

Etude de l'endoribonucléase de restriction RegB.

Fakhri Saïda

▶ **To cite this version:**

Fakhri Saïda. Etude de l'endoribonucléase de restriction RegB.. Sciences du Vivant [q-bio]. Ecole Polytechnique X, 2003. Français. NNT : . pastel-00000707

HAL Id: pastel-00000707
https://pastel.hal.science/pastel-00000707

Submitted on 21 Jul 2010

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.

ÉCOLE POLYTECHNIQUE

THÈSE de DOCTORAT

Spécialité : Biochimie

présentée par

Fakhri SAÏDA

Pour obtenir le grade de DOCTEUR EN SCIENCES
de l'ÉCOLE POLYTECHNIQUE

Étude de l'endoribonucléase de restriction RegB

Directeur de thèse : M. François BONTEMS

Soutenue le 29 octobre 2003 devant la commission d'examen :

M.	Marc Uzan	Président
M.	Harald Putzer	Rapporteur
Mme	Sophie Zinn-Justin	Rapporteur
M.	Jean-Yves Lallemand	Examinateur
M.	François Bontems	Examinateur

Remerciements

Ce travail a été réalisé dans le laboratoire ICSN-RMN de l'Ecole polytechnique dirigé par le professeur Jean-Yves Lallemand. Je tiens à lui exprimer toute ma gratitude pour m'avoir donné la chance de faire cette thèse ainsi que pour son soutien constant durant toute la période passée au laboratoire.

Le travail présenté a été dirigé par le docteur François Bontems. Sa culture pluridisciplinaire et sa bonne humeur constante ont été d'un grand secours pour surmonter certains des moments difficiles de cette thèse. Qu'il en soit vivement remercié.

J'exprime ma reconnaissance envers le docteur Marc Uzan (Institut Jacques Monod, Paris) pour sa rigueur scientifique, pour ses critiques nombreuses mais très utiles, pour avoir accepté d'examiner cette thèse et enfin pour avoir joué un rôle déterminant dans le choix de mon programme post-doctoral.

Je remercie le docteur Harald Putzer (IBPC, Paris) et le docteur Sophie Zinn-Justin (CEA, Saclay) d'avoir accepté de rapporter cette thèse.

Je remercie les personnes que j'ai côtoyées à la paillasse : Benoît Odaert, Jean-Baptiste Duvignaud, Francis Duffieux, Olivier Pamlard, John K. Smith, Jean-Bernard Créchet, Eric Jacquet et Philippe Lopes.

Je remercie également Alain Louis-Joseph, Nadia Brocard, Bayesté Leclaire, Soumaya Laalami, Paulo Loreiro de Sousa, Maurice Guéron, Jean-Louis Leroy et Karim Snoussi.

Enfin, je reste reconnaissant envers le docteur Gérard Faye (Institut Curie, Orsay) qui, le premier, m'a fait découvrir avec passion la recherche biologique.

Je dédie cette thèse à ma famille.

Palaiseau, le 30 octobre 2003.

Table des matières

Introduction 7

1.1 L'activité ribonucléase est associée à quatre niveaux d'organisation macromoléculaire 9

1.1.1 Une simple chaîne ribonucléique 9

1.1.2 Une chaîne ribonucléique associée à une chaîne peptidique : la RNase P 11

1.1.3 Une seule chaîne polypeptidique 13

1.1.4 Un complexe multiprotéique : Exemple du dégradosome d'*E. coli* et de l'exosome de la *S. cerevisiae* 15

1.2 La spécificité de coupure chez les ribonucléases 17

1.2.1 Clivage d'une classe particulière d'ARN 17

1.2.2 Clivage d'un ARN simple brin ou d'un ARN double brin 17

1.2.3 Clivage d'un ARN dans une structure nucléique hybride 19

1.2.4 Clivage spécifique après une base particulière 19

1.2.5 Clivage au niveau d'un contexte séquentiel particulier 20

1.2.5.1 La colicine E5 : une RNAse de restriction qui reconnaît le motif Q(G)U 21

1.2.5.2 La restrictocine 22

1.3 L'endoribonucléase de restriction RegB 24

1.3.1 Le cycle lytique du phage T4 dans *E. coli* 24

1.3.2 La découverte du gène *regB* 25

1.3.3 RegB est une endoribonucléase séquence-spécifique 26

1.3.4 L'activité de RegB est stimulée *in vitro* par la protéine ribosomique S1 27

1.3.5 Recherche de l'information supplémentaire portée par l'ARN et identifiable par RegB 27

1.4 Présentation du travail de recherche que nous avons mené sur RegB 29

Chapitre 1 Expression de la ribonucléase RegB 30

1.1 Le système d'expression pET 31

1.2 RegB ne peut être maintenue *in vivo* dans un vecteur pET classique 33

1.2.1 Preuve expérimentale 33
1.2.2 Des systèmes d'expression plus fermés sont nécessaires 35

1.3 Expression de RegB à partir d'un vecteur pET contenant un RBS Phi 10 modifié 36
1.3.1 RegB se maintient dans un vecteur pET contenant un RBS peu efficace 36
1.3.2 Adaptation du système RBS Phi 10 atténué / λCE6 à la culture en fermenteur 39

1.4 Expression de RegB en couplant la stratégie pET à une origine de réplication thermorégulable .....
1.4.1 Construction du vecteur d'expression pOURegB 41
1.4.1.1 Le nombre de copies du plasmide pOU61 est thermorégulable 41
1.4.1.2 Caractéristique du plasmide pOURegB 44
1.4.2 Expression de la ribonucléase RegB avec le système pOURegB 45

1.5 Expression de RegB dans un système eucaryote 49
1.5.1 Justification du choix 49
1.5.2 Construction des vecteurs d'expression pYRegB1 et pYRegB2 49

1.6 Avantages et inconvénients des différents systèmes d'expression de RegB 53

1.7 Mise au point d'un nouvel outil d'ingénierie génétique exploitant le système RegB 54
1.7.1 Principe 54
1.7.2 Particularité du système de sélection positive utilisant RegB 54
1.7.3 Construction du vecteur pTOXR-1 55
1.7.4 Le système de sélection positive de pTOXR-1 : Application au clonage du bimodule 4-5 de la protéine ribosomique S1 57
1.7.5 Le système d'expression duale de pTOXR-1 59
1.7.6 Perspectives d'utilisation du plasmide pTOXR-1 60

Chapitre 2 Caractérisation du chemin réactionnel catalysé par RegB 61

2.1 La liaison phosphodiester 63

2.2 Clivage de la liaison O3'-P 64

2.3 Clivage de la liaison P-O5' 66

2.3.1 Hydrolyse de la liaison P-O5' : Exemple de la nucléase staphylococcale 67
2.3.2 Clivage de la liaison P-O5' par transfert du groupement phosphate : la RNase A 68

2.4 RegB libère une extrémité 3' Phosphate 70

2.5 Suivi cinétique par RMN du ^{31}P de la réaction catalysée par RegB 71
2.5.1 Principe 71
2.5.2 Choix d'un modèle expérimental de coupure par RegB 73
2.5.3 Identification du produit de coupure 75
2.5.4 Conclusion et conséquences 77

Chapitre 3 Investigation par mutagenèse aléatoire de la structure primaire de RegB 78

3.1 La stratégie de mutagenèse aléatoire par PCR 80

3.2 Mutagenèse aléatoire de RegB et stratégie de sélection 81

3.3 Résultats de la mutagenèse de la partie 5' du gène *regB* (partie N-terminale de RegB) 85

3.4 Résultats de la mutagenèse de la partie 3' du gène *regB* (partie C-terminale de RegB) 88

Chapitre 4 Investigation par mutagenèse dirigée de la partie N-terminale de RegB 91

4.1 Résidus aminés impliqués dans le mécanisme catalytique des ribonucléases de la famille des transphosphorylases 93
4.1.1 Le groupe de la RNase A 93
4.1.2 Le groupe de la RNase T1 94
4.1.3 Le groupe de la RNase T2 96
4.1.4 Le groupe de Barnase 97

