



HAL
open science

**DISSECTION DU MÉCANISME DE
TERMINAISON/ANTITERMINAISON AU NIVEAU
DU TERMINATEUR TRI DU PHAGE LAMBDA :
APPLICATION A L'ÉTUDE DES COMPLEXES
ARN-PROTÉINE IN VIVO**

Erwann Vieu

► **To cite this version:**

Erwann Vieu. DISSECTION DU MÉCANISME DE TERMINAISON/ANTITERMINAISON AU NIVEAU DU TERMINATEUR TRI DU PHAGE LAMBDA : APPLICATION A L'ÉTUDE DES COMPLEXES ARN-PROTÉINE IN VIVO. domain_other. Université d'Orléans, 2004. Français. NNT: . tel-00009769v1

HAL Id: tel-00009769

<https://theses.hal.science/tel-00009769v1>

Submitted on 25 Jul 2005 (v1), last revised 5 Aug 2005 (v2)

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

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



THESE

Présentée

A L'UNIVERSITE D'ORLEANS

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE D'ORLEANS

Discipline : Aspect moléculaire et cellulaire de la Biologie

PAR

Erwann VIEU

DISSECTION DU MECANISME DE
TERMINAISON/ANTITERMINAISON AU NIVEAU DU
TERMINATEUR TR1 DU PHAGE LAMBDA : APPLICATION A
L'ETUDE DES COMPLEXES ARN-PROTEINE *IN VIVO*.

Soutenu le xx septembre 2004

Devant le jury composé de :

- Dr. Lionello BOSSI	Rapporteur
- Dr. Annie KOLB	Rapporteur
- Pr. Tobias HEVOR	Examineur
- Pr. Jacques PAOLETTI	Examineur
- Dr. Rachid RAHMOUNI	Directeur de Thèse

AVANT PROPOS

Mon travail de thèse s'inscrit dans l'étude des signaux régulant l'élongation de la transcription chez *Escherichia Coli* (*E. Coli*), il a été réalisé au sein du laboratoire "Mécanismes moléculaires de la régulation des gènes" sous la direction du Dr. A. Rachid Rahmouni, au Centre de Biophysique Moléculaire, et mené à bien grâce à une allocation de la Ligue Nationale Contre Le Cancer.

Avant de débiter, je souhaite dédier ces quelques lignes au bactériophage lambda (λ , [Figure 1](#)) grand artisan du développement de la biologie moléculaire au cours des cinquante dernières années. En effet, lorsqu'on parcourt les différentes revues consacrées à λ on réalise combien l'étude de ce virus bactérien est intimement liée à l'histoire et aux grandes découvertes qui ont fait notre jeune discipline.

L'étude du bactériophage λ a fourni les clés de nombreux mécanismes biologiques fondamentaux. Ses deux modes de développement, lytique et lysogénique, sa façon d'utiliser son hôte *E. Coli* pour se développer, ont fait du phage λ un modèle pour comprendre un grand nombre de processus biologiques : l'isolation de la ligase (et la capacité de cloner des fragments d'ADN), la régulation positive ou négative des gènes, la recombinaison, la réplication de l'ADN... Toutes ces connaissances découlent de l'étude, durant les cinquante dernières années, de la génétique du phage λ (pour revue : [Gottesman, 1999](#)). Aussi, c'est en utilisant l'ADN de λ comme substrat que Roberts découvre, en 1969, le facteur de terminaison bactérien Rho, ainsi que le phénomène d'antiterminaison ([Roberts, 1969](#)).

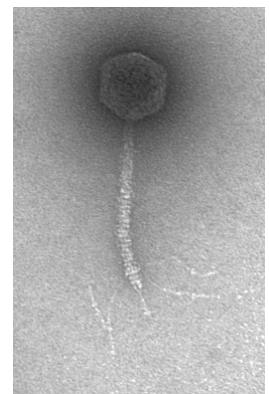


Figure 1 : Photographie du bactériophage λ en microscopie électronique.

Mais le bactériophage λ ne constitue pas seulement un modèle ayant aidé à appréhender de nombreux processus biologiques, il permet, encore actuellement, de révéler de nouveaux principes en biologie moléculaire. Ainsi, au cours de la première partie de mon travail de thèse, je me suis consacré à l'étude du phénomène d'antiterminaison intervenant au niveau du terminateur tR1 du phage λ , l'un des terminateurs les plus utilisés pour étudier le fonctionnement du facteur de terminaison Rho. Les résultats obtenus au cours de ce travail

Avant propos.

nous ont, dans un second temps, permis de développer un système pour cribler, directement dans la cellule, des interactions entre ARN et protéines.

Afin d'introduire ces travaux, je commencerai par rappeler les différentes étapes de la transcription chez *E. Coli*, en décrivant plus particulièrement les mécanismes permettant le relargage de l'ARN en fin de transcription. Je développerai certaines stratégies utilisées par les bactéries, ou les phages, pour moduler ce processus de terminaison. Je décrirai plus précisément l'antiterminaison N-dépendante chez le phage λ et les raisons pour lesquelles nous nous sommes intéressés à ce phénomène. Je développerai enfin les systèmes permettant d'analyser les interactions ARN-protéine impliquées, entre autre, dans les mécanismes de régulation précédemment décrits, et les raisons qui nous ont poussés à développer notre propre système.

Je conclurai en replaçant dans la discussion les résultats obtenus durant cette thèse.

Note : Afin de faciliter la lecture, les noms des gènes ou des séquences ADN seront représentés en *minuscules italiques*. Les éléments ou les motifs ARN seront représentés en MAJUSCULES.

TABLE DES MATIERES.

INTRODUCTION

- page 1 -

Chapitre I. LA TRANSCRIPTION. 1

1. De l'ADN à l'ARN : La transcription chez <i>E. Coli</i> .	1
2. Réguler l'élongation de la transcription.	2
2. 1. Les signaux de pause.	3
2. 2. Les signaux d'arrêt.	4
2. 3. Les signaux de terminaison.	4
3. La terminaison intrinsèque.	5
4. La terminaison facteur dépendante.	8
4. 1. TRCF.	9
4. 2. Rho.	9

Chapitre II. RHO. 10

1. Caractéristiques des terminateurs Rho-dépendants.	10
1. 1. La séquence rut.	12
1. 2. Les sites de terminaison.	14
2. Organisation structurale.	16
3. Quand Rho rencontre l'ARN.	18
4. Activité ATPase.	26
5. Activité hélicase.	27
6. NusG : cofacteur de la terminaison Rho dépendante.	29
7. NusA : Antagoniste de Rho?	29
8. Résumé de l'action de Rho.	30
9. Rôles biologiques de Rho.	32
10. Un facteur Rho eucaryote?	33

Chapitre III. REGULATION DE LA TERMINAISON. 35

1. Atténuation.....	35
2. L'antiterminaison processive.	37
2. 1. Antiterminaison dans les opérons ribosomiaux.	38
2. 2. Antiterminaison chez le phage λ	39
2. 2. 1. <i>Antiterminaison dirigée par N.</i>	40
2. 2. 2. <i>Antiterminaison dirigée par le facteur Q.</i>	47
2. 3. Terminaison et antiterminaison chez le phage HK022.....	48
2. 4. Antiterminaison chez les eucaryotes.....	49
3. Importance des interactions ARN-protéine dans la régulation des gènes.....	50

Chapitre IV. LES INTERACTIONS ARN-PROTEINE. 51

1. Des ARNs comme agents thérapeutiques?	51
1. 1. Inhibition de l'expression des gènes par un ARN.....	52
1. 2. Réparation d'une anomalie génétique grâce à un ARN.....	54
1. 3. Des ARNs inhibiteurs des fonctions protéiques.	54
2. Sélection et analyse des interactions ARN-protéine <i>in vitro</i>	55
2. 1. La méthode SELEX.....	55
2. 2. La méthode T7 phage display.....	56
3. Sélection et analyse des interactions ARN-protéine <i>in vivo</i>	58
3. 1. Activation de la transcription : Le système triple hybride.....	58
3. 2. Répression de la traduction par un complexe ribonucléoprotéique.....	60
3. 2. 1. <i>Système procaryote.</i>	60
3. 2. 2. <i>Système eucaryote.</i>	62
3. 3. Sélection de complexes ARN-protéine par antiterminaison de la transcription.....	63
3. 3. 1. <i>Système procaryote.</i>	63
3. 3. 2. <i>Système eucaryote.</i>	64
4. Intérêt d'un nouveau système de sélection?	64

RESULTATS

- page 66 -

Chapitre V. ARTICLE. 66

1. Introduction à l'article "Dual role of boxB RNA motif in the mechanisms of termination/antitermination at the lambda tR1 terminator revealed in vivo".	66
2. Article.	69
3. Conclusion.	80

Chapitre VI. MISE AU POINT D'UN SYSTEME DE SELECTION D'APTAMERS IN VIVO. 81

1. Hypothèse de départ.	82
1. 1. Le terminateur tR1 n'est pas adapté à l'élaboration du crible.	82
1. 2. Description de la séquence d'entrée synthétique.	83
2. Principe du système.	85
3. Design du système.	87
3. 1. Plasmide rapporteur.	87
3. 2. Plasmide protéique.	88
4. Résultats préliminaires.	88
5. Optimisation du système.	90
5. 1. Changement de système rapporteur.	90
5. 2. Résultats.	92
5. 3. Amélioration de l'excision de l'ARNt.	94
5. 4. Nouveaux couples ARN-protéine testés.	94
5. 5. Optimisation du site d'insertion du motif d'intérêt.	96
6. Résultats obtenus avec les nouvelles constructions.	96

DISCUSSION-PERSPECTIVE

- page 99 -

1. Un modèle complémentaire de l'antitermination au niveau de tR1.....	99
2. Un crible génétique pour étudier les interactions ARN-protéine <i>in vivo</i>	102
3. Autres utilisations du système génétique.....	104

MATERIEL ET METHODES

- page 108 -

REFERENCES

- page 114 -

INTRODUCTION

1. De l'ADN à l'ARN : La transcription chez *E. coli*.

La transcription constitue la première étape de la transmission de l'information génétique. Ce passage de l'ADN vers l'ARN est exécuté par des enzymes connues sous le nom d'ARN polymérases. Chez *E. coli* l'ARN polymérase est composée de cinq sous-unités : α_2 , β , β' , ω , pour une masse totale de 450kDa (Burgess et Travers, 1970), et le cycle transcriptionnel peut être délimité en quatre phases bien distinctes (Figure 2) :

- La reconnaissance de la séquence promotrice.
- L'initiation de la transcription.
- L'élongation de la molécule d'ARN.
- La terminaison et le relargage de la molécule d'ARN.

Chacune de ces étapes représente une cible potentielle pour des signaux de régulation.

Le processus de transcription débute lorsque l'ARN polymérase, s'associe à un facteur d'initiation de la famille sigma (σ), pour reconnaître spécifiquement, sur l'ADN, une série de motifs qui compose le promoteur (pour revue : de Haseth et al., 1998). L'ARN polymérase déroule l'ADN sur un segment d'environ 17 paires de bases (pb) qui constitue la bulle de transcription. Cette étape lui permet d'accéder à l'ADN simple brin qui doit lui servir de matrice (de Haseth et Helmann, 1995; Saeker et al., 2002). La transcription débute alors avec la formation de la première liaison phosphodiester et l'incorporation des premiers nucléotides. Au cours de cette phase, l'ARN polymérase se trouve sous une forme peu stable pouvant aboutir à la synthèse et au relargage de transcrits ARN courts (2 à 8 nucléotides), un phénomène nommé initiation abortive (Chamberlin et Hsu, 1996; Vo et al, 2003).

Si la polymérisation se poursuit, l'ARN polymérase quitte la phase d'initiation pour entrer dans la phase d'élongation. Cette conversion se caractérise par :

- La perte du facteur σ , qui est éventuellement recyclé pour d'autres cycles transcriptionnels (Hansen et McClure, 1980; Stackhouse et al., 1989).

- Le déplacement de l'ARN polymérase hors de la séquence promotrice, accompagné d'une perte des interactions avec l'ADN qui compose cette séquence (Krummel et Chamberlin, 1989).
- Un changement conformationnel de l'ARN polymérase vers une forme hautement stable et processive (Krummel et Chamberlin, 1992).

Le complexe ternaire, ARN polymérase-ADN-ARN, est alors capable de synthétiser à une vitesse de 30 à 100 nucléotides/sec (parfois plus; Mooney et al., 1998; Vogel et Jensen, 1994) des transcrits extrêmement longs (jusqu'à 20 000 nucléotides), permettant ainsi la transcription d'opérons entiers (Levin et al., 1987).

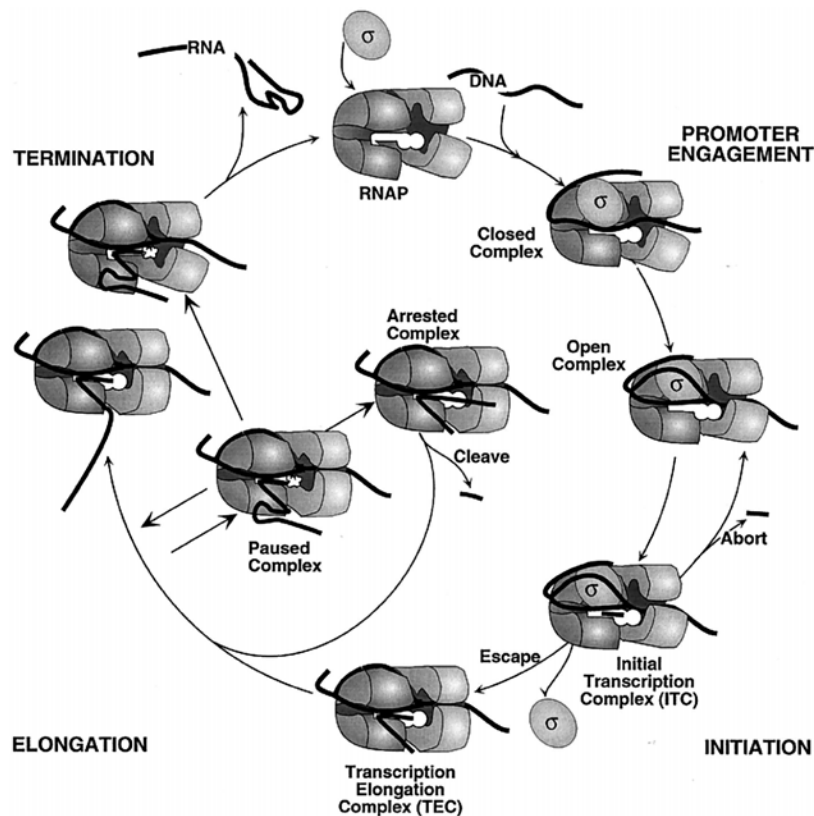


Figure 2 : Le cycle transcriptionnel chez *E. coli* incluant les différents signaux de régulation de la phase d'élongation (d'après : Mooney et al., 1998).

2. Réguler l'élongation de la transcription.

L'extrême stabilité des complexes d'élongation implique l'existence de signaux spécifiques permettant de moduler la processivité de ces complexes (Mooney et al., 1998).

Ces signaux sont :

- Les signaux de pauses où les complexes effectuent une pause plus ou moins longue en cours d'élongation avant de reprendre la transcription.
- Les signaux d'arrêt où la ARN polymérase est stoppée et nécessite l'intervention de facteurs extérieurs pour reprendre la transcription.
- Enfin les signaux de terminaison qui provoquent la dissociation du complexe ternaire d'élongation et le relargage de la ARN polymérase et du transcrit.

2. 1. Les signaux de pause.

Le phénomène de pause transcriptionnelle peut intervenir au niveau de sites spécifiques : codant des motifs ARN structurés (site de classe I) ou induisant l'apparition d'un hybride ADN:ARN instable (site de classe II) (Artsimovitch et Landick, 2000). Ces signaux provoquent une série de réarrangements dans le site actif du complexe ternaire (rupture de l'alignement entre le 3'OH du transcrit et le nucléotide entrant; Touloukhonov et Landick, 2003; Figure 3). Le complexe d'élongation passe alors de sa forme hautement processive vers une forme alternative (réversible) dont la capacité d'extension du transcrit est fortement réduite (d'un facteur 10^2 à 10^4). Ce mécanisme peut être modulé par certains facteurs externes (Artsimovitch et Landick, 2000 et chapitre III).

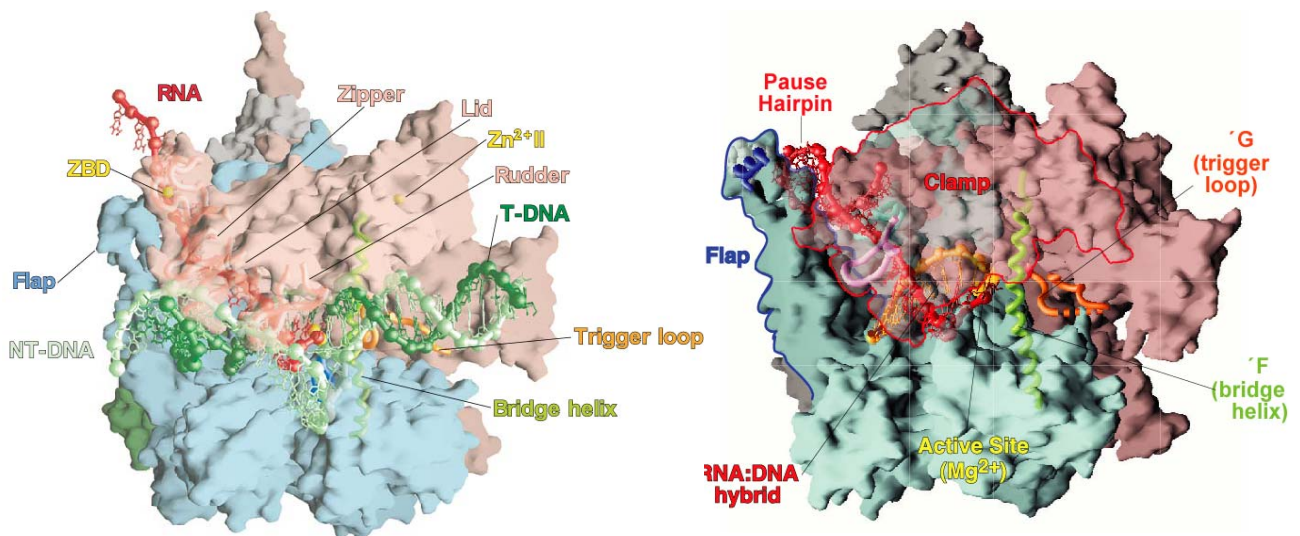


Figure 3 : Comparaison du complexe ternaire en phase d'élongation et en état de pause (d'après : Touloukhonov et Landick, 2003).

Les signaux de pause servent essentiellement à synchroniser la transcription et la traduction chez les procaryotes (Gong et Yanofsky, 2002; Landick et al., 1985). Cette fonction permet d'éviter l'apparition, entre l'ARN polymérase et les ribosomes, d'ARN nu sur lequel viendraient se fixer des protéines capables de moduler la transcription ou la stabilité de l'ARNm. Cette synchronisation est utilisée dans un grand nombre d'opérons bactériens pour rendre l'ARN polymérase sensible à des signaux dit d'atténuation (Yanofsky, 2000; Landick et al., 1996; et chapitre III). De plus les signaux de pause, en modulant la structure et la vitesse de l'ARN polymérase, peuvent la rendre plus sensible à l'action de structures ou facteurs régulateurs. Cette chute de processivité de l'ARN polymérase, caractéristique du phénomène de pause, constitue apparemment la première étape de l'arrêt et de la terminaison de la transcription (Artsimovitch et Landick, 2000).

2. 2. Les signaux d'arrêt.

L'arrêt de la transcription, contrairement à la pause, est un phénomène irréversible. Ce mécanisme n'a cependant été observé qu'*in vitro* (Davenport et al., 2000). On suppose que ce phénomène trouve sa signification biologique lors d'une carence en nucléosides triphosphates (NTP), lors d'une erreur d'incorporation ou encore lorsque l'ARN polymérase est bloquée par un obstacle protéique. Des études réalisées *in vitro* semblent indiquer que, dans ce type de situation, l'ARN polymérase effectue un glissement vers l'arrière permettant l'extraction de l'extrémité 3' du transcrit (Landick, 1997; von Hippel, 1998; Nudler, 1999). Ce retour en arrière rend l'ARN polymérase incapable d'assurer l'élongation du transcrit. Elle peut cependant reprendre la transcription en provoquant l'hydrolyse du fragment d'ARN extrudé (Orlova et al., 1995). Des travaux, réalisés entre autre au laboratoire, ont montré que les facteurs d'élongation GreA et GreB favorisaient largement cette hydrolyse *in vitro* (Borukhov et al., 1993) et *in vivo* (Toulmé et al., 2000).

2. 3. Les signaux de terminaison.

C'est le troisième type de signaux de régulation de l'élongation qui va nous intéresser ici.

L'étude des signaux de terminaison est en effet primordiale, car tout comme la reconnaissance d'un promoteur et l'initiation, la terminaison de la transcription est étroitement modulée, et la faculté de relarguer proprement le transcrit constitue un outil de régulation extrêmement important pour l'expression du génome bactérien (pour revue: [Henkin et Yanofsky, 2002](#)).

Les terminateurs transcriptionnels ont trois fonctions dans l'expression des gènes :

- Tout d'abord, ils définissent la fin de l'unité de transcription d'un gène ou d'un groupe de gènes (opéron). Cela permet d'assurer l'indépendance transcriptionnelle de deux séquences d'ADN adjacentes.
- Les terminateurs peuvent également être trouvés entre les gènes exprimés, à partir d'un même promoteur, ou bien entre le promoteur et le premier gène de l'opéron. Ces terminateurs intergéniques permettent de moduler l'expression des gènes d'un même opéron, comme dans les phénomènes d'atténuation (voir : chapitre III).
- Enfin les terminateurs peuvent se trouver aussi au sein même des gènes. Ce sont alors des terminateurs latents qui ne fonctionnent que lorsque la transcription et la traduction sont désynchronisées par une mutation, ou lors de certaines formes de stress comme une carence en acides aminés ([Ruterhouser et Richardson, 1989](#)).

Chez *E. coli* la terminaison de la transcription peut intervenir selon deux mécanismes bien distincts. Elle peut être intrinsèque, ou nécessiter l'intervention d'un facteur protéique (terminaison facteur-dépendante). Dans les pages suivantes j'aborderai brièvement le mécanisme de la terminaison intrinsèque, puis je décrirai, cette fois plus en détail, la terminaison facteur-dépendante, et plus particulièrement le mécanisme complexe de la terminaison dépendante du facteur Rho, qui a été mon partenaire indirect pendant ces années de thèse.

3. La terminaison intrinsèque.

Comme je l'ai précisé précédemment ce mode de terminaison ne nécessite l'intervention d'aucun facteur extérieur. *In vitro* elle peut intervenir simplement en présence de NTP, de l'ARN polymérase et d'une matrice ADN contenant un terminateur intrinsèque. Un terminateur intrinsèque est caractérisé par une séquence d'ADN palindromique interrompue,

riche en bases G et C, et suivie d'une répétition de bases T (7 à 8 en général) (Brendel et Trifonov, 1984; Brendel et al., 1986). C'est la transcription de cette séquence par l'ARN polymérase qui conduit à la terminaison de la transcription. En effet, ce signal aboutit à la formation d'une tige boucle extrêmement stable au sein de l'ARN naissant (Figure 4). Cette structure est couplée, au sein de la bulle de transcription, à une région hybride ARN/ADN instable résultant de la présence d'une répétition d'uraciles en 3' de la tige boucle (Martin et Tinoco, 1980; Richardson et Greenblatt, 1996).

La stabilité du complexe de transcription en cours d'élongation repose sur une panoplie d'interactions au sein du complexe ternaire d'élongation (Nudler, 1999; Nudler et al., 1998). Ces interactions font intervenir trois domaines de l'ARN polymérase décrits récemment (pour revue : Gelles et Landick, 1998; von Hippel, 1998), et qui lui permettent d'interagir avec l'ADN double brin (DBS), l'hybride ARN/ADN (HBS) ou encore l'ARN néo-synthétisé (RBS). Le processus de terminaison intrinsèque réside alors dans la faculté du terminateur à perturber rapidement et de façon irréversible ces interactions. Le mécanisme précis permettant de briser le réseau d'interactions au sein du complexe ternaire d'élongation n'est cependant pas encore totalement caractérisé.

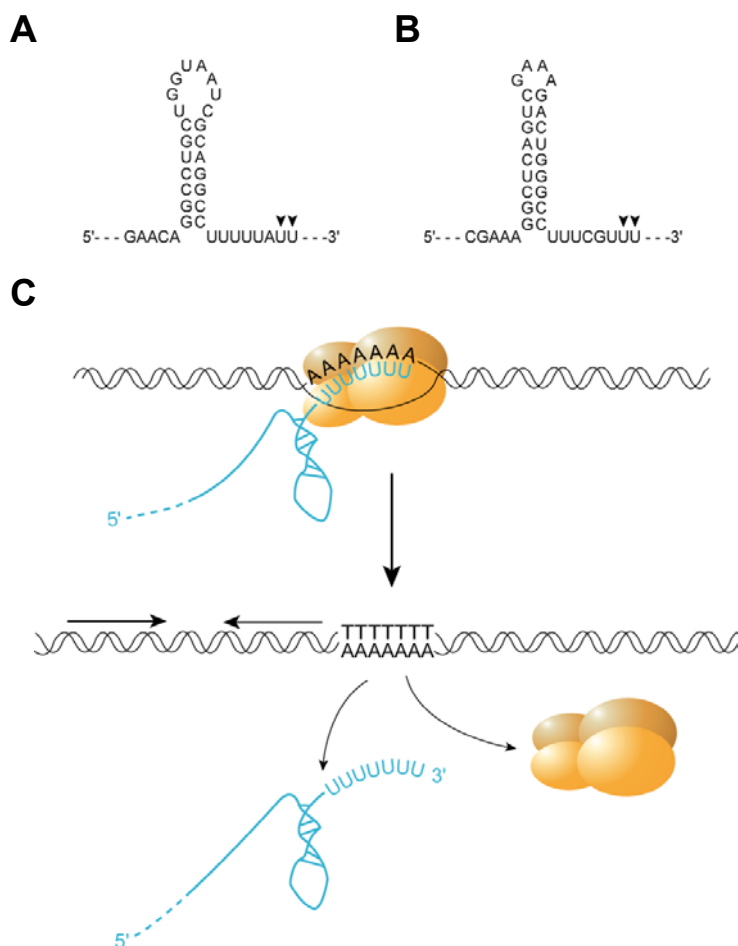


Figure 4 : Principe de la terminaison intrinsèque : Séquences de deux terminateurs intrinsèques naturels λ R2 (A) et *rrnBT2* (B) (les flèches indiquent les sites de terminaison). (C) Représentation schématique du mécanisme de terminaison intrinsèque.

Plusieurs modèles sont proposés pour expliquer le rôle de chaque élément du terminateur dans la dissociation du complexe d'élongation. Tous ces modèles s'accordent pour dire que la formation d'une structure en tige boucle stable, dans le canal de sortie de l'ARN au sein de l'ARN polymérase, déstabilise l'hybride ADN/ARN déjà affaibli par la forte concentration en paires d'ArU. Le motif en tige boucle entrerait en compétition avec la formation de certaines de ces paires d'ArU selon l'une des hypothèses ([Komissarova et al., 2002](#); [Gusarov et Nudler, 1999](#)), il provoquerait l'extraction d'une partie de l'hybride hors de la bulle de transcription selon un autre modèle ([Yarnell et Roberts, 1999](#)). Certaines études suggèrent que des interactions entre l'ARN néo-synthétisé et l'ARN polymérase interviendraient dans un changement allostérique de l'ARN polymérase favorisant le relargage de l'ARN ([Landick, 2001](#)). Par exemple, des travaux réalisés au laboratoire ont montré que des groupements fonctionnels des résidus du bras droit de l'hairpin sont impliqués dans des interactions avec l'ARN polymérase et favorisent le processus de terminaison ([MacDonald et al., 1993](#); [Schwartz et al., 2003](#)).

Bien que très importante pour l'efficacité de la terminaison, la structure ARN et la répétition de U ne sont apparemment pas les seuls éléments qui entrent en jeu dans la terminaison intrinsèque. Selon certains travaux les séquences ADN situées de part et d'autre du terminateur, y compris les séquences non transcrites en aval du terminateur, semblent jouer un rôle dans la terminaison. Ryder et Roberts ont récemment montré qu'en réduisant la réassociation des deux brins d'ADN à la sortie de la bulle de transcription (par des mutations sur le brin non codant), la terminaison était beaucoup moins efficace, voire pratiquement supprimée ([Ryder et Roberts, 2003](#)). Ce résultat met en évidence un couplage éventuel entre la dissociation de l'hybride ADN/ARN riche en AU et la réassociation des deux brins d'ADN dans le processus de terminaison intrinsèque. Ce mécanisme pourrait permettre d'accentuer le phénomène de terminaison. Cependant certains travaux semblent contredire cette hypothèse. En particulier, Uptain et Chamberlin ont montré que la terminaison intrinsèque pouvait être mise en place sur une matrice ADN simple brin ([Uptain et Chamberlin, 1997](#)).

La terminaison intrinsèque peut également être modulé par la présence d'une seconde ARN polymérase sur l'ADN. En effet un phénomène de coopérativité entre deux complexes d'élongation, initiés au niveau du même promoteur, a récemment été mis en évidence ([Epshtein et al., 2003](#)). Ce processus permet à la première polymérase de franchir un obstacle tel qu'un terminateur intrinsèque, ou un complexe ADN-protéine.

On peut noter que la formation d'une structure en tige boucle extrêmement stable en 3' du transcrit peut avoir une fonction autre que celle de terminer la transcription. En effet, contrairement à ce qui se passe chez les eucaryotes, chez les procaryotes l'extrémité 3' des ARNm est la conséquence directe de la terminaison de la transcription par l'ARN polymérase. Cette structure semble favoriser la stabilité de l'ARNm et diminuer sa sensibilité aux nucléases. Ainsi ce motif ARN, riche en GC, participerait à la régulation d'un gène, à la fois en assurant la terminaison correcte de la transcription de ce gène mais également en stabilisant son ARNm contre les dégradations (Cisneros et al., 1996).

Certains terminateurs intrinsèques présentent une séquence pouvant varier du consensus décrit précédemment, certains par exemple ne contiennent pas de répétitions d'uraciles (Unniraman et al., 2002). Toutefois, la relative conservation du motif des terminateurs intrinsèques a rendu possible la mise au point de logiciels algorithmiques permettant de prédire leurs positions au sein de la séquence génomique (d'Aubenton Carafa et al., 1990; Ermolaeva et al., 2000; Lesnik et al., 2001). Ces logiciels ont permis de déterminer que les signaux intrinsèques sont responsables de la terminaison de la moitié des unités de transcription codantes (ARNm) et de 70% des unités de transcription non codantes (ARNr, ARNt) chez *E. coli*. Cela signifie que la bactérie a développé un autre type de mécanisme pour réguler l'expression de ses gènes à des endroits où la présence d'une structure stable est incompatible, soit avec le niveau de régulation voulu, soit avec une fonction biologique donnée. Par exemple, la séquence hautement spécifique et limitée d'un terminateur intrinsèque est incompatible avec la présence d'une séquence codante au milieu d'un gène. Les autres événements de terminaison chez *E. coli* sont donc médiés par des facteurs protéiques et en grande majorité par la protéine Rho.

4. La terminaison facteur dépendante.

Bien que des protéines phagiques, comme la protéine Nun du phage HK022, soient également capables de provoquer le relargage du transcrit chez *E. coli* (j'en reparlerais plus tard dans ce manuscrit), on connaît principalement deux facteurs protéiques d'*E. coli* qui permettent la terminaison de la transcription : TRCF et Rho.

4. 1. TRCF.

Le premier de ces facteurs, TRCF (pour "Transcription Repair Coupling Factor"; [Selby et Sancar, 1993](#)) intervient lorsque l'ARN polymérase rencontre un dommage dans l'ADN (causé par une exposition aux UV par exemple). L'ARN polymérase est arrêtée au contact de cette lésion. La réparation de l'ADN endommagé nécessite l'accès de l'ADN aux enzymes de réparation. Le facteur TRCF (produit du gène *mfd*) permet alors de dissocier le complexe arrêté afin de rendre l'ADN disponible pour l'action des ces enzymes. En fait TRCF agit comme un facteur de terminaison spécifique des complexes de transcription arrêtés ([Selby et Sancar, 1995a](#) et [b](#)).

4. 2. Rho.

Le second facteur de terminaison est un complexe homohexamérique connu sous le nom de facteur Rho ([Roberts, 1969](#)), et dont les caractéristiques sont décrites dans le chapitre suivant.

C'est en recherchant, dans un extrait bactérien, des facteurs capables de modifier la synthèse *in vitro* d'ARN à partir d'ADN du phage λ , que Roberts a découvert la protéine Rho (Roberts, 1969). Ses expériences ont révélé que Rho (pour "release") permettait de réduire, de façon conservée, la taille des transcrits synthétisés à partir des promoteurs pL et pR des opérons précoces de λ , sans pour autant réduire la quantité de transcrits initiée par les deux promoteurs. Roberts en a conclu que Rho agissait en provoquant la terminaison précoce de la transcription à des sites précis.

Depuis sa découverte, Rho est devenu un modèle permettant de comprendre les mécanismes utilisés par certains facteurs protéiques pour réguler la terminaison de la transcription. On a longtemps attribué un rôle mineur au facteur Rho dans la régulation de l'expression des gènes chez les procaryotes (Richardson et Greenblatt, 1996). Pourtant, bien que certaines en soient dépourvues (Quirk et al., 1993), la majorité des bactéries ont besoin (de façon plus ou moins importante) de l'action du facteur de terminaison Rho (Zheng et Friedman, 1994). Chez *E. coli* en particulier, Rho est indispensable à la survie de la bactérie et les mutants de Rho présentent une activité soit résiduelle soit conditionnelle (Platt et Richardson, 1992; Das et al., 1976). L'impossibilité d'obtenir des mutants nuls, mais aussi l'abondance de la protéine dans la bactérie (1-10 μ M; Geiselmann et al., 1992) ont alors rapidement suggéré un rôle nettement plus important de Rho.

On pense aujourd'hui que Rho est responsable d'une grande partie des événements de terminaison qui interviennent chez *E. coli*. Vraisemblablement, quasiment tous les phénomènes de terminaison qui ne sont pas provoqués par des terminateurs intrinsèques le sont par le facteur Rho. On peut penser que la bactérie a mis en place ce mécanisme afin de permettre le relargage du transcrit à des sites où le complexe ternaire d'élongation est trop stable pour se dissocier spontanément.

1. Caractéristiques des terminateurs Rho-dépendants.

Contrairement aux terminateurs intrinsèques (voir précédemment), les terminateurs Rho-dépendants ne présentent pas de séquence consensus clairement définie (Figure 5). Ils

sont donc difficiles à repérer au sein du génome bactérien et très peu de terminateurs Rho-dépendants ont été caractérisés à ce jour. C'est en partie pour cette raison que Rho a longtemps été considéré comme un facteur minoritaire pour la terminaison de la transcription chez *E. coli*.

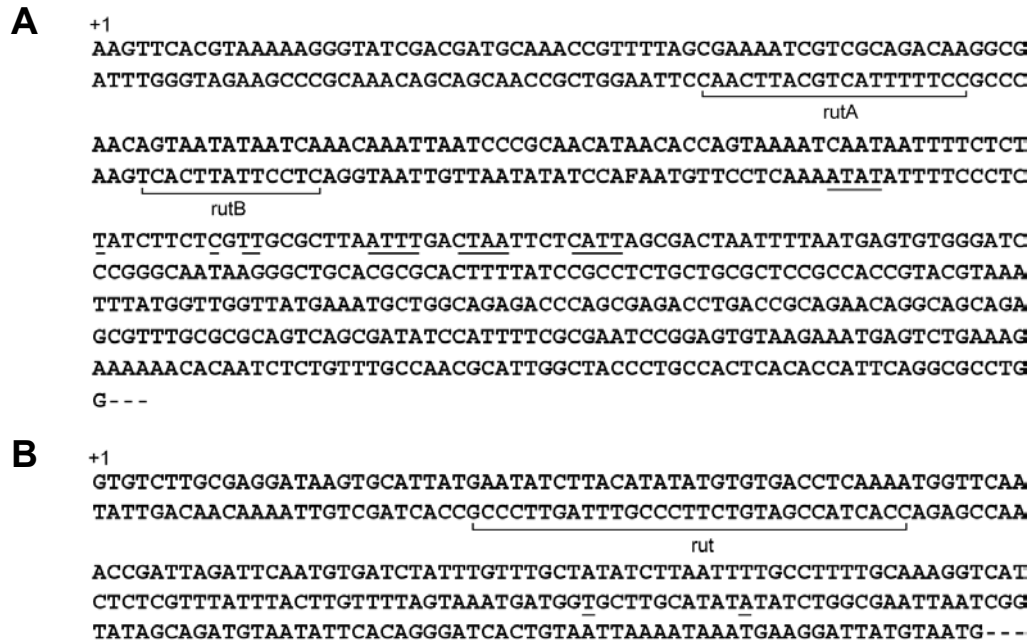


Figure 5 : Séquences des terminateurs *trp*' (A) et *tna* (B) d'*E. coli*.

Les éléments *rut* sont indiqués et les sites de terminaison sont soulignés.

Malgré leur diversité de séquence, les terminateurs Rho-dépendants présentent quelques caractéristiques communes. On sait, en particulier, qu'ils sont composés de deux éléments distincts répartis dans une séquence ADN de 150 à 200 pb environ (Lau et al., 1983; Richardson, 1990; Zalatan et al., 1993) :

- Une séquence essentielle nommée *rut* (pour "Rho utilization") et située en amont du site de terminaison.
- Le site de terminaison proprement dit, où intervient le relargage du transcrit, aussi appelé *tsp* (pour "transcription stop point").

La zone de terminaison située en aval du site *rut* peut être très étendue (jusqu'à 100pb, Figure 5). Le relargage du transcrit peut en effet intervenir à plusieurs endroits au sein de cette séquence (Lau et al., 1982). Cette dispersion contraste fortement avec la précision de la

terminaison intrinsèque pour laquelle le relargage intervient au niveau de deux ou trois bases adjacentes.

1. 1. La séquence *rut*.

La séquence *rut* code pour le site d'entrée de Rho sur l'ARN. Nous allons le voir, Rho possède une forte affinité pour certains transcrits ARN. Cette affinité particulière est corrélée à la présence d'un site *rut* dans la séquence ADN qui donne naissance au transcrit. Ainsi Rho ne fixe pas les transcrits terminés en amont du site *rut* ni ceux issus d'une séquence où le site *rut* a été déleté (Faus et Richardson, 1989). Enfin un oligonucléotide, complémentaire de l'élément RUT ARN, est capable de bloquer la terminaison Rho-dépendante, probablement parce qu'il interfère avec la fixation de Rho sur le transcrit (Chen et Richardson, 1987; Guérin et al., 1998).

Des expériences d'empreintes (Bear et al., 1988; Zhu et al., 1998a), ont révélées que Rho protège une zone de 70-80 nucléotides sur le transcrit d'une dégradation à la RNaseA. Cependant, certains terminateurs naturels (λ tR1, tna) ou artificiels (Guérin et al., 1998) peuvent être beaucoup plus courts (entre 40pb et 60pb). La grande variété de séquence des terminateurs Rho-dépendants rend très difficile la détermination d'une séquence-type minimale pour fixer Rho à l'ARN, les séquences adjacentes non spécifiques pouvant prendre part à l'interaction (Hart et Roberts, 1994).

L'étude de l'affinité de Rho pour les séquences ARN a montré que celui-ci possédait une plus forte affinité pour les séquences riches en cytosine (C), et une affinité très faible pour les séquences hautement structurées (Galluppi et Richardson, 1980; Richardson, 2002). L'analyse des séquences *rut* connues montre ainsi qu'elles correspondent à des régions peu structurées et possédant une proportion de C supérieure à la moyenne (Morgan et al., 1983; Figures 5 et 6).

Une autre caractéristique des séquences *rut* est une quantité de guanine (G) nettement inférieure à la moyenne (Alifano et al., 1991). La faible proportion de résidus G est révélatrice en raison de leur forte aptitude à s'apparier avec les autres bases pour former des structures secondaires stables dans l'ARN.

Pour résumer, on peut décrire un site *rut* comme une séquence ADN, codant pour un segment du transcrit d'au moins 40 à 60 nucléotides, peu ou pas structuré et présentant un ratio C/G supérieur à la moyenne.

Ces caractéristiques ont permis la mise au point d'un programme informatique permettant de rechercher, au sein du génome, des "bulles" riches en C et pauvres en G situées en amont de sites de terminaison connus, et pouvant servir de séquence *rut* pour induire la terminaison. Ce système a tout d'abord été vérifié sur deux terminateurs connus (Figure 6), puis a permis de définir les séquences du terminateur *hisG* nécessaire à la terminaison Rho-dépendante (Alifano et al., 1991; Rivellini et al., 1991).

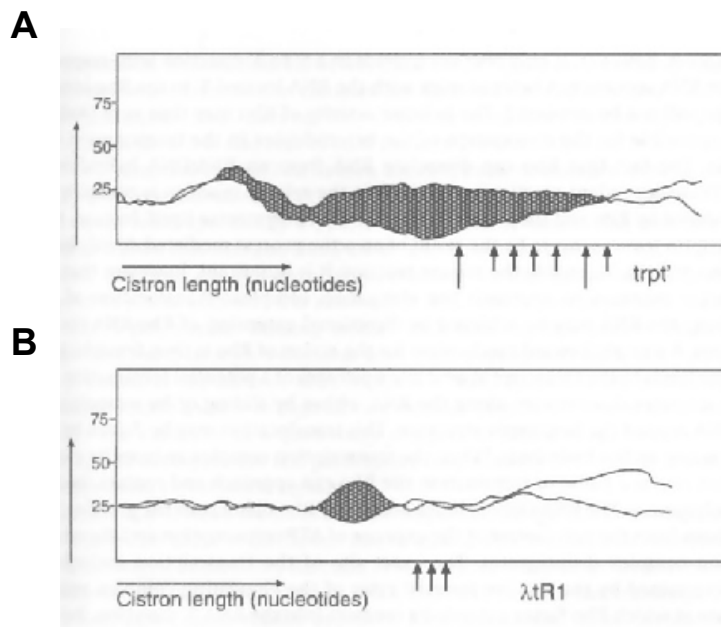


Figure 6 : Analyses informatiques des terminateurs *trpL'* (A) et λ tR1 (B).

Les graphiques représentent les pourcentages de C et de G, au sein du brin codant, en fonction de la longueur en nucléotides. Les flèches indiquent les sites de terminaison observés. Les zones hachurées correspondent à des régions riches en C et pauvres en G (d'après : Alifano et al., 1991 et Rivellini et al., 1991).

Quelques études ont été menées pour déterminer quelles cytosines sont impliquées dans le phénomène de terminaison. Différents terminateurs ont été mutés afin de mesurer l'influence de la perte de résidus C sur l'efficacité de terminaison. Bien que la majorité des sites *rut* étudiés doivent comporter un certain nombre de résidus C, aucune étude n'a pu

identifier de nucléotides essentiels à la terminaison (Hart et Roberts, 1991; Zalatan et Platt, 1992). Apparemment, une séquence suffisamment longue, capable de générer un ARN dépourvu de structure secondaire et présentant une répartition homogène de résidus C, est capable d'initier la terminaison Rho-dépendante (Guérin et al., 1998).

Dans la cellule, Rho se structure sous la forme d'un anneau hexamérique (voir : Organisation structurale). D'après cette structure, il semble qu'un polynucléotide simple brin de 12-13 nucléotides minimum soit nécessaire pour relier les domaines de fixation à l'ARN de deux monomères successifs. D'autres études vont dans le même sens et suggèrent qu'une répétition de C tout les 12 ou 13 nucléotides, sur un transcrit non structuré, est suffisante pour servir de site primaire de fixation permettant à Rho d'initier la terminaison (Hart et Roberts, 1991). Cet espacement est généralement assez bien retrouvé chez les terminateurs étudiés jusqu'à présent, comme le terminateur tR1 du phage λ (Figure 7).

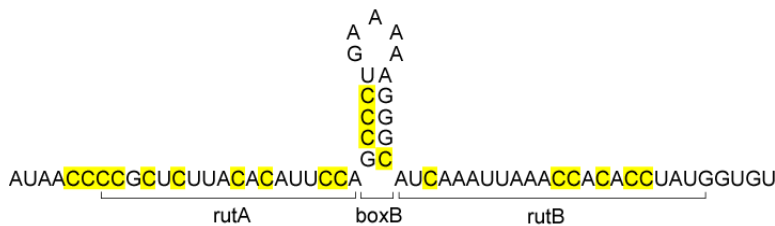


Figure 7 : Séquence ARN du terminateur tR1 du phage λ . Les cytosines sont surlignées.

1. 2. Les sites de terminaison.

Les sites de terminaison sont encore moins bien définis que le site *rut*. Cependant, les études cinétiques de la transcription en absence de Rho ont montré que le relargage de l'ARN par Rho intervient en général au niveau des sites qui induisent une pause de l'ARN polymérase en cours d'élongation (Lau et al., 1984; Morgan et al., 1983 a et b; Guérin et al., 1998).

En 1992, Jin et ses collaborateurs ont montré qu'un mutant de Rho possédant une vitesse de translocation réduite pouvait provoquer une terminaison sauvage lorsque la vitesse d'élongation de l'ARN polymérase était réduite, *in vitro*, par de faibles concentration en nucléotides ou *in vivo* par une mutation (mutation *rpoB8*; Fisher et Yanofsky, 1983; Jin et al., 1992). Ces résultats ont révélés l'existence d'un couplage cinétique entre Rho et l'ARN

polymérase. Rho se fixe en général à l'ARN 60 nucléotides en amont du premier site de terminaison. Il doit donc être capable de combler cette distance, et de rattraper le complexe d'élongation afin de permettre le relargage du transcrit. L'efficacité de la terminaison dépendrait donc de ce couplage cinétique entre : la fixation puis la translocation de Rho le long du transcrit d'une part et la vitesse d'élongation de l'ARN polymérase d'autre part. Les facteurs ou les événements capables de réguler cette vitesse vont alors jouer un rôle prédominant pour déterminer la position des sites de relargage et l'efficacité de la terminaison au niveau de ces sites (Burns et al., 1998; Landick et al., 1987). Dans cette hypothèse, la pause de l'ARN polymérase va favoriser la terminaison, et à l'inverse les facteurs favorisant la processivité de l'ARN polymérase vont réguler négativement la terminaison Rho-dépendante (voir chapitre suivant).

Malgré leur rôle apparemment indispensable, ce ne sont pas les sites *tsp* qui assurent la spécificité d'action des terminateurs Rho-dépendant. En 1996, Richardson et Richardson montrent que Rho peut induire le relargage du transcrit dans une séquence non terminatrice située en aval de la partie supérieure (sites *rut*) du terminateur λ , suggérant ainsi que la séquence *rut* est l'élément prédominant dans la terminaison induite par Rho et que les sites *tsp* n'ont qu'une importance relative (Richardson et Richardson, 1996). Cette fois encore les sites de relargage coïncident avec des sites de pause de l'ARN polymérase au sein de ces séquences non terminatrices. Par contre, il semble dans cette étude que l'efficacité de terminaison n'est pas corrélée avec la force du signal de pause au niveau de ces sites. Les auteurs de ce travail suggèrent que ce n'est pas tant la durée de la pause, mais plutôt un changement conformationnel de l'ARN polymérase au contact du site de pause qui est responsable de la sensibilité du complexe d'élongation vis-à-vis du facteur Rho. Selon eux, le signal de pause ne représenterait donc pas une composante cinétique mais plutôt une composante structurale du phénomène de terminaison Rho-dépendante.

Quoiqu'il en soit, l'ensemble des résultats obtenus jusqu'à ce jour montre que le site *rut* est l'élément le plus important d'un terminateur Rho-dépendant. La position et l'efficacité de la terminaison vont dépendre de la spécificité de Rho pour cette séquence et de sa capacité à initier la réaction au niveau du site RUT (Richardson et Richardson, 1996; Graham et Richardson, 1998; Guerin et al., 1998). Nous venons de le voir, les sites de terminaison sont beaucoup moins spécifiques et peuvent être remplacés par des séquences non terminatrices, même si la plupart du temps ces sites de terminaison sont corrélés à des sites de pause de

l'ARN polymérase. On ne sait pas si leur action est cinétique, structurale, ou les deux. L'hypothèse "structurale" pourrait s'expliquer par la ressemblance entre certains sites de pause (classe I) et les terminateurs intrinsèques, bien que d'autres sites de pause (classe II) soient peu structurés (Artsimovitch et Landick, 2000). Les sites de pauses induisent peut être un remaniement structural de l'ARN polymérase, la rendant alors plus sensible à l'action de Rho.

Sans pour autant réfuter l'hypothèse d'une compétition cinétique entre Rho et l'ARN polymérase, les résultats décrits précédemment semblent suggérer que le véritable contrôle de cette compétition cinétique se jouerait au niveau de la fixation de Rho sur le transcrit et de l'initiation de la translocation. On peut en effet penser que Rho, sous réserve d'une fixation efficace sur l'ARN, se retrouverait très rapidement dans une position propice à la terminaison, c'est-à-dire assez proche de l'ARN polymérase. Cependant le complexe d'élongation serait alors trop stable pour permettre la terminaison de la transcription. D'après l'hypothèse décrite précédemment sa conformation serait modifiée au contact du site de pause, permettant ainsi à Rho d'achever la déstabilisation du complexe.

2. Organisation structurale.

Il a fallu attendre 1983 et l'isolement puis le séquençage du gène *rho* par Pinkman et Platt, pour obtenir des informations sur l'organisation structurale et les propriétés physico-chimiques du facteur de terminaison (Pinkhan et Platt, 1983). La comparaison de la structure primaire putative de Rho avec celles d'autres protéines, a permis de mettre en évidence deux domaines fonctionnels distincts séparés par un site de coupure à la trypsine (Bear et al., 1985; Dolan et al., 1990; Figure 8).

- Un motif caractéristique des ATPases (Yager et von Hippel, 1987; Dombroski et Platt, 1988).
- Et un motif similaire au motif de fixation à l'ARN de certaines ribonucléoprotéines eucaryotes (Query et al., 1989; Kenan et al., 1991; Figure 8).

Comme nous l'avons déjà vu Rho possède, en effet, la capacité de fixer l'ARN, il est également capable d'hydrolyser l'ATP, ainsi que les autres NTPs (Lowery et Richardson, 1977).

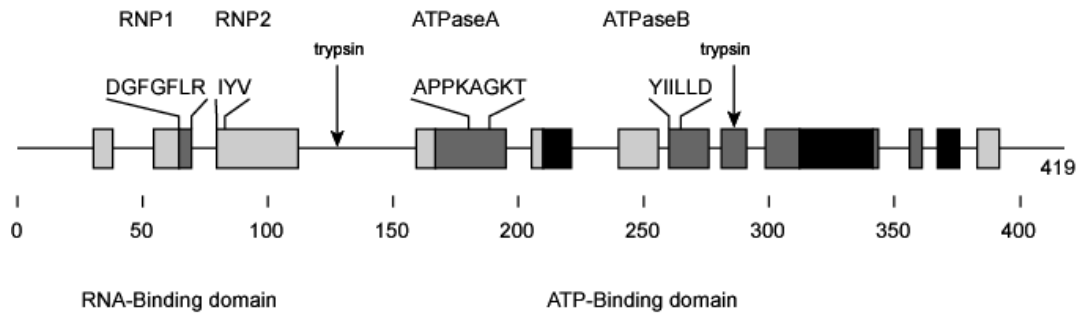


Figure 8 : Organisation structurale du polypeptide Rho (d'après : [Richardson, 1996](#)). Les différents domaines sont représentés ainsi que les sites d'action de la trypsine. Les zones de séquences conservées, entre les différentes bactéries, sont représentées par des rectangles, les segments les plus foncés correspondent aux zones les mieux conservées.

Dans les conditions physiologiques, le facteur Rho existe sous la forme d'un assemblage de six sous-unités identiques de 419 acides aminés, qui sont le produit du gène *rho* ([Finger et Richardson, 1982](#); [Geiselman et von Hippel, 1992](#)). Les études structurales, essentiellement en microscopie électronique ([Bear et al., 1988](#); [Gogol et al., 1991](#); [Yu et al., 2000](#)), ont montré que cet hexamère se structure sous la forme d'un anneau de 120Å de diamètre avec une ouverture centrale de 30 Å. Jusqu'à l'été 2003 les données structurales de Rho provenaient uniquement :

- Des structures RMN et cristallographiques du domaine de fixation à l'ARN seul ([Briercheck et al., 1998](#); [Allison et al., 1998](#)).
- De modèles moléculaires obtenus par homologie avec certaines protéines, en particulier avec la F₁ ATP synthase, dont la sous-unité β est fortement homologue au domaine C-terminal de Rho ([Opperman et Richardson, 1994](#)).
- De reconstructions tridimensionnelles réalisées à partir de clichés de microscopie électronique de protéines Rho fixées sur une grille de carbone ([Yu et al., 2000](#); [Gogol et al., 1991](#)).

Récemment Skordalakes et Berger ont dévoilé la structure cristallographique de l'hexamère entier à 3Å de résolution ([Skordalakes et Berger, 2003](#); [Figure 9](#)), confirmant la plupart des informations dévoilées par les études structurales citées précédemment.

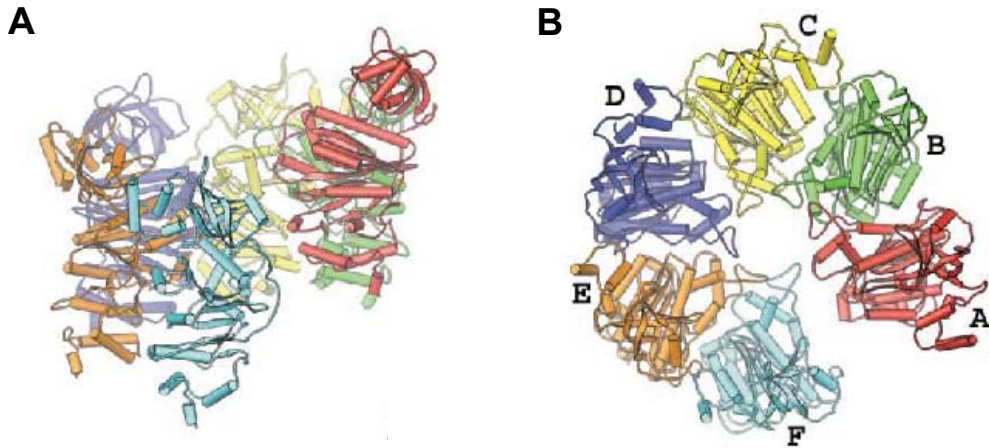


Figure 9 : *Structure de l'hexamère Rho. (A) Vue de côté et (B) vue du dessus de la même structure. Chaque monomère est représenté dans une couleur et par une lettre différente (d'après : Skordalakes et Berger, 2003).*

3. Quand Rho rencontre l'ARN.

Il est rapidement apparu que la fixation à l'ARN est essentielle pour la fonction de Rho. L'interaction entre Rho et l'ARN initie la réaction de terminaison. Elle est nécessaire pour activer l'hydrolyse de l'ATP par Rho (Lowery-Goldhammer et Richardson, 1974), hydrolyse qui est essentielle à la translocation du facteur le long de l'ARN et au relargage du transcrit hors du complexe ternaire d'élongation. Cependant, si l'activité ATPase de Rho est ARN dépendante, la fixation de Rho à l'ARN peut se faire en absence d'ATP (Galluppi et Richardson, 1980; McSwiggen et al., 1988).

Le site de fixation de Rho à l'ARN correspond aux 150 premiers résidus N-terminaux du monomère. Un fragment contenant les 129 premiers résidus a été isolé et cristallisé, seul (Modrack et Richardson, 1994; Allison et al., 1998) ou complexé à de l'ARN (Bogden et al., 1999). Les structures obtenues, ainsi que des alignements de séquences, ont révélés la présence d'un motif OB (pour "oligonucleotide/oligosaccharide binding") caractéristique des domaines de reconnaissance des polynucléotides simple brin. C'est ce motif qui est responsable de l'affinité de Rho pour les séquences ARN simple brin (Geiselman et al., 1992; Richardson et Richardson, 1992). Cette reconnaissance n'implique pas l'hydroxyle 2' de l'ARN, Rho peut donc également reconnaître l'ADN simple brin grâce à son motif OB.

Rho possède une préférence pour les ARN simple brin riches en C; il se fixe, par exemple, très fortement à des oligonucléotides poly(rC). L'affinité de Rho pour son substrat ARN a été mesurée par de nombreuses techniques (filtration, gel retard, etc...), et avec des substrats variés (TR1, TRPT', (rC)_n, (rUrC)_n; Wang et von Hippel, 1993b; Steinmetz et Platt, 1994; Gan et Richardson, 1999). Par exemple la constante de dissociation de Rho pour l'ARN du terminateur λtR1 peut varier de 0,02nM (en présence d'ATP) à 0,2nM (en absence d'ATP) (Gan et Richardson, 1999). L'affinité de Rho pour l'ARN peut également être fortement diminuée lorsque celui-ci est dépourvu de C (Cerruzi, 1985).

Des oligonucléotides rC courts (moins de 13 nucléotides) ont été utilisés pour analyser l'affinité de chaque monomère au sein de l'hexamère (Geiselman et von Hippel, 1992; Wang et von Hippel, 1993b). Ces études d'interaction avec l'ARN montrent qu'il existe un site de fixation à l'ARN au niveau de l'extrémité N-terminale de chaque monomère. Au sein de l'hexamère trois de ces sites d'interaction présentent une forte affinité et les trois autres présentent une faible affinité pour l'ARN. De plus, l'affinité d'une sous-unité isolée pour un oligo rC (Modrack et Richardson, 1994) est identique à celle d'un site faible. Ceci indique que les sous-unités n'ont pas toutes la même conformation au sein de l'hexamère (Wang et von Hippel, 1993b; Geiselman et al., 1992) et que des interactions protéine-protéine entre les sous-unités de Rho permettent de faire passer certains monomères d'une conformation à faible affinité vers une conformation à forte affinité.

L'hexamère, aussi bien que le domaine N-terminal d'un monomère, sont capables de fixer d'autres types de polynucléotides dépourvus de C, mais leur affinité pour ce type de séquence est nettement inférieure à leur affinité pour le poly(rC). Par exemple, le poly(rU), qui est le polynucléotide (après le poly(rC)) qui fixe le mieux Rho, présente une affinité dix fois plus faible que le poly(rC) (Wang et von Hippel, 1993b). Une structure cristallographique du domaine N-terminal complexé à un oligonucléotide (rC)₉ a permis de comprendre les raisons moléculaires de l'affinité de Rho pour les pyrimidines en général, et pour les résidus C en particulier (Richardson et Richardson, 1992; Bogden et al., 1999) :

- Deux résidus, Arg-66 et Asp-78, forment trois liaisons hydrogène spécifiques (de type Watson-Crick) avec la base d'une cytosine.
- De plus la base située en 5' de ce résidu C se fixe au niveau d'une poche dont la taille ne permet d'accueillir qu'une pyrimidine.

Ainsi la partie N-terminale d'un monomère Rho présente au moins deux domaines voisins qui sont fortement cytosine- ou pyrimidine-spécifiques. Rho peut aussi fixer l'ADN simple brin (Modrack et Richardson, 1994; Galluppi et Richardson, 1980), mais son affinité pour des oligo(dC) est environ 20 fois plus faible que pour les oligo(rC) (Briercheck et al., 1996; Richardson, 1982). Par ailleurs, la fixation d'ADN ne permet pas l'hydrolyse de l'ATP (Galluppi et Richardson, 1980).

Cette dernière observation est extrêmement importante puisqu'elle a permis de découvrir que le domaine N-terminal ne constituait pas l'unique site de fixation à l'ARN présent chez Rho. Les mesures de l'hydrolyse de l'ATP en fonction du type de substrat ADN ou ARN ont en effet distingué deux types de sites de fixation :

- Le premier que nous venons de décrire fait intervenir le motif OB dans la partie N-terminale de la protéine. Il est capable de fixer à la fois l'ADN simple brin et l'ARN, avec une nette préférence pour un transcrit non structuré et riche en résidus C. La fixation d'ARN ou d'ADN au niveau de ce site n'est pas liée à l'activité ATPasique de Rho.
- Le second site, au contraire, est spécifique de l'ARN mais ne nécessite pas de séquence particulière (même si les résidus rC semblent préférés; Richardson, 1982). C'est la fixation du transcrit au niveau de ce site qui induit l'hydrolyse de l'ATP.

Jusqu'à récemment la position du site secondaire était inconnue. Certains travaux suggéraient même que ce site résultait d'un changement de conformation du site primaire (Geiselmann et al., 1993). Une série d'études menées ces dernières années montre que les deux sites sont indépendants (Burgess et Richardson, 2001; Wei et Richardson, 2001). Le site de fixation secondaire à l'ARN est en fait localisé au niveau du domaine C-terminal de Rho. Plus précisément, il fait intervenir des boucles Q et R qui sont disposées vers le centre de l'hexamère dans la structure basée sur l'homologie avec la F₁ ATP synthase (positionnement confirmé par la suite grâce aux structures de Skordalakes et Berger, Figure 10).

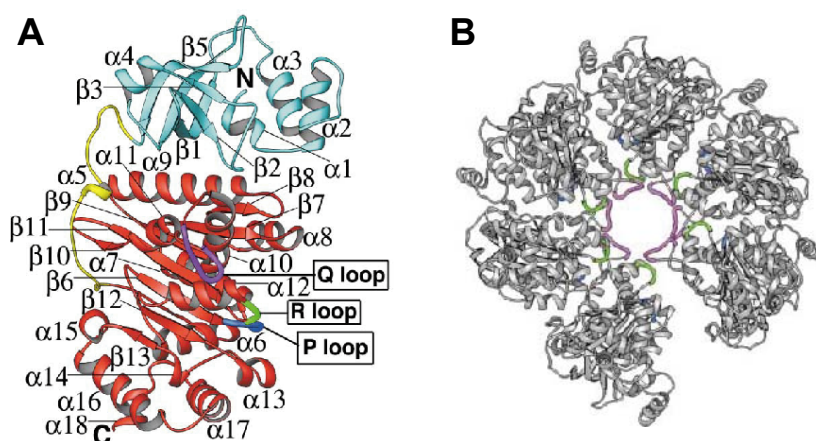


Figure 10 : Structure et topologie de Rho :

(A) monomère, (B) hexamère. On distingue les boucles Q (violet) et R (vert) situées dans le centre de l'anneau et qui correspondent au site de fixation secondaire à l'ARN (d'après : Skordalakes et Berger 2003).

Des expériences récentes de pontage chimique ont mis en évidence l'interaction physique entre des résidus de la boucle R et un ARN portant un site RUT prolongé en 3' par une séquence aléatoire. Par contre, si cette séquence est absente et que l'extrémité 3' de l'ARN correspond au 3' de la séquence RUT, l'interaction n'implique que le site primaire localisé dans la partie N terminale de Rho (Burgess et Richardson, 2001a).

Une seconde étude a révélé que la boucle Q se trouvait protégée de l'attaque de radicaux peroxydes (OH^\cdot) en présence d'ARN mais pas en présence d'ADN. De plus, la mutation de résidus de la boucle Q a permis de mettre en évidence quatre acides aminés (S281, K283, T286, D290) essentiels pour l'interaction ARN spécifique au niveau du site secondaire et pour l'activité ATPasique qui en résulte (Wei et Richardson, 2001a).

Ces résultats indiquent qu'après avoir fixé le site primaire grâce à sa séquence RUT, le transcrit pénètre au centre de l'hexamère et interagit avec des résidus localisés au niveau du domaine de fixation et d'hydrolyse de l'ATP.

Cette révélation soulève la question suivante : comment l'ARN pénètre-t-il dans le centre de l'hexamère pour interagir avec le site de fixation secondaire?

Le passage d'un polynucléotide simple brin dans le centre de l'anneau est une caractéristique des hélicases hexamériques qui sont structurellement et mécaniquement proches de Rho (Egelman et al., 1995). Chez ces hélicases on propose généralement trois modèles différents (Figure 11) pour expliquer l'entrée d'un polynucléotide simple brin (ADN ou ARN) dans le centre de l'hexamère (pour revue voir : Patel et Picha, 2000).

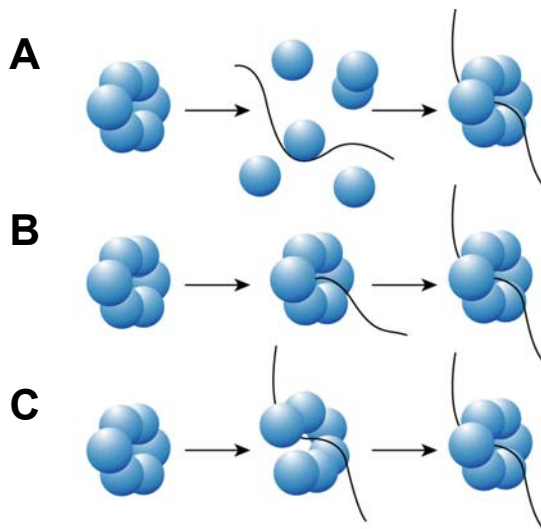


Figure 11 : Modèles de fixation d'un polynucléotide simple brin dans le centre d'une hélicase hexamérique : (A) Formation de l'hexamère autour du polynucléotide. (B) Entrée du polynucléotide par une de ses extrémités. (C) Ouverture de l'anneau pour permettre l'entrée du polynucléotide (d'après : Patel et Picha, 2000).

- Selon la première hypothèse, l'hexamère pourrait se former directement autour du polynucléotide à partir de la fixation initiale d'une ou plusieurs sous-unités sur la molécule ADN ou ARN (Figure 11A). Cependant, comme je l'ai déjà précisé, la forme hexamérique de Rho est majoritaire dans les conditions physiologiques, et même si l'interaction avec l'ARN stabilise l'hexamère, elle n'est pas responsable de sa formation (Geiselmann et al., 1992).

- La deuxième hypothèse suppose que l'entrée du polynucléotide se ferait grâce à l'une de ses extrémités qui passerait dans le centre d'un hexamère déjà formé (Figure 11B). Cependant en utilisant un transcrit circularisé comme substrat Burgess et Richardson ont récemment montré que Rho pouvait fonctionner sans disposer d'une extrémité libre de l'ARN (Burgess et Richardson, 2001b). Ce résultat est assez cohérent avec la fonction de Rho *in vivo*. En effet dans le contexte d'un terminateur situé à la fin d'un gène ou d'un opéron, Rho doit être capable de fixer un ARN dont l'extrémité 5' est bloqué par les ribosomes en cours de traduction, et l'extrémité 3' est bloquée par l'ARN polymérase en cours d'élongation de la transcription.

- Finalement le troisième modèle (Figure 11C) suggère que le fonctionnement de l'enzyme nécessite une ouverture de l'anneau suffisante pour permettre le passage d'un simple brin dans le centre de l'hexamère et, dans le cas de Rho, l'interaction de l'ARN avec le site de fixation secondaire. En effet, même si la majorité des clichés de microscopie électronique (Gogol et al., 1991; Yu et al., 2000) représentent Rho sous sa forme fermée, une petite partie correspond à une forme interrompue de l'anneau.

Cette interruption peut être la conséquence, soit d'une ouverture de l'hexamère, soit de la perte d'une sous-unité (Figure 12A). La structure de Rho révélée par Skordalakes et Berger a permis de conforter cette observation (Figure 12).

Deux structures de Rho complexé avec des oligonucléotides en présence ou en absence d'AMP-PNP (un analogue non hydrolysable de l'ATP), ont été proposées. Les deux structures révèlent un hexamère ouvert ressemblant à la portion d'une hélice équivalant à un tour d'hélice. Le dernier monomère (qui interagit avec la partie 3' du transcrit) se retrouvant ainsi 45Å plus bas que le premier monomère (Figure 12B). Cette disposition aboutit à une ouverture d'à peu près 12Å, suffisante pour permettre l'entrée d'un brin unique d'ARN dans le centre de l'anneau (Skordalakes et Berger, 2003). On ignorait jusqu'à présent si cette conformation, minoritaire sur les clichés de microscopie électronique, avait un sens biologique ou non. L'isolement de cette conformation par Skordalakes et Berger suggère fortement qu'elle possède une fonction réelle dans le mécanisme de terminaison.

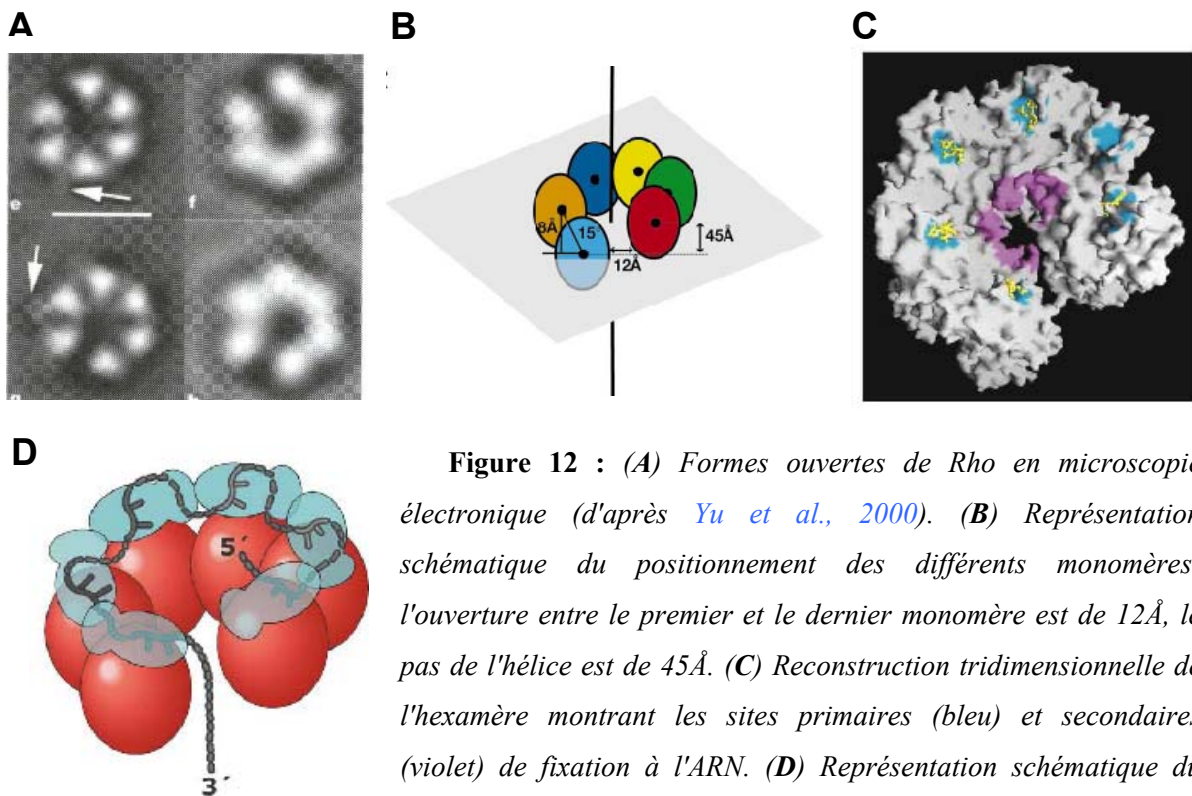


Figure 12 : (A) Formes ouvertes de Rho en microscopie électronique (d'après Yu et al., 2000). (B) Représentation schématique du positionnement des différents monomères; l'ouverture entre le premier et le dernier monomère est de 12Å, le pas de l'hélice est de 45Å. (C) Reconstruction tridimensionnelle de l'hexamère montrant les sites primaires (bleu) et secondaires (violet) de fixation à l'ARN. (D) Représentation schématique du positionnement de l'ARN par les domaines N-terminaux des différents protomères vers le centre de l'hexamère (d'après : Skordalakes et Berger, 2003).

Les premières structures 3D de Rho issues de la microscopie électronique laissent penser que l'ARN, en se fixant au site primaire, encerclait l'hexamère (Yu et al., 2000). La structure de Skordalakes et Berger a révélé que les motifs OB de l'extrémité N-terminale des sous-unités de Rho, sont en fait dirigés vers le centre de l'hexamère (Figure 12C). Les auteurs suggèrent que cette particularité ainsi que la structure hélicoïdale de l'hexamère favorise le positionnement de l'ARN vers le centre de Rho, orientation propice à l'entrée de l'extrémité 3' du transcrit dans le centre de l'anneau (Figure 12D).

On ne sait pas précisément ce qui induit l'ouverture de l'anneau; il ne s'agit peut être que d'un événement aléatoire. Il est en effet envisageable que l'orientation de l'ARN induite par le site primaire de fixation permette la pénétration du transcrit au cœur d'un anneau déjà ouvert ou bien dès l'ouverture aléatoire de celui-ci. Cependant il reste possible que la fixation du transcrit au niveau du site primaire de Rho provoque un changement conformationnel conduisant à l'ouverture de l'anneau.

L'interaction de l'ARN avec les boucles Q et R du domaine C-terminal de Rho n'est plus spécifique de la nature des bases présentes dans l'ARN, mais de l'ARN lui-même. En effet, comme je l'ai déjà précisé, si le site primaire est capable de fixer aussi bien l'ARN que l'ADN, le site secondaire lui est spécifique de l'ARN. On ignore actuellement tout du détail des interactions qui assurent cette spécificité entre l'ARN et le site secondaire de Rho.

Une fois l'ARN fixé au niveau du site secondaire, l'hexamère se referme probablement. Cette fermeture n'a pas été observée dans des conditions dynamiques. Cependant, les clichés de microscopie électronique réalisés en présence d'oligonucléotides courts ((rC)₂₃) dévoilent les formes ouvertes et fermées de l'hexamère; alors qu'en présence d'oligo longs ((rC)₁₀₀) seule la forme fermée de l'hexamère est visible (Gogol et al., 1991). On constate que la forme fermée de l'anneau est favorisée en présence d'oligonucléotides suffisamment longs pour interagir simultanément avec le site primaire et le site secondaire.

Par ailleurs, les différentes structures de Rho suggèrent que, dans la conformation ouverte, les éléments responsables de l'hydrolyse de l'ATP ne sont pas dans des positions propices à la catalyse (Skordalakes et Berger, 2003; Gan et Richardson, 1999). La fermeture de l'anneau, après interaction entre l'ARN et le site secondaire, semble donc nécessaire pour initier l'étape de translocation. Le site secondaire va être impliqué dans la translocation ATP dépendante de Rho le long du transcrit. La localisation de ce site au sein du motif responsable

de la fixation et de l'hydrolyse de l'ATP suggère qu'un couplage entre la fixation de l'ARN et l'hydrolyse d'ATP, au travers de changements conformationnels et de l'alternance entre sites de faible et de forte affinité, serait responsable de la translocation de Rho le long du transcrit.

Ces informations sont confortées par l'étude d'un antibiotique, la bicyclomycine (Figure 13) qui bloque la translocation de Rho *in vitro* (Magyar et al., 1996). On pense qu'il agit soit en entrant en compétition avec la fixation de l'ARN au niveau du site secondaire, soit en découplant la translocation de l'activité ATPase de Rho. En effet certaines prédictions placent les sites de fixation de la bicyclomycine à proximité des sites de fixation et d'hydrolyse de l'ATP.

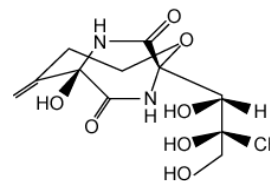


Figure 13 : Structure de la bicyclomycine

Si on résume les informations sur l'interaction de Rho avec l'ARN, on constate que le site primaire sert : d'une part, à initier la réaction de terminaison en permettant à Rho de reconnaître sa séquence de fixation initiale, et d'autre part à positionner l'ARN dans le centre de l'anneau pour favoriser l'interaction avec le site secondaire. Si le passage du transcrit dans le centre de l'anneau est maintenant prouvé, on ne sait pas, à ce jour, si l'interaction entre le site primaire et la séquence RUT est maintenue au cours de la translocation de Rho (Figure 14).

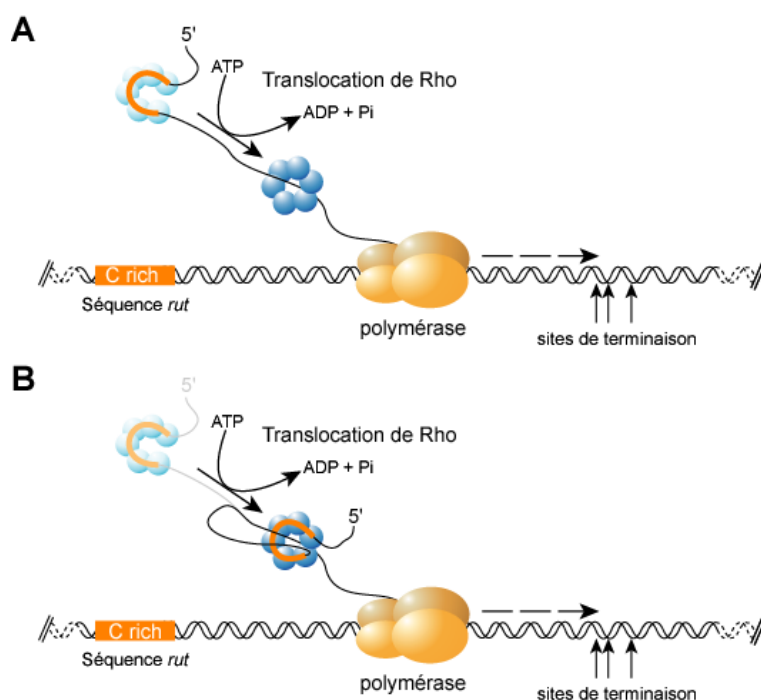


Figure 14 : Modèles de translocation de Rho.

(A) "Tracking" : Rho transloque le long de l'ARN sans maintenir son interaction initiale avec le site RUT. (B) "Tethered tracking" : Rho transloque le long de l'ARN en maintenant une interaction entre le site primaire et le transcrit.

Deux modèles s'opposent sur ce mécanisme de translocation:

- La première hypothèse suggère que Rho migre le long du transcrit sans maintenir sa fixation au site RUT (Geiselmann et al., 1993; Figure 14A).
- L'autre hypothèse soutient que Rho conserve son interaction initiale avec le site RUT pendant toute la réaction de translocation. Ce dernier modèle repose sur une expérience ayant montré que Rho, au cours de la translocation, pouvait maintenir un contact avec deux parties éloignées du transcrit (Faus et Richardson, 1990; Steinmetz et Platt, 1994; Figure 14B).

D'après les structures tridimensionnelles, l'orientation des domaines N-terminaux de Rho semble compatible avec le passage d'un brin ARN au centre de l'anneau tout en maintenant leur interaction avec RUT. Cependant, il est difficile de prouver qu'une molécule unique de Rho soit capable d'interagir, simultanément, avec deux sites du transcrit assez éloignés. Les études enzymatiques (Geiselmann et al., 1993; Walstrom et al., 1997) semblent indiquer par ailleurs que Rho ne nécessite pas le maintien de son interaction primaire pour orienter sa translocation dans le sens 5' vers 3'.

Aucune des deux hypothèses ne peut donc être favorisée. On peut alors penser que, faute de réelle fonction au cours de la translocation, le maintien de cette première interaction est aléatoire ou du moins transitoire. En fonction de la force de l'interaction, ou de contraintes apportées par d'autres facteurs, l'interaction primaire sera maintenue ou non. On peut également imaginer que cette interaction évolue au cours du temps (en fonction de la composition de l'ARN en C). Mais il semblerait au regard des différents travaux cités précédemment que celle-ci soit facultative.

4. Activité ATPase.

Le domaine de fixation et d'hydrolyse de l'ATP correspond à la partie C-terminale de Rho (résidus 167 à 419) (Richardson, 1996). La structure de la partie C-terminale n'avait, jusqu'à récemment, pas été déterminée. L'étude mécanistique de Rho a pendant longtemps nécessité d'établir un modèle moléculaire de ce domaine basé sur ses homologies avec certaines protéines. La protéine la plus proche de Rho est la F₁ ATP synthase, plus exactement la sous-unité β fortement homologue au domaine de fixation à l'ATP de Rho (Dombroski et

al., 1988). Ainsi, jusqu'à la publication de la structure de l'hexamère entier par Skordalakes et Berger, la structure de la F₁ ATP synthase a longtemps servi de base pour simuler la structure du site de fixation ATP de Rho (sous forme monomérique ou hexamérique).

Bien que l'ATP soit le substrat préférentiel de Rho, celui-ci est capable d'hydrolyser les trois autres NTPs (Lowery et Richardson, 1977) mais avec une efficacité moindre. Apparemment cette différence d'activité est liée à une conformation différente de l'hexamère en fonction du NTP fixé (Brennan et al., 1990). Comme nous l'avons vu, la fixation d'ARN au niveau du site secondaire de Rho est nécessaire à l'hydrolyse de l'ATP. Cependant, même en absence d'ARN, Rho est capable de fixer trois molécules d'ATP. Il semblerait en effet que chaque hexamère possède trois sites forts de liaison à l'ATP (Stitt et Xu, 1998). Deux points de vue s'opposent actuellement concernant l'existence de trois autres sites de liaison faible (Kim et Patel, 1999; Stitt, 2001). Mais l'existence de ces derniers est difficile à vérifier et leur fonction, si fonction il y a, reste inconnue.

La plupart des études montrent qu'en présence d'ARN, l'interaction entre Rho et l'ATP est modifiée, de même la fixation d'ATP influence l'interaction entre l'ARN et le site secondaire de Rho (Kim et Patel., 2001). Lorsque l'ARN entre en contact avec le site secondaire, il induit une série de changements conformationnels au sein des domaines C-terminaux de l'hexamère. Ces modifications structurales entraînent l'hydrolyse séquentielle des trois molécules d'ATP (Stitt et Xu, 1998). Cette hydrolyse aboutit à son tour à des modifications conformationnelles qui vont affecter l'interaction entre l'ARN et le site secondaire afin de permettre la translocation.

Les détails moléculaires de ce mécanisme restent assez flous. La résolution d'une structure de Rho en interaction avec l'ARN, au niveau de son site secondaire, permettra sans doute de dévoiler le mécanisme complet du couplage entre l'hydrolyse de l'ATP et la translocation de Rho le long de l'ARN.

5. Activité hélicase.

Il reste également à savoir comment, en fin de réaction, Rho est capable de dissocier l'ARN du complexe ternaire d'élongation.

Rho possède de nombreuses homologies structurales et fonctionnelles avec les hélicases hexamériques (pour revue : [Patel et Picha, 2000](#)). La réaction de dissociation Rho-dépendante n'est pas spécifique du complexe ternaire d'élongation. En effet Rho peut induire, *in vitro*, la dissociation d'un oligonucléotide ADN hybridé au transcrit en 3' d'un site RUT ([Brennan et al., 1987](#); [Walstrom et al., 1997](#)). Cette activité ARN/ADN hélicase est dépendante de l'hydrolyse de l'ATP. C'est un mécanisme unidirectionnel, la dissociation de l'hybride se faisant dans l'orientation 5'→3' du transcrit ([Walstrom et al., 1997](#)).

La vitesse de la translocation est dépendante du degré de structuration secondaire le long du transcrit ([Walstrom et al., 1998](#)). En effet, la présence de structures secondaires stables en aval du site RUT déplace la position des sites de terminaison en aval ([Zhu et von Hippel, 1998a](#)). Néanmoins l'activité hélicase de Rho n'est pas bloquée par la présence de structure en tige boucle dans l'ARN ([Steinmetz et al., 1990](#)).

L'activité hélicase de Rho est souvent avancée pour expliquer la dissociation du complexe ternaire d'élongation ([Walstrom, 1997a; b et 1998](#)). Rho provoquerait la dissociation du complexe en détruisant l'hybride ADN/ARN au sein de la bulle de transcription. Cependant, lorsqu'on observe la structure du complexe ternaire d'élongation ([Murakami et Darst, 2003](#); [Darst, 2001](#); [Toulokhonov et Landick, 2003](#); [Figure 3](#)), on peut se demander comment Rho peut accéder à cet hybride enfoui au sein d'un complexe multiprotéique extrêmement stable. L'une des hypothèses est qu'un site de pause structuré pourrait déstabiliser le complexe ternaire d'élongation et le rendre plus sensible à l'action de Rho. Un autre modèle suggère que Rho entre en contact avec l'ARN polymérase par des interactions non spécifiques (ou grâce à un cofacteur cellulaire) et que ces contacts provoquent des changements allostériques capables de déstabiliser l'ARN polymérase. Enfin une troisième hypothèse suggère que la dissociation du complexe d'élongation n'est pas due à la fonction hélicase de Rho. Dans ce dernier modèle, Rho n'accéderait pas à l'hybride ADN/ARN, mais il provoquerait son extraction hors de la bulle de transcription grâce à la force motrice de son activité translocase. En brisant ainsi le réseau de coordination au sein du site actif de l'ARN polymérase, Rho provoquerait la dissociation du complexe ternaire d'élongation. Les signaux déstabilisant l'hybride ADN/ARN (site de pause de classe II) favoriseraient alors ce dernier mécanisme.

6. NusG : cofacteur de la terminaison Rho-dépendante.

Comme je l'ai expliqué précédemment, le processus de terminaison Rho-dépendante peut être vu, en partie, comme une compétition cinétique entre la vitesse d'élongation de l'ARN polymérase et la vitesse de translocation de Rho le long du transcrit natif.

Dans ce contexte, NusG, une protéine de 21kDa essentielle pour *E. coli* (Burns et Richardson, 1995; Burns et al., 1998), régule la terminaison Rho-dépendante de façon assez paradoxale. NusG a tout d'abord été découverte pour sa participation au mécanisme d'antiterminaison N-dépendante chez le bactériophage λ (voir chapitre suivant). On a ensuite découvert que NusG agissait comme un facteur d'élongation de la transcription permettant à l'ARN polymérase de résister à certains signaux de pause (Artsimovitch et Landick, 2000). Cependant NusG semble aussi essentielle pour certains terminateurs Rho-dépendants (Pasman et von Hippel, 2000) : NusG déplace les sites de terminaison Rho-dépendants à des positions situées en amont sur le substrat (Nehrke et Platt, 1994; Burns et Richardson, 1995).

NusG fixe, via des interactions faibles, à la fois l'ARN polymérase (Burns et Richardson, 1995; Pasman et von Hippel, 2000) et Rho (Li et al., 1993, Sullivan et Gottesman, 1992). On ne sait pas cependant s'il est capable de fixer les deux simultanément.

Deux hypothèses sont avancées pour expliquer le rôle de NusG.

- La première hypothèse suggère que NusG en se fixant à l'ARN polymérase favorise une conformation de cette dernière plus sensible à l'action de Rho et au relargage du transcrit (Nehrke et Platt, 1994; Pasman et von Hippel, 2000).
- D'un autre côté si NusG fixe simultanément l'ARN polymérase et Rho, on peut penser qu'il permet d'orienter Rho par rapport à l'ARN polymérase de façon à favoriser son interaction avec le transcrit à la sortie du complexe ternaire d'élongation, cependant cette hypothèse semble structurellement peu probable (voir : Richardson, 2002).

7. NusA : antagoniste de Rho?

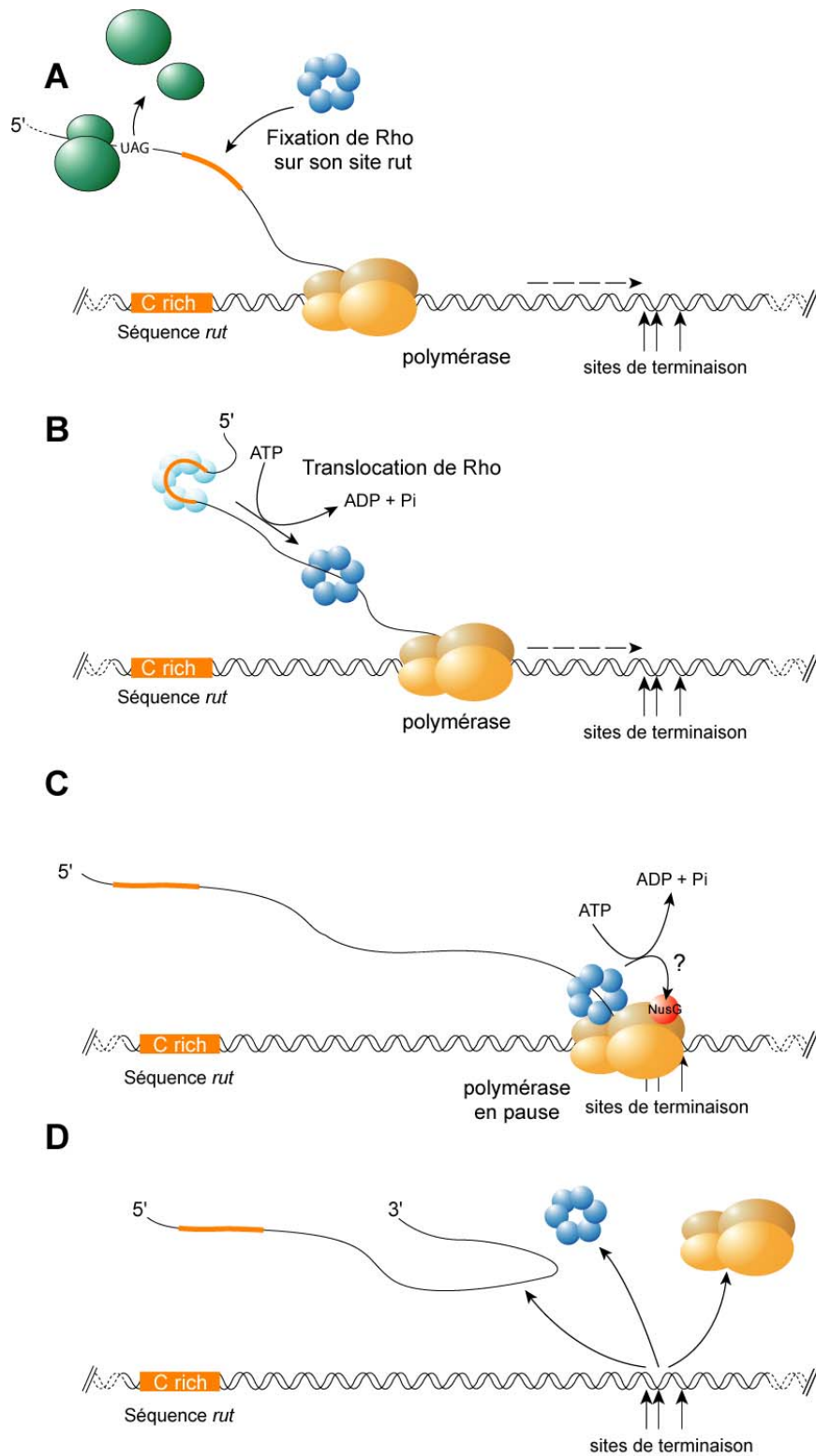
NusA, comme NusG, est capable de fixer Rho et l'ARN polymérase, ces fixations peuvent se faire simultanément (Schmidt et Chamberlain, 1984). Comme NusG, c'est un

facteur nécessaire à l'antitermination chez le phage λ . C'est également un facteur d'élongation mais, à l'inverse de NusG, NusA favorise la pause de l'ARN polymérase (Artsimovitch et Landick, 2000; Lau et al., 1983). Son action vis-à-vis de Rho est à première vue assez ambiguë. Selon les études NusA favoriserait, ou défavoriserait, la terminaison par le facteur Rho. NusA est également essentielle pour la viabilité de la cellule (Kainz et Gourse, 1998). Cependant, un mutant conditionnel de NusA (*nusA10(Cs)*; Craven et Friedman, 1991) peut être compensé par certaines mutations de Rho, pour lesquelles le facteur de terminaison possède une activité réduite, suggérant que les deux protéines ont des fonctions opposées (Zheng et Friedman, 1994).

On sait aujourd'hui que l'un des rôles majeur de NusA est de favoriser le couplage entre transcription et traduction (Kung et al., 1975; Zheng et Friedman, 1994). La présence d'un ribosome, traduisant l'ARNm à la vitesse où celui-ci est synthétisé, bloque l'accès de Rho à l'ARN naissant, jusqu'à la fin d'un gène ou d'un opéron. Si le couplage entre transcription et traduction est interrompu, Rho peut se fixer, s'il trouve un segment adapté sur le transcrit nu, et initier le processus de terminaison. NusA, en retardant l'ARN polymérase au niveau de certains sites de pause, évite aux ribosomes de prendre du retard sur le complexe d'élongation et de laisser l'ARN libre d'accès pour le facteur Rho. Dans ce contexte NusA, bien que réduisant la processivité de l'ARN polymérase, agit contre la terminaison Rho-dépendante.

8. Résumé de l'action de Rho.

En cours de transcription, Rho, sous la forme d'un hexamère, reconnaît et fixe un segment du transcrit, déstructuré, dépourvu de ribosomes et riche en résidus C, nommé RUT (Figure 15A). Cette reconnaissance se fait grâce à son site primaire d'interaction localisé dans les extrémités N-terminales des différents protomères de Rho. La conformation ouverte de l'hexamère permet l'entrée de l'ARN dans le centre de l'anneau où il interagit avec un site secondaire. Cette seconde fixation, couplée à des changements conformationnels, entraîne l'hydrolyse de l'ATP. L'énergie issue de cette hydrolyse permet à Rho de se déplacer de 5' en 3' le long du transcrit (Figure 15B). Quand Rho entre en contact avec l'ARN polymérase arrêtée au niveau d'un site de pause (Figure 15C) il provoque la dissociation du complexe ternaire d'élongation (Figure 15D). Actuellement, les détails précis de cette dernière étape restent inconnus.



Terminaison de la transcription

Figure 15 : Représentation schématique du mécanisme de terminaison Rho-dépendante chez *E. coli*. (A) Rho reconnaît et fixe son site RUT au niveau du transcrit, puis transloque de 5' vers 3' (B); rejoint l'ARN polymérase arrêtée (ou déstabilisée) au niveau d'un site de pause et dissocie le complexe d'élongation (C) et (D).

9. Rôles biologiques de Rho.

On retrouve chez un grand nombre de bactéries, dont certaines proviennent de lignées extrêmement divergentes, des gènes homologues à *rho* (Opperman et Richardson, 1994; voir Figure 8, page 16). Si Rho s'est maintenu aussi longtemps au cours de l'évolution c'est qu'il apporte un certain nombre d'avantages à la bactérie.

Chez *E. coli* les terminateurs transcriptionnels ont trois principales fonctions dans la régulation de l'expression des gènes :

- Ils définissent tout d'abord la fin de gènes ou d'opérons, de façon à maintenir l'indépendance transcriptionnelle de séquences génomiques adjacentes.
- Positionnés entre deux gènes d'un même opéron, ou entre le promoteur et le premier gène d'un opéron, ils permettent de moduler le niveau relatif d'expression des différents gènes d'un même opéron. Dans ce cas, on les appelle des atténuateurs transcriptionnels.
- Enfin, des terminateurs sont aussi présents dans les gènes, toutefois ils sont latents et sont activés seulement lorsque la traduction et la transcription se retrouvent découplées.

Les deux premières catégories de terminateurs, extragéniques et intergéniques, peuvent être à la fois intrinsèques ou dépendants de Rho. Les terminateurs intragéniques eux sont majoritairement, si ce n'est exclusivement, Rho-dépendants. Ils ne fonctionnent pas dans les conditions normales d'expression, probablement parce que la présence de ribosomes sur le transcrit empêche la fixation de Rho. Par contre, lorsque le ribosome est bloqué ou ralentit (par une mutation non sens ou une carence en acide aminé), l'ARN est alors dépourvu de protéines et peut être reconnu par Rho, qui induit alors la terminaison. Ainsi dans un opéron, une mutation nonsense dans un gène provoque la diminution de l'expression des gènes situés plus en aval. Ce mécanisme est appelé "polarité transcriptionnelle" (Stanssens et al., 1986; Richardson, 1991). L'un de ses avantages est, qu'en condition de stress, Rho prévient la perte d'énergie issue de la transcription d'un ARNm qui ne pourrait être traduit entièrement. Par exemple chez *E. coli*, en condition de carence en acides aminés, Rho termine précocement la transcription au sein de l'opéron *lac* (Richardson, 1991; Ruteshouser et Richardson, 1989) : cet opéron contient en effet plusieurs terminateurs Rho latents. Ce phénomène est rendu possible grâce à la diversité de séquence des terminateurs Rho-dépendants.

Rho peut avoir un autre rôle que la terminaison de la transcription : comme la structure en tige boucle d'un terminateur intrinsèque (voir chapitre I), Rho peut moduler la stabilité et la durée de vie des ARNm. En effet les ARNs fixés par le facteur Rho sauvage sont dégradés beaucoup plus rapidement dans des mutants *nusD* (Sozhamannan et Stitt, 1997). Ces mutants produisent des protéines Rho dont la fixation à l'ARN est altérée (Stitt et Mosig, 1989; Washburn et Stitt, 1996). Ces résultats suggèrent que Rho stabilise les ARNs auxquels il se fixe. Cette faculté n'est pas le propre de Rho : le même effet protecteur a été montré pour d'autres protéines, par exemple le facteur de réplication Rop du pBR322 (Sozhamannan et Stitt, 1997).

10. Un facteur Rho eucaryote?

Il semble que le facteur Rho n'ait pas été conservé au cours de l'évolution des cellules procaryotes vers les cellules eucaryotes : aucun analogue de Rho n'a encore été découvert chez les cellules nucléées, ni en terme de séquence, ni en terme de fonction.

Contrairement aux procaryotes, les eucaryotes utilisent trois ARN polymérases différentes pour synthétiser les ARNr (ARN polymérase I), les ARNm (ARN polymérase II) et les ARNt (ARN polymérase III). Des études essentiellement réalisées chez la levure suggèrent toutefois que les mécanismes de terminaison utilisés par les ARN polymérase eucaryotes sont apparentés aux mécanismes de terminaison intrinsèque et Rho-dépendante d'*E. coli*. Par exemple, l'ARN polymérase I effectue une pause induite par une protéine fixant l'ADN et nommée Reb1p (Lang et al., 1994; Liu et al., 1998), puis termine la transcription au niveau d'une séquence riche en T (Lang et Reeder, 1995). La RNAP III, elle, ne nécessite apparemment qu'une répétition de T sur le brin codant (Campbell et Setzer, 1992; Matsuzaki et al., 1994; Allisson et Hall, 1985). Ces deux ARN polymérases adoptent donc un mécanisme proche de la terminaison intrinsèque des procaryotes, même si la présence d'une structure en tige boucle n'est pas requise.

A l'inverse la façon dont l'ARN polymérase II termine la transcription suggère un mécanisme plus proche de la terminaison Rho-dépendante (Proudfoot, 1989). L'ARN polymérase II doit d'abord passer par une séquence spécifique avant de terminer la transcription au niveau de sites multiples le long d'une séquence étendue. En 1998, Reeder et ses collaborateurs ont montré que Rho peut induire, *in vitro*, le relargage du transcrit par

Introduction.

l'ARN polymérase II de levure (terminaison toutefois dépendante de Reb1p; [Lang et al., 1998](#); [Wu et Platt, 1993](#)). Par contre, Rho n'a aucun effet sur les ARN polymérases I et III. Ces observations signifient que l'action de Rho est spécifique de l'ARN polymérase II et laissent imaginer que l'ARN polymérase II termine la transcription en utilisant un mécanisme proche de la terminaison Rho-dépendante des procaryotes ([Proudfoot, 1989](#)). Ceci reste cependant à vérifier.

Chapitre III. REGULATION DE LA TERMINAISON.

L'initiation de la transcription est considérée, à juste titre, comme la cible privilégiée des mécanismes de régulation de la majorité des organismes, toutefois, d'autres étapes intervenant en aval de l'initiation sont également soumises à une régulation très précise. En particulier, la régulation de l'expression génique par la terminaison précoce de la transcription est une stratégie fréquemment utilisée chez la bactérie. En fait, la terminaison est rarement totale, et le niveau d'expression des gènes situés en aval du terminateur peut être contrôlée en modifiant l'efficacité du terminateur.

Au cours de leur évolution, les bactéries, mais aussi les phages, ont mis en place un très grand nombre de signaux permettant de réguler la terminaison. Il serait impossible d'en faire une liste exhaustive, mais on peut cependant les classer en deux catégories bien distinctes :

- L'atténuation de la transcription (pour revue : [Yanofsky, 2000](#)).
- L'antiterminaison de la transcription ([Nudler et Gottesman, 2002](#)).

Dans les paragraphes suivants je présenterai brièvement quelques exemples représentatifs de ces deux classes de signaux.

1. Atténuation.

L'atténuation correspond au contrôle de la transcription d'un ou plusieurs gènes proches les uns des autres par la régulation de la terminaison. Ce type de signal met en jeu la formation dans l'ARN d'un motif alternatif, nommé antiterminateur, au voisinage du terminateur. Les deux structures, antiterminatrice et terminatrice, partagent une partie de leur séquence, de sorte que leur formation est mutuellement exclusive ([Lee et Yanofsky, 1977](#); [Landick et al., 1996](#); [Yanofsky, 2000](#)). Aussi, les facteurs (protéines ou ARN) qui favorisent ou défavorisent la formation de l'une ou l'autre des structures permettent de faire varier l'efficacité de la terminaison au niveau de ces sites.

La plupart des opérons impliqués dans la synthèse ou l'utilisation des acides aminés sont régulés au niveau transcriptionnel par ce type de mécanisme. Dans ces opérons, la terminaison est régulée par la concentration intracellulaire de l'acide aminé en question. Ces opérons sont généralement caractérisés par une région "leader" localisée en amont du premier gène structural (Henkin, 1996; 2000; Henkin et Yanofsky, 2002). Cette région leader contient une séquence codant pour un peptide court. Cette séquence codante est localisée à une position cruciale dans le transcrit et porte un ou plusieurs codons de l'acide aminé en rapport avec l'opéron concerné. Par exemple : l'opéron *his* de *Salmonella typhimurium* contient une région leader qui peut former les deux structures mutuellement exclusives, terminateur et antitermineur. Cette région contient également une séquence codant un peptide leader et portant sept codons Histidine (His) répétés en tandem (Landick et al., 1996). Lorsque la séquence leader est transcrite, une carence en ARNt^{His} chargés conduit à l'arrêt des ribosomes au niveau de ces codons His. L'antitermineur se forme alors et inhibe la formation du terminateur intrinsèque situé en aval, permettant ainsi à l'ARN Polymérase de poursuivre la transcription de l'opéron. A l'inverse, lorsque l'His est présente dans la cellule en quantité suffisante, les ribosomes poursuivent la traduction du peptide leader et empêchent la formation de l'antitermineur, favorisant ainsi la terminaison. Le même type de mécanisme est responsable de l'atténuation au niveau de l'opéron *trp* d'*E. coli* (Landick et al., 1996).

Les ribosomes ne sont pas les seuls facteurs protéiques capables de réguler la formation des structures antiterminatrices. Chez *B. subtilis*, le facteur TRAP (pour "Trp RNA Attenuation Protein") est une protéine organisée en onze sous-unités identiques de 8kDa. Ce facteur possédant 11 sites de fixation au Tryptophane (Trp) est activé par la présence de Trp dans la cellule. Cette activation favorise la fixation de TRAP à la région leader du transcrit, l'ARN s'enroule alors autour de la protéine et la formation de l'antitermineur est inhibée. Le facteur TRAP favorise donc la terminaison de la transcription de l'opéron *trp* de *B. subtilis* (Antson et al., 1995 et 1999).

Chez *E. coli*, l'opéron *bgl*, qui contient trois gènes : *bglG*, *bglF* et *bglM*, est impliqué dans l'utilisation des sucres aromatiques. En présence de ces sucres, la protéine BglG est phosphorylée par le produit du gène *bglF*, elle se fixe alors au transcrit au niveau de la structure antiterminatrice et, contrairement à l'exemple précédent, stabilise cette structure qui est moins forte que celle du terminateur. La protéine BglG favorise ainsi l'antiterminaison chez *E. coli* (Amster-Choder et Wright, 1993).

Enfin, l'atténuation peut également être régulé par des facteurs nucléotidiques; ainsi, chez les bactéries gram-positives, les opérons codant pour certaines aminoacyl transférases sont régulés par des ARN de transfert (ARNt) non chargés qui stabilisent la formation de l'antitermineur (Grundy et Henkin, 1993; Grundy et al., 2002).

Les mécanismes décrits ci-dessus concernent exclusivement les terminateurs intrinsèques, pourtant les terminateurs Rho-dépendants sont également sujets à ce type de contrôle. Nous l'avons vu, l'étape primordiale de la terminaison Rho-dépendante est la reconnaissance et la fixation de Rho sur une séquence spécifique du transcrit. Cette fixation nécessite que l'ARN soit nu, c'est-à-dire dépourvu de protéine. Ainsi, la présence d'un ribosome en cours de traduction sur le transcrit empêche l'initiation de la terminaison. Chez *E. coli* la transcription de l'opéron tryptophanase (*tna*) est régulée par la terminaison Rho-dépendante (Landick et al., 1996; Konan et Yanofsky, 2000). L'opéron contient une région leader nommée *tnaC*, située juste en amont du terminateur Rho-dépendant, et codant pour un peptide de 24 acides aminés contenant un unique et essentiel Trp. L'excès de Trp dans la cellule inhibe le clivage du peptide TnaC naissant et bloque le ribosome au niveau du codon stop de TNAC, or le site de fixation de Rho est adjacent à ce codon stop. La présence du ribosome interfère donc avec la fixation de Rho à l'ARN et inhibe la terminaison (Gong et al., 2001).

L'ensemble des mécanismes décrits précédemment met en jeu la fixation d'une protéine ou d'un ARN qui va favoriser, défavoriser, ou masquer, une structure ou une séquence terminatrice située à proximité de son site de fixation sur le transcrit.

Une deuxième classe de mécanismes, l'antiterminaison, fait cette fois intervenir la modification directe du complexe ternaire d'élongation par des protéines ou des acides nucléiques. Cette modification peut affecter la terminaison intrinsèque et/ou moduler l'action du facteur Rho. Contrairement à l'atténuation, cette antiterminaison peut être processive, c'est-à-dire que l'ARN polymérase peut résister à des terminateurs situés loin en aval du signal d'antiterminaison.

2. L'antiterminaison processive.

Le principe de l'antiterminaison processive a été découvert chez le phage λ (Roberts,

1969). Depuis d'autres exemples ont été mis en évidence dans des opérons bactériens, en particulier dans les opérons *rrn* qui codent pour les ARN ribosomaux et certains ARN de transfert (Morgan, 1980; Aksoy et al., 1984; Holben et Morgan, 1984; Albrechsten et al., 1990).

2. 1. Antiterminaison dans les opérons ribosomaux.

Chez *E. coli*, la biosynthèse des ribosomes consomme environ 50% de l'énergie métabolique disponible chez la bactérie en phase de croissance exponentielle (Bremer et Denis, 1987). La synthèse des ARNs et des protéines qui composent les ribosomes doit donc être extrêmement bien régulée.

Bien qu'ils ne soient pas traduits et qu'ils contiennent de nombreux terminateurs Rho-dépendants, les opérons ribosomaux ne sont pas soumis au phénomène de polarité (Morgan, 1980; Brewster et Morgan, 1981). Ceci permet d'assurer par exemple la stœchiométrie de synthèse des ARNr 16S, 23S, et 5S. Cette absence de polarité peut être attribuée à un mécanisme d'antiterminaison. Ce mécanisme d'*E. coli* fait intervenir : l'ARN polymérase, les protéines NusA (Friedman et Baron, 1974), NusB (Keppel et al., 1974), NusG (Downing et al., 1990), et la protéine ribosomal S10 (NusE; Friedman et al., 1981). NusA, NusB, NusG et S10 sont indispensables (Vogel et Jensen, 1997), mais ne sont pas suffisantes pour induire l'antiterminaison (Squires et al., 1993), suggérant l'intervention d'un ou plusieurs autres facteurs (probablement des protéines ribosomales) dans ce mécanisme (Squires et al., 1993; Torres et al., 2001).

NusB reconnaît et fixe une séquence de 12 nucléotides (UGCUCUUUAACA) nommée BOXA, située dans la région 5' du transcrit des ARNr (Nodwell et Greenblatt, 1993) et absolument conservée dans chacun des sept opérons ribosomaux d'*E. coli*. Cette fixation nécessite l'interaction entre NusB et S10 (Nodwell et Greenblatt, 1993; Lüttgen et al., 2002) et initie la formation d'un complexe multiprotéique qui stabilise le complexe ternaire d'élongation. L'ARN polymérase ainsi modifiée peut franchir les terminateurs Rho-dépendants mais pas les terminateurs intrinsèques (Albrechtsen et Squires, 1990; Condon et al., 1995).

Le mécanisme précis qui induit cette antitermination est encore inconnu. La nature des interactions entre les différents composants du complexe d'antitermination est résumée schématiquement sur la [Figure 16](#).

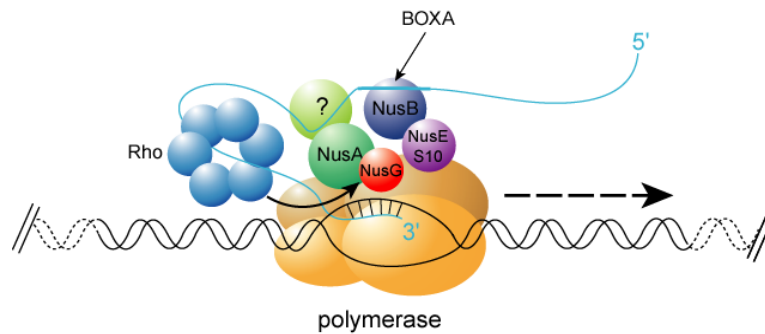


Figure 16 : Conformation hypothétique du complexe d'antitermination d'*E. coli* (modifié selon : [Greenblatt et al., 1993](#)). Le point d'interrogation représente un ou plusieurs facteurs supplémentaires, probablement des protéines ribosomales.

Certains bactériophages, tel λ , se sont appropriés le mécanisme d'antitermination d'*E. coli* et en utilisent les composants pour réguler la transition entre les différentes étapes de leur propagation ([Das, 1993](#)).

2. 2. Antitermination chez le phage λ .

Le bactériophage λ est le prototype de toute une catégorie de phages, dits tempérés, qui possèdent deux cycles de vie bien particuliers :

- D'une part ils peuvent infecter les cellules, utiliser la machinerie cellulaire pour générer de nouvelles particules et provoquer la lyse de la cellule hôte (comme le fait le bactériophage T7 par exemple). Cette première phase est nommée cycle lytique.
- D'autre part, ils peuvent infecter la cellule, s'intégrer dans le génome de leur hôte et demeurer ainsi, se répliquant comme partie intégrante du chromosome, jusqu'à l'apparition d'un signal (un stress environnemental par exemple) qui va activer la production de nouvelles particules et finalement la lyse de la cellule hôte. Ce cycle est nommé lysogénique.

Après infection d'*E. coli*, le phage λ est maintenu en phase lysogénique par la présence d'un répresseur transcriptionnel, produit du gène *cI*. Ce gène est le seul exprimé au cours de cette période de la vie du phage. Le cycle lytique débute lorsque la répression est levée (en fonction des conditions environnementales), l'ARN polymérase de l'hôte initie alors la transcription précoce du génome du phage λ à partir de deux promoteurs divergents nommés pL et pR (voir : [Figure 17](#)). Les premiers gènes exprimés sont *cro* et *N*, puis la transcription est stoppée au niveau de deux terminateurs Rho-dépendants nommés tL1 et tR1 ([Richardson et Greenblatt, 1996](#)).

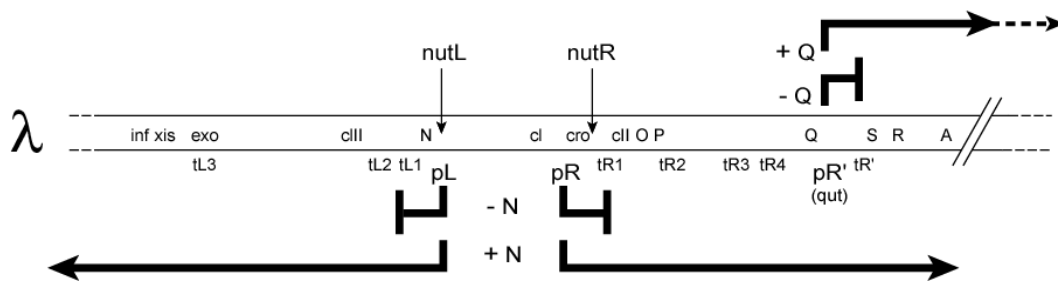


Figure 17 : Carte génétique partielle du bactériophage λ , indiquant les effets transcriptionnels des protéines *N* et *Q*. (d'après : [Greenblatt et al., 1993](#)).

Le produit du gène *cro* entre en compétition avec le répresseur *cI* pour la fixation aux opérateurs *oL* et *oR* et permet d'activer la transcription au niveau des promoteurs pL et pR. Il favorise ainsi sa propre expression, et celle de la protéine *N*. Une fois produite, la protéine *N* rend l'ARN polymérase résistante à tous les terminateurs présents dans les deux opérons précoces (tL1-tL3 et tR1-tR4). Ceci permet à l'ARN polymérase de la cellule hôte de transcrire plus efficacement les gènes situés en aval, entre autre : les gènes *O* et *P* impliqués dans la réplication de l'ADN phagique, et le gène *Q* situé plus loin dans l'opéron précoce pR. La protéine *Q* régule ensuite la transcription tardive du génome de λ , permettant l'expression de gènes impliqués dans la lyse de la cellule hôte et la morphogenèse du bactériophage ([Roberts, 1988](#)).

2. 2. 1. Antitermination dirigée par *N*.

Bien que composée de seulement 107 acides aminés (environ 12kDa), et totalement

déstructurée en solution (Mogridge et al., 1998; Van Gilst et von Hippel, 1997), N est la clé de l'élongation de la transcription du génome du phage λ (Greenblatt et al., 1993; Friedman et Court, 1995; Weisberg et Gottesman, 1999).

Sa capacité à inhiber la terminaison de la transcription repose sur sa participation à un complexe d'antiterminaison proche de celui décrit précédemment pour la régulation des opérons *rrn*. Ce complexe comprend quatre facteurs cellulaires plus l'ARN polymérase de l'hôte *E. coli*. La formation de ce complexe est liée à la présence, dans le génome de λ , d'une séquence cis-régulatrice appelée *nut* (pour "N-utilization"; Richardson et Greenblatt, 1996). Cette séquence a été localisée au niveau des deux opérons précoces, droit et gauche, du phage λ . Il existe donc deux sites *nut* dans le génome de λ , *nutL* et *nutR* (Figure 17).

On a découvert que N pouvait supprimer la polarité causée par Rho dans les opérons *trp* et *gal* portant des mutations nonsense, mais seulement si la transcription de ces deux opérons était initiée au niveau des promoteurs précoces de λ . Ceci a suggéré que ces séquences précoces contenaient un élément nécessaire à l'action de N (Adhya, 1974; Franklin, 1974; Segawa et Imamoto, 1974).

L'isolement et le séquençage de mutations du site *nutL* de λ pour lesquelles N est inactif ont permis d'isoler les séquences, situées en aval du promoteur pL, requises pour l'action de N (Salstrom et Szybalsky, 1978). Puis, par homologie, le site *nutR* a été identifié (Rosenberg et al., 1978). Ces séquences, placées entre un promoteur et un terminateur, sont capables, en présence de N, d'activer l'expression d'un gène rapporteur situé en aval du terminateur (de Crombrughe et al., 1979). On a ensuite découvert que ces sites *nut* agissaient au niveau du transcrit ARN. En effet, N ne se complexe avec l'ARN polymérase qu'après transcription des sites *nut*, et la traduction par des ribosomes des séquences NUT bloque l'action de N (Olson et al., 1982). Par ailleurs Nodwell et Greenblatt ont montré, un peu plus tard, que le complexe d'antiterminaison protège le transcrit NUT de la digestion par les RNases (Nodwell et Greenblatt, 1991).

Les sites *nut* sont formés par deux éléments distincts, *boxA* et *boxB*, séparés par une séquence aléatoire ("linker") (Friedman et Court, 1995). La *boxA* est une cassette de 12 nucléotides dont la séquence est hautement conservée, non seulement chez les phages de type λ , mais également au niveau de l'élément antitermineur des gènes des ARN ribosomiaux d'*E. coli* (Figure 18). La séquence de *boxB*, absente du système *rrn*, est beaucoup moins

conservée mais correspond toujours à une séquence répétée-inversée interrompue qui aboutit à la formation d'une structure en tige-boucle au niveau du transcrit.

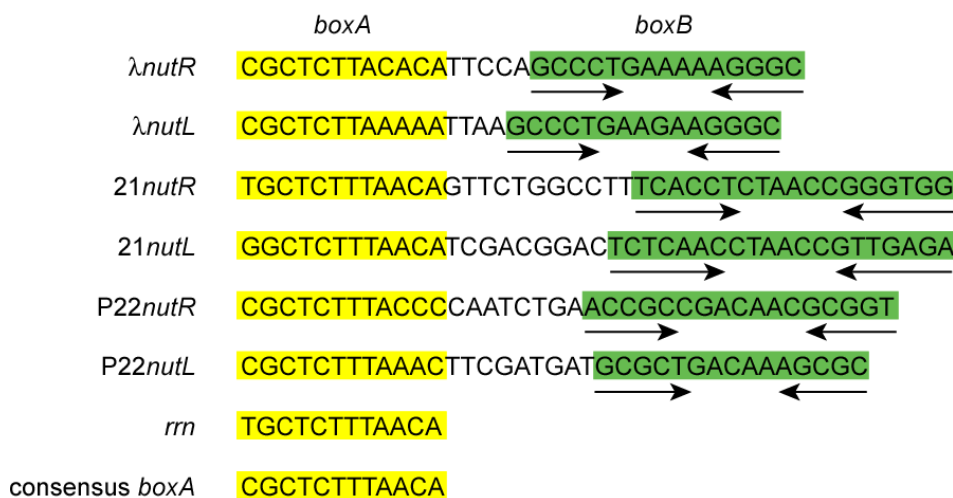


Figure 18 : Comparaison des séquences *nut* de certains phages de type λ et de la *boxA* des opérons ribosomiaux. Les flèches correspondent à des séquences répétées-inversées. Pour λ et P22 ces séquences forment des structures en tige boucle dans l'ARN (Su et al., 1997; Cai et al., 1998; Legault et al., 1998)

N est une petite protéine basique (12kDa), qui possède un domaine riche en Arginine (ARM) présent chez certaines protéines qui fixent spécifiquement les ARN. La protéine N se fixe ainsi au transcrit de λ au niveau du motif en tige-boucle BOXB. Sa structure, relativement désorganisée en solution, devient beaucoup mieux définie au contact de l'ARN, en particulier son extrémité N-terminale (résidus 1 à 22) se structure en hélice α au contact de l'ARN. Cette interaction favorise la formation d'une boucle de type GNRA dans le motif BOXB, par l'extrusion du quatrième nucléotide de la boucle (Legault et al., 1998). La région C-terminale de N (résidus 23 à 107) est également requise pour l'antiterminaison processive (Mogridge et al., 1998).

On retrouve des homologues du système N/NUT (Franklin, 1985) dans tous les autres phages de type λ (à l'exception d'un seul dont je reparlerai plus loin dans ce chapitre). Cependant, l'action de la protéine N d'un bactériophage est spécifique des séquences *nut* de ce même bactériophage (Friedman et al., 1973).

- Rôles des facteurs Nus :

Des souches mutantes d'*E. coli* chez lesquelles le phage λ ne peut pas se développer ont été isolées. Quatre de ces mutants, (*nusA*, *nusB*, *nusE* et *nusG*), présentent un défaut de fonctionnement de N. Ces gènes codent pour les facteurs cellulaires qui modulent la terminaison et l'antiterminaison chez *E. coli* (voir précédemment). Des études *in vitro* ont ensuite montré que ces facteurs cellulaires étaient nécessaires à l'action de N (Das et Wolska, 1984; Horwitz et al., 1987).

- NusA est une protéine de 56kDa essentielle à la survie de la bactérie. Elle affecte la processivité de l'ARN polymérase et intervient dans le contrôle des phénomènes de pause, de terminaison et d'antiterminaison (Kainz et Gourse, 1998; Artsimovitch et Landick, 2000).
- NusB semble être essentielle seulement à basse température (Friedman et al., 1976; Georgopoulos et al., 1980).
- NusE correspond à la protéine ribosomale S10 (Friedman et al., 1981).
- Enfin NusG (Sullivan et Gottesman, 1992) est impliquée dans les interactions entre Rho et l'ARN polymérase. Son action est encore floue puisqu'elle est nécessaire à l'antiterminaison alors que, comme nous l'avons vu, elle favorise la terminaison Rho-dépendante.
- Deux autres mutants, *nusC* et *nusD*, ont été isolés. Ils correspondent respectivement : à un mutant de la sous-unité β de l'ARN polymérase incapable d'interagir avec la protéine N (Georgopoulos, 1971; Ghysen et Pironio; 1972); et à un mutant de Rho résistant à l'action de N (cependant, chez ce dernier mutant, N agit toujours sur les terminateur intrinsèques; Das et al., 1983).

La fixation de N sur la boxB provoque la formation d'un complexe incluant : N, l'ARN polymérase et l'ensemble des facteurs Nus. L'assemblage du complexe repose sur de nombreuses interactions protéine-protéine au sein du complexe ternaire d'élongation.

La base du complexe réside dans l'interaction entre, N et NusA d'une part, NusA et l'ARN polymérase d'autre part (Greenblatt et Li 1981a et b). Certains travaux suggèrent que deux molécules de NusA participent au complexe (Horwitz et al., 1987). NusA peut interagir avec la partie C-terminale de N (Greenblatt et Li, 1981a et b), avec des composants de BOXA et BOXB (Mogridge et al., 1995; Friedman et Olson, 1983) et avec la sous-unité α de l'ARN

polymérase (Mah et al., 1999; Gill et al., 1991). NusB et S10 interagissent l'un avec l'autre *in vitro* (Mason et Greenblatt 1992). Le complexe NusB-S10 reconnaît et fixe BOXA par l'intermédiaire de NusB, et l'ARN polymérase par l'intermédiaire de S10 (Mason et Greenblatt, 1991). Cependant, contrairement à ce qui se passe dans les opérons ribosomiaux, le complexe NusB-S10 n'est pas capable de fixer la BOXA du phage λ en l'absence de N et de BOXB (Mason et Greenblatt, 1992; Nodwell et Greenblatt, 1993). N est indispensable à la fixation des facteurs Nus sur les sites *nut* du phage λ , probablement parce que la séquence *boxA* du phage dévie légèrement de la *boxA* originelle des opérons *rrn*. Cette différence a vraisemblablement été mise en place par le phage λ pour maintenir l'expression précoce de ses gènes sous le contrôle de N. Enfin, NusG interagit avec l'ARN polymérase (Li et al., 1992). La fixation de NusG et S10 sur l'ARN polymérase est faible en l'absence du site NUT et des autres facteurs. La Figure 19 schématise l'ensemble des interactions intervenant au sein du complexe d'antiterminaison processive de λ .

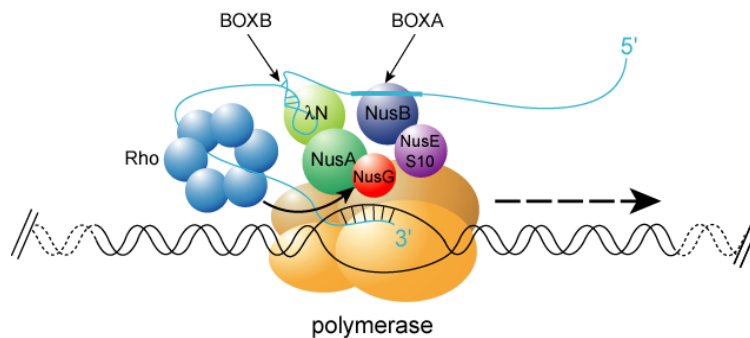


Figure 19 : Conformation hypothétique du complexe d'antiterminaison du phage λ
(modifié selon Greenblatt et al., 1993).

Comme pour les opérons *rrn*, différents modèles proposent qu'une boucle dans l'ARN permet d'amener le complexe N/NUT/Nus au contact du complexe ternaire d'élongation, de façon à permettre l'interaction entre certains facteurs Nus et l'ARN polymérase (Richardson et Greenblatt, 1996). Les modifications structurales apportées par les différentes interactions au sein du complexe sont encore mal définies. En particulier on ignore encore le rôle précis de NusA et NusG dans le complexe d'antiterminaison. Seuls, ces facteurs sont plutôt prompts à favoriser la terminaison. Certaines hypothèses suggèrent que leur rôle dans l'antiterminaison consiste plutôt à stabiliser l'ensemble des interactions au sein du complexe d'élongation. On ne sais pas non plus si le site NUT maintient son interaction avec le complexe une fois celui-

ci formé. Cependant, certains travaux suggèrent que l'ARN polymérase renforce les interactions N/BOXB au sein du complexe (Mah et al., 2000).

En présence de N, la transcription se poursuit donc au-delà des terminateurs du phage λ , et des gènes très éloignés peuvent être exprimés. Ainsi, si λ est intégré dans le chromosome bactérien, la transcription peut s'étendre plusieurs dizaines de kilobases à l'intérieur du génome de la cellule hôte. Contrairement au complexe d'antiterminaison d'*E. coli*, le complexe d'élongation modifié par N est capable de résister, à la fois aux terminateurs Rho-dépendants et aux terminateurs intrinsèques.

De récentes hypothèses ont été avancées pour expliquer l'action de N au niveau des terminateurs intrinsèques. Gusarov et Nudler ont utilisé un système minimal comprenant l'ARN polymérase, N, et NusA, pour étudier, *in vitro*, l'antiterminaison au niveau du terminateur intrinsèque tR2 du phage λ (Gusarov et Nudler, 2001). En se basant sur des travaux antérieurs (Horwitz et al., 1987) suggérant que le complexe d'antiterminaison contient deux molécules de NusA, les auteurs avancent que, dans le complexe d'antiterminaison, N interagit avec l'ARN polymérase et les deux molécules de NusA. L'une d'elles fixe le site NUT, tandis que l'autre, dont la conformation est modifiée par N, interagit avec l'ARN à sa sortie du complexe transcriptionnel. Au niveau d'un terminateur intrinsèque, N et la deuxième molécule de NusA séquestrent le bras droit de la tige-boucle et empêchent la formation de la structure terminatrice, permettant à l'ARN polymérase de résister au signal de terminaison (Figure 20).

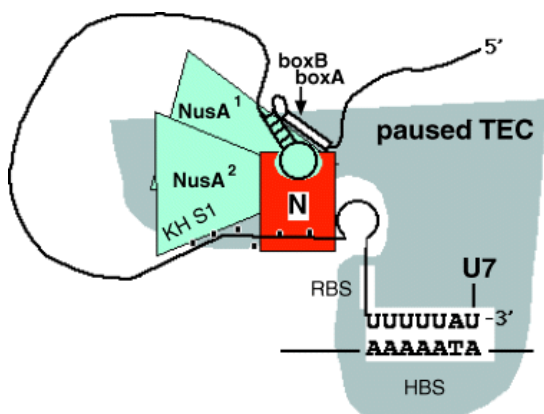


Figure 20 : Modèle de Gusarov et Nudler du contrôle de la terminaison intrinsèque par N. NusA2 complexée à N interagit avec le bras gauche de la structure en tige-boucle terminatrice et inhibe sa formation (KH et S1 sont les domaines de fixation à l'ARN de NusA, d'après : Gusarov et Nudler, 2001).

En revanche le mécanisme permettant à N de réduire la terminaison Rho-dépendante est encore mal connu. La plupart des modèles suggèrent que le complexe, induit par la fixation de N au niveau du site NUT, favorise la processivité du complexe d'élongation. Les

signaux de pause qui jouent un rôle prépondérant dans la terminaison dépendante de Rho (voir chapitre précédent) présentent de fortes homologues avec les terminateurs intrinsèques. On peut donc penser que le modèle de Gusarov et Nudler, c'est-à-dire la déstabilisation de l'hairpin par N et NusA, s'applique aussi aux signaux de pause qui spécifient le site de la terminaison Rho-dépendante. N, couplé à NusA, donnerait dans ce cas un avantage à l'ARN polymérase dans la compétition cinétique qui l'oppose à Rho. Dans ce modèle, les autres facteurs serviraient simplement de stabiliser le complexe pour lui permettre d'agir sur de longues distances.

Cependant, le complexe d'antiterminaison au niveau des opérons *rrn*, qui ne contient pas N mais tous les autres facteurs Nus, n'affecte que les terminateurs Rho-dépendants (Albrechtsen et al., 1990; Condon et al., 1995). Ceci laisse suggérer l'existence de deux mécanismes différents : l'un mettant en jeu N et NusA pour bloquer la formation des terminateurs intrinsèques, alors que l'autre serait basé sur l'interaction entre les facteurs Nus et l'ARN polymérase, avec une spécificité d'action contrôlée : par N pour les gènes du phage λ , et par des facteurs encore indéterminés pour les opérons *rrn*. Nous avons vu dans le chapitre précédent que, pour certains terminateurs, NusG est nécessaire à l'action de Rho. On peut penser, qu'au cœur du complexe d'antiterminaison, l'action de NusG en faveur de Rho est bloquée par la présence des autres facteurs Nus. Ce mécanisme serait identique chez λ et chez *E. coli*. Au vu de la faible affinité du couple NusB-S10 pour la BOXA de λ , N agirait simplement pour assurer la spécificité d'action du complexe d'antiterminaison au niveau des opérons précoces de λ .

En plus de cette antiterminaison processive, N peut inhiber la terminaison de façon non processive (c'est-à-dire au niveau de sites de terminaison proches du site *nut*). NusA est le seul facteur cellulaire nécessaire à cet événement (Mason et al., 1992); c'est vraisemblablement un autre mécanisme qui est utilisé par N dans ce phénomène. Au cours de son évolution le bactériophage λ a su adapter le dispositif d'antiterminaison d'*E. coli* à son profit (Richardson et Greenblatt, 1996). On peut penser qu'un mécanisme beaucoup plus simple ait été mis en place plus tôt au cours de l'évolution. Le complexe non processif N-NUT-NusA serait alors un vestige de ce mécanisme primitif. Puis, au fur et à mesure de son évolution, λ aurait augmenté les contraintes de son système de régulation génique (apparition de nouveaux terminateurs) et aurait, en parallèle, développé un système de régulation plus performant en détournant le mécanisme d'antiterminaison d'*E. coli* à son profit. Ces

informations laissent penser que, durant l'évolution, le phage s'est adapté aux mécanismes d'*E. coli* et les a modifiés pour réguler l'expression précoce et tardive de ses propres gènes. Cela signifie donc que, plus tôt dans l'évolution, les bactériophages avaient mis en place des systèmes de régulation de la terminaison beaucoup plus simples et ne faisant intervenir, ni la BOXA, ni les facteurs Nus. Or, récemment, Neely et Friedman ont décrit le système d'antiterminaison chez H-19B, un phage de type λ pour lequel l'antiterminaison au niveau des sites NUT ne requiert ni la BOXA, ni les facteurs cellulaires NusB et S10 (Neely et Friedman, 2000). Il s'agit bien ici d'un exemple d'un mécanisme primitif faisant agir essentiellement N et NusA. Les résultats que j'ai obtenus au cours de la première partie de ma thèse permettent d'éclaircir en partie ce mécanisme.

2. 2. 2. Antiterminaison dirigée par le facteur Q.

La protéine Q antitermine la transcription initiée au niveau du promoteur pR' (Roberts et al., 1998). Contrairement à N, Q agit en se fixant à l'ADN au niveau d'un site *qut* (pour Q utilisation) situé légèrement en amont du promoteur pR' (Somasekhar et Szybalski, 1983; Yang et al, 1987). Puis Q entre en contact avec l'ARN polymérase arrêtée sur un site de pause en position +16 après le promoteur (Grayhack et al., 1985; Yarnell et Roberts, 1992; Roberts et al., 1998). Cette pause est dépendante du facteur $\sigma 70$ (Ring et al., 1996; Ko et al., 1998). Cette dernière information constitue un exemple remarquable de l'apport constant de λ à nos connaissances en biologie moléculaire. En effet il s'agit du premier exemple connu d'une interaction entre σ et l'ARN polymérase maintenue après l'initiation, et d'une fonction régulatrice de σ dans l'élongation de la transcription, fonction confirmée par la suite (Bar-Nahum et Nudler, 2001). L'ARN polymérase modifiée par Q devient résistante aux terminateurs intrinsèques et Rho-dépendants situés au-delà du site *qut*. On sait cependant peu de chose sur le mécanisme de cette modification, si ce n'est qu'apparemment Q est capable d'inhiber la formation d'hybride ARN/ARN à la sortie de l'ARN polymérase. Q agit donc probablement en interdisant la formation des structures en tige boucle caractéristiques des terminateurs intrinsèques et de certains sites de pauses (Yarnell et Roberts, 1999). *In vitro* NusA favorise l'action de Q, mais son importance n'a pas été démontrée *in vivo* (Barik et Das, 1990; Yarnell et Roberts, 1999).

2. 3. Terminaison et antiterminaison chez le phage HK022.

Curieusement les sites NUT de λ peuvent être impliqués dans des phénomènes d'antiterminaison mais également dans des phénomènes de terminaison. HK022, un des homologues et compétiteurs du phage λ , bloque l'expression et le développement de λ en inhibant l'élongation de la transcription des opérons pL et pR (Robert et al., 1987).

Cette particularité est liée à une protéine de 109 acides aminés (13 kDa) connue sous le nom de Nun. Bien que très différente de N, Nun contient un motif riche en Arginine (ARM) dans sa partie N-terminale (Lazinski et al., 1989). Comme pour N, ce motif lui permet de se fixer sur les sites NUTL et NUTR des transcrits de λ (Chattopadhyay et al., 1995). La partie C-terminale de Nun, quand à elle, fixe le facteur NusA, l'ARN polymérase et l'ADN (Watnick et Gottesman, 1998 et 1999, Watnick et al., 2000). *In vitro*, Nun, et le complexe d'élongation comprenant un site NUT au niveau du transcrit, sont les seuls éléments nécessaires à la réaction (Hung et Gottesman, 1995). Cependant ces travaux montrent que *in vitro*, Nun ne provoque que l'arrêt de l'ARN polymérase sans relargage du transcrit. Or de nombreuses études montrent que, *in vivo*, Nun induit la terminaison de la transcription et le relargage de l'ARN et de l'ARN polymérase (Sloan et Weisberg, 1993; Robert et al., 1987). Des travaux récents dans le groupe de Gottesman semblent indiquer que la protéine TRCF (ou MFD), dont nous avons parlé dans le chapitre I, participe, *in vitro* et *in vivo*, à la dissociation des complexes d'élongation arrêtés par Nun (Washborn et al., 2003). Ainsi TRCF agirait comme cofacteur pour permettre à Nun de terminer la transcription précoce de λ . Comme pour l'antiterminaison N-dépendante, la terminaison par Nun est favorisée par l'ensemble des facteurs Nus (Robledo et al., 1991).

Non seulement Nun induit la terminaison de la transcription du génome du phage λ , mais elle inhibe également l'action de N en interférant avec sa fixation au niveau des sites NUT. La décision de terminer ou d'antiterminer la transcription de λ va donc dépendre du rapport Nun/N présent dans la cellule (Hung et Gottesman, 1995).

En plus de ces mécanismes, Nun interfère avec la synthèse de la protéine N. La traduction de N est régulée par une structure en tige boucle localisée entre NUTL et le début du gène N. Cette structure, qui inhibe la traduction de N, est le substrat de la RNaseIII, une endonucléase qui dégrade les doubles brins présents dans les tiges boucles ARN. Le clivage de cette structure spécifique contrôle l'autorégulation par N de sa propre traduction (Wilson et

al., 2002). Nun en se fixant au site NUTL réprime la traduction de N, probablement en interférant avec le chargement des ribosomes sur l'ARNm (Kim et al., 2003). L'ensemble de ces mécanismes permet à HK022, par l'intermédiaire de Nun, de bloquer l'expression d'un compétiteur au sein de la cellule.

Gottesman et ses collaborateurs ont récemment isolé des mutations ponctuelles dans la région C-terminale de Nun (Kim et al., 2003). L'un de ces mutants, K106/107D, n'est plus capable de terminer la transcription de λ bien qu'il soit toujours capable de fixer les sites NUT. Curieusement, ce mutant provoque l'antiterminaison au niveau du terminateur tR1, mais pas au niveau de terminateurs plus éloignés. Les auteurs de ce travail suggèrent que le mutant Nun bloque l'action de Rho en interférant avec sa fixation sur l'ARN.

Parallèlement, HK022, comme λ , régule l'expression de ses gènes en supprimant la terminaison de la transcription. L'expression tardive des gènes chez HK022 dépend d'une protéine Q similaire à celle du phage λ (Atkinson et Gottesman, 1992). Par contre, contrairement aux autres phages de type λ , HK022 ne possède pas d'homologue de la protéine N ou du système N/NUT. L'expression des gènes précoces d'HK022 repose sur un mécanisme d'antiterminaison ARN-dépendant (King et al., 1996), deux séquences régulatrices dans le transcrit de HK022, PUTL et PUTR (pour "polymerase utilization"), sont responsables de ce phénomène. Ces séquences, hautement structurées, modifient l'ARN polymérase, augmentent sa vitesse d'élongation, et suppriment la terminaison intrinsèque et la terminaison facteur-dépendante (King et Weisberg, 2003). Cette action ne requiert ni protéine phagique, ni facteur cellulaire (Sen et al., 2001).

2. 4. Antiterminaison chez les eucaryotes.

Pour finir, il est important de noter que l'antiterminaison n'est pas un phénomène propre aux procaryotes. Bien que cette forme de régulation génique ait été assez peu étudiée chez les cellules eucaryotes, l'élongation de la transcription représente une cible pour la régulation des gènes, en particulier dans les systèmes viraux.

Chez le VIH par exemple, la transcription initiée au niveau des séquences LTR ("Long Terminal Repeat") est atténuée au niveau d'un motif ARN, nommé TAR, situé environ 60 nucléotides en aval de l'initiation de la transcription. La poursuite de la transcription nécessite

l'action de la protéine virale Tat ([Kao et al., 1987](#); [Cullen, 1990](#)). Tat se fixe à l'ARN rétroviral au niveau de la séquence TAR et augmente la processivité de l'ARN polymérase II, résultant en des transcrit beaucoup plus longs ([Laschia et al., 1993](#)). Tat, comme N, fixe son motif ARN grâce à un domaine riche en Arginine, et recrute des facteurs cellulaires pour modifier l'ARN polymérase ([Marciniak et Sharp, 1991](#); [Kato et al., 1992](#)), le mécanisme précis de cette action reste cependant inconnue.

3. Importance des interactions ARN-protéine dans la régulation des gènes.

La structure de l'ARN, contrairement à l'ADN, est extrêmement variable et flexible, permettant ainsi de former une grande variété de motifs d'interactions et de mettre en place un système de régulation complexe par des protéines se fixant à ces motifs. Les interactions ARN-protéine jouent ainsi un rôle clé dans la régulation de la transcription comme illustré dans cette introduction. Une multitude de complexe ribonucléoprotéique sont également impliqués dans des processus post-transcriptionnels incluant : l'épissage des ARNm précurseurs, la polyadénylation, l'editing, le routage intracellulaire, la traduction ou encore la dégradation des ARNm (pour revue : [Stülke, 2002](#)). Ces complexes sont de plus impliqués dans un grand nombre de dérèglement et de pathologies. Comprendre les mécanismes moléculaires impliqués dans les interactions ARN-protéine et rechercher des molécules capable d'inhiber la formation de ces complexes constituent depuis maintenant plus de dix ans un enjeu particulièrement important (pour revue : [Frankel, 2000](#)).

Chapitre IV. LES INTERACTIONS ARN-PROTEINE.

Depuis les années cinquante, le dogme de la biologie voulait que l'information génétique soit transmise de l'ADN vers l'ARN, puis de l'ARN vers les protéines. On a ensuite découvert que l'ARN ne constituait pas qu'un simple intermédiaire entre l'ADN et les protéines, et que l'organisation de la machinerie de biosynthèse des protéines reposait, en fait, sur trois familles d'ARN :

- Les ARNs ribosomiaux (ARNr), qui constituent les ribosomes en tant que machine de synthèse des protéines.
- Les ARNs messagers (ARNm), qui fournissent l'information génétique, portée par l'ADN, vers la machinerie de synthèse protéique.
- Les ARNs de transfert (ARNt), qui fournissent la machinerie en acides aminés activés et prêts à être incorporés dans les protéines.

Depuis, de nombreuses études ont montré que l'ARN possède bien d'autres fonctions que celles des trois espèces majoritaires (ARNr, ARNm, ARNt). Ces découvertes ont permis, ces dernières années, de redéfinir complètement le rôle des ARNs dans la cellule (voir : [Nature](#), (2002). Vol 418, pp122-235; [Finnegan et Matzke](#), 2003).

Des études récentes suggèrent que, chez les mammifères, les ARNs non codants représentent plus de 90% des transcrits du génome ([Mattick et Gagen](#), 2001). Difficile de croire que tant de molécules ARN soient dépourvues de fonction au sein d'une cellule. En fait ces ARNs "minoritaires" sont impliqués dans : la duplication de l'ADN, la division cellulaire, la synthèse d'ARNm, la régulation de la traduction, les modifications post-traductionnelles, le transport des protéines, la différenciation cellulaire, etc... (pour revue : [Storz](#), 2002). Et chaque année, de nouvelles espèces d'ARNs non codants, et possédant de nouvelles fonctions dans la cellule, sont découverts ([Mattick](#), 2004).

1. Des ARNs comme agents thérapeutiques?

Les molécules d'ARN peuvent adopter une grande variété de formes spatiales. Leur repliement est gouverné par des contacts spécifiques mettant en jeu des interactions entre les

différentes bases de l'ARN. Ces interactions ont récemment été répertoriées, au sein de douze familles, selon : (i) le type de liaison entre les nucléosides (WatsonCrick ou Hoogsteen) ou entre les sucres, et (ii) selon l'orientation des sucres (*cis* ou *trans*) (Leontis et Westhof, 2001). Une fois combinées ces interactions peuvent former des structures hélicoïdales, des triplets ou des quadruplets de bases, des boucles à l'extrémité de tiges (boucle GNRA par exemple), des boucles internes (symétriques ou asymétriques), ou encore des boucles de jonctions (pour revue : Leontis et Westhof, 2003). Ces motifs peuvent également interagir entre eux dans l'espace. Certains ARNs peuvent alors former des machineries complexes capables d'activités catalytiques, d'autres peuvent former des interactions, extrêmement spécifiques avec l'ADN, les protéines, ou même d'autres ARNs.

Ces caractéristiques ont incité les chercheurs à exploiter les ARNs à des fins médicales. Aujourd'hui ces ARNs thérapeutiques peuvent être répertoriés en trois catégories principales : (i) les inhibiteurs de gènes, (ii) les correcteurs d'anomalie génétiques, (iii) les inhibiteurs de protéines.

1. 1. Inhibition de l'expression des gènes par un ARN.

L'inactivation de l'expression d'un gène par un ARN complémentaire à un ARNm cible a été pour la première fois mise en évidence par Pestka et ses collaborateurs (Petska et al., 1984) et Inouye et ses collaborateurs (Coleman et al., 1984) chez la bactérie. Ces ARNs complémentaires dit "antisens" reconnaissent et fixent de façon séquence-spécifique un ARNm cible. La formation d'une double hélice entre les deux partenaires conduit alors à la dégradation de l'ARNm ou à l'inhibition de sa traduction. La stratégie antisens est rapidement apparue comme un moyen efficace de réprimer spécifiquement un gène d'intérêt.

La découverte, au début des années 80, de structures ARN possédant une activité catalytique (Kruger et al., 1982), a permis la mise au point d'outils thérapeutiques nommés "trans-cleaving ribozyme" (Uhlenbeck, 1988; Cech, 1988). Ces molécules (ribozymes) se fixent sur un ARNm cible par des appariements complémentaires, puis coupent cet ARN cible qui devient alors inutilisable. Le ribozyme, lui, est relargué et peut être recyclé et réutilisé pour de nouveaux cycles de dégradation. Plusieurs essais cliniques de phase I et II ont été menés, en utilisant cette classe d'ARNs catalytiques, sur des patients atteints de maladies infectieuses ou de cancers (pour revue : Doudna et Cech, 2002).

L'efficacité des deux systèmes décrits précédemment est toutefois limitée. En effet les premières études ont montré qu'une répression efficace nécessite une quantité considérable d'ARNs inhibiteurs dans la cellule ([van der Krol et al., 1988](#)). De plus ces techniques sont limitées par la faible pénétration du matériel dans la cellule et par la stabilité limitée des ARNs une fois dans le milieu cellulaire. Enfin il semblerait que les effets de ces molécules observés *in vivo* soient la conséquence de mécanismes alternatifs incontrôlés (réaction immunitaire) consécutifs de la présence d'une grande quantité d'ARNs "étrangers" dans la cellule, plutôt que d'une action ciblée de ces ARNs ([Robinson, 2004](#)).

Récemment, deux nouvelles stratégies permettant d'inhiber l'expression des gènes, dans les cellules de mammifères, ont été décrites.

Tout d'abord dans de nombreuses cellules eucaryotes, l'expression d'un gène peut être inhibée par un ARN double brin dont la séquence est complémentaire de l'ARNm qui est ciblé. Ce processus, désigné "découverte de l'année 2002" par la revue *Science* ([Couzin, 2002](#)), est connu sous le nom d'interférence par l'ARN ("RNA interference" ou "post-transcriptionnal gene silencing"; [Bernstein et al., 2001a](#); [Hannon, 2002](#)). Ce système a tout d'abord été découvert chez *Caenorhabditis elegans* ([Fire et al., 1998](#)), puis mis en évidence chez la drosophile ([Pal-Bhadra et al., 1997](#); [Hammond et al., 2000](#)) et chez les mammifères ([Elbashir et al., 2001](#); [Paddison et al., 2002](#)). Son fonctionnement repose sur :

- L'expression par l'ARN polymérase III d'ARNs qui s'organisent ensuite sous forme de doubles brins.
- Une ribonucléase de la famille des RNaseIII, nommée DICER ([Bernstein et al., 2001b](#)), qui coupe ces ARNs et les transforme en petits ARNs interférents (siRNA).
- Enfin un complexe multimoléculaire, nommé RISC, qui incorpore les siRNA, et les utilise pour cibler un ARNm et le dégrader ([Hammond et al., 2000](#)).

Au-delà des avantages fondamentaux que ce processus peut apporter dans le domaine de la recherche (invalidation de gène par exemple), on imagine facilement les avantages thérapeutiques que l'on pourrait tirer d'un tel mécanisme.

Enfin on peut noter l'utilisation d'introns mobiles du groupe II, cette fois pour s'insérer dans des ADN pathogènes et les réprimer. L'avantage consiste à cibler l'ADN plutôt que l'ARN, ainsi l'effet est maintenu dans la descendance de la cellule; il n'est donc pas nécessaire de maintenir le matériel ARN dans la cellule ([Guo et al., 2000](#)).

Il reste désormais à déterminer si ces deux nouvelles stratégies d'inhibition de l'expression des gènes peuvent être efficacement utilisés à des fins thérapeutiques ([Robinson, 2004](#)).

1. 2. Réparation d'une anomalie génétique grâce à un ARN.

L'information génétique peut être modifiée lors du passage de l'ADN à la protéine. Chez les eucaryotes cette modification a lieu principalement au niveau de l'ARN, lorsque le phénomène d'épissage supprime les introns et lie entre eux les exons pour former un ARNm mature. Caractéristique intéressante, la machinerie responsable de la maturation de ces ARNs est elle-même constituée de molécules ARN. Une nouvelle stratégie s'est ainsi basée sur cette particularité biologique pour réparer un défaut ou une mutation dans un gène, directement au niveau de son ARNm. En utilisant les introns du groupe I ([Jones et al., 1996](#)) ou le complexe d'épissage ([Puttaraju et al., 1999](#)), il devient possible de remplacer un intron muté ou défectueux par un intron portant une séquence correctrice (pour revue : [Sullenger et Cech, 1994](#)). Cette méthode est appelée "trans splicing". Cependant il semble que la spécificité de cette technique soit faible ([Jones et al., 1996](#)); l'amélioration du ciblage constitue, à l'heure actuelle, la priorité des recherches dans ce type de système.

1. 3. Des ARNs inhibiteurs des fonctions protéiques.

De nombreuses protéines sont impliquées dans les dérèglements des fonctions cellulaires, elles représentent alors des cibles pharmacologiques intéressantes, entraînant la recherche de composés capables de se fixer spécifiquement à ces cibles et d'inhiber leur action.

Par exemple, la plupart des virus, comme le VIH ou le bactériophage λ , se servent de motifs ARNs hautement structurés pour recruter des protéines, virales ou appartenant à la cellule hôte, afin de réguler l'expression de leurs gènes. Comme nous l'avons vu, le phage λ utilise des sites NUT, contenant le motif BOXB, pour recruter la protéine N ainsi que certains facteurs cellulaires, et activer l'expression de ses gènes précoces. De même, le VIH utilise des éléments ARN, nommés TAR (pour "Trans Activation Response") ou RRE (pour "Rev Response Element"), pour recruter des protéines régulatrices, Tat et Rev respectivement, et contrôler l'expression du génome viral.

L'une des premières stratégies pour inhiber la fonction d'une protéine (ici Tat), fut d'utiliser directement l'élément TAR du VIH. Ce motif, exprimé dans des lymphocytes infectés par le VIH, entre alors en compétition avec le motif TAR rétroviral pour la fixation de Tat et réprime ainsi le développement du virus (Sullenger et al., 1990). Ces premiers résultats ont alors suggérés que d'autres petits ARNs structurés seraient capable de fixer des protéines d'intérêt thérapeutique et d'inhiber leur(s) fonction(s). Des travaux ont donc été menés pour développer des méthodes permettant d'isoler, *in vitro*, des ARNs de haute affinité à partir d'un vaste répertoire de séquences aléatoires.

2. Sélection et analyse des interactions ARN-protéine *in vitro*.

2. 1. La méthode SELEX.

Dans une collection d'ARNs de séquence aléatoire, un grand nombre de formes spatiales peuvent être représentées et sélectionnées pour leur capacité à interagir avec une protéine cible. En 1990, deux groupes décrivent une méthode permettant, à partir d'une vaste bibliothèque de séquences, de sélectionner, *in vitro*, des ARNs ou des ADNs hautement spécifiques d'une protéine cible donnée (Ellington et Szostak, 1990; Tuerk et Gold, 1990). Ces ARNs/ADNs ligands furent nommés aptamers (Ellington et Szostak, 1990) et la technique fut baptisée SELEX (pour "Systematic Evolution of Ligands by Exponential Enrichment"; Tuerk et Gold, 1990; Figure 21).

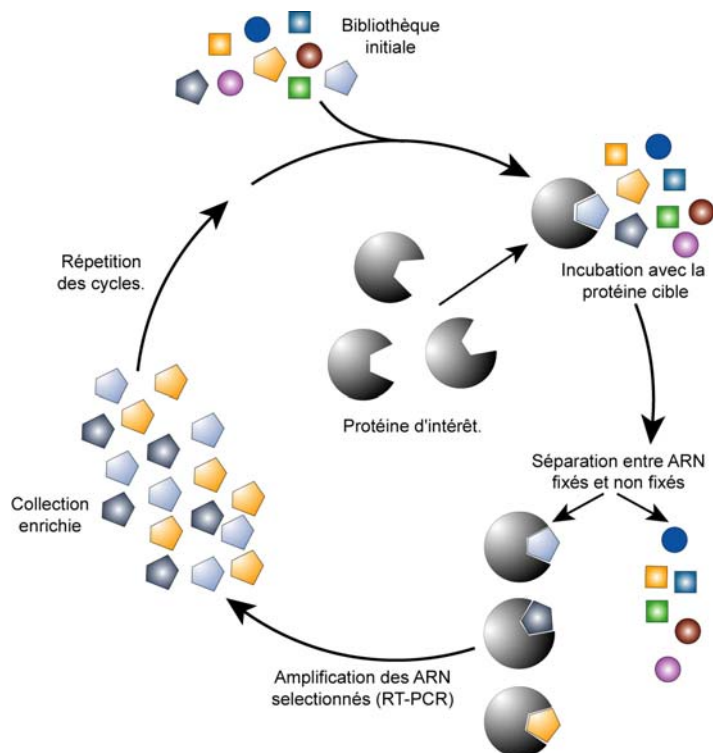


Figure 21 : Principe du système SELEX. Un répertoire d'acides nucléiques est incubé avec les protéines d'intérêt. Les molécules qui fixent la protéine cible sont séparées des autres motifs de la bibliothèque. Les séquences fixées sont alors éluées et amplifiées pour générer un nouveau répertoire enrichi en séquences qui fixent la protéine d'intérêt (d'après : Ellington et Szostak, 1990; Tuerk et Gold, 1990)

Grâce à cette démarche, plusieurs aptamers possédant de très fortes affinités pour des protéines ont été sélectionnés (pour revue : Gold et al., 1997). Certains d'entre eux présentent des affinités équivalentes à celles des anticorps monoclonaux ($10^{-12}\text{M} < K_d < 10^{-9}\text{M}$). Par ailleurs, ils peuvent être synthétisés chimiquement en grande quantité et même modifiés pour devenir plus résistants aux dégradations en milieu cellulaire. Les oligonucléotides ainsi sélectionnés, peuvent être utilisés directement pour un travail de diagnostic *in vitro*. De plus la connaissance de la structure tridimensionnelle des aptamers fournit les éléments indispensables (forme spatiale de la molécule, positionnement des groupements fonctionnels) pour la recherche et la synthèse de composés biologiquement actifs. Plusieurs molécules issues de ces sélections sont actuellement en phase d'essais cliniques (pour revue : Sullenger et Gilboa, 2002).

2. 2. La méthode T7 phage display.

Contrairement au SELEX, cette méthode permet de sélectionner, à partir d'une banque d'ADNs complémentaires (ADNc), les séquences codants des protéines, ou des motifs peptidiques, capables de fixer spécifiquement un motif ARN cible (Figure 22).

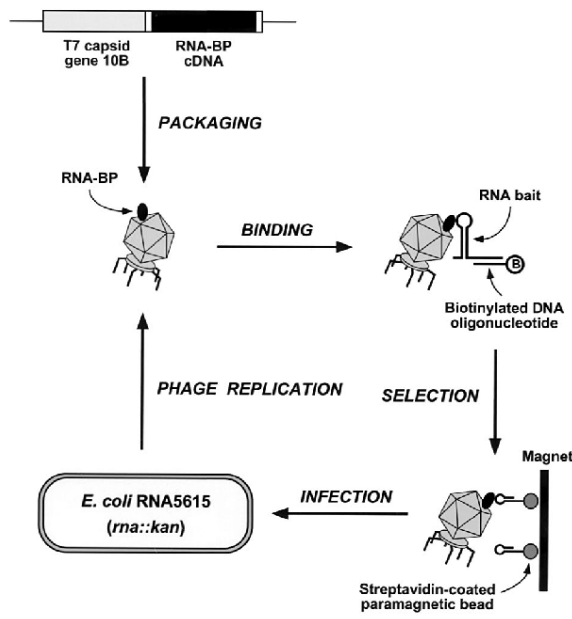


Figure 22 : Principe de sélection, *in vitro*, de protéines par T7 phage display (d'après : [Danner et Belasco, 2001](#)). Les phages recombinants portent la protéine d'intérêt couplée aux protéines de la capside virale. Ils sont ensuite soumis à l'interaction avec l'ARN cible qui est hybridé à un ADN biotinylé. Les phages qui ont fixé l'ARN sont piégés par des billes magnétiques couplées à de la streptavidine, ces phages sont isolés du reste de la réaction grâce à un aimant. Ils sont ensuite relargués et utilisés pour infecter des cellules d'*E. coli* avant de subir d'autres cycles de purification.

Le système repose sur la création de phages recombinants présentant à leur surface une protéine hétérologue fusionnée à la protéine 10B de la capside virale du phage T7 ([Figure 22](#)). Les phages ainsi produits peuvent alors reconnaître un ARN cible fixé sur une colonne grâce à un système de billes magnétiques. Les phages fixés sont utilisés pour infecter des cellules d'*E. coli*. Après réplique, les phages obtenus sont soumis à de nouveaux cycles de sélection ([Danner et Belasco, 2001](#)). On obtient ainsi un enrichissement de la population en phages, portant à leur surface une protéine spécifique d'un motif ARN d'intérêt.

Les auteurs de ce système l'ont utilisé afin d'étudier l'interaction entre la protéine spliceosomale U1A et son motif de reconnaissance sur l'ARN, U1hpII. En mutant le domaine RRM (pour "RNA Recognition Motif") responsable de la fixation d'U1A à l'ARN, ils ont mis en évidence plusieurs résidus importants pour cette fixation, ainsi qu'un variant de la protéine U1A présentant une affinité plus forte pour son motif de reconnaissance ([Laird-Offringa et Belasco, 1995](#)).

Dès les premières utilisations de ces méthodes, il est clairement apparu que de grandes améliorations pourraient leur être apportées si les étapes de sélection pouvaient être réalisées directement dans la cellule. Un tel crible permettrait de s'affranchir de la purification de la protéine cible et de réduire le très grand nombre de cycles de sélection-amplification, caractéristique de la méthode SELEX, qui peuvent occasionner des pertes de séquences

intéressantes parmi le répertoire. Ainsi, des études menées *in vivo*, ont mis en évidence des aptamers laissés de côté en cours de cycles de SELEX, et qui pourtant présentent *in vivo* une affinité plus forte pour leur protéine cible que les molécules sélectionnées *in vitro* en fin de cycles SELEX (Cassiday et Maher, 2003). De plus, la capacité d'un aptamer ou d'un motif peptidique à fixer *in vivo* sa cible spécifique va dépendre de sa stabilité, de son bon repliement, et de sa localisation cellulaire correcte, conditions qui ne sont pas garantis dans le cas d'une sélection faite *in vitro*. Une fois encore une sélection directement dans la cellule permettrait de s'affranchir de ces contraintes.

3. Sélection et analyse des interactions ARN-protéine *in vivo*.

Différents systèmes ont donc été développés, au cours de la décennie passée, pour permettre de sélectionner et d'analyser les interactions ARN-protéine *in vivo*. Ces systèmes sont généralement basés sur des méthodes génétiques. Ils reposent, dans la majorité des cas, sur la transformation (ou la transfection) de cellules bactériennes, de levure, ou de mammifères par deux constructions plasmidiques :

- Un premier plasmide exprimant la protéine cible (ou une banque d'ADNc), sous le contrôle d'un promoteur inductible.
- Un second plasmide portant un système rapporteur sous le contrôle d'un promoteur constitutif; cette construction peut, dans certains cas (triple hybride), être insérée directement dans le génome de la cellule.

Le principe de ces systèmes est de mettre l'expression d'un gène rapporteur sous le contrôle de l'interaction entre un motif ARN et sa protéine cible. Ce contrôle se fait par l'intermédiaire de mécanismes inspirés de processus de régulation déjà existants chez les procaryotes et les eucaryotes et qui sont: (i) la trans-activation de la transcription, (ii) la répression de la traduction, (iii) l'antiterminaison de la transcription.

3. 1. Activation de la transcription : le système triple hybride.

Dans le système triple hybride, une interaction ARN-protéine, dans le noyau de la levure *Saccharomyces cerevisiae*, entraîne l'expression d'un gène rapporteur. L'interaction

peut être détectée via la croissance cellulaire, la couleur des colonies, ou encore quantifié via l'activité spécifique d'une enzyme. Cette approche est dérivée du système double hybride de la levure, qui permet de détecter les interactions protéine-protéine (Fields et Song, 1989).

La stratégie générale du système triple hybride est schématisée par la Figure 23A. Des sites de fixation ADN (ici l'opérateur LexA en multicopie) sont placés dans le génome de la levure, en amont d'un gène rapporteur (*HIS3* ou *lacZ*). Une première protéine hybride est constituée d'un domaine de fixation à l'ADN (la protéine LexA d'*E. coli*) lié à un domaine de fixation à l'ARN (la coat protein du phage MS2). Ce domaine de fixation à l'ARN interagit avec son site spécifique (motif MS2) au sein d'un ARN bifonctionnel (hybride; Figure 23B). La deuxième partie de cet ARN interagit avec une seconde protéine hybride constituée, elle aussi, d'un site de fixation à l'ARN (protéine ou peptide d'intérêt) couplé à un domaine d'activation de la transcription (le facteur de transcription Gal4 de la levure). L'activation du gène rapporteur, détectée par des essais phénotypiques (survie en milieu His⁻, en présence de 3-aminotriazol) ou quantifiée par des essais biochimiques (dosage β -galactosidase), témoigne alors d'une interaction entre le deuxième site de l'hybride ARN et le domaine de fixation à l'ARN du second hybride protéique. Ces dernières années, une vingtaine de protéines fixant spécifiquement des motifs ARN ont été découvertes grâce à cette méthode (pour revue : Jaeger et al., 2004).

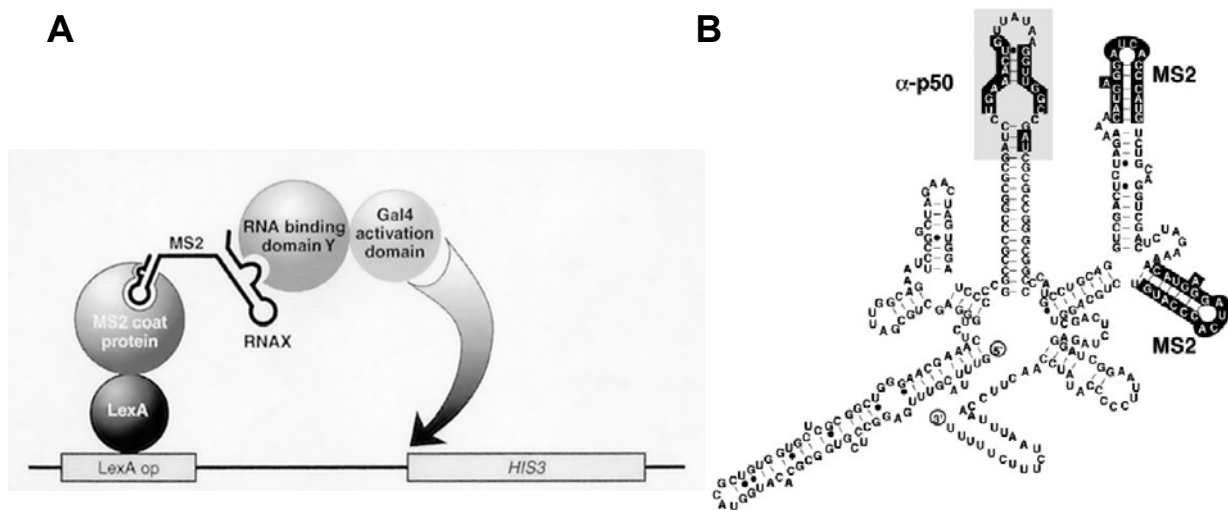


Figure 23 : (A) Principe du système triple hybride chez la levure (d'après : Bernstein et al., 2002). Les éléments notés ici sont les plus couramment utilisés pour ce système, ils peuvent cependant être remplacés par d'autres hybrides. (B) Exemple d'hybride ARN utilisé (d'après : Cassidy et Maher, 2003).

Le triple hybride permet également de disséquer les interactions ARN-protéine au sein d'un complexe déjà caractérisé. Enfin, ce procédé est plus généralement utilisé pour vérifier, *in vivo*, l'affinité de composés isolés *in vitro*. Cassidy et Maher ont, par exemple, utilisé le triple hybride pour étudier l'affinité d'un aptamers sélectionné par SELEX pour sa spécificité vis-à-vis de l'homodimère p50 de NFκB (Cassiday et Maher, 2001).

Malheureusement, comme pour le double hybride, les principales limitations du système viennent de l'apparition de faux positifs (résultant d'interactions parasites entre les différents hybrides) qui doivent être éliminés par des étapes de sélection supplémentaires. En effet, le repliement et l'interaction des différents partenaires, au sein des molécules chimériques utilisées par le triple hybride, peuvent être perturbés par les éléments présents dans la cellule; l'hybride ARN, par exemple, peut être la cible de l'attaque des nucléases cellulaires. Par ailleurs son repliement extrêmement complexe (voir Figure 23B) est difficile à assurer *in vivo*. Ces contraintes techniques peuvent réduire l'utilisation du triple hybride à l'étude d'un nombre limité de couples ARN-protéines.

3. 2. Répression de la traduction par un complexe ribonucléoprotéique.

Une deuxième stratégie pour mettre en évidence et analyser une interaction ARN-protéine repose sur la répression de la traduction. Cette méthode consiste à placer un motif de reconnaissance ARN à proximité du site de fixation des ribosomes sur le transcrit d'un gène rapporteur. En absence d'interaction, le gène est traduit, mais si une protéine se fixe à son site de reconnaissance sur l'ARNm, le complexe formé bloque l'entrée des ribosomes et la traduction du gène rapporteur est inhibée.

Deux modèles différents ont ainsi été développés, l'un pour étudier les interactions ARN-protéines chez les procaryotes (Jain et Belasco, 1996), et l'autre pour analyser ces interactions chez les eucaryotes (Hentze et Kühn, 1996).

3. 2. 1. Système procaryote.

La stratégie développée par Jain et Belasco s'inspire largement d'un mécanisme de régulation existant naturellement chez le bacteriophage MS2. A la fin de l'infection d'*E. coli*

par le bactériophage MS2, la protéine "coat protein" virale (MS2CP) fixe une structure en tige boucle dans l'ARN phagique. Cette fixation représente le signal d'encapsidation du génome ARN du bactériophage, mais elle possède également une fonction génétique : en effet ce motif (MS2) contient le site d'initiation de la traduction du gène de la réplicase du phage, en s'y fixant, la "coat protein" réprime l'expression de la réplicase du phage MS2 (Bernadi et Spahr, 1972).

En 1990, D. Peabody met à profit cette particularité pour étudier en détail l'interaction entre cet opérateur traductionnel MS2 et la MS2CP (Peabody, 1990). Dans un article de 1992 il présente ce mécanisme comme un nouveau système génétique d'analyse des interactions ARN-protéine (Peabody et Ely, 1992). Mais ce sont Jain et Belasco, en 1996, qui vont adapter ce système à l'étude systématique, *in vivo*, des interactions entre ARN et protéines chez les procaryotes (Jain et Belasco, 1996).

Chez *E. coli* deux régions sont critiques pour la traduction d'un ARNm en protéine. Il s'agit : du codon d'initiation de la traduction et d'une séquence légèrement en amont, nommée Shine-Dalgarno (S/D), et reconnue par le ribosome grâce à une complémentarité de séquence avec l'extrémité 3' de l'ARN ribosomal 16S. Dans ce système bactérien, un motif ARN, cible d'une protéine d'intérêt, est placé sur un premier plasmide, juste en amont du S/D d'un gène rapporteur *lacZ*. La protéine d'intérêt, exprimée à partir d'un second plasmide, se fixe sur son motif de reconnaissance ARN et, ainsi, bloque stériquement la fixation des ribosomes et inhibe la traduction de *lacZ* (Figure 24). L'interaction ARN-protéine provoque ici la réduction de l'activité β -galactosidase dans la cellule.

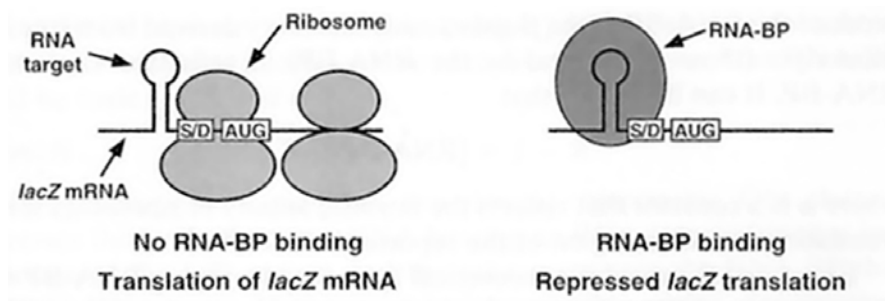


Figure 24 : Principe de répression de la traduction par une protéine hétérologue chez *E. coli* (d'après : Jain et Belasco, 2000).

Jain et Belasco ont tout d'abord utilisé ce système pour étudier la reconnaissance de l'ARN RRE par la protéine Rev du HIV. L'identification de mutants de Rev, qui présentent une altération de la reconnaissance du motif ARN, mais aussi de mutants capables de palier des mutations dans le motif RRE, a permis d'établir un modèle structural de Rev fixé à son ARN cible (Jain et Belasco, 1996). La confirmation de ce modèle par les structures RMN du complexe Rev-RRE (Battiste et al., 1996; Ye et al., 1996; Ellington et al., 1996) a prouvé l'intérêt et la puissance de ce genre d'approches.

3. 2. 2. *Système eucaryote.*

Chez les eucaryotes, la fixation d'une protéine près de la structure CAP d'un ARNm peut inhiber l'association des ribosomes et réprimer la traduction de l'ARNm. Ce type de mécanisme a été décrit pour le gène de la Ferritine dont l'ARNm contient un motif nommé IRE (pour "Iron Response Element") près de la CAP (Hentze et Kühn, 1996). Une protéine nommée IRP-1 (pour "Iron Regulatory Protein 1") se fixe sur IRE et réprime la traduction de la Ferritine.

En se basant sur ce principe Hentze et collaborateurs (Paraskeva et al., 1998) ont développé une méthode nommée TRAP (pour "Translational Repression Assays Procedure") afin d'étudier les interactions ARN-protéine chez la levure *Saccharomyces cerevisiae*. Dans ce système, le motif de reconnaissance est inséré dans la région 5'UTR de l'ARNm du gène rapporteur (ici la GFP). L'expression de la protéine d'intérêt et sa fixation sur son motif ARN cible réprime l'expression du gène rapporteur. Le niveau de répression de la GFP peut alors être mesuré par cytométrie de flux.

Ce type de stratégie peut être appliquée, *in vivo*, pour réprimer spécifiquement l'expression d'un gène d'intérêt, en clonant un aptamer dans le 5'UTR de ce gène. L'expression du produit du gène visé se retrouve alors sous le contrôle de la protéine cible de cet aptamer, exprimé soit intrinsèquement soit à partir d'un plasmide introduit dans la cellule (Werstuck et Green, 1998).

3. 3. Sélection de complexes ARN-protéine par antiterminaison de la transcription.

Parmi toutes les stratégies développées pour rechercher des motifs ARN/peptidiques spécifiques de protéines/ARN d'intérêt, l'une d'elle fait intervenir le mécanisme d'antiterminaison décrit au chapitre précédent (chapitre III). Là encore, deux systèmes sont disponibles, permettant la sélection chez les procaryotes et chez les eucaryotes. Ces deux systèmes ont été développés dans le laboratoire de A. D. Frankel.

3. 3. 1. Système procaryote.

Le système bactérien repose sur le mécanisme d'antiterminaison N-dépendant (décrit au chapitre III) permettant l'expression des gènes précoces du phage λ . N reconnaît et fixe un motif BOXB, au sein des sites NUT des transcrit précoces de λ , puis induit la formation d'un complexe permettant à l'ARN polymérase de résister aux terminateurs situés en aval des sites NUT. Ici le motif BOXB est remplacé par le motif ARN cible. La reconnaissance de ce motif ARN se fait par l'intermédiaire d'un domaine de fixation à l'ARN fusionné à l'extrémité N-terminale de N. L'interaction ARN-protéine, via ce motif peptidique hétérologue, induit la formation du complexe d'antiterminaison N-dépendant. L'ARN polymérase alors modifiée peut franchir une série de terminateurs intrinsèques pour transcrire le gène *lacZ* situé au-delà de ces terminateurs (Figure 25), permettant ainsi de visualiser l'interaction via la coloration des bactéries ou via l'activité enzymatique de la β -galactosidase (Harada et al., 1996; Peled-Zehavi et al., 2000).

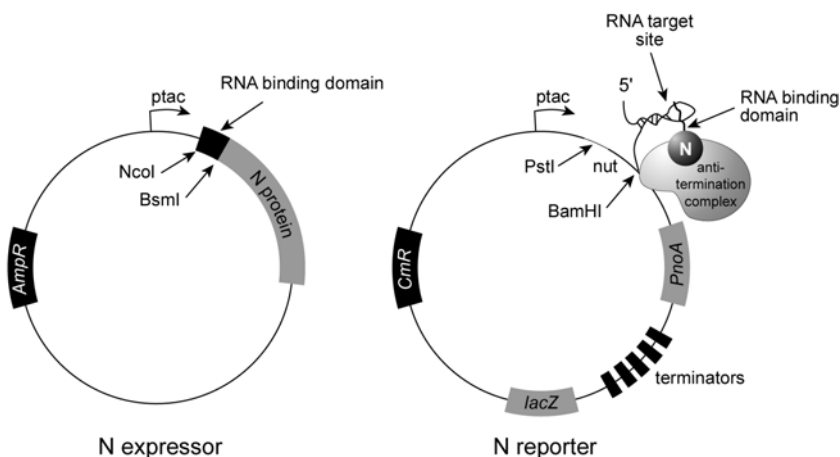


Figure 25 : Représentation des constructions plasmidiques du système d'antiterminaison par un peptide hétérologue ARN-spécifique fusionné à la protéine N du phage λ (d'après : Peled-Zehavi et al., 2000).

Ce système a été testé sur trois couples ARN-protéine : U1A-U1hpII, Rev-RRE et MS2CP-MS2. Le gène rapporteur *lacZ* peut également être remplacé par un gène de résistance à la kanamycine, permettant de cribler un nombre de clone beaucoup plus important (Peled-Zehavi et al., 2003).

Le principal inconvénient de ce système réside précisément dans la fusion entre N et le peptide d'intérêt. Le repliement et l'action de chaque partenaire de cette protéine hybride peuvent, en effet, être modifiés par la présence de l'autre partenaire. De plus, le motif ARN d'intérêt peut interférer avec les autres éléments nécessaires à la formation du complexe d'antiterminaison, comme la BOXA située juste en amont du site d'insertion du motif d'intérêt.

3.3.2. Système eucaryote.

Le système Tat fusion est basé sur les propriétés d'activateur transcriptionnel de la protéine Tat du VIH (Landt et al., 2000). Tat active la transcription, essentiellement en amplifiant la processivité des complexes transcriptionnels de l'ARN polymérase II initiés au niveau des LTR du VIH. Pour fonctionner Tat doit se fixer à TAR, un motif ARN en tige-boucle situé à l'extrémité 5' du transcrit naissant. Le domaine d'activation transcriptionnel de Tat est indépendant du domaine de fixation à l'ARN, Tat peut donc fonctionner en fixant l'ARN via une interaction ARN-protéine hétérologue (Selby et Peterlin, 1990) ou même en étant fixé à de l'ADN (Southgate et Green, 1991). Le système utilise donc une fusion entre Tat et une protéine d'intérêt (ou une bibliothèque). La fixation est mesurée par la capacité de la fusion à activer la transcription d'un rapporteur sous le contrôle d'un LTR du HIV où TAR a été remplacé par le motif ARN d'intérêt. L'interaction est mesurée via l'activation du gène de la GFP et visualisée par cytométrie de flux. Ce système a été utilisé essentiellement pour disséquer les mécanismes régissant les interactions entre l'ARN et les motifs riches en Arginine de certaines protéines interagissant avec l'ARN (Tan et Frankel, 1998).

4. Intérêt d'un nouveau système de sélection?

Aucun des systèmes décrits précédemment n'est dépourvu de faiblesses (voir chapitre VI), cette observation, ainsi que les résultats obtenus dans la première partie de mon travail de

Introduction.

thèse (voir chapitre V) nous ont conduits à élaborer un nouveau système de sélection d'aptamers *in vivo*. La conception de ce système est décrite dans la deuxième partie des "résultats".

RESULTATS

1. Introduction à l'article "Dual role of boxB RNA motif in the mechanisms of termination/antitermination at the lambda tR1 terminator revealed in vivo".

Cet article présente les résultats expérimentaux que j'ai obtenus en étudiant les mécanismes de terminaison et d'antiterminaison intervenant au niveau du terminateur tR1 de l'opéron précoce droit du phage λ (opéron *pR*).

Ce travail se base principalement sur les nombreux travaux réalisés sur ce terminateur. En effet, depuis de nombreuses années, tR1 constitue le terminateur le plus utilisé pour étudier le mécanisme de terminaison Rho-dépendante chez *E. coli*, *in vitro* comme *in vivo* (Chen et Richardson, 1987; Graham et Richardson, 1998; Richardson et Richardson, 1996). Ce terminateur présente la particularité d'être le siège des séquences responsables de la terminaison Rho-dépendante, mais aussi des séquences responsables du phénomène d'antiterminaison dépendant de la protéine N du phage λ . En effet au niveau de λ tR1, les sites *rut* ("Rho-utilisation") et les sites *nut* ("N-utilization") sont superposés (Figure 26) :

- La partie supérieure du terminateur tR1 est en effet composée de deux sites *rut* nommés *rutA* et *rutB*. Ces sites sont séparés par la séquence *boxB* de *nutR*.
- De plus *boxA*, le deuxième élément du site *nutR*, fait partie de *rutA*.

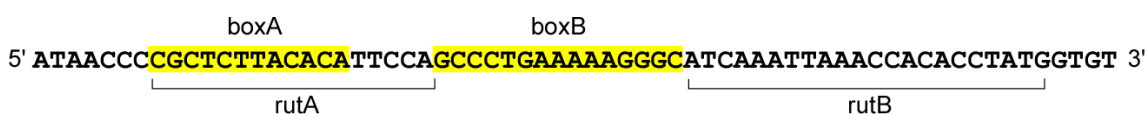


Figure 26 : Séquence ADN de la partie 5' du terminateur λ tR1. Les deux sites *rut* sont indiqués, et les éléments du site *nutR* sont surlignés en jaunes.

Des travaux antérieurs ont montrés que la séquence composée de *rutA*, *boxB*, et *rutB* est suffisante pour induire la terminaison Rho-dépendante au sein d'une séquence non terminatrice placée en aval (Richardson et Richardson, 1996). De plus, des expériences de délétions ont montré que *rutA* et *rutB* sont indispensables à la fonction de Rho au niveau de

tR1, alors que la séquence *boxB* peut être totalement déléetée sans pour autant affecter la terminaison induite par ce terminateur (Graham et Richardson, 1998; Hart et Roberts, 1994).

La séquence *boxB* est, cependant, indispensable au mécanisme d'antiterminaison N-dépendant. En effet, comme nous l'avons vu précédemment, la fixation de la protéine N au niveau du motif ARN BOXB induit la formation d'un complexe, contenant : N, le site NUT, l'ARN polymérase et un ensemble de facteurs cellulaires (facteurs Nus). Ce complexe permet à la transcription de se poursuivre au-delà des terminateurs intrinsèques et Rho-dépendants situés en aval des sites *nut*. Le mécanisme permettant cette résistance à la terminaison est encore mal connu. Il semblerait que le complexe d'antiterminaison adopte des mécanismes distincts pour résister, d'une part aux terminateurs intrinsèques, d'autre part aux terminateurs Rho-dépendants (Albrechtsen et Squires, 1990; Condon et al., 1995; Nudler et Gottesman, 2002). D'autres études, chez le phage λ (Mason et al., 1992) ou chez le phage H-19B (Neely et Friedman, 2000), semblent indiquer que dans certains cas N peut induire l'antiterminaison, sur une distance limitée, sans requérir l'action d'une majorité de facteurs Nus, le mécanisme précis de ce phénomène reste toutefois inconnu.

Dans le travail suivant nous avons utilisé un système génétique dans lequel la séquence *rutA-boxB-rutB* a été clonée entre un promoteur constitutif et un opéron artificiel composé respectivement des gènes *cat* et *ARNtArg⁵*. L'*ARNtArg⁵* sert ici de gène reporter. Le gène *cat*, quant à lui, fournit les sites de pauses transcriptionnels nécessaires à la terminaison Rho-dépendante. A l'aide de ce système, nous avons cherché à analyser le rôle du motif BOXB dans le phénomène de terminaison/antiterminaison au niveau de tR1. Pour cela nous avons construit différents dérivés de tR1 pour lesquels la séquence *boxB* est soit mutée soit remplacée par une séquence hétérologue codant pour un motif ARN en tige boucle.

- Les mutations de BOXB (au niveau des bases C et G de la tige) déstabilisant le motif, réduisent fortement le fonctionnement de Rho au niveau du terminateur tR1. En revanche, les mutations permettant de conserver la structure en tige boucle de BOXB n'affectent pas l'efficacité de la terminaison.
- L'efficacité de la terminaison n'est pas non plus affectée lorsque BOXB est remplacé par un motif hétérologue : ici l'aptamer MS2F5. Ce motif a été sélectionné, par SELEX, pour son affinité vis-à-vis de la coat protein du phage MS2 (Rowell et al., 1998).

Résultats.

- Enfin, de façon intéressante nous avons été en mesure de mimer le phénomène d'antitermination N-dépendant en remplaçant le couple BOXB/N par le couple MS2F5/MS2 coat protein, qui ne possède pas de fonction antiterminatrice intrinsèque.

Les conclusions de cette étude et ses implications pour la suite de mon travail de thèse seront développées dans la troisième partie de ce chapitre, à la suite de l'article.

2. Article.

Dual Role of boxB RNA Motif in the Mechanisms of Termination/Antitermination at the Lambda tR1 Terminator Revealed *in vivo*

Erwann Vieu and A. Rachid Rahmouni*

Centre de Biophysique
Moléculaire, CNRS, rue
Charles Sadron, 45071 Orléans
France

Rho-dependent transcription termination at the phage lambda tR1 terminator is governed primarily by the upstream rut element that encodes two RNA regions rutA and rutB. The two regions are separated by the boxB RNA motif, which is believed to be dispensable for Rho activity but serves as a binding site for λ N protein in the antitermination process. By using a minimal *in vivo* termination system, we show that the intervening boxB RNA motif has a double function in the mechanisms of termination/antitermination at λ tR1. As a folded hairpin structure, it acts as a clamp that holds rutA and rutB side by side for optimal interactions with Rho leading to efficient termination. Conversely, the binding of N protein to boxB induces antitermination at λ tR1 by preventing access of Rho to the rut sequences. This dual role was clearly shown *in vivo* by studying the effects of multiple mutations within the boxB hairpin stem on transcription termination and by substituting the N/boxB couple with the unrelated coat protein of phage MS2 and its stem-loop RNA binding site.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: antitermination; phage lambda; Rho-dependent termination; tR1 terminator; transcription regulation

*Corresponding author

Introduction

The progression of bacteriophage lambda development relies on a sequential expression of its genes. Upon infection of *Escherichia coli* cells, transcription of the phage leftward and rightward early operons initiated from pL and pR promoters by the host RNA polymerase (RNAP) is restricted to the two proximal genes producing only two proteins N and Cro, respectively. Subsequent expression of downstream genes is controlled by a succession of antitermination events that allow RNAP to read through a series of transcription terminators.^{1–4} The antitermination process requires binding of N protein to a specific *cis*-acting RNA sequence within the transcript. This RNA sequence called nut site (for N-utilization) is formed by a 15 nt stem-loop structure (boxB) and an adjacent 12 nt motif (boxA). N binding to boxB nucleates the formation of a large complex containing a set of host-encoded proteins (NusA, NusB, NusG and ribosomal protein S10) that interacts

with RNAP converting it into a termination-resistant form. For the rightward operon, the first termination signal to be overcome by RNAP is the Rho-dependent terminator tR1, which is followed by the more distal intrinsic terminator tR2.⁵ Recent *in vitro* studies with a minimal antitermination system have provided some insights into the mechanism by which N, in association with NusA, suppresses intrinsic termination.^{6,7} However, the molecular details of the N-mediated antitermination process at the Rho-dependent terminator are still unknown.

The two types of bacterial transcription termination signals differ by their mechanisms and DNA sequences. Transcript release at intrinsic terminators is a spontaneous process that arises when RNAP encounters a relatively well conserved and precisely localized sequence on the template. Alternatively, transcription termination at Rho-dependent signals occurs within a less defined sequence and requires the participation of Rho protein. A variety of Rho-dependent terminators has been identified in prokaryotic genomes. However, only a few examples have been analyzed in detail, due to their complexity and the absence of any obvious consensus sequence. In general,

Abbreviations used: RNAP, RNA polymerase.
E-mail address of the corresponding author:
rahmouni@cns-orleans.fr

Rho-dependent terminators are composed of two overlapping parts extending over a region of about 150–200 bp of DNA.^{8–10} The upstream part, called rut (for Rho-utilization), encodes a segment of the nascent RNA that is rich in cytosine residues and has a relatively low level of secondary structure. The downstream part contains several clusters of sites where termination occurs. Some of these sites have been correlated with the positions where RNAP pauses during transcript elongation. Current views suggest that Rho protein causes transcription termination in a multistep process. The reaction initiates with the binding of the protein as a hexamer to the rut site on the transcript. This binding activates the RNA-dependent ATPase activity of Rho leading to its translocation along the transcript in a 5' to 3' direction. When Rho catches up with a transcription elongation complex delayed at a downstream pause site, it brings about transcript release in a manner that remains unclear.^{8–10}

The lambda tR1 terminator is the classical paradigm for Rho-dependent termination that has been well studied both *in vitro* and *in vivo*. Deletion experiments have shown that the upstream rut element and especially two regions designated rutA and rutB are the primary determinants for Rho activity.^{11–13} Apparently, the function of the downstream release sites reflects merely their close proximity to the rut element since their replacement with foreign non-terminator DNA sequences does not prevent the termination process.¹² One intriguing feature in the λ tR1 terminator is that the N-utilization sequence is embedded within the rut element: boxA is a part of rutA and boxB is surrounded by rutA and rutB. This observation led us to consider the possibility that the control of termination/antitermination at λ tR1 may reside in a competition between Rho and N for binding to overlapping sequences on the transcript. Here we have explored this hypothesis *in vivo* by using plasmid-constructs harboring derivatives of the rut element and a downstream tRNA gene as a transcriptional reporter. Our results show that the boxB RNA motif plays a dual role in the mechanisms of termination/antitermination at λ tR1. In the absence of N protein, its folding serves as a structural hinge that brings rutA and rutB side by side for efficient Rho-dependent termination. Conversely, the binding of N to boxB induces antitermination at tR1 by preventing access of Rho to the rut sequence. This dual role of the RNA structural motif was further demonstrated by substituting boxB/N with the unrelated MS2 stem-loop/MS2 coat protein complex.

Results

A minimal *in vivo* system to analyze termination/antitermination at λ tR1

Previous studies have shown that the rut

element of λ tR1 is sufficient to induce Rho-dependent transcription termination within downstream non-terminator DNA sequences.¹² We therefore designed plasmid constructs harboring an artificial operon in which the rut element was inserted between a constitutive promoter and the *cat* gene to study the mechanisms of termination/antitermination *in vivo* (Figure 1(A)). Different derivatives of

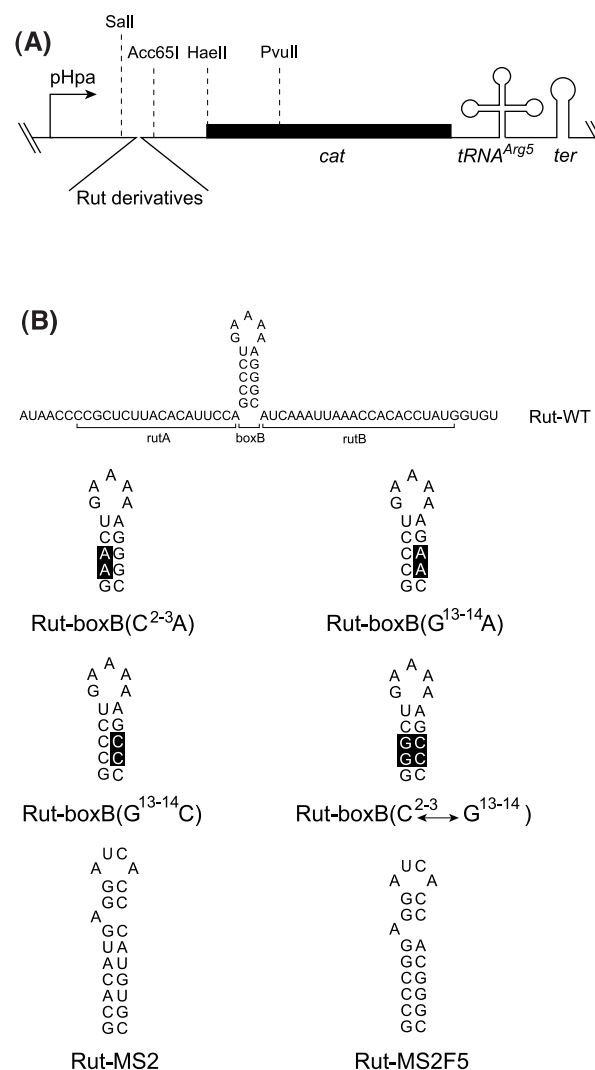


Figure 1. Schematic representation of the minimal *in vivo* termination system. (A) Relevant part of the plasmid constructs showing the artificial *cat*-tRNA^{Arg5} operon and the insertion site for the rut elements. (B) RNA sequences of the different rut elements used here. Rut-boxB(C²⁻³A) and Rut-boxB(G¹³⁻¹⁴A) correspond to the constructs where the two C bases at positions 2 and 3 or the two G bases at positions 13 and 14 of the boxB sequence, respectively, were replaced by A. In the Rut-boxB(G¹³⁻¹⁴C) variant, the two G bases at positions 13 and 14 were replaced by C and an additional replacement within this derivative of the two C bases at positions 2 and 3 by G led to the construct Rut-boxB(C²⁻³ ↔ G¹³⁻¹⁴) where base-pairing was restored. Rut- Δ boxB is a Rut-WT derivative with a complete deletion of the boxB motif. The control construct Δ Rut corresponds to the parent plasmid lacking a rut element.

the rut sequence with the boxB motif either mutated or substituted by other hairpin motifs were also investigated (Figure 1(B)). The tRNA^{Arg5} gene was inserted downstream from the *cat* gene in all the constructs to serve as a transcriptional reporter. In effect, mutant tRNAs have been used in several systems to measure gene expression *in vivo* and it was shown that their accumulation inside the cell, upon excision and maturation, is a reliable indication for the level of transcription in the DNA region where they are inserted.^{14–16} The amounts of tRNA^{Arg5} produced *in vivo* by the different constructs were determined by northern blotting analyses using a specific oligonucleotide as a probe. The blots were normalized by hybridization with a second oligonucleotide complementary to the 5 S rRNA.

We first compared the relative level of tRNA^{Arg5} produced by the plasmid harboring the wild-type rut sequence (Rut-WT) to that obtained with the parent plasmid vector lacking rut (Δ Rut). The Δ Rut plasmid was taken as a reference construct (100% level of tRNA^{Arg5}) throughout this study to calculate the termination efficiencies obtained with the different rut-derivatives. As shown in Figure 2(A), the presence of the λ tR1 rut element within the operon reduced the accumulation of tRNA^{Arg5} by more than 90%, indicating the high efficiency of transcription termination upstream from the tRNA gene (about 93% termination

efficiency, compare lanes 1 and 2). The termination efficiency of the rut-containing plasmid was reduced from 93% to 30% when the two constructs were analyzed in the Rho-15 temperature-sensitive mutant strain grown under non-permissive conditions (Figure 2(A), lanes 4 and 5). These results highlight the involvement of Rho protein in a termination event arising upstream from the tRNA^{Arg5} gene within the rut-containing plasmid.

As a further step to characterize the Rho-dependent RNA release induced by the minimal rut element, we analyzed the termination endpoints by S₁ mapping experiments. Cellular RNAs isolated from cells harboring the plasmids were hybridized to a 3'-end-labeled single-stranded DNA probe (the template strand of the Acc65I-PvuII DNA fragment, see Figure 1) and digested with S₁ nuclease as described in Materials and Methods. The S₁ mapping data in Figure 2(B) show that the RNAs extracted from cells transformed with the Rut-WT plasmid contain probe-specific truncated transcripts having 3' ends clustered over the sequence 5'GCCU with the major RNA 3' end being at the U residue. Presumably, this cluster of RNA 3' ends detected by S₁ mapping corresponds to the sites on the DNA template where Rho protein brings about transcript release. The major release site is located 76 bp downstream from the rut insertion site in the plasmid. However, the possibility that the 3' ends

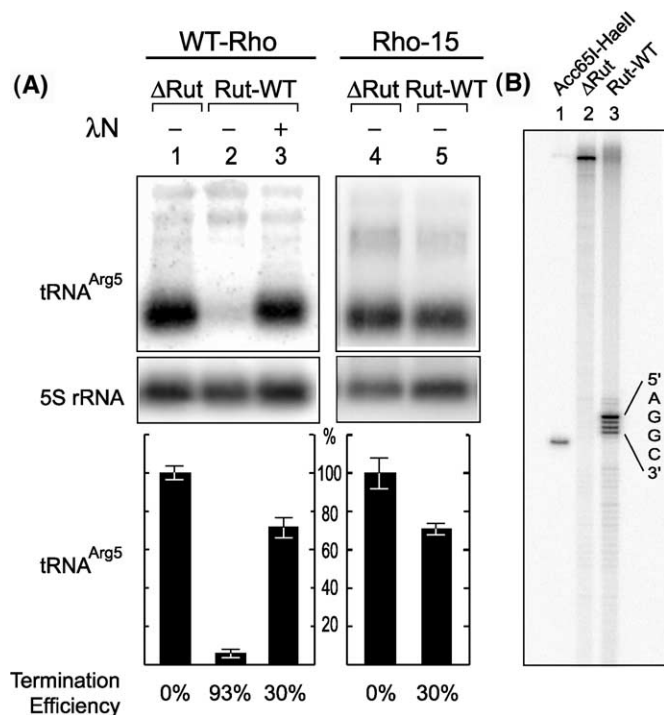


Figure 2. Termination and anti-termination *in vivo* with the minimal system. (A) Northern blot analyses of the *in vivo* accumulation of tRNA^{Arg5} in the absence (–) or the presence (+) of λ N protein. The results obtained in the Rho-15 mutant strain are also shown in parallel with those obtained with the Rho wild-type cells. The blots were first hybridized with an oligonucleotide probe that recognizes specifically tRNA^{Arg5} then stripped out and re-hybridized with a probe complementary to 5 S rRNA. For each lane, the tRNA^{Arg5} signal was normalized with the value obtained for 5 S rRNA. The levels of tRNA^{Arg5} detected with the different constructs were compared to the arbitrary 100% level obtained with the control plasmid lacking rut (Δ Rut). The error bars represent the standard deviation from a mean of three independent RNA extractions and northern blot analyses. The blots were quantified on a STORM PhosphorImager and

the data processed with Imagequant software version 5.1. (B) S₁ nuclease mapping of the Rho-dependent termination end-points. RNAs extracted from *E. coli* cells harboring the indicated plasmids were hybridized to the template strand of the Acc65I/PvuII DNA fragment labeled at the 3' end, digested with S₁ nuclease and analyzed on an 8% denaturing polyacrylamide gel. The molecular mass marker Acc65I/HaeII was obtained by digesting the labeled Acc65I/PvuII DNA fragment with HaeII. The sequence of the 5' tip of the DNA probe protected from S₁ digestion is shown on the right.

detected by S_1 mapping experiments may result from *in vivo* processing of transcripts released by Rho at more distal positions cannot be completely excluded. Such possibility was suggested for termination at tR1 in its natural context in the phage genome.¹⁷

Having shown that the *rut* element is sufficient to induce Rho-dependent transcription termination *in vivo* within our artificial operon, we investigated whether N protein is able to suppress the termination process. *E. coli* cells harboring the Rut-WT construct were transformed with a second plasmid that expresses the full length N protein from a constitutive promoter. The measurements of tRNA^{Arg5} accumulation in the co-transformant show that the production of λ N protein reduced the termination efficiency within the Rut-WT plasmid from 93% to 30% (compare lanes 2 and 3 in Figure 2(A)). Thus, in addition to Rho-dependent termination, our minimal system is suitable to study N-mediated antitermination *in vivo*.

The folding of boxB within the *rut* RNA is required for efficient Rho-dependent termination

The *rut* element of λ tR1 contains the N-utilization boxB motif located between the *rutA* and *rutB* sequences. Previous *in vivo* studies with a complete λ tR1 terminator have shown that in contrast to *rutA* or *rutB* deletions, the removal of the boxB motif does not have any significant effect on the efficiency of termination.¹³ To test whether boxB can have any contribution in our minimal system, we analyzed the transcriptional behavior of a construct harboring a *rut* element that lacks the boxB sequence (Rut- Δ boxB). The tRNA^{Arg5} measurements shown in Figure 3(A) are consistent with the previous report, they indicate that the boxB sequence contributes for only about 10% to the efficiency of termination (82% efficiency with Rut- Δ boxB as compared to 93% efficiency with Rut-WT, compare lanes 2 and 3). The analysis of the *in vivo* termination products by S_1 mapping revealed that Rho brings about transcript release in the Rut- Δ boxB construct at the same major release site (Figure 3(B), lane 2 and Figure 2(B), lane 3). Note, however, that the other minor RNA 3' ends are now located on the downstream side of the major endpoint. This may indicate a slightly different interaction of Rho with the boxB-less *rut* RNA sequence. Moreover, this difference observed for the minor release sites argues against the aforementioned possible *in vivo* processing of the 3' ends of the transcripts.

During the initial step of the termination process, Rho acts as an RNA binding protein. The appropriate binding site for hexameric Rho has been shown to involve at least 40–50 contiguous nucleotides of RNA.^{18,19} The size of *rutA* and *rutB* sequences in our minimal system is 25 nucleotides each. The fact that deletion of the intervening boxB sequence (15 nt) has a negligible effect on the efficiency of

termination suggests that the *rutA* and *rutB* RNA sequences may contribute collectively for optimal interaction with Rho at the initial step. This is conceivable if the folded boxB stem-loop structure serves as a structural hinge that brings the two *rut* RNA segments side by side as a continuous sequence. To test this possibility, we constructed a series of *rut* variant plasmids having mutations in either the ascending or the descending portions of the boxB hairpin stem (Figure 1(B)) that reduce considerably its folding potential. RNA secondary structure predictions were conducted at the Mfold web server²⁰ to explore the potential of the mutated sequences to form stable stem structures on their own or promote alternative folding of the whole *rut* sequence. As shown in Figure 3(A), the tRNA^{Arg5} measurements obtained with the plasmid Rut-boxB(C²⁻³A) revealed that two mutations in the left arm of the stem have a drastic effect on Rho activity with the termination efficiency falling down to 27% (lane 4). A similar result was observed with the construct Rut-boxB(G¹³⁻¹⁴A) in which the two destabilizing mutations are now located on the opposite strand of the stem (31% termination efficiency, lane 5). Conceivably, the termination defect of the two plasmid mutants could result merely from a change in the sequence context such as the creation of a C-poor sequence in Rut-boxB(C²⁻³A). To exclude this possibility, we constructed a third *rut* variant plasmid (Rut-boxB(G¹³⁻¹⁴C)) in which the G bases at positions 13 and 14 were replaced by C (Figure 1(B)). These mutations increase the cytosine contents of the boxB RNA sequence while maintaining an unfavorable folding potential. Again, these two mutations in the right arm of the stem showed also a termination defect (41% termination efficiency, lane 6). Remarkably, this termination defect was largely suppressed by additional replacements of the C²⁻³ on the opposite strand with G bases in Rut-boxB(C²⁻³ \leftrightarrow G¹³⁻¹⁴) which restore base-pairing of the stem (72% termination efficiency, lane 7). Thus, in contrast to the deletion experiments described above, these results strongly suggest that the boxB RNA motif plays an important role in Rho-dependent termination at λ tR1 but its contribution seems to be at the structural level. Apparently, the boxB hairpin acts by clamping together the *rutA* and *rutB* RNA segments for efficient interaction with Rho.

To provide additional evidence that the clamp-like structure of the intervening RNA motif is the primary determinant for Rho activity, we analyzed the thermodynamic stability of the different stem-loops used here. We have noticed in previous works with intrinsic terminators^{21,22} that the strength of a folded RNA hairpin structure embedded within a transcript correlates very well with its ability to interrupt the extension of a primer by M-MLV reverse transcriptase. Thus, we analyzed the RNAs produced *in vivo* from the different constructs by primer extension using a 5' end-labeled oligonucleotide that hybridizes to a

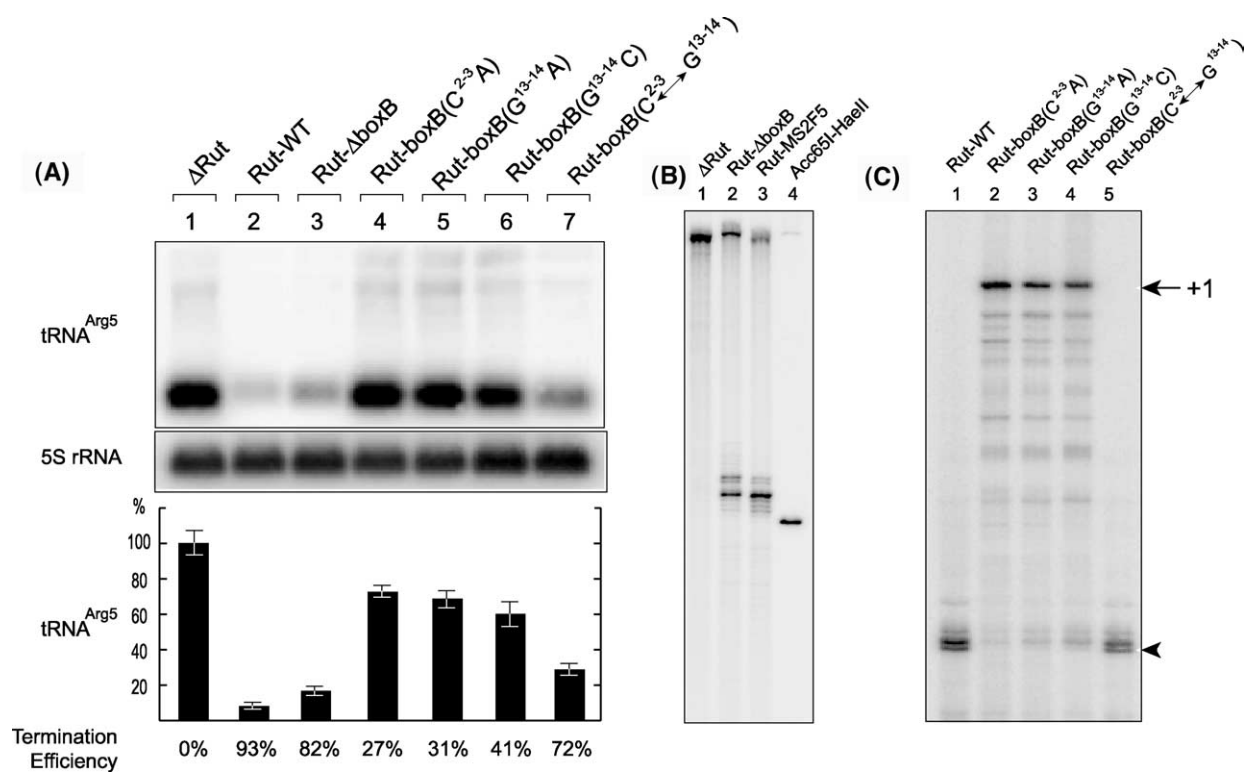


Figure 3. Structural contribution of boxB RNA motif to Rho-dependent termination. (A) Northern blot analyses of the *in vivo* accumulation of tRNA^{Arg5} in *E. coli* cells harboring different Rut derivative plasmids. The experiments were performed and quantified as described in the legend to Figure 2(A). (B) S₁ nuclease mapping of the Rho-dependent termination end-points within the indicated plasmid constructs. The experiments were carried out as in Figure 2(B). (C) Autoradiogram of the extension products obtained after *in vitro* reverse transcription of the RNAs extracted from cells harboring the indicated Rut-constructs. The reverse transcription reactions were performed using a ³²P-end-labeled primer that hybridizes to a region between the rut insertion site and the termination end-points and analyzed on an 8% denaturing polyacrylamide gel. The positions of the runoff products at the +1 position of the transcripts are indicated by an arrow. The arrowhead shows the position of the reverse transcriptase stops at the base of the hairpin stem.

region between the rut element and the termination site. As shown in Figure 3(C), most reverse transcription products obtained with either rutWT or rut-boxB(C²⁻³ ↔ G¹³⁻¹⁴) transcripts correspond to the stops at the base of the hairpin stem (lanes 1 and 5). In contrast, reverse transcriptase was able to efficiently read through the intervening hairpins that contain the destabilizing mutations (lanes 2–4). Apparently, these results show a correlation between the thermodynamic stability of an intervening RNA hairpin and its ability to organize a productive Rho binding site within the rut sequence.

Termination/antitermination at λ tR1 with a heterologous RNA/protein complex

Since the boxB RNA motif appears to contribute to the termination process simply as a clamp that positions rutB next to rutA for efficient interactions with Rho at the initial binding step, we reasoned that a similar contribution could be obtained with an unrelated well-folded RNA structural motif. We therefore constructed Rut-derivative plasmids in which the boxB sequence was replaced with

either the wild-type operator hairpin of bacteriophage MS2 or the SELEX-derived MS2F5 aptamer variant (Figure 1(B)). The two RNA stem-loops have been well characterized by biochemical and structural studies either alone or in complex with MS2 coat protein.^{23,24} *In vitro*, the MS2F5 RNA aptamer has a higher thermal stability and binds to coat protein with a better affinity than the wild-type MS2 operator hairpin.

To verify that in the context of our system the MS2F5 aptamer and the wild-type MS2 operator have also different thermodynamic stabilities, we analyzed the transcripts harboring the two intervening RNA motifs by reverse transcription experiments. As shown in Figure 4(A), the extension of the primer with reverse transcriptase was completely halted at the base of the MS2F5 aptamer motif (lane 3). Conversely, under the same conditions, the enzyme was able to synthesize through the wild-type MS2 sequence in a similar way to the boxB destabilizing mutants (lane 4). For the transcriptional behavior of the two MS2 derivative plasmids, the northern blotting analyses of tRNA^{Arg5} accumulation in cells transformed with the constructs are in complete agreement with the

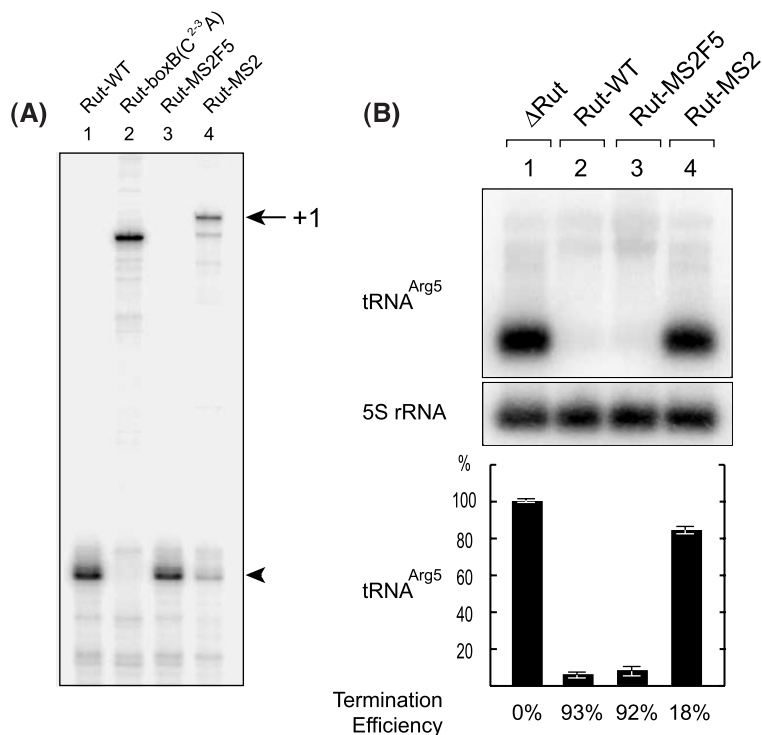


Figure 4. Rho-dependent termination with a heterologous RNA motif. (A) Analysis of the primer extension products obtained after *in vitro* reverse transcription of the RNAs extracted from cells harboring the indicated plasmid derivatives. The reverse transcription reactions were carried out as in Figure 3(C). The positions of the +1 of the transcript and the base of the hairpin stem are indicated by the arrow and the arrowhead, respectively. Note, however, that the transcript harboring MS2 hairpin is eight nucleotides longer than the Rut-boxB derivative. (B) Northern blot analyses of the *in vivo* accumulation of tRNA^{Arg5} in *E. coli* cells harboring plasmids with the MS2 derivative rut elements. The experiments were performed and quantified as described in the legend to Figure 2(A).

thermodynamic stability results (Figure 4(B)). The rut sequence containing the MS2F5 aptamer induces Rho-dependent transcription termination with the same efficiency as the rutWT element (compare lanes 2 and 3 in Figure 4(B)). Furthermore, the S₁ mapping experiments indicate that Rho-dependent transcript release occurs at the same DNA location as in the Rut-WT construct (Figure 3(B), lane 3). In contrast, the rut sequence containing the wild-type MS2 operator exhibits a low (about 18%, Figure 4(B), lane 4) Rho-dependent termination, reflecting the lower thermal stability of the hairpin stem as in the case of the boxB destabilizing mutations. Again, these results imply that the primary function of the intervening boxB RNA motif in Rho activity is simply to bring rutA and rutB in close proximity so that the two RNA segments form an uninterrupted binding site. Indeed, such structural function is better fulfilled when the intervening RNA stem-loop is stably folded.

The structural role of the boxB RNA motif in Rho-dependent termination at λ tR1 suggested to us that the N-mediated antitermination process may result from a competition between the two proteins for binding to neighboring sequences at the rut element wherein the N/boxB complex formation interferes with Rho interaction. To test this idea, we examined whether the antitermination process could be achieved by replacing N/boxB with the MS2F5/MS2 coat protein complex. *E. coli* cells harboring the Rut-MS2F5 construct were transformed with a second plasmid that overproduces the MS2 coat protein from a constitutive promoter. The northern blotting analyses of the RNAs extracted from the co-transformant show

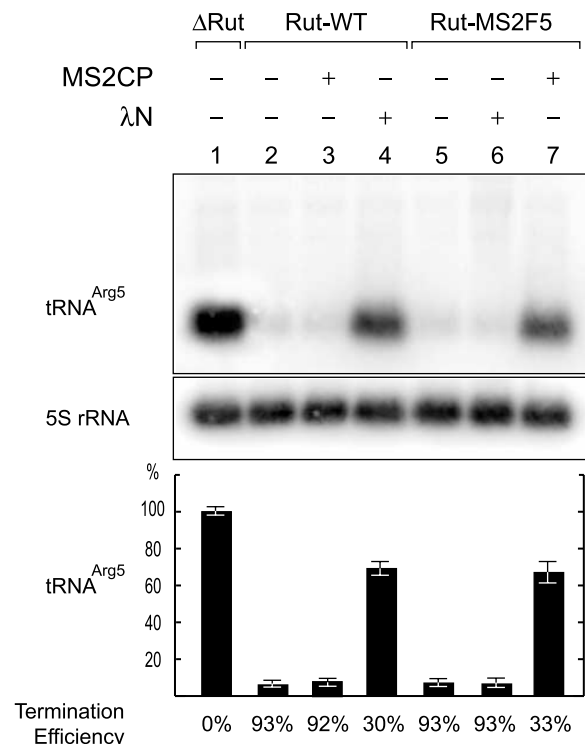


Figure 5. Antitermination with a heterologous RNA/protein complex. Northern blot analyses of the levels of tRNA^{Arg5} produced *in vivo* by the plasmids Rut-WT and Rut-MS2F5 in the absence (-) or the presence (+) of a second plasmid that overexpresses either λ N protein or MS2 coat protein (MS2CP). The experiments were performed and quantified as described in the legend to Figure 2(A).

clearly that the *in vivo* expression of MS2 coat protein causes antitermination within the Rut-MS2F5 construct (compare lanes 5 and 7 in Figure 5). The control co-transformation experiments in which λ N and MS2 coat protein expression vectors were interchanged demonstrate the specificity of each protein for its RNA target site in the antitermination process (Figure 5). This specificity was also observed when using other RNA/protein complexes (a test for RNA/protein interactions *in vivo* to be published elsewhere). Altogether, these results suggest that the formation of a ribonucleoprotein complex at the junction between *rutA* and *rutB* RNA segments interferes with the access of Rho to its binding site. Presumably, such interference is at the basis of the mechanism of antitermination at λ tR1.

Discussion

The molecular mechanisms by which Rho brings about transcript release at Rho-dependent termination sites are still not completely understood. It is clear, however, that the interactions between Rho protein and the nascent transcript are essential to the termination process.^{8–10} The protein first binds as a hexamer to a segment of the transcript (loading site) encoded by the upstream part of the terminator (*rut* element) and acts to promote transcript release downstream from its initial binding site. In λ tR1 terminator, which is the prototype for Rho-dependent termination, the *rut* element encodes two C-rich RNA regions, called *rutA* and *rutB*, that have been shown by deletion experiments as well as oligonucleotide competitions to be important for efficient Rho-dependent termination.^{11–13,25,26} The two regions are separated by the *boxB* RNA hairpin that serves as a binding site for λ N protein in the antitermination process and which is supposed to be dispensable for Rho activity.¹³ Here, we present *in vivo* data demonstrating that the *boxB* stem-loop does in fact play an important role in Rho-dependent transcription termination at λ tR1. However, the intervening RNA motif does not contribute to Rho activity as a sequence *per se* but acts as a structural clamp that organizes the loading site by holding *rutA* and *rutB* side by side for a productive interaction with Rho protein. A similar contribution to Rho activity is readily obtained when the *boxB* stem-loop is replaced by an unrelated RNA hairpin provided that the structural motif is thermodynamically stable.

A model that summarizes our findings with the minimal system and their relevance to the mechanisms of termination/antitermination at λ tR1 is illustrated in Figure 6. In the context of the phage genome, the tR1 terminator is located between the *cro* and *cII* genes, just downstream from the open reading frame of the Cro protein. Thus, the termination process should start as soon as sufficient naked RNA, following ribosome release at the *cro*

stop codon, emerges from the transcription elongation complex and becomes available to recruit a Rho hexamer and secure it on the transcript. Since the deletion experiments have shown that the *rutA* region contributes for about 64% of the termination efficiency,¹³ a large part of this RNA segment is probably involved in this initial interaction. Recent crystal structures of hexameric Rho bound to nucleic acid substrates revealed that a 12 nt RNA is the minimal segment required to span two adjacent protomers within the ring-shaped hexameric Rho, each protomer interacting with a cytosine and an adjacent pyrimidine residue through its N-terminal domain.^{27,28} Accordingly, the entire *rutA* RNA segment (about 33 nt in size, between the *cro* stop codon and the base of the *boxB* hairpin stem) would cover at most three protomers. Thus, a fully productive interaction between hexameric Rho and the transcript will require binding of additional C-rich RNA sequences such as *rutB*, which will bulge out the intervening RNA hairpin motif (Figure 6(A)). Considering the kinetic competition between Rho and the transcribing polymerase in the termination process,^{29,30} factors that would favor timely folding of a rather stable intervening hairpin should facilitate a fully productive interaction, thereby enabling Rho to catch up with the elongation complex before it enters the open reading frame of the *cII* gene. It is, therefore, tempting to speculate that apart from constraints specific to N binding, it may not be surprising that *boxB* is made of a short GC-rich stem capped by a GNRA-like loop, a combination that should favor rapid hairpin folding.^{31,32} This could explain the more pronounced termination defect observed with the construct harboring the wild-type MS2 operator hairpin, which has a longer stem and is not capped with a GNRA loop, as compared to the constructs with the *boxB* having destabilizing mutations.

Unlike tR1, the *E. coli trp t'* terminator, another relatively well studied Rho-dependent termination signal, has a more extended upstream *rut* element and remains partially functional upon multiple significant deletions within the *rut* sequences. These results suggested that the *trp t'* *rut* element may contain a greater redundancy of RNA segments necessary for productive interaction with Rho.^{29,30,33} Consistent with this suggestion, RNA repeats containing regularly spaced cytosine residues have been shown to serve as loading sites for Rho leading to efficient termination.^{34–36} Thus, our results do not exclude alternative interactions between Rho and the *rut* RNA sequence that may account for a minor part of the termination efficiencies observed with the different constructs. Such alternative interactions may involve additional C residues within the intervening RNA motif or within sequences flanking the *rut* element in our minimal system. This could explain the less pronounced termination defect observed with the construct Rut-*boxB*(G^{13–14}C) where alternative interactions with the two extra

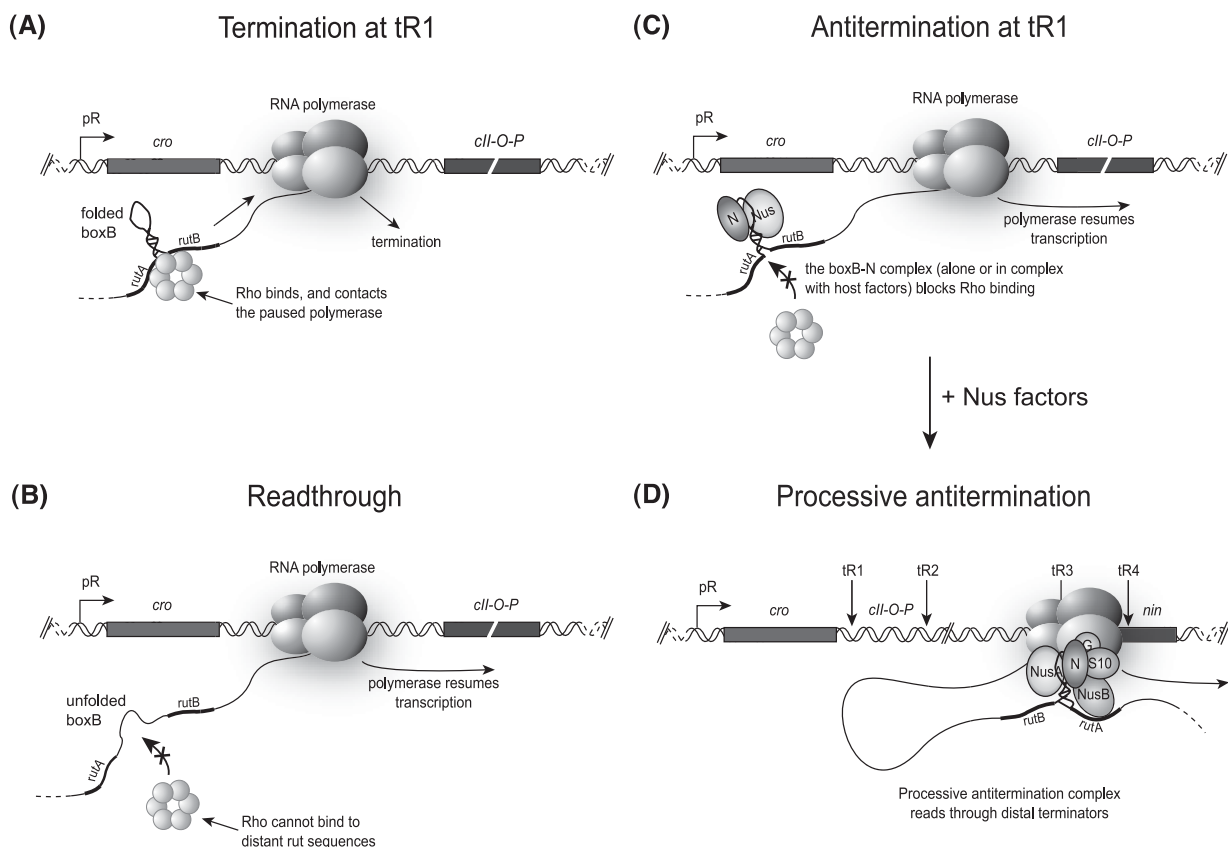


Figure 6. Model for the mechanisms of termination/antitermination at λ tR1 terminator. (A) In the absence of λ N protein, Rho binds to the *cro* transcript at *rutA* and *rutB* attachment sites that are clamped together by the intervening well-folded *boxB* hairpin. Upon productive binding, Rho causes transcription termination at downstream positions. (B) If the intervening hairpin is unfolded, Rho cannot bind productively to the distant *rut* RNA sequences. This will lead to transcriptional read through. (C) When λ N protein is present, its binding (either alone or in complex with the Nus factors) to *boxB* hairpin interferes with the access of Rho to the *rut* sequences which will prevent termination. (D) Additional Nus factors associate with N and RNA polymerase to promote the formation of a processive antitermination complex able to read through the more distal terminators (tR2, tR3, tR4).

C residues within the unfolded *boxB* mutant could be favored.

The specific structural constraints of λ tR1, i.e. size-limited *rutA* and *rutB* elements separated by *boxB*, are presumably dictated by its physiological regulation by the N-mediated antitermination process. Our co-transformation experiments show that λ N protein or the unrelated MS2 coat protein suppress Rho-dependent termination in the minimal system when the *Rut* construct encodes their cognate RNA binding site. Apparently, the binding of the proteins to their RNA target hairpin in the middle of the *rut* sequences interferes with the access of Rho to the loading site. These results imply that such interference is probably at the basis of the N-mediated antitermination at λ tR1 (Figure 6(C)). A similar antitermination process was recently reported for the regulation of the *E. coli* tryptophanase operon (*tna*) in which a translating ribosome stalled at a stop codon was shown to suppress termination by interfering with the access of Rho to its loading site.^{37,38} Interestingly, this Rho-dependent terminator, which lies at the end of the leader region of the *tna* operon, has a

relatively short *rut* element, a constraint shared with λ tR1. Additional support to our interpretation comes from studies on the phage HK022 Nun protein, which excludes λ from *E. coli* by binding to the *nut* RNA sequences and blocking transcription through interactions with RNAP and the DNA template. Mutant Nun proteins that are unable to block transcription, due to a defect in contacting the DNA template, have been shown to stimulate transcriptional read through at λ tR1.^{39–41} These results suggested that Nun binding to *boxB* prevents access of Rho to its loading site.

The suppression of termination at λ tR1 by a physical obstruction to Rho binding, shown here, is independent from the subsequent modification of RNAP by N protein, in association with the host Nus factors, into a form that is resistant to downstream terminators. The formation of such a processive antitermination complex is required to read through the more distal Rho-dependent terminators tR3 and tR4 within the *nin* region of the phage genome (Figure 6(D)).⁴² Previous studies with the closely related ribosomal RNA anti-termination system^{43,44} have also shown that once

the antitermination complex is formed, RNAP is able to read through Rho-dependent terminators placed at positions distal to the nut sequence. The molecular mechanisms enabling the transcription complex to bypass distant Rho-dependent termination sites are not well understood. However, since N-binding to RNAP is known to increase the rate of transcript elongation, a more favorable kinetic competition in which the transcribing polymerase translocates faster than Rho is often invoked as a likely mechanism. In the case of λ tR1, whether the complex formed by N and boxB is sufficient on its own to interfere with Rho binding or whether it requires the loading of the Nus factors to increase the size of the steric hindrance cannot be answered unambiguously in the present work and will need additional investigations with Nus mutants. Nevertheless, our results with the MS2 coat protein (which has a similar size to N but binds as a dimer) suggest that a stable RNA/protein complex containing only the two partners is sufficient to prevent Rho binding. In this regard, our minimal system constitutes a powerful method to study RNA/protein interactions *in vivo* (results to be published).

Materials and Methods

Enzymes, chemicals and oligonucleotides

Restriction enzymes, T4 polynucleotide kinase and Klenow fragment of DNA polymerase I were obtained from New England Biolabs. T4 DNA ligase, S₁ nuclease and Superscript RNase H⁻ Reverse Transcriptase from M-MLV were bought from Invitrogen. Most chemicals including antibiotics were from Sigma. Unlabeled dNTPs were obtained from Boehringer Mannheim, whereas [α -³²P]dATP and [γ -³²P]ATP were from Amersham Biosciences. All oligonucleotides were synthesized by MWG-Biotech.

Bacterial strains and plasmids

The Rut derivative plasmids are based on a pACYC177 parent vector and contain a hisR^{HPa} constitutive promoter⁴⁵ that drives transcription of the *cat* gene (deleted of its Shine-Dalgarno sequence and ATG translation initiation codon) followed by the *tRNA*^{Arg5} gene.¹⁵ The *tRNA*^{Arg5} gene along with the downstream intrinsic terminator were subcloned from pOM10.¹⁶ The rut-derivative sequences were cloned between Sall and Acc65I restriction sites using oligonucleotides. The Δ Rut plasmid corresponds to the parent construct digested by Sall-Acc65I, filled-in and re-ligated. The correct constructs were selected by dideoxy sequencing. The plasmids were maintained in *E. coli* strain SU1675 (our laboratory collection) by selecting with 30 μ g/ml kanamycin. For the co-transformation experiments, the second plasmid that overproduces N protein or MS2 coat protein is a compatible pBR322-based vector that was maintained by selecting with ampicillin (100 μ g/ml). The N coding sequence was amplified by PCR from phage λ DNA using the following oligonucleotides: 5'GCGGGATCCATGTGCCAATC-GCGGGG and 5'AGATCTATCAGATAAGAGGAATCGATTTCC. The MS2

coat protein coding sequence was amplified from the plasmid pCT119⁴⁶ using the following oligonucleotides: 5'GGAGTTGGATCCATGGCTTCTAACTTTAC and 5'CGGCCGAGATCTATTAGTAGATGCCGGAG. Each PCR fragment was cloned into the pBR322-derivative pZE12-luc plasmid⁴⁷ under the control of a plac variant promoter. In the strain SU1675 (Δ lac) the two proteins are expressed constitutively. *E. coli* rho-15 mutant IT1631⁴⁸ harboring the indicated plasmid was grown at 30 °C up to an $A_{600\text{ nm}}$ of 0.2–0.4 and then shifted to 42 °C for 45 minutes before RNA extraction.

RNA isolation and northern blotting analyses

Total bacterial RNAs were extracted from the cells harboring the indicated plasmids as described.⁴⁹ For the northern blot analyses, 15 μ g of total RNAs were separated on a 1.2% (w/v) agarose gel containing 6% (v/v) formaldehyde and electro-blotted (TRANS-BLOT SD from Biorad) onto a nylon membrane Hybond-N⁺ (Q-biogen). The blots were probed and washed as described.¹⁵ A negative control with no tRNA^{Arg5} was always included to ensure the specificity of the tRNA^{Arg5} signal. Hybridization signals were quantified with a Storm PhosphorImager (Molecular Dynamics) and the data processed with Imagequant software version 5.1.

S₁ mapping and primer extension

To prepare the single-stranded DNA probe, the plasmid DNA was digested by Acc65I and PvuII, then labeled at the 3' end by filling-in with [α -³²P]dATP and the Klenow fragment of DNA polymerase I. Separation of the two strands of the Acc65I/PvuII fragment and S₁ mapping experiments were performed as described⁵⁰ except that the hybridization of total RNAs with the labeled DNA strand was carried out in a buffer containing 75% formamide. Primer extension experiments with M-MLV reverse transcriptase were carried out exactly as described.⁵¹

Acknowledgements

We are highly grateful to our laboratory members for stimulating discussions and technical advice. We thank Marc Boudvillain for reading the manuscript and D. Peabody, H. Bujard & F. Boccard for the gift of plasmids. This work was supported in part by la Ligue contre le Cancer, l'ARC and l'ANRS. É.V. is a recipient of a fellowship from la Ligue contre le Cancer.

References

1. Das, A. (1993). Control of transcription termination by RNA-binding proteins. *Annu. Rev. Biochem.* **62**, 893–930.
2. Roberts, J. W. (1993). RNA and protein elements of *E. coli* and lambda transcription antitermination complexes. *Cell*, **72**, 653–655.
3. Friedman, D. I. & Court, D. L. (1995). Transcription antitermination: the lambda paradigm updated. *Mol. Microbiol.* **18**, 191–200.
4. Greenblatt, J., Mah, T. F., Legault, P., Mogridge, J., Li,

- J. & Kay, L. E. (1998). Structure and mechanism in transcriptional antitermination by the bacteriophage lambda N protein. *Cold Spring Harb. Symp. Quant. Biol.* **63**, 327–336.
5. Roberts, J. W. (1969). Termination factor for RNA synthesis. *Nature*, **224**, 1168–1174.
 6. Gusarov, I. & Nudler, E. (2001). Control of intrinsic transcription termination by N and NusA: the basic mechanisms. *Cell*, **107**, 437–449.
 7. Nudler, E. & Gottesman, M. E. (2002). Transcription termination and anti-termination in *E. coli*. *Genes Cells*, **7**, 755–768.
 8. Platt, T. (1994). Rho na dRNA: models for recognition and response. *Mol. Microbiol.* **11**, 983–990.
 9. Richardson, J. P. & Greenblatt, J. (1996). Control of RNA chain elongation and termination. *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Neidhardt, F. C., et al.), pp. 822–848, American Society for Microbiology, Washington, DC.
 10. Richardson, J. P. (2002). Rho-dependent termination and ATPases in transcript termination. *Biochim. Biophys. Acta*, **1577**, 251–260.
 11. Chen, C. Y. & Richardson, J. P. (1987). Sequence elements essential for rho-dependent transcription termination at lambda tR1. *J. Biol. Chem.* **262**, 11292–11299.
 12. 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. *J. Biol. Chem.* **271**, 21597–21603.
 13. Graham, J. E. & Richardson, J. P. (1998). Rut sites in the nascent transcript mediate Rho-dependent transcription termination *in vivo*. *J. Biol. Chem.* **273**, 20764–20769.
 14. Sloan, S. B. & Weisberg, R. A. (1993). Use of a gene encoding a suppressor tRNA as a reporter of transcription: analyzing the action of the Nun protein of bacteriophage HK022. *Proc. Natl Acad. Sci. USA*, **90**, 9842–9846.
 15. Lopez, P. J., Iost, I. & Dreyfus, M. (1994). The use of a tRNA as a transcriptional reporter: the T7 late promoter is extremely efficient in *Escherichia coli* but its transcripts are poorly expressed. *Nucl. Acids Res.* **22**, 1186–1193.
 16. Espéli, O., Moulin, L. & Boccard, F. (2001). Transcription attenuation associated with bacterial repetitive extragenic BIME elements. *J. Mol. Biol.* **314**, 375–386.
 17. Roberts, E. A., Eisenbraun, T. L., Andrews, C. L. & Bear, D. G. (1991). 3'-end formation at the phage lambda tR1 rho-dependent transcription termination site. *Biochemistry*, **30**, 5429–5437.
 18. Morgan, W. D., Bear, D. G., Litchman, B. L. & von Hippel, P. H. (1985). RNA sequence and secondary structure requirements for rho-dependent transcription termination. *Nucl. Acids Res.* **13**, 3739–3754.
 19. Alifano, P., Rivellini, F., Limauro, D., Bruni, C. B. & Carlomagno, M. S. (1991). A consensus motif common to all Rho-dependent prokaryotic transcription terminators. *Cell*, **64**, 553–563.
 20. Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucl. Acids Res.* **31**, 3406–3415.
 21. Epshtein, V., Toulmé, F., Rahmouni, A. R., Borukhov, S. & Nudler, E. (2003). Transcription through the roadblocks: the role of RNA polymerase cooperation. *EMBO J.* **22**, 4719–4727.
 22. Schwartz, A., Rahmouni, A. R. & Boudvillain, M. (2003). The functional anatomy of an intrinsic transcription terminator. *EMBO J.* **22**, 3385–3394.
 23. Rowsell, S., Stonehouse, N. J., Convery, M. A., Adams, C. J., Ellington, A. D. & Hirao, I. (1998). Crystal structures of a series of RNA aptamers complexed to the same protein target. *Nature Struct. Biol.* **5**, 970–975.
 24. Parrot, A. M., Lago, H., Adams, C. J., Ashcroft, A. E., Stonehouse, N. J. & Stockley, P. G. (2000). RNA aptamers for the MS2 bacteriophage coat protein and the wild-type RNA operator have similar solution behaviour. *Nucl. Acids Res.* **28**, 489–497.
 25. Chen, C. Y., Galluppi, G. R. & Richardson, J. P. (1986). Transcription termination at lambda tR1 is mediated by interaction of rho with specific single-stranded domains near the 3' end of cro mRNA. *Cell*, **46**, 1023–1028.
 26. Faus, I. & Richardson, J. P. (1990). Structural and functional properties of the segments of lambda cro mRNA that interact with transcription termination factor Rho. *J. Mol. Biol.* **212**, 53–66.
 27. Skordalakes, E. & Berger, J. M. (2003). Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell*, **114**, 135–146.
 28. Richardson, J. P. (2003). Loading Rho to terminate transcription. *Cell*, **114**, 157–159.
 29. Zhu, A. Q. & von Hippel, P. H. (1998). Rho-dependent termination within the trp t' terminator. I. Effects of rho loading and template sequence. *Biochemistry*, **37**, 11202–11214.
 30. Zhu, A. Q. & von Hippel, P. H. (1998). Rho-dependent termination within the trp t' terminator. II. Effects of kinetic competition and rho processivity. *Biochemistry*, **37**, 11215–11222.
 31. Legault, P., Li, J., Mogridge, J., Kay, L. E. & Greenblatt, J. (1998). NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell*, **93**, 289–299.
 32. Sorin, E. J., Engelhardt, M. A., Herschlag, D. & Pande, V. S. (2002). RNA simulations: probing hairpin unfolding and the dynamics of a GNRA tetraloop. *J. Mol. Biol.* **317**, 493–506.
 33. Zalatan, F., Galloway-Salvo, J. & Platt, T. (1993). Deletion analysis of the *Escherichia coli* rho-dependent transcription terminator trp t'. *J. Biol. Chem.* **268**, 17051–17056.
 34. Hart, C. M. & Roberts, J. W. (1991). Rho-dependent transcription termination. Characterization of the requirement for cytidine in the nascent transcript. *J. Biol. Chem.* **266**, 24140–24148.
 35. Hart, C. M. & Roberts, J. W. (1994). Deletion analysis of the lambda tR1 termination region. Effect of sequences near the transcript release sites, and the minimum length of rho-dependent transcripts. *J. Mol. Biol.* **237**, 255–265.
 36. Guérin, M., Robichon, N., Geiselman, J. & Rahmouni, A. R. (1998). A simple polypyrimidine repeat acts as an artificial Rho-dependent terminator *in vivo* and *in vitro*. *Nucl. Acids Res.* **26**, 4895–4900.
 37. Konan, K. V. & Yanofsky, C. (2000). Rho-dependent transcription termination in the tna operon of *Escherichia coli*: roles of the boxA sequence and the rut site. *J. Bacteriol.* **182**, 3981–3988.
 38. Gong, F. & Yanofsky, C. (2002). Analysis of tryptophanase operon expression *in vitro*: accumulation of TnaC-peptidyl-tRNA in a release factor 2-depleted

- S-30 extract prevents Rho factor action, simulating induction. *J. Biol. Chem.* **277**, 17095–17100.
39. Chattopadhyay, S., Hung, S. C., Stuart, A. C., Palmer, A. G., Garcia-Mena, J., Das, A. & Gottesman, M. E. (1995). Interaction between the phage HK022 Nun protein and the nut RNA of phage lambda. *Proc. Natl Acad. Sci. USA*, **92**, 12131–12135.
 40. Hung, S. C. & Gottesman, M. E. (1995). Phage HK022 Nun protein arrests transcription on phage lambda DNA *in vitro* and competes with the phage lambda N antitermination protein. *J. Mol. Biol.* **247**, 428–442.
 41. Kim, H. C., Zhou, J. G., Wilson, H. R., Mogilnitskiy, G., Court, D. L. & Gottesman, M. E. (2003). Phage HK022 Nun protein represses translation of phage lambda N (transcription termination/translation repression). *Proc. Natl Acad. Sci. USA*, **100**, 5308–5312.
 42. Cheng, S. C., Court, D. L. & Friedman, D. I. (1995). Transcription termination signals in the nin region of bacteriophage lambda: identification of Rho-dependent termination regions. *Genetics*, **140**, 875–887.
 43. Squires, C. L., Greenblatt, J., Li, J., Condon, C. & Squires, C. L. (1993). Ribosomal RNA antitermination *in vitro*: requirement for Nus factors and one or more unidentified cellular components. *Proc. Natl Acad. Sci. USA*, **90**, 970–974.
 44. Torres, M., Condon, C., Balada, J. M., Squires, C. & Squires, C. L. (2001). Ribosomal protein S4 is a transcription factor with properties remarkably similar to NusA, a protein involved in both non-ribosomal and ribosomal RNA antitermination. *EMBO J.* **20**, 3811–3820.
 45. Figueroa-Bossi, N., Guérin, M., Rahmouni, A. R., Leng, M. & Bossi, L. (1998). The supercoiling sensitivity of a bacterial tRNA promoter parallels its responsiveness to stringent control. *EMBO J.* **17**, 2359–2367.
 46. Peabody, D. S. (1990). Translational repression by bacteriophage MS2 coat protein expressed from a plasmid. A system for genetic analysis of a protein–RNA interaction. *J. Biol. Chem.* **265**, 5684–5689.
 47. Lutz, R. & Bujard, H. (1997). Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/11-I2 regulatory elements. *Nucl. Acids Res.* **25**, 1203–1210.
 48. Das, A., Court, D. & Adhya, S. (1976). Isolation and characterization of conditional lethal mutants of *Escherichia coli* defective in transcription termination factor rho. *Proc. Natl Acad. Sci. USA*, **73**, 1959–1963.
 49. Guérin, M., Leng, M. & Rahmouni, A. R. (1996). High resolution mapping of *E. coli* transcription elongation complex *in situ* reveals protein interactions with the non-transcribed strand. *EMBO J.* **15**, 5397–5407.
 50. Toulmé, F., Guérin, M., Robichon, N., Leng, M. & Rahmouni, A. R. (1999). *In vivo* evidence for back and forth oscillations of the transcription elongation complex. *EMBO J.* **18**, 5052–5060.
 51. Toulmé, F., Mosrin-Huaman, C., Sparkowski, J., Das, A., Leng, M. & Rahmouni, A. R. (2000). GreA and GreB proteins revive backtracked RNA polymerase *in vivo* by promoting transcript trimming. *EMBO J.* **19**, 6853–6859.

Edited by M. Gottesman

(Received 12 February 2004; received in revised form 2 April 2004; accepted 6 April 2004)

3. Conclusion.

L'ensemble des résultats obtenus au cours de ce travail suggère que le motif ARN BOXB possède une fonction double au sein du terminateur λ tR1 :

- Tout d'abord, en se structurant sous forme d'une tige boucle extrêmement stable, il rapproche RUTA et RUTB l'un à côté de l'autre pour former un site RUT unique dont l'interaction avec Rho est optimale.
- Par ailleurs, la fixation de la protéine N sur le motif BOXB, induit le mécanisme d'antiterminaison au niveau de tR1, en bloquant l'accès de Rho à sa séquence RUT.

Les résultats que nous avons obtenus avec le couple MS2F5/MS2 coat protein suggèrent qu'un complexe ARN-protéine hétérologue stable, formé au sein de l'élément RUT d'un terminateur Rho dépendant, interfère avec la fixation du facteur de terminaison sur l'ARN et avec le relargage de l'ARN en aval de cet élément RUT.

De façon intéressante, le système minimal que nous avons développé au cours de ce travail semble adapté à la détection, *in vivo*, de complexes ribonucléoprotéiques se formant au sein de l'élément RUT d'un terminateur. Dans cette optique ce système d'analyse génétique, basé sur l'interférence avec la terminaison Rho-dépendante, pourrait s'avérer une méthode puissante pour étudier les interactions ARN-protéine chez *E. coli*.

Chapitre VI. DEVELOPPEMENT D'UN SYSTEME DE SELECTION D'APTAMERS *IN VIVO*.

Un grand nombre de systèmes existent actuellement, aussi bien chez les procaryotes que chez les eucaryotes, pour sélectionner des petites molécules (aptamers) spécifiques d'une protéine donnée, où pour analyser une interaction ARN-protéine connue (voir : chapitre IV). Cependant ces systèmes ne sont pas dépourvus d'inconvénients majeurs :

- Certains, comme le triple hybride, mais aussi les systèmes d'antiterminaison par des fusions avec les protéines N du phage λ et Tat du VIH, reposent sur l'utilisation de molécules chimériques. Leur fonctionnement dépend donc du bon comportement, une fois dans le milieu cellulaire, des différents composants utilisés. De mauvais repliements des différents partenaires, des interactions parasites les impliquant, peuvent malheureusement aboutir à l'apparition de nombreux faux positifs.
- Les systèmes reposant sur l'inhibition de la traduction d'un gène reporter sont beaucoup plus simples au niveau de leur utilisation. Ils reposent sur la baisse d'activité d'un gène consécutive à la fixation d'une protéine à proximité des sites nécessaires à la traduction de son ARNm. Cette baisse d'activité peut être visualisée, via la coloration ou plutôt la décoloration des colonies, en utilisant, par exemple, le gène *lacZ* comme reporter. Ce type de sélection, efficace pour étudier un complexe ARN-protéine connu, n'est malheureusement pas adapté à un crible à grande échelle. Celui-ci requiert plutôt l'utilisation d'un gène permettant aux cellules de survivre, soit en milieu minimal, soit en présence d'un antibiotique. Or, un système reposant sur la chute d'activité d'un reporter n'est pas compatible avec ce type de sélection. Pour résoudre ce problème il faudrait, par exemple, utiliser un gène codant pour un répresseur transcriptionnel; la chute d'activité de ce répresseur permettrait alors d'activer un gène reporter secondaire adapté, mais, avouons le, cette solution serait un peu compliquée.

C'est pour ces diverses raisons que nous avons décidé de mettre au point un système de sélection (crible génétique) qui devrait s'avérer beaucoup plus simple d'utilisation tout en étant aussi performant. Ce crible est basé sur le phénomène de terminaison Rho-dépendante de la transcription chez *E. coli*.

1. Hypothèse de départ.

La terminaison Rho-dépendante (voir : chapitre II) intervient lorsque Rho reconnaît et fixe un site nommé RUT au niveau du transcrit néosynthétisé. Cette fixation initiale amorce une cascade d'évènements aboutissant à la dissociation du complexe ternaire d'élongation et au relargage du transcrit. Cette fixation constitue donc une étape déterminante pour la terminaison Rho-dépendante. Les facteurs (nucléotidiques ou protéiques) qui sont susceptibles d'empêcher cette fixation, vont alors inhiber la terminaison.

Dans le travail précédent, nous avons montré que la formation d'un complexe ARN-protéine hétérologue (MS2F5-MS2 coat protein), entre les séquences RUTA et RUTB, était capable d'induire l'antiterminaison de la transcription au niveau du terminateur tR1 du phage λ . De même, les travaux réalisés dans le laboratoire de C. Yanofsky montrent que des ribosomes, bloqués sur le transcrit à proximité de la séquence RUT de l'opéron *tna*, interfèrent avec la terminaison dépendante de Rho (Konan et Yanofsky, 2000; Gong et Yanofsky, 2002). Ces observations montrent que la formation d'un complexe ribonucléoprotéique, à proximité ou au sein de la séquence d'entrée du facteur Rho, inhibe la fixation de Rho sur le transcrit et le relargage de l'ARN. L'expression d'un gène reporter en aval du terminateur tR1 témoignerait donc de la formation d'un complexe ARN-protéine, entre RUTA et RUTB, suffisamment stable pour bloquer la fixation de Rho sur l'ARN.

J'ai mis à profit cette particularité pour élaborer notre système de sélection et d'analyse *in vivo* des interactions ARN-protéine chez *E. coli*. En effet, en remplaçant systématiquement *boxB* par une séquence courte aléatoire, il deviendrait possible de cribler les interactions éventuelles entre l'ARN codé par cette séquence et une protéine d'intérêt, au travers du niveau de transcription d'un gène rapporteur inséré en aval.

1. 1. Le terminateur tR1 n'est pas adapté à l'élaboration du crible.

L'un des résultats obtenus précédemment, c'est-à-dire la chute de l'efficacité de terminaison avec le dérivé portant le motif MS2 sauvage entre RUTA et RUTB, semble indiquer que l'utilisation de tR1 serait limitée à l'étude des seuls motifs ARN dont la structure en tige boucle agit comme une pince permettant l'assemblage correct d'un élément RUT fonctionnel à partir des séquences RUTA et RUTB. L'insertion d'un motif juste à proximité

des séquences *rut* (comme dans l'opéron *tna*) semble également déconseillé. En effet, un motif ARN en tige boucle repose généralement sur la présence de nombreuses paires GC. Il est possible alors que sa séquence entre en interaction avec les résidus de la BOXB pour générer des structures alternatives incompatibles avec la fixation de Rho sur le transcrit.

Ces différentes observations montrent que l'utilisation du terminateur tR1 n'est pas adaptée à l'élaboration du crible; c'est pour cette raison que nous avons utilisé une séquence *rut* synthétique, développée au laboratoire, qui induit une terminaison Rho-dépendante quasi-totale, *in vitro* et *in vivo*.

1. 2. Description de la séquence d'entrée synthétique.

En 1998, un travail réalisé au laboratoire (Guérin et al., 1998) décrit l'élaboration d'un terminateur Rho-dépendant artificiel, par insertion d'une séquence ADN $(AG/TC)_n$ au sein d'une unité de transcription (Figure 27).

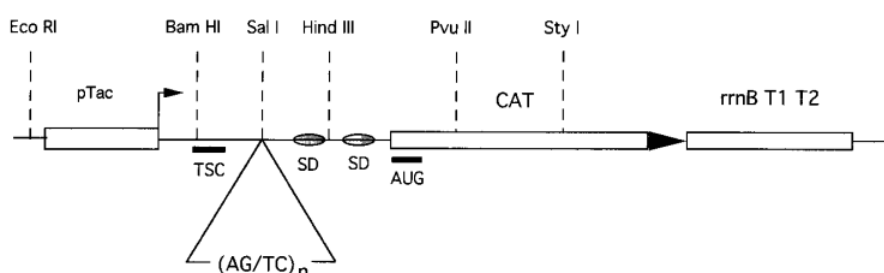


Figure 27 : Représentation schématique de la construction plasmidique utilisée pour étudier l'implication de la séquence $(AG/TC)_n$ sur la transcription chez *E. coli* (d'après : Guérin et al., 1998).

Les expériences menées *in vitro* et *in vivo* montrent que la transcription de cette séquence, dans l'orientation aboutissant aux répétitions $(rUrC)_n$ sur l'ARN, induit le relargage du transcrit au niveau d'une séquence située dans le gène *cat*, 130pb en aval. Comme démontré par les expériences réalisées chez un mutant thermosensible (Rho-111, Figure 28C), cette terminaison est dépendante de Rho. Elle est observée pour des inserts de 30 ou 44 pb de long (constructions pTC30 et pTC44; Figure 28), et elle est quasiment totale pour la répétition de 44pb. En revanche, l'orientation inverse (constructions pAG30 et pAG44, Figure 28),

généralisant une répétition (rArG)_n sur le transcrit, ne permet pas d'activer les fonctions de Rho. Cette construction sert de référence permettant d'attribuer la terminaison à la présence de la séquence (rUrC)_n dans l'ARN.

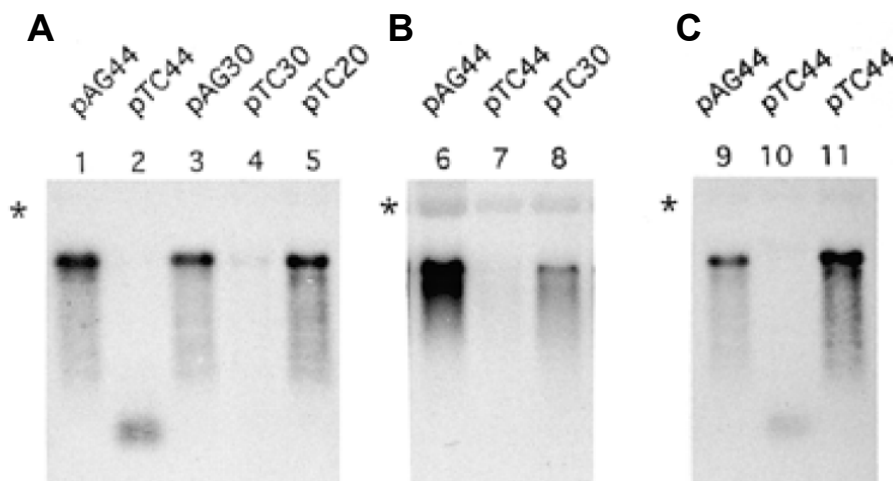


Figure 28 : Analyse, par Northern blot, du niveau de transcription en fonction du type de séquence insérée entre le promoteur et le gène *cat*. Les ARNs totaux sont extraits à partir de cellules, d'*E. coli* (A et B), ou de *S. typhimurium* (C; souche sauvage : lignes 9 et 10, ou mutant *rho-111* : ligne 11). Les membranes sont révélées grâce à une sonde HindIII-PvuII (A et C) ou une sonde PvuII-StyI (B).

Au cours de ce travail, l'implication de la répétition (rUrC)_n, dans la fixation de Rho sur le transcrit a été confirmée par des expériences de compétition à l'aide d'un oligonucléotide AG₁₈ complémentaire de la séquence (rURC)_n. En effet, l'hybridation de l'oligonucléotide sur le transcrit, au niveau de la séquence (rUrC), interfère avec la terminaison en bloquant l'accès de Rho à la séquence d'entrée synthétique.

Les expériences d'empreintes (Bear et al., 1988), ainsi que la structure tridimensionnelle de Rho (Skordalakes et Berger, 2003), suggèrent que la fixation initiale de Rho sur le transcrit implique une séquence ARN de 70-80 nucléotides. Ici, pourtant, la séquence générant une terminaison optimale est beaucoup plus courte (44pb); on ne peut pas exclure que les séquences adjacentes soient également impliquées dans l'interaction. Toutefois, la répétition de 44pb, lorsqu'elle est insérée dans le gène *cat* (au site PvuII), entraîne une terminaison aussi efficace que lorsqu'elle est insérée plus en amont (au site Sall,

voir [Figure 27](#)). De plus, les séquences entourant chacun des sites d'insertion de la séquence synthétique ne présentent pas de zones particulièrement riches en résidus cytosines.

L'ensemble de ces travaux a montré que cette séquence de 44pb était suffisante pour induire une terminaison efficace de la transcription, et cela quel que soit son environnement. L'alternance des bases C et T au sein de cette séquence a ensuite été légèrement modifiée pour obtenir une nouvelle séquence notée $(CTT)_n$ et qui, cette fois, provoque une terminaison totale de la transcription. Cette nouvelle séquence représente donc l'outil idéal pour élaborer notre crible génétique.

2. Principe du système.

Un complexe ribonucléoprotéique lorsqu'il se forme à proximité d'une séquence RUT, est capable de bloquer la fixation de Rho sur l'ARN et d'interférer avec la terminaison. Le principe du crible met à profit cette faculté pour permettre de repérer la formation d'un tel complexe *in vivo*. La [Figure 29](#) schématise le principe du crible.

- La séquence ADN correspondant au motif ARN d'intérêt (représentée en rouge) est insérée juste en aval de la séquence *rut* synthétique (notée $(CTT)_n$), et en amont d'un gène rapporteur. Le tout est placé sous le contrôle d'un promoteur constitutif ([Figure 29A](#)).
- Le motif ARN, seul, même très stable, ne permet de bloquer ni la fixation, ni la translocation de Rho sur l'ARN ([Steinmetz et al., 1990](#); [Figure 27B](#)).
- Par contre, si on produit, en trans, la protéine cible de ce motif ARN, le complexe formé par les deux partenaires va alors bloquer la fixation et/ou la translocation de Rho sur le transcrit ([Figure 29C](#)).
- L'ARN polymérase peut alors poursuivre la transcription du gène rapporteur situé en aval (par exemple un gène codant pour un ARNt; [Figure 29D](#)).

Dans notre système de sélection, l'activation de l'expression du gène rapporteur va donc témoigner de la fixation de la protéine d'intérêt sur son motif ARN cible.

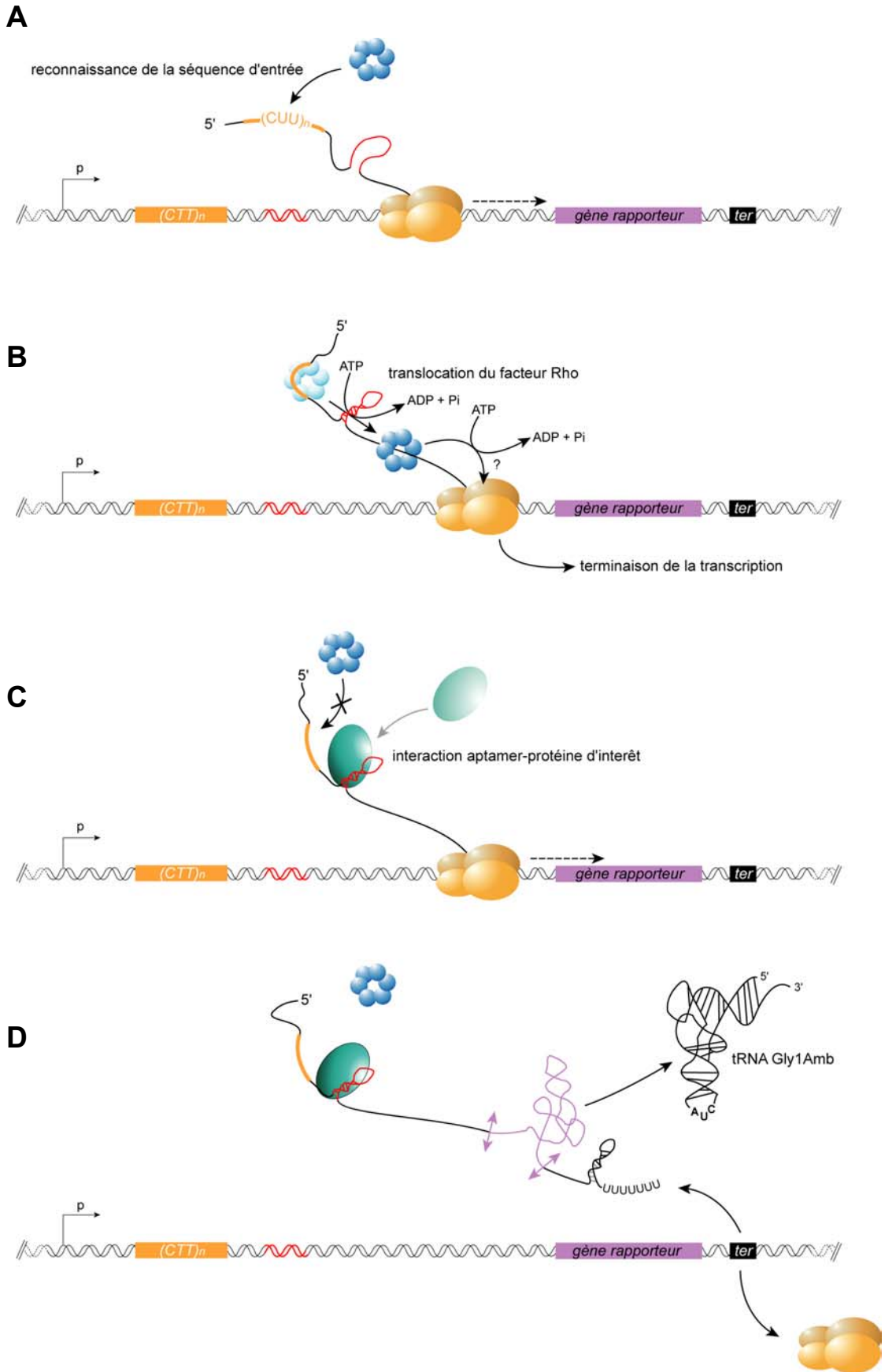


Figure 29 : Principe du crible génétique basé sur l'interférence d'un complexe ribonucléoprotéique vis-à-vis de la terminaison Rho-dépendante.

3. Design du système.

Le crible génétique par antiterminaison de la transcription repose sur la transformation de cellules d'*E. coli* par deux constructions plasmidiques compatibles (Figure 30).

- Un vecteur plasmidique rapporteur (test ou contrôle) portant le système de sélection génétique sous le contrôle d'un promoteur constitutif.
- Un vecteur de production de la protéine d'intérêt sous le contrôle d'un promoteur inductible.

3. 1. Plasmide rapporteur.

Le plasmide rapporteur utilisé est un dérivé du pACYC177, il porte un gène de résistance à la kanamycine, et le système de sélection génétique. Ce dernier se présente sous la forme d'un opéron artificiel, composé respectivement :

- De la séquence *rut* développée au laboratoire, dans l'orientation aboutissant à la répétition (rUrC)_n sur le transcrit. Chaque construction "test" est associée à une construction "contrôle" qui porte la séquence *rut* inversée ((rArG)_n sur le transcrit) et qui sert de référence dans cette étude (Figure 30).
- De la séquence codant le motif ARN d'intérêt, qui est insérée 18pb en aval de l'extrémité 3' de la séquence *rut* synthétique (positionnement optimal pour interférer avec la fixation de Rho). Cette insertion est réalisée entre les sites BglII et Acc65I.
- Du gène *cat*, qui apporte les sites potentiels de relargage du transcrit, nécessaires au bon fonctionnement du système (Guérin et al., 1998; Vieu et Rahmouni, 2004).
- D'un rapporteur génétique, permettant de sélectionner l'interaction ARN-protéine intervenant en aval.
- D'un terminateur intrinsèque fort, permettant l'indépendance transcriptionnelle de l'opéron.

Le tout est placé sous le contrôle d'un promoteur constitutif hisR^{Hpa} (Figuerola-Bossi et al., 1998). Chacun des éléments du système est entouré de sites de restriction uniques permettant de remplacer les différentes cassettes (séquence d'entrée, motif de reconnaissance, etc...).

3. 2. Plasmide protéique.

La séquence codante de la protéine d'intérêt est insérée sur un dérivé pBR322, sous le contrôle d'un promoteur dérivé du promoteur lac dont l'activité est régulée par la quantité de répresseur lac (*lacR*) produit dans une souche DH5 α Z1. L'activité du promoteur est alors sensible à la présence d'IPTG (Lutz et Bujard, 1997; Figure 30). Le vecteur protéique porte, de plus, un gène de résistance à l'ampicilline, permettant sa cotransformation avec le plasmide rapporteur.

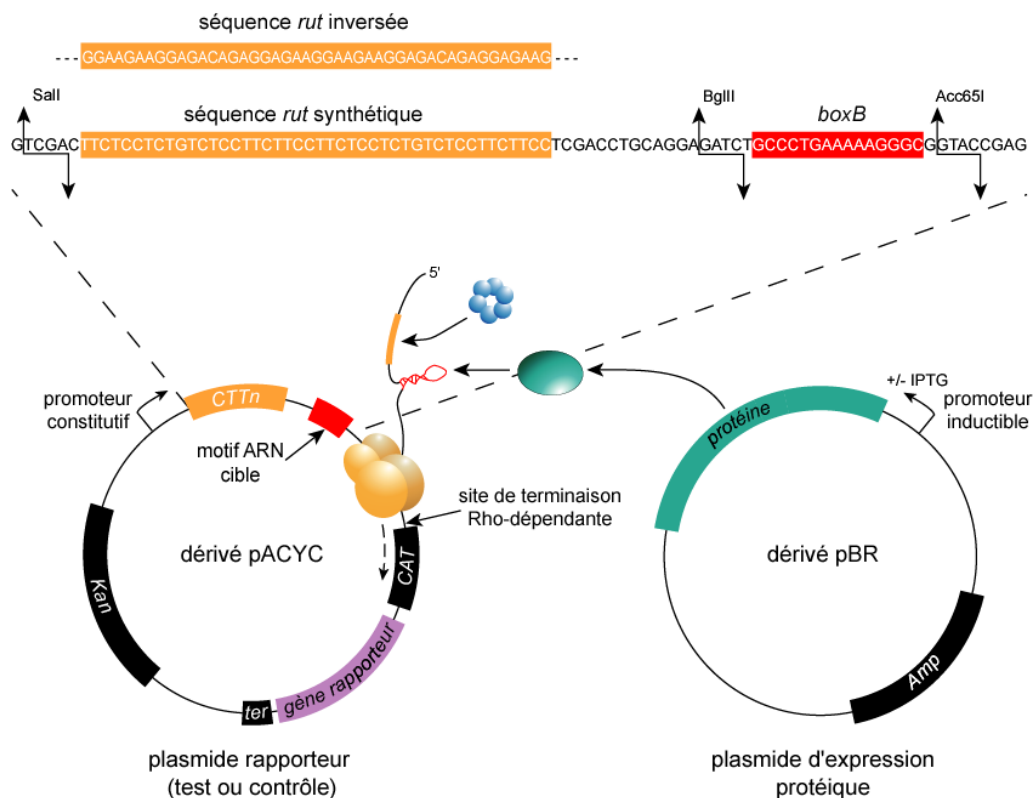


Figure 30 : Représentation schématique des plasmides : rapporteur (test ou contrôle) et d'expression protéique. Les séquences des éléments régulateurs : la séquence *rut* (orange) et la séquence codant le motif d'intérêt (ici la *boxB* en rouge) sont représentées plus haut.

4. Résultats préliminaires.

Les premières étapes de validation du système ont été réalisées en utilisant le gène *lacZ* comme gène rapporteur et le couple BOXB-N comme couple test. Les vecteurs pRR02

(pCT-boxB-lacZ) et pMB08 (pZ-λN; [Vieu et Rahmouni, 2004](#)) ont été transformés dans des cellules *E. coli* DH5αZ1 ([Lutz et Bujard, 1997](#)). L'activité β-galactosidase a été mesurée, en culture liquide en présence ou en absence de la protéine N, comme décrit dans la partie "Matériel et Méthodes". Les résultats sont présentés sur la [Figure 31](#).

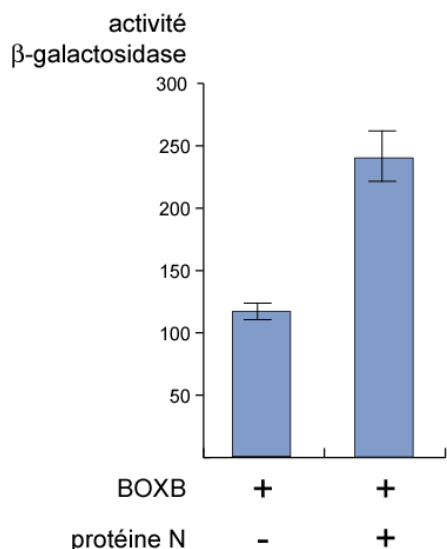


Figure 31 : Niveau d'expression de la β-galactosidase (en unités Miller) dans les cellules *E. coli* DH5αZ1 co-transformées par les plasmides pRR02 (pCT-boxB-lacZ) et pMB08 (pZ-λN), en présence (+) ou en absence (-) de la protéine N (+ ou - IPTG). Les barres d'erreur ont été déterminées à partir de trois séries de mesures distinctes.

On observe que l'expression de la protéine N, en présence du vecteur test capable de transcrire son motif de reconnaissance spécifique, induit l'activation de l'expression du gène *lacZ* d'un facteur 2,4 environ. Ces résultats montrent effectivement que le système fonctionne; conformément à l'hypothèse de départ, la fixation de la protéine d'intérêt sur son motif ARN cible interfère bien avec la terminaison Rho-dépendante, ce qui se traduit par l'activation de l'expression du gène rapporteur situé en aval. Cependant, ces résultats mettent à jours certaines imperfections du système :

- Tout d'abord, on peut observer que cette activation, d'un facteur 2,4 environ, est relativement faible, or, l'activation transcriptionnelle observée en northern blot, à l'aide d'une sonde nucléotidique complémentaire du 5' du transcrit LACZ, est beaucoup plus importante. Comme nous l'avons vu dans le travail précédent, le phénomène d'antiterminaison, provoqué par la formation d'un complexe ribonucléoprotéique à proximité du site RUT, n'est pas processif, or le gène *lacZ* est extrêmement long (plus de 3000pb). On peut supposer qu'au bout d'un moment le complexe ribonucléoprotéique se dissocie (spontanément ou sous l'action de Rho), la grande taille du *lacZ* donne alors l'opportunité à Rho de rattraper l'ARN polymérase

et de terminer la transcription avant la fin du gène. Ce phénomène aboutirait alors à des transcrits abortifs conduisant à la production de protéines tronquées inactives. Ceci expliquerait le décalage entre la transcription du *lacZ* observée en Northern et l'expression de la β -galactosidase mesurée en culture liquide.

- Par ailleurs, le bruit de fond, c'est-à-dire le niveau d'expression de base en absence de protéine, est trop élevé (environ 100 unités) pour pouvoir permettre une sélection via la coloration des colonies en milieu agrémenté de bleu X-Gal. Le gène reporter *lacZ* est, ici, présent directement sur un plasmide en mult copie dans la cellule, ce qui peut expliquer ce fort niveau d'expression basale.

- Enfin, le gène *lacZ* n'est pas adapté pour isoler des aptamers à partir d'une grande bibliothèque de séquences, il serait préférable d'utiliser un gène reporter permettant une sélection par mort/vie des cellules (résistance à un antibiotique ou survie en milieu minimum).

Pour toutes ces raisons le gène *lacZ* plasmidique a été remplacé.

5. Optimisation du système.

5. 1. Changement de système rapporteur.

Il s'est donc avéré nécessaire de choisir un nouveau système rapporteur pour notre test, le nouveau gène rapporteur devant présenter les caractéristiques suivantes :

- Il doit être court, pour être transcrit efficacement lorsqu'un complexe ARN-protéine est formé à proximité de la séquence *rut*.
- Il doit permettre de baisser le bruit de fond, de façon à favoriser une sélection fine des séquences intéressantes.
- Il doit être adapté à une sélection par mort/vie des cellules, permettant le crible d'un répertoire de séquences beaucoup plus large.
- Enfin le produit de ce gène doit être parfaitement corrélé à la quantité de transcrits présents dans la cellule, afin de permettre une plus grande amplitude de l'activité du gène reporter en présence ou en absence de protéine, de façon à pouvoir différencier des niveaux d'affinité différents.

J'ai alors choisi de remplacer le gène *lacZ* par un gène codant un ARN de transfert. Nous avons déjà utilisé cette stratégie dans le travail précédent, ici cependant, je n'ai pas utilisé l'ARNt^{Arg5} mais un ARNt suppresseur de mutations Amber (McClain et al., 1991). Ce système présente un grand nombre des avantages recherchés:

- Tout d'abord, le produit de ce gène est court et très stable. Cette stabilité est la conséquence du fort repliement de la structure secondaire de l'ARNt mature (pour revue voir : Mörl et Marchfelder, 2001; Figure 31).
- De plus, des travaux antérieurs ont montré que l'accumulation d'un ARNt dans la cellule est un bon marqueur de l'activité transcriptionnelle (Lopez et al., 1994).
- Une fois excisé du transcrit primaire, cet ARNt peut être détecté, par Northern blot, grâce à une sonde nucléotidique complémentaire (voir : Matériels et Méthodes).
- Son expression peut également être visualisée au travers de la traduction d'un gène reporter secondaire, chromosomique, porteur d'une mutation Amber (Sloan et Weisberg, 1993). Ce rapporteur secondaire étant chromosomique (donc en monocopie dans la cellule) le bruit de fond devrait chuter fortement.

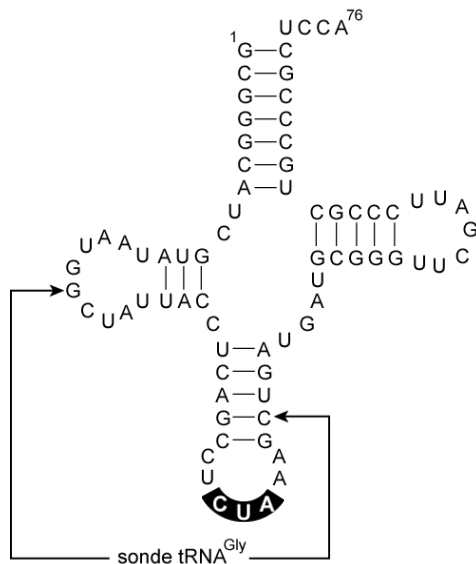


Figure 32 : Séquence nucléotidique de l'ARNt^{Gly1am} mature. Cette séquence diffère de l'ARNt^{Gly} sauvage d'*E. coli* par les substitutions U34C, C35U, et C36A dans l'anticodon (les positions sont encadrées). La sonde oligonucléotidique tRNA^{Gly} est complémentaire des positions 40-18, comme indiqué par les flèches (voir Matériel et Méthodes).

Pour ce travail nous avons utilisé l'ARNt^{Gly1am} (Figure 32). Cette ARNt est celui qui possède la plus forte efficacité de suppression parmi les suppresseurs de mutations Amber (30%; McClain et al., 1991). Ainsi, lorsqu'il est produit, cet ARNt reconnaît le codon UAG au

sein d'un ARNm muté et introduit une Glycine dans la séquence peptidique permettant ainsi d'achever la synthèse de la protéine. Celle-ci agit alors comme un marqueur secondaire de la transcription de l'ARNt. Ce système est très avantageux puisqu'il suffirait de transporter les plasmides dans une souche présentant une mutation Amber sur le marqueur génétique de notre choix pour pouvoir utiliser ce dernier comme gène de sélection. Il est également possible d'utiliser une souche portant plusieurs marqueurs génétiques mutés, qui seront activés par le même ARNtGly1am, afin d'éviter l'apparition de faux positifs résultant d'une réversion d'un des marqueurs.

Le gène *ARNtGly1am* a donc été cloné à la place du gène rapporteur précédent (*lacZ*), et le système de sélection a de nouveau été testé en utilisant ce nouveau marqueur et le couple BOXB-N.

5. 2. Résultats.

Le nouveau plasmide rapporteur (pCT-boxB-Gly1am) a été transformé dans une souche nommée KL335 (*lacZam*, *lacI*⁻; Birge et Low, 1974). Cette souche étant dépourvue de répresseur lac (*lacI*), l'expression de la protéine N est constitutive. L'expérience a été réalisée en comparant des cellules transformées par le vecteur rapporteur seul (monotransformants) avec des cellules transformées par le vecteur rapporteur et le vecteur protéique (cotransformants).

L'activation de la transcription de l'ARNtGly1am, par la formation du complexe BOXB-N, a été mesurée selon deux types d'analyses différentes :

- Tout d'abord l'ARNtGly1am a été dosé en Northern blot (Figure 33A), par hybridation avec une sonde radiomarquée complémentaire d'une portion de l'ARNt comprenant l'anticodon (voir : Figure 32).
- Ensuite l'expression de l'ARNtGly1am a été dosée au travers de l'activité β -galactosidase (en milieu liquide et en milieu solide Figure 33B) de la souche KL335 transformée par les différents vecteurs. En effet cette souche porte la mutation *lacZ608*, une mutation Amber présente sur le gène *lacZ*, l'activité β -galactosidase va donc témoigner de la quantité d'ARNt supresseur présent dans le milieu cellulaire. Les résultats obtenus sont présentés sur la Figure 33.

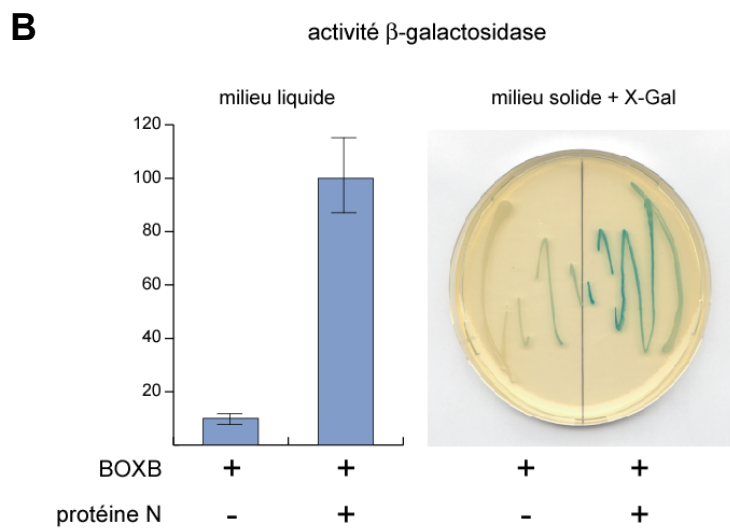
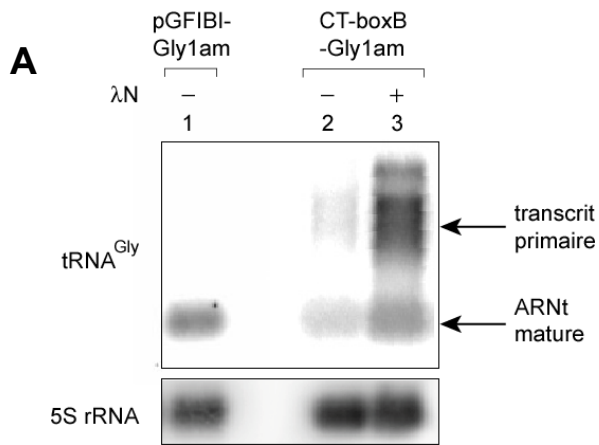


Figure 33 : Expression du tRNA^{Gly1am} en

présence ou en absence de la protéine N de λ .

(A) Analyse par Northern blot de l'expression de l' ARNt^{Gly1am}, en absence (ligne 2) ou en présence (ligne 3) de la protéine N. Les membranes ont tout d'abord été hybridées avec une sonde spécifique de l'ARNt^{Gly1am} (tRNA^{Gly}), cette spécificité est contrôlée grâce à la construction pGFIBI-Gly1am (ligne 1) qui exprime le tRNA^{Gly1am} de manière constitutive (Normanly et al., 1986).

Les membranes sont ensuite lavées et ré-hybridées avec une sonde spécifique de l'ARN ribosomal 5S (5S rRNA). (B) Analyse de l'expression de l'ARNt^{Gly1am} par l'intermédiaire du dosage de l'activité β -galactosidase en culture liquide d'une part et en milieu solide (LB-agar) complémenté en X-Gal d'autre part.

On peut constater que le nouveau système reporter a permis de répondre à la majorité des questions posées précédemment, en effet l'amplitude de l'activation a fortement augmenté (passant d'un facteur 2 à un facteur 8-10) et le bruit de fond a considérablement diminué (une dizaine d'unités Miller contre une centaine d'unités environ lorsque le gène *lacZ* était le reporter plasmidique; voir Figure 31).

Cependant, comme on peut le voir sur la Figure 31A, l'excision et la maturation de l'ARNt est incomplète. Or, pour être fonctionnel, l'ARNt^{Gly1am} doit être parfaitement excisé du transcrit primaire, aussi bien en 5' qu'en 3'. Cela signifie que les résultats observés en expression β -galactosidase (Figure 31B) ne sont le reflet que de la faible minorité des ARNt

correctement maturés. L'efficacité du système peut donc encore être augmentée en améliorant l'excision et la maturation de l'ARNtGly1am.

La coupure d'un ARNt en 5' repose sur l'action de la RNaseP, une enzyme présente chez tous les organismes (pour revue : [Mörl et Marchfelder, 2001](#)). Le mécanisme aboutissant à la maturation de l'extrémité 3', qui est indispensable à l'aminoylation, est beaucoup moins bien connu. Deux enzymes, la RNaseE ([Roy et Apirion, 1983 a et b](#)) et la RNaseIII ([Apirion et Miczak, 1993](#)), semblent être impliquées dans ce processus ([Li et Deutscher, 1994 et 1996](#)). L'action de ces enzymes peut être gênée par la présence de protéines fixées au niveau de l'ARN précurseur. Or pour des commodités de clonage, le gène *tRNAGly1am* a été cloné au milieu du gène *cat* (au site EcoRI). L'excision est probablement gênée par la présence des ribosomes traduisant l'ARNm CAT. Il est également possible que les séquences adjacentes du site d'insertion du gène *tRNAGly1am* interfèrent avec la maturation de l'ARNt. Quoiqu'il en soit, il semble que ce positionnement soit incompatible avec la maturation correcte de l'ARNt suppresseur et donc avec le bon fonctionnement du système.

5. 3. Amélioration de l'excision de l'ARNt.

Pour régler ce problème, le gène *tRNAGly1am* a été déplacé une centaine de paires de bases, environ, en aval du gène *cat*, et juste en amont d'un terminateur intrinsèque fort. Le gène *cat*, quant à lui, a été délété de ses Shine Dalgarno et de son ATG d'initiation afin d'éviter toute interférence entre la traduction de son ARNm et la maturation de l'ARNt.

5. 4. Nouveaux couples ARN-protéine testés.

Jusqu'à présent nous utilisons le couple BOXB-N pour tester et valider notre système. L'utilisation de ce couple est ambiguë dans la mesure où N peut provoquer l'antiterminaison par d'autres mécanismes que l'interférence avec la fixation directe de Rho sur le transcrit (voir chapitre III). Même si la *boxA* est absente de notre système, on ne peut, en effet, pas écarter la possibilité que les résultats observés soient le fruit d'un autre mécanisme utilisé par N pour bloquer la terminaison Rho-dépendante. De plus la validité d'un tel système nécessite l'utilisation de plusieurs couples ARN-protéine différents.

Nous avons donc testé notre crible sur deux nouveaux couples ARN-protéine bien caractérisés :

- Le couple formé par la protéine d'épissage humaine U1A et son motif de reconnaissance : U1hpII (pour "U1 small nuclear RNA hairpin II"). U1A reconnaît son motif ARN grâce à son extrémité N-terminale (acides aminés 1-102) contenant deux motifs RNP (pour "Ribonucleoprotein") aussi appelés RRM (pour "RNA Recognition Motif") (Scherly et al., 1990; Oubridge et al., 1994; Figure 34A).
- Le couple formé par la coat protein du phage MS2 et son motif de reconnaissance naturel : l'opérateur traductionnel MS2 (Beckett et Uhlenbeck, 1988; Figure 34B). La coat protein se fixe à l'ARN, sous la forme d'un dimère, grâce à deux feuillets β (Peabody, 1993). Elle agit comme la principale protéine structurale de la capside virale mais également comme répresseur de la traduction de la réplicase virale.

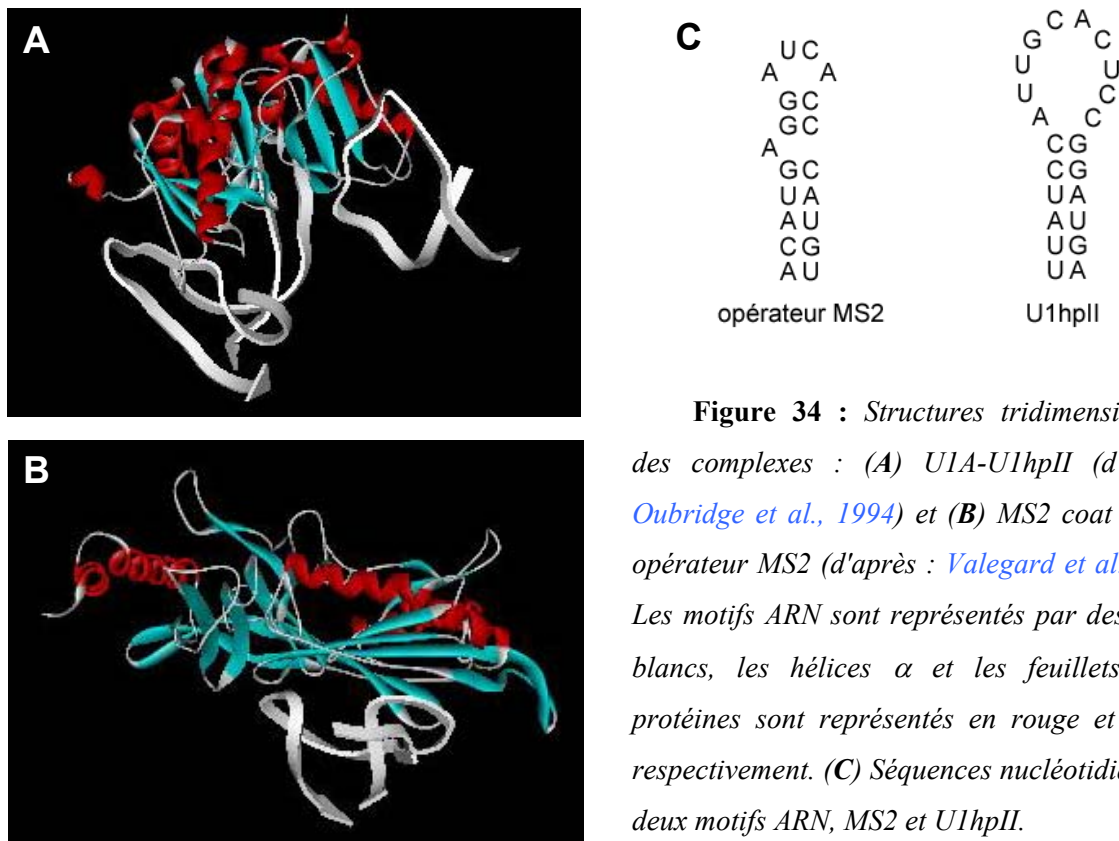


Figure 34 : Structures tridimensionnelles des complexes : (A) U1A-U1hpII (d'après : Oubridge et al., 1994) et (B) MS2 coat protein-opérateur MS2 (d'après : Valegard et al., 1997). Les motifs ARN sont représentés par des rubans blancs, les hélices α et les feuillets β des protéines sont représentés en rouge et en vert respectivement. (C) Séquences nucléotidiques des deux motifs ARN, MS2 et U1hpII.

Les séquences codant la protéine MS2 entière d'une part et les 102 acides aminés N-terminaux de la protéine U1A d'autre part ont donc été clonées sur le vecteur protéique, sous le contrôle d'un promoteur inductible à l'IPTG (voir : [Figure 30](#)).

5. 5. Optimisation du site d'insertion du motif d'intérêt.

Les séquences codant pour les deux motifs ARN d'intérêt ont, elles, été clonées entre les sites BglIII et Acc65I du plasmide rapporteur (test ou contrôle). La séquence d'intérêt (en rouge sur la [Figure 35](#)) a chaque fois été flanquée, en 5' et en 3', d'un doublet GC (en noire sur la [Figure 35](#)), afin de stabiliser la formation du motif d'intérêt et d'éviter que celui-ci ne soit, une fois dans l'ARN, séquestré dans la formation de structures alternatives mettant en jeu les séquences adjacentes.



Figure 35 : Séquences des motifs *U1hpII* et *MS2* insérées entre les sites *BglIII* et *Acc65I* des plasmides test et contrôle. Les plasmides correspondants sont nommés : *pCT-MS2-tRNAGly*, *pAG-MS2-tRNAGly*, *pCT-U1hpII-tRNAGly* et *pAG-U1hpII-tRNAGly*.

6. Résultats obtenus avec les nouvelles constructions.

Les nouvelles constructions sont transformées avec leurs plasmides protéiques respectifs dans des cellules DH5 α Z1 ([Lutz et Bujard, 1997](#)). Le bon fonctionnement du système est vérifié en analysant, par Northern blot, la quantité d'ARNtGly1am produit dans les différents transformants ([Figure 36](#)).

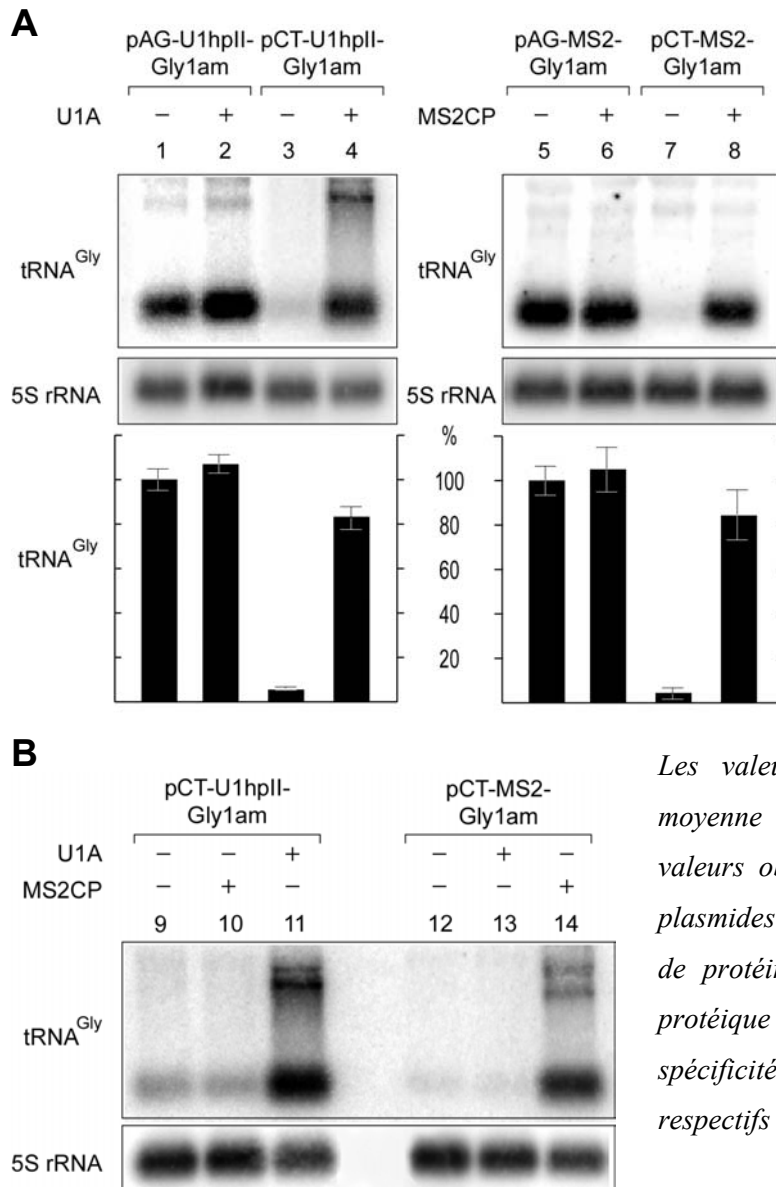


Figure 36 : Activation de l'expression de l'ARNtGly1am par la fixation des protéines U1A et MS2CP sur leurs motifs ARN respectifs. (A) Analyse par Northern blot de l'accumulation de l'ARNtGly1am, en absence (-) ou en présence (+) des protéines U1A ou MS2CP. L'hybridation est réalisée comme décrit dans la légende de la Figure 33. Le signal est chaque fois normalisé par le signal obtenu avec l'ARN 5S.

Les valeurs du graphique représentent la moyenne de trois extractions différentes. Les valeurs obtenues sont comparées à celle des plasmides contrôle (pAG) obtenues en absence de protéine. (B) Les plasmides d'expression protéique sont intervertis pour vérifier la spécificité des protéines pour leurs motifs ARN respectifs dans le processus de sélection

La Figure 36 montre que le système fonctionne parfaitement. En effet, la production de la protéine (U1A ou MS2CP) active l'expression de l'ARNtGly1am (Figure 36A, comparer les lignes 3 et 4 et les lignes 7 et 8). Cette activation par une protéine nécessite la transcription de son motif ARN spécifique entre la séquence d'entrée de Rho et le gène rapporteur. En effet aucune activation de la synthèse de tRNA n'est observée dans les souches où les couples ont été intervertis (Figure 36B, lignes 10 et 13). Enfin cette activation est bien due à l'interférence du couple ARN-protéine vis-à-vis de la fixation de Rho au transcrit. En effet la production de la protéine dans les souches portant un vecteur contrôle, ne provoque quasiment pas d'augmentation de l'expression du tRNA (Figure 36A, comparer les lignes 1 et 2 et les lignes

5 et 6). On observe tout de même une légère augmentation du signal (environ 10%) en présence de protéine. Probablement un léger effet de stabilisation de l'ARNm par la fixation de la protéine. Si cet effet est également observé en complémentation il faudra en tenir compte pour normaliser l'activation obtenue avec les constructions test.

Au-delà de la validation du système par deux nouveaux couples ARN-protéine, on constate que l'excision de l'ARNtGly1am est ici beaucoup plus complète. Cette élévation du taux d'ARNtGly1am mature devrait accentuer le signal observé en complémentation (Figure 33) permettant encore d'augmenter l'amplitude du signal. Ceci devrait autoriser l'utilisation du système pour analyser des gammes assez larges d'interactions ARN-protéine.

DISCUSSION
PERSPECTIVE

Le travail décrit dans cette thèse a été entrepris afin d'appréhender les bases des mécanismes de terminaison/antiterminaison intervenant au niveau du terminateur Rho-dépendant tR1 du phage λ . Les résultats obtenus ont ensuite permis d'initier la mise au point d'un crible génétique, basé sur la terminaison Rho-dépendante, permettant de sélectionner et d'analyser des interactions entre ARNs et protéines chez la bactérie *E. coli*.

1. Un modèle complémentaire de l'antiterminaison au niveau de tR1.

Le terminateur tR1 du phage λ est très certainement le terminateur Rho-dépendant le mieux caractérisé à ce jour. Cela tient en partie à des raisons historiques puisque c'est l'utilisation de ce terminateur qui a permis de mettre au jour Rho et le phénomène d'antiterminaison. Ce terminateur possède la particularité de porter, au niveau de ses séquences en 5', à la fois les éléments nécessaires à la terminaison et ceux nécessaires à l'antiterminaison. Parmi ces éléments, la séquence *boxB* décrite auparavant comme inutile pour la terminaison, est en revanche indispensable à l'antiterminaison. En effet, cette séquence aboutit sur l'ARN à la formation d'une structure en tige boucle correspondant au motif de reconnaissance de la protéine N du phage λ . La fixation de la protéine N sur le motif BOXB, puis l'association de facteurs cellulaires (facteurs Nus), conduisent à une modification de l'ARN polymérase la rendant résistante (de manière processive) aux terminateurs situés en aval de tR1.

Certains travaux réalisés chez le phage λ ([Mason et al., 1992](#)), ou chez le phage H-19B qui lui est apparenté ([Neely et Friedman, 2000](#)), décrivent l'existence de mécanismes d'antiterminaison également dépendants de N mais non processifs et faisant intervenir un nombre réduit de facteurs Nus. Ceci laisse entendre qu'un mécanisme différent de ce qui est écrit dans la littérature pourrait intervenir au niveau de tR1.

Au cours de ce travail nous avons mis en évidence le double rôle du motif BOXB dans les deux mécanismes de régulation de la transcription qui interviennent au niveau de tR1, la terminaison Rho-dépendante d'une part, et le phénomène d'antiterminaison N-dépendante d'autre part.

- Nous avons d'abord montré qu'en tant qu'isolateur structural, BOXB favorise la terminaison en permettant la formation d'un site RUT unique à partir de RUTA et RUTB. La réorganisation d'une séquence régulatrice par des motifs ARN est un phénomène relativement fréquent chez la bactérie. L'atténuation en est un exemple, tout comme le mécanisme décrit récemment par Neely et Friedman : selon cette étude, des réarrangements structuraux, intervenant entre la BOXA et la BOXB des sites NUT du phage H-19B, seraient responsables de l'antiterminaison par la protéine N intervenant au niveau de ces sites (Neely et Friedman, 2000). Dans cette étude, comme dans la notre, la délétion de ces séquences n'affecte pas le mécanisme, confirmant que ce n'est pas la composition nucléotidique des structures mais les structures elle-même et les réarrangements qu'elles induisent qui sont importants.
- Dans un second temps, nous avons montré que le couple BOXB/N favorise l'antiterminaison au niveau de tR1 (qui contient le site *nutR*) en formant un obstacle stérique à la fixation de Rho sur l'ARN.

Cette dernière observation est confortée par certains travaux ayant montré que : un ribosome bloqué à proximité du site RUT de l'opéron *tna* (Konan et Yanofsky, 2000; Gong et Yanofsky, 2002) ou encore un mutant de Nun en se fixant à la BOXB de NUTR (Kim et al., 2003) étaient capables d'inhiber la fixation de Rho sur le transcrit et de provoquer une antiterminaison non processive. Tous ces résultats témoignent de l'existence de mécanismes d'antiterminaison alternatifs.

Il est nécessaire toutefois de distinguer le mécanisme décrit dans notre étude de celui permettant à l'ARN polymérase de résister, de façon processive, à tous les types de terminateurs localisés en aval du site *nutR*. On peut cependant supposer que ces deux mécanismes sont complémentaires et que le phénomène décrit dans cet article constitue une étape préliminaire au mécanisme processif :

- En bloquant la fixation de Rho, le complexe BOXB/N donne, tout d'abord, un avantage cinétique à l'ARN polymérase. Ce gain de temps lui permet alors de franchir les premiers sites de terminaison (ceux de tR1 ou de notre système, Figure 37A).
- En effet, chez λ , contrairement à ce qui se produit dans les opérons *rrn*, les facteurs Nus nécessitent la fixation préliminaire de N sur BOXB avant de modifier l'ARN polymérase et de lui permettre de franchir les terminateurs localisés plus loin en

aval. Le mécanisme d'interférence décrit ici donnerait donc le temps nécessaire à la formation complète du complexe d'antiterminaison N-dépendant (Figure 37B).

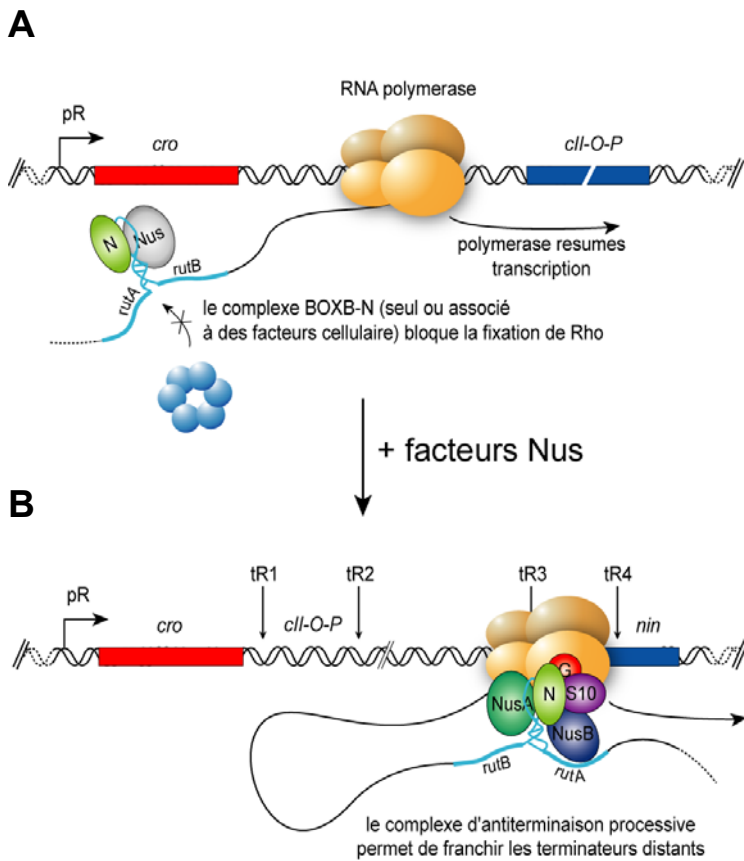


Figure 37 : *Modèle du mécanisme de terminaison/antiterminaison au niveau du terminateur λ tR1. (A) La fixation de la protéine N (seule ou associée à des facteurs Nus) au niveau de BOXB interfère avec l'accès de Rho à l'ARN. (B) Des facteurs Nus supplémentaires s'associent au couple BOXB-N et à l'ARN polymérase pour former un complexe processif d'antiterminaison capable de franchir des terminateurs plus éloignés*

D'autres travaux seront nécessaires pour déterminer si des liens existent entre les deux mécanismes. En particulier il reste à déterminer si certains facteurs Nus participent au mécanisme non processif intervenant au niveau de tR1 (pour augmenter la taille du complexe d'interférence).

Indirectement, ce travail a montré qu'il était possible de détecter, *in vivo*, la formation d'un complexe ARN-protéine hétérologue lorsque celle-ci intervient à proximité d'un site RUT. Cette observation nous a conduit à la mise au point d'un test, utilisant la terminaison Rho-dépendante et l'un de ses mécanismes de régulation, permettant de détecter et d'analyser des interactions ARN-protéines chez *E. coli*.

2. Un crible génétique pour étudier les interactions ARN-protéine *in vivo*.

La majeure partie de mon travail a consisté à la mise au point, l'optimisation, et la validation de notre système de sélection :

Le terminateur tR1 a tout d'abord été remplacé par une séquence *rut* synthétique, développée au laboratoire, qui induit une terminaison quasi-totale *in vivo* (Guerin et al., 1998). Contrairement à tR1, l'efficacité de ce terminateur ne repose pas sur la structuration correcte d'un motif ARN, elle est également indépendante de l'environnement génétique de la séquence synthétique. En effet nous avons pu constater au cours de ce travail que l'efficacité de la terminaison, en absence de la protéine d'intérêt, est peu ou pas modifiée par la présence du motif ARN d'intérêt en aval de la séquence synthétique. Cela permet un niveau transcriptionnel de base (bruit de fond) constant quel que soit le motif d'intérêt.

Le plasmide rapporteur a été élaboré pour permettre le remplacement rapide des différentes séquences d'intérêt. En particulier le site d'insertion de la séquence codant le motif ARN cible a été placé entre deux sites de restriction BglIII et Acc65I, et flanqué, de part et d'autre, d'un doublet GC. Des séquences aléatoires peuvent alors être amplifiées, par PCR, entre des amorces contenant les deux sites de restriction et les doublets GC. Après digestion par les enzymes BglIII/Acc65I, ces séquences peuvent aisément être clonées dans le vecteur rapporteur digéré par les mêmes enzymes. Les doublets GC donnent naissance sur le transcrit à un isolateur structural permettant de favoriser le repliement du motif ARN et d'augmenter son accessibilité. Cet isolateur devrait limiter le risque de perte de motifs, intéressants mais trop instables, au cours du crible.

Le système de sélection génétique a été optimisé en utilisant le gène *tRNAGlyIam* comme rapporteur. Le principal avantage est qu'une fois transcrit et correctement excisé cet ARNt suppresseur peut activer l'expression d'un marqueur génétique secondaire porteur d'une mutation Amber (par exemple un gène permettant de sélectionner les colonies en fonction de leur survie sur un milieu particulier). Le plus de ce système est, qu'une fois isolés, les différents clones pourront être analysés sans avoir à re-cloner les séquences d'intérêt dans d'autres plasmides portant un marqueur génétique adapté à un dosage quantitatif de l'interaction. Il suffira juste de changer de souche bactérienne réceptrice, ce qui représente un gain de temps appréciable par rapport aux méthodes de sélection actuelles.

Ainsi, si l'on souhaite réaliser un crible à grande échelle, on effectuera une première série de sélections dans une souche portant la mutation Amber sur un gène de résistance à un antibiotique par exemple. On sélectionnera alors les clones d'intérêt en fonction de leur survie sur un milieu contenant l'antibiotique en question. Une sélection semi-quantitative pourra même être effectuée en modulant la concentration en antibiotique dans le milieu. Si ensuite on souhaite doser la force des interactions isolées précédemment, il nous suffira d'extraire les plasmides à partir des clones résistants à l'antibiotique et de les transformer dans une nouvelle souche portant la mutation Amber sur un gène dont l'activité est facilement quantifiable, par exemple le gène *lacZ*. L'interaction décelée pourra même être quantifiée précisément, *in vivo*, en comparant l'activité du rapporteur chez le clone isolé à celle d'une gamme de complexe ARN-protéine d'affinité connue, on pourra alors mesurer les constantes d'affinité du motif ARN sélectionné pour sa protéine d'intérêt.

Les deux étapes de sélection et d'analyse quantitative pourraient même être réalisées dans une seule et même souche qui porterait une mutations Amber sur au moins deux marqueurs génétiques : l'un permettant de détecter une interaction par mort/vie des cellules, l'autre permettant de quantifier la force de l'interaction au sein des complexes isolés. La construction d'une telle souche constitue l'une des perspectives de cette étude.

La validation définitive du système dans ce type de souche va représenter la suite immédiate de ce travail. Pour cela il sera nécessaire de modifier le système d'induction des plasmides protéiques. En effet le système d'induction de l'expression de protéines d'intérêt repose actuellement sur le système opérateur/répresseur de l'opéron lactose. Le système n'est pas compatible avec toutes les souches bactériennes (en particulier les souches *lacI⁻*) et avec certains rapporteurs secondaires (le gène *lacZ* en particulier). Afin d'assurer l'adaptabilité du système, le principe d'induction doit être modifié. Le remplacement du système d'induction par une induction à l'Arabinose (Guzman et al., 1995) est actuellement en cours au laboratoire.

Le fonctionnement de notre crible génétique a jusqu'à présent été validé sur les couples MS2-MS2CP et U1hpII-U1A en utilisant l'ARNtGly1am comme unique rapporteur transcriptionnel. Les résultats concordent avec notre hypothèse de départ puisque la production de la protéine active l'expression de l'ARNtGly1am dans la cellule, et que cette expression est corrélée à la présence de son motif ARN spécifique au sein du transcrit rapporteur.

Une fois totalement optimisé et validé notre système pourra être utilisé dans différentes stratégies :

- Tout d'abord pour sélectionner, à partir d'une bibliothèque de séquences nucléotidiques aléatoires, les motifs ARN (aptamers) reconnus spécifiquement par une protéine d'intérêt (par exemple une protéine virale ou oncogénique).
- Pour isoler à partir d'une bibliothèque d'ADNc, ceux codant pour un peptide ou une protéine fixant spécifiquement un motif ARN d'intérêt (par exemple un motif ARN rétroviral comme Tar ou RRE).
- Pour déterminer au sein d'un complexe ARN-protéine connu les éléments, nucléotidiques ou peptidiques, nécessaires à l'interaction. Pour cela nous pourrions utiliser des méthodes de mutagenèse aléatoire par PCR (Cadwell et Joyce, 1992), afin d'introduire chez l'un des deux partenaires des mutations ponctuelles et d'observer l'effet de ces mutations sur l'interaction.
- Enfin le test pourra servir à la recherche de drogues (molécules naturelles ou synthétiques) capables d'interférer *in vivo* avec la formation et la fonction de complexes ribonucléoprotéiques d'intérêts.

Finalement, sachant que Rho est capable de terminer la transcription par l'ARN polymérase II de *Saccharomyces cerevisiae*, on pense que l'ARN polymérase II serait sensible à un mécanisme proche de la terminaison Rho-dépendante d'*E. coli*. La description de ce mécanisme pourra peut être permettre de transposer notre crible génétique des bactéries vers les cellules eucaryotes.

3. Autres utilisations du système génétique.

Au-delà de ces différentes utilisations, les résultats obtenus au cours de mon travail de thèse suggèrent que la séquence *rut* synthétique représente un outil extrêmement intéressant pour l'étude du fonctionnement de Rho *in vivo*. Elle est d'ailleurs déjà utilisée, au laboratoire, pour des études de son fonctionnement *in vitro* (Walmacq et al., à paraître).

Dans notre système, on ignore encore si le complexe ribonucléoprotéique bloque la fixation ou bien la translocation de Rho le long du transcrit. Au vu de la proximité du motif par rapport à la séquence d'entrée, il semblerait que ce soit plutôt la fixation de Rho qui soit

ici en jeu. Toutefois l'effet d'un complexe ribonucléoprotéique sur la translocation de Rho le long de l'ARN est actuellement inconnu et pourrait représenter une étude intéressante. En effet il existe des exemples d'hélicases capables de dissocier un complexe ARN-protéine selon un mécanisme dépendant de l'hydrolyse d'ATP (Jankowsky et al, 2001, Fairman et al., 2004). Si un tel complexe est capable de bloquer la translocation de Rho le long du transcrit, cette particularité pourrait être utilisée afin d'étudier, *in vivo*, les paramètres cinétiques de la translocation du facteur de terminaison, en plaçant par exemple cet obstacle à différentes positions du transcrit par rapport au site RUT. Dans le cas contraire, l'étude de la dissociation d'un complexe ribonucléoprotéique par l'action hélicase (ou plutôt translocase) ATP dépendante de Rho, pourrait donner des informations quant au mécanisme permettant à l'hexamère de dissocier un complexe multiprotéique tel que l'ARN polymérase.

D'autres questions importantes restent posées quant au fonctionnement du facteur Rho et aux mécanismes par lesquels son action est régulée :

- Le rôle des sites de pauses reste ambiguë : constituent-ils une composante uniquement cinétique (permettant à Rho de rattraper l'ARN polymérase) ou agissent-ils également de façon structurale, en provoquant la déstabilisation du complexe d'élongation?
- Comment le facteur NusG agit-il en faveur de la terminaison tout en étant nécessaire à l'antiterminaison?
- Enfin, par quel mécanisme Rho réussit-il, en fin de réaction, à dissocier le complexe ternaire d'élongation?

Afin d'apporter certaines réponses à ces interrogations, je propose un modèle dérivé du système génétique que nous avons élaboré au cours de ma thèse :

- Sur un plasmide, on place un promoteur inductible (par exemple un promoteur inductible à l'arabinose), puis la séquence *rut* synthétique (dans l'orientation CT), suivie d'une ou plusieurs séquences opératrices (opérateur lac) et enfin un gène rapporteur (Figure 38). Le vecteur portera également le répresseur lacI^q exprimé constitutivement.
- En présence d'arabinose, l'ARN polymérase qui transcrit cet opéron artificiel se retrouve bloquée par la présence d'un ou plusieurs complexes opérateur-répresseur sur son parcours. Ce genre de stratégie a déjà été utilisée au laboratoire pour étudier

l'effet de signaux régulateurs sur l'ARN polymérase (Toulmé et al., 2000; Mosrin Huaman et al., 2004; Epshtein et al., 2003).

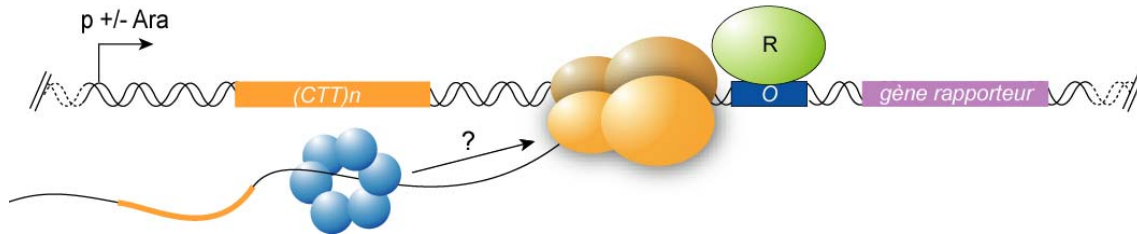


Figure 38 : *Modèle d'étude du fonctionnement de Rho in vivo. La même construction portant la séquence rut inversée sera utilisé pour définir la proportion de complexe d'élongation se dissociant spontanément (sans l'action de Rho).*

- Dans ce cas de figure, l'obstacle protéique va mimer l'effet cinétique d'un site de pause et permettre au facteur Rho de rattraper l'ARN polymérase. Ces complexes seront ensuite isolés en bloquant l'activité du promoteur (en remplaçant l'arabinose par du glucose). Il suffira alors de briser le complexe opérateur-répresseur par ajout d'IPTG, pour mesurer, au travers de l'expression du gène rapporteur, l'effet de Rho sur une ARN polymérase stoppée par un obstacle stérique.

Si dans ces conditions Rho n'a aucun effet sur l'ARN polymérase, cela signifie que les sites de pauses favorisent la terminaison autrement qu'en réduisant la vitesse du complexe d'élongation. L'ajout d'un site de pause juste en amont de la séquence opératrice permettra alors de vérifier si la déstabilisation du complexe ternaire d'élongation est nécessaire pour favoriser la terminaison Rho-dépendante. En fonction des différents résultats obtenus, l'utilisation de souches mutantes en certains facteurs Nus permettra de donner des informations sur l'effet des facteurs NusA et NusG en faveur ou non des différents mécanismes.

Ce type de stratégie pourra également être utilisée pour étudier l'antiterminaison processive dépendante de N. En effet la plupart des hypothèses supposent que N et les facteurs Nus protègent l'ARN polymérase de la terminaison en augmentant sa processivité. Il serait intéressant d'observer le comportement de Rho face à une ARN polymérase, modifiée

par N et les facteurs Nus, mais stoppée par la présence du complexe opérateur-répresseur sur son parcours (Figure 39).

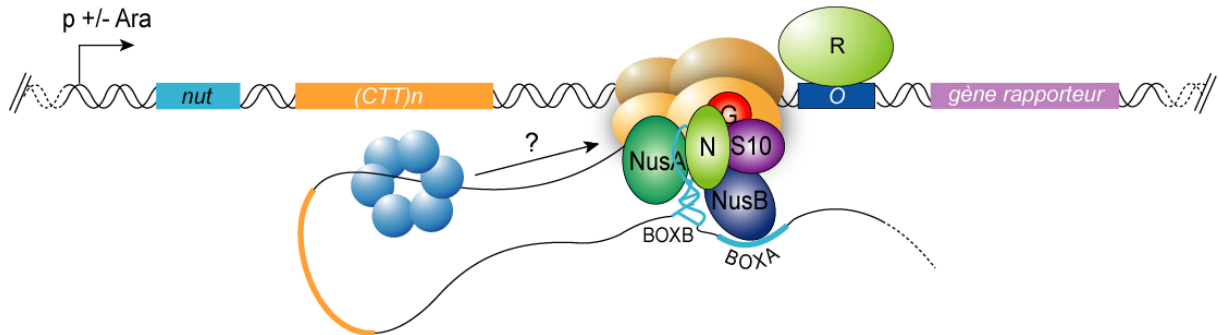


Figure 39 : *Modèle d'étude de l'antitermination processive, in vivo.*

Les résultats permettrons peut être de savoir si le complexe d'antitermination inhibe la terminaison en empêchant Rho de rattraper l'ARN polymérase (effet cinétique), en stabilisant l'ARN polymérase (effet structural), ou encore en bloquant l'accès de Rho au facteur NusG qui lui serait indispensable pour induire la dissociation du complexe d'élongation.

Quoiqu'il en soit l'ensemble des informations recueillies dans ce manuscrit montre que, 35 ans après sa découverte, le facteur Rho demeure un outils fondamental dans la compréhension des mécanismes impliqués dans la régulation de l'expression des gènes, et un sujet d'étude toujours aussi vaste et intéressant.

MATERIEL ET
METHODES

MATERIELS ET METHODES.

1. Souches.

DH5 α Z1 : (ϕ 80*dlacZ* Δ M15) *endA1*, *recA1*, *hsdR17*, *supE4*, *4thi-1*, *gyrA*, *relA1*, Δ (*lacZYA-argF*)U169, (λ *att*, *lacI^q*, *tetR*, *Sp^r*); (Lutz et Bujard, 1997)

KL335 : *lacI22*, *lacZ608*(Am), λ^- ; (Birge et Low, 1974).

2. Enzymes, produits chimiques et oligonucléotides.

Les enzymes de restriction, la polynucléotide kinase du phage T4, et le fragment Klenow de l'ADN polymérase I ont été fourni par New England Biolabs. L'ADN ligase du phage T4 a été obtenu auprès d'Invitrogen. La majorité des produits chimiques, incluant les antibiotiques, ont été fourni par Sigma. Les dNTPs non marqués ont été obtenus auprès de Boehringer Mannheim, tandis que le [γ ³²P]dATP, les [α ³³P]ddNTPs, et la thermostable ADN polymérase ont été fourni par Amersham Biosciences. Les oligonucléotides utilisés au cours de ce travail ont tous été synthétisés chez MWG-Biotech.

3. Plasmides.

3. 1. Plasmides rapporteurs.

Les plasmides rapporteurs (tests et contrôles) sont basés sur un plasmide parent p7745pLboxB qui est dérivé du pACYC177 (origine de répllication p15A). Ce plasmide porte un gène de résistance à la kanamycine, il contient le promoteur constitutif *hisR*^{Hpa} de *Salmonella typhimurium* (Figueroa Bossi et al., 1998), suivi par la séquence *rut* synthétique (Guerin et al., 1998; Figure 30), la séquence *boxB* clonée entre les sites BglII et Acc65I, et le gène *cat* (chloramphenicol acetyl transferase).

- **pRR02 (pCT-boxB-lacZ; Figure 40A)** : Ce vecteur est obtenu en clonant le gène *lacZ* (fragment Sall/XbaI rempli de pOM10 : Espeli et al., 2001) dans le site EcoRI rempli de p7745pLboxB.

- **pEV06 (pCT-boxB-Gly1am; Figure 40B)** : le gène *tRNAGly1am* est amplifié à partir du plasmide pGFIBI-Gly1am (Sloan et Weisberg, 1993) grâce aux amorces RNAup : 5' CTA TGA GAA TTC GCG GGC 3' et RNAlow : 5' CCA TGG CTG CAG TGG 3'. Le produit PCR est digéré par EcoRI/NcoI (sites contenus dans les amorces) et cloné dans les mêmes sites de p7745pLboxB.

- **pEV38 (pCT-MS2-tRNAGly1am; Figure 40C)** : cette construction a été obtenue en plusieurs étapes :

- Le gène *tRNAGly1am* à tout d'abord été amplifié à partir du plasmide pGFIBI-Gly1am (Sloan et Weisberg, 1993) entre les amorces tRNAup2 : 5' CTA TGA TTA ATG AAT TCG CGG GC 3' et tRNAlow2 : 5' CAA TGA TTA ATC TGC AGT GGA GC 3'; puis digéré par AseI et cloné dans le site AseI du plasmide pOM10 (Espeli et al., 2001) pour donner le vecteur pEV30.

- Le fragment NcoI/PstI de pEV30 (extrémité 3' du *cat*, *ARNtGly1am* et terminateur fort du gène 32 du phage T4) est introduit dans les mêmes sites de pRUTWT (Vieu et Rahmouni, 2004), pour donner le vecteur pEV33 qui porte le gène *cat* entier suivi du *tRNAGly1am*.

- Le vecteur pEV38 est finalement obtenu par échange de fragments HindIII-Acc65I entre la construction pEV33 (récepteur) et p7745pLboxB (donneur), puis par l'insertion de la séquence de l'opérateur MS2 entre les sites BglII et Acc65I en utilisant le duplex d'oligonucléotides : MS2GC1 : 5' GAT CTG CAC ATG AGG ATC ACC CAT GTG CG 3' et MS2GC2 5' : GTA CCG CAC ATG GGT GAT CCT CAT GTG CA 3'.

- **pAA01 (pAG-MS2-tRNAGly1am)** : ce plasmide est identique à pEV38 excepté pour la séquence d'entrée qui est inversée.

- **pMG26 (pCT-U1hpII-tRNAGly1am; Figure 40D)** : ce plasmide est obtenu à partir du plasmide pEV38, la séquence codant l'opérateur MS2 a été remplacée par la séquence codant le motif U1hpII en utilisant le duplex d'oligonucléotides : U1hpIII : 5' GAT CTG

CTT ATC CAT TGC ACT CCG GAT GAG CG 3' et U1hpII2 : 5'GTA CCG CTC ATC CGG AGT GCA ATG GAT AAG CA 3'.

- **pMG27 (pAG-U1hpII-tRNAGly1am)** : ce plasmide est identique à pMG16 excepté pour la séquence d'entrée qui est inversée.

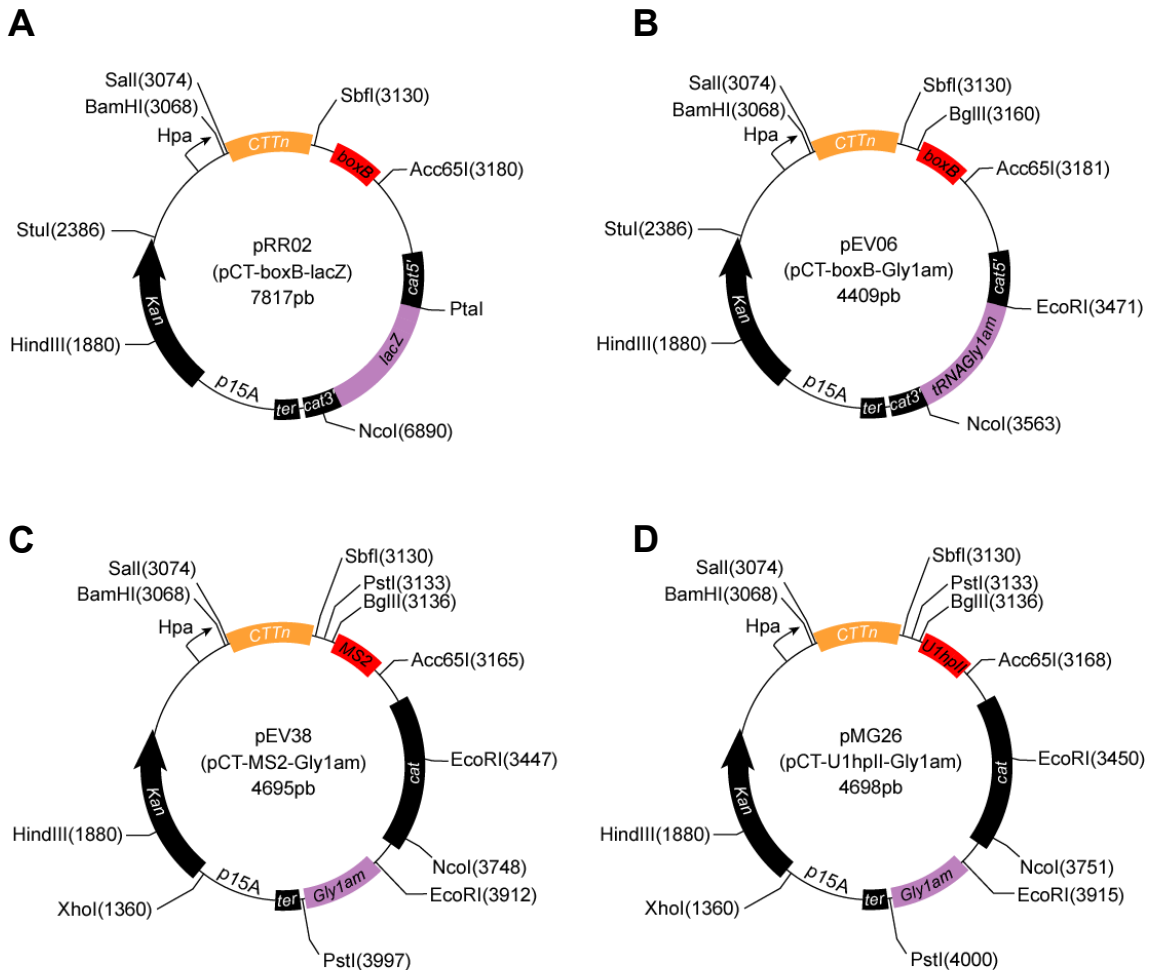


Figure 40 : Cartes génétiques des plasmides rapporteurs : (A) pRR02 (pCT-boxB-lacZ), (B) pEV06 (pCT-boxB-Gly1am), (C) pEV38 (pCT-MS2-Gly1am), (D) pMG26 (pCT-U1hpII-Gly1am).

3. 2. Plasmides protéiques.

Les plasmides protéiques sont dérivés du plasmide pEZ12-luc (ColE1, Lutz et Bujard, 1997), ils portent un gène de résistance à l'ampicilline.

- **pMB08 (pZ- λ N)** : (voir : [Vieu et Rahmouni, 2004](#)).

- **pMG19 (pZ-MS2CP)** : (voir : [Vieu et Rahmouni, 2004](#)).

- **pMB18 (pZ-U1A)** : Le plasmide pMB18 a été obtenu par échange de fragment XhoI/NcoI entre les vecteurs pU1A1 (donneur; [Jain et Belasco, 1996](#)) et pMB08 (accepteur).

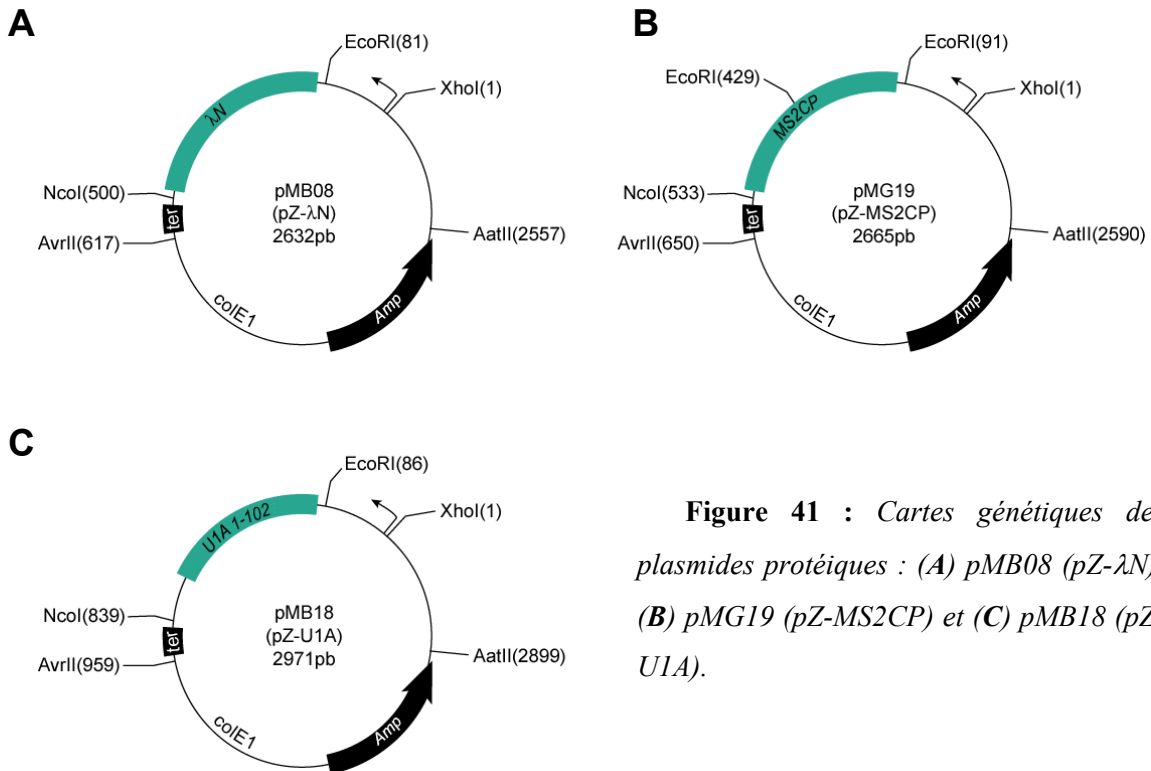


Figure 41 : Cartes génétiques des plasmides protéiques : (A) pMB08 (pZ- λ N), (B) pMG19 (pZ-MS2CP) et (C) pMB18 (pZ-U1A).

Chaque construction a été contrôlée par séquençage aux dideoxynTP (kit thermosequenase d'Amersham Biosciences).

4. Détection et quantification de l'activité β -galactosidase.

4. 1. Détection de l'activité β -galactosidase en milieu solide.

L'activité β -galactosidase des bactéries est mise en évidence par la coloration bleue des colonies en milieu LB-Agar complémenté à 40 μ g/ml en X-Gal (5-Bromo-4-Chloro-3-Indolyl- β -D-Galactopyranoside; Qbiogen).

4. 2. Essai β -galactosidase en culture liquide (méthode de Miller).

8ml de LB sontensemencés au 1/1000, à partir d'une culture de la nuit, puis incubés à 37°C sous agitation jusqu'à $A_{600} = 0,3-0,7$. 100 μ l de cette culture cellulaire sont alors mélangés à 900 μ l de tampon Z (60mM Na_2HPO_4 , 40mM NaH_2PO_4 , 10mM MgSO_4 , 50mM β -mercaptoethanol). On ajoute ensuite 60 μ l de chloroforme et 30 μ l de SDS 0,1%, le mélange est vortexé 10s pour lyser les cellules, puis placé à 28°C pendant 5min. Au temps $t = 0$ on ajoute 200 μ l d'ONPG (Ortho-nitrophenyl- β -D-galactopyranoside : Sigma) à 8mg/ml dans du tampon A (120mM K_2HPO_4 , 70mM KH_2PO_4 , 3,3mM $\text{C}_6\text{H}_8\text{O}_7\text{Na}_3$, 14mM $(\text{NH}_4)_2\text{SO}_4$) et le mélange est incubé à 28°C. Au temps t on arrête la réaction par ajout de 500 μ l de Na_2CO_3 (1M). La solution est centrifugée, on récupère le surnageant et on mesure la DO à 420 et 550nm. L'activité β -galactosidase, en unités Miller, est alors calculée selon la formule suivante (Miller, 1972):

$$1000 \times (\text{DO}_{420} - 1.75 \times \text{DO}_{550}) / (t(\text{min}) \times 0.1 \times \text{DO}_{600})$$

5. Dosage de l'accumulation d'ARNt dans les cellules.

5. 1. Extraction d'ARN.

4ml de LB sontensemencés au 1/1000, à partir d'une culture de la nuit, jusqu'à $\text{DO}_{600} = 0,6-0,7$. Les cellules sont alors culottées puis remises en solution dans 400 μ l de LB frais. La solution est alors transférée dans un tube contenant 500 μ l de REB (20mM acétate de sodium pH 5,2; 2% SDS; 0,3M sucrose) préalablement chauffés à 100°C, et le tout est incubé 30s à 100°C. On ajoute alors 1ml de phénol, on agite à température ambiante et on laisse 30min dans la glace. La solution est centrifugée 10min à 4°C, le surnageant est récupéré (700 μ l) et mélangé à 700 μ l de phénol-chloroforme. Le mélange est une nouvelle fois centrifugé à 4°C, on ajoute 50 μ l d'acétate de sodium (3M; pH 4,5) et 1ml d'éthanol froid. La précipitation est réalisée une nuit à -20°C (ou 30min à -80°C). Le précipité est lavé, séché et repris dans 25 μ l d'eau traitée au DEPC.

5. 2. Northern blot.

15µg d'ARN totaux sont chargés sur un gel d'agarose 1,2%, contenant 14% de formaldéhyde. La migration est réalisée pendant 2h à 110V environ, en tampon MOPS1X pH7 (95mM MOPS, 50nM acétate de sodium, 10nM EDTA). Le gel est découpé et lavé une première fois 30min dans l'eau (pour enlever le formaldéhyde), puis à nouveau 30min dans du TBE 0,5X. Le transfert des ARNs est réalisé à l'aide d'un électroblotteur (transblot SD de Biorad), en présence de TBE 0,5X pendant 30min, sur une membrane de nylon chargée positivement (Hybond-N⁺, Q-biogen) et préalablement lavée dans 0,5X TBE. Après le transfert, les membranes sont séchées 2h à 80° (fixation irréversible des ARNs).

5. 3. Hybridation.

Les membranes sont tout d'abord préhybridées 4h à 65°C dans un tampon de préhybridation (5X SSC, SDS 2%, Denhart 5X, NaHPO₄ 2mM, ADN de sperme de saumon 100µg/ml). L'hybridation est effectuée une nuit à 65°C en ajoutant la sonde tRNA^{Gly} : 5' GCTTTAGACCGTGAGGTAATAGCC 3' (complémentaire des nucléotides 18 à 40 de l'ARNtGly1am) ou la sonde 5S : 5' ACT ACC ATC GGC GCT ACG GC 3' (complémentaire des nucléotides 61 à 80 de l'ARNr 5S) marquées avec du dATPα³²P, par phosphorylation de l'extrémité 5' à la polynucléotide kinase. Les membranes sont ensuite lavées 15 min dans du 2XSSC-0,1% SDS puis 15 min dans du 1XSSC-0,1% SDS. Les signaux d'hybridation sont révélés avec un Storm Phosphorimager (Molecular Dynamics) et les données sont traitées avec le logiciel Imagequant 5.1.

REFERENCES

REFERENCES

1. Adhya, S., Gottesman, M. & De Crombrughe, B. (1974). Release of polarity in *Escherichia coli* by gene N of phage lambda: termination and antitermination of transcription. *Proc. Natl. Acad. Sci. U S A*, **71**, 2534-2538.
2. Aksoy, S., Squires, C. L. & Squires, C. (1984). Evidence for antitermination in *Escherichia coli* RRNA transcription. *J. Bacteriol.* **159**, 260-264.
3. Albrechtsen, B., Squires, C. L., Li, S., Squires, C. (1990). Antitermination of characterized transcriptional terminators by the *Escherichia coli* rrnG leader region. *J. Mol. Biol.* **213**, 123-134.
4. Alifano, P., Rivellini, F., Limauro, D., Bruni, C. B. & Carlomagno, M. S. (1991). A consensus motif common to all Rho-dependent prokaryotic transcription terminators. *Cell*, **64**, 553-563.
5. Allison, L. A., Moyle, M., Shales, M. & Ingles, C. J. (1985). Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell*, **42**, 599-610.
6. Allison, T. J., Wood, T. C., Briercheck, D. M., Rastinejad, F., Richardson, J. P. & Rule, G. S. (1998). Crystal structure of the RNA-binding domain from transcription termination factor rho. *Nat. Struct. Biol.* **5**, 352-356.
7. Amster-Choder, O. & Wright, A. (1993). Transcriptional regulation of the bgl operon of *Escherichia coli* involves phosphotransferase system-mediated phosphorylation of a transcriptional antiterminator. *J. Cell. Biochem.* **51**, 83-90.
8. Antson, A. A., Dodson, E. J., Dodson, G., Greaves, R. B., Chen, X. & Gollnick, P. (1999). Structure of the trp RNA-binding attenuation protein, TRAP, bound to RNA. *Nature*, **401**, 235-242.
9. Antson, A. A., Otridge, J., Brzozowski, A. M., Dodson, E. J., Dodson, G. G., Wilson, K. S., Smith, T. M., Yang, M., Kurecki, T. & Gollnick, P. (1995). The structure of trp RNA-binding attenuation protein. *Nature*, **374**, 693-700.
10. Apirion, D. & Miczak, A. (1993). RNA processing in prokaryotic cells. *Bioessays*, **15**, 113-20.
11. Artsimovitch, I. & Landick, (2000). R. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl. Acad. Sci. U S A*, **97**, 7090-7095.

Références.

12. Atkinson, B. L. & Gottesman, M. E. (1992). The Escherichia coli rpoB60 mutation blocks antitermination by coliphage HK022 Q-function. *J. Mol. Biol.* **227**, 29-37.
13. d'Aubenton Carafa, Y., Brody, E. & Thermes, C. (1990). Prediction of rho-independent Escherichia coli transcription terminators. A statistical analysis of their RNA stem-loop structures. *J. Mol. Biol.* **216**, 835-858.
14. Barik, S. & Das, A. (1990). An analysis of the role of host factors in transcription antitermination in vitro by the Q protein of coliphage lambda. *Mol. Gen. Genet.* **222**, 152-156.
15. Bar-Nahum, G. & Nudler, E. (2001). Isolation and characterization of sigma(70)-retaining transcription elongation complexes from Escherichia coli. *Cell*, **106**, 443-451.
16. Battiste, J. L., Mao, H., Rao, N. S., Tan, R., Muhandiram, D. R., Kay, L. E., Frankel, A. D. & Williamson, J. R. (1996). Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science*, **273**, 1547-1551.
17. Bear, D. G., Andrews, C. L., Singer, J. D., Morgan, W. D., Grant, R. A., von Hippel, P. H. & Platt, T. (1985). Escherichia coli transcription termination factor rho has a two-domain structure in its activated form. *Proc. Natl. Acad. Sci. U S A*, **82**, 1911-1915.
18. Bear, D. G., Hicks, P. S., Escudero, K. W., Andrews, C. L., McSwiggen, J. A. & von Hippel, P. H. (1988). Interactions of Escherichia coli transcription termination factor rho with RNA. II. Electron microscopy and nuclease protection experiments. *J. Mol. Biol.* **199**, 623-635.
19. Bernardi, A. & Spahr, P. F. (1972). Nucleotide sequence at the binding site for coat protein on RNA of bacteriophage R17. *Proc. Natl. Acad. Sci. U S A*, **69**, 3033-3037.
20. Bernstein, D. S., Buter, N., Stumpf, C. & Wickens, M. (2002). Analyzing mRNA-protein complexes using a yeast three-hybrid system. *Methods*, **26**, 123-141.
21. Bernstein, E., Denli, A. M. & Hannon, G. J. (2001a). The rest is silence. *RNA*, **7**, 1509-1521.
22. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. (2001b). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, **409**, 363-366.
23. Beckett, D. & Uhlenbeck, O. C. (1988). Ribonucleoprotein complexes of R17 coat protein and a translational operator analog. *J. Mol. Biol.* **204**, 927-938.
24. Birge, E. A. & Low, K. B. (1974). Detection of transcribable recombination products following conjugation in rec⁺, recB⁻ and recC⁻ strains of Escherichia coli K12. *J. Mol. Biol.* **83**, 447-457.

Références.

25. Bogden, C. E., Fass, D., Bergman, N., Nichols, M. D. & Berger, J. M. (1999). The structural basis for terminator recognition by the Rho transcription termination factor. *Mol. Cell*, **3**, 487-493.
26. Borukhov, S., Sagitov, V. & Goldfarb, A. (1993). Transcript cleavage factors from *E. coli*. *Cell*, **72**, 459-466.
27. Bremer & Denis (1987). Modulation of chemical composition and other parameters of the cell by growth rate. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* vol. 2 (Neidhardt, F. C. et al, eds), pp. 1527-1542, American Society for Microbiology, Washington, D. C.
28. Brennan, C. A., Dombroski, A. J. & Platt, T. (1987). Transcription termination factor rho is an RNA-DNA helicase. *Cell*, **48**, 945-952.
29. Brennan, C. A., Steinmetz, E. J., Spear, P. & Platt, T. (1990). Specificity and efficiency of rho-factor helicase activity depends on magnesium concentration and energy coupling to NTP hydrolysis. *J. Biol. Chem.* **265**, 5440-5447.
30. Brendel, V., Hamm, G. H. & Trifonov, E. N. (1986). Terminators of transcription with RNA polymerase from *Escherichia coli*: what they look like and how to find them. *J. Biomol. Struct. Dyn.* **3**, 705-723.
31. Brendel, V. & Trifonov, E. N. (1984). A computer algorithm for testing potential prokaryotic terminators. *Nucleic Acids Res.* **12**, 4411-4427.
32. Brewster, J. M. & Morgan, E. A. (1981). Tn9 and IS1 inserts in a ribosomal ribonucleic acid operon of *Escherichia coli* are incompletely polar. *J. Bacteriol.* **148**, 897-903.
33. Briercheck, D. M., Allison, T. J., Richardson, J. P., Ellena, J. F., Wood, T. C. & Rule, G. S. (1996). ¹H, ¹⁵N and ¹³C resonance assignments and secondary structure determination of the RNA-binding domain of *E. coli* rho protein. *J. Biomol. NMR*, **8**, 429-44.
34. Briercheck, D. M., Wood, T. C., Allison, T. J., Richardson, J. P. & Rule, G. S. (1998). The NMR structure of the RNA binding domain of *E. coli* rho factor suggests possible RNA-protein interactions. *Nat. Struct. Biol.* **5**, 393-399.
35. Burgess, B. R. & Richardson, J. P. (2001a). RNA passes through the hole of the protein hexamer in the complex with the *Escherichia coli* Rho factor. *J. Biol. Chem.* **276**, 4182-4189.
36. Burgess, B. R. & Richardson, J. P. (2001b). Transcription factor Rho does not require a free end to act as an RNA-DNA helicase on an RNA. *J. Biol. Chem.* **276**, 17106-17110.

Références.

37. Burgess, R. R. & Travers, A. A. (1970). Escherichia coli RNA polymerase: purification, subunit structure, and factor requirements. *Fed. Proc.* **29**, 1164-1169.
38. Burns, C. M. & Richardson, J. P. (1995). NusG is required to overcome a kinetic limitation to Rho function at an intragenic terminator. *Proc. Natl. Acad. Sci. U S A*, **92**, 4738-4742.
39. Burns, C. M., Richardson, L. V. & Richardson, J. P. (1998). Combinatorial effects of NusA and NusG on transcription elongation and Rho-dependent termination in Escherichia coli. *J. Mol. Biol.* **278**, 307-316.
40. Cadwell, R. C. & Joyce, G. F. (1992). Randomization of genes by PCR mutagenesis. *PCR Methods Appl.* **2**, 28-33.
41. Cai, Z., Gorin, A., Frederick, R., Ye, X., Hu, W., Majumdar, A., Kettani, A. & Patel, D. J. (1998). Solution structure of P22 transcriptional antitermination N peptide-boxB RNA complex. *Nat. Struct. Biol.* **5**, 203-212.
42. Campbell, F. E. Jr. & Setzer, D. R. (1992). Transcription termination by RNA polymerase III: uncoupling of polymerase release from termination signal recognition. *Mol. Cell. Biol.* **12**, 2260-2272.
43. Cassidy, L. A. & Maher, L. J. 3rd. (2003). Yeast genetic selections to optimize RNA decoys for transcription factor NF-kappa B. *Proc. Natl. Acad. Sci. U S A*, **100**, 3930-3935.
44. Cassidy, L. A. & Maher, L. J. 3rd. (2001). In vivo recognition of an RNA aptamer by its transcription factor target. *Biochemistry*, **40**, 2433-2438.
45. Cech, T. R. (1988). Conserved sequences and structures of group I introns: building an active site for RNA catalysis--a review. *Gene*, **73**, 259-271.
46. Ceruzzi, M. A., Bektesh, S. L. & Richardson, J. P. (1985). Interaction of rho factor with bacteriophage lambda cro gene transcripts. *J. Biol. Chem.* **260**, 9412-9418.
47. Chamberlin, M. & Hsu, L. (1996). RNA chain initiation and promoter clearance by E. Coli RNA polymerase. In *Regulation of gene expression in E. Coli* (Lin, E. C. C. & Lynch, A. S. ed). pp. 7-25. Landes, R. G. Co., Austin, Tex.
48. Chattopadhyay, S., Hung, S. C., Stuart, A. C., Palmer, A. G., Garcia-Mena, J., Das, A. & Gottesman, M. E. (1995). Interaction between the phage HK022 Nun protein and the nut RNA of phage lambda. *Proc. Natl Acad. Sci. USA*, **92**, 12131-12135.

Références.

49. Chen, C. Y., Galluppi, G. R. & Richardson, J. P. (1986). Transcription termination at lambda tR1 is mediated by interaction of rho with specific single-stranded domains near the 3' end of cro mRNA. *Cell*, **46**, 1023-1028.
50. Chen, C. Y. & Richardson, J. P. (1987). Sequence elements essential for rho-dependent transcription termination at lambda tR1. *J. Biol. Chem.* **262**, 11292-11299.
51. Cisneros, B., Court, D., Sanchez, A. & Montanez, C. (1996). Point mutations in a transcription terminator, lambda tI, that affect both transcription termination and RNA stability. *Gene*, **181**, 127-33.
52. Coleman, J., Green, P. J. & Inouye, M. (1984). The use of RNAs complementary to specific mRNAs to regulate the expression of individual bacterial genes. *Cell*, **37**, 429-436.
53. Condon, C., Squires, C. & Squires, C. L. (1995). Control of rRNA transcription in Escherichia coli. *Microbiol. Rev.* **59**, 623-645.
54. Couzin, J. (2002). Breakthrough of the year. Small RNAs make big splash. *Science*, **298**, 2296-2297.
55. de Crombrughe, B., Mudryj, M., DiLauro, R. & Gottesman, M. (1979). Specificity of the bacteriophage lambda N gene product (pN): nut sequences are necessary and sufficient for antitermination by pN. *Cell*, **18**, 1145-1151.
56. Craven, M. G. & Friedman, D. I. (1991). Analysis of the Escherichia coli nusA10(Cs) allele: relating nucleotide changes to phenotypes. *J. Bacteriol.* **173**, 1485-1491.
57. Cullen, B. R. (1990). The HIV-1 Tat protein: an RNA sequence-specific processivity factor? *Cell*, **63**, 655-657.
58. Danner, S. & Belasco, J. G. (2001). T7 phage display: a novel genetic selection system for cloning RNA-binding proteins from cDNA libraries. *Proc. Natl. Acad. Sci. U S A*, **98**, 12954-12959.
59. Darst, S. A. (2001). Bacterial RNA polymerase. *Curr. Opin. Struct. Biol.* **11**, 155-162.
60. Das, A. (1993). Control of transcription termination by RNA-binding proteins. *Annu. Rev. Biochem.* **62**, 893-930.
61. Das, A., Gottesman, M. E., Wardwell, J., Trisler, P. & Gottesman, S. (1983). lambda mutation in the Escherichia coli rho gene that inhibits the N protein activity of phage lambda. *Proc. Natl. Acad. Sci. U S A*, **80**, 5530-5534.

Références.

62. Das, A., Court, D. & Adhya, S. (1976). Isolation and characterization of conditional lethal mutants of *Escherichia coli* defective in transcription termination factor rho. *Proc. Natl Acad. Sci. USA*, **73**, 1959-1963.
63. Das, A. & Wolska, K. (1984). Transcription antitermination in vitro by lambda N gene product: requirement for a phage nut site and the products of host nusA, nusB, and nusE genes. *Cell*, **38**, 165-173.
64. Davenport, R. J., Wuite, G. J., Landick, R. & Bustamante, C. (2000). Single-molecule study of transcriptional pausing and arrest by *E. coli* RNA polymerase. *Science*, **287**, 2497-2500.
65. Dolan, J. W., Marshall, N. F. & Richardson, J. P. (1990). Transcription termination factor rho has three distinct structural domains. *J. Biol. Chem.* **265**, 5747-5754.
66. Dombroski, A. J. & Platt, T. (1988). Structure of rho factor: an RNA-binding domain and a separate region with strong similarity to proven ATP-binding domains. *Proc. Natl. Acad. Sci. U S A*, **85**, 2538-2542.
67. Doudna, J. A. & Cech, T. R. (2002). The chemical repertoire of natural ribozymes. *Nature*, **418**, 222-228.
68. Downing, W. L., Sullivan, S. L., Gottesman, M. E. & Dennis, P. P. (1990). Sequence and transcriptional pattern of the essential *Escherichia coli* secE-nusG operon. *J. Bacteriol.* **172**, 1621-1627.
69. Egelman, E. H., Yu, X., Wild, R., Hingorani, M. M. & Patel, S. S. (1995). Bacteriophage T7 helicase/primase proteins form rings around single-stranded DNA that suggest a general structure for hexameric helicases. *Proc. Natl. Acad. Sci. U S A*, **92**, 3869-3873.
70. Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. & Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, **411**, 494-498.
71. Ellington, A. D. & Szostak, J. W. (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature*, **346**, 818-822.
72. Ellington, A. D., Leclerc, F., Cedergren, R. (1996). An RNA groove. *Nat. Struct. Biol.* **3**, 981-984.
73. Epshtein, V., Toulmé, F., Rahmouni, A. R., Borukhov, S. & Nudler, E. (2003). Transcription through the roadblocks: the role of RNA polymerase cooperation. *EMBO J.* **22**, 4719-4727.
74. Ermolaeva, M. D., Khalak, H. G., White, O., Smith, H. O. & Salzberg, S. L. (2000). Prediction of transcription terminators in bacterial genomes. *J. Mol. Biol.* **301**, 27-33.

Références.

75. Espéli, O., Moulin, L. & Boccard, F. (2001). Transcription attenuation associated with bacterial repetitive extragenic BIME elements. *J. Mol. Biol.* **314**, 375-386.
76. Fairman, M. E., Maroney, P. A., Wang, W., Bowers, H. A., Gollnick, P., Nilsen, T. W. & Jankowsky, E. (2004). Protein displacement by DExH/D "RNA helicases" without duplex unwinding. *Science*, **304**, 730-734.
77. Faus, I. & Richardson, J. P. (1989). Thermodynamic and enzymological characterization of the interaction between transcription termination factor rho and lambda cro mRNA. *Biochemistry*, **28**, 3510-3517.
78. Faus, I. & Richardson, J. P. (1990). Structural and functional properties of the segments of lambda cro mRNA that interact with transcription termination factor Rho. *J. Mol. Biol.* **212**, 53-66.
79. Fields, S. & Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature*, **340**, 245-246.
80. Figueroa-Bossi, N., Guérin, M., Rahmouni, A.R., Leng, M. & Bossi, L. (1998). The supercoiling sensitivity of a bacterial tRNA promoter parallels its responsiveness to stringent control. *EMBO J.* **17**, 2359-2367.
81. Finger, L. R. & Richardson, J. P. (1982). Stabilization of the hexameric form of Escherichia coli protein rho under ATP hydrolysis conditions. *J. Mol. Biol.* **156**, 203-219.
82. Finnegan, E. J. & Matzke M. A. (2003). The small RNA world. *J. Cell Sci.* **116**, 4689-4693.
83. Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391**, 806-811.
84. Fisher, R. F. & Yanofsky, C. (1983). Mutations of the beta subunit of RNA polymerase alter both transcription pausing and transcription termination in the trp operon leader region in vitro. *J. Biol. Chem.* **258**, 8146-8150.
85. Franklin, N. C. (1974). Altered reading of genetic signals fused to the N operon of bacteriophage lambda: genetic evidence for modification of polymerase by the protein product of the N gene. *J. Mol. Biol.* **89**, 33-48.
86. Franklin, N. C. (1985). Conservation of genome form but not sequence in the transcription antitermination determinants of bacteriophages lambda, phi 21 and P22. *J. Mol. Biol.* **181**, 75-84.
87. Frankel, A. D. (2000). Fitting peptides into the RNA world. *Curr. Opin. Struct. Biol.* **10**, 332-340.

Références.

88. Friedman, D. I. & Baron, L. S. (1974). Genetic characterization of a bacterial locus involved in the activity of the N function of phage lambda. *Virology*, **58**, 141-148.
89. Friedman, D. I., Baumann, M. & Baron, L. S. (1976). Cooperative effects of bacterial mutations affecting lambda N gene expression. I. Isolation and characterization of a nusB mutant. *Virology*, **73**, 119-127.
90. Friedman, D. I. & Court, D. L. (1995). Transcription antitermination: the lambda paradigm updated. *Mol. Microbiol.* **18**, 191-200.
91. Friedman, D. I., Jolly, C. T. & Mural, R. J. (1973). Interference with the expression of the N gene function of phage lambda in a mutant of Escherichia coli. *Virology*, **51**, 216-26.
92. Friedman, D. I. & Olson, E. R. (1983). Evidence that a nucleotide sequence, "boxA," is involved in the action of the NusA protein. *Cell*, **34**, 143-149.
93. Friedman, D. I., Schauer, A. T., Baumann, M. R., Baron, L. S. & Adhya, S. L. (1981). Evidence that ribosomal protein S10 participates in control of transcription termination. *Proc. Natl. Acad. Sci. U S A*, **78**, 1115-1118.
94. Galluppi, G. R. & Richardson, J. P. (1980). ATP-induced changes in the binding of RNA synthesis termination protein Rho to RNA. *J. Mol. Biol.* **138**, 513-539.
95. Gan, E. & Richardson, J. P. (1999). ATP and other nucleotides stabilize the Rho-mRNA complex. *Biochemistry*, **38**, 16882-16888.
96. Geiselman, J., Wang, Y., Seifried, S. E. & von Hippel, P. H. (1993). A physical model for the translocation and helicase activities of Escherichia coli transcription termination protein Rho. *Proc. Natl. Acad. Sci. U S A*, **90**, 7754-7758
97. Geiselman, J., Yager, T. D., Gill, S. C., Calmettes, P. & von Hippel, P. H. (1992). Physical properties of the Escherichia coli transcription termination factor rho. 1. Association states and geometry of the rho hexamer. *Biochemistry*, **31**, 111-121.
98. Gelles, J. & Landick, R. (1998). RNA polymerase as a molecular motor. *Cell*, **93**, 13-16.
99. Georgopoulos, C. P. (1971). Bacterial mutants in which the gene N function of bacteriophage lambda is blocked have an altered RNA polymerase. *Proc. Natl. Acad. Sci. U S A*, **68**, 2977-2981.
100. Georgopoulos, C. P., Swindle, J., Keppel, F., Ballivet, M., Bisig, R. & Eisen, H. (1980). Studies on the E. coli groNB (nusB) gene which affects bacteriophage lambda N gene function. *Mol. Gen. Genet.* **179**, 55-61.

Références.

101. Ghysen, A. & Pironio, M. (1972). Relationship between the N function of bacteriophage lambda and host RNA polymerase. *J. Mol. Biol.* **65**, 259-272.
102. Gill, S. C., Weitzel, S. E. & von Hippel, P. H. (1991). Escherichia coli sigma 70 and NusA proteins. I. Binding interactions with core RNA polymerase in solution and within the transcription complex. *J. Mol. Biol.* **220**, 307-324.
103. Gogol, E. P., Seifried, S. E. & von Hippel, P. H. (1991). Structure and assembly of the Escherichia coli transcription termination factor rho and its interaction with RNA. I. Cryoelectron microscopic studies. *J. Mol. Biol.* **221**, 1127-1138.
104. Gold, L., Brown, D., He, Y., Shtatland, T., Singer, B. S. & Wu, Y. (1997). From oligonucleotide shapes to genomic SELEX: novel biological regulatory loops. *Proc. Natl. Acad. Sci. U S A*, **94**, 59-64.
105. Gold, L., Singer, B., He, Y. Y. & Brody, E. (1997). SELEX and the evolution of genomes. *Curr. Opin. Genet. Dev.* **7**, 848-851.
106. Gong, F., Ito, K., Nakamura, Y. & Yanofsky, C. (2001). The mechanism of tryptophan induction of tryptophanase operon expression: tryptophan inhibits release factor-mediated cleavage of TnaC-peptidyl-tRNA(Pro). *Proc. Natl. Acad. Sci. U S A*, **98**, 8997-9001.
107. Gong, F. & Yanofsky, C. (2002). Analysis of tryptophanase operon expression in vitro: accumulation of TnaC-peptidyl-tRNA in a release factor 2-depleted S-30 extract prevents Rho factor action, simulating induction. *J. Biol. Chem.* **277**, 17095-17100.
108. Gottesman, M. (1999). Bacteriophage lambda: the untold story. *J. Mol. Biol.* **293**, 177-180.
109. Graham, J.E. & Richardson, J. P. (1998). Rut sites in the nascent transcript mediate Rho-dependent transcription termination in vivo. *J. Biol. Chem.* **273**, 20764-20769.
110. Grayhack, E. J., Yang, X. J., Lau, L. F. & Roberts, J. W. (1985). Phage lambda gene Q antiterminator recognizes RNA polymerase near the promoter and accelerates it through a pause site. *Cell*, **42**, 259-269.
111. Greenblatt, J. & Li, J. (1981a). Interaction of the sigma factor and the nusA gene protein of E. coli with RNA polymerase in the initiation-termination cycle of transcription. *Cell*, **24**, 421-428. Greenblatt, J. & Li, J. (1981b). The nusA gene protein of Escherichia coli. Its identification and a demonstration that it interacts with the gene N transcription anti-termination protein of bacteriophage lambda. *J. Mol. Biol.* **147**, 11-23.

Références.

113. Greenblatt, J., Mah, T. F., Legault, P., Mogridge, J., Li, J. & Kay, L. E. (1998). Structure and mechanism in transcriptional antitermination by the bacteriophage lambda N protein. *Cold. Spring. Harb. Symp. Quant. Biol.* **63**, 327-336.
114. Greenblatt, J., Nodwell, J. R. & Mason, S. W. (1993). Transcriptional antitermination. *Nature*, **364**, 401-406.
115. Grundy, F. J. & Henkin, T. M. (1993). tRNA as a positive regulator of transcription antitermination in *B. subtilis*. *Cell*, **74**, 475-482.
116. Grundy, F. J., Winkler, W. C. & Henkin, T. M. (2002). tRNA-mediated transcription antitermination in vitro: codon-anticodon pairing independent of the ribosome. *Proc. Natl. Acad. Sci. USA*, **99**, 11121-11126.
117. Guérin, M., Leng, M. & Rahmouni, A. R. (1996). High resolution mapping of E.coli transcription elongation complex in situ reveals protein interactions with the non-transcribed strand. *EMBO J.* **15**, 5397-5407.
118. Guérin, M., Robichon, N., Geiselman, J. & Rahmouni, A. R. (1998). A simple polypyrimidine repeat acts as an artificial Rho-dependent terminator in vivo and in vitro. *Nucl. Acids Res.* **26**, 4895-4900.
119. Guo, H., Karberg, M., Long, M., Jones, J. P. 3rd., Sullenger, B. & Lambowitz, A. M. (2000). Group II introns designed to insert into therapeutically relevant DNA target sites in human cells. *Science*, **289**, 452-457.
120. Gusarov, I. & Nudler, E. (1999). The mechanism of intrinsic transcription termination. *Mol. Cell*, **3**, 495-504.
121. Gusarov, I. & Nudler, E. (2001). Control of intrinsic transcription termination by N and NusA: the basic mechanisms. *Cell*, **107**, 437-449.
122. Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* **177**, 4121-4130.
123. Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature*, **404**, 293-296.
124. Hannon, G. J. (2002). RNA interference. *Nature*, **418**, 244-251.
125. Hansen, U. M. & McClure, W. R. (1980). Role of the sigma subunit of *Escherichia coli* RNA polymerase in initiation. II. Release of sigma from ternary complexes. *J. Biol. Chem.* **255**, 9564-9570.

Références.

126. Harada, K., Martin, S. S. & Frankel, A. D. (1996). Selection of RNA-binding peptides in vivo. *Nature*, **380**, 175-179.
127. Hart, C. M. & Roberts, J. W. (1991). Rho-dependent transcription termination. Characterization of the requirement for cytidine in the nascent transcript. *J. Biol. Chem.* **266**, 24140-24148.
128. Hart, C. M. & Roberts, J. W. (1994). Deletion analysis of the lambda tR1 termination region. Effect of sequences near the transcript release sites, and the minimum length of rho-dependent transcripts. *J. Mol. Biol.* **237**, 255-265.
129. de Haseh, P. L. & Helmann, J. D. (1995). Open complex formation by Escherichia coli RNA polymerase: the mechanism of polymerase-induced strand separation of double helical DNA. *Mol. Microbiol.* **16**, 817-824.
130. de Haseh, P. L. Zupancic, M. L. & Record, M. T. Jr. (1998). RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. *J. Bacteriol.* **180**, 3019-3025.
131. Henkin, T. M. (1996). Control of transcription termination in prokaryotes. *Annu. Rev. Genet.* **30**, 35-57.
132. Henkin, T. M. (2000). Transcription termination control in bacteria. *Curr. Opin. Microbiol.* **3**, 149-153.
133. Henkin, T. M. & Yanofsky, C. (2002). Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. *Bioessays*, **24**, 700-707.
134. Hentze, M. W. & Kuhn, L. C. (1996). Molecular control of vertebrate iron metabolism: mRNA-based regulatory circuits operated by iron, nitric oxide, and oxidative stress. *Proc. Natl. Acad. Sci. U S A*, **93**, 8175-8182.
135. Holben, W. E. & Morgan, E. A. (1984). Antitermination of transcription from an Escherichia coli ribosomal RNA promoter. *Proc. Natl. Acad. Sci. U S A*, **81**, 6789-6793.
136. Horwitz, R. J., Li, J. & Greenblatt, J. (1987). An elongation control particle containing the N gene transcriptional antitermination protein of bacteriophage lambda. *Cell*, **51**, 631-641.
137. Hung, S. C. & Gottesman, M. E. (1995). Phage HK022 Nun protein arrests transcription on phage lambda DNA in vitro and competes with the phage lambda N antitermination protein. *J. Mol. Biol.* **247**, 428-442.
138. Jaeger, S., Eriani, G. & Martin, F. (2004). Results and prospects of the yeast three-hybrid system. *FEBS Lett.* **556**, 7-12.

Références.

139. Jain, C. & Belasco, J. G. (1996). A structural model for the HIV-1 Rev-RRE complex deduced from altered-specificity rev variants isolated by a rapid genetic strategy. *Cell*, **87**, 115-125.
140. Jain, C. & Belasco, J. G. (2000). Rapid genetic analysis of RNA-protein interactions by translational repression in *Escherichia coli*. *Methods Enzymol.* **318**, 309-332.
141. Jankowsky, E., Gross, C. H., Shuman, S. & Pyle, A. M. (2001). Active disruption of an RNA-protein interaction by a DExH/D RNA helicase. *Science*, **291**, 121-125.
142. Jin, D. J., Burgess, R. R., Richardson, J. P. & Gross, C. A. (1992). Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho. *Proc. Natl. Acad. Sci. U S A*, **89**, 1453-1457.
143. Jones, J. T., Lee, S. W. & Sullenger, B. A. (1996). Tagging ribozyme reaction sites to follow trans-splicing in mammalian cells. *Nat. Med.* **2**, 643-648.
144. Kainz, M. & Gourse, R. L. (1998). The C-terminal domain of the alpha subunit of *Escherichia coli* RNA polymerase is required for efficient rho-dependent transcription termination. *J. Mol. Biol.* **284**, 1379-1390.
145. Kao, S. Y., Calman, A. F., Luciw, P. A. & Peterlin, B. M. (1987). Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature*, **330**, 489-493.
146. Kato, H., Sumimoto, H., Pognonec, P., Chen, C. H., Rosen, C. A. & Roeder, R. G. (1992). HIV-1 Tat acts as a processivity factor in vitro in conjunction with cellular elongation factors. *Genes Dev.* **6**, 655-666.
147. Kenan, D. J., Query, C. C. & Keene, J. D. (1991). RNA recognition: towards identifying determinants of specificity. *Trends. Biochem. Sci.* **16**, 214-220.
148. Keppel, F., Georgopoulos, C. P. & Eisen, H. (1974). Host interference with expression of the lambda N gene product. *Biochimie*, **56**, 1505-1509.
149. Kim, D. E. & Patel, S. S. (1999). The mechanism of ATP hydrolysis at the noncatalytic sites of the transcription termination factor Rho. *J. Biol. Chem.* **274**, 32667-32671.
150. Kim, D. E., Patel, S. S. (2001). The kinetic pathway of RNA binding to the *Escherichia coli* transcription termination factor Rho. *J. Biol. Chem.* **276**, 13902-13910.
151. Kim, H. C., Zhou, J. G., Wilson, H. R., Mogilnitskiy, G., Court, D. L. & Gottesman, M. E. (2003). Phage HK022 Nun protein represses translation of phage lambda N (transcription termination/translation repression). *Proc. Natl Acad. Sci. USA*, **100**, 5308-5312.

Références.

152. King, R. A., Banik-Maiti, S., Jin, D. J. & Weisberg, R. A. (1996). Transcripts that increase the processivity and elongation rate of RNA polymerase. *Cell*, **87**, 893-903.
153. King, R.A. & Weisberg, R. A. (2003). Suppression of factor-dependent transcription termination by antiterminator RNA. *J. Bacteriol.* **185**, 7085-7091.
154. Ko, D. C., Marr, M. T., Guo, J. & Roberts, J. W. (1998). A surface of Escherichia coli sigma 70 required for promoter function and antitermination by phage lambda Q protein. *Genes Dev.* **12**, 3276-3285.
155. Komissarova, N., Becker, J., Solter, S., Kireeva, M. & Kashlev, M. (2002). Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. *Mol. Cell*, **10**, 1151-1162.
156. Konan, K. V. & Yanofsky, C. (2000). Rho-dependent transcription termination in the tna operon of Escherichia coli: roles of the boxA sequence and the rut site. *J. Bacteriol.* **182**, 3981-3988.
157. Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E. & Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell*, **31**, 147-157.
158. Krummel, B. & Chamberlin, M. J. (1989). RNA chain initiation by Escherichia coli RNA polymerase. Structural transitions of the enzyme in early ternary complexes. *Biochemistry*, **28**, 7829-7842.
159. Krummel, B. & Chamberlin, M. J. (1992). Structural analysis of ternary complexes of Escherichia coli RNA polymerase. Deoxyribonuclease I footprinting of defined complexes. *J. Mol. Biol.* **225**, 239-250.
160. Kung, H., Spears, C. & Weissbach, H. (1975). Purification and properties of a soluble factor required for the deoxyribonucleic acid-directed in vitro synthesis of beta-galactosidase. *J. Biol. Chem.* **250**, 1556-62.
161. Laird-Offringa, I. A. & Belasco, J. G. (1995). Analysis of RNA-binding proteins by in vitro genetic selection: identification of an amino acid residue important for locking U1A onto its RNA target. *Proc. Natl. Acad. Sci. U S A*, **92**, 11859-11863.
162. Laspia, M. F., Wendel, P. & Mathews, M. B. (1993). HIV-1 Tat overcomes inefficient transcriptional elongation in vitro. *J. Mol. Biol.* **232**, 732-746.
163. Lau, L. F., Roberts, J. W. & Wu, R. (1983). RNA polymerase pausing and transcript release at the lambda tR1 terminator in vitro. *J. Biol. Chem.* **258**, 9391-9397.

Références.

164. Landick, R. (1997). RNA polymerase slides home: pause and termination site recognition. *Cell*, **88**, 741-744.
165. Landick, R. (2001). RNA polymerase clamps down. *Cell*, **105**, 567-570.
166. Landick, R., Carey, J. & Yanofsky, C. (1985). Translation activates the paused transcription complex and restores transcription of the trp operon leader region. *Proc. Natl. Acad. Sci. U S A*, **82**, 4663-4667.
167. Landick, R. C., Turnbough, Jr. & Yanofsky, C. (1996). Transcription attenuation. In *Escherichia coli and Salmonella: Cellular and Molecular Biology* vol. 1 (Neidhardt, F. C. et al, eds), pp. 1263-1286, American Society for Microbiology, Washington, D. C.
168. Landt, S. G., Tan, R., Frankel, A. D. (2000). Screening RNA-binding libraries using Tat-fusion system in mammalian cells. *Methods Enzymol.* **318**, 350-363.
169. Lang, W. H., Morrow, B. E., Ju, Q., Warner, J. R. & Reeder, R. H. (1994). A model for transcription termination by RNA polymerase I. *Cell*, **79**, 527-534.
170. Lang, W. H. & Reeder, R. H. (1995). Transcription termination of RNA polymerase I due to a T-rich element interacting with Reb1p. *Proc. Natl. Acad. Sci. U S A*, **92**, 9781-9785.
171. Lang, W. H., Platt, T. & Reeder, R. H. (1998). Escherichia coli rho factor induces release of yeast RNA polymerase II but not polymerase I or III. *Proc. Natl. Acad. Sci. U S A*, **95**, 4900-4905.
172. Lau, L. F., Roberts, J. W. & Wu, R. (1982). Transcription terminates at lambda tR1 in three clusters. *Proc. Natl. Acad. Sci. U S A*, **79**, 6171-6175.
173. Lau, L. F., Roberts, J. W. & Wu, R. (1983). RNA polymerase pausing and transcript release at the lambda tR1 terminator in vitro. *J. Biol. Chem.* **258**, 9391-9397.
174. Lau, L. F., Roberts, J. W., Wu, R., Georges, F. & Narang, S. A. (1984). A potential stem-loop structure and the sequence CAAUCAA in the transcript are insufficient to signal rho-dependent transcription termination at lambda tR1. *Nucleic Acids Res.* **12**, 1287-1299.
175. Lazinski, D., Grzadzińska, E. & Das, A. (1989). Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. *Cell*, **59**, 207-218.
176. Lee, F. & Yanofsky, C. (1977). Transcription termination at the trp operon attenuators of Escherichia coli and Salmonella typhimurium: RNA secondary structure and regulation of termination. *Proc. Natl. Acad. Sci. U S A*, **74**, 4365-4369.

Références.

177. Legault, P., Li, J., Mogridge, J., Kay, L. E. & Greenblatt, J. (1998). NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell*, **93**, 289-299.
178. Leontis, N. B. & Westhof, E. (2001). Geometric nomenclature and classification of RNA base pairs. *RNA*, **7**, 499-512.
179. Leontis, N. B. & Westhof, E. (2003). Analysis of RNA motifs. *Curr. Opin. Struct. Biol.* **13**, 300-308.
180. Lesnik, E. A., Sampath, R., Levene, H. B., Henderson, T. J., McNeil, J. A. & Ecker, D. J. (2001). Prediction of rho-independent transcriptional terminators in Escherichia coli. *Nucleic Acids Res.* **29**, 3583-3594.
181. Levin, J. R., Krummel, B. & Chamberlin, M. J. (1987). Isolation and properties of transcribing ternary complexes of Escherichia coli RNA polymerase positioned at a single template base. *J. Mol. Biol.* **196**, 85-100.
182. Li, J., Horwitz, R., McCracken, S. & Greenblatt, J. (1992). NusG, a new Escherichia coli elongation factor involved in transcriptional antitermination by the N protein of phage lambda. *J. Biol. Chem.* **267**, 6012-9.
183. Li, J., Mason, S. W. & Greenblatt, J. (1993). Elongation factor NusG interacts with termination factor rho to regulate termination and antitermination of transcription. *Genes Dev.* **7**, 161-172.
184. Li, Z. & Deutscher, M. P. (1994). The role of individual exoribonucleases in processing at the 3' end of Escherichia coli tRNA precursors. *J. Biol. Chem.* **269**, 6064-6071.
185. Li, Z. & Deutscher, M. P. (1996). Maturation pathways for E. coli tRNA precursors: a random multienzyme process in vivo. *Cell*, **86**, 503-512.
186. Liu, M., Xie, Z. & Price, D. H. (1998). A human RNA polymerase II transcription termination factor is a SWI2/SNF2 family member. *J. Biol. Chem.* **273**, 25541-25544.
187. Lopez, P. J., Iost, I. & Dreyfus, M. (1994). The use of a tRNA as a transcriptional reporter: the T7 late promoter is extremely efficient in Escherichia coli but its transcripts are poorly expressed. *Nucl. Acids Res.* **22**, 1186-1193.
188. Lowery, C. & Richardson, J. P. (1977). Characterization of the nucleoside triphosphate phosphohydrolase (ATPase) activity of RNA synthesis termination factor p. II. Influence of synthetic RNA homopolymers and random copolymers on the reaction. *J. Biol. Chem.* **252**, 1381-1385.
189. Lowery-Goldhammer, C. & Richardson, J. P. (1974). An RNA-dependent nucleoside triphosphate phosphohydrolase (ATPase) associated with rho termination factor. *Proc. Natl. Acad. Sci. U S A*, **71**, 2003-2007.

Références.

190. Luttgén, H., Robelek, R., Muhlberger, R., Diercks, T., Schuster, S. C., Kohler, P., Kessler, H., Bacher, A. & Richter, G. (2002). Transcriptional regulation by antitermination. Interaction of RNA with NusB protein and NusB/NusE protein complex of *Escherichia coli*. *J. Mol. Biol.* **316**, 875-885.
191. Lutz, R. & Bujard, H. (1997). Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucl. Acids Res.* **25**, 1203-1210.
192. Macdonald, L. E., Zhou, Y. & McAllister, W. T. (1993). Termination and slippage by bacteriophage T7 RNA polymerase. *J. Mol. Biol.* **232**, 1030-47.
193. Magyar, A., Zhang, X., Kohn, H. & Widger, W. R. (1996). The antibiotic bicyclomycin affects the secondary RNA binding site of *Escherichia coli* transcription termination factor Rho. *J. Biol. Chem.* **271**, 25369-25374
194. Mah, T. F., Kuznedelov, K., Mushegian, A., Severinov, K. & Greenblatt J. (2000). The alpha subunit of *E. coli* RNA polymerase activates RNA binding by NusA. *Genes Dev.* **14**, 2664-2675.
195. Mah, T. F., Li, J., Davidson, A. R. & Greenblatt, J. (1999). Functional importance of regions in *Escherichia coli* elongation factor NusA that interact with RNA polymerase, the bacteriophage lambda N protein and RNA. *Mol. Microbiol.* **34**, 523-537.
196. Marciniak, R. A. & Sharp, P. A. (1991). HIV-1 Tat protein promotes formation of more-processive elongation complexes. *EMBO J.* **10**, 4189-4196.
197. Martin, F. H. & Tinoco, I. Jr. (1980). DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. *Nucleic Acids Res.* **8**, 2295-9.
198. Matsuzaki, H., Kassavetis, G. A. & Geiduschek, E. P. (1994). Analysis of RNA chain elongation and termination by *Saccharomyces cerevisiae* RNA polymerase III. *J. Mol. Biol.* **235**, 1173-1192.
199. Mason, S. W., Li, J. & Greenblatt, J. (1992). Host factor requirements for processive antitermination of transcription and suppression of pausing by the N protein of bacteriophage lambda. *J. Biol. Chem.* **267**, 19418-19426.
200. Mason, S. W., Li, J. & Greenblatt, J. (1992). Direct interaction between two *Escherichia coli* transcription antitermination factors, NusB and ribosomal protein S10. *J. Mol. Biol.* **223**, 55-66.
201. Mason, S. W. & Greenblatt, J. (1991). Assembly of transcription elongation complexes containing the N protein of phage lambda and the *Escherichia coli* elongation factors NusA, NusB, NusG, and S10. *Genes Dev.* **5**, 1504-1512.

Références.

202. Mattick, J. S. (2004). RNA regulation: a new genetics? *Nat. Rev. Genet.* **5**, 316-323.
203. Mattick, J. S. & Gagen, M. J. (2001). The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. *Mol. Biol. Evol.* **18**, 1611-1630.
204. McClain, W. H., Foss, K., Jenkins, R. A. & Schneider, J. (1991). Rapid determination of nucleotides that define tRNA(Gly) acceptor identity. *Proc. Natl. Acad. Sci. U S A*, **88**, 6147-6151.
205. McSwiggen, J. A., Bear, D. G. & von Hippel, P. H. (1988). Interactions of Escherichia coli transcription termination factor rho with RNA. I. Binding stoichiometries and free energies. *J. Mol. Biol.* **199**, 609-622.
206. Miller, J. H. (1972). In *Experiments in Molecular Genetics*. (Miller, J. H. ed), pp352-355, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.
207. Modrak, D. & Richardson, J. P. (1994). The RNA-binding domain of transcription termination factor rho: isolation, characterization, and determination of sequence limits. *Biochemistry*, **33**, 8292-8299.
208. Mogridge, J., Legault, P., Li, J., Van Oene, M. D., Kay, L. E. & Greenblatt, J. (1998). Independent ligand-induced folding of the RNA-binding domain and two functionally distinct antitermination regions in the phage lambda N protein. *Mol. Cell.* **1**, 265-275.
209. Mogridge, J., Mah, T. F. & Greenblatt, J. (1995). A protein-RNA interaction network facilitates the template-independent cooperative assembly on RNA polymerase of a stable antitermination complex containing the lambda N protein. *Genes Dev.* **9**, 2831-2845.
210. Mooney, R. A., Artsimovitch, I. & Landick, R. (1998). Information processing by RNA polymerase: recognition of regulatory signals during RNA chain elongation. *J Bacteriol.* **180**, 3265-3275.
211. Morgan, E. A. (1980). Insertions of Tn 10 into an E. coli ribosomal RNA operon are incompletely polar. *Cell*, **21**, 257-265.
212. Morgan WD, Bear DG, von Hippel PH. (1983). Rho-dependent termination of transcription. I. Identification and characterization of termination sites for transcription from the bacteriophage lambda PR promoter. *J. Biol. Chem.* **258**, 9553-9564.
213. Morgan, W. D., Bear, D. G. & von Hippel, P. H. (1983). Rho-dependent termination of transcription. II. Kinetics of mRNA elongation during transcription from the bacteriophage lambda PR promoter. *J. Biol. Chem.* **258**, 9565-9574.

Références.

214. Morgan, W. D., Bear, D. G., Litchman, B. L. & von Hippel, P. H. (1985). RNA sequence and secondary structure requirements for rho-dependent transcription termination. *Nucl. Acids Res.* **13**, 3739-3754.
215. Mörl, M. & Marchfelder, A. (2001). The final cut. The importance of tRNA 3'-processing. *EMBO Rep.* **2**, 17-20.
216. Mosrin-Huaman, C., Turnbough, C. L. Jr. & Rahmouni, A.R. (2004). Translocation of Escherichia coli RNA polymerase against a protein roadblock in vivo highlights a passive sliding mechanism for transcript elongation. *Mol. Microbiol.* **51**, 1471-1481.
217. Murakami, K. S. & Darst, S. A. (2003). Bacterial RNA polymerases: the whole story. *Curr. Opin. Struct. Biol.* **13**, 31-39.
218. Neely, M. N. & Friedman, D. I. (2000). N-mediated transcription antitermination in lambdoid phage H-19B is characterized by alternative NUT RNA structures and a reduced requirement for host factors. *Mol. Microbiol.* **38**, 1074-1085.
219. Nehrke, K. W. & Platt, T. A. (1994). Quaternary transcription termination complex. Reciprocal stabilization by Rho factor and NusG protein. *J. Mol. Biol.* **243**, 830-839.
220. Nodwell, J. R. & Greenblatt, J. (1991). The nut site of bacteriophage lambda is made of RNA and is bound by transcription antitermination factors on the surface of RNA polymerase. *Genes Dev.* **5**, 2141-51.
221. Nodwell, J. R. & Greenblatt, J. (1993). Recognition of boxA antiterminator RNA by the E. coli antitermination factors NusB and ribosomal protein S10. *Cell*, **72**, 261-268.
222. Normanly, J., Masson, J. M., Kleina, L. G., Abelson, J., Miller, J. H. (1986). Construction of two Escherichia coli amber suppressor genes: tRNAPheCUA and tRNACysCUA. *Proc. Natl. Acad. Sci. U S A*, **83**, 6548-6552.
223. Nudler, E. (1999). Transcription elongation: structural basis and mechanisms. *J. Mol. Biol.* **288**, 1-12.
224. Nudler, E., Gusarov, I., Avetissova, E., Kozlov, M. & Goldfarb, A. (1998). Spatial organization of transcription elongation complex in Escherichia coli. *Science*, **281**, 424-8.
225. Nudler, E. & Gottesman, M. E. (2002). Transcription termination and anti-termination in E. coli. *Genes Cells*, **7**, 755-768.
226. Opperman, T. & Richardson, J. P. (1994). Phylogenetic analysis of sequences from diverse bacteria with homology to the Escherichia coli rho gene. *J. Bacteriol.* **176**, 5033-5043.

Références.

227. Olson, E. R., Flamm, E. L. & Friedman, D. I. (1982). Analysis of nutR: a region of phage lambda required for antitermination of transcription. *Cell*, **31**, 61-70.
228. Oubridge, C., Ito, N., Evans, P. R., Teo, C. H. & Nagai, K. (1994). Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature*, **372**, 432-438.
229. Orlova, M., Newlands, J., Das, A., Goldfarb, A. & Borukhov, S. (1995). Intrinsic transcript cleavage activity of RNA polymerase. *Proc. Natl. Acad. Sci. U S A*, **92**, 4596-4600.
230. Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J. & Conklin, D. S. (2002). Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev.* **16**, 948-958.
231. Pal-Bhadra, M., Bhadra, U. & Birchler, J. A. Cosuppression in *Drosophila*: gene silencing of Alcohol dehydrogenase by white-Adh transgenes is Polycomb dependent. *Cell*, **90**, 479-490.
232. Paraskeva, E., Atzberger, A. & Hentze, M. W. (1998). A translational repression assay procedure (TRAP) for RNA-protein interactions in vivo. *Proc. Natl. Acad. Sci. U S A*, **95**, 951-956.
233. Parrot, A. M., Lago, H., Adams, C. J., Ashcroft, A. E., Stonehouse, N. J. & Stockley, P. G. (2000). RNA aptamers for the MS2 bacteriophage coat protein and the wild-type RNA operator have similar solution behaviour. *Nucl. Acids Res.* **28**, 489-497.
234. Paman, Z., von Hippel, P. H. (2000). Regulation of rho-dependent transcription termination by NusG is specific to the Escherichia coli elongation complex. *Biochemistry*, **39**, 5573-5585.
235. Patel, S. S. & Picha, K. M. (2000). Structure and function of hexameric helicases. *Annu. Rev. Biochem.* **69**, 651-697.
236. Peabody, D. S. (1990). Translational repression by bacteriophage MS2 coat protein expressed from a plasmid. A system for genetic analysis of a protein-RNA interaction. *J. Biol. Chem.* **265**, 5684-5689.
237. Peabody, D. S. (1993). The RNA binding site of bacteriophage MS2 coat protein. *EMBO J.* **12**, 595-600.
238. Peabody, D. S. & Ely, K. R. (1992). Control of translational repression by protein-protein interactions. *Nucleic Acids Res.* **20**, 1649-1655.
239. Peled-Zehavi, H., Horiya, S., Das, C., Harada, K. & Frankel, A. D. (2003). Selection of RRE RNA binding peptides using a kanamycin antitermination assay. *RNA*, **9**, 252-261.

Références.

240. Peled-Zehavi, H., Smith, C. A., Harada, K. & Frankel, A. D. (2000). Screening RNA-binding libraries by transcriptional antitermination in bacteria. *Methods Enzymol.* **318**, 297-308.
241. Pestka, S., Daugherty, B. L., Jung, V., Hotta, K. & Pestka, R. K. (1984). Anti-mRNA: specific inhibition of translation of single mRNA molecules. *Proc. Natl. Acad. Sci. U S A*, **81**, 7525-7528.
242. Pinkham, J. L. & Platt, T. (1983). The nucleotide sequence of the rho gene of E. coli K-12. *Nucleic Acids Res.* **11**, 3531-3545.
243. Platt, T. (1994). Rho and RNA: models for recognition and response. *Mol. Microbiol.* **11**, 983-990.
244. Platt, T. & Richardson, J. P. (1992). Escherichia coli Rho factor: protein and enzyme of transcription termination. In *transcriptional regulation* (McKnight, S. L. & Yamamoto, K. R. eds), pp. 365-388, Cold Spring Harbor Laboratory Press, Cold spring Harbor, N. Y.
245. Proudfoot, N. J. (1989). How RNA polymerase II terminates transcription in higher eukaryotes. *Trends Biochem. Sci.* **14**, 105-110.
246. Puttaraju, M., Jamison, S. F., Mansfield, S. G., Garcia-Blanco, M. A. & Mitchell, L. G. (1999). Spliceosome-mediated RNA trans-splicing as a tool for gene therapy. *Nat. Biotechnol.* **17**, 246-252.
247. Query, C. C., Bentley, R. C. & Keene, J. D. (1989). A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein. *Cell*, **57**, 89-101.
248. Quirk, P. G., Dunkley, E. A. Jr., Lee, P. & Krulwich, T. A. (1993). Identification of a putative Bacillus subtilis rho gene. *J. Bacteriol.* **175**, 647-654.
249. Richardson, J. P. (1982). Activation of rho protein ATPase requires simultaneous interaction at two kinds of nucleic acid-binding sites. *J. Biol. Chem.* **257**, 5760-5766.
250. Richardson, J. P. (1990). Rho-dependent transcription termination. *Biochim. Biophys. Acta.* **1048**, 127-138.
251. Richardson, J. P. (1991). Preventing the synthesis of unused transcripts by Rho factor. *Cell*, **64**, 1047-1049.
252. Richardson, J. P. (1996). Structural organization of transcription termination factor Rho. *J. Biol. Chem.* **271**, 1251-1254.

Références.

253. Richardson, J. P. (2002). Rho-dependent termination and ATPases in transcript termination. *Biochim. Biophys. Acta.* **1577**, 251-260.
254. Richardson, J. P. (2003). Loading Rho to terminate transcription. *Cell*, **114**, 157-159.
255. Richardson, J. P. & Greenblatt, J. (1996). Control of RNA chain elongation and termination. In *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Neidhardt, F. C. et al, eds), pp. 822-848, American Society for Microbiology, Washington, D. C.
256. Richardson, L. V. & Richardson, J. P. (1992). Cytosine nucleoside inhibition of the ATPase of *Escherichia coli* termination factor rho: evidence for a base specific interaction between rho and RNA. *Nucleic Acids Res.* **20**, 5383-5387.
257. 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. *J. Biol. Chem.* **271**, 21597-21603.
258. Ring, B. Z., Yarnell, W. S. & Roberts, J. W. (1996). Function of *E. coli* RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. *Cell*, **86**, 485-493.
259. Rivellini, F., Alifano, P., Piscitelli, C., Blasi, V., Bruni, C. B. & Carlomagno, M. S. (1991). A cytosine-over guanosine-rich sequence in RNA activates rho-dependent transcription termination. *Mol. Microbiol.* **5**, 3049-3054.
260. Robert, J., Sloan, S. B., Weisberg, R. A., Gottesman, M. E., Robledo, R. & Harbrecht, D. (1987). The remarkable specificity of a new transcription termination factor suggests that the mechanisms of termination and antitermination are similar. *Cell*, **51**, 483-492.
261. Roberts, J. W. (1969). Termination factor for RNA synthesis. *Nature*, **224**, 1168-1174.
262. Roberts, J. W. (1988). Phage lambda and the regulation of transcription termination. *Cell*, **52**, 5-6.
263. Roberts, J. W. (1993). RNA and protein elements of *E. coli* and lambda transcription antitermination complexes. *Cell*, **72**, 653-655.
264. Roberts, E. A., Eisenbraun, T. L., Andrews, C. L. & Bear D. G. (1991). 3'-end formation at the phage lambda tR1 rho-dependent transcription termination site. *Biochemistry*, **30**, 5429-5437.
265. Roberts, J. W., Yarnell, W., Bartlett, E., Guo, J., Marr, M., Ko, D. C., Sun, H. & Roberts, C. W. (1998). Antitermination by bacteriophage lambda Q protein. *Cold Spring Harb. Symp. Quant. Biol.* **63**, 319-25.

Références.

266. Robinson, R. (2004). RNAi Therapeutics: How Likely, How Soon? *PloS. Biol.* **2**,:18-20.
267. Robledo, R., Atkinson, B. L. & Gottesman, M. E. (1991). Escherichia coli mutations that block transcription termination by phage HK022 Nun protein. *J. Mol. Biol.* **220**, 613-619.
268. Rosenberg, M., Court, D., Shimatake, H., Brady, C. & Wulff, D. L. (1978). The relationship between function and DNA sequence in an intercistronic regulatory region in phage lambda. *Nature*, **272**, 414-423.
269. Rowsell, S., Stonehouse, N. J., Convery, M. A., Adams, C. J., Ellington, A. D., Hirao, I., Peabody, D. S., Stockley, P. G. & Phillips, S. E. (1998). Crystal structures of a series of RNA aptamers complexed to the same protein target. *Nat. Struct. Biol.* **5**, 970-975.
270. Roy, M. K. & Apirion, D. (1983). Purification and properties of ribonuclease E, an RNA-processing enzyme from Escherichia coli. *Biochim. Biophys. Acta.* **747**, 200-208.
271. Roy, M. K., Singh, B., Ray, B. K. & Apirion, D. (1983). Maturation of 5-S rRNA: ribonuclease E cleavages and their dependence on precursor sequences. *Eur. J. Biochem.* **131**, 119-127.
272. Ruteshouser, E. C. & Richardson, J. P. (1989). Identification and characterization of transcription termination sites in the Escherichia coli lacZ gene. *J. Mol. Biol.* **208**, 23-43.
273. Ryder, A. M. & Roberts, J. W. (2003). Role of the non-template strand of the elongation bubble in intrinsic transcription termination. *J. Mol. Biol.* **334**, 205-213.
274. Saecker, R. M., Tsodikov, O. V., McQuade, K. L., Schlax, P. E. Jr., Capp, M. W. & Record, M. T. Jr. (2002). Kinetic studies and structural models of the association of E. coli sigma(70) RNA polymerase with the lambdaP(R) promoter: large scale conformational changes in forming the kinetically significant intermediates. *J. Mol. Biol.* **319**, 649-671.
275. Salstrom, J. S. & Szybalski, W. (1978). Coliphage lambda^{nutL-}: a unique class of mutants defective in the site of gene N product utilization for antitermination of leftward transcription. *J. Mol. Biol.* **124**, 195-221.
276. Scherly, D., Boelens, W., Dathan, N. A., van Venrooij, W. J. & Mattaj, I. W. (1990). Major determinants of the specificity of interaction between small nuclear ribonucleoproteins U1A and U2B" and their cognate RNAs. *Nature*, **345**, 502-506.
277. Schmidt, M. C. & Chamberlin, M. J. (1984). Binding of rho factor to Escherichia coli RNA polymerase mediated by nusA protein. *J. Biol. Chem.* **259**, 15000-15002.

Références.

278. Schwartz, A., Rahmouni, A. R. & Boudvillain, M. (2003). The functional anatomy of an intrinsic transcription terminator. *EMBO J.* **22**, 3385-3394.
279. Segawa, T. & Imamoto, F. (1974). Diversity of regulation of genetic transcription. II. Specific relaxation of polarity in read-through transcription of the translocated trp operon in bacteriophage lambda trp. *J. Mol. Biol.* **87**, 741-754.
280. Selby, M. J. & Peterlin, B. M. (1990). Trans-activation by HIV-1 Tat via a heterologous RNA binding protein. *Cell*, **62**, 769-776.
281. Selby, C. P. & Sancar, A. (1993). Molecular mechanism of transcription-repair coupling. *Science*, **260**, 53-58.
282. Selby, C. P. & Sancar, A. (1995a). Structure and function of transcription-repair coupling factor. I. Structural domains and binding properties. *J. Biol. Chem.* **270**, 4882-4889.
283. Selby, C. P. & Sancar, A. (1995b). Structure and function of transcription-repair coupling factor. II. Catalytic properties. *J. Biol. Chem.* **270**, 4890-4895.
284. Sen, R., King, R. A. & Weisberg, R. A. (2001). Modification of the properties of elongating RNA polymerase by persistent association with nascent antiterminator RNA. *Mol. Cell*, **7**, 993-1001
285. Sloan, S. B. & Weisberg, R. A. (1993). Use of a gene encoding a suppressor tRNA as a reporter of transcription: analyzing the action of the Nun protein of bacteriophage HK022. *Proc. Natl Acad. Sci. USA*, **90**, 9842-9846.
286. Skordalakes, E. & Berger, J. M. (2003). Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell*, **114**, 135-146.
287. Somasekhar, G. & Szybalski, W. (1983). Mapping of the Q-utilization site (qut) required for antitermination of late transcription in bacteriophage lambda. *Gene*, **26**, 291-294.
288. Sorin, E. J., Engelhardt, M. A., Herschlag, D. & Pande, V. S. (2002). RNA simulations: probing hairpin unfolding and the dynamics of a GNRA tetraloop. *J. Mol. Biol.* **317**, 493-506.
289. Southgate, C. D. & Green, M. R. (1991). The HIV-1 Tat protein activates transcription from an upstream DNA-binding site: implications for Tat function. *Genes Dev.* **5**, 2496-2507.
290. Sozhamannan, S. & Stitt, B. L. (1997). Effects on mRNA degradation by Escherichia coli transcription termination factor Rho and pBR322 copy number control protein Rop. *J. Mol. Biol.* **268**, 689-703.

Références.

291. Squires, C. L., Greenblatt, J., Li, J., Condon, C. & Squires C. L. (1993). Ribosomal RNA antitermination in vitro: requirement for Nus factors and one or more unidentified cellular components. *Proc. Natl Acad. Sci. USA*, **90**, 970-974.
292. Stackhouse, T. M., Telesnitsky, A. P. & Meares, C. F. (1989). Release of the sigma subunit from Escherichia coli RNA polymerase transcription complexes is dependent on the promoter sequence. *Biochemistry*, **28**, 7781-7788.
293. Stanssens, P. Remaut, E. & Fiers, W. (1986). Inefficient translation initiation causes premature transcription termination in the lacZ gene. *Cell*, **44**, 711-718.
294. Steinmetz, E. J., Brennan, C. A. & Platt, T. (1990). A short intervening structure can block rho factor helicase action at a distance. *J. Biol. Chem.* **265**, 18408-18413.
295. Steinmetz, E. J. & Platt, T. (1994). Evidence supporting a tethered tracking model for helicase activity of Escherichia coli Rho factor. *Proc. Natl. Acad. Sci. U S A*, **91**, 1401-1405.
296. Stitt, B. L. (2001). Escherichia coli transcription termination factor Rho binds and hydrolyzes ATP using a single class of three sites. *Biochemistry*, **40**, 2276-2281.
297. Stitt, B. L. & Mosig, G. (1989). Impaired expression of certain prereplicative bacteriophage T4 genes explains impaired T4 DNA synthesis in Escherichia coli rho (nusD) mutants. *J. Bacteriol.* **171**, 3872-3880.
298. Stitt, B. L. & Xu, Y. (1998). Sequential hydrolysis of ATP molecules bound in interacting catalytic sites of Escherichia coli transcription termination protein Rho. *J. Biol. Chem.* **273**, 26477-26486.
299. Storz, G. (2002). An expanding universe of noncoding RNAs. *Science*, **296**, 1260-1263.
300. Stulke, J. (2002). Control of transcription termination in bacteria by RNA-binding proteins that modulate RNA structures. *Arch. Microbiol.* **177**, 433-440.
301. Su, L., Radek, J. T., Labeets, L. A., Hallenga, K., Hermanto, P., Chen, H., Nakagawa, S., Zhao, M., Kates, S. & Weiss, M. A. (1997). An RNA enhancer in a phage transcriptional antitermination complex functions as a structural switch. *Genes Dev.* **11**, 2214-2226.
302. Sullenger, B. A. & Cech, T. R. (1994). Ribozyme-mediated repair of defective mRNA by targeted, trans-splicing. *Nature*, **371**, 619-622.
303. Sullenger, B. A., Gallardo, H. F., Ungers, G. E. & Gilboa, E. (1990). Overexpression of TAR sequences renders cells resistant to human immunodeficiency virus replication. *Cell*, **63**, 601-608

Références.

304. Sullenger, B. A. & Gilboa, E. (2002). Emerging clinical applications of RNA. *Nature*, **418**, 252-258.
305. Sullivan, S. L. & Gottesman, M. E. (1992). Requirement for E. coli NusG protein in factor-dependent transcription termination. *Cell*, **68**, 989-994.
306. Tan, R. & Frankel, A. D. (1998). A novel glutamine-RNA interaction identified by screening libraries in mammalian cells. *Proc. Natl. Acad. Sci. U S A*, **95**, 4247-4252.
307. Torres, M., Condon, C., Balada, J. M., Squires, C. & Squires, C. L. (2001). Ribosomal protein S4 is a transcription factor with properties remarkably similar to NusA, a protein involved in both non-ribosomal and ribosomal RNA antitermination. *EMBO J.* **20**, 3811-3820.
308. Toulmé, F., Guérin, M., Robichon, N., Leng, M. & Rahmouni, A. R. (1999). In vivo evidence for back and forth oscillations of the transcription elongation complex. *EMBO J.* **18**, 5052-5060.
309. Toulmé, F., Mosrin-Huaman, C., Sparkowski, J., Das, A., Leng, M. & Rahmouni, A. R. (2000). GreA and GreB proteins revive backtracked RNA polymerase in vivo by promoting transcript trimming. *EMBO J.* **19**, 6853-6859.
310. Toulokhonov, I. & Landick, R. (2003). The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination. *Mol. Cell*, **12**, 1125-1136.
311. Tuerk, C. & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, **249**, 505-510.
312. Uhlenbeck, O. C. (1987). A small catalytic oligoribonucleotide. *Nature*, **328**, 596-600.
313. Unniraman, S., Prakash, R. & Nagaraja, V. (2002). Conserved economics of transcription termination in eubacteria. *Nucleic Acids Res.* **30**, 675-684.
314. Uptain, S. M. & Chamberlin, M. J. (1997). Escherichia coli RNA polymerase terminates transcription efficiently at rho-independent terminators on single-stranded DNA templates. *Proc. Natl. Acad. Sci. U S A*, **94**, 13548-13553.
315. Valegard, K., Murray, J. B., Stonehouse, N. J., van den Worm, S., Stockley, P. G. & Liljas, L. (1997). The three-dimensional structures of two complexes between recombinant MS2 capsids and RNA operator fragments reveal sequence-specific protein-RNA interactions. *J. Mol. Biol.* **270**, 724-738.
316. van der Krol, A. R., Mol, J. N. & Stuitje, A. R. (1988). Modulation of eukaryotic gene expression by complementary RNA or DNA sequences. *Biotechniques*, **6**, 958-976.

Références.

317. Van Gilst, M. R. & von Hippel, P. H. (1997). Assembly of the N-dependent antitermination complex of phage lambda: NusA and RNA bind independently to different unfolded domains of the N protein. *J. Mol. Biol.* **274**, 160-173.
318. View, E. & Rahmouni, A. R. (2004). Dual Role of boxB RNA Motif in the Mechanisms of Termination/Antitermination at the Lambda tR1 Terminator Revealed in vivo. *J. Mol. Biol.* **339**, 1077-1087.
319. Vo, N. V., Hsu, L. M., Kane, C. M. & Chamberlin, M. J. (2003). In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 3. Influences of individual DNA elements within the promoter recognition region on abortive initiation and promoter escape. *Biochemistry*, **42**, 3798-3811.
320. Vogel, U. & Jensen, K. F. (1994). The RNA chain elongation rate in Escherichia coli depends on the growth rate. *J. Bacteriol.* **176**, 2807-2813.
321. Vogel, U. & Jensen, K. F. (1997). NusA is required for ribosomal antitermination and for modulation of the transcription elongation rate of both antiterminated RNA and mRNA. *J. Biol. Chem.* **272**, 12265-12271.
322. von Hippel, P. H. (1998). An integrated model of the transcription complex in elongation, termination, and editing. *Science*, **281**, 660-665.
323. Walstrom, K. M., Dozono, J. M. & von Hippel, P. H. (1998). Effects of reaction conditions on RNA secondary structure and on the helicase activity of Escherichia coli transcription termination factor Rho. *J. Mol. Biol.* **279**, 713-726.
324. Walstrom, K. M., Dozono, J. M., Robic, S. & von Hippel, P. H. (1997a). Kinetics of the RNA-DNA helicase activity of Escherichia coli transcription termination factor rho. 1. Characterization and analysis of the reaction. *Biochemistry*, **36**, 7980-7992.
325. Walstrom, K. M., Dozono, J. M. & von Hippel, P. H. (1997b). Kinetics of the RNA-DNA helicase activity of Escherichia coli transcription termination factor rho. 2. Processivity, ATP consumption, and RNA binding. *Biochemistry*, **36**, 7993-8004.
326. Wang, Y. & von Hippel, P. H. (1993a). Escherichia coli transcription termination factor rho. I. ATPase activation by oligonucleotide cofactors. *J. Biol. Chem.* **268**, 13940-13946.
327. Wang, Y. & von Hippel, P. H. (1993b). Escherichia coli transcription termination factor rho. II. Binding of oligonucleotide cofactors. *J. Biol. Chem.* **268**, 13947-13955.
328. Watnick, R. S. & Gottesman, M. E. (1998). Escherichia coli NusA is required for efficient RNA binding by phage HK022 nun protein. *Proc. Natl. Acad. Sci. U S A*, **95**, 1546-1551.

Références.

329. Watnick, R. S. & Gottesman, M. E. (1999). Binding of transcription termination protein nun to nascent RNA and template DNA. *Science*. **286**, 2337-2339. Watnick, R. S., Herring, S. C., Palmer, A. G. 3rd. & Gottesman, M. E. (2000). The carboxyl terminus of phage HK022 Nun includes a novel zinc-binding motif and a tryptophan required for transcription termination. *Genes Dev.* **14**, 731-739.
331. Washburn, R. S. & Stitt, B. L. (1996). In vitro characterization of transcription termination factor Rho from *Escherichia coli* rho(nusD) mutants. *J. Mol. Biol.* **260**, 332-346.
332. Washburn, R. S., Wang, Y. & Gottesman, M. E. (2003). Role of *E. coli* transcription-repair coupling factor Mfd in Nun-mediated transcription termination. *J. Mol. Biol.* **329**, 655-662.
333. Wei, R. R. & Richardson, J. P. (2001b). Mutational changes of conserved residues in the Q-loop region of transcription factor Rho greatly reduce secondary site RNA-binding. *J. Mol. Biol.* **314**, 1007-1015.
334. Wei, R. R. & Richardson, J. P. (2001a). Identification of an RNA-binding Site in the ATP binding domain of *Escherichia coli* Rho by H₂O₂/Fe-EDTA cleavage protection studies. *J. Biol. Chem.* **276**, 28380-28387.
335. Weisberg, R. A. & Gottesman, M. E. (1999). Processive antitermination. *J. Bacteriol.* **181**, 359-367.
336. Werstuck, G. & Green, M. R. (1998). Controlling gene expression in living cells through small molecule-RNA interactions. *Science*, **282**, 296-298.
337. Wilson, H. R., Yu, D., Peters, H. K. 3rd., Zhou, J. G. & Court, D. L. (2002). The global regulator RNase III modulates translation repression by the transcription elongation factor N. *EMBO. J.* **21**, 4154-4161
338. Wu, S. Y. & Platt, T. (1993). Transcriptional arrest of yeast RNA polymerase II by *Escherichia coli* rho protein in vitro. *Proc. Natl. Acad. Sci. USA*, **90**, 6606-6610.
339. Yager, T. D. & von Hippel, P. H. (1987). Transcript elongation in *Escherichia Coli*. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (Neidhardt, F. C. et al, eds), pp. 1241-1275, American Society for Microbiology, Washington, D. C.
340. Yang, X. J., Hart, C. M., Grayhack, E. J. & Roberts, J. W. (1987). Transcription antitermination by phage lambda gene Q protein requires a DNA segment spanning the RNA start site. *Genes Dev.* **1**, 217-226.
341. Yanofsky, C. (2000). Transcription attenuation: once viewed as a novel regulatory strategy. *J. Bacteriol.* **182**, 1-8.

Références.

342. Yarnell, W. S. & Roberts, J. W. (1992). The phage lambda gene Q transcription antiterminator binds DNA in the late gene promoter as it modifies RNA polymerase. *Cell*, **69**, 1181-1189.
343. Yarnell, W. S. & Roberts, J. W. (1999). Mechanism of intrinsic transcription termination and antitermination. *Science*, **284**, 611-615.
344. Ye, X., Gorin, A., Ellington, A. D. & Patel, D. J. (1996). Deep penetration of an alpha-helix into a widened RNA major groove in the HIV-1 rev peptide-RNA aptamer complex. *Nat. Struct. Biol.* **3**, 1026-1033.
345. Yu, X., Horiguchi, T., Shigesada, K. & Egelman, E. H. (2000). Three-dimensional reconstruction of transcription termination factor rho: orientation of the N-terminal domain and visualization of an RNA-binding site. *J. Mol. Biol.* **299**, 1279-1287.
346. Zalatan, F., Galloway-Salvo, J. & Platt, T. (1993). Deletion analysis of the Escherichia coli rho-dependent transcription terminator trp t'. *J. Biol. Chem.* **268**, 17051-17056.
347. Zheng, C. & Friedman, D. I. (1994). Reduced Rho-dependent transcription termination permits NusA-independent growth of Escherichia coli. *Proc. Natl. Acad. Sci. U S A*, **91**, 7543-7547.
348. Zhu, A. Q. & von Hippel, P. H. (1998a). Rho-dependent termination within the trp t' terminator. I. Effects of rho loading and template sequence. *Biochemistry*, **37**, 11202-11214.
349. Zhu, A. Q. & von Hippel, P. H. (1998b). Rho-dependent termination within the trp t' terminator. II. Effects of kinetic competition and rho processivity. *Biochemistry*, **37**, 11215-11222.
350. Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucl. Acids Res.* **31**, 3406-3415.