleic acids research*, 31(23), pp. 6748–57. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14627808> (Accessed: 11 January 2019).

Santangelo, T. J. and Artsimovitch, I. (2011) 'Termination and antitermination: RNA polymerase runs a stop sign.', *Nature reviews. Microbiology*. NIH Public Access, 9(5), pp. 319–29. doi: 10.1038/nrmicro2560.

Santangelo, T. J. and Roberts, J. W. (2004) 'Forward translocation is the natural pathway of RNA release at an intrinsic terminator.', *Molecular cell*, 14(1), pp. 117–26. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15068808> (Accessed: 10 September 2018).

Sedlyarova, N. *et al.* (2016a) 'sRNA-Mediated Control of Transcription Termination in *E. coli*', *Cell*, 167(1), p. 111–121.e13. doi: 10.1016/j.cell.2016.09.004.

Sedlyarova, N. *et al.* (2016b) 'sRNA-Mediated Control of Transcription Termination in *E. coli*', *Cell*, 167(1), p. 111–121.e13. doi: 10.1016/j.cell.2016.09.004.

Seehafer, C. *et al.* (2011) 'From alpaca to zebrafish: Hammerhead ribozymes wherever you look', *RNA*. doi: 10.1261/rna.2429911.

Serganov, A. *et al.* (2006) 'Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch', *Nature*, 441(7097), pp. 1167–1171. doi: 10.1038/nature04740.

Serganov, A., Huang, L. and Patel, D. J. (2008) 'Structural insights into amino acid binding and gene control by a lysine riboswitch', *Nature*, 455(7217), pp. 1263–1267. doi: 10.1038/nature07326.

Skordalakes, E. *et al.* (2005) 'Structural mechanism of inhibition of the Rho

transcription termination factor by the antibiotic bicyclomycin.', *Structure (London, England : 1993)*, 13(1), pp. 99–109. doi: 10.1016/j.str.2004.10.013.

Skordalakes, E. and Berger, J. M. (2003) 'Structure of the Rho Transcription Terminator: Mechanism of mRNA Recognition and Helicase Loading', *Cell*, 114(1), pp. 135–146. doi: 10.1016/S0092-8674(03)00512-9.

Soukup, G. A. and Breaker, R. R. (1999) 'Design of allosteric hammerhead ribozymes activated by ligand-induced structure stabilization.', *Structure (London, England : 1993)*, 7(7), pp. 783–91. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10425680> (Accessed: 4 January 2019).

Stitt, B. L. (1988) 'Escherichia coli transcription termination protein rho has three hydrolytic sites for ATP.', *The Journal of biological chemistry*, 263(23), pp. 11130–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/3042765> (Accessed: 14 January 2019).

Sudarsan, N. *et al.* (2003) 'An mRNA structure in bacteria that controls gene expression by binding lysine', *Genes & Development*, 17(21), pp. 2688–2697. doi: 10.1101/gad.1140003.

Symmons, M. F. *et al.* (2002) 'Running rings around RNA: a superfamily of phosphate-dependent RNases.', *Trends in biochemical sciences*, 27(1), pp. 11–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11796219> (Accessed: 11 January 2019).

Thèze, J. *et al.* (1974) 'Mapping of the structural genes of the three aspartokinases and of the two homoserine dehydrogenases of Escherichia coli K-12.', *Journal of bacteriology*. American Society for Microbiology (ASM), 117(1), pp. 133–43. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/4148765> (Accessed: 21 August 2018).

Tolic', S. F. *et al.* (2003) 'Diversity in the Rates of Transcript Elongation by Single RNA Polymerase Molecules* □ S'. doi: 10.1074/jbc.M310290200.

Vanzo, N. F. *et al.* (1998) 'Ribonuclease E organizes the protein interactions in the Escherichia coli RNA degradosome.', *Genes & development*, 12(17), pp. 2770–81. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9732274> (Accessed: 11 January 2019).

Wang, J. X. and Breaker, R. R. (2008) 'Riboswitches that sense S -adenosylmethionine and S -adenosylhomocysteine This paper is one of a selection of papers published in this Special Issue, entitled CSBMCB — Systems and Chemical Biology, and has undergone the Journal's usual peer review process.', *Biochemistry and Cell Biology*, 86(2), pp. 157–168. doi: 10.1139/O08-008.

WATSON, J. D. and CRICK, F. H. C. (1953) 'Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid', *Nature*. Nature Publishing Group, 171(4356), pp. 737–738. doi: 10.1038/171737a0.

Wei, R. R. and Richardson, J. P. (2001a) 'Identification of an RNA-binding Site in the ATP Binding Domain of *Escherichia coli* Rho by H₂O₂/Fe-EDTA Cleavage Protection Studies', *Journal of Biological Chemistry*, 276(30), pp. 28380–28387. doi: 10.1074/jbc.M102444200.

Wei, R. R. and Richardson, J. P. (2001b) 'Mutational changes of conserved residues in the Q-loop region of transcription factor Rho greatly reduce secondary site RNA-binding', *Journal of Molecular Biology*, 314(5), pp. 1007–1015. doi: 10.1006/jmbi.2000.5207.

Weinberg, Z. *et al.* (2008) 'The aptamer core of SAM-IV riboswitches mimics the ligand-binding site of SAM-I riboswitches.', *RNA (New York, N.Y.)*, 14(5), pp. 822–8. doi: 10.1261/rna.988608.

Winkler, W. C. *et al.* (2003) 'An mRNA structure that controls gene expression by binding S-adenosylmethionine', *Nature Structural & Molecular Biology*, 10(9), pp. 701–707. doi: 10.1038/nsb967.

Winkler, W. C. *et al.* (2004) 'Control of gene expression by a natural metabolite-responsive ribozyme.', *Nature*, 428(6980), pp. 281–6. doi: 10.1038/nature02362.

Winkler, W., Nahvi, A. and Breaker, R. R. (2002) 'Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression.', *Nature*, 419(6910), pp. 952–6. doi: 10.1038/nature01145.

Wösten, M. M. (1998) 'Eubacterial sigma-factors.', *FEMS microbiology reviews*, 22(3), pp. 127–50. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9818380> (Accessed: 29 August 2018).

Xu, Y., Kohn, H. and Widger, W. R. (2002) 'Mutations in the Rho Transcription Termination Factor That Affect RNA Tracking*'. in Press. doi: 10.1074/jbc.M111009200.

Yanofsky, C. (1981) 'Attenuation in the control of expression of bacterial operons.', *Nature*, 289(5800), pp. 751–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7007895> (Accessed: 4 January 2019).

Zuo, Y. and Steitz, T. A. (2015) 'Crystal structures of the E. coli transcription initiation complexes with a complete bubble.', *Molecular cell*. Elsevier, 58(3), pp. 534–40. doi: 10.1016/j.molcel.2015.03.010.