4.2 Sélection des résidus Histidines à muter 99

4.3 Mutagenèse dirigée des Histidines 42, 48, 68 et 73 100

4.4 Caractérisation *in vivo* de l'effet des mutations histidine-en-alanine 101

4.5 Etude des cinétiques de coupure de l'ARN Selex22tb par les différents mutants de RegB 102

4.6 Etude par RMN de la structure globale des mutants H48A et H68A 104

4.7 Quantification par Résonance Plasmonique de Surface de la liaison des mutants de RegB à l'ARN substrat 106

4.8 Discussion des rôles des Histidines 48 et 68 111

Chapitre 5 Evolution naturelle du gène *regB* dans un virus apparenté au phage T4 : Etude du phage RB49 112

5.1 Séquençage du gène *regB* du phage RB49 114

5.2 Comparaison des séquences des protéines RegB du phage T4 et du phage RB49 115

Chapitre 6 Etude de la structure tridimentionelle de RegB par RMN : Mise au point de la stratégie 119

6.1 Choix de la version de RegB 121

6.2 Détermination des conditions spectroscopiques d'étude de RegB H48A 123

6.2.1 La température 123

6.2.2 pH, force ionique et molécule tampon 124

6.3 Double marquage $^{13}C/^{15}N$ de RegB H48A 125

6.4 Spectres RMN préliminaires 126

Chapitre 7 Etude de la structure tridimentionelle de RegB par RMN : Début de l'attribution séquentielle de la protéine 129

7.1 La procédure globale 132

7.2 Les expériences RMN utilisées 132

7.3 Recherche des enchaînements séquentiels 134

7.3.1 Principe 134

7.3.2 Exemple d'enchaînement 135

7.4 Projection d'un enchaînement séquentiel sur la structure primaire de RegB H48A 136

7.5 Limites de la stratégie d'attribution choisie 138

Conclusion 140

Matériels et méthodes 143

Références bibliographiques 166

Annexes 184

Publications 188

Introduction

Introduction

L'acide ribonucléique (ARN) est une macromolécule biologique polymère des bases azotées uracile U ($C_4H_4N_2O_2$), cytosine C ($C_4H_5N_3O$), guanine G ($C_5H_5N_5O$) et adénine A ($C_5H_5N_5$). Cette macromolécule occupe une place centrale dans le processus de transfert de l'information génétique où elle remplie trois fonctions fondamentales distinctes. L'ARN messager (ARNm) sert de support labile de l'information génétique transcrite à partir d'un support stable (l'acide désoxyribonucléique ADN) puis traduite en protéine. L'ARN de transfert (ARNt) participe au décodage de l'information génétique portée par l'ARNm en permettant le couplage de l'unité informationnelle (le codon) à un résidu aminé particulier. Enfin, l'ARN ribosomique (ARNr) constitue près de 50 % du complexe macromoléculaire siège de la synthèse protéique : le ribosome.

L'étude du métabolisme de cette macromolécule n'a cessé de nous surprendre par sa complexité et surtout par son originalité. Avant d'illustrer cette complexité et cette originalité, il est important de définir d'abord ce que nous entendons par le terme « métabolisme » des ARN. Nous exclurons délibérément de ce terme la synthèse proprement dite de l'ARN par transcription. Nous limiterons le terme métabolisme des ARN aux processus permettant l'obtention d'une molécule d'ARN biologiquement active à partir d'un précurseur inactif (*maturation*), l'*inactivation* d'une molécule d'ARN fonctionnelle par clivage intra-séquentiel et la *dégradation* d'un ARN par clivage processif orienté (à partir d'une extrémité) aboutissant à la restitution des nucléotides de base. Autrement dit, le métabolisme des ARN sera lié à deux activités : le clivage endonucléolitique (à l'intérieur de la séquence de l'ARN) aboutissant à la maturation d'un précurseur inactif ou à l'inactivation d'un ARN fonctionnel et le clivage exonucléolitique (dégradation à partir des extrémités 3' ou 5'). On rassemblera ces deux activités sous le terme global d'activité « ribonucléase ».

L'objet de cette introduction n'est pas de faire une description qui catalogue les activités endoribonucléasiques et exoribonucléasiques qui existent chez les différents organismes procaryotes et eucaryotes. Des ouvrages de référence tels « Ribonucleases structures and fonctions » *(D'Alessio et Riordan ed., 1997)* ainsi que les volumes 341 et 342 de la série « Methods in Enzymology » présentent un panorama quasi complet des activités ribonucléases connues.

Nous nous proposons dans ce qui suit d'illustrer avec des exemples clés deux aspects importants de l'activité ribonucléase. Nous étudierons dans une première partie l'organisation macromoléculaire support de cette activité. Nous montrerons que cette dernière peut être associée à quatre niveaux d'organisation de complexité croissante. Dans une deuxième partie, nous étudierons le problème de la spécificité de coupure chez les ribonucléases. Nous terminerons, enfin, par présenter la ribonucléase de restriction RegB objet de la présente dissertation.

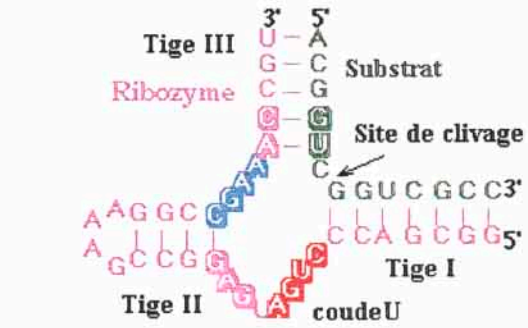
1.1 L'activité ribonucléase est associée à quatre niveaux d'organisation macromoléculaire

1.1.1 Une simple chaîne ribonucléique

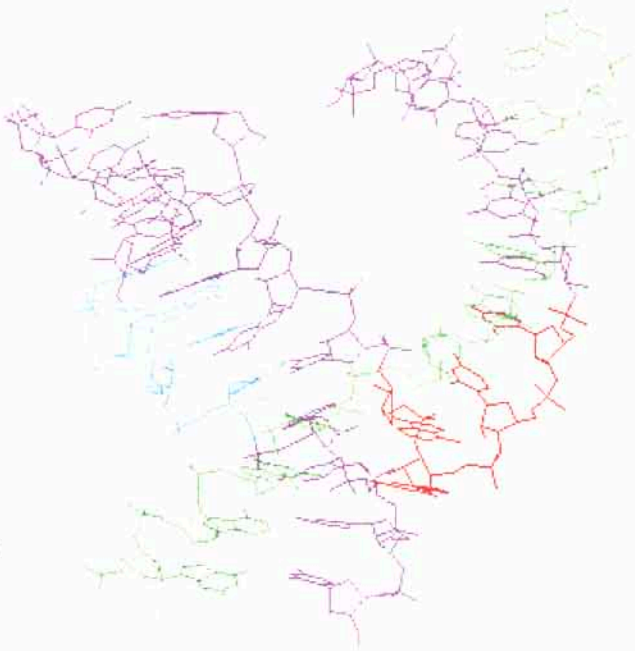
Pendant plusieurs années, l'activité enzymatique a été exclusivement associée aux protéines. On expliquait alors ce phénomène par le fait que, parmi toutes les macromolécules biologiques connues, seules les protéines possédaient des propriétés physico-chimiques qui répondent aux exigences de la catalyse chimique. En 1981, Thomas Cech (Cech *et al.*, 1981) alors qu'il étudiait la maturation d'un l'ARN ribosomique de protozoaire *(Tetrahymena thermophila)* découvrit que la réaction de coupure des introns et la mise bout à bout des exons se produit *in vitro* même en absence d'enzyme. Il conclut que c'est l'ARN précurseur lui même qui est capable de catalyser son propre épissage (coupure des introns et ligature des exons). Cette découverte mit fin au dogme « enzyme = protéine ». Depuis, plusieurs exemples d'ARN catalysant des étapes importantes dans le métabolisme des ARN ont été découverts et ont été appelés « ribozymes » par contraction des termes « ribonucléique » et « enzyme ».

Etudions un exemple de ces ribozymes : le ribozyme à tête de marteau (Forster *et al.*, 1987). Il s'agit d'un ARN de petite taille capable de s'auto-cliver en présence d'un ion divalent, en général l'ion magnésium Mg^{2+}, au niveau d'une liaison phosphodiester particulière libérant une extrémité 5'OH (Hertel *et al.*, 1992). Ce type de ribozyme existe, entre autres, dans certaines des entités biologiques les moins complexes de la nature : les viroïdes (Hutchins *et al.*, 1986). Il s'agit de « parasites moléculaires » formés uniquement de petits ARN circulaires (250-450 nucléotides) hautement structurés et dépourvus de toute fonction de messager (Diener, T. 1991). Ces parasites répliquent leurs ARN selon le modèle du cercle roulant en utilisant la machinerie enzymatique de la cellule hôte. Les ribozymes portés par ces ARN réalisent eux mêmes le clivage des brins oligomériques générés lors de cette réplication. La structure d'un exemple de ribozyme, dit à tête de

marteau, est présentée dans la Figure I (Scott *et al.*, 1995). Ce ribozyme est caractérisé par un cœur formé de nucléotides hautement conservés qui constituent le site actif. Ce cœur contient un coude appelé coude U dont le rôle principal est de maintenir l'ARN substrat (en vert sur la Figure I) dans une conformation favorable au clivage. Le site actif du ribozyme est par ailleurs entouré par trois doubles hélices (tiges I, II et III).



A



B

Figure I : Structure d'un ribozyme à tête de marteau. (A) Structure secondaire montrant en vert le substrat ARN et en rose l'essentiel de l'ARN catalytique. Les bases ombrées sont celles définies comme indispensables à l'activité. Le coude U maintient l'ARN substrat dans une conformation favorable au clivage. La liaison phosphodiester clivée est indiquée avec une flèche. (B) Structure tridimensionnelle utilisant le même code couleur que (A) (code PDB utilisé : 1MME).

Il est important de noter ici que grâce à son appariement avec l'ARN substrat, le ribozyme à tête de marteau clive très spécifiquement une liaison phosphodiester et une seule. Il démontre ainsi qu'une chaîne ribonucléique simple peut regrouper les qualités d'un catalyseur à la fois efficace et très spécifique. Dans l'exemple que nous allons présenter ci-après, nous allons voir que la catalyse et la spécificité vont être disjointes. Elles seront assurées par deux macromolécules différentes.

1.1.2 Une chaîne ribonucléique associée à une chaîne peptidique : la RNase P

Les ARN de transfert sont initialement transcrits en ARN précurseurs (pré-ARNt) qui doivent subir des étapes de clivage au niveau des extrémités 3' et 5' pour donner naissance à un ARNt fonctionnel pour la traduction. L'enzyme responsable du clivage de la partie 5' des pré-ARNt est la ribonucléase P (RNase P).

Les premières tentatives de purification de cette ribonucléase dans le laboratoire de Sydney Altman (prix Nobel de chimie 1989) aboutissaient toujours à un protéine inactive. On comprit plus tard qu'un ARN co-purifiait avec la protéine d'intérêt et que cette dernière ne présentait une activité ribonucléase *in vitro* qu'en présence de son acide ribonucléique (Guerrier-Takada *et al.*, 1983).

Chez *E. coli*, la RNase P est constituée d'une sous-unité ribonucléique appelée ARN M1 et d'une petite protéine basique, la protéine C5 qui comprend 119 résidus aminés (14 kDa). Il est établi que la catalyse de la réaction de clivage est assurée par l'ARN. Le clivage de plusieurs pré-ARNt peut se produire *in vitro* en absence de la sous-unité C5 (Guerrier-Takada *et al.*, 1983). *In vivo*, en revanche, l'activité de la RNase P nécessite la présence de la sous-unité C5 (Kirsebom L., 2001). Il est étonnant de remarquer que comparée à toutes les autres RNases connues, seule la RNase P est retrouvée chez tous les eucaryotes, chez toutes les *archaea* et chez toutes les eubactéries (à l'exception d'*Aquifex aeolicus*) (Condon *et al.*, 2002).

L'avantage précis que procure la présence de la chaîne polypeptidique à l'ARN catalytique n'est pas complètement compris. Il a été initialement proposé que la partie protéique de la RNase P réduit les répulsions électrostatiques le long de la chaîne polyanionique de l'ARN. Cette hypothèse se base sur le fait que « l'utilité » de la partie protéique de la RNase P décroît quand la force ionique de la solution augmente (Reich *et al.*, 1988). Des études ultérieures ont montré que la stabilisation de l'ARN catalytique n'est pas l'unique effet de la partie protéique. Des études cinétiques, thermodynamiques et de liaisons croisées (cross-linking) ont montré que la sous-unité protéique augmente la vitesse de relargage du produit de la réaction et augmente l'affinité de l'enzyme pour le substrat ARNt pré-mature. Ces hypothèses expliquent en partie l'augmentation de la vitesse de la

réaction de clivage observée lorsque la partie protéique est ajoutée à l'ARN catalytique (Kurz *et al.*, 1998). Par ailleurs, des expériences de délétions au niveau de la partie 5' du pré-ARNt ont suggéré que l'augmentation de l'affinité de la RNase P pour son substrat est liée à la liaison directe de la partie protéique à l'extrémité 5' du substrat (Crary *et al.*, 1998 ; Niranjanakumari *et al.*, 1998).

Il n'existe à ce jour aucune structure publiée du complexe ribonucléoprotéique de la RNase P ni de l'ARN catalytique entier isolé. Seules des structures cristallographiques de la partie protéique de quelques RNase P ont été publiées. Celle de *Bacillus subtilis* est présentée dans la Figure II.

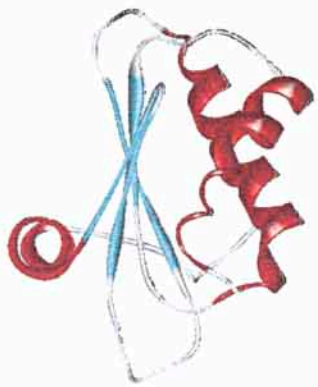


Figure II : Structure de la partie protéique de la RNase P de *Bacillus Subtilis*. Figure construite à partir du code PDB 1A6F (Stams *et al.*, 1998).

Il s'agit d'une structure en « sandwich α-β » étonnamment similaire à la protéine ribosomique S5. Ce résultat suggère que la RNase P a évolué de façon étroitement liée à la machinerie de la traduction d'où elle a « tiré profit » des propriétés physico-chimiques que procurent les chaînes polypeptidiques.

Il est important de noter que contrairement au ribozyme à tête de marteau décrit précédemment et qui dévient inactif après qu'il ait catalysé la rupture de l'une de ses liaisons phosphodiesters, l'ARN catalytique de la RNase P n'est pas détruit à l'issue du clivage de son substrat pré-ARNt et peut donc agir de façon répétitive sur de nouvelles molécules. Cette propriété est extrêmement importante car elle définit désormais l'ARN catalytique de la RNase P comme une « vraie enzyme ». Il est très probable que ce soit l'association de la chaîne polypeptidique à l'ARN catalytique qui ait permis de transformer ce dernier d'un simple catalyseur « suicide » en un catalyseur enzymatique. Pourquoi ?

D'après la Figure I du paragraphe précédent, il apparaît clairement que si l'ARN a réussi à catalyser son propre auto-clivage, c'est en grande partie grâce à la structuration de cet ARN, structuration qui est dictée en grande partie par l'appariement des bases. Cette structuration a permis de résoudre plusieurs problèmes qu'une enzyme « ordinaire » rencontre habituellement. Elle a permis en effet le recrutement du substrat à couper (celui qui s'apparie à une séquence au voisinage du site catalytique) et par la même occasion le positionnement de ce substrat dans le site actif dans une conformation telle que la liaison scissile est désormais au voisinage des groupements chimiques qui vont catalyser la coupure. Par ailleurs, le fait que la réaction soit intra-moléculaire augmente fortement la probabilité que la partie substrat de l'ARN rencontre la partie catalytique. Ce sont ces deux points fondamentaux qui expliquent comment un ARN a su jouer seul le rôle d'une ribonucléase. A ce niveau, le recrutement du substrat et la catalyse de la coupure étaient deux fonctions chevauchantes assurées par la même molécule. Il est très légitime de penser que c'est l'association de la chaîne polypeptidique qui a su séparer ces deux fonctions. Grâce à la sous-unité C5 de la RNase P, il est à présent possible de recruter non seulement des substrats libres c'est à dire non liés covalemment à l'ARN catalytique (la réaction devient alors inter-moléculaire) mais aussi de favoriser le contact entre l'ARN catalytique et l'ARN substrat en atténuant les répulsions électrostatiques des squelettes polyphosphatés. La chaîne peptidique a remplacé l'importance exclusive des interactions ARN/ARN (appariements base-base) par des interactions protéines/ARN entre la chaîne polypeptique et le substrat d'une part et entre la chaîne polypeptidique et l'ARN catalytique d'autre part. Nous pensons que c'est ce changement dans le type d'interactions qui a donné à la RNase P la propriété d'une enzyme vraie capable de catalyser les réactions de coupure de façon récursive.

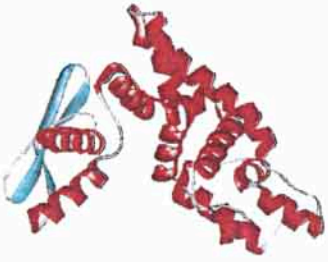
Nous verrons dans les deux paragraphes qui suivent que le niveau d'organisation (ARN + protéine) que nous venons de décrire est largement dominé dans le paysage ribonucléasique actuel par une organisation macromoléculaire basée exclusivement sur des protéines.

1.1.3 Une seule chaîne polypeptidique

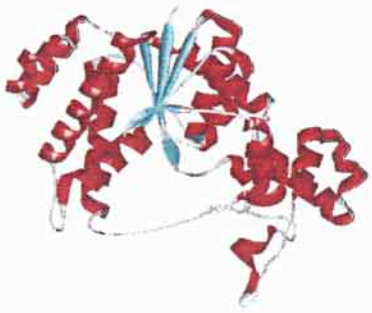
La grande majorité des activités ribonucléases connues actuellement est assurée par des protéines. Chez les bactéries, par exemple, 22 protéines différentes porteuses de 19 activités ribonucléases distinctes ont été identifiées (Condon *et al.*, 2002). Des activités RNases semblent avoir adopté toutes les structures secondaires de base mais avec une nette préférence pour les structures de type α/β (sous-unités α et β successives) et les structures de type $\alpha + \beta$ (sous-unités α et sous-unités

β isolées) (Aravind *et al.*, 2001). Ces structures combinent en effet la flexibilité des hélices α et la rigidité des feuillets β et offrent donc une charpente optimale pour la construction d'une activité catalytique.

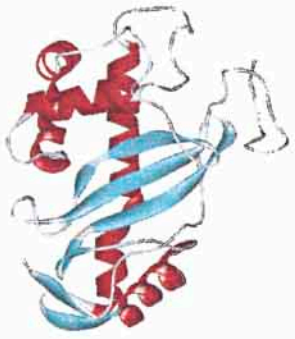
La Figure III présente quelques exemples de repliements adoptés par les ribonucléases. L'étude des stratégies catalytiques adoptées par ces protéines sera faite dans le chapitre 4 de cette dissertation.



RNase H du phage T4 (1TF3)



RNase III de Thermotoga Maritima (1O0W)



RNase Rh de Rhizopus niveus (1BOL)



RNase A bovine (Bos taurus) (1RTB)

Figure III : 4 Exemples de repliements de ribonucléases. Les codes PDB utilisés pour construire ces images sont entre parenthèses.

En rapport avec l'analyse du paragraphe précédent, il est important de constater à ce niveau que le recrutement du substrat et la catalyse de la coupure sont réalisés désormais exclusivement par la chaîne polypeptidique.

1.1.4 Un complexe multiprotéique : Exemple du dégradosome d'*E. coli* et de l'exosome de la *Saccharomyces cerevisiae*

Le niveau d'organisation macromoléculaire le plus complexe connu actuellement et associé à des activités ribonucléases est celui des complexes multiprotéiques. Ce niveau est rencontré aussi bien chez les procaryotes que chez les eucaryotes.

Chez *E. coli*, un complexe nucléolytique portant le nom de *dégradosome* est formé de quatre composants : la ribonucléase E (RNase E), la polynucléotide phosphorylase (PNPase), l'ARN hélicase B (RhlB) et l'énolase (Ehretsmann *et al.*, 1992 ; Miczak *et al.*, 1996) (d'autres protéines comme la polyphosphate kinase et des protéines chaperonnes semblent avoir une participation moindre à ce complexe) (Blum *et al.*, 1997). Ce complexe a été découvert lors de la purification de la RNase E, une endoribonucléase impliquée dans la dégradation de certains ARN ribosomiques, ARN de transferts ainsi que d'une partie importante des ARN messagers de la cellule (Cohen *et al.*, 1997 ; Régnier *et al.*, 2000). La RNase E est une ribonucléase spécifique du simple brin. Elle comprend 1061 résidus aminés contenant plusieurs domaines. La catalyse de la réaction de clivage est assurée par la partie N-terminale alors qu'une région centrale, riche en arginines, joue le rôle de domaine de liaison à l'ARN. La partie C-terminale sert de support pour l'assemblage des autres composants du dégradosome (PNPase + RhlB + Enolase) (Kaberdin *et al.*, 1998). La PNPase possède une activité exoribonucléase 3'-5'. L'hélicase RhlB pourrait dérouler les structures secondaires qui bloqueraient la progression de la PNPase (Coburn *et al.*, 1999).

Chez les eucaryotes, plusieurs processus cellulaires nécessitant des activités exo-ribonucléases 3'-5' sont réalisés par des complexes multiprotéiques qui portent le nom d'*exosome* (Mitchell *et al.*, 1997). L'exosome de la levure *S. cerevisiae* est le complexe le mieux caractérisé. 10 des 11 composants de cet exosome ont soit une activité exoribonucléase 3'-5' établie sur la base de preuves expérimentales directes soit une activité exoribonucléase 3'-5' prédite sur la base d'homologies de séquences (Allmang *et al.*, 1999 ; Burkard *et al.*, 2000). Des analyses génétiques ont montré que l'exosome de *S. cerevisiae* intervient dans la maturation de la partie 3' de plusieurs ARN stables dont l'ARN ribosomique 5.8 S, les petits ARN nucléaires (snARN) et les petits ARN nucléolaires (snoARN) (Allmang *et al.*, 1999 ; Kufel *et al.*, 2000). Il intervient également dans le métabolisme des ARNm en dégradant les précurseurs non épissés qui persistent dans le noyau

(Bousquet-Antonelli *et al.*, 2000) ainsi que les ARNm ayant perdu leurs queues poly(A) (Jacobs *et al.*, 1998). La Figure IV ci-après présente un modèle tridimensionnel de la structure de l'exosome de la levure. Ce modèle a été construit en utilisant des images de microscopie électronique et en exploitant la similitude qui existe entre certains domaines de l'exosome de la levure et un composant du dégradosome bactérien (la PNPase) dont la structure tridimensionnelle a été publiée en 2000 (Symmons *et al.*, 2000).

