



LE DOUBLE RÔLE DE LA STRUCTURE PSEUDONOEUD CHEZ LE RIBORÉGULATEUR  
*BTUB D'ESCHERICHIA COLI*

par

Antony Lussier

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

FACULTÉ DES SCIENCES  
UNIVERSITÉ DE SHERBROOKE

Sherbrooke, Québec, Canada, Juin 2015

Le 22 juin 2015

*le jury a accepté le mémoire de Monsieur Antony Lussier  
dans sa version finale.*

Membres du jury

Professeur Daniel Lafontaine

Directeur de recherche

Département de biologie

Professeur Vincent Burrus

Évaluateur interne

Département de biologie

Professeur Sébastien Rodrigue

Président-rapporteur

Département de biologie

## SOMMAIRE

Chez les organismes procaryotes, l'adaptation rapide aux changements de l'environnement est synonyme de survie. Un des moyens utilisés pour s'adapter rapidement est un contrôle strict et efficace de l'expression génétique. Par exemple, un changement rapide de pH entraînera l'arrêt de synthèse de certains canaux dans le but de conserver un pH cellulaire interne stable se situant entre 7,2 et 7,8 chez la bactérie *Escherichia coli*.(Slonczewski et al. 1981) Alors, quand ces changements de l'environnement surviennent, la bactérie se doit de posséder des détecteurs efficaces qui peuvent réguler l'expression des gènes rapidement. Un des nombreux moyens utilisés chez les procaryotes pour détecter et contrôler l'expression génétique est l'utilisation des riborégulateurs. Ces ARN structurés sont localisés dans le 5' non codant de certains gènes bactériens et permettent une régulation efficace de l'opéron avec lequel ils sont associés. Les riborégulateurs lient directement un métabolite, appelé ligand, ce qui entraîne un changement dans la structure d'ARN et permet une régulation du (des) gène (s) en aval du riborégulateur. Pour accomplir ces deux étapes, soit la détection et la régulation, les riborégulateurs se doivent d'être très structurés pour effectuer une reconnaissance spécifique du ligand et pour réguler l'expression génétique. Un des riborégulateurs les plus structurés est le riborégulateur *btuB* chez la bactérie *E. coli*. En effet, ce riborégulateur lie l'adénosylcobalamine (AdoCbl) et est constitué de quatorze tiges boucles ainsi que d'une structure particulière nommée pseudonoeud. Cette structure a été montrée par le passé comme étant importante pour la réponse au ligand dans un contexte *in vivo*.(Nahvi et al. 2002) Par contre, son rôle dans le changement de structure lors de la liaison du ligand est encore méconnu. Le but ce projet de maîtrise était de déterminer le rôle de ce pseudonoeud dans le changement de structure du riborégulateur lors de la liaison de l'AdoCbl à ce dernier. Les résultats de cette étude ont non seulement permis de déterminer que le pseudonoeud est important pour la communication de la liaison du ligand, mais qu'il est aussi important pour la liaison. De plus, nous avons aussi confirmé que le contexte transcriptionnel est essentiel pour que le pseudonoeud puisse être formé, comme suggéré par les précédents travaux de Perdriquet et al, (2012).

Mots clés : *E. coli*, ARN, riborégulateur, *btuB*, adénosylcobalamine, pseudonoëud

## REMERCIEMENTS

J'aimerais, tout d'abord, remercier le professeur Daniel Lafontaine pour m'avoir laissé la chance de faire ma maîtrise dans son laboratoire. Grâce à lui, j'ai pu développer un esprit scientifique critique et éclairé dans une très belle ambiance de laboratoire. De plus, j'aimerais le remercier de m'avoir inclus dans son équipe même si je possédais une moyenne limite pour entrer à la maîtrise. Merci de m'avoir laissé cette chance.

Je voudrais aussi remercier mes conseillers Vincent Burrus et Sébastien Rodrigue pour leurs précieux conseils dans l'avancement de mon projet de maîtrise. Ils m'ont aidé à jeter un regard critique sur certains résultats de mon projet ce qui m'a permis de faire des expériences clés pour l'écriture de mon article scientifique.

Je veux aussi remercier toute l'équipe du laboratoire passé et présent qui m'ont permis de passer de très bons moments et de m'avoir aidé avec mes difficultés tout au long de mes stages et ma maîtrise. Je remercie donc, Maxime, Laurène, Adrien, Sébastien, Patrick, Mélanie, Anne-Marie, Julien, Jean-François, Pierre et Fatima. Merci à toute la gang!

Enfin, j'aimerais remercier les membres de ma famille mon père, ma mère et mon frère qui m'ont toujours fourni un support moral irremplaçable tout au long de mes études universitaires.

## TABLE DES MATIÈRES

SOMMAIRE.....	iii
REMERCIEMENTS .....	v
TABLE DES MATIÈRES.....	vi
LISTE DES ABRÉVIATIONS .....	viii
LISTE DES TABLEAUX .....	x
LISTE DES FIGURES .....	xi
CHAPITRE 1.....	1
INTRODUCTION GÉNÉRALE.....	1
1.1 Transcription bactérienne.....	1
1.1.1 Initiation de la transcription.....	2
1.1.2 Élongation de la transcription.....	3
1.1.3 Terminaison de la transcription.....	3
1.2 Les riborégulateurs.....	5
1.2.1 Découverte des riborégulateurs.....	5
1.2.2 Les classes de riborégulateur.....	8
1.2.3 Modes de régulation des riborégulateurs.....	11
1.2.4 Riborégulateurs coenzyme B <sub>12</sub> .....	14
1.2.5 Riborégulateur btuB d'Escherichia coli.....	19
1.3 Les pseudonoeds d'ARN.....	22
1.3.1 Définitions.....	22
1.3.2 Classifications des pseudonoeds.....	23
1.3.3 Fonctions des pseudonoeds dans la régulation génétique.....	25
CHAPITRE 2.....	30
IMPORTANCE OF A KISSING-LOOP FOR BTUB RIBOSWITCH LIGAND SENSING AND REGULATORY CONTROL .....	30
2.1 Résumé de l'article.....	31
2.2 Abstract.....	32
2.3 Introduction.....	33
2.4 Results.....	38
2.5 Discussion.....	60
2.6 Experimental procedures.....	64

2.7 References .....	70
2.8 Footnotes .....	73
CHAPITRE 3.....	74
DISCUSSION.....	74
CONCLUSION .....	81
BIBLIOGRAPHIE .....	84



## LISTE DES ABRÉVIATIONS

2'-dG	2'-deoxyguanosine
A	Adénine
Acétyl-CoA	Acétyl-coenzyme A
ADN	Acide désoxyribonucléique
AdoCbl	Adénosylcobalamine
ATP	Adénosine tri-phosphate
ARN	Acide ribonucléique
ARNm	ARN messenger
ARNpol	ARN polymérase ADN dépendante
ARNr	ARN ribosomal
ARNt	ARN de transfert
B12	Vitamine B <sub>12</sub>
C	Cytosine
c-di-AMP	di-adénosine monophosphate cyclique
c-di-GMP	di-guanosine monophosphate cyclique
F <sup>-</sup>	Ion fluore
FMN	Flavine mononucléotide
G	Guanine
GlcN6P	Glucosamine-6-phosphate
HDV	Virus de l'hépatite delta ( <i>Hepatitis delta virus</i> )
Mg <sup>+</sup>	Ion magnésium
Mono	Cofacteur molybdène

nt	Nucléotide
pH	Potentiel d'hydrogène
PN et PK	Pseudonoed ( <i>Pseudoknot</i> )
PreQ <sub>1</sub>	Pre-queuosine <sub>1</sub>
RBS	Site de liaison du ribosome ( <i>ribosome binding site</i> )
<i>rut</i>	<i>Rho-utilization site</i>
SAM	S-adénosylméthionine
SAH	S-adénosylhomocystéine
SD	Séquence Shine-Dalgarno
THF	Tetrahydrofolate
TPP	Thiamine pyrophosphate
Tuco	Cofacteur tungstène
U	Uracile
UTR	Région non-traduite (untranslated region)

## LISTE DES TABLEAUX

1. Classification des motifs d'ARN considérés ou pressentis comme étant des riborégulateurs. .... 9-11
2. Classification par gène des pseudonoeuds recensés ayant un rôle déterminé dans la régulation génétique bactérienne et de certains bactériophages ..... 27-28

## LISTE DES FIGURES

1.	Représentation générale d'un gène codant pour une protéine et des régions d'intérêts en amont et en aval de ce dernier .....	5
2.	Représentation schématique du mécanisme de régulation transcriptionnel par un riborégulateur .....	7
3.	Voies de synthèse de l'adénosylcobalamine déterminées à l'aide d'essais enzymatiques <i>in vitro</i> chez les bactéries <i>Pseudomonas denitrificans</i> et <i>Salmonella typhimurium</i> .....	16
4.	Structure secondaire de la famille de riborégulateurs B <sub>12</sub> .....	18
5.	Structures secondaires du riborégulateur <i>btuB</i> d' <i>Escherichia coli</i> .....	21
6.	Structure secondaire du pseudonœud présent dans le 3' terminal du virus de la mosaïque jaune du navet .....	23
7.	Classification des pseudonœuds selon leurs arrangements 2D .....	24

# CHAPITRE 1

## INTRODUCTION GÉNÉRALE

### 1.1 Transcription bactérienne

Dans tous les domaines du vivant, le métabolisme d'une cellule est constitué des mêmes processus biologiques essentiels. Tout d'abord, il y a l'ADN qui est transcrit en ARN messager (ARNm) lequel est traduit en protéines. Ces constituants de bases sont présents dans toute cellule depuis la bactérie jusqu'aux cellules eucaryotes. Comme la transcription est une des étapes critiques du métabolisme pour la survie et la reproduction, il est important de bien la comprendre. Il n'est cependant pas aisé de l'étudier puisque plusieurs acteurs protéiques sont impliqués dans ce processus extrêmement rapide. Aujourd'hui, c'est chez la bactérie Gram- *E. coli* que nous avons le plus d'information fonctionnelle sur l'enzyme responsable de la transcription. En effet, les promoteurs reconnus par cette enzyme sont connus ainsi que plusieurs acteurs protéiques interagissant avec celle-ci.(Feklistov 2014) Cette ARN polymérase ADN dépendante (ARNpol) a été la première à être isolée.(Furth et al. 1962) Cette enzyme est constituée d'un cœur de cinq sous-unités polypeptidiques ( $\alpha_2\beta\beta'\omega$ ) et d'un facteur de transcription accessoire sigma ( $\sigma$ ). Chez les organismes bactériens, une seule ARNpol est responsable de la production totale de l'ARN de la cellule.(A Darst 2001) Le processus de la transcription est aussi responsable en grande partie de la structure de l'ARN suite à la transcription. En effet, il existe des sites de pauses de la transcription qui permettent à des structures d'ARN (tige boucle, pseudonoeud, etc.) de prendre forme même si celles-ci ne sont pas nécessairement les plus favorisées au niveau de la thermodynamique.(Perdrizet et al. 2012) Il faut donc considérer l'importance du phénomène de la transcription lors d'études fonctionnelles et structurales des ARN.(Koduvayur & Woodson 2004)(Wickiser et al. 2005) Enfin, le processus complet de la transcription peut être divisé en trois grandes étapes qui seront

discutées dans les prochaines pages soit, l'initiation, l'élongation et la terminaison de la transcription.

### **1.1.1 Initiation de la transcription**

Tout d'abord, pour que la transcription de l'ADN en ARN puisse avoir lieu dans une cellule bactérienne, il faut que l'ARN polymérase passe par l'étape essentielle d'initiation de la transcription. Cette étape consiste en la recherche d'une séquence particulière appelée promoteur par l'ARNpol sous forme d'holoenzyme, c'est-à-dire la polymérase associée avec un facteur sigma. Ces séquences, situées en amont des gènes, sont constituées de plusieurs dizaines de nucléotides. Les promoteurs les plus communs sont caractérisés par deux éléments de séquence conservés appelés boîte -35 (TTGACA) et boîte -10 (TATAAT) (voir Figure 1). Ces deux boîtes font parties des éléments reconnus par un facteur sigma qui est le plus souvent le facteur sigma 70. Cependant, seul l'élément -10 est absolument essentiel pour l'initiation de la transcription.(Feklistov et al. 2006) Une fois le complexe transcriptionnel présent au niveau du promoteur (ARNpol et facteur sigma), il y a ouverture de l'ADN double brin par interaction de charge entre l'ADN chargé négativement et le site actif de l'ARN polymérase chargé positivement. Par la suite, le brin codant d'ADN est introduit dans le site actif de la polymérase et il y a formation de la bulle de transcription qui se propage en direction 3' du brin d'ADN engagé. Par contre, la bulle de transcription peut changer de taille et se modifier pour permettre une initiation correcte au site d'initiation de la transcription.(Robb et al. 2013) Il y aura alors incorporation du premier nucléotide au site +1 de la transcription. Il est communément accepté que c'est à la suite de ce phénomène que se termine l'initiation de la transcription et que débute la phase d'élongation de la transcription.(Hsu 2002)

### 1.1.2 Élongation de la transcription

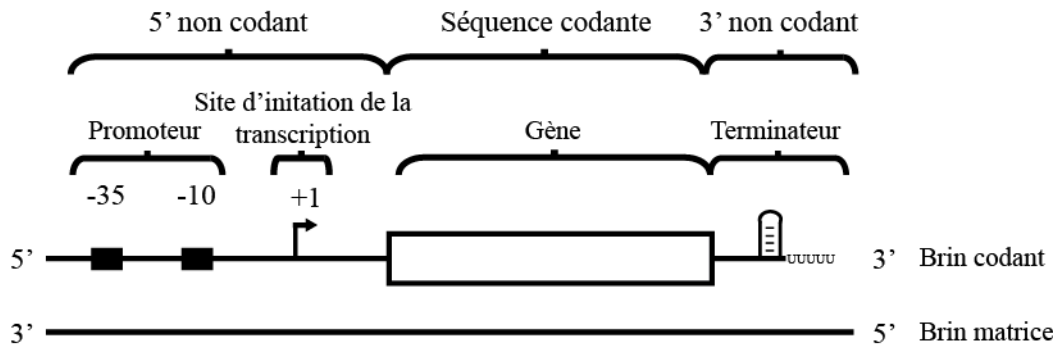
Lorsque l'ARNpol est engagée sur le promoteur, il y a ajout du premier ribonucléotide au site +1 de la transcription. L'ajout de nucléotides s'effectue selon un procédé de complémentarité très précis, cytosine-guanine et adénine-uridine. C'est grâce à ce principe de complémentarité que la synthèse de l'ARN se produit à un taux de fidélité très élevé. Lorsque le premier ribonucléotide est en place, le deuxième ribonucléotide se positionne dans le site actif et la polymérase catalyse alors la formation d'un lien phosphodiester avec le premier ribonucléotide. Une fois le lien phosphodiester produit, il y a libération d'une molécule de pyrophosphate inorganique qui est produit lors de la polymérisation du brin d'ARN. Après le relâchement de cette molécule, le segment d'ARN synthétisé et l'ADN sont décalés de 1 nucléotide sur l'ADN en direction 3' par le processus de translocation. À ce moment, le prochain nucléotide d'ARN est ajouté à la chaîne naissante.(Robb et al. 2013) Ce cycle se répète plusieurs fois jusqu'à ce que l'ARN atteigne une taille entre 7 et 10 nucléotides, parfois plus. Cependant, la polymérase reste fixée au promoteur pendant les premières étapes de la polymérisation jusqu'au moment où la tension entre l'ADN et la polymérase devient si forte, qu'il peut alors se produire deux phénomènes : le premier phénomène possible est celui de la tension sur l'ARNpol qui brise l'hybride ADN-ARN en formation avortant ainsi la transcription. Le second phénomène est celui du facteur sigma décrochant du promoteur et permettant ainsi à l'ARNpol de poursuivre la transcription jusqu'au bout de la séquence codante.(Revyakin et al. 2006)(Margeat et al. 2006) Il est connu qu'en conditions *in vitro*, les quantités de sels et la présence d'autres facteurs reliés à la transcription (GreA et GreB) peuvent influencer la proportion de transcrit continuant dans la phase d'élongation.(Susa et al. 2006) Cependant, à ce jour, la proportion de transcrits achevés et inachevés n'a pas encore été déterminée *in vivo*.

### 1.1.3 Terminaison de la transcription

Le plus souvent, la transcription se termine à la fin de l'opéron ou du gène transcrit, par la présence d'un terminateur intrinsèque de transcription sur l'ARN. Ce type de terminateur se

forme sur l'ARN naissant et est constitué d'une tige boucle riche en paires de bases guanine-cytosine ainsi que d'une courte séquence d'uracile (voir Figure 1). Ce type de terminateur induit une pause de la transcription et une déstabilisation de l'hybride ARN-ADN dans la bulle de transcription ce qui mène à l'arrêt de la transcription et à la dissociation du transcrit et de l'ARNpol.(Wilson & von Hippel 1995) D'autre part, il existe un autre mécanisme d'arrêt de la transcription qui est très répandu chez les procaryotes et qui consiste en la terminaison de transcription dépendant de Rho.(Boudvillain et al. 2013) Jusqu'à ce jour, très peu d'informations sont disponibles quant à la nature des séquences des terminateurs de transcription Rho. Ces séquences, nommées *rut* (*Rho utilization site*), sont peu structurées, riches en cytosines et pauvres en guanines et ne sont pas conservées entre les terminateurs de cette famille. Concernant le mécanisme d'action de ce terminateur, il est connu que la protéine Rho interagit avec l'ARN au site *rut*. De plus, ce processus de terminaison dépendant de Rho a lieu entre 10 et 100 nucléotides après le site *rut*.(Richardson & Richardson 1996)(Chae et al. 2011) À ce jour, plusieurs modèles sont proposés quant au mécanisme d'action de Rho, mais aucun n'est encore confirmé. Il est connu que cette protéine a la capacité de se déplacer sur l'ARN grâce à son activité ATPase et qu'elle se déplace pour se rapprocher de l'ARNpol.(Koslover et al. 2012) Par contre, ayant atteint la position de l'ARNpol, la protéine Rho pourrait cesser la transcription de plusieurs manières. Elle pourrait agir soit par contact allostérique avec la polymérase, ce qui perturberait le complexe d'élongation, soit par contact physique, tel un béliet,(Park & Roberts 2006) déstabilisant ainsi l'hybride ARN-ADN et menant à l'arrêt de la transcription.(Epshtein et al. 2010) La terminaison de la transcription peut être dépendante de plusieurs mécanismes de régulation tels que l'action de facteurs de transcription, l'action des protéines N et Q et les riborégulateurs, pour n'en nommer que quelques-uns.(Henkin 1996)





**Figure 1. Représentation générale d'un gène codant pour une protéine et des régions d'intérêts en amont et en aval de ce dernier.** Dans la région 5' non codante d'un gène, on retrouve le promoteur, situé entre les nucléotides -35 et -10 en amont du site d'initiation de la transcription (+1). Ensuite, la transcription du gène a lieu tout au long de la séquence codante jusqu'au terminateur de transcription, dans la région 3' non codante. Il est à noter que la séquence du terminateur se trouve sur l'ADN mais que la structure tige boucle est formée sur l'ARN.

## 1.2 Les riborégulateurs

### 1.2.1 Découverte des riborégulateurs

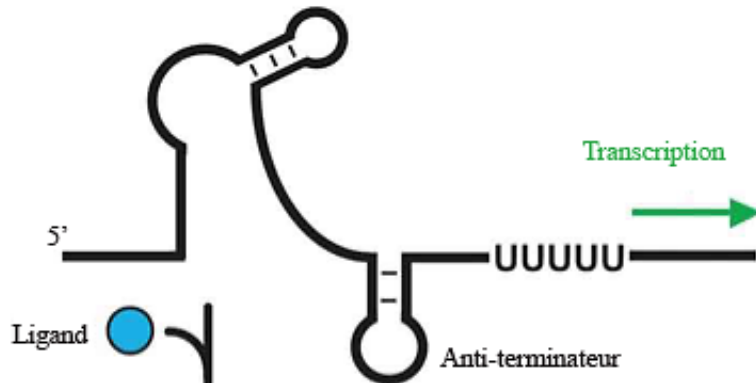
Ces éléments de régulation génétique sont de découverte récente (2002). (Nahvi et al. 2002) Avant leur identification, il était supposé qu'un ARN ne pouvait pas exercer le double rôle de détecteur et de régulateur. En effet, avant leur découverte, il était admis dans la communauté scientifique qu'un élément protéique senseur devait absolument interagir avec l'ARN ou l'ADN pour induire une régulation génétique. Ce paradigme a été mis en doute dans un article de Xiangwu Nou et Robert Kadner (Nou & Kadner 2000) où il a été démontré que les ribosomes n'avaient plus accès à l'ARNm du gène *btuB* d'*E.coli* en présence d'adénosylcobalamine (AdoCbl). Les auteurs ont conclu l'article sans déterminer si l'ARN était capable de lier directement l'AdoCbl ou bien si l'AdoCbl interagissait avec une protéine ribosomale qui ensuite se liait à l'ARNm. Ce n'est que deux ans plus tard que l'interaction entre l'AdoCbl et l'ARN de *btuB* fut clairement confirmée. De plus, une première structure secondaire de ce riborégulateur

fut proposée, ce qui en fit le premier riborégulateur découvert.(Nahvi et al. 2002) Dans cet article, la liaison spécifique du ligand est clairement montrée à l'aide d'essais *in vitro*. Cet article dénote aussi l'importance de certaines structures clés dans le changement de conformation de l'ARN lors de la liaison du ligand dans un contexte bactérien. Depuis, environ une vingtaine de classes de riborégulateurs ont été cataloguées chez les procaryotes ainsi que chez certains eucaryotes.(Peselis & Serganov 2014b) Les riborégulateurs se doivent d'adopter un grand éventail de structures différentes pour permettre la liaison de leur ligand de manière spécifique. De plus, deux riborégulateurs liant le même ligand chez deux bactéries différentes peuvent être très différents au point de vue structural.(Batey) Dans la même voie, le mode de régulation de deux riborégulateurs AdoCbl, par exemple, peut varier d'un genre bactérien à un autre. Aujourd'hui, l'existence et l'importance des riborégulateurs sont bien établies et les défis actuels consistent essentiellement à trouver de nouvelles classes de riborégulateurs et à déterminer leurs structures tridimensionnelles détaillées. De cette manière, il sera possible d'établir leurs mécanismes d'action et de les exploiter comme cibles antibiotiques potentielles.

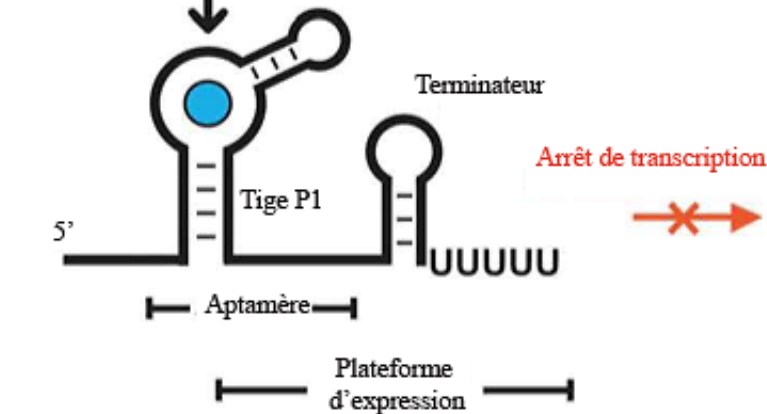
Généralement, un riborégulateur est une structure d'ARN interagissant avec un métabolite appelé ligand et qui, au travers de cette liaison spécifique, régule l'expression génétique d'un opéron. Le ligand associé au riborégulateur est habituellement relié au métabolisme du gène régulé par ce même riborégulateur. Ils sont généralement retrouvés chez les bactéries et sont localisés dans la section 5' non codante d'un ARNm mais ils peuvent aussi se retrouver dans la section 3' non codante ou dans un intron chez les organismes eucaryotes. Ces structures d'ARN sont délimitées en deux sections plus ou moins distinctes : l'aptamère et la plateforme d'expression (voir Figure 2). L'aptamère est la région responsable de la reconnaissance du ligand tandis que la plateforme d'expression est responsable de la régulation génétique. Lors de la liaison du ligand, le riborégulateur régule le ou les gène(s) en aval par l'entremise d'un changement de structure de la plateforme d'expression (Figure 2B).(Winkler & Breaker 2003) Ce changement de conformation crée, dans cet exemple, un terminateur intrinsèque de la transcription. Le seul exemple connu ne suivant pas ce schéma est le riborégulateur-ribozyme *glmS*.(Tinsley et al. 2007) Ce dernier ne change pas de conformation lors de la liaison du ligand

mais accomplit son rôle de régulateur grâce à son activité ribozyme dont le ligand (glucose-6-phosphate) est un cofacteur.

A



B



**Figure 2. Représentation schématique du mécanisme de régulation transcriptionnel par un riborégulateur.** Nous pouvons remarquer que lorsque le ligand (bleu) se lie à l'aptamère, il y a formation de la tige P1. Cet événement participe au changement de conformation de la plateforme d'expression formant un terminateur de transcription intrinsèque, ce qui arrêtera la transcription et fait décrocher l'ARNpol. Adapté de Jane N. Kim et Ronald R. Breaker (2008)(Kim & Breaker 2008).

### 1.2.2 Les classes de riborégulateur

De nos jours, de plus en plus de riborégulateurs sont identifiés. Plusieurs types de riborégulateurs sont répertoriés ou prédits à l'aide de techniques bio-informatiques, ce qui contribue à la découverte de nouveaux riborégulateurs (voir Tableau 1). (Sun et al. 2013) Il devient alors important d'avoir un système de classification pour les comparer. La méthode la plus utilisée consiste à les classer en fonction du ligand reconnu. (Breaker 2011) On peut les regrouper par grandes familles : sont réportoriées celle des acides nucléiques (adénine, guanine, ATP, PreQ<sub>1</sub>, c-di-GMP, c-di-AMP, 2'-dG), celle des acides aminés (lysine, glycine, glutamine), celle des cofacteurs et leurs dérivés (AdoCbl, TPP, FMN, SAM, SAH, Mono/Tuco, THF), celle des ions (Mg<sup>+</sup>, Mn<sup>+</sup> et F<sup>-</sup>) et d'autres types de molécule de nature variée (GlcN6P, ARNt et aminoglycoside). (Serganov & Nudler 2013) Il est à noter que certains ligands peuvent être reconnus par plusieurs motifs d'ARN différents et donc plusieurs aptamères différents, comme dans les cas des riborégulateurs SAM, c-di-GMP et PreQ<sub>1</sub>. Les riborégulateurs pourraient aussi être classés selon leur mode de régulation. Par contre, comme le mode de régulation d'un grand nombre de riborégulateurs est prédit et non confirmé, il est alors plus simple de les classer selon leur ligand potentiel ou confirmé.

**Tableau 1. Classification des motifs d'ARN considérés ou pressentis comme étant des riborégulateurs.**

Nom du motif d'ARN	Ligands (riborégulateurs)	Nombre estimé	Validation expérimentale	NoRfam
Cobalamine ( <i>B12</i> element)	Adenosylcobalamine	536	Oui	RF00174
FMN ( <i>RFN</i> element)	Flavine mononucléotide	233	Oui	RF00050
GEMM	di-GMP cyclique	89	Oui	RF01051
glmS	Glucosamine-6-phosphate	44	Oui	RF00234
glnA	Glutamine	13	Oui	RF01739
Glycine	Glycine	324	Oui	RF00504
Lysine (L-box)	Lysine	186	Oui	RF00168
mini-ykkC	?	67	Non	RF01068
MoCo	Molybdène ou tungstène	62	Oui	RF01055
PreQ1	Pre-queuosine <sub>1</sub>	72	Oui	RF00522
PreQ1-II	Pre-queuosine <sub>1</sub>	17	Oui	RF01054
Purine (G-Box)	Guanine, adénine	141	Oui	RF00167
SAH	S-adenosylhomocystéine	27	Oui	RF01057
SAM (S-box)	S-adénosylméthionine	257	Oui	RF00163
SAM-Alpha	S-adénosylméthionine?	39	Oui	RF00521
SAM-Chlorobi	S-adénosylméthionine	11	Non	RF01724

**Tableau 1 (suite). Classification des motifs d'ARN considérés ou pressentis comme étant des riborégulateurs.**

Nom du motif d'ARN	Ligands (riborégulateurs)	Nombre estimé	Validation expérimentale	NoRfam
SAM-IV	S-adénosylméthionine	13	Oui	RF00634
SAM-SAH	S-adénosylméthionine,S-adenosylhomocystéine	13	Oui	RF01727
Senseur Mg	Magnésium	5	Oui	RF01056
serC	?	32	Non	RF00517
Smk-box	S-adénosylméthionine	13	Oui	RF01767
speF	?	12	Non	RF00518
sucA	?	14	Non	RF01070
T-box	ARNt	1134	Oui	RF00230
THF	Tetrahydrofolate	23	Oui	RF01831
TPP ( <i>THI</i> element)	Thiamine pyrophosphate	564	Oui	RF00059
ybhL	?	8	Non	RF00520
ydaO-yuaA	di-AMP cyclique	59	Oui	RF00379
ykkC-yxkD	?	58	Non	RF00442
ykoK (M-box)	Magnésium	48	Oui	RF00380
ylbH	?	10	Non	RF00516

**Tableau 1 (suite). Classification des motifs d'ARN considérés ou pressentis comme étant des riborégulateurs**

Nom du motif d'ARN	Ligands (riborégulateurs)	Nombre estimé	Validation expérimentale	NoRfam
yybP-ykoY	Manganèse	232	Oui	RF00080

Y sont indiqués : le ligand connu pour les motifs riborégulateurs, le nombre de motifs trouvés par analyse bio-informatique, la confirmation expérimentale de l'existence des motifs et le numéro de référence des motifs associés de la banque de données Rfam pour référence. Adapté de Sun, E. I. et al., (2013)(Sun et al. 2013).

### 1.2.3 Modes de régulation des riborégulateurs

D'abord, le premier mode de régulation découvert a été celui de la régulation traductionnelle.(Nahvi et al. 2002) La régulation traductionnelle se produit lorsque le ligand se lie à l'aptamère et qu'il y a séquestration de la séquence Shine-Dalgarno (SD) et/ou du premier codon, dans la plateforme d'expression sous forme d'une tige-boucle. De cette manière, le ribosome n'a plus accès à la séquence SD pour lier l'ARNm et la traduction du gène associé est alors grandement diminuée. Plusieurs riborégulateurs utilisent cette méthode de régulation comme le riborégulateur AdoCbl(Nahvi et al. 2002) ainsi qu'un riborégulateur TPP(Serganov et al. 2006) chez *E. coli*. De plus, ce mode de régulation est parmi les plus répandus chez les bactéries Gram-.

Ensuite, la régulation transcriptionnelle Rho-indépendante est aussi utilisée, chez les bactéries Gram+, comme mécanisme de régulation de l'expression génétique. Tout comme la régulation traductionnelle, la liaison du ligand à l'aptamère entraîne la formation d'une tige-boucle dans la plateforme d'expression (voir Figure 2). Par contre, celle-ci se situe en amont, ou proche de la séquence SD, et agit comme terminateur de la transcription. En effet, la structure adoptée est

très similaire à un terminateur intrinsèque de la transcription, c'est-à-dire une tige-boucle suivie d'une séquence poly-U. De cette manière, le complexe transcriptionnel est déstabilisé, dissocié, et il y a production d'un transcrit tronqué qui est rapidement dégradé.(Winkler et al. 2003) Ce type de régulation est parmi les plus répandus et il est retrouvé, par exemple, chez les riborégulateurs adénine(Mandal & Breaker 2004), guanine(Mandal et al. 2003) et lysine(Sudarsan et al. 2003) de la bactérie *Bacillus subtilis*.

De plus, il existe des riborégulateurs qui régulent la transcription de manière Rho-dépendante. Ces riborégulateurs utilisent l'accessibilité d'un site *rut* qui devient disponible lors de la liaison du ligand. C'est alors que la protéine Rho peut se lier au site *rut* et effectuer un arrêt de la transcription par Rho, tel que décrit à la section 1.1.3. Ce mode de régulation, encore peu courant, est retrouvé chez le riborégulateur magnésium de la bactérie *Salmonella enterica*.(Hollands et al. 2012) Ce mécanisme est aussi présent chez le riborégulateur FMN de la bactérie *E. coli*, mais la régulation imposée par le riborégulateur est exécutée conjointement avec une régulation traductionnelle ce qui en fait un riborégulateur à double mécanisme de régulation.(Pedrolli et al. 2015)

Il existe un autre mécanisme qui utilise les capacités catalytiques de l'ARN. À ce jour, il n'y a qu'un riborégulateur connu qui utilise ce mécanisme, le riborégulateur-ribozyme *glmS* retrouvé chez plusieurs bactéries Gram+. Ce riborégulateur particulier se sert de la liaison du ligand comme coenzyme pour initier le clivage de l'ARN. Il a été démontré(Tinsley et al. 2007) qu'il n'y a pas de changement de conformation du riborégulateur lors de la liaison du ligand, mais que ce riborégulateur se sert du ligand pour exécuter la réaction de clivage. Ainsi, lorsque le ligand est présent en quantité suffisante, le riborégulateur provoque l'auto-clivage de son propre ARN entraînant ainsi sa dégradation et empêchant par le fait même la traduction du gène *glmS*.(Klein & Ferré-D'Amaré 2006)

Chez les organismes eucaryotes contenant des riborégulateurs connus, un tout autre mécanisme de régulation génétique est utilisé. Effectivement, le type de régulation utilisé par ces



riborégulateurs est la régulation de l'épissage alternatif. À la différence des organismes procaryotes, les riborégulateurs présents chez les eucaryotes sont, la plupart du temps, retrouvés à l'intérieur d'un intron. Pour ces riborégulateurs, le changement de conformation de la plateforme d'expression rend accessible un site d'épissage différent qui est reconnu préférentiellement et qui produit un ARNm instable. Cet ARNm instable est alors rapidement dégradé, diminuant l'expression du gène concerné.(Cheah et al. 2007) Aujourd'hui seulement quelques organismes eucaryotes possèdent des riborégulateurs confirmés. Ils sont tous des riborégulateurs reconnaissant le TPP et ils sont présents chez certaines espèces de plantes(Wachter et al. 2007)(Bocobza et al. 2007), chez deux espèces d'algues (*Chlamydomonas reinhardtii* et *Volvox carteri*)(Croft et al. 2007) et chez deux espèces de champignons (*Aspergillus oryzae* et *Neurospora crassa*).(Cheah et al. 2007)

Un nouveau mécanisme a été découvert récemment chez le riborégulateur *lysC* d'*E. coli*, appelé la régulation nucléolytique.(Caron et al. 2012) Le mécanisme consiste à promouvoir la dégradation rapide de l'ARNm à l'aide d'un complexe protéique appelé le dégradosome. Lorsque le ligand se lie au riborégulateur, deux sites de clivage du dégradosome deviennent alors accessibles. Ce complexe protéique clive alors l'ARNm, ce qui le rend instable et amène une dégradation rapide de ce dernier.(Caron et al. 2012)

Enfin, un mécanisme de régulation a été identifié, où l'aptamère d'un riborégulateur SAM peut agir sur un autre gène en tant que petit ARN chez la bactérie *Listeria monocytogenes*.(Loh et al. 2009) En effet, ce riborégulateur transcriptionnel forme un petit transcrit stable suite à la liaison du ligand et à l'arrêt de la transcription. Ce transcrit, constitué essentiellement de l'aptamère SAM, lie l'ARNm d'un gène de virulence au niveau de la séquence SD. Cette dernière est alors masquée par le petit ARN hybridé sur l'ARNm et le ribosome n'a donc plus accès à l'ARN pour initier la traduction. Il y a alors régulation d'un gène en *trans* par un petit ARN constitué d'un aptamère de riborégulateur. L'explication proposée quant au lien physiologique entre le facteur de virulence et un gène du métabolisme du SAM, serait que la virulence de la bactérie est influencée par la disponibilité des nutriments, dans ce le S-adenosylméthionine. Il existe aussi

des exemples où un riborégulateur AdoCbl module la transcription d'un petit ARN à l'aide d'une régulation transcriptionnelle.(Mellin et al. 2013)(DebRoy et al. 2014) Dans ce cas, quand le ligand est lié au riborégulateur, il y a inhibition de la transcription du petit ARN à l'aide d'un terminateur de transcription intrinsèque. Ce petit ARN, quant à lui, régule la production d'un facteur de transcription appelé *aspocR* en agissant comme ARN anti-sens sur l'ARNm de *aspocR*.(Mellin et al. 2013) Par conséquent, le riborégulateur régule de façon indirecte l'expression d'un facteur de transcription en permettant ou non la production d'un petit ARN régulateur.

Le plus souvent, les riborégulateurs régulent l'expression génétique à l'aide d'un seul mécanisme. Par contre, certains riborégulateurs peuvent utiliser la combinaison de deux mécanismes de régulation. Ces cas de double mécanismes étant rares(Caron et al. 2012)(DebRoy et al. 2014) et étant composés de mécanismes connus et mentionnés précédemment, ils ne seront donc pas décrits plus en détail. De plus, le mécanisme de régulation n'est pas nécessairement le même pour deux riborégulateurs reconnaissant le même ligand. Par exemple, un riborégulateur TPP chez *E. coli* contrôlant la traduction(Rentmeister et al. 2007) régule dans un autre cas l'épissage alternatif chez certaines plantes.(Cheah et al. 2007) Avec l'aide d'outils expérimentaux et bio-informatiques qui ne cessent de s'affiner, nous pouvons nous attendre à découvrir, dans les prochaines années, de nouveaux mécanismes de régulation chez les riborégulateurs.

#### **1.2.4 Riborégulateurs coenzyme B<sub>12</sub>**

La molécule utilisée comme ligand par les riborégulateurs B<sub>12</sub> est l'adénosylcobalamine. Ce métabolite aussi appelé coenzyme B<sub>12</sub>, ado-cobalamine ou vitamine B<sub>12</sub>, fait partie de la famille des cobalamines. Cette famille de molécule est assez restreinte et est composée de grandes molécules de plus de 1500 daltons qui peuvent utiliser près de vingt-cinq enzymes pour leur synthèse (voir Figure 3). Comme ce sont des molécules imposantes, elles doivent absolument utiliser des transporteurs membranaires pour pénétrer à l'intérieur des micro-

organismes.(Bradbeer 1993) En effet, la récupération de ces molécules sous forme de cobinamides ou de cobalamine dans l'environnement est bien moins coûteuse en énergie pour une cellule qu'une synthèse complète. Certains organismes comme *E. coli* ne peuvent pas synthétiser la cobalamine *de novo*, le principal intermédiaire de synthèse de la vitamine B<sub>12</sub>.(Warren et al. 2002) Ils doivent donc la récupérer dans l'environnement à partir des corrinoides ou bien utiliser une voie de synthèse alternative plus longue pour produire de la vitamine B<sub>12</sub>. Cette vitamine est nécessaire à la plupart des organismes vivants depuis la bactérie jusqu'à l'Homme.(Yamada 2013) Par contre, seulement certaines espèces de bactéries et d'archéobactéries peuvent synthétiser la vitamine B<sub>12</sub> *de novo*, les autres organismes doivent donc récupérer la cobalamine dans l'alimentation. Ce coenzyme est nécessaire pour plusieurs réactions physiologiques chez les bactéries. Il sert d'abord à balancer les réactions d'oxydoréduction lors de la fermentation et à fournir un aldéhyde pour la synthèse d'adénosine triphosphate.(Roth et al. 1996) Ensuite, l'AdoCbl permet un grand éventail de réactions enzymatique comme la synthèse de la méthionine(González et al. 1992), un acide aminé, ainsi que la décomposition du glycérol(Forage & Foster 1982) et du propanediol(Toraya et al. 1979) pour la production d'acétyl-CoA.(Blackwell & Turner 1978)

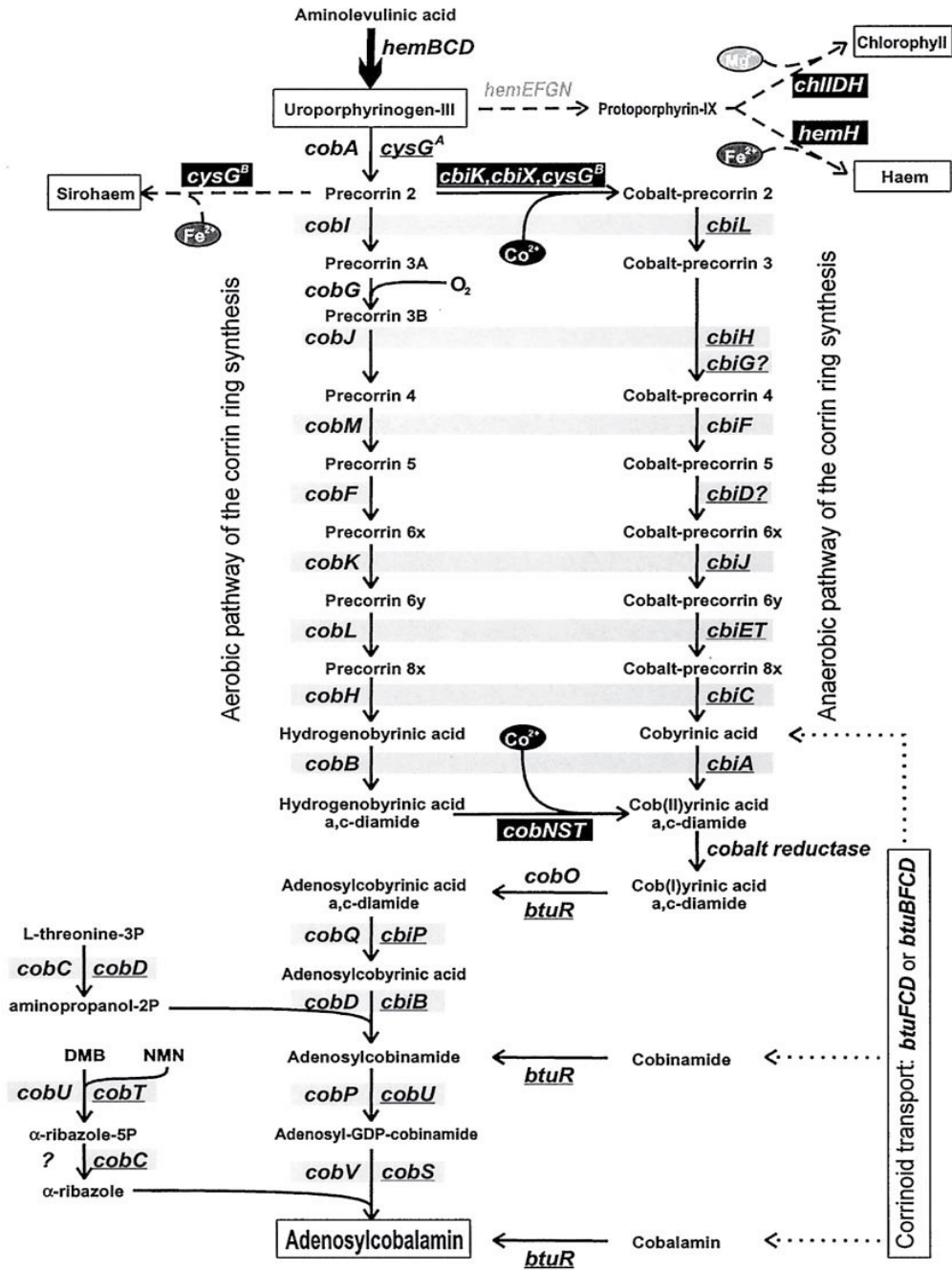
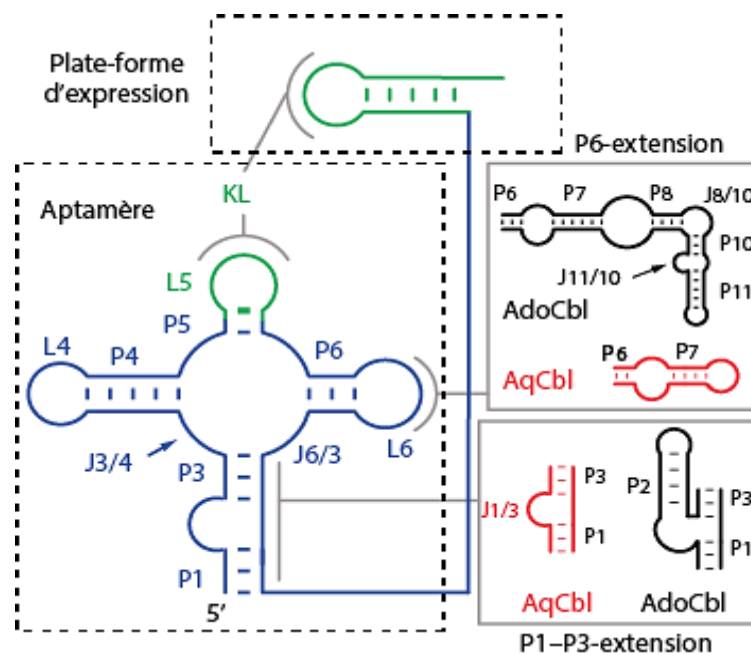


Figure 3. Voies de synthèse de l'adénosylcobalamine déterminées à l'aide d'essais enzymatiques *in vitro* chez les bactéries *Pseudomonas denitrificans* (pour la voie aérobie) et *Salmonella typhimurium* (pour la voie anaérobie). À gauche est représentée la voie de

synthèse aérobie et à droite la synthèse anaérobie *de novo* de la vitamine B<sub>12</sub>. Les similarités génétiques majeures entre les deux voies de synthèse sont surlignées en gris tandis que les gènes en noir peuvent différer selon l'organisme. Tiré de Rodionov, D et al.,(2003)(D. a Rodionov et al. 2003)

Comme la vitamine B<sub>12</sub> est un métabolite essentiel, elle se doit d'être régulée de manière stricte et efficace. Chez les procaryotes, les riborégulateurs sont parfois utilisés pour réguler certains gènes importants pour la synthèse ou le transport de cette vitamine. Ce type de riborégulateur est très répandu dans la nature et est surclassé seulement par les riborégulateurs TPP et T-box (voir Tableau 1). Le mode de régulation des riborégulateurs AdoCbl peut varier entre la régulation transcriptionnelle,(Butzin et al. 2013) traductionnelle(Nahvi et al. 2002) ou bien la régulation en *trans*(Mellin et al. 2013) selon l'espèce bactérienne dans laquelle le riborégulateur se retrouve. De plus, une structure consensus a été déterminée par alignement de séquence (voir Figure 5). Cette structure consensus appelée «B12 box» est utilisée pour identifier de nouveaux riborégulateurs AdoCbl.(Barrick & Breaker 2007) Elle peut être très variable en complexité, en composition de nucléotide ainsi qu'en taille ce qui rend parfois l'identification de nouveaux riborégulateurs difficiles sans confirmations expérimentales. Plusieurs éléments importants ont cependant été identifiés chez les riborégulateurs AdoCbl : le «B12 box» ou cœur conservé, l'extension structurale P6 et un pseudonoeud d'ARN (KL) (voir Figure 5). Le cœur est le plus conservé en termes de structure, de séquence et est proposé pour être le module essentiel à la reconnaissance du ligand. De plus, il est abondamment retrouvé lors des alignements bio-informatique.(Barrick & Breaker 2007) L'extension P6 a été identifiée comme étant une structure facultative, mais procurant de la spécificité à la liaison de l'AdoCbl.(Johnson et al. 2012) Cet élément permettrait une liaison distinctive de l'AdoCbl au riborégulateur par rapport à des analogues proches comme le méthylcobalamine, l'aquocobalamine et l'hydroxycobalamine. Le dernier élément conservé, le pseudonoeud, est lui aussi souvent retrouvé chez les riborégulateurs AdoCbl. Cet élément a été observé pour la première fois chez le riborégulateur AdoCbl d'*E. coli* par Nahvi et al.(Nahvi et al. 2002) L'équipe de recherche a constaté que cette structure, qui est formée en présence du ligand, est essentielle pour la

régulation génétique par le riborégulateur. De plus, cette structure occupe une place particulière entre la tige P1 et la tige séquestratrice du RBS située dans la plateforme d'expression. Cette position est particulière puisque habituellement, chez les riborégulateurs, c'est la formation de la tige P1 qui provoque directement le changement de structure de la plateforme d'expression. Cette communication se fait normalement sans l'intervention d'une région structurée entre ces deux éléments. Grâce à l'analyse des structures cristallines de riborégulateurs B<sub>12</sub>, il a été proposé que le pseudonoeud serait important pour la formation de la poche de liaison du ligand.(Johnson et al. 2012)(Peselis & Serganov 2012) En effet, dans les cristaux, l'AdoCbl se retrouve à proximité de cette structure et il a été suggéré que l'AdoCbl pourrait faire des interactions avec le groupe phosphate des nucléotides composant le pseudonoeud. De plus, dans les cristaux, le pseudonoeud se referme par-dessus le cœur du riborégulateur, ce qui pourrait participer à piéger le ligand à l'intérieur du riborégulateur. Finalement, il semble que le pseudonoeud des riborégulateurs B<sub>12</sub> possède divers rôles tels que la stabilisation du ligand et le changement de conformation de la plateforme d'expression lors de la liaison du ligand.(Nahvi et al. 2002)(Johnson et al. 2012)



**Figure 4. Structure secondaire de la famille de riborégulateurs B<sub>12</sub>.** Le cœur conservé appelé «B12 box» est représenté en bleu et le pseudonoeud (KL) facultatif en vert. Les extensions de

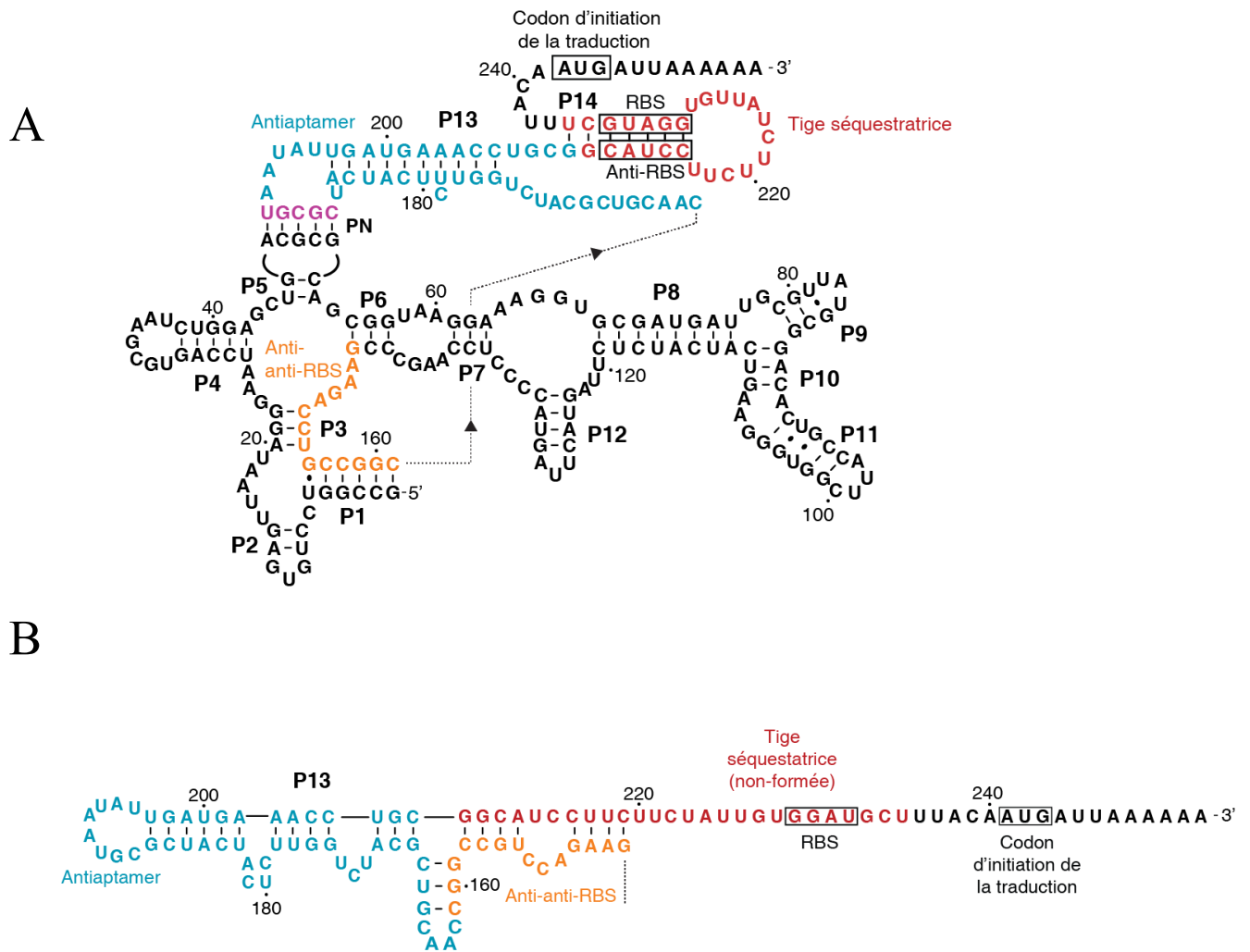
tiges périphériques, qui sont aussi facultatives, sont illustrées en rouge et noir pour les classes de riborégulateurs aquocobalamines et adénosylcobalamines respectivement. Tiré de Johnson, J. E et al, (2012).(Johnson et al. 2012)

### **1.2.5 Riborégulateur *btuB* d'*Escherichia coli***

Comme mentionné précédemment, le riborégulateur *btuB* de la bactérie *E. coli* a été le premier riborégulateur répertorié. Le groupe de recherche de Robert Kadner a remarqué à la fin des années 80 que l'expression du gène *btuB* chez *E.coli* était réprimée lorsque la bactérie était en croissance dans un milieu riche en cobalamines.(Lundrigan & Kadner 1989) Environ dix ans plus tard, la même équipe démontrait que cette régulation par les cobalamines était spécifique à l'AdoCbl. De plus, ils ont aussi démontré que la régulation de *btuB* avait lieu à la limite entre le 5' non codant et la séquence codante de *btuB* au niveau de la répression de la traduction.(Nou & Kadner 2000)(Kadner 1998) De plus, cette équipe a déterminé que la répression de la traduction de *btuB* avait lieu grâce à la séquestration de la séquence RBS dans une tige boucle prédite par bio-informatique.(Nou & Kadner 2000) Deux ans plus tard, l'équipe de Ronald Breaker(Nahvi et al. 2002) proposa le terme riborégulateur pour définir la structure d'ARN de *btuB* se situant dans le 5' non-codant. Cette structure, qui a été proposée à l'aide de cartographie chimique en ligne et de bio-informatique, lie directement l'AdoCbl. La liaison de l'AdoCbl au riborégulateur *btuB* provoque ultimement la formation de la tige séquestratrice du RBS (tige P14). Depuis, la structure du riborégulateur a été obtenue(Perdrizet et al. 2012) et une structure secondaire vraisemblable de celui-ci est disponible en ce qui concerne la forme liée au ligand (voir Figure 6A). On peut remarquer la formation de la tige P1 et du pseudonoeud en présence du ligand. De plus, on remarque la formation de la tige séquestratrice qui se forme par interaction entre l'anti-RBS et le RBS. En ce qui concerne la forme non liée (voir Figure 6B), la structure disponible qui est proposée ne comprend que les régions allant de l'anti-anti-RBS jusqu'au codon d'initiation de la traduction. Les raisons de ce manque d'information sur la forme non liée sont multiples. D'abord, il n'est pas possible d'en obtenir une structure cristallographique puisqu'il est probable que cette forme n'est pas suffisamment stable. Ensuite, la cartographie

chimique et les autres méthodes d'analyse biochimique ne nous permettent pas de tracer une structure claire et fiable. Par contre, les informations disponibles nous montrent que l'anti-anti-RBS et l'anti-RBS interagissent en l'absence d'AdoCbl afin que le RBS soit accessible à la liaison des ribosomes. Une étude récente(Perdrizet et al. 2012) concernant ce riborégulateur a démontré qu'il existe trois sites de pause de la transcription dans la séquence du riborégulateur *btuB* lorsque celui-ci est transcrit à partir de l'ARNpol d'*E. coli*. Ces sites de pause semblent importants pour le repliement de la plateforme d'expression lors de la liaison du ligand. Ces résultats montrent donc que l'utilisation d'une polymérase bactérienne devrait être préconisée, lors de l'étude de l'ARN, afin de tenir compte de l'influence des sites de pauses sur le repliement final du riborégulateur.





**Figure 5. Structures secondaires du riborégulateur *btuB* d'*Escherichia coli*.** Sur la structure liée en A : l'anti-anti-RBS (orange) qui forme une partie de la tige P1, l'antiaptamère (turquoise) contenant le pseudonoeud (PN) et la tige séquestratrice du RBS (rouge) qui est formée par l'interaction de l'anti-RBS et du RBS. Nous remarquons que même si le RBS est séquestré, le codon d'initiation de la traduction reste en principe accessible. La structure non liée en B est incomplète, car seules les sections de l'anti-anti-RBS jusqu'au codon d'initiation de la traduction ont été cartographiées. Nous pouvons malgré tout observer que l'anti-anti-RBS a

comme rôle de lier l'anti-RBS pour empêcher la formation de la tige séquestratrice et de rendre le RBS accessible pour la traduction. Ces structures sont déterminées par des expériences biochimiques,(Perdrizet et al. 2012)(Nahvi et al. 2002) et des prédictions informatiques.(Winkler & Breaker 2003),<sup>46</sup>

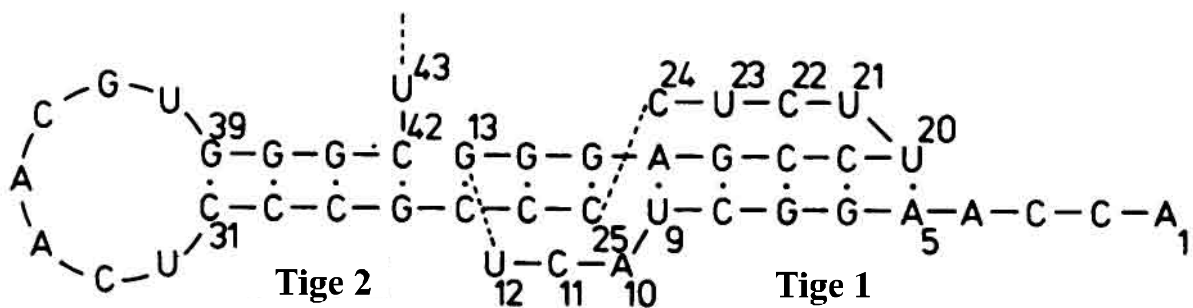
Au niveau structural, ce riborégulateur est l'un des plus complexes connus à ce jour. La complexité de cet ARN peut être en partie justifiée par la taille de son ligand. En effet, l'AdoCbl est une molécule de grande taille et très complexe que l'ARN doit accommoder autant physiquement qu'au niveau des charges. Concernant ce riborégulateur, le pseudonoeud n'a pas encore de rôle clair, mais il semble important pour la communication entre la partie aptamère et la partie plateforme d'expression lors de la liaison du ligand. En effet, lorsque le pseudonoeud est muté, il y a perte de régulation de l'ARN de *btuB* par la liaison du ligand.(Nahvi et al. 2002) Cependant, l'effet de cette mutation sur la liaison du ligand et sur la structure de la plateforme d'expression est méconnu.

## **1.3 Les pseudonoeuds d'ARN**

### **1.3.1 Définitions**

Avec les nouvelles découvertes réalisées sur les rôles fonctionnels de l'ARN, celui-ci ne consiste plus à être seulement un intermédiaire entre l'ADN et la protéine, mais il remplit aussi des rôles régulateurs très puissants. Dans la cellule, la flexibilité(Gautheret & Cedergren 1993) de l'ARN est exploitée afin que ce dernier adopte des structures qui influence l'expression génétique. Parmi les structures d'ARN qui permettent une régulation génétique, il y a le pseudonoeud. Le premier pseudonoeud d'ARN fut identifié par Rietveld, K et al., chez le virus de la mosaïque jaune du navet (voir Figure 7).(Rietveld et al. 1982) Cette structure particulière est présente dans le 3' terminal de l'ARN viral. À ce moment-là, ce nouveau type de structure

était inconnu dans le monde des acides nucléiques et ce n'est que quelques années plus tard que la même équipe introduisit la définition du pseudonoeud.(Pleij et al. 1985) C'est-à-dire une structure tridimensionnelle d'ARN qui forme, via une tige boucle, des paires de base avec une séquence complémentaire en dehors de la boucle. Depuis cette découverte, plus de deux-cents pseudonoeuds ont été répertoriés dans la nature. Aujourd'hui, quelques cristaux de pseudonoeuds sont disponibles dont celui du virus de la mosaïque jaune du navet.(Canady et al. 1996) De plus, avec l'aide de techniques plus sensibles comme la spectroscopie à résonance magnétique nucléaire, de plus en plus de pseudonoeuds sont découverts. Il y a aussi des pseudonoeuds qui participent à des structures d'ARN plus importantes en taille comme les riborégulateurs.

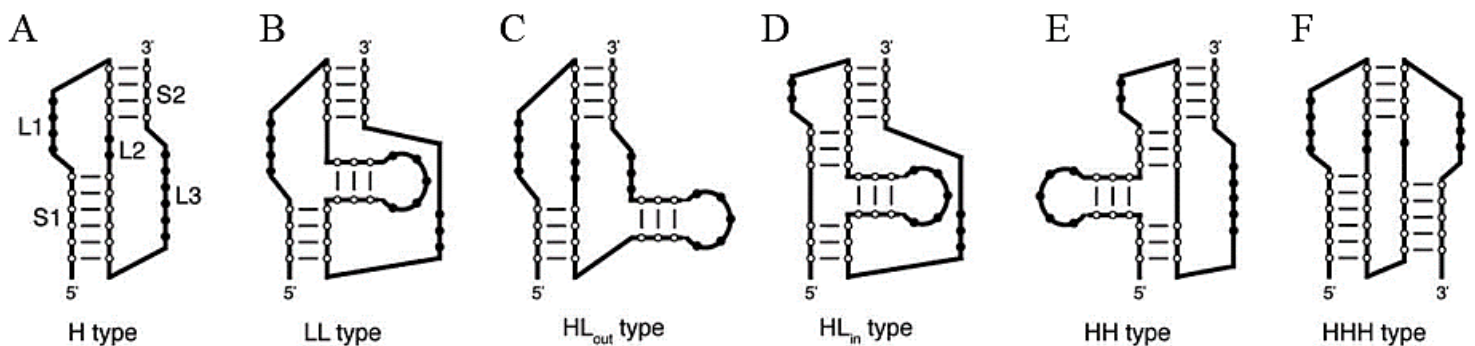


**Figure 6. Structure secondaire du pseudonoeud présent dans le 3' terminal du virus de la mosaïque jaune du navet.** On peut remarquer la formation du pseudonoeud par les interactions entre les bases G13, G14, G15 de la tige 1 et les bases C25, C26, C27 de la tige 2. Tiré de Rietveld, K et al.,(Rietveld et al. 1982)

### 1.3.2 Classifications des pseudonoeuds

Depuis leurs découvertes, plusieurs types de pseudonoeuds ont été identifiés et avec l'accumulation des données il est nécessaire de pouvoir les regrouper en catégorie pour mieux

les comparer.(Han & Byun 2003) Il y a maintenant six types de pseudonoeds et cette caractérisation est basée sur l'environnement structural immédiat du pseudonoed. Tout d'abord, il y a les pseudonoeds sous leur forme la plus simple classée de type H. Ce type de pseudonoed est le plus simple et consiste en une interaction entre l'extrémité d'une tige boucle et l'intérieur d'une tige d'ARN (Figure 7A). Ensuite, si une tige boucle est présente entre le début et la fin de la structure pseudonoed, un L est ajouté dans la dénomination du type de pseudonoed. Alors, selon l'emplacement de la tige boucle, on peut classer le pseudonoed comme étant LL, HL<sub>out</sub> ou HL<sub>in</sub> (Figure 8B, 8C, 8D). Si une tige boucle est présente en amont d'un pseudonoed de type H, on le classe de type HH (Figure 9E). Finalement, si le pseudonoed prend la forme d'une interaction entre l'extrémité de deux tiges boucles, on le classifie HHH (Figure 9F). Le pseudonoed de type H est la forme de pseudonoed la plus répandue dans la nature jusqu'à présent.(Han & Byun 2003)(Kim et al. 2004)



**Figure 7. Classification des pseudonoeds selon leurs arrangements 2D.** Ils sont classés en six groupes (H, LL, HL<sub>out</sub>, HL<sub>in</sub>, HH et HHH) selon leur constitution et selon la présence ou non d'une tige boucle dans le pseudonoed. Les pseudonoeds les plus simples sont de type H. Ensuite, selon la présence d'une tige boucle, on peut classer le pseudonoed comme étant LL, HL<sub>out</sub> ou HL<sub>in</sub> (B, C, D). Si une tige boucle est présente en amont d'un pseudonoed de type H, on le classe HH (E). Le dernier type de pseudonoed prend la forme d'une interaction entre l'extrémité de deux tiges boucles qui est appelé le type HHH (F) Cette classification a été établie par Han, K. & Byun, Y (2003)(Han & Byun 2003). Adaptée de Peselis, A. & Serganov A.(2014)(Peselis & Serganov 2014a)

### 1.3.3 Fonctions des pseudonoeds dans la régulation génétique

Les pseudonoeds sont présents sur plusieurs ARN et ils ont des rôles variés dans la régulation génétique. D'abord, ils peuvent servir à réguler la traduction d'un gène comme dans l'exemple de la protéine gp32 chez les bactériophages de type-T.(McPheeters et al. 1988) Il s'agit d'un mécanisme d'auto-inhibition assez simple, où la protéine gp32 peut lier un pseudonoed situé en amont de la séquence SD du gène gp32. Cette protéine liée à l'ARN est alors utilisée comme point de nucléation pour d'autres protéines qui viennent masquer la séquence SD et il y a répression de la traduction de ce gène.(McPheeters et al. 1988)(Shamoo et al. 1993) Un mode de régulation similaire existe avec les ARN ribosomiaux (ARNr) de certaines bactéries, dont *E. coli*. Durant la production des protéines ribosomiales, certaines de ces protéines, dont S4, S15 et L20, peuvent lier le SD de leur propre ARN messenger.(Tang & Draper 1989) Cette liaison se produit au niveau des pseudonoeds naturellement formés. Elles créent un encombrement qui inhibe l'initiation de la traduction de l'ARNm. Il est intéressant de noter que ces trois pseudonoeds sont modérément distincts. Effectivement, le pseudonoed S15 de type HL<sub>out</sub> n'est constitué que d'environ 60 nucléotides, en comparaison avec le pseudonoed L20 qui est aussi de type HL<sub>out</sub> mais qui est composé de plus de 280 nucléotides. En ce qui concerne le pseudonoed S4, il n'est toujours pas caractérisé.(Serganov et al. 2002)(Haentjens-Sitri et al. 2008)

Comme nous avons vu précédemment, les riborégulateurs sont des structures d'ARN complexes pouvant adopter un grand nombre de conformations différentes. En effet, ces ARN complexes peuvent comporter dans leur structure des pseudonoeds, comme dans le cas du riborégulateur *btuB*, ou bien être constitués principalement d'un seul pseudonoed. Par exemple, les riborégulateurs PreQ<sub>1</sub> sont principalement constitués d'un pseudonoed de type H ou HL<sub>out</sub> selon leur classe.(Kang et al. 2014)(Klein et al. 2009)(Spitale et al. 2009) Ce pseudonoed sera stabilisé lorsque le ligand, preQ<sub>1</sub>, se lie au riborégulateur. Cette stabilisation influence l'expression du gène en abrogeant la transcription de manière prématurée.(Kang et al. 2009) De manière similaire, le riborégulateur SAM-II est aussi constitué principalement d'un

pseudonoëud qui est aussi stabilisé par le ligand. Par contre, dans ce cas, c'est la traduction qui est inhibée par la séquestration de la séquence SD.(Gilbert et al. 2008)

Dans d'autres cas, il existe des riborégulateurs plus volumineux en terme de structure qui peuvent comporter une structure de type pseudonoëud dont le rôle n'est pas toujours défini. Ils ont alors comme rôle probable d'établir des interactions à longue distance dans le but de stabiliser la structure générale de l'ARN. Par exemple, les riborégulateurs lysine semblent être stabilisés globalement à l'aide d'un pseudonoëud.(D. A. Rodionov et al. 2003)

Enfin, il y a aussi le rôle des pseudonoëuds dans les ribozymes. Dans certains ribozymes et introns auto-catalytiques, il y a un pseudonoëud qui possède le rôle de cœur catalytique pour l'auto-clivage de l'ARN (voir Tableau 2). C'est le rôle des pseudonoëuds identifiés chez le riborégulateur-ribozyme *glmS*. Ce riborégulateur-ribozyme comporte trois pseudonoëuds distincts où les pseudonoëuds 1 et 2 forment le cœur catalytique et où le troisième n'a pas de rôles définis à ce jour.(Klein & Ferré-D'Amaré 2006)(Peselis & Serganov 2014a) De plus, d'autres ribozymes possèdent ces structures d'ARN comme les ribozymes HDV(Webb et al. 2009) et des variantes du ribozyme hammerhead.(Perreault et al. 2011) Comme on peut le constater, les pseudonoëuds sont des structures d'ARN répandues et qui possèdent des rôles cruciaux, que ce soit au niveau structural ou bien au niveau catalytique. De plus, comme discuté précédemment, plusieurs riborégulateurs contiennent ou sont constitués de pseudonoëuds qui se trouvent à être essentielle pour le bon fonctionnement du riborégulateur. C'est pour cette raison que l'étude de la fonction des pseudonoëuds présents dans les riborégulateurs devient intéressante. Alors, pour ce projet de maîtrise, nous avons voulu étudier le rôle du pseudonoëud présent chez le riborégulateur *btuB* d'*Escherichia coli* dans le but de déterminer le rôle de ce pseudonoëud dans la communication de la liaison du ligand entre l'aptamère et la plateforme d'expression. Donc, les objectifs de ce projet sont de définir l'importance du pseudonoëud pour le changement de conformation du riborégulateur et de clarifier l'importance de ce pseudonoëud pour la liaison du ligand.

**Tableau 2. Classification par gène des pseudonoeux recensés ayant un rôle déterminé dans la régulation génétique bactérienne et de certains bactériophages.**

Nom du gène	Type de régulation	Type de pseudonoeux
Adénine	Riborégulateur	HHH
Adénosylcobalamine	Riborégulateur	HL <sub>out</sub>
C-di-GMP-II	Riborégulateur	H-like
Désoxyguanosine	Riborégulateur	HHH
Fluoride	Riborégulateur	HL <sub>out</sub>
glmS		
PK1		LL
PK2	Ribozyme- Riborégulateur	HL <sub>out</sub>
PK3		HL <sub>out</sub>
gp32	Site de liaison de protéine régulatrice	H
Guanine	Riborégulateur	HHH
Hammerhead-variant	Ribozyme	H, HL <sub>out</sub>
HDV-variant	Ribozyme	LL, H

**Tableau 2 (suite). Classification par gène des pseudonoceuds recensés ayant un rôle déterminé dans la régulation génétique bactérienne et de certains bactériophages.**

Nom du gène	Type de régulation	Type de pseudonoceud
L20	Site de liaison de protéine régulatrice	HL <sub>out</sub>
Lysine	Riborégulateur	H
PreQ <sub>1</sub> -I	Riborégulateur	H
PreQ <sub>2</sub> -II	Riborégulateur	HL <sub>out</sub>
S15	Site de liaison de protéine régulatrice	HL <sub>out</sub>
S4	Site de liaison de protéine régulatrice	-
SAH	Riborégulateur	LL
SAM-I	Riborégulateur	HL <sub>out</sub>
SAM-II	Riborégulateur	H
SAM-IV	Riborégulateur	HL <sub>out</sub>
SAM-V	Riborégulateur	H
THF	Riborégulateur	HH

Dans ce tableau, l'identification du pseudonoceud est attribuée par le motif d'ARN régulé par ce dernier. Il y a ensuite, le mécanisme par lequel le pseudonoceud régule le gène associé et



finalement le type de pseudonoëud. Adapté de Peselis, A. & Serganov A. (2014).(Peselis & Serganov 2014a)

## CHAPITRE 2

**Titre:** Importance of a Kissing-Loop for btuB Riboswitch Ligand Sensing and Regulatory Control

**Auteurs:** Antony Lussier, Laurène Bastet, Adrien Chauvier et Daniel Lafontaine

Département de Biologie, Faculté des Sciences, Groupe ARN/RNA Group, Université de Sherbrooke, Sherbrooke, Québec, Canada, J1K 2R1

**Participation :** J'ai participé à la rédaction de l'article et j'ai produit la plupart des résultats et des figures de l'article ainsi que la section des méthodes. Laurène Bastet a contribué à l'article en dessinant les graphiques des essais  $\beta$ -galactosidase des figures 2B, D et 7B. De plus, elle a fourni les gels de SHAPE des figures 4A et 4B. Enfin, Adrien a fourni les résultats de ribonucléase T1 de la figure 6.

**Publication :** Il est prévu que cet article sera soumis dans le journal suivant : *The Journal of biological chemistry*.

## 2.1 RÉSUMÉ DE L'ARTICLE

Les riborégulateurs sont des éléments de régulation génétique présents sur l'ARN. Ils sont généralement retrouvés chez les procaryotes, en amont des gènes, dans la région 5' non codante. Les riborégulateurs sont constitués de deux domaines : l'aptamère et la plateforme d'expression. L'aptamère effectue la liaison d'un métabolite appelé ligand au riborégulateur, tandis que la plateforme d'expression régule l'expression génétique en réponse à la liaison du ligand au domaine de l'aptamère. Une des principales caractéristiques des riborégulateurs est leur capacité à adopter une structure alternative lors de la liaison du ligand au riborégulateur. De manière générale, c'est la formation de la tige P1, chez un riborégulateur, qui permet à la plateforme d'expression de modifier sa structure dans le but de réguler l'expression génétique. À l'aide de nos travaux chez le riborégulateur *btuB* d'*E.coli*, nous avons pu déterminer qu'une seconde structure en plus de la tige P1 est nécessaire pour que le riborégulateur effectue la régulation génétique. Cette structure est un pseudonoeud d'ARN se formant lors de la liaison du ligand, à l'image de la tige P1. Par contre, la formation de ces deux structures est nécessaire pour que le riborégulateur effectue une régulation génétique. Ensuite, nous avons déterminé que le pseudonoeud est important pour la liaison du ligand. En effet, l'absence ou la déstabilisation du pseudonoeud provoque une baisse réponse au ligand dans une région du riborégulateur considérée importante pour sa liaison. Enfin, certains résultats nous indiquent que le contexte transcriptionnel est nécessaire pour la formation du pseudonoeud. Nous avons donc caractérisé pour la première fois un rôle double pour une structure de type pseudonoeud chez un riborégulateur. Ces résultats pourront faire avancer l'étude des riborégulateurs comportant un pseudonoeud au sein d'une structure d'ARN complexe comme c'est le cas pour les riborégulateurs lysine.

## 2.2 ABSTRACT

RNA-based genetic regulation is exemplified by metabolite-binding riboswitches that modulate gene expression through conformational changes. Crystal structures show that the *Escherichia coli* *btuB* riboswitch contains a kissing-loop interaction that is in close proximity to the bound ligand. To analyze the role of the kissing-loop interaction in the riboswitch regulatory mechanism, we employed RNase H cleavage assays probing the structure of nascent riboswitch transcripts produced by the *E. coli* RNA polymerase. By interrogating the folding of the aptamer, kissing-loop and riboswitch expression platform, we establish the conformation of each structural component in the absence or presence of bound adenosylcobalamin. We find that the kissing-loop interaction is not essential for ligand binding. However, we show that kissing-loop formation improves ligand binding efficiency and is required to couple ligand binding to riboswitch conformational changes that are involved in regulating gene expression. We also provide evidence showing that the kissing-loop needs to be performed co-transcriptionally, consistent with the importance of the transcriptional context for *btuB* gene regulation. These results implicate that the *btuB* riboswitch modulates the formation of a tertiary structure to perform metabolite sensing and regulate gene expression.

## 2.3 INTRODUCTION

Riboswitches are highly structured regulatory elements that are located in untranslated regions of bacterial mRNAs and that are involved in the control of gene expression by modulating their structure upon metabolite binding (1). These RNA switches exert a regulatory control at different genetic levels such as transcription, translation and mRNA decay (1-3). Riboswitches are composed of two domains consisting of an aptamer implicated in metabolite binding and an expression platform involved in gene expression regulation. While riboswitch aptamer domains are highly conserved to allow specific interaction with cellular metabolites, expression platforms exhibit both sequence and structure diversities to coordinate the state of the aptamer and genetic modulation. The vast majority of riboswitches negatively regulates the expression of downstream gene(s) upon metabolite binding (1). In the case of such transcriptional switches, ligand binding promotes the formation of an anti-antiterminator—the P1 stem—that allows the folding of a terminator structure involved in transcription termination. However, for translationally-controlling riboswitches, P1 stem formation promotes the sequestration of the ribosome binding site (RBS) sequence leading to the inhibition of translation initiation. Thus, for both transcription- and translation-controlling riboswitches, modulation of the P1 stem secondary structure is central for the outcome of genetic expression (3,4).

Some riboswitches have been reported to rely on the modulation of tertiary interactions such as pseudoknot structures to control gene expression (5). For example, the *S*-adenosylmethionine (SAM)-II riboswitch modulates the formation of a pseudoknot interaction that is involved in the sequestration of the RBS sequence when bound to SAM (6,7). However, in the absence of ligand, the RBS is accessible for ribosome binding thus ensuring translation initiation. Another example consists in the adenosylcobalamin (AdoCbl) sensing riboswitch that represents one of the most intricate and complex riboswitch structures including a kissing-loop (KL) interaction (Fig. 1A) (8,9). This riboswitch is involved in the control of the *btuB* gene encoding an outer membrane protein used for B<sub>12</sub> derivatives transport, which is part of the TonB-dependent active transport system (10-12).

Figure 1

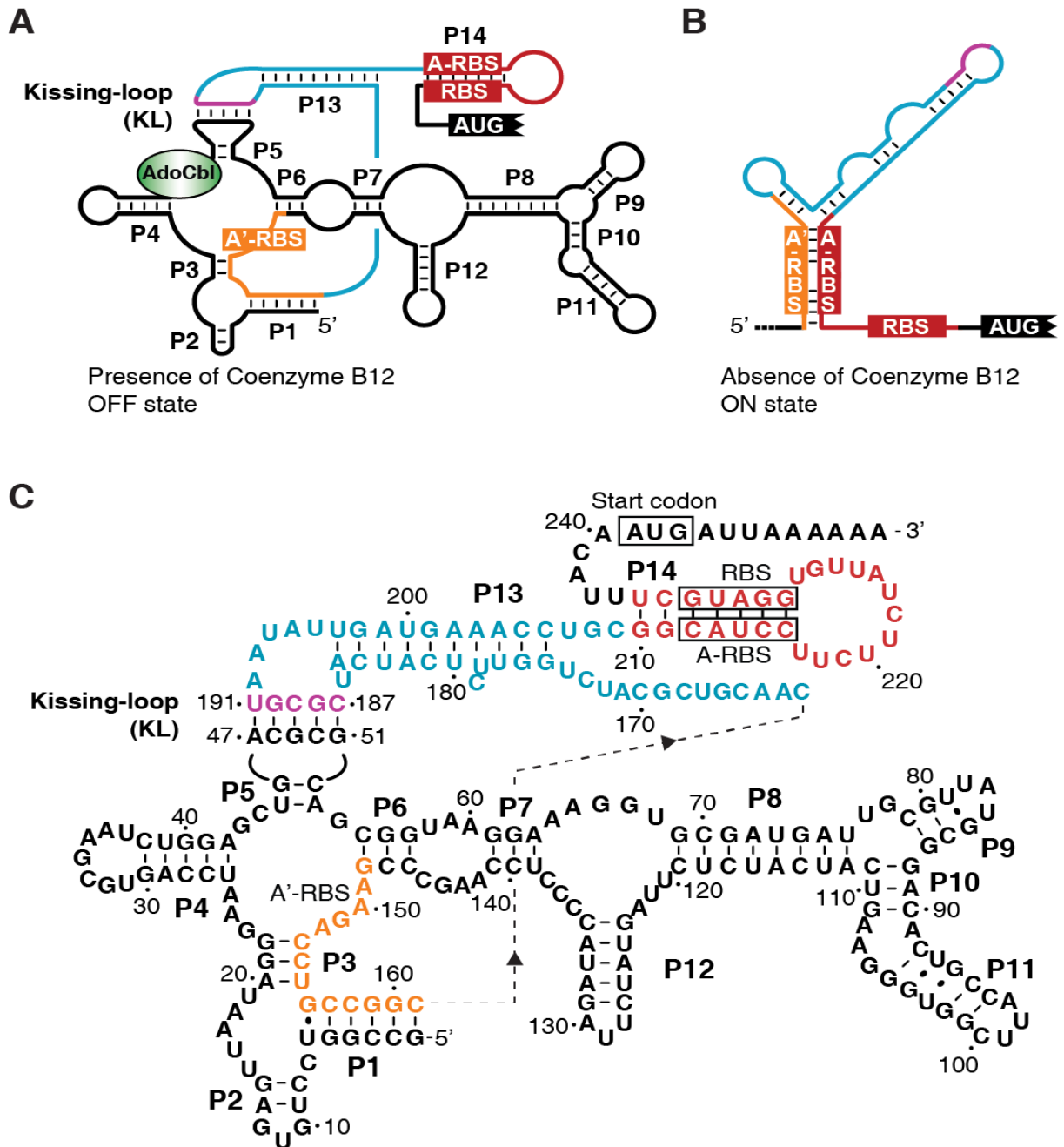


FIGURE 1. *A*, schematic representing the secondary structure of the *btuB* riboswitch when bound to adenosylcobalamin (AdoCbl). The ribosome binding site (RBS) is shown as interacting

with the anti-ribosome binding site (A-RBS) to form the sequestering stem, thereby inhibiting translation initiation (indicated in red). The anti-anti-ribosome binding site (A'-RBS) is shown in yellow. The kissing-loop (KL) tertiary interaction involving P5 and P13 helices is indicated. Paired helical domains (P1-P14) are indicated for each stem and is based according to recent phylogenetic analysis (32). *B*, schematic representing the expression platform in absence of AdoCbl (18). The helical interaction involving the A'-RBS and A-RBS sequences is shown. The RBS sequence is represented as not interacting with the rest of the riboswitch, consistent with the activation of translation in absence of AdoCbl. *C*, detailed representation of the *btuB* riboswitch secondary structure in the ligand bound state (15). The A'-RBS, A-RBS, RBS and start codon are indicated using their respective color scheme as shown in panel *A*. The dashed line represents the phosphodiester backbone connection between residues 161 and 162 and the arrows show the strand polarity.

As observed for the SAM-II riboswitch, AdoCbl binding to *btuB* inhibits translation initiation by precluding ribosome access to the RBS sequence (13). A recent phylogenetic analysis has proposed a switching mechanism for the *Escherichia coli btuB* riboswitch involving anti-RBS (A-RBS) and anti-anti-RBS (A'-RBS) sequences that selectively interact together to allow ribosome binding in absence of AdoCbl (Figs. 1*A* and 1*B*). The *E. coli btuB* aptamer domain is organized around a four-way junction exhibiting highly conserved residues (B<sub>12</sub>-element) that coordinates a highly complex secondary structure composed of twelve helical domains (14,15) and that encompasses the A'-RBS sequence as part of the P1 stem (Fig. 1*A*). The aptamer domain is also involved in the formation of a kissing-loop structure through the interaction of L5 and L13 stem-loop structures (Fig. 1*C*). Mutagenesis showed that residues involved in the kissing-loop interaction are important for ligand binding (10), consistent with the kissing-loop being involved in genetic control. Recent crystal structures revealed that the bound ligand interacts weakly with the kissing-loop ribose-phosphate backbone through van der Waals contacts (8,9). The kissing-loop interaction was also shown to form in absence of ligand at elevated Mg<sup>2+</sup> ions concentrations (9). However, AdoCbl binding to the riboswitch was found to reduce the required Mg<sup>2+</sup> concentration to physiological levels (9,16). Interestingly, an additional interaction

involves the AdoCbl adenosyl moiety and a highly conserved adenine residue located in the A'-RBS region of the four-way junction (8,9), providing a mechanism distinguishing between AdoCbl and related cellular metabolites such as methylcobalamin and aquocobalamin.

Although peripheral extensions (P6-P10 helices) are involved in ligand recognition (8,9,17), there is relatively little information regarding the role of the kissing-loop structure in the folding and ligand-induced genetic regulation of the *btuB* riboswitch. A recent study reported that transcriptional pause sites are present upstream of the kissing-loop (C184), RBS (U225) and AUG sequences (U235) (18). Given that nascent mRNA exiting from paused RNA polymerases coincides with sequences allowing aptamer formation (C184) and kissing-loop interaction (U225), it suggests that transcriptional pausing is important for the folding of both domains (18,19), consistent with the inability of pausing-deficient *E. coli* RNA polymerase to support riboswitch regulation (18). The importance of the kissing-loop interaction is also apparent from recently solved crystal structures showing that although cobalamin riboswitches from different organisms markedly differ in their peripheral regions and ligand binding specificity (8,9), the underlying ligand binding mechanism is common to all of them and rely on a kissing-loop interaction.

In addition to the kissing-loop interaction, the *E. coli btuB* riboswitch contains additional helical domains (P13 and P14) that are directly involved in the selective sequestration of the RBS sequence (Fig. 1A). Despite available crystal structures and biochemical probing data (8-10,16,17), a clear mechanism describing the relationship between the kissing-loop interaction and RBS control still remains to be elucidated. In this study, we have characterized the role of the kissing-loop in ligand-induced conformational changes occurring in both the aptamer domain and riboswitch expression platform. Because the kissing-loop is located at the interface of the aptamer and expression platform, we speculated that it would be a key element coordinating the aptamer bound state to riboswitch conformational changes. Similarly to residues involved in the riboswitch P1 stem, we determined that kissing-loop mutations strongly perturb AdoCbl-dependent gene repression *in vivo*, suggesting that the kissing-loop is a key



element involved in *btuB* regulation. By structurally probing nascent riboswitches, we found that disruption of the kissing-loop structure effectively uncouples ligand binding from conformational changes occurring in the riboswitch expression platform. These results suggest that although ligand binding can be achieved using only the aptamer domain, kissing-loop formation further stabilizes the ligand-bound state of the *btuB* riboswitch. The kissing-loop was found to not be the only factor influencing riboswitch conformational changes since the comparison of truncated *btuB* riboswitches revealed that the stability of the P13 helical domain is also of crucial importance. Lastly, the transcriptional context was found to be important for kissing-loop formation, consistent with previous findings (18). Together, our study suggests that while the kissing-loop is important for ligand binding, it is also implicated as a transmitting module between the aptamer domain and the expression platform to ensure efficient genetic control.

## 2.4 RESULTS

The identity of the P1 stem is crucial for *btuB* riboswitch in vivo regulation. In most riboswitches, the P1 stem is directly involved in controlling gene expression regulation through riboswitch conformational changes (32). We have recently shown for *Bacillus subtilis* riboswitches responding to adenine, lysine and SAM (26,33,34), and for the *E. coli* lysine-sensing riboswitch (31), that ligand-dependent gene regulation is highly dependent on the formation of P1 stem base pairs, but not on the identity of involved residues. Since riboswitch crystal structures show that AdoCbl-sensing is achieved through kissing-loop tertiary interaction rather than using the P1 stem (8,9), we investigated the role of the P1 stem in gene regulation by engineering various riboswitches exhibiting P1 stem mutations designed to disrupt or allow the formation of P1 stem base pairs (Fig. 2A). We constructed BtuB-LacZ translational constructs containing the riboswitch domain with the first 70 *btuB* codons, and assayed riboswitch activity in a defined growth medium in the absence or presence of AdoCbl. The riboswitch construct was fused to a *lacUV5* promoter to avoid any promoter-specific genetic regulation (23). When assaying the  $\beta$ -galactosidase activity of the wild-type riboswitch in the presence of 5  $\mu$ M AdoCbl, the enzymatic activity was decreased by  $\sim$ 15-fold compared to the absence of ligand (Fig. 2B). These results suggest that the *btuB* riboswitch represses gene expression in presence of AdoCbl, in agreement with previous results (10,11,13). To establish the importance of P1 stem formation, we introduced destabilizing P1 stem mutations on either the 5' (M1.1) or 3' side (M1.2) of the P1 stem (Fig. 2A). The presence of AdoCbl completely failed to reduce  $\beta$ -galactosidase activity in both cases (Fig. 2B), indicating that the integrity of the P1 stem is crucial for genetic regulation. The simultaneous introduction of M1.1/M1.2 complementary mutations allowed a partial AdoCbl regulation ( $\sim$ 1.6-fold) (Fig. 2B), in agreement with the genetic control being perturbed compared to the wild-type construct. Given that the P1 stem contains three base pairs exhibiting strict phylogenetic conservation at the sequence level (32), we engineered a second set of riboswitch mutants in which only non-conserved base pairs were altered (M1.3 and M1.4) (Fig. 2A). Similarly to what observed for M1 and M2 mutants, a complete loss of ligand-dependent genetic regulation was also obtained (Fig. 2B). Moreover, a construct combining both complementary mutations (M1.3/M1.4)

resulted in a ligand-induced regulation of ~2.9-fold (Fig. 2B), indicating that regulation is not as efficient as in the wild-type (~15-fold). Together, our results suggest that the identity of the P1 stem is important for the AdoCbl-induced *btuB* riboswitch regulation. Furthermore, since we observed low  $\beta$ -galactosidase activities for all tested mutants, it suggests that the P1 stem sequence is important for *btuB* gene expression in conditions of low AdoCbl abundance.

The kissing-loop interaction is required for *btuB* riboswitch control. According to cobalamin riboswitch crystal structures (8,9), the kissing-loop interaction is in proximity to the bound metabolite, suggesting that the interaction is important for gene regulation. Previous work has shown that mutation of the kissing-loop interaction resulted in a complete loss of AdoCbl-dependent control and that complementary mutations only partially restored regulation (10), all of which suggesting that the identity of kissing-loop residues is important. However, since crystal structures do not show extensive interaction between the bound ligand and the kissing-loop (8,9), it indicates that the identity of residues is not crucial for riboswitch function. Thus, to determine the role of kissing-loop residues in genetic regulation, we introduced several mutations in the kissing-loop interaction and assessed their effect on riboswitch regulation using  $\beta$ -galactosidase assays. When introducing C187G/G188C mutations inhibiting the formation of two kissing-loop base pairs (K1 mutant, Fig. 2C), high  $\beta$ -galactosidase activities were observed both in absence and presence of AdoCbl (Fig. 2D), indicating that the *btuB* riboswitch control is strongly perturbed. The introduction of compensatory mutations restoring the formation of both base pairs (K1c, Fig. 2C) resulted in a significant decreased *btuB* expression regardless of the presence of ligand (Fig. 2D), consistent with previous results (18). A small degree of ligand-induced regulation (~1.7-fold) was obtained for the K1c mutant, showing that the regulation is altered compared to the wild-type (Fig. 2D). A similar analysis was also performed for other base pairs of the kissing-loop (K2 and K3 mutants, Fig. 2C). In each case, the presence of AdoCbl failed to modulate the  $\beta$ -galactosidase activity (Fig. 2D), consistent with these base pairs also being important for the *btuB* riboswitch regulation. Compensatory mutations restored riboswitch control only to a limited extent since AdoCbl-dependent repression of ~1.2-fold and ~1.9-fold were obtained for K2c and K3c mutants, respectively (Fig. 2D). Thus, our results

suggest that kissing-loop residues are important for riboswitch gene regulation. However, in contrast to the P1 stem (Fig. 2B), kissing-loop mutations did not result in very low  $\beta$ -galactosidase activities ( $< 0.25$  unit) in absence of ligand, suggesting that both structures are differently involved in *btuB* regulation.

**Figure 2**

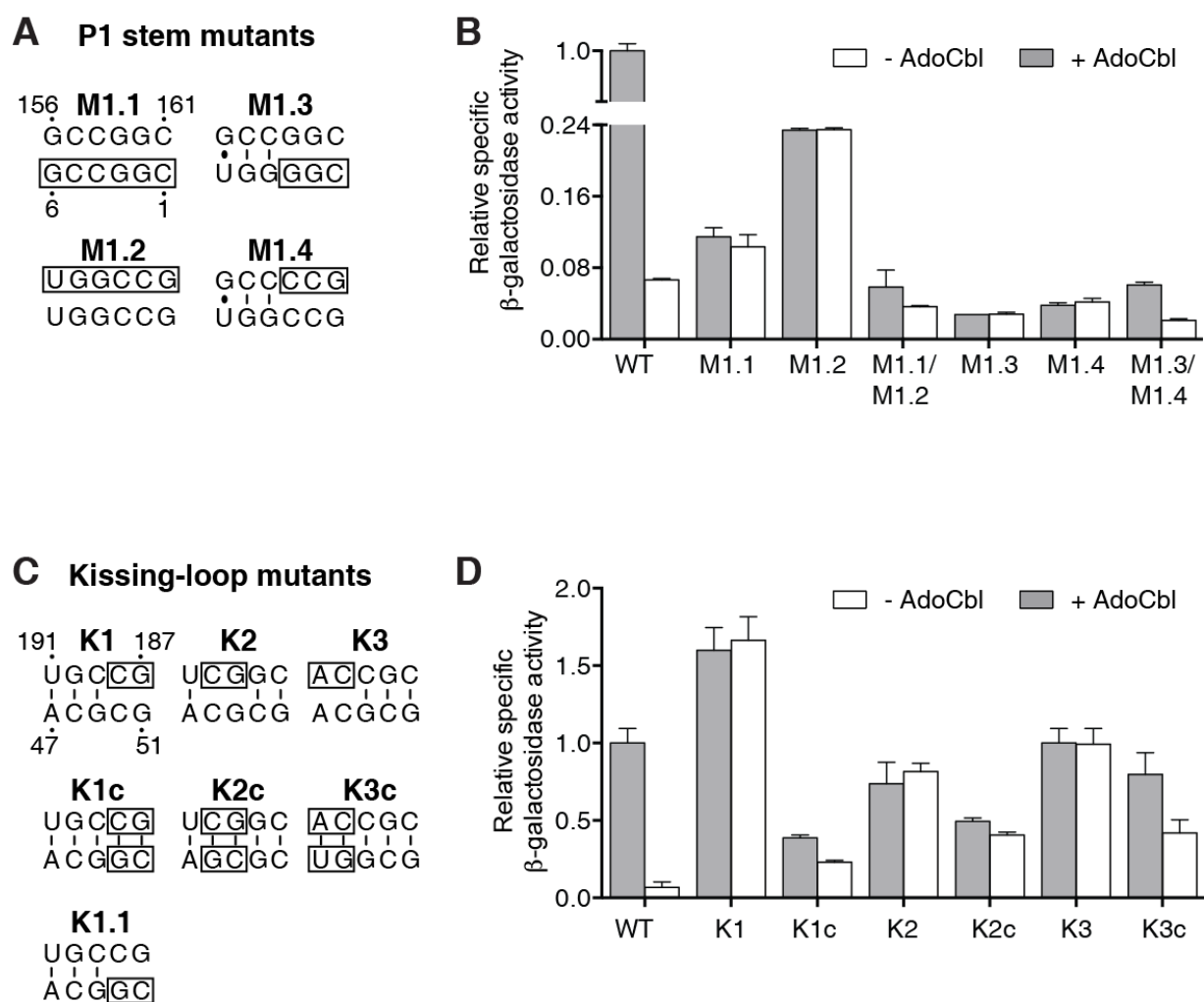


FIGURE 2. A, representations of characterized P1 stem mutants. Boxes indicate mutated regions corresponding to the Watson-Crick complement of the wild-type sequence. The wild-type P1

stem is shown in Fig. 1C. *B*,  $\beta$ -galactosidase assays of translational BtuB-LacZ fusions for the wild-type and P1 stem mutants. Enzymatic activities were measured in the absence and presence of 5  $\mu$ M AdoCbl. Values are normalized to the enzyme activity obtained for the wild-type and the absence of ligand. The average values of three independent experiments with SDs are shown.

*C*, representation of kissing-loop mutants. Mutations are boxed and correspond to the Watson-Crick complement of the wild-type kissing-loop interaction (shown in Fig. 1C). *D*,  $\beta$ -galactosidase assays of translational BtuB-LacZ fusions for the wild-type and kissing-loop mutants. Enzymatic activities were measured in the absence and presence of 5  $\mu$ M AdoCbl. Values are normalized to the enzyme activity obtained for the wild-type and the absence of ligand. The average values of three independent experiments with SDs are shown.

SHAPE chemical probing of the *btuB* riboswitch. To gain insight into AdoCbl-induced *btuB* structural changes, we performed SHAPE analysis using the electrophile reagent *N*-methylisatoic anhydride (NMIA), which reacts with 2'-OH groups located in RNA flexible regions (35). In this assay, purified riboswitch transcripts were renatured and analyzed for their ability to undergo structural changes in presence of ligand. SHAPE analysis showed that increasing AdoCbl concentrations resulted in several *btuB* riboswitch conformational changes (Fig. 3A), suggesting that ligand binding is attained in our experimental conditions. As previously reported (9), AdoCbl-dependent protections were observed in regions corresponding to stems P10-P13 and single strand regions J1/13, J6/7 and J7/8 (Fig. 3A). Furthermore, ligand-dependent protections were also obtained in the kissing-loop region (C50 residue, Fig. 3A and 3B), consistent with kissing-loop formation upon AdoCbl binding. However, given that the folding of the *btuB* riboswitch expression platform was recently reported to be highly dependent on the transcriptional context, (18), we sought to investigate AdoCbl-induced riboswitch conformational changes on nascent transcripts.

**Figure 3**

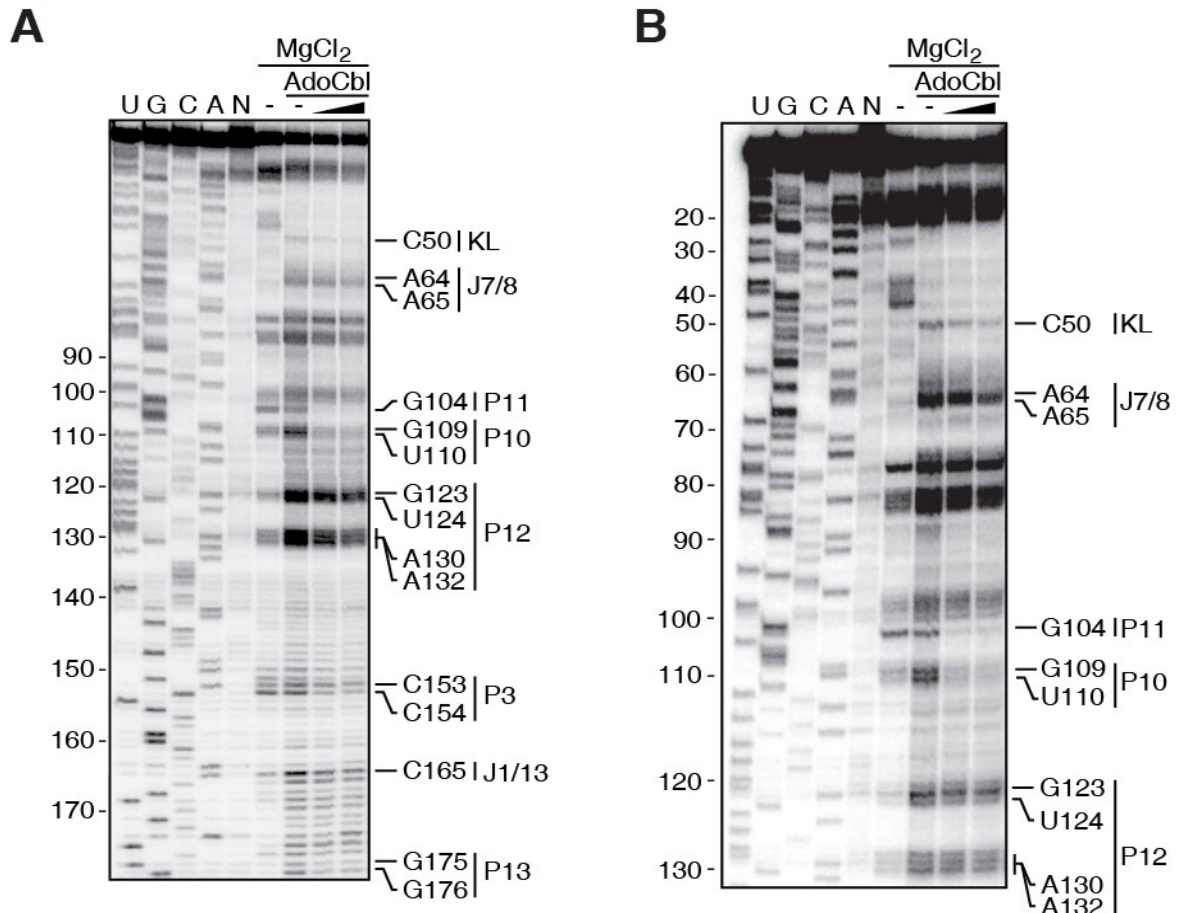


FIGURE 3. *A* and *B*, SHAPE analysis of the *btuB* riboswitch performed in the absence (-Mg) or presence of 10 mM Mg<sup>2+</sup> ions as a function of AdoCbl concentration (20 and 50 μM). Nucleotides of the *btuB* riboswitch exhibiting AdoCbl-induced modifications are indicated on the right of the gel. SHAPE modifications were detected through reverse transcription reactions using DNA oligonucleotides targeting regions 1-202 nucleotides (*A* and *B*). Reactions performed in presence of DMSO (N) and sequencing lanes (U, G, C or A) are indicated.

Probing the *btuB* riboswitch native structure using RNase H assays. The RNase H cleavage assay is a well-established technique whereby a DNA oligonucleotide is used to target an RNA sequence that is cleaved upon RNase H recognition (18,36,37). The native structure of the *btuB* riboswitch was previously investigated using a combination of single-round transcription and RNase H assays (18). In that study, nascent *btuB* riboswitches transcribed using the *E. coli* RNA polymerase were probed using a series of DNA oligonucleotides covering the entire riboswitch (18). Thus, to characterize the influence of the kissing-loop interaction on *btuB* riboswitch structural changes, we employed a similar approach in which three DNA oligonucleotides were used to target the aptamer, kissing-loop and P13 stem (Fig. 4A and Table 3). All oligonucleotides were selected based on their ability to report riboswitch conformational changes upon AdoCbl binding, according to SHAPE (Fig. 3) and RNase H data (18).

Single-round transcription assays were performed using a DNA construct comprising the *lacUV5* promoter and the *btuB* riboswitch fused to the first 70 *btuB* codons. In absence of RNase H, transcription reactions yielded a major product corresponding to the full-length transcript (FL) (see lane N, Fig. 4B). However, when RNase H cleavage assays were performed using a DNA probe targeting the aptamer domain, nascent riboswitch RNAs were cleaved at low AdoCbl concentrations and generated a specific cleavage product (see product P, Fig. 4B). However, when experiments were performed over a range of AdoCbl concentrations, a protection from RNase H cleavage was detected (Fig. 4B), consistent with the formation of an AdoCbl-riboswitch complex being resistant to RNase H cleavage, as previously observed (18). As a result, we have quantified RNase H cleavage products and fitted data to a two-state binding model (26), thereby allowing to determine the AdoCbl concentration at which half of the conformational switch is taking place ( $K_{\text{switch}}$ ). By fitting our data, we determined a value of  $53 \pm 4$  nM, which is similar to that previously reported ( $\sim 93$  nM) (18). We obtained a very similar value when using a concentration of the DNA probe 2.5 times higher ( $K_{\text{switch}} = 58 \pm 7$  nM, Fig. 4C), indicating that the probe is used at saturation. The  $K_{\text{switch}}$  value agrees well with the dissociation constant of the AdoCbl-aptamer complex determined using in-line probing ( $\sim 300$

nM) (10), indicating that our RNase H cleavage assay provides a sensitive tool to study ligand-induced *btuB* riboswitch conformational changes in a transcriptional context.

To establish the effect of AdoCbl binding on the formation of the kissing-loop interaction, we performed RNase H cleavage assays using a DNA probe targeting the kissing-loop structure (KL probe, Fig. 4A). As observed for the aptamer probe, we detected an efficient RNase H cleavage activity in low AdoCbl concentrations, which was progressively reduced when transcriptions were performed at higher ligand concentrations (Fig. 4D). Fitting analysis yielded a  $K_{\text{switch}}$  value of  $115 \pm 4$  nM that corresponds to a higher value than what was obtained using the aptamer probe ( $53 \pm 4$  nM). Such variations are expected between  $K_{\text{switch}}$  values since they are most likely affected by the ability of a given DNA probe to hybridize to its RNA sequence, which can markedly differ due to the stability of targeted regions. Interestingly, the ligand-dependent protection of the kissing-loop structure suggests that ligand binding stabilizes the formation of the kissing-loop, consistent with crystal structures showing that the bound AdoCbl is closely located to the kissing-loop structure (8,9). We also investigated the riboswitch expression platform by targeting the P13 stem (P13 probe, Fig. 4A). As observed for both aptamer and the kissing-loop probes, RNase H cleavage assays showed that the P13 helical domain is protected in presence of AdoCbl ( $K_{\text{switch}} = 224 \pm 6$  nM, Fig. 4E), consistent with the P13 stem being more stable in the context of the AdoCbl-riboswitch complex. Together, our results suggest that structural changes are observed in three different regions of the *btuB* riboswitch in presence of AdoCbl, consistent with the formation of a more stable structure resistant to RNase H cleavage activity. Furthermore, even though  $K_{\text{switch}}$  values show variations between DNA oligonucleotides used, they nevertheless provide a powerful way to probe specific regions of the *btuB* riboswitch, thus allowing characterization of ligand-induced conformational changes of nascent riboswitch transcripts.



Figure 4

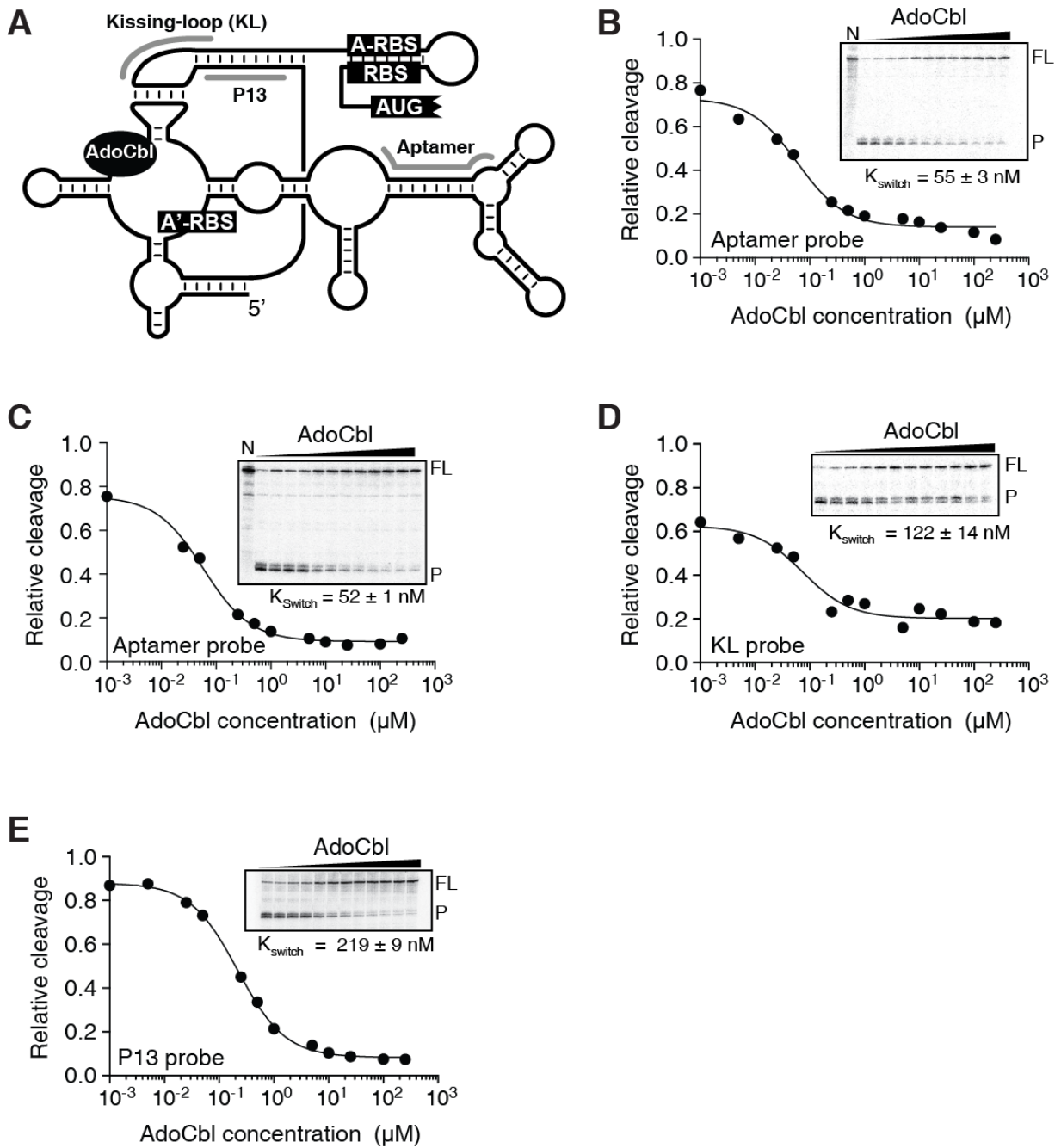


FIGURE 4. *A*, schematic representing targeted positions of DNA oligonucleotides (shown in gray) used to probe the aptamer, the kissing-loop and the P13 helical domain. The riboswitch is

shown as when bound to AdoCbl. Probes were selected on their ability to report on AdoCbl binding (18). *B*,  $K_{\text{switch}}$  determination using 20  $\mu\text{M}$  aptamer probe in the context of the wild-type riboswitch. *In vitro* transcriptions were performed in presence of increasing AdoCbl concentrations ranging from 1 nM to 250  $\mu\text{M}$ . RNase H cleavage assays show that a specific cleavage product (P) is obtained in low AdoCbl concentrations that is progressively reduced at high concentrations, thus yielding uncleaved full-length riboswitch molecules (FL). Transcriptions performed in absence of RNase H show only the FL species (lane N). A plot depicting normalized fractions of cleaved RNA *versus* AdoCbl concentrations yields a  $K_{\text{switch}}$  value of  $53.1 \pm 4$  nM, which is in good agreement with previous in-line probing analysis (10,15). Experiments were done in triplicate and reported data on the graph correspond to a representative result. *C*,  $K_{\text{switch}}$  determination using a higher concentration of the aptamer probe (50  $\mu\text{M}$ ) in the wild-type riboswitch. Transcriptions were performed as a function of AdoCbl concentration ranging from 1 nM to 250  $\mu\text{M}$ . The cleavage product (P) is decreased, thus yielding the full-length transcript (FL), upon increasing AdoCbl concentration. Absence of RNase H results in only the production of the FL product (lane N). A plot depicting normalized fractions of cleaved RNA *versus* AdoCbl concentrations yields a  $K_{\text{switch}}$  value of  $58 \pm 7$  nM. Experiments were done in triplicate and reported data on the graph correspond to a representative result. *D* and *E*,  $K_{\text{switch}}$  determinations using 20  $\mu\text{M}$  kissing-loop (*D*) and P13 stem (*E*) probes in the wild-type riboswitch. Transcriptions were performed as a function of AdoCbl concentration ranging from 1 nM to 250  $\mu\text{M}$ . In presence of low AdoCbl concentrations, a cleavage product (P) is observed and decreased at high ligand concentrations, thus yielding the full-length transcript (FL). Plots depicting normalized fractions of cleaved RNA *versus* AdoCbl concentrations yield  $K_{\text{switch}}$  value of  $115 \pm 4$  nM for the kissing-loop probe (*D*) and  $224 \pm 6$  nM for the P13 stem probe (*E*). Experiments were done in triplicate and reported data on the graph correspond to a representative result.

The kissing-loop is crucial to coordinate ligand binding and riboswitch conformational changes. According to *in vivo* data obtained using translational BtuB-LacZ constructs (Fig. 2D)(10), the K1 kissing-loop mutant strongly perturbs the ligand-dependent riboswitch regulatory control. To characterize the role of the kissing-loop structure in *btuB* riboswitch regulation, we investigated the structure of the K1 riboswitch mutant and its ability to perform ligand-induced structural changes using RNase H cleavage assays. When using the aptamer probe, we determined that AdoCbl-dependent protection could be achieved only at high ligand concentrations (Fig. 5A), yielding a  $K_{\text{switch}}$  value of  $12000 \pm 3000$  nM. This high value suggests that although the destabilization of the kissing-loop interaction clearly affects ligand binding activity, the K1 mutant riboswitch can still bind AdoCbl at high concentration. We also performed RNase H assays using a riboswitch variant containing mutations C50G/G51C, which destabilize the same base pairs as in the K1 mutant but through the corresponding pairing partners (K1.1 mutant, Fig. 2C). A similar result to the K1 riboswitch was obtained ( $K_{\text{switch}} = 10700 \pm 1783$  nM), consistent with the importance of these residues for AdoCbl binding. However, when both K1 and K1.1 mutations were simultaneously introduced to allow base pair formation (K1c mutant, Fig. 2C), the ligand-induced protection to RNase H cleavage was observed at significantly lower AdoCbl concentrations (Fig. 5A). The  $K_{\text{switch}}$  value ( $368 \pm 32$  nM) was found to be only ~7-fold higher than the wild-type, indicating that kissing-loop base pairs C50-G188 and G51-C187 are important to form a stable AdoCbl-riboswitch complex.

The characterization of kissing-loop riboswitch mutants was also performed using the probe targeting the kissing-loop structure. To assess the formation of the kissing-loop in the context of the K1 mutant, we employed a different DNA probe (KLc probe, Table 3) matching introduced mutations in the L13 loop. When RNase H cleavage assays were done using the K1 mutant, the relative cleavage was not modulated by the presence of AdoCbl, even at high concentrations (Fig. 5B). These findings suggest that the kissing-loop interaction is not stabilized in the context of the K1 mutant. However, when using the K1c riboswitch variant in which the kissing-loop is allowed to form, ligand-dependent protection of RNase H cleavage activity was recovered and exhibited a  $K_{\text{switch}}$  value of  $261 \pm 85$  nM (Fig. 5B), directly showing

the formation of the kissing-loop interaction in presence of compatible Watson-Crick pairing residues.

We also performed RNase H cleavage assays using the P13 probe to characterize the influence of the kissing-loop on conformational changes taking place in the expression platform. As observed for the kissing-loop probe, AdoCbl titration completely failed to reduce RNase H cleavage activity in the context of the K1 mutant (Fig. 5C), showing the inability of the expression platform to undergo AdoCbl-dependent structural changes. However, a significant decrease in RNase H cleavage was detected using the K1c mutant ( $K_{\text{switch}} = 356 \pm 37$  nM, Fig. 5C), consistent with a functional kissing-loop structure being important to induce conformational changes in the expression platform. Together, these results clearly show that the presence of the kissing-loop is very important, but not vital, for the *btuB* riboswitch to bind AdoCbl in a transcriptional context. However, our data also indicate that the kissing-loop structure is crucial in coordinating the aptamer ligand binding activity to *btuB* riboswitch conformational changes taking place in the expression platform.

**Figure 5**

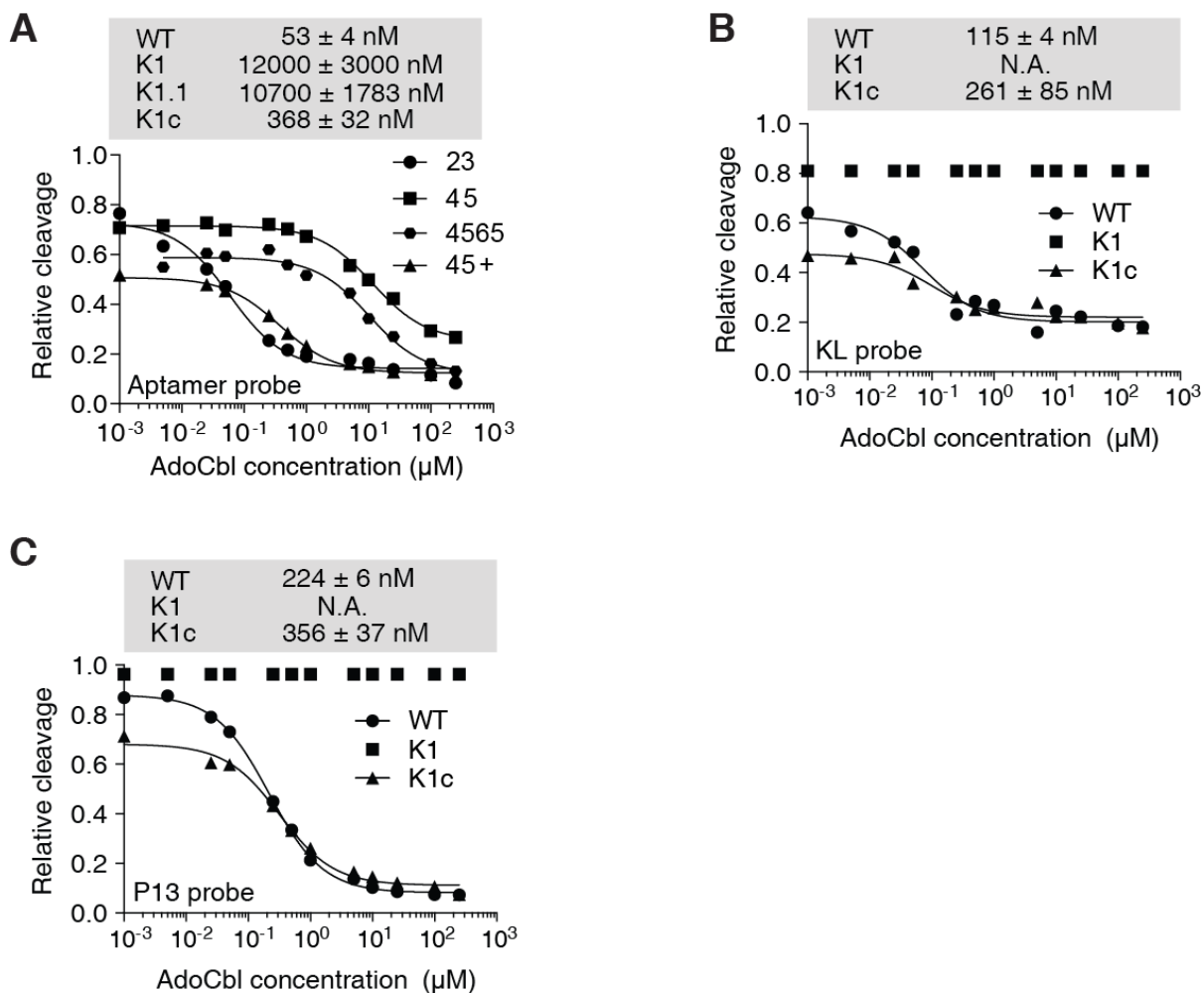


FIGURE 5. *A*,  $K_{\text{switch}}$  determinations of the wild-type, K1, K1.1 and K1c riboswitch mutants. In each case, transcription reactions were performed as a function of AdoCbl concentrations (1 nM to 250  $\mu$ M) and RNase H cleavage assays were done using the aptamer probe at 20  $\mu$ M. The normalized fractions of cleaved riboswitches species *versus* AdoCbl concentrations are shown for each analyzed riboswitch construct. Fitting analysis reveals  $K_{\text{switch}}$  values for the wild-type ( $53 \pm 4$  nM), K1 mutant ( $12000 \pm 3000$  nM), K1.1 mutant ( $10700 \pm 1783$  nM) and K1c mutant ( $368 \pm 32$  nM). Data of the wild-type riboswitch are taken from Fig. 3*B* and are represented for comparison purpose. Experiments were done in triplicate and reported data on the graph correspond to a representative result. *B* and *C*,  $K_{\text{switch}}$  determinations of the wild-type, K1 and

K1c riboswitch mutants. Transcription reactions were performed as a function of AdoCbl concentrations (1 nM to 250  $\mu$ M) and RNase H cleavage assays were done using the kissing-loop (*B*) or P13 stem probe (*C*) at 20  $\mu$ M. Normalized fractions of cleaved riboswitch species *versus* AdoCbl concentrations are shown for each analyzed riboswitch construct. Fitting analysis using the kissing-loop probe reveals  $K_{\text{switch}}$  values for the wild-type ( $115 \pm 4$  nM) and K1c mutant ( $261 \pm 85$  nM). A similar fitting analysis using the P13 stem probe data yield  $K_{\text{switch}}$  values for the wild-type ( $224 \pm 6$  nM) and K1c mutant ( $356 \pm 37$  nM). Note that the absence of significant variation does not allow a  $K_{\text{switch}}$  determination for the K1 mutant when using kissing-loop (*B*) or P13 stem (*C*) probe. Data of the wild-type riboswitch for the kissing-loop (*B*) and P13 stem (*C*) probes are taken from Fig. 3*D* and Fig. 3*E*, respectively, and are represented for comparison purpose. Experiments were done in triplicate and reported data on the graph correspond to a representative result.

Probing the *btuB* riboswitch native structure using RNase T1 assays. To seek additional information about AdoCbl-induced structural changes occurring in *btuB* nascent transcripts, we employed a biotin-based solid-phase transcription approach to probe riboswitches using partial RNase T1 digestion (38). By performing the assays in absence or in presence of AdoCbl, several ligand-dependent changes were detected throughout the riboswitch structure, notably in stems P8-P11 and in regions J1/13 and J3/6 (Fig. 6*A*). RNase T1 protection was also observed for residues G188 and G190 that are involved in the kissing-loop (Fig. 6*A*). No structural change could be observed in the A-RBS or RBS region. By monitoring the accessibility of the J1/13 region over a range of ligand concentration, a  $K_{\text{switch}}$  of  $220 \pm 59$  nM was obtained (Fig. 6*B*), which is similar to values determined using RNase H probes targeting the kissing-loop or P13 structures (Fig. 4*D* and 4*E*). No ligand-induced protection was observed for the K1 mutant (Fig. 6*A*). Notably, the high accessibility of kissing-loop residues G188 and G190 in the K1 mutant is in agreement with the high RNase H cleavage activity obtained using the KL probe (Fig. 5*B*). Thus, these results are in good agreement with RNase H data showing that AdoCbl promotes the stabilization of the aptamer and kissing-loop structures in the wild-type riboswitch.

**Figure 6**

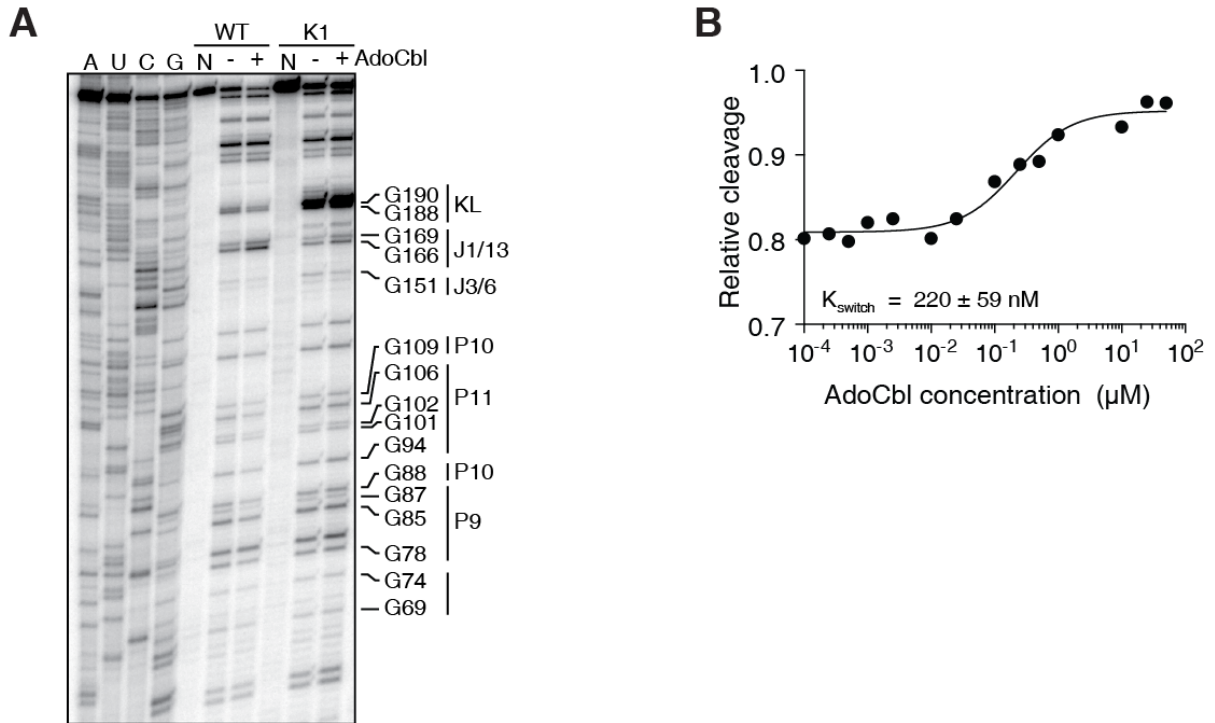


FIGURE 6. *A*, partial RNase T1 digestion analysis of the *btuB* riboswitch performed in absence (-) or presence of 100  $\mu$ M AdoCbl. Reactions were performed with wild-type (WT) and K1 riboswitch mutant (K1). Nucleotides of the *btuB* riboswitch exhibiting AdoCbl-induced modifications are indicated on the right of the gel. Sequencing lanes (U, G, C or A) are indicated. *B*,  $K_{\text{switch}}$  determination of the wild-type riboswitch using RNase T1 digestion was performed as a function of AdoCbl concentration (10 nM to 250  $\mu$ M). The normalized fraction of cleavage for J1/13 is shown for each concentration. Fitting analysis yields a  $K_{\text{switch}}$  value of  $92 \pm 42$  nM).

The stability of the AdoCbl-riboswitch complex is improved through the kissing-loop interaction. It was recently reported that RNA polymerase pause sites are located in *btuB* riboswitch expression platform (18). Among these, a pause site was mapped in the P13 stem at position C184 (Fig. 1C). Taking into account that ~12 nucleotides are present in the RNA polymerase exit channel (38,39), it suggests that transcriptional complexes paused at C184 would allow ~172 nucleotides of the *btuB* riboswitch to fold, therefore allowing aptamer domain formation. However, in such a case, the kissing-loop interaction is not formed since it has not been yet transcribed, thereby potentially precluding ligand binding. To investigate whether the aptamer domain can perform AdoCbl binding, we engineered a truncated version of the *btuB* riboswitch containing only the first 161 nucleotides, effectively corresponding to the aptamer domain (Fig. 7A). Using such a construct, we performed RNase H cleavage assays as a function of AdoCbl concentration (Fig. 7B). Protection from RNase H cleavage was found to occur only at high ligand concentrations with the aptamer probe, indicating that the aptamer domain can perform ligand binding although with reduced efficiency ( $K_{\text{switch}} = 62100 \pm 6000$  nM). This finding is consistent with the importance of the kissing-loop for AdoCbl binding (Fig. 5A).

Given that an additional pause site was also found at position U225 (18), it suggests that nascent *btuB* riboswitches in such paused transcription complexes exhibit ~213 nucleotides, allowing to form both the kissing-loop and P13 helical domain. We thus engineered a transcription construct in which the riboswitch domain was truncated at position 202 and assessed the ability to perform AdoCbl sensing using RNase H assays. This particular RNA size was chosen since it is typically used in in-line probing experiments to study the binding features of the *btuB* aptamer (10,15-17). The ligand-induced stabilization of the AdoCbl-riboswitch complex was significantly improved in the context of the 202 nt construct as compared to the shorter 161 nt construct ( $K_{\text{switch}} = 232 \pm 25$  nM, Fig. 7B). Since the *btuB* secondary structure model suggests that the P13 stem contains at least four more base pairs (Fig. 1C), we also used a construct truncated at position 206 enabling the complete formation of the predicted P13 stem. RNase H cleavage assays revealed that such *btuB* riboswitch transcripts exhibit improved ligand-dependent



protection from RNase H cleavage ( $K_{\text{switch}} = 115 \pm 31$  nM, Fig. 7B), showing that the stability of the P13 helix is important for the formation of a stable riboswitch-ligand complex.

The influence of P13 stability on the AdoCbl-dependent formation of kissing-loop was assayed using kissing-loop and P13 stem DNA probes (Figs. 7C and 7D). In these experiments, the 161 nt construct was not tested due to the absence of the kissing-loop interaction. RNase H assays performed on the 202 and 206 nt constructs showed that the ligand-induced formation of the kissing-loop structure is highly dependent on the stability of the P13 helical domain given that  $K_{\text{switch}}$  values of  $3800 \pm 250$  nM and  $109 \pm 41$  nM were obtained for the 202 and 206 nt constructs, respectively (Fig. 7C). As expected, we also found that the stability of the P13 stem is important for AdoCbl to induce conformational changes in the expression platform since RNase H cleavage assays using the P13 probe showed that the 202 and 206 nt constructs yielded  $K_{\text{switch}}$  values of  $1700 \pm 200$  nM and  $404 \pm 30$  nM, respectively (Fig. 7D). Taken together, our results suggest that although ligand binding can be achieved at the first transcriptional pause site (C184), the stabilization of the RNA-ligand complex is strongly increased at the second pause site (U225) when the P13 stem is transcribed.

**Figure 7**

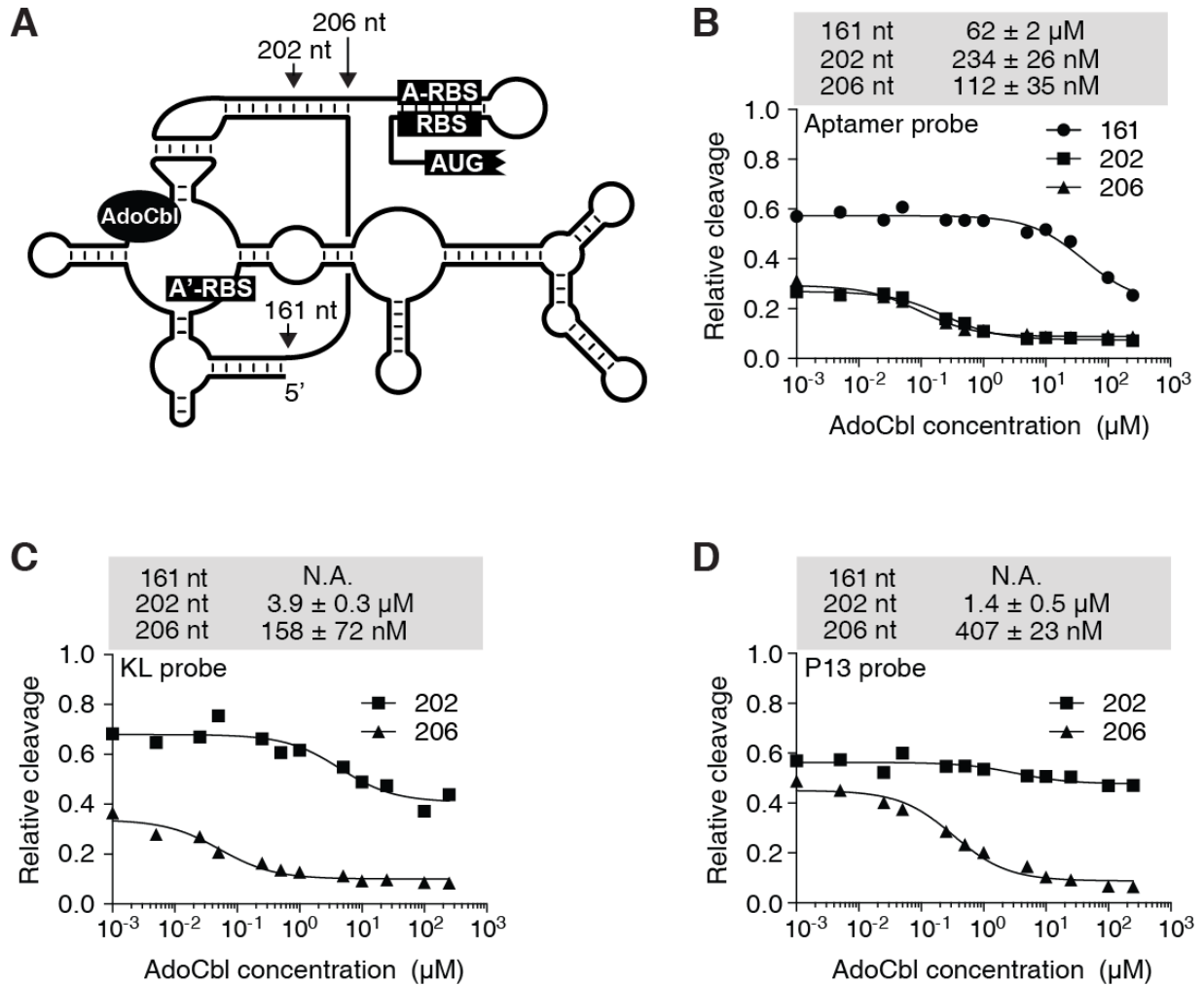


FIGURE 7. *A*, schematic representing riboswitch molecular sizes of 161, 202 and 206 nt used for RNase H assays. The riboswitch is shown as when bound to AdoCbl. *B*,  $K_{\text{switch}}$  determinations of the 161, 202 and 206 nt riboswitch constructs. In each case, transcription reactions were performed as a function of AdoCbl concentrations (1 nM to 250  $\mu\text{M}$ ) and RNase H cleavage assays were done using the aptamer probe at 20  $\mu\text{M}$ . The normalized fractions of cleaved riboswitches species *versus* AdoCbl concentrations are shown for each analyzed riboswitch construct. Fitting analysis yields  $K_{\text{switch}}$  values for constructs containing 161 nt

( $62100 \pm 5000$  nM), 202 nt ( $232 \pm 25$  nM) and 206 nt ( $115 \pm 31$  nM). Experiments were done in triplicate and reported data on the graph correspond to a representative result. *C* and *D*,  $K_{\text{switch}}$  determinations of the 161, 202 and 206 nt riboswitch constructs using kissing-loop (*C*) and P13 stem (*D*) probes. Transcription reactions were performed as a function of AdoCbl concentrations (1 nM to 250  $\mu$ M) and RNase H cleavage assays were done using the kissing-loop (*C*) or P13 stem probe (*D*) at 20  $\mu$ M. Normalized fractions of cleaved riboswitch species *versus* AdoCbl concentrations are shown for each analyzed riboswitch construct. Fitting analysis using the kissing-loop probe (*C*) reveals  $K_{\text{switch}}$  values for constructs containing 202 nt ( $3800 \pm 250$  nM) and 206 nt ( $56 \pm 51$  nM). A similar fitting analysis using the P13 stem probe data (*D*) yield  $K_{\text{switch}}$  values for constructs containing 202 nt ( $1700 \pm 200$  nM) and 202 nt ( $404 \pm 30$  nM). Note that the aptamer probe was not used to target constructs containing 202 and 206 nt since the binding site of the DNA oligonucleotide is not presence in these constructs. Experiments were done in triplicate and reported data on the graph correspond to a representative result.

The P13 stem is important for the adoption of the ligand-free conformation. Based on predicted *btuB* structures (Fig. 8A)(18), ligand binding to the riboswitch rearranges the P13 stem to allow the kissing-loop interaction. The A'-RBS sequence appears to be the dominant feature to control gene expression by selectively sequestering the RBS in presence of AdoCbl (Fig. 8A). However, since a structural reorganization of the P13 stem has been proposed based on phylogenetic analysis (18), it suggests that the kissing-loop interaction has a role in controlling the access to the RBS region (Fig. 8A). To decipher the role of the P13 stem in the riboswitch regulation, we engineered *btuB* riboswitch constructs harboring mutations in P13 and P14 helical domains. We first disrupted the stability of the stem by substituting residues 173-176 for their Watson-Crick complement (M13.1, Fig. 8A). When assaying the effect of these mutations in a translational *BtuB-LacZ* construct, the  $\beta$ -galactosidase activity in absence and presence of AdoCbl was found to be dramatically decreased (Fig. 8B), indicating that the P13.1 mutant is constitutionally repressed. We also mutated base pairs U187-G198 and G188-U197 that are predicted to form only in absence of AdoCbl (Fig. 8A). When changing the identity of these nucleotides for their Watson-Crick complement (M13.2, Fig. 8A), very low *btuB* expression was observed both in

absence and presence of ligand (Fig. 8B). Similarly, introduction of U200A/G201A mutations resulted in relatively low  $\beta$ -galactosidase activities (M13.3, Fig. 8A), indicating that P13 stability is crucial for *btuB* regulation. As a control, we altered the sequence in the P14 loop that is not predicted to be involved in *btuB* gene regulation (M14.1, Fig. 8A). A decreased regulation efficiency of ligand-dependent repression was observed for this mutant (~3.5-fold), indicating that the riboswitch can still perform AdoCbl-dependent regulation. These results clearly indicate that the presence of the P13 stem not only serve to stabilize the AdoCbl-riboswitch complex (Fig. 5B), but is also important for *btuB* expression in absence of ligand by maintaining the A-RBS/A'-RBS interaction (Fig. 8A).

**Figure 8**

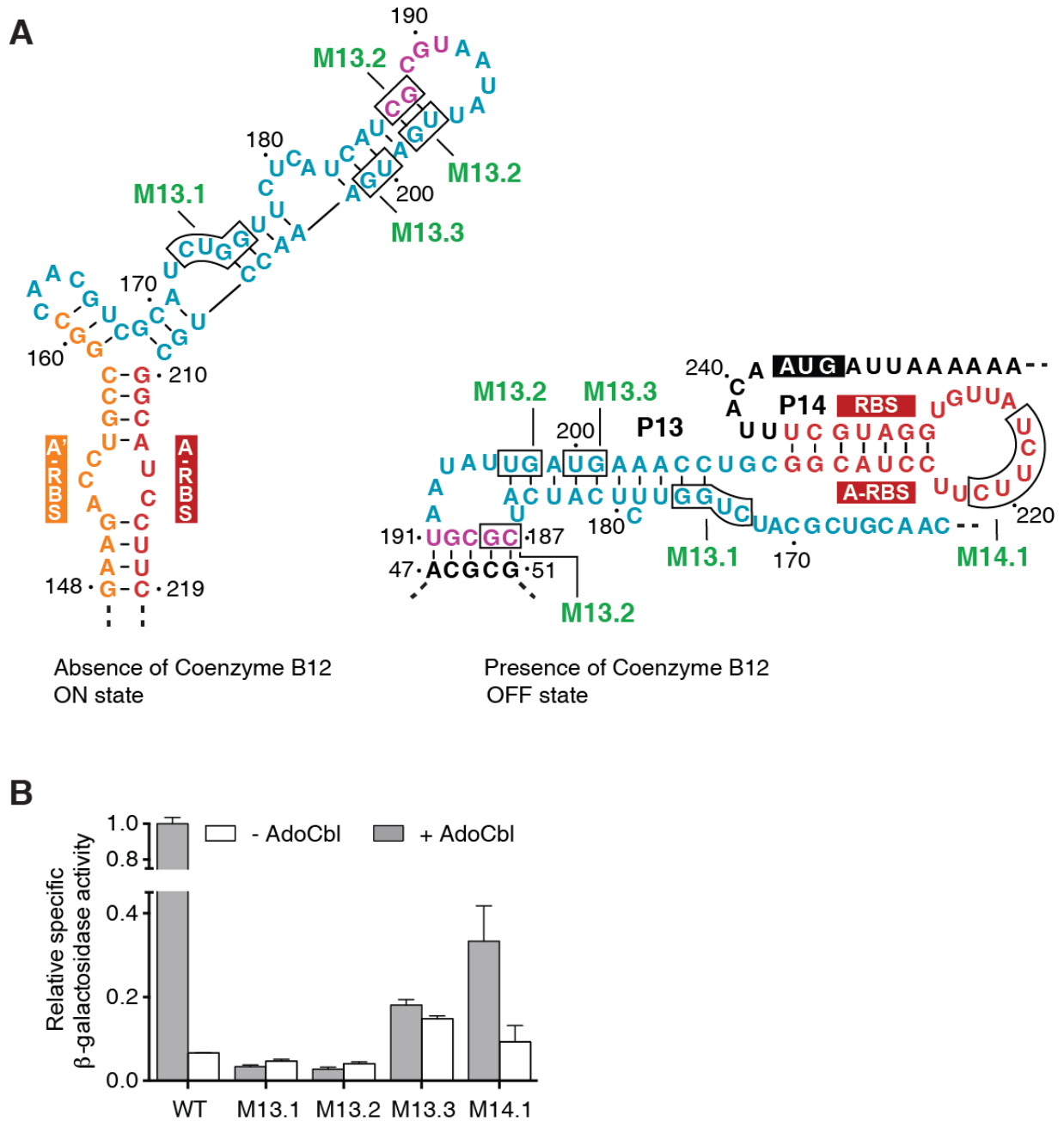


FIGURE 8. *A*, secondary structures representing the *btuB* expression platform when the riboswitch is in the absence or in presence of bound AdoCbl. Analyzed mutations are boxed and

are shown in both structures. The conformations of both structures are represented according to previous studies (15,18,32). The color scheme corresponds to the one used in Fig. 1C. While the A'-RBS and the A-RBS are shown interacting in absence of AdoCbl, the P14 stem comprising the A-RBS and RBS sequences is represented in presence of AdoCbl. The kissing-loop interaction is also shown in presence of ligand. The AUG start codon is shown in a black box in the AdoCbl-bound structure. *B*,  $\beta$ -galactosidase assays of translational BtuB-LacZ fusions for the wild-type and P13/P14 stem mutants. Enzymatic activities were measured in the absence and presence of 5  $\mu$ M AdoCbl. Values are normalized to the enzyme activity obtained for the wild-type and the absence of ligand. The average values of three independent experiments with SDs are shown.

Formation of kissing-loop interaction relies on transcriptional context. The presence of the U225 pause site allowing kissing-loop formation (18) suggests a folding mechanism in which the pause site allows sufficient time for the kissing-loop interaction to form, which is important for AdoCbl sensing (Fig. 5A). To explore this folding mechanism, we engineered a *trans*-acting version of the *btuB* riboswitch in which an isolated aptamer domain and an expression platform are allowed to interact together through the kissing-loop, thereby mimicking a kissing-loop interaction taking place in the paused transcription complex (Fig. 9A). In these experiments, the 161 nt aptamer construct was transcribed using the *E. coli* RNA polymerase in presence of increasing concentration of AdoCbl. In each reaction, a transcript corresponding to the expression platform was incubated to reconstitute the kissing-loop structure through intermolecular interaction with the aptamer domain. RNase H assays were first performed on transcription reactions containing the aptamer construct incubated with an equimolar molar ratio of the expression platform transcript. In such a context, RNase H assays yielded a  $K_{\text{switch}}$  value of  $71010 \pm 80$  nM (Fig. 9B), which is similar to that obtained in absence of the kissing-loop ( $K_{\text{switch}} = 62100 \pm 6000$  nM, Fig. 7B). These results suggest that a *trans*-acting version of the *btuB* riboswitch was not obtained in these conditions. Similar results were obtained when using a 10-fold molar excess of the expression platform (data not shown). Together, these results indicate that the addition in *trans* of the expression platform does not allow the kissing-loop

interaction, in agreement with the transcriptional context being crucial for the formation of the kissing-loop structure.

**Figure 9**

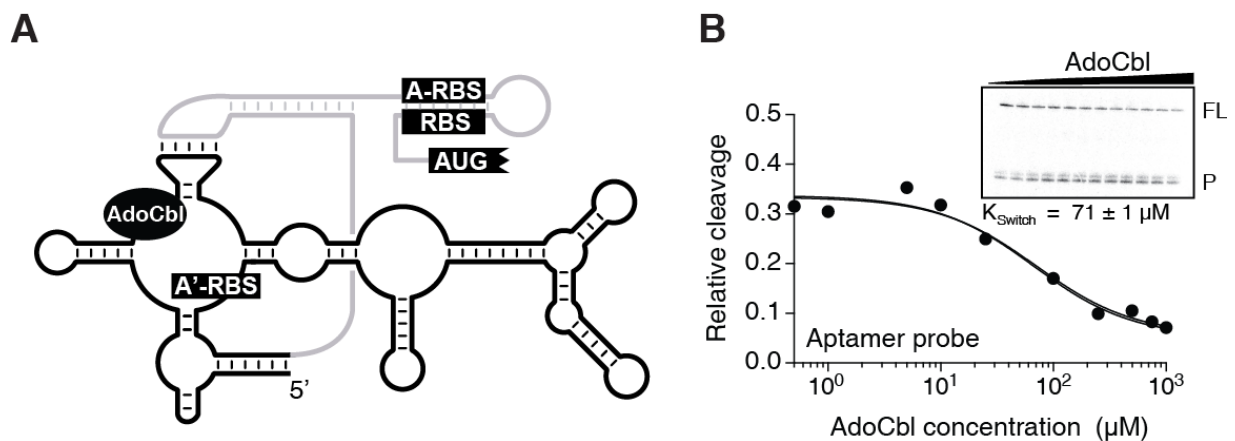


FIGURE 9. *A*, schematic representing a *trans*-acting version of the *btuB* riboswitch where a 161 nt riboswitch aptamer construct is allowed to interact with a transcript corresponding to the riboswitch expression platform (gray). The riboswitch is shown as when bound to AdoCbl. *B*,  $K_{\text{switch}}$  determination is performed as a function of AdoCbl concentration ranging from 1 nM to 250  $\mu\text{M}$  and using 20  $\mu\text{M}$  of the aptamer probe. In presence of low AdoCbl concentrations, a cleavage product (P) is observed and is decreased at high ligand concentrations, thus yielding the full-length transcript (FL). The plot depicting normalized fractions of cleaved RNA *versus* AdoCbl concentrations yields a  $K_{\text{switch}}$  value of  $71010 \pm 80$  nM. Experiments were done in triplicate and reported data on the graph correspond to a representative result.

## 2.5 DISCUSSION

RNA-based gene regulation mechanisms inherently rely on the correct folding of intervening RNA molecules, ranging from simple productive base pairing to complex tertiary structure formation (2,5,40-42). As found for large RNA molecules such as the group I intron and RNase P (42), the transcription process is an integral part of the RNA folding pathway since the polarity of transcription allows incremental availability of newly synthesized RNA portions, therefore modulating the formation of upstream RNA structure. This polarity-driven folding process is further controlled through the strategic locations of transcriptional pause sites (43) that presumably allow sufficient time for secondary and tertiary interactions to take place, often being of high importance in the folding of complex RNA structures (44,45). Transcriptional pausing has also been identified in metabolite-binding riboswitches responsive to flavin mononucleotide (46), adenine (26) and AdoCbl (18), and in the pH-responsive riboregulator (47).

The strong requirement of the *btuB* riboswitch to rely on the transcription process to attain its native folding (18) requires that structural studies aimed at characterizing *btuB* folding should preferentially be done using nascent transcripts. However, low transcription efficiencies of the *E. coli* RNA polymerase *in vitro*, as compared to T7 RNA polymerase, renders more technically challenging to probe RNA structure as conventional probing techniques (e.g., SHAPE analysis) often require larger RNA amount to detect chemical modifications through reverse transcription. Although SHAPE analysis allows to monitor *btuB* structural changes upon ligand binding (Fig. 3) (9), it has been shown that transcriptional pausing is critical to coordinate ligand binding and riboswitch conformational changes taking place in the expression platform, thus requiring to use of nascent transcripts. RNase H cleavage assays are well suited to study nascent transcripts, as they require low RNA amount and can be specifically targeted on different transcript domains. However, in contrast to other probing techniques directly assessing ligand binding, RNase H cleavage activity relies on the hybridization of an oligonucleotide that could potentially introduce a bias in the study. In the present case,  $K_{\text{switch}}$  values obtained by targeting



different riboswitch domains range from 53 to 224 nM, which are all in the range of the reported  $\sim 300$  nM  $K_d$  obtained using in-line probing assays (10) and RNase T1 analysis (Fig. 6B). Thus, these values suggest that oligonucleotide binding does not significantly perturb the formation of the riboswitch-ligand complex. Since variations between  $K_{\text{switch}}$  values are observed between different oligonucleotides used, it suggests that not only the ligand binding affinity is reflected in the determined  $K_{\text{switch}}$  value. Nevertheless, the ability to directly assess defined riboswitch structural features coupled to mutagenesis is a powerful approach to comprehend the riboswitch folding mechanism.

In this study, we further characterized the folding of the cobalamin-sensing *btuB* riboswitch that is one of the largest known riboswitches exhibiting complex secondary and tertiary structures (8,9,14,15). Such an intricate RNA architecture is presumably important to specifically recognize the relatively large AdoCbl metabolite, which is mainly achieved through shape complementarity (8,9). Combined with previous findings (18), our results support a model in which transcriptional pausing at C184 allows time for *btuB* aptamer domain formation and where the presence of AdoCbl results in the sequestering of the A'-RBS region via the formation of the P1 stem. Although such a paused complex contains all necessary elements to bind AdoCbl, RNase H data suggest that the free concentration of AdoCbl would need to be in the micromolar range to generate the riboswitch-ligand complex (Fig. 7B). However, our results indicate that further synthesis of the *btuB* riboswitch leading to the second transcriptional pause site (U225) significantly improves AdoCbl affinity. The dominant structural feature contained in the newly transcribed *btuB* portion corresponds to the kissing-loop interaction, therefore implying that the formation of the kissing-loop is important for AdoCbl binding, in agreement with its close proximity to the ligand in crystal structures (8,9). Since RNase H data show that *btuB* transcripts containing 202 and 206 nucleotides respectively exhibit  $\sim 267$  and 540-fold decrease in  $K_{\text{switch}}$  values (Fig. 7B), it suggests that both kissing-loop formation as well as P13 stability are crucial elements in AdoCbl recognition. Whether the formation of the kissing-loop is achieved through a conformational capture or induced-fit mechanism remains to be determined. However, since the kissing-loop can be formed in absence of AdoCbl at elevated

Mg<sup>2+</sup> concentrations (9), it suggests that the kissing-loop exhibits structural dynamics allowing to sense the presence of AdoCbl, which would be consistent with a conformational capture mechanism as previously observed for other riboswitches (7,34,48).

Our results indicate that the kissing-loop interaction is also important to coordinate AdoCbl binding and riboswitch conformational changes taking place in the expression platform (Figs. 5B and 5C). Such an unconventional role for a kissing-loop interaction differs from previous studies reporting that kissing-loop structures are usually involved in stabilizing the ligand-riboswitch complex, as found for purine and lysine riboswitches (33,49). The presence of multiple transcriptional pause sites located across the expression platform is indicative that the folding process is highly controlled in this region, consistent with the decrease folding response in presence of ligand when the riboswitch is transcribed using T7 RNA polymerase (18). Because of such strategically located pause sites, it is likely that the formation of the kissing-loop structure is directing the folding of P13 and P14 helical domains that are involved in riboswitch regulation, consistent with RNase H data (Fig. 5B). This regulatory control involves a "helix sliding" mechanism rearranging part of P13 base pairs (Fig. 1A)(18), which is presumably important for P14 folding. Cobalamin-responsive riboswitches controlling translation typically exhibit L13 loop elements containing the RBS sequence, thus providing a direct mechanism modulating translation initiation through kissing-loop formation (8,9). This is in contrast to the *E. coli btuB* riboswitch containing P13 and P14 structural elements (Fig. 1A), the latter being directly implicated in RBS selective sequestration. Secondary structure predictions suggest that ligand-binding to the *btuB* riboswitch results in a major rearrangement of the P13 helical domain in which L13 residues become available for kissing-loop interaction (Fig. 8A). Surprisingly,  $\beta$ -galactosidase assays show that destabilization of P13 base pairs (M13 mutants) severely limits AdoCbl-induced gene regulation by markedly repressing gene expression (Fig. 8B). These results imply that the stability of P13 is mostly important to maintain *btuB* expression in the ligand-free form. In particular, mutations destabilizing base pairs C187-G198 and G188-U197 (M13.2 mutant) strongly alter AdoCbl-dependent gene repression (Fig. 8B) indicating that the stability of P13 base pairs comprising kissing-loop residues is very

important for riboswitch activity. Precedence for such a helical conformational change has been observed for the VS ribozyme where an intramolecular kissing-loop interaction induces a helical rearrangement of the substrate stem-loop that is required for ribozyme catalysis (50). Although this mechanism presents similar characteristics to the *btuB* riboswitch, there is still no direct evidence indicating whether P13 conformational changes are performed as a consequence of P1 stem formation or through the interaction of the kissing-loop, or both. However, half-life measurements performed in the absence or presence of AdoCbl indicate that ligand binding to *btuB* mRNA can occur post-transcriptionally (11), suggesting that pre-existing riboswitches are able to undergo conformational rearrangements to repress gene expression upon ligand binding. Clearly, additional work will need to be done to understand the molecular basis of the structural switching.

Similarly to M13 mutants, sequence changes introduced in the P1 stem perturb *btuB* genetic control by strongly repressing expression regardless of AdoCbl presence (Fig. 2B). These findings are in contrast to other riboswitches where mutations introduced in the 5' side of the P1 stem usually result in the de-repression of gene expression (26,31). Given that M1.2 and M1.4 mutants directly weaken base pairs formation of the A'-RBS/A-RBS interaction, the A-RBS/RBS helical domain (P14) is readily favored in such a context thereby resulting in low  $\beta$ -galactosidase activities. However, for M1.1 and M1.3 mutants, no obvious deleterious effect can be deduced since the current regulation model does not take into account the 5' side of the P1 stem (Fig. 1A). However, a secondary structure analysis performed using the program *mfold* predicts an interaction between the 5' side of the P1 stem and L5 residues (48-52 nt) in absence of AdoCbl. Thus, in the context of M1.1 and M1.3 mutants, it is possible that the P1-L5 interaction is disabled, which would therefore promote the kissing-loop interaction and *btuB* gene repression.

In summary, our study shows that the formation of the long-range kissing-loop interaction is important for efficient ligand binding and is vital to coordinate conformational changes in the riboswitch expression platform. Given that *btuB* transcriptional pause sites are strategically

located to allow formation of riboswitch key structural elements such as the kissing-loop (18), it suggests that the transcription elongation is a highly controlled process in which discontinuous movements of the RNA polymerase are an integral part of the riboswitch gene regulatory control. The presence of transcriptional pause sites in a wide range of bacterial transcripts (51) suggests that pause sites are highly implicated in the regulation of bacterial RNAs.

## 2.6 EXPERIMENTAL PROCEDURES

**Bacterial Strains**—Bacterial strains used in this study were derived from *E. coli* MG1655. All described *btuB* mutations were performed as previously described (20). Translational BtuB-LacZ fusions (Table 1) were made using a PCR reaction procedure (21,22).

**Templates for In Vitro Transcription**—Templates were produced using a *lacUV5* promoter (23) and were amplified from *E. coli* MG1655 genomic DNA. Kissing-loop riboswitch mutants were amplified using bacterial colony PCR from  $\beta$ -galactosidase strains (Table 2). Templates for the expression platform (residues 162-218) were made by using the T7 promoter (24).

**Radioactive Labeling of Trinucleotide Initiator**—To ensure specific end-labeling of the *btuB* riboswitch in single-round transcriptions, a GCC trinucleotide (Oligos etc.) was radiolabeled using a polynucleotide kinase T4 (Thermo scientific). Briefly, 5 pmol of trinucleotide was labeled according to the manufacturer's protocol using [ $\gamma$ - $^{32}$ P] ATP at 37°C for 1 h, and was directly added to the transcription mixture without further purification.

**Single-round In Vitro Transcription Assays**—*btuB* transcripts were produced using *in vitro* transcription using the *E. coli* RNA polymerase (Epicentre). Briefly, transcription reactions were prepared using 300 fmol of DNA template, 300 nM of CTP and GTP, 1 pmol of  $^{32}$ P-GCC trinucleotide and 1X TB (20 mM Tris-HCl pH 8.0, 20 mM NaCl, 20 mM MgCl<sub>2</sub>, 14 mM  $\beta$ -mercaptoethanol and 0.1 mM EDTA). The transcription mixture was incubated at 37°C for 5 min. A second reaction mixture containing 5  $\mu$ M UTP and 0.2 U/ $\mu$ L *E. coli* RNA polymerase

was then added and the resulting mixture was incubated at 37°C for 15 min to allow transcription initiation. A third mixture containing 100 μM NTP, 1X TB and 1 mg/mL heparin was added to the reaction vessel and incubated 15 min at 37°C for transcription elongation in single-round conditions. AdoCbl was added to the required concentration. Experiments described in Fig. 9 were performed by further addition of 150 fmol of a purified transcript corresponding to the riboswitch expression platform (162-218 nt). The resulting mixture was incubated at 37°C for 5 min.

**RNase H Cleavage Analysis**—To completed transcription reactions, a mixture containing 1X cleavage buffer (5 mM Tris-HCl, pH 8.0, 20 mM MgCl<sub>2</sub>, 100 mM KCl, 50 μM EDTA and 10 mM β-mercaptoethanol) was added together with 20 μM of an oligonucleotide probe for 5 min at 37°C. To the reacting vessel, 0.1 U/μL of Ribonuclease H (Ambion) was added and incubated for 5 min at 37°C. Where indicated, RNase H reactions were performed using 50 μM of an oligonucleotide. Reactions were stopped using an equal volume of formamide loading dye.  $K_{\text{switch}}$  values were obtained using a similar procedure to that used for  $T_{50}$  (25,26). The quantification assumes a simple 1:1 stoichiometry between the aptamer and AdoCbl, as expected from in-line probing data (10,15). Reported values represent an average of at least three independent experiments.

**RNase T1 Cleavage Analysis**—Transcription reactions were performed on solid surface using 5'-biotinylated DNA template (27). Reactions were prepared using 2 pmol of DNA template, 1X TB1 (5 mM Tris-HCl pH 8.0, 100 mM KCl, 10 mM and MgCl<sub>2</sub>), 1 pmol RNAP-β'his-tag and 2 pmol of sigma 70 factor, and incubated for 5 min at 37°C. A solution of 10 μM GCC trinucleotide, 25 μM of CTP and GTP and 15 μCi [ $\alpha$ -<sup>32</sup>P] UTP was then added and incubated for 8 min at 37°C. A mixture of 10 μL of magnetic avidin beads was added to the reaction and incubated for 10 min at 37°C. Resulting transcription complexes attached to beads were washed two times using a solution of 5 mM Tris-HCl pH 8.0, 1 M KCl and 10 mM MgCl<sub>2</sub> and resuspended in TB buffer. Transcriptions were completed by addition of 100 μM NTPs and 1 mg/mL heparin in absence or presence of AdoCbl, and incubated for 15 min at 37°C. The

supernatant was subjected to RNase T1 (0.2 U/ $\mu$ L) for 1 min at 37°C. Reactions were stopped using PCI extraction and mixed with equal volume of formamide loading dye. Ladders were performed using 3'-O-methyl NTP as previously described (26).

Preparation of Transcripts Corresponding to the Expression Platform—A *btuB* transcript corresponding to the expression platform (positions 162 to 218) was produced using T7 RNA polymerase transcription. A GCG sequence was added at the beginning of the transcript to minimize the 5'-heterogeneity of the RNA population (28). Transcription and RNA purifications steps were performed as previously described (29).

SHAPE Analysis—The SHAPE reaction was performed as described previously (30). RNA samples were slow-cooled in a solution containing 100 mM K-HEPES pH 8.0, 100 mM NaCl, 3 mM MgCl<sub>2</sub> and the indicated concentration of AdoCbl. *N*-methylisatoic anhydride (NMIA) dissolved in dimethyl sulfoxide (DMSO) was added to react for 80 min at 37 °C. NMIA-treated RNA were precipitated, resuspended and reverse transcribed as previously described (30). Two DNA oligonucleotides were used: BtuB202 (179-202 nt) and BtuB234 (212-234 nt) to cover the entire sequence of the *btuB* riboswitch.

$\beta$ -Galactosidase Assays— $\beta$ -galactosidase experiments were performed as previously described (31). Briefly, the bacterial culture was grown overnight in M63 0.2% glycerol minimal medium and diluted 50X into fresh medium. Arabinose (0.1%) was added to induce expression of *lacZ* constructs and AdoCbl (5  $\mu$ M) was added where indicated.  $\beta$ -galactosidase activities were calculated as previously described (31) and were relativized to the specific activity of the wild-type construct obtained in absence of ligand. Reported activities represent data of at least three independent experimental trials.

**Table 1.** Summary of strains/plasmids used in this study.

<b>Strains</b>	<b>Relevant marker</b>	<b>References</b>
AL1	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub>	This study
AL2	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> M1.1	This study
AL3	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> M1.2	This study
AL4	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> M1.3	This study
AL5	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> M1.4	This study
AL6	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> M1.1/1.2	This study
AL7	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> M1.3/1.4	This study
AL8	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> K1	This study
AL9	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> K1c	This study
AL10	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> K2	This study
AL11	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> K2c	This study
AL12	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> K3	This study
AL13	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> K3c	This study
AL14	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> K1.1	This study
AL15	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> KM13.1	This study
AL16	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> KM13.2	This study
AL17	PM1205 lacI':::PBAD-BtuB <sub>27cd</sub> KM13.3	This study
AL18	PM1205 lacI':::PBAD-BtuB <sub>70cd</sub> KM14.1	This study

**Table 2.** PCR constructs used for *in vitro* RNA synthesis.

<b>Constructions</b>	<b>Oligonucleotides</b>
<b><i>In vitro</i> transcription assays</b>	
WT construct	5AL-765LB (MG1655 genomic DNA)
K1 mutant	5AL-765LB (K1 strain genomic DNA)
K1c mutant	5AL-765LB (K1c strain genomic DNA)
K1.1 mutant	5AL-765LB (K1.1 strain genomic DNA)
161 nt variant	5AL-1121LB (MG1655 genomic DNA)
202 nt variant	5AL-TrxbtuB2 (MG1655 genomic DNA)
206 nt variant	5AL-1025AL (MG1655 genomic DNA)
Expression platform	1123LB-1125LB (MG1655 genomic DNA)
WT construct for SHAPE	ToebtuB2-TrxbtuB1 (MG1655 genomic DNA)



**Table 3.** Summary of oligonucleotides used for *in vitro* probing.

Oligonucleotides	Sequence 5'-3'
5 AL	GGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATGT GTGGGCCGGTCCTGTGAGTTAATAG
765 LB	CGGGCTGGTATCCTGTG
1025 AL	CGCAGGTTTCATCAATATTACGCGA
1121 LB	GCCGGCAGGTCTTCGGGCTTGG
1123 LB	AATAATATCGACTCACTATAGGGCCAACGTCGCATCTGGTTCTCA TCATCG
1125 LB	AAGGATGCCGCAGGTTTCATCAATATTACGC
Aptamer probe	CGCAATCATCGCAC
P13 probe	ATGATGAGAACCAGATGC
KL probe	TCATCAATATTACGCGATGATG
KLc probe	TCATCAATATTACGGCATGATG
ToebtuB2	CGGGCTGGTATCCTGTG
TrxbtuB1	AATAATATCGACTCACTATAGGGGCCGGTCCTGTGAGTTAATAG
TrxbtuB2	TCATCAATATTACGCGATGATGAG
BtuB202	TCATCAATATTACGCGATGATGAG
BtuB234	GCATCCACAATAGAAGAAGGATG

## 2.7 REFERENCES

1. Serganov, A., and Nudler, E. (2013) *Cell* **152**, 17-24
2. Breaker, R. R. (2012) *Cold Spring Harb Perspect Biol* **4**, 1-15
3. Batey, R. T. (2012) *Quarterly reviews of biophysics* **45**, 345-381
4. Garst, A. D., and Batey, R. T. (2009) *Biochim Biophys Acta* **1789**, 584-591
5. Peselis, A., and Serganov, A. (2014) *Biochim Biophys Acta* **1839**, 908-918
6. Gilbert, S. D., Rambo, R. P., Van Tyne, D., and Batey, R. T. (2008) *Nat Struct Mol Biol* **15**, 177-182
7. Haller, A., Rieder, U., Aigner, M., Blanchard, S. C., and Micura, R. (2011) *Nat Chem Biol* **7**, 393-400
8. Peselis, A., and Serganov, A. (2012) *Nat Struct Mol Biol* **19**, 1182-1184
9. Johnson, J. E., Jr., Reyes, F. E., Polaski, J. T., and Batey, R. T. (2012) *Nature* **492**, 133-137
10. Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L., and Breaker, R. R. (2002) *Chem Biol* **9**, 1043
11. Nou, X., and Kadner, R. J. (1998) *J Bacteriol* **180**, 6719-6728
12. Bradbeer, C., Kenley, J. S., Di Masi, D. R., and Leighton, M. (1978) *J Biol Chem* **253**, 1347-1352
13. Nou, X., and Kadner, R. J. (2000) *Proc Natl Acad Sci U S A* **97**, 7190-7195
14. Vitreschak, A. G., Rodionov, D. A., Mironov, A. A., and Gelfand, M. S. (2003) *Rna* **9**, 1084-1097
15. Nahvi, A., Barrick, J. E., and Breaker, R. R. (2004) *Nucleic Acids Res* **32**, 143-150
16. Choudhary, P. K., and Sigel, R. K. (2014) *RNA* **20**, 36-45
17. Gallo, S., Oberhuber, M., Sigel, R. K., and Krautler, B. (2008) *Chembiochem* **9**, 1408-1414
18. Perdrizet, G. A., 2nd, Artsimovitch, I., Furman, R., Sosnick, T. R., and Pan, T. (2012) *Proc Natl Acad Sci U S A* **109**, 3323-3328

19. Souliere, M. F., Haller, A., Santner, T., and Micura, R. (2013) *Angewandte Chemie* **52**, 1874-1877
20. Yu, D., Ellis, H. M., Lee, E. C., Jenkins, N. A., Copeland, N. G., and Court, D. L. (2000) *Proc Natl Acad Sci U S A* **97**, 5978-5983
21. Mandin, P., and Gottesman, S. (2009) *Mol Microbiol* **72**, 551-565
22. Masse, E., and Gottesman, S. (2002) *Proc Natl Acad Sci U S A* **99**, 4620-4625
23. Hollands, K., Proshkin, S., Sklyarova, S., Epshtein, V., Mironov, A., Nudler, E., and Groisman, E. A. (2012) *Proc Natl Acad Sci U S A* **109**, 5376-5381
24. Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) *Nucleic Acids Res* **15**, 8783-8798
25. Blouin, S., and Lafontaine, D. A. (2007) *RNA* **13**, 1256-1267
26. Lemay, J. F., Desnoyers, G., Blouin, S., Heppell, B., Bastet, L., St-Pierre, P., Masse, E., and Lafontaine, D. A. (2011) *PLoS Genet* **7**, e1001278
27. Nudler, E., Gusarov, I., and Bar-Nahum, G. (2003) *Methods Enzymol* **371**, 160-169
28. Pleiss, J. A., Derrick, M. L., and Uhlenbeck, O. C. (1998) *Rna* **4**, 1313-1317
29. Dalgarno, P. A., Bordello, J., Morris, R., St-Pierre, P., Dube, A., Samuel, I. D., Lafontaine, D. A., and Penedo, J. C. (2013) *Nucleic Acids Res* **41**, 4253-4265
30. Tremblay, R., Lemay, J. F., Blouin, S., Mulhbacher, J., Bonneau, E., Legault, P., Dupont, P., Penedo, J. C., and Lafontaine, D. A. (2011) *J Biol Chem* **286**, 27406-27415
31. Caron, M. P., Bastet, L., Lussier, A., Simoneau-Roy, M., Masse, E., and Lafontaine, D. A. (2012) *Proc Natl Acad Sci U S A* **109**, E3444-3453
32. Barrick, J. E., and Breaker, R. R. (2007) *Genome Biol* **8**, R239
33. Blouin, S., Chinnappan, R., and Lafontaine, D. A. (2010) *Nucleic Acids Res* **39**, 3373-3387
34. Heppell, B., Blouin, S., Dussault, A. M., Mulhbacher, J., Ennifar, E., Penedo, J. C., and Lafontaine, D. A. (2011) *Nature Chem Biol* **7**, 384-392
35. Merino, E. J., Wilkinson, K. A., Coughlan, J. L., and Weeks, K. M. (2005) *J Am Chem Soc* **127**, 4223-4231
36. Treiber, D. K., and Williamson, J. R. (2000) *Methods Enzymol* **317**, 330-353
37. Zarrinkar, P. P., and Williamson, J. R. (1994) *Science* **265**, 918-924

38. Komissarova, N., and Kashlev, M. (1998) *Proc Natl Acad Sci U S A* **95**, 14699-14704
39. Monforte, J. A., Kahn, J. D., and Hearst, J. E. (1990) *Biochemistry* **29**, 7882-7890
40. Waters, L. S., and Storz, G. (2009) *Cell* **136**, 615-628
41. Breaker, R. R. (2011) *Mol Cell* **43**, 867-879
42. Pan, T., and Sosnick, T. (2006) *Annu Rev Biophys Biomol Struct* **35**, 161-175
43. Landick, R. (2006) *Biochem Soc Trans* **34**, 1062-1066
44. Pan, T., Artsimovitch, I., Fang, X. W., Landick, R., and Sosnick, T. R. (1999) *Proc Natl Acad Sci U S A* **96**, 9545-9550
45. Wong, T. N., Sosnick, T. R., and Pan, T. (2007) *Proc Natl Acad Sci U S A* **104**, 17995-18000
46. Wickiser, J. K., Winkler, W. C., Breaker, R. R., and Crothers, D. M. (2005) *Mol Cell* **18**, 49-60
47. Nechooshtan, G., Elgrably-Weiss, M., Sheaffer, A., Westhof, E., and Altuvia, S. (2009) *Genes Dev* **23**, 2650-2662
48. Hammond, M. C. (2011) *Nat Chem Biol* **7**, 342-343
49. Batey, R. T., Gilbert, S. D., and Montange, R. K. (2004) *Nature* **432**, 411-415
50. Andersen, A. A., and Collins, R. A. (2000) *Mol Cell* **5**, 469-478
51. Larson, M. H., Mooney, R. A., Peters, J. M., Windgassen, T., Nayak, D., Gross, C. A., Block, S. M., Greenleaf, W. J., Landick, R., and Weissman, J. S. (2014) *Science* **344**, 1042-1047
53. Larson, M. H., Mooney, R. A., Peters, J. M., Windgassen, T., Nayak, D., Gross, C. A., Block, S. M., Greenleaf, W. J., Landick, R., and Weissman, J. S. (2014) *Science* **344**, 1042-1047

## **2.8 FOOTNOTES**

We thank members of the Lafontaine laboratory for helpful discussions and Alain Lavigneur for critical reading of the manuscript. The Canadian Institutes for Health Research (CIHR) supported this work. D.A.L. is a Chercheur-Boursier Senior of the Fonds de la Recherche en Santé du Québec.

## CHAPITRE 3

### DISCUSSION

L'objectif de ce projet de maîtrise était de déterminer le rôle du pseudonoeud dans le changement de conformation du riborégulateur *btuB* d'*E. coli*. Classiquement, c'est la formation de la tige P1 à elle seule, lors de la liaison du ligand, qui permet à la plateforme d'expression de changer de conformation pour effectuer la régulation génétique.(Garst & Batey 2009)(Batey 2012) Par contre, puisque le pseudonoeud de *btuB* se situe entre la tige P1 et la tige P14 séquestratrice du RBS (Figure 1C, Lussier et al, manuscrit en préparation), il est logique de présumer que cette structure pourrait avoir un certain rôle à jouer lors du changement de conformation entre la forme non liée et liée au ligand. Comme discuté lors de l'introduction, c'est avec les travaux de Nahvi et al.(Nahvi et al. 2002) que cette structure a été proposée et que son importance pour la régulation génétique a été démontrée. Par la suite, Perdrizet et al.(Perdrizet et al. 2012) ont proposé un modèle où le pseudonoeud n'est pas formé en l'absence du ligand, formant ainsi un lien entre les modèles non lié et celui lié au ligand. Cette structure pourrait alors fonctionner de concert avec la tige P1 pour permettre la formation de la tige séquestratrice P14 (Figure 1C Lussier et al, manuscrit en préparation). De plus, à l'aide des données cristallographiques,(Johnson et al. 2012)(Peselis & Serganov 2012) on peut supposer que le pseudonoeud participe à la liaison du ligand par sa proximité spatiale avec ce dernier. Mes travaux, ainsi que certains résultats d'essais  $\beta$ -galactosidases obtenus par Laurène Bastet, ont démontré l'importance du pseudonoeud dans la régulation génétique par sa séquence et sa structure.

Tout d'abord, nous avons tenté de déterminer si la tige P1 jouait un rôle clé dans la réponse au ligand, comme c'est le cas pour la plupart des riborégulateurs. Pour ce faire, nous avons utilisé des essais  $\beta$ -galactosidases avec différents mutants de la tige P1 (Figure 2A Lussier et al, manuscrit en préparation). À l'aide de ces mutants, nous avons déstabilisé la tige P1 et nous

avons tenté de la rétablir à l'aide de mutants compensatoires. Nous avons alors constaté que les mutants déstabilisant la tige P1 perdent leur capacité à répondre au ligand (Figure 2B Lussier et al, manuscrit en préparation). Les mutations compensatoires, quant à elles, permettent de récupérer un certain niveau de régulation par le ligand mais ne permettent pas de reconstituer une régulation comparable au type sauvage. Donc, cette tige est sensible aux modifications de séquence puisque même les mutants compensatoires ne permettent pas de récupérer une régulation de type sauvage. À la lumière de ces résultats, nous pouvons affirmer que la tige P1 de ce riborégulateur joue un rôle clé dans la régulation génétique comme la plupart des riborégulateurs connus.

De manière similaire à la tige P1, le pseudonoeud ne peut pas être modifié au niveau de la séquence sans affecter la régulation génétique et/ou le repliement du riborégulateur. En effet, lorsque des mutations sont introduites dans cette structure en vue de la déstabiliser, le riborégulateur perd sa capacité de régulation génétique à l'AdoCbl dans les essais  $\beta$ -galactosidase. Ce phénomène a été observé pour différents mutants du pseudonoeud (Figure 2C, D Lussier et al, manuscrit en préparation). De plus, lorsque le pseudonoeud est muté dans le but de rétablir sa formation à l'aide de mutants complémentaires, nous ne retrouvons pas le même niveau de régulation que le type sauvage, et ce, pour tous les mutants complémentaires (Figure 2D Lussier et al, manuscrit en préparation). Ces résultats suggèrent que la présence du pseudonoeud est nécessaire à la régulation et que la nature des nucléotides formant cette structure affecte l'efficacité de cette régulation. Nous avons aussi observé le changement de conformation du riborégulateur à l'aide de la cartographie à la ribonucléase H. À l'aide de cet essai, nous avons déterminé une valeur que nous avons appelée le  $K_{\text{switch}}$ . En bref, cette valeur représente la concentration de ligand nécessaire pour que la moitié des molécules d'ARN changent de conformation *in vitro*. Par exemple, une valeur élevée de  $K_{\text{switch}}$  signifie qu'une grande concentration de ligand est nécessaire pour que le riborégulateur bascule de la forme non liée à la forme liée. Comme cet essai se base sur le clivage d'un hybride ADN-ARN, nous avons utilisé trois sondes d'ADN s'hybridant dans les régions d'intérêts suivantes : P8/P9, la tige P13 et le pseudonoeud. Alors, nous avons obtenus, pour chacune des sondes, des valeurs de  $K_{\text{switch}}$

pour chaque région ciblée (Figure 3A Lussier et al, manuscrit en préparation). D'abord, nous avons choisi une sonde s'hybridant dans la région P8/P9 parce qu'elle nous rapporte indirectement la liaison du ligand par sa proximité avec une zone étant fortement modifiée lors de la liaison du ligand (Figure 4C Lussier et al, manuscrit en préparation). Ensuite, les sondes de la tige P13 et du pseudonoeud nous permettent de déterminer le repliement de l'ARN à la frontière entre l'aptamère et la plateforme d'expression et de rendre compte de la formation du pseudonoeud. De cette manière, nous pouvons observer le repliement de l'ARN à trois endroits différents pour une même molécule d'ARN. Avec cette méthode nous avons calculé la valeur de  $K_{switch}$  des trois sondes pour le type sauvage, un mutant du pseudonoeud (K1) et le mutant complémentaire à ce dernier (K1c). Tout d'abord, nous avons obtenu pour le type sauvage des valeurs de  $K_{switch}$  de  $53 \pm 4$  nM,  $224 \pm 6$  nM,  $115 \pm 4$  nM pour les sondes aptamère, P13 et pseudonoeud respectivement (Figure 3B, D et E Lussier et al, manuscrit en préparation). Comme les mutants du pseudonoeud ont montré des résultats comparables en essais  $\beta$ -galactosidase, nous avons présumé que ces différents mutants avaient tous un impact similaire sur la régulation. Alors, nous nous sommes concentrés sur la caractérisation d'une seule paire de mutant du pseudonoeud (K1 et K1c) en ce qui concerne la détermination du  $K_{switch}$ . Ensuite, nous avons calculé la valeur de  $K_{switch}$  du mutant K1 pour la sonde aptamère qui est de  $12 \pm 3$   $\mu$ M (Figure 5A Lussier et al, manuscrit en préparation). Enfin, aucun  $K_{switch}$  n'a pu être calculé pour les 2 autres sondes pour le mutant K1 puisqu'il n'y avait aucune variation de clivage induite par le ligand (Figure 5B, C Lussier et al, manuscrit en préparation). Ces résultats obtenus pour le mutant K1 sont en accord avec les essais  $\beta$ -galactosidase puisque les mutants du pseudonoeud ne présentaient plus de régulation au ligand, d'où le besoin d'une plus grande concentration de ligand pour faire changer la conformation de l'ARN, résultant ainsi en un  $K_{switch}$  élevé. Par la suite, nous avons obtenus des valeurs plus similaires au type sauvage avec le mutant complémentaire (K1c) pour les trois sondes. Pour ce mutant complémentaire, nous avons obtenu des valeurs de  $K_{switch}$  de  $368 \pm 32$  nM pour la sonde aptamère,  $356 \pm 37$  nM pour la sonde P13 et  $261 \pm 85$  nM pour la sonde pseudonoeud (Figure 5 Lussier et al, manuscrit en préparation). On observe donc pour ce mutant des valeurs de  $K_{switch}$  qui se retrouvent plus près des valeurs du type sauvage. Ces résultats indiquent que ce mutant complémentaire permet de récupérer une



certaine capacité de régulation par rapport au mutant K1. Cela pourrait indiquer un rétablissement de la formation de la structure pseudonoëud, d'où la diminution des valeurs de  $K_{\text{switch}}$  qui sont plus comparables aux valeurs du type sauvage.

Ensuite, nous avons l'importance de la séquence du pseudonoëud et de la stabilité. Pour ce faire, nous avons utilisé différentes longueurs de transcrit (Figure 6A Lussier et al, manuscrit en préparation) que nous avons cartographiés à la ribonucléase H à l'aide des trois sondes utilisées précédemment. Nous avons alors conçu une construction du riborégulateur *btuB*, sans pseudonoëud, comprenant seulement 161 nucléotides (nt) et pour laquelle nous avons obtenu une valeur de  $K_{\text{switch}}$  de  $62 \pm 6 \mu\text{M}$  avec la sonde aptamère (Figure 6B Lussier et al, manuscrit en préparation). Cette valeur est environ 1000 fois plus élevée que la valeur du type sauvage pour la même sonde indiquant une perte importante d'efficacité de liaison du ligand pour cette construction. On peut remarquer que l'absence du pseudonoëud influence de manière importante la liaison du ligand. Dans la même voie, les mutants du pseudonoëud qui inhibent la formation de cette structure, ont aussi affiché un effet important sur la liaison du ligand, tel que constaté avec le mutant du pseudonoëud K1 décrit précédemment. Ces deux cas nous montrent que la formation du pseudonoëud est importante pour obtenir une liaison efficace du ligand par le riborégulateur. Pour observer l'effet de la présence du pseudonoëud sur la liaison du ligand, nous avons allongé la précédente construction (161 nt) à 202 nt ce qui permet la formation du pseudonoëud avec une tige P13 incomplète (Figure 6A Lussier et al, manuscrit en préparation). Cette construction a été utilisée dans plusieurs études sur ce riborégulateur et elle est considérée comme étant la séquence complète de l'aptamère. Pour cette construction, nous avons remarqué une liaison du ligand plus près du type sauvage pour la sonde aptamère avec une valeur de  $K_{\text{switch}}$  de  $232 \pm 25 \text{ nM}$  (Figure 6B Lussier et al, manuscrit en préparation). En ce qui concerne les sondes P13 et du pseudonoëud, nous avons obtenus des valeurs de  $K_{\text{switch}}$  de  $1,7 \pm 0,2 \mu\text{M}$  et de  $3,8 \pm 0,25 \mu\text{M}$  (Figure 6C, D Lussier et al, manuscrit en préparation). Ces valeurs nous indiquent que la présence du pseudonoëud rétablit une certaine mesure de liaison du ligand au riborégulateur. En effet, même si cette construction ne semble pas permettre une liaison du ligand comparable au type sauvage, elle montre que le ligand se lie plus efficace dans ce cas

qu'avec la construction 161nt sans pseudoneud. Par la suite, nous avons considéré une construction de 206 nt pour déterminer si l'intégrité de la tige P13 influence vraiment la liaison du ligand à l'ARN de *btuB* (Figure 6A Lussier et al, manuscrit en préparation). Nous avons observé pour les sondes aptamère et P13 des valeurs de  $K_{\text{switch}}$  de  $115 \pm 31$  nM et de  $404 \pm 30$  nM (Figure 6B, D Lussier et al, manuscrit en préparation). En ce qui concerne la sonde du pseudoneud, la valeur de  $K_{\text{switch}}$  serait même plus basse que le type sauvage à  $109 \pm 41$  nM (Figure 6C Lussier et al, manuscrit en préparation). Nous pouvons alors constater que les différences de valeur de  $K_{\text{switch}}$  entre les constructions 202 nt et 206 nt sont appréciables. En effet, il semble que lorsque la tige P13 est complète (206 nt), nous observons un repliement de l'ARN comparable au type sauvage, en comparaison d'une construction où cette tige est tronquée (202 nt). Nous pouvons donc proposer que l'intégrité de la tige P13 est un facteur important pour une liaison optimale du ligand au riborégulateur. En résumé, nos résultats indiquent que le pseudoneud est essentiel pour la liaison du ligand. En effet, la différence de  $K_{\text{switch}}$  entre une construction sans pseudoneud (161 nt) et une autre avec un pseudoneud (202nt) est assez notable. Ces résultats sont en accord avec les observations précédentes des mutants du pseudoneud (K1 et K1c) qui montrent eux aussi le rôle important de la formation du pseudoneud dans la liaison à l'AdoCbl. De plus, nous avons aussi remarqué que l'intégrité de la tige P13 est essentielle pour que le riborégulateur puisse se replier de manière comparable au type sauvage. Effectivement, nous avons observé des différences de  $K_{\text{switch}}$  entre les constructions 202 et 206 nucléotides qui pourrait supporter cette hypothèse. Ces résultats pourraient donc indiquer que lorsque la tige P13 est incomplète, la stabilité des régions de la tige P13 et du pseudoneud est affectée. La déstabilisation de la tige P13 pourrait donc indirectement influencer la formation du pseudoneud. Puis, lorsque la tige P13 est complète et donc plus stable, le riborégulateur peut alors lier le ligand de manière beaucoup plus comparable au type sauvage. Ce phénomène pourrait alors expliquer la différence des valeurs de  $K_{\text{switch}}$  entre les constructions 202 nt et 206 nt pour la sonde du pseudoneud. L'importance de la stabilité de la tige P13 a aussi été démontrée dans un contexte *in vivo* avec un mutant de la tige P13 (M13.1) en essais  $\beta$ -galactosidase (Figure 7A Lussier et al, manuscrit en préparation). Ce mutant de la tige P13, qui avait pour but de déstabiliser cette dernière, perd toute capacité de régulation en présence du ligand. Il semble alors que la stabilité de cette région est aussi importante dans un

contexte *in vivo* que *in vitro*. Donc, en combinant les résultats *in vivo* et *in vitro*, nous pouvons affirmer que la séquence du pseudonoeud est essentielle pour la liaison efficace du ligand au riborégulateur. Pour expliquer ce phénomène, il existe deux explications possibles basées sur la structure tridimensionnelle : soit le pseudonoeud interagit directement avec le ligand, soit la composition des bases du pseudonoeud crée une structure bien spécifique qui ne peut être altérée sans avoir un effet sur la liaison du ligand. Malheureusement, les techniques utilisées ne permettent pas d'affirmer avec précision, laquelle des deux hypothèses représente le mieux la réalité. Pour y répondre, il faudrait avoir à notre disposition un essai qui serait capable de fournir une résolution comparable à la cristallographie et pouvoir alors muter chaque nucléotide du pseudonoeud pour observer comment la liaison du ligand est affectée par rapport à chacune de ces mutations. À l'heure actuelle, la technique de cristallographie n'est pas assez avancée pour produire la quantité de cristaux nécessaire à une telle étude et aucune autre technique existante ne permet d'obtenir une telle résolution avec de l'ARN.

Nous avons aussi déterminé que l'intégrité de la plateforme d'expression est importante pour l'expression du riborégulateur. En effet, avec les essais  $\beta$ -galactosidase des mutants des tiges P13 et P14, nous avons remarqué que si ces régions sont déstabilisées à l'aide de certaines mutations (Figure 7B Lussier et al, manuscrit en préparation), l'expression du gène *lacZ* diminue de manière significative. Il semblerait donc que, dans une certaine limite, cette région structurée doit adopter une conformation particulière permettant l'expression génétique. Ces résultats sont en accord avec les modèles de structures secondaires disponibles qui montrent que la plateforme d'expression est hautement structurée que ce soit en l'absence ou en présence de ligand et que visiblement cette région structurée est importante pour l'expression génétique (Figure 7A Lussier et al, manuscrit en préparation).

Dans un article de Perdrizet et al. (2012), il a été démontré que le repliement co-transcriptionnel de l'ARN est crucial au bon repliement de la plateforme d'expression du riborégulateur *btuB*. Comme le pseudonoeud fait partie de la plateforme d'expression, nous avons voulu déterminer si le pseudonoeud peut se former en l'absence du contexte transcriptionnel. Pour ce faire, nous

avons produit un aptamère (1-161nt) du riborégulateur *btuB* transcrit avec l'ARNpol d'*E. coli*. Ensuite, nous avons produit séparément le domaine de la plateforme d'expression (162-218nt) avec la polymérase du phage T7 qui est utilisée couramment dans l'étude des riborégulateurs (Nahvi et al. 2002) (Winkler & Breaker 2003) (Kang et al. 2009) (Lafontaine 2007) (Choudhary & Sigel 2014) (Figure 8 Lussier et al, manuscrit en préparation). L'aptamère et la plateforme d'expression sont donc transcrits séparément et ensuite incubés ensemble pour essayer de déterminer s'il y a formation du pseudonoeud sans l'aide du contexte transcriptionnel. Alors, si on incube les deux ARN ensemble et que le pseudonoeud se forme, c'est que l'interaction de type pseudonoeud est favorisée thermodynamiquement et qu'il ne nécessite probablement pas le contexte transcriptionnel pour se former. Si cette structure ne se forme pas, c'est que le pseudonoeud n'est pas thermodynamiquement favorisé et que le contexte transcriptionnel est essentiel pour le repliement et la formation de cette structure. Il est à noter que pour cette expérience nous n'avons utilisé qu'une sonde s'hybridant dans l'aptamère pour rapporter la liaison du ligand. L'utilisation de cette seule sonde est justifiée par l'hypothèse qu'une sonde s'hybridant dans le pseudonoeud, ou la tige P13, pourrait empêcher une formation partielle du pseudonoeud par compétition entre la sonde et l'ARN étudié. Alors, nous n'aurions pas pu déterminer de différence au niveau du  $K_{switch}$  entre un pseudonoeud faiblement formé et l'absence de ce dernier. En résumé, le résultat de cette expérience est que le pseudonoeud n'a pas réussi à se former entre le transcrite de l'aptamère et celui de la plateforme d'expression. En effet, la valeur de  $K_{switch}$  obtenue est du même ordre de grandeur que la valeur de l'aptamère seul, sans pseudonoeud, soit  $71 \pm 1 \mu\text{M}$ . (Figure 6B Lussier et al, manuscrit en préparation). Le contexte co-transcriptionnel semble donc nécessaire pour permettre la formation du pseudonoeud, comme suggéré par Perdrizet et al. (2012) (Perdrizet et al. 2012)

Finalement, à l'aide des essais  $\beta$ -galactosidase et de cartographie à la ribonucléase H, nous avons pu préciser, dans une certaine mesure, le rôle du pseudonoeud dans le fonctionnement du riborégulateur *btuB* d'*E. coli*. D'abord, nous avons montré, à l'aide des essais *in vivo* en  $\beta$ -galactosidase, que l'intégrité du pseudonoeud est essentielle pour la régulation génétique par le ligand. De plus, avec la cartographie *in vitro* à la ribonucléase H, nous avons aussi montré que

l'identité des nucléotides du pseudonoeud est importante pour obtenir un repliement de type sauvage. Ensuite, nous avons constaté que les régions de la tige P13 et du pseudonoeud sont importantes pour la liaison du ligand et l'expression génétique. En effet, la présence du pseudonoeud contribue grandement à la liaison du ligand et la stabilité de la tige P13 serait importante, non seulement pour la formation du pseudonoeud, mais aussi pour obtenir une expression génétique efficace. Enfin, nous avons suggéré que la formation de pseudonoeud nécessite le contexte transcriptionnel pour se former correctement. Par conséquent, nous pouvons affirmer que le pseudonoeud contribue à deux phénomènes essentiels, soit la liaison du ligand et le changement de conformation de la plateforme d'expression de manière ligand-dépendante. Ce qui veut dire qu'en plus de la tige P1, ce riborégulateur utilise un pseudonoeud pour modifier sa structure lors de la liaison du ligand et ainsi réprimer l'expression génétique.

## CONCLUSION

Dans le domaine des riborégulateurs, de plus en plus de structures et de ligands sont découverts chaque année. Puisque la diversité des riborégulateurs s'accroît, nous découvrons de nouvelles structures d'ARN qui peuvent à l'occasion se révéler très complexes. Pour déterminer le repliement d'un nouveau riborégulateur, il serait idéal de pouvoir en produire un cristal. Par contre, cette technique d'analyse structurale ne permet pas d'observer des changements de conformation entre les formes non liées et liées au ligand puisque la forme non liée n'est habituellement pas assez stable pour permettre une cristallisation. Alors, pour déterminer la structure d'un nouveau riborégulateur, il faut se baser sur les structures répertoriées chez des riborégulateurs semblables. Il devient donc important d'étudier le repliement des riborégulateurs connus dans le but d'en avoir un schéma de repliement complet qui pourrait aider à comprendre comment d'autres ARN peuvent se replier en présence de structures particulières comme les pseudonoeuds. Mon projet qui portait sur la détermination du rôle du pseudonoeud dans le repliement du riborégulateur *btuB* chez *E. coli*, avait pour but de servir cet objectif.

Ce riborégulateur n'est pas le seul à comporter une structure de type pseudonoeud, mais son rôle au niveau du repliement et son importance pour la régulation génétique n'était pas clairement expliquée. En combinant des essais *in vivo* et des essais de cartographies *in vitro*, et à l'aide de cristaux d'autres riborégulateurs coenzyme B<sub>12</sub> nous avons réussi à déterminer un double rôle pour ce pseudonoeud. Son premier rôle est de servir d'intermédiaire structural entre la formation de la tige P1 et de la tige séquestatrice P14. En effet, lorsque la tige P1 se forme, l'anti-anti-RBS est séquestré dans cette dernière, ce qui pourrait rapprocher la tige P5 et la tige P13 pour former le pseudonoeud. Le pseudonoeud pourrait alors promouvoir la formation de la tige séquestatrice du RBS (tige P14). Nous avons indirectement observé que le pseudonoeud est essentiel pour la formation de la tige P14, mais nous n'avons aucun modèle permettant d'expliquer clairement son action. Pour ce faire, il faudrait pouvoir cartographier la plateforme d'expression au nucléotide près avec des mutants du pseudonoeud pour observer les différences de repliement causées par ces mutations. Par contre, nous pouvons suspecter que la formation du pseudonoeud créerait un déplacement de nucléotide qui rendrait favorable la formation de la tige P14.

Le deuxième rôle de ce pseudonoeud consiste à participer de manière importante à la liaison de l'AdoCbl au riborégulateur. L'effet du pseudonoeud lors de la liaison du ligand a été démontré à l'aide de plusieurs constructions, dont l'une où le pseudonoeud et la plateforme d'expression étaient absents. Cette construction, ne comprenant que l'aptamère, liait le ligand de manière très faible ne suggérant même aucune liaison du ligand, si l'on se réfère aux concentrations d'AdoCbl physiologiques (Fowler et al. 2010). De plus, l'intégrité de la tige P13 est aussi importante pour la liaison du ligand puisque, si cette dernière n'est pas complète, la liaison du ligand en est affectée. À l'aide de ces résultats et des structures cristallines disponibles, nous suggérons que le pseudonoeud pourrait interagir avec le ligand dans une certaine mesure et alors participer à sa liaison dans l'aptamère. Ce résultat était assez inattendu puisque de manière habituelle, on peut soustraire la plateforme d'expression à un riborégulateur sans grandement affecter sa capacité à lier son ligand. Par contre, ici, le pseudonoeud est considéré en partie dans la plateforme d'expression et sa localisation est mal définie dans la littérature. Alors, à la lumière

de ces résultats, je crois qu'il est juste d'affirmer que le pseudonoeud et la tige P13 font partie de l'aptamère qui se définit comme le domaine minimal permettant la liaison du ligand. La plateforme d'expression quant à elle ne serait constituée que de la tige P14. Enfin, nous avons aussi montré que la formation du pseudonoeud ne pouvait se produire en utilisant un ARN de la plateforme d'expression en *trans*, ajoutant ainsi aux travaux de Perdrizet et al (Perdrizet et al. 2012), qui soulignent l'importance du contexte co-transcriptionnel pour le repliement de ce riborégulateur.

Finalement, même si nous ne sommes pas parvenus à établir un modèle de repliement précis pour le pseudonoeud, nous avons pu établir deux rôles précis pour ce dernier. Ces deux rôles, qui sont l'aide à la liaison du ligand et la régulation génétique, permettent d'expliquer l'importance de cette structure pour le bon fonctionnement du riborégulateur. Par contre, cette combinaison de fonctions particulières semble, à ce jour, limitée aux riborégulateurs B<sub>12</sub>. Il y a bien entendu des pseudonoeuds chez d'autres riborégulateurs, mais, à ce jour, ils ne semblent pas avoir ce double rôle retrouvé chez *btuB*. Dans le premier cas, ils servent à créer un site catalytique et pour le second, il fait partie intégrante du riborégulateur. Ces rôles sont donc assez éloignés de ceux retrouvés chez le riborégulateur *btuB*. Beaucoup de riborégulateurs B<sub>12</sub> retrouvés par alignement de séquence pourraient posséder un pseudonoeud localisé dans la même région que celui d'*E. coli*. Ils pourraient alors avoir des rôles semblables et il se pourrait que de nouvelles classes de riborégulateur comportent des pseudonoeuds remplissant les mêmes rôles que ceux mis en lumière lors de mon projet.

## BIBLIOGRAPHIE

- A Darst, S., 2001. Bacterial RNA polymerase. *Current Opinion in Structural Biology*, 11, pp.155–162.
- Barrick, J.E. & Breaker, R.R., 2007. The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome biology*, 8(11), p.R239.
- Batey, R.T., Recognition of S-adenosylmethionine by riboswitches. *Wiley interdisciplinary reviews. RNA*, 2(2), pp.299–311.
- Batey, R.T., 2012. Structure and mechanism of purine binding riboswitches. *Q Rev Biophys*, 45, pp.345–381.
- Blackwell, C.M. & Turner, J.M., 1978. Microbial metabolism of amino alcohols. Formation of coenzyme B12-dependent ethanolamine ammonia-lyase and its concerted induction in *Escherichia coli*. *The Biochemical journal*, 176, pp.751–757.
- Bocobza, S. et al., 2007. Riboswitch-dependent gene regulation and its evolution in the plant kingdom. *Genes & development*, 21(22), pp.2874–9.
- Boudvillain, M., Figueroa-Bossi, N. & Bossi, L., 2013. Terminator still moving forward: expanding roles for Rho factor. *Current opinion in microbiology*, 16(2), pp.118–24.
- Bradbeer, C., 1993. The proton motive force drives the outer membrane transport of cobalamin in *Escherichia coli*. *Journal of bacteriology*, 175(10), pp.3146–50.
- Breaker, R.R., 2011. Prospects for riboswitch discovery and analysis. *Molecular cell*, 43(6), pp.867–79.
- Butzin, N.C. et al., 2013. *Thermotoga lettingae* can salvage cobinamide to synthesize vitamin B12. *Applied and environmental microbiology*, 79(22), pp.7006–12.
- Canady, M.A. et al., 1996. Crystal structure of turnip yellow mosaic virus. *Nature Structural Biology*, 3(9), pp.771–781.



- 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..
- Chae, H. et al., 2011. Rho-dependent termination of *ssrS* (6S RNA) transcription in *Escherichia coli*: Implication for 3' processing of 6S RNA and expression of downstream *ygfA* (putative 5-formyl-tetrahydrofolate cyclo-ligase). *Journal of Biological Chemistry*, 286, pp.114–122.
- Cheah, M.T. et al., 2007. Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature*, 447(7143), pp.497–500.
- Choudhary, P.K. & Sigel, R.K.O., 2014. Mg(2+)-induced conformational changes in the *btuB* riboswitch from *E. coli*. *RNA (New York, N.Y.)*, 20(1), pp.36–45.
- Croft, M.T. et al., 2007. Thiamine biosynthesis in algae is regulated by riboswitches. *Proceedings of the National Academy of Sciences of the United States of America*, 104(52), pp.20770–5.
- DebRoy, S. et al., 2014. Riboswitches. A riboswitch-containing sRNA controls gene expression by sequestration of a response regulator. *Science (New York, N.Y.)*, 345(6199), pp.937–40.
- Epshtein, V. et al., 2010. An allosteric mechanism of Rho-dependent transcription termination. *Nature*, 463, pp.245–249.
- Feklistov, A. et al., 2006. A Basal Promoter Element Recognized by Free RNA Polymerase  $\sigma$  Subunit Determines Promoter Recognition by RNA Polymerase Holoenzyme. *Molecular Cell*, 23, pp.97–107.
- Feklistov, A., 2014. RNA polymerase : in search of promoters.
- Forage, R.G. & Foster, M.A., 1982. Glycerol fermentation in *Klebsiella pneumoniae*: Functions of the coenzyme B12-dependent glycerol and diol dehydratases. *Journal of Bacteriology*, 149, pp.413–419.

- Fowler, C.C., Brown, E.D. & Li, Y., 2010. Using a riboswitch sensor to examine coenzyme B(12) metabolism and transport in *E. coli*. *Chemistry & biology*, 17(7), pp.756–65.
- Furth, J., Hurwitz, J. & Anders, M., 1962. The Role of Deoxyribonucleic Acid in Ribonucleic Acid Synthesis: I. THE PURIFICATION AND PROPERTIES OF RIBONUCLEIC ACID POLYMERASE. *The Journal of biological chemistry*.
- Garst, A.D. & Batey, R.T., 2009. A switch in time: Detailing the life of a riboswitch. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1789, pp.584–591.
- Gautheret, D. & Cedergren, R., 1993. Modeling the three-dimensional structure of RNA. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 7, pp.97–105.
- Gilbert, S.D. et al., 2008. Structure of the SAM-II riboswitch bound to S-adenosylmethionine. *Nature structural & molecular biology*, 15, pp.177–182.
- González, J.C. et al., 1992. Comparison of cobalamin-independent and cobalamin-dependent methionine synthases from *Escherichia coli*: two solutions to the same chemical problem. *Biochemistry*, 31, pp.6045–6056.
- Haentjens-Sitri, J. et al., 2008. A Competition Mechanism Regulates the Translation of the *Escherichia coli* Operon Encoding Ribosomal Proteins L35 and L20. *Journal of Molecular Biology*, 375, pp.612–625.
- Han, K. & Byun, Y., 2003. PSEUDOVIEWER2: Visualization of RNA pseudoknots of any type. *Nucleic Acids Research*, 31, pp.3432–3440.
- Henkin, T., 1996. Control of transcription termination in prokaryotes. *Annual review of genetics*, (57), pp.35–57.
- Hollands, K. et al., 2012. Riboswitch control of Rho-dependent transcription termination. *Proceedings of the National Academy of Sciences*, 109(14), pp.5376–5381.
- Hsu, L.M., 2002. Promoter clearance and escape in prokaryotes. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression*, 1577(2), pp.191–207.

- Johnson, J.E. et al., 2012. B12 cofactors directly stabilize an mRNA regulatory switch. *Nature*, 492(7427), pp.133–7.
- Kadner, R.J., 1998. Coupled Changes in Translation and Transcription during Cobalamin-Dependent Regulation of *btuB* Expression in *Escherichia coli*. , 180(24), pp.6719–6728.
- Kang, M., Eichhorn, C.D. & Feigon, J., 2014. Structural determinants for ligand capture by a class II preQ1 riboswitch. *Proceedings of the National Academy of Sciences of the United States of America*, 111, pp.E663–71.
- Kang, M., Peterson, R. & Feigon, J., 2009. Structural Insights into Riboswitch Control of the Biosynthesis of Queuosine, a Modified Nucleotide Found in the Anticodon of tRNA. *Molecular cell*, 33(6), pp.784–90.
- Kim, J.N. & Breaker, R.R., 2008. Purine sensing by riboswitches *Biology of the Cell*. , 100, pp.1–11.
- Kim, N. et al., 2004. Candidates for novel RNA topologies. *Journal of Molecular Biology*, 341, pp.1129–1144.
- Klein, D.J., Edwards, T.E. & Ferré-D'Amaré, A.R., 2009. Cocrystal structure of a class I preQ1 riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. *Nature structural & molecular biology*, 16, pp.343–344.
- Klein, D.J. & Ferré-D'Amaré, A.R., 2006. Structural basis of *glmS* ribozyme activation by glucosamine-6-phosphate. *Science (New York, N.Y.)*, 313(5794), pp.1752–6.
- Koduvayur, S.P. & Woodson, S.A., 2004. Intracellular folding of the *Tetrahymena* group I intron depends on exon sequence and promoter choice. *RNA (New York, N.Y.)*, 10, pp.1526–1532.
- Koslover, D.J. et al., 2012. Binding and translocation of termination factor Rho studied at the single-molecule level. *Journal of Molecular Biology*, 423, pp.664–676.
- Lafontaine, D.A., 2007. Ligand recognition determinants of guanine riboswitches. , 35(16), pp.5568–5581.

- Loh, E. et al., 2009. A trans-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. *Cell*, 139(4), pp.770–9.
- Lundrigan, M.D. & Kadner, R.J., 1989. Altered cobalamin metabolism in *Escherichia coli* btuR mutants affects btuB gene regulation. *Journal of Bacteriology*, 171, pp.154–161.
- Mandal, M. et al., 2003. Riboswitches Control Fundamental Biochemical Pathways in *Bacillus subtilis* and Other Bacteria. *Cell*, 113(5), pp.577–586.
- Mandal, M. & Breaker, R.R., 2004. Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nature structural & molecular biology*, 11(1), pp.29–35..
- Margeat, E. et al., 2006. Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes. *Biophysical journal*, 90(4), pp.1419–31.
- McPheeters, D.S., Stormo, G.D. & Gold, L., 1988. Autogenous regulatory site on the bacteriophage T4 gene 32 messenger RNA. *Journal of molecular biology*, 201, pp.517–535.
- Mellin, J.R. et al., 2013. A riboswitch-regulated antisense RNA in *Listeria monocytogenes*. *Proceedings of the National Academy of Sciences of the United States of America*, 110(32), pp.13132–7.].
- Nahvi, A. et al., 2002. Genetic Control by a Metabolite Binding mRNA. , 9(02), pp.1043–1049.
- Nou, X. & Kadner, R.J., 2000. Adenosylcobalamin inhibits ribosome binding to btuB RNA. *Proceedings of the National Academy of Sciences of the United States of America*, 97(13), pp.7190–5.
- Park, J.-S. & Roberts, J.W., 2006. Role of DNA bubble rewinding in enzymatic transcription termination. *Proceedings of the National Academy of Sciences of the United States of America*, 103, pp.4870–4875.
- Pedrolli, D. et al., 2015. The ribB FMN riboswitch from *Escherichia coli* operates at the transcriptional and translational level and regulates riboflavin biosynthesis. *FEBS Journal*, p.n/a–n/a..

- Perdrizet, G.A. et al., 2012. Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. *Proceedings of the National Academy of Sciences of the United States of America*, 109(9), pp.3323–8..
- Perreault, J. et al., 2011. Identification of Hammerhead Ribozymes in All Domains of Life Reveals Novel Structural Variations. *PLoS Computational Biology*, 7.
- Peselis, A. & Serganov, A., 2012. Structural insights into ligand binding and gene expression control by an adenosylcobalamin riboswitch. *Nature structural & molecular biology*, 19(11), pp.1182–4.
- Peselis, A. & Serganov, A., 2014a. Structure and function of pseudoknots involved in gene expression control. *Wiley interdisciplinary reviews. RNA*, 5(6), pp.803–22.
- Peselis, A. & Serganov, A., 2014b. Themes and variations in riboswitch structure and function. *Biochimica et biophysica acta*, 1839(10), pp.908–918..
- Pleij, C.W.A., Rietveld, K. & Bosch, L., 1985. A new principle of RNA folding based on pseudoknotting. *Nucleic Acids Research*, 13, pp.1717–1731.
- Rentmeister, A. et al., 2007. Conformational changes in the expression domain of the Escherichia coli thiM riboswitch. *Nucleic acids research*, 35(11), pp.3713–22.
- Revyakin, A. et al., 2006. Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science (New York, N.Y.)*, 314(5802), pp.1139–43.
- Richardson, L. V. & Richardson, J.P., 1996. Rho-dependent termination of transcription is governed primarily by the upstream Rho utilization (rut) sequences of a terminator. *Journal of Biological Chemistry*, 271, pp.21597–21603.
- Rietveld, K. et al., 1982. The tRNA-Uke structure at the 3' terminus of turnip yellow mosaic virus RNA. Differences and similarities with canonical tRNA. *Nucleic Acids Research*, 10, pp.1929–1946.
- Robb, N.C. et al., 2013. The transcription bubble of the RNA polymerase-promoter open complex exhibits conformational heterogeneity and millisecond-scale dynamics:

- Implications for transcription start-site selection. *Journal of Molecular Biology*, 425, pp.875–885.
- Rodionov, D. a et al., 2003. Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes. *The Journal of biological chemistry*, 278(42), pp.41148–59.
- 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.
- Roth, J.R., Lawrence, J.G. & Bobik, T.A., 1996. COBALAMIN ( COENZYME B 12 ): Synthesis and Biological Significance.
- Serganov, A. et al., 2002. Do mRNA and rRNA binding sites of E.coli ribosomal protein S15 share common structural determinants? *Journal of molecular biology*, 320, pp.963–978.
- Serganov, A. et al., 2006. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature*, 441(7097), pp.1167–71.
- Serganov, A. & Nudler, E., 2013. A decade of riboswitches. *Cell*, 152(1-2), pp.17–24.
- Shamoo, Y. et al., 1993. Translational repression by the bacteriophage T4 gene 32 protein involves specific recognition of an RNA pseudoknot structure. *Journal of molecular biology*, 232, pp.89–104.
- Slonczewski, J.L. et al., 1981. pH homeostasis in Escherichia coli: measurement by <sup>31</sup>P nuclear magnetic resonance of methylphosphonate and phosphate. *Proceedings of the National Academy of Sciences of the United States of America*, 78(10), pp.6271–5..
- Spitale, R.C. et al., 2009. The structural basis for recognition of the preQ0 metabolite by an unusually small riboswitch aptamer domain. *Journal of Biological Chemistry*, 284, pp.11012–11016.
- Sudarsan, N. et al., 2003. An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes & development*, 17(21), pp.2688–97.
- Sun, E.I. et al., 2013. Comparative genomics of metabolic capacities of regulons controlled by cis-regulatory RNA motifs in bacteria. *BMC genomics*, 14(1), p.597.

- Susa, M., Kubori, T. & Shimamoto, N., 2006. A pathway branching in transcription initiation in *Escherichia coli*. *Molecular microbiology*, 59(6), pp.1807–17.
- Tang, C.K. & Draper, D.E., 1989. Unusual mRNA pseudoknot structure is recognized by a protein translational repressor. *Cell*, 57, pp.531–536.
- Tinsley, R.A., Furchak, J.R.W. & Walter, N.G., 2007. Trans-acting glmS catalytic riboswitch: locked and loaded. *RNA (New York, N.Y.)*, 13(4), pp.468–77.
- Toraya, T., Honda, S. & Fukui, S., 1979. Fermentation of 1,2-propanediol and 1,2-ethanediol by some genera of Enterobacteriaceae, involving coenzyme B12-dependent diol dehydratase. *Journal of Bacteriology*, 139, pp.39–47.
- Wachter, A. et al., 2007. Riboswitch control of gene expression in plants by splicing and alternative 3' end processing of mRNAs. *The Plant cell*, 19(11), pp.3437–50.
- Warren, M.J. et al., 2002. The biosynthesis of adenosylcobalamin (vitamin B12). *Natural Product Reports*, 19(4), pp.390–412.
- Webb, C.-H.T. et al., 2009. Widespread occurrence of self-cleaving ribozymes. *Science (New York, N.Y.)*, 326, p.953.
- Wickiser, J.K. et al., 2005. The Speed of RNA Transcription and Metabolite Binding Kinetics Operate an FMN Riboswitch. , 18, pp.49–60.
- Wilson, K.S. & von Hippel, P.H., 1995. Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proceedings of the National Academy of Sciences of the United States of America*, 92(19), pp.8793–7.
- Winkler, W.C. et al., 2003. An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nature structural biology*, 10(9), pp.701–7.
- Winkler, W.C. & Breaker, R.R., 2003. Genetic control by metabolite-binding riboswitches. *ChemBioChem*, 4(10), pp.1024–1032..

Yamada, K., 2013. Cobalt: its role in health and disease. *Metal ions in life sciences*, 13, pp.295–320.