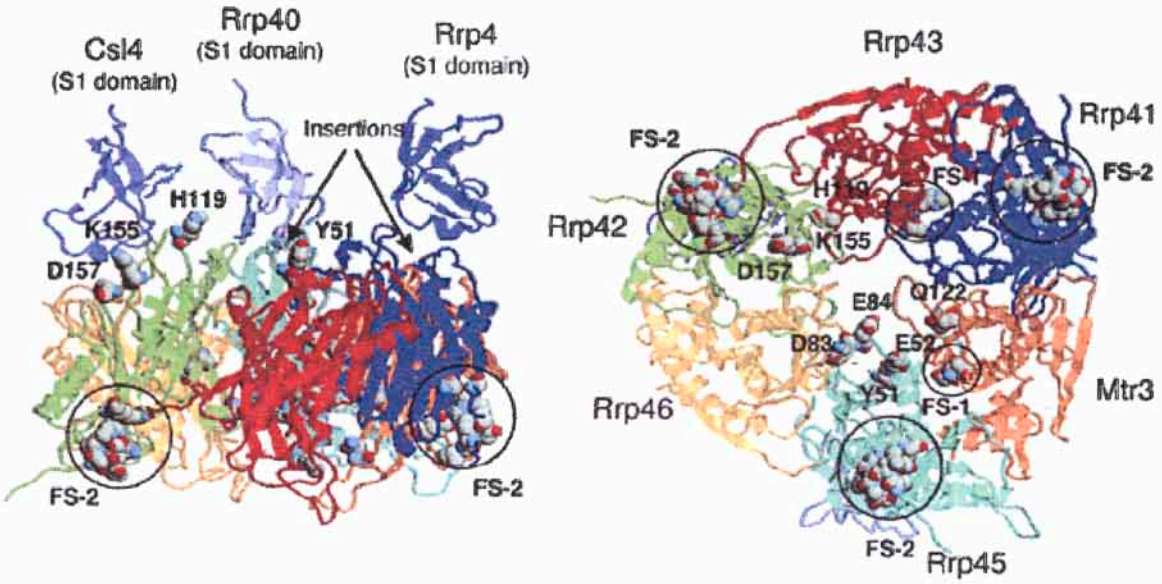


Figure IV : Deux vues différentes du modèle tridimensionnel de l'exosome de *S. cerevisiae*. Rrp40, Rrp41, Rrp42, Rrp43, Rrp45, Rrp46, Csl4, Rrp4 et Mtr3 désignent les sous-unités de l'exosome contenant un ou plusieurs domaines de la PNPase. Les résidus numérotés sont les résidus conservés dans les molécules orthologues. Les cercles indiquent les sites fonctionnels prédits (FS). D'après Aloy *et al.*, 2002.

Les quatre niveaux d'organisation macromoléculaire que nous venons d'exposer suggèrent que l'activité ribonucléase, supportée aux origines de la vie par des acides ribonucléiques, a évolué en remplaçant progressivement la chaîne ribonuléique par une chaîne polypeptidique. La meilleure adéquation des chaînes polypeptidiques avec les exigences de la catalyse enzymatique est certainement à l'origine de ce remplacement. La RNase P serait un vestige de l'une des étapes de ce remplacement progressif.

Abordons à présent le deuxième point important qui caractérise les ribonucléases : la spécificité de coupure.

1.2 La spécificité de coupure chez les ribonucléases

La spécificité de coupure chez les ribonucléases est un terme général qui peut être étudié à différentes échelles et selon différents critères. Plusieurs de ces critères sont chevauchants. Aussi, nous informons le lecteur que le bref aperçu de la spécificité que nous allons présenter ne peut être organisé que de façon artificielle.

1.2.1 Clivage d'une classe particulière d'ARN

La spécificité de coupure peut être liée au type d'ARN (ARNm, ARNt ou ARNr). La RNase P que nous avons présenté précédemment reconnaît spécifiquement les précurseurs des ARN de transfert se trouvant dans la cellule. Les bases de cette reconnaissance ne sont pas complètement connues mais plusieurs expériences suggèrent que la spécificité est liée à la structure tridimensionnelle particulière de l'ARNt ainsi qu'à la présence du motif RCC(A/C) à l'extrémité 3' de l'ARN substrat (Kirsebom L., 2001). Cette suggestion explique pourquoi certaines molécules comme le précurseur de l'ARN ribosomique 4.5 S et l'ARN du bactériophage M3 sont clivés par RNase P. Ces ARN possèdent en fait soit des structures semblables à celle des ARN de transferts soit des structures en tige-boucle supportant un motif RCC(C/A) à leur extrémité 3' terminale (Altman *et al.*, 1999).

1.2.2 Clivage d'un ARN simple brin ou d'un ARN double-brins

La RNase E présente une activité ribonucléase spécifique de l'ARN simple brin. Bien qu'on ait montré expérimentalement que plusieurs des sites de clivage soient précédés ou suivis de structures secondaires simples de type tige-boucles, leur présence ne semble pas être indispensable pour que l'ARN soit clivé par la RNase E (Cohen *et al.*, 1997).

La RNase III d'*E. coli* est, en revanche, une endoribonucléase spécifique de l'ARN double-brins. Cette RNase magnésium dépendante est impliquée dans plusieurs processus dont la maturation

des précurseurs des ARN ribosomiques et des précurseurs des ARN messager du phage T7 (Dunn *et al.*, 1973 ; Court D., 1993). Le clivage par la RNase III dépend d'une information portée à la fois par la séquence du substrat et par sa structure. Les substrats ayant des structures en doubles hélices sont clivés au niveau des deux brins de façon analogue aux endonucléases de restriction. Les substrats contenant des boucles internes sont en revanche clivés uniquement sur un seul brin (Nicholson A.W., 1999). La spécificité de la RNase III pour l'ARN double-brins est liée en partie à la particularité structurale de cette protéine au niveau de sa partie carboxy-terminale. Cette RNase contient en effet un motif appelé « domaine de liaison à l'ARN double-brins » (dsRBD) (St Johnston *et al.*, 1992) que l'on retrouve dans plusieurs protéines liant les ARN double-brins. La structure tridimensionnelle du domaine dsRBD de la RNase III a été résolue par RMN (Kharrat *et al.*, 1995). Elle montre un aspect compact ellipsoïde de type αββββα. Par ailleurs la structure cristallographique du domaine dsRBD que l'on retrouve dans la protéine A (fixatrice d'ARN) de *Xenopus laevis* a été résolue en présence d'un substrat ARN double-brins (Ryter *et al.*, 1998). Cette structure, présentée dans la Figure V, montre que l'interaction entre la protéine et l'ARN se fait sur une longueur de 16 paires de bases environ. Le nombre important de contacts impliquant les groupes 2'OH des deux brins de l'ARN avec le domaine dsRBD explique la sélectivité de reconnaissance de l'ARN double-brins.

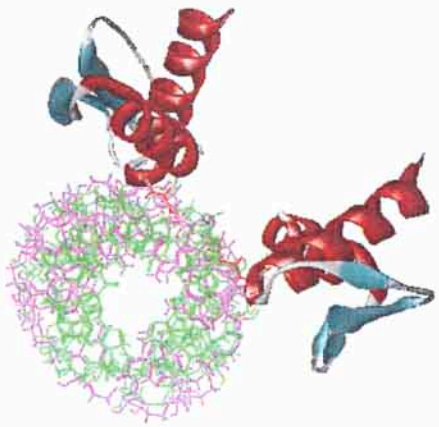


Figure V : Structure cristallographique de la protéine A (*Xenopus Leavis*) en complexe avec son ligand ARN double-brins. Le code PDB utilisé est 1DI2.

Signalons que des travaux relativement récents ont mis en évidence l'implication directe et originale de la reconnaissance spécifique des ARN double-brins dans la régulation de l'expression génétique. Cette implication fait partie d'un mécanisme global qui porte le nom d'*interférence ARN* (ARNi). Dans ce mécanisme, des ARN double-brins dirigent l'extinction ciblée de l'expression de certains gènes. Chez les animaux comme chez les plantes, le phénomène ARNi est caractérisé par la présence d'ARN de 22 nucléotides environ homologues aux gènes cibles. Ces ARN servent de « guides » pour instruire la destruction de l'ARNm messager par une nucléase multi-composant qui porte le nom de RISC. Parmi les enzymes impliquées dans la production des ARN guides, l'enzyme « Dicer » est une enzyme de la famille de la Rase III qui clive spécifiquement l'ARN double-brins (Bernstein *et al.*, 2001).

1.2.3 Clivage d'un ARN dans une structure nucléique hybride

La Ribonucléase H hydrolyse spécifiquement le brin ARN se trouvant dans un duplex ADN/ARN (Hostomsky *et al.*, 1993). Cette enzyme se retrouve dans plusieurs organismes. Chez *E. coli*, la RNase HI produit du gène *rnhA*, agit comme une phosphodiestérase Mg^{2+} dépendante en libérant une extrémité 3'OH. Elle serait impliquée dans le métabolisme de l'ADN où elle clive lors de la réplication une amorce ARN pour fournir une extrémité 3' nécessaire à l'initiation de la synthèse de l'ADN (Hostomsky *et al.*, 1993). Un modèle a été proposé pour expliquer comment la RNase H discrimine entre un duplex ARN/ADN et un duplex ARN/ARN pure (non clivé par la RNase H). Ce modèle s'appuie sur la résolution, par résonance magnétique nucléaire, de la structure tridimensionnelle en solution d'un duplex ARN/ADN substrat (Fedoroff *et al.*, 1993). Selon ce modèle, la RNase H reconnaît spécifiquement le duplex hybride car ce dernier n'adopte ni la forme A (caractéristique des duplex ARN/ARN) ni la forme B (caractéristique des duplex ADN/ADN) mais une forme intermédiaire. Parmi les particularités de cette forme intermédiaire citons l'existence d'un sillon mineur plus étroit que celui de la forme A mais plus large que celui de la forme B et l'adoption par le désoxyribose de l'ADN d'une conformation originale appelée O4'-endo.

1.2.4 Clivage spécifique après une base particulière

Il existe dans la littérature plusieurs exemples de ribonucléases qui clivent des liaisons phosphodiesters se trouvant en 3' d'une base particulière mais ces clivages ne présentent jamais un

caractère de spécificité absolue. En fait, ces RNases ne clivent pas exclusivement la base en question mais présentent simplement une préférence pour cette dernière.

La RNase T1 est l'un des rares exemples de ribonucléases qui clivent avec une grande spécificité la liaison phosphodiester consécutive à une guanine (à un nombre limité d'exceptions près) (Irie M., 1997 ; Yoshida H., 2001). Les bases moléculaires de cette spécificité relative sont bien connues. La base guanine est en effet recrutée dans le site actif de l'enzyme par l'intermédiaire de plus liaisons hydrogènes. L'azote N(7) et l'oxygène O(6) de la base acceptent, par exemple, respectivement des liaisons hydrogènes de la part des groupes amides NH de l'asparagine 43 et de l'asparagine 44. Par ailleurs, les résidus tyrosines 42 et 45 s'empilent parfaitement de part et d'autre du cycle de guanine (voir Figure VI). Le segment peptidique compris entre la tyrosine 42 et le glutamate 46 est défini comme étant le site de fixation de la guanine (Yoshida H., 2001 ; Langhorst *et al.*, 1999).

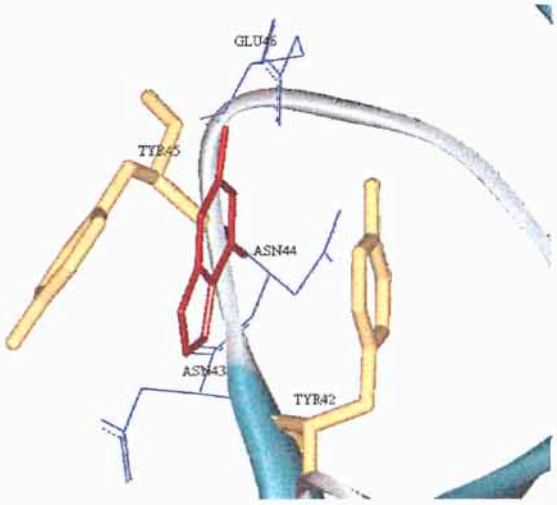


Figure VI : Structure de la RNase T1 en complexe avec le 2' GMP. Les deux Tyrosines (en jaune) s'empilent de part et d'autre du 2'GMP (en rouge). La figure a été construite à partir du code PDB 1BVI.

1.2.5 Clivage au niveau d'un contexte séquentiel particulier

Parmi toutes les ribonucléases caractérisées jusqu'à nos jours, seul un nombre très restreint nécessite la présence d'un contexte séquentiel particulier pour cliver la liaison phosphodiester. Par ailleurs, il est établi que ce contexte est souvent nécessaire mais non suffisant pour diriger la coupure.

Un déterminant structural supplémentaire est en général requis. A notre connaissance, il existe actuellement quatre groupes de ribonucléases spécifiques d'une séquence contenant au moins deux nucléotides. Ces quatre groupes sont : le groupe des colicines (exemple de la colicine E5), le groupe de la restrictocine, le groupe de la ribonucléase spécifique de la queue polyA (PARN) et le groupe de l'endoribonucléase RegB.

Avant de présenter quelques exemples, Il est très surprenant de constater qu'à l'exception de la ribonucléase spécifique de la queue polyA (PARN), toutes les ribonucléases spécifiques d'une séquence nucléotidique sont liées de très près à la machinerie de la traduction.

1.2.5.1 La colicine E5 : une RNAse de restriction qui reconnaît le motif QU

Le domaine C-terminal de la colicine E5 (115 résidus aminés) clive les boucles anticodons de différents ARN de transfert chez *E. coli* aussi bien *in vivo* qu'*in vitro* (Ogawa *et al.*, 1999). Les ARNt cibles de la colicine E5 sont ceux des acides aminés tyrosine, histidine, asparagine et acide aspartique. Comme le montre la Figure VII, le clivage se fait spécifiquement au milieu du motif QU (Q désigne la queuine, une guanine modifiée) (Masaki *et al.*, 2002).

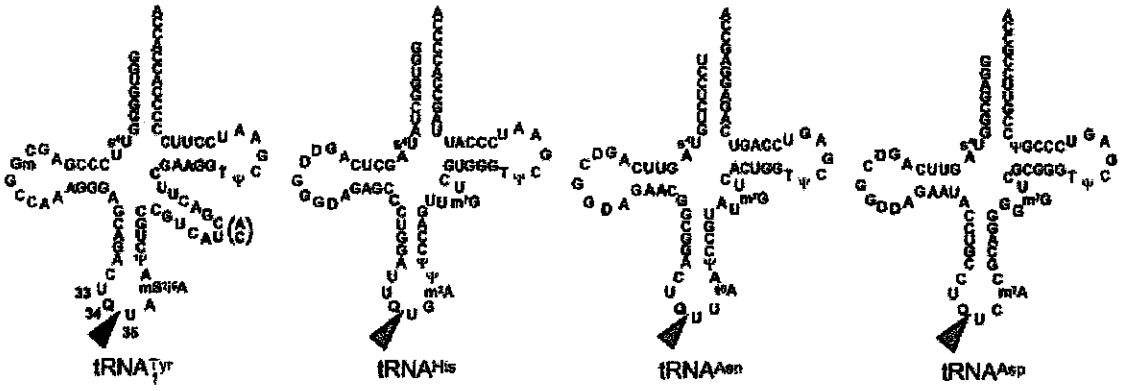


Figure VII : Structures secondaires des ARNt substrats de la colicine E5. Le triangle sombre indique que le site de clivage a été déterminé par séquençage ou par analyse de la base terminale, le triangle clair indique que le site de clivage a été déterminé par l'étude de la taille des produits de coupure. D'après Masaki *et al.*, 2002.

La spécificité de coupure de cette RNase pour le motif QU (et GU) a été confirmée *in vitro*. Les bases structurales de cette spécificité ne sont pas encore établies. La structure

tridimensionnelle du domaine C-terminal de la colicine E5 en complexe avec le dinucléotide non clivable dGpdU est en cours d'analyse (Masaki H., communication personnelle).

1.2.5.2 La restrictocine

La restrictocine est une RNase produite par le champignon *Aspergillus restrictus* (Rodriguez *et al.*, 1982). Elle clive spécifiquement une liaison phosphodiester unique parmi les milliers de liaisons que comprend la partie ribonucléique des ribosomes. Cette protéine possède 86 % d'identité de séquence avec son homologue fonctionnel, l'α-sarcine, produite par *Aspergillus giganteus* (Rodriguez *et al.*, 1982). La liaison phosphodiester clivée est celle qui lie la base G4325 à la base A4326 (numérotation relative à l'ARNr 28 S du rat). Cette liaison appartient à une séquence très conservée qui porte le nom de boucle sarcine/ricine (Sarcin/Ricin loop : SRL) et qui se trouve dans l'ARN ribosomique 23 S (procaryote) et 28 S (eucaryote) (Munishkin *et al.*, 1997). C'est cette séquence SRL que reconnaît la restrictocine. Des oligonucléotides (29-mer environ) mimant ce motif ont été synthétisés (Yang *et al.*, 2001). La Figure VIII en donne un exemple.

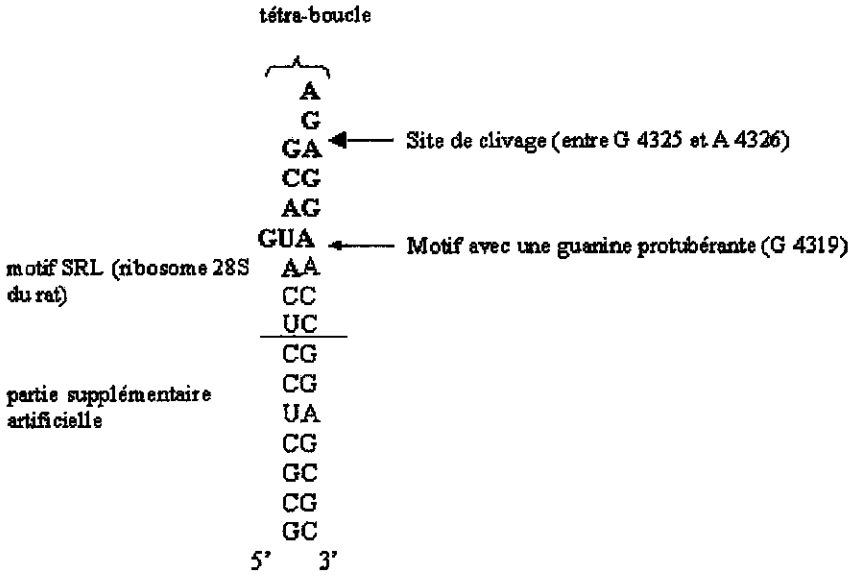


Figure VIII : Structure d'un ARN de synthèse portant le motif SRL de l'ARNr 28S. La séquence du motif SRL est 5' UCAGUACGAGAGGAACC 3'. Les nucléotides conservés sont en gras. D'après Yang *et al.*, 2001.

Les structures tridimensionnelles de la restrictocine en complexe avec différents oligonucléotides synthétiques (29-mer) reproduisant la séquence et la structure du motif SRL ont été résolues (Yang *et al.*, 2001). Les auteurs montrent que l'ARN adopte au niveau de la liaison scissile une structure du type « tétra-boucle ». Ils suggèrent que la protéine positionne convenablement le substrat grâce à deux interactions spécifiques. La première implique une structure de coude très particulier dit coude « S », la deuxième implique une guanine protubérante extrêmement déterminante pour la spécificité de coupure (la guanine G 4319) (voir Figure IX).

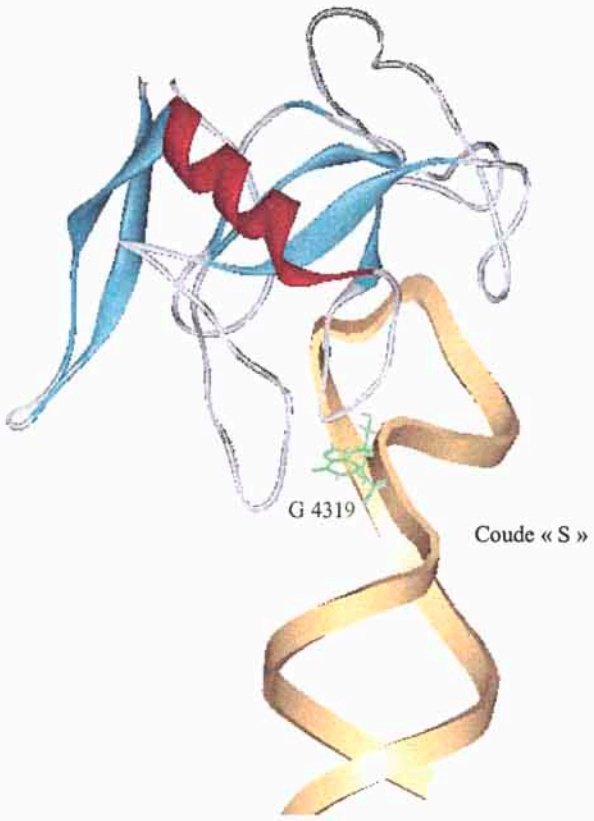


Figure IX : Bases structurales de la spécificité de coupure de la restrictocine. L'interaction avec la guanine G 4319 et la structure de coude « S » positionne l'ARN dans la conformation favorable au clivage. Figure construite à partir du code PDB 1JBT et la référence Yang *et al.*, 2001.

Un troisième exemple de ribonucléase qui clive une liaison phosphodiester unique dans un contexte séquentiel particulier est celui de l'endoribonucléase RegB du phage T4. Nous verrons que cette ribonucléase présente une spécificité de séquence telle qui justifie l'emploi du terme

« ribonucléase de restriction ». C'est cette RNase qui sera l'objet de toute notre dissertation. La partie qui suit résume l'état des connaissances relatives à RegB au moment où on a entrepris de travailler sur cette enzyme.

1.3 L'endoribonucléase de restriction RegB

La RNase RegB est codée dans le génome du bactériophage T4. Avant d'aborder la spécificité de cette ribonucléase, il convient d'abord de rappeler le cycle de développement de ce virus et l'historique de la découverte de RegB.

1.3.1 Le cycle lytique du phage T4 dans *E. coli*

Le bactériophage T4 appartient à la famille des *Myoviridae.* Il s'agit d'un virus à capside icosaédrique possédant une queue à symétrie hélicoïdale (Figure X). Son ADN double-brins de 168 903 paires de bases est complètement séquencé.

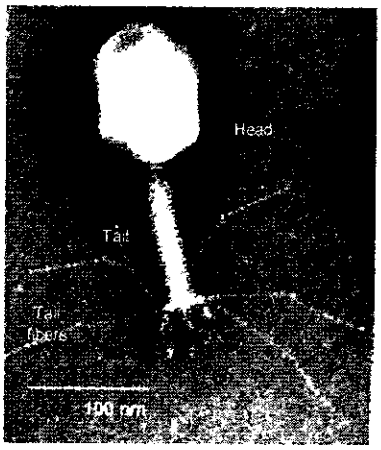


Figure X : Photo du bactériophage T4 obtenue par microscopie électronique. *(Source : Abedon S.T., Ohio State University).*

Le cycle de développement de ce phage nécessite l'activation puis l'inactivation successives de différents groupes de gènes, à des moments précis (Karam *et al.*, 1994). Trois classes de promoteurs sont successivement utilisées lors du développement.

Les promoteurs précoces sont reconnus, immédiatement après l'infection, par l'ARN polymérase non modifiée de la bactérie (sous-unités $\alpha_2.\beta.\beta'$/ facteur σ^{70}). Parmi les protéines précoces produites, on distingue des nucléases qui dégradent l'ADN de la bactérie et deux enzymes catalysant la synthèse de la 5-hydroxyméthyl-cytosine (nucléotide modifié qui protège l'ADN du phage contre le système de restriction de la bactérie).

Deux à trois minutes après l'infection à 30 °C, *les promoteurs moyens* sont utilisés. L'activation de ces promoteurs nécessite la fixation de la protéine précoce *MotA* ainsi que l'intervention du facteur anti-sigma *AsiA*. Les protéines moyennes sont essentiellement des enzymes de la réplication de l'ADN du phage selon le modèle du cercle roulant.

Sept minutes après l'infection, *les promoteurs tardifs* sont activés à la suite du remplacement du facteur σ^{70} (majeur dans *E. coli*) par le facteur σ^{55} du bactériophage T4. Les protéines tardives comprennent les protéines de la tête et de la queue du phage et les enzymes impliquées dans la lyse de la paroi de la bactérie (lysozyme) et la libération des virions.

1.3.2 La découverte du gène *regB*

Des mutants du bactériophage T4, capables de pousser sur un milieu contenant des analogues de l'acide folique (inhibiteurs du métabolisme des nucléotides), ont été isolés il y a 28 ans (Chace *et al.*, 1975). L'analyse du génome de ces mutants a révélé l'existence de larges délétions sur le chromosome phagique. Ces délétions conduisent à une dérégulation post-transcriptionnelle aboutissant à la surproduction de plusieurs enzymes dont la dihydrofolate réductase. Les auteurs ont suggéré que ces mutants avaient perdu au moins un gène appelé *regB* (reg pour régulation) impliqué notamment dans la régulation négative de l'expression génétique du phage T4.

En 1988, Marc Uzan et ses collaborateurs ont identifié une activité endoribonucléase dans *E. coli* infectée par le phage T4. Cette activité introduit spécifiquement des coupures au milieu de la séquence GGAG d'au moins deux ARNm du phage T4 libérant, ainsi des ARNm avec une extrémité 5' OH. Les auteurs ont montré que cette nouvelle activité dépendait de la synthèse des protéines du virus. En effet, l'addition au milieu de culture d'inhibiteurs de la synthèse protéique, *avant* l'infection par le phage T4, empêche la dégradation des ARNm étudiés. Le concept de nucléase codée par le génome du phage T4 et provoquant l'inhibition de l'expression d'une classe particulière de protéines du phage a été alors proposé (Uzan *et al.*, 1988).

Un an plus tard, il s'est avéré que cette activité endoribonucléase coupait également l'ARNm polycistronique de la dihydrofolate réductase, enzyme qui était surproduite dans les mutants du phage T4 présentés ci-dessus (Ruckman *et al.*, 1989). Les auteurs ont montré, par ailleurs, que les mutants de délétions de Chace et Hall (voir ci-dessus) étaient dépourvus de l'activité nucléase précitée. Ils ont en conclu que le gène support de cette activité endoribonucléase est codé par les parties délétées dans les mutants du phage T4. En utilisant la méthode de suppression de phénotype par complémentation, Judy Ruckman et ses collaborateurs ont identifié de façon formelle la région du génome du phage T4 située entre la coordonnée 62 276 et 62 737 comme étant la phase ouverte de lecture du gène *regB*, dont l'existence avait été postulée 15 ans auparavant.

Ce gène est situé immédiatement après un promoteur précoce. Ce résultat est en accord avec l'apparition rapide de l'activité endoribonucléase soulignée en 1988 par Marc Uzan et ses collaborateurs (90 secondes après l'infection à 30 °C). Par ailleurs, Judy Ruckman et ses collaborateurs ont constaté que le produit de ce gène est létal pour la bactérie quand il est surexprimé dans un vecteur à haut nombre de copies. Ils ont remarqué également que RegB coupe son propre ARNm au niveau du motif GGAG de la séquence Shine-Dalgarno (Shine *et al.*, 1975), ainsi qu'en trois positions de sa séquence codante.

1.3.3 RegB est une endoribonucléase séquence spécifique

L'alignement des séquences de plusieurs ARNm du phage T4, ayant subi *in vivo* des tests de coupure par RegB, a montré que dans la quasi-totalité des coupures a lieu au niveau du tétranucléotide GGAG (Sanson *et al.*, 1995). Un seul cas de coupure au milieu de la séquence GGAU est rapporté. Le motif GGAG coupé est toujours localisé dans des régions intergéniques et appartient souvent à une séquence Shine-Dalgarno (séquence de liaison de l'ARNm à l'ARN ribosomique 16 S). Les motifs GGAG situés dans la séquence codante sont pour la plupart peu, voire pas coupés.

Par ailleurs, il a été découvert que parmi les motifs GGAG des régions intergéniques, certains étaient résistants à RegB. C'est le cas de certains ARNm précoces du phage T4, de la plupart des ARNm moyens et de tous les ARNm tardifs. Cette absence de coupure n'est pas due à une disparition rapide de la protéine précoce RegB. En effet, l'activité endoribonucléase est détectée jusqu'à 10 mn après l'infection à 30°C, c'est à dire jusqu'à la fin de la phase moyenne (Sanson *et al.*, 2000). L'absence de coupure n'est pas due non plus à une protéine (phagique ou bactérienne) protégeant les GGAG ou inhibant RegB. L'intervention d'une protéine phagique est exclue puisque des bactéries transformées par deux plasmides, l'un portant le gène RegB, l'autre codant pour un ARN cible,

reproduisent le même profil de sélectivité que celui observé lors de l'infection par le phage T4. De même, l'intervention d'une protéine bactérienne est exclue puisque des tests de coupure *in vitro* avec des ARN purifiés en présence de la protéine RegB purifiée montrent une conservation de la sélectivité observée *in vivo*.

Ces résultats ont conduit aux conclusions suivantes : d'une part la protéine RegB est *nécessaire et suffisante* pour reconnaître et cliver spécifiquement son ARN substrat au niveau de la séquence GGAG, d'autre part le motif GGAG est *nécessaire* mais *non suffisant* pour définir la sensibilité d'un ARN à RegB. Cette dernière conclusion conduit émettre l'hypothèse qu'un ARN contenant le motif GGAG porte une *information supplémentaire* nécessaire à sa reconnaissance par RegB.

1.3.4 L'activité de RegB est stimulée *in vitro* par la protéine ribosomique S1

Des tests enzymatiques réalisés avec la protéine RegB purifiée ont montré que l'efficacité de coupure (et non pas sa spécificité) était très inférieure à celle observée *in vivo*. Judy Ruckman et ses collaborateurs ont montré que cette efficacité peut être stimulée d'un facteur pouvant atteindre 100 en présence de la protéine ribosomique S1 (Ruckman *et al.*, 1994).

S1 est une protéine de 557 acides aminés (Subramanian A.R., 1983). Elle fait partie de la petite sous-unité 30 S du ribosome. Elle comprend 6 motifs d'environ 70 acides aminés chacun, appelés motifs S1. L'endoribonucléase séquence-spécifique RegB requiert donc (au moins *in vitro*) un partenaire : la protéine S1.

1.3.5 Recherche de l'information supplémentaire portée par l'ARN et identifiable par RegB

L'information supplémentaire portée par l'ARN et identifiable par RegB doit résider soit dans la structure primaire de l'ARN (un biais dans la séquence en amont ou en aval du motif GGAG), soit dans sa structure secondaire qui pourrait favoriser ou au contraire inhiber le clivage par RegB. L'étude de la séquence des ARNm du phage T4 clivés par RegB a révélé une région riche en A+U du coté 3' du motif GGAG (Uzan M., 2001). Plus surprenant encore, cette partie 3' est dépourvue de cytosines (voir Figure XI).

Par ailleurs, l'utilisation de la méthode SELEX *(Systematic Evolution of Ligands by EXponential enrichment)* a permis de sélectionner *in vitro*, dans une banque combinatoire, des ARN clivés par RegB en présence de la protéine S1 (Jayasena *et al.*, 1996). En analysant la quarantaine de séquences obtenues, les auteurs ont retrouvé le motif GGAG avec une préférence pour le motif GGAGG (biais qui n'existe pas dans les ARNm du phage T4). Par ailleurs, ils ont constaté que la région en 3' du GGAG était anormalement riche en A+C. La différence qui existe entre les séquences, en 3' du motif GGAG, des ARNm naturels du phage T4 et des ARN artificiels suggère que l'enchaînement des ribonucléotides n'est probablement pas un déterminant majeur.

Région riche en A+U mais dépourvue de C

Séquences clivées

P	***regB***	**AUCAGUUAAGAGGAGAAUAACAUGACUAUCAAU**
P	***motA***	**CUACUGAUACAGAAUUACGGAGAUUAUAAAAUGUCUAAAG**
P	**ORF *frd1***	**UUUGAUUAAAGCUUGCUAAGGAGAAUAAAAUGAGAUUACAA**
P	***5' cef***	**UACUGAUACAGAAAACAACUUGGAGAAUAAAAUGGAAAAUUC**
P	***motB***	**CCCUAUCAAAACUAAUGGAGAAAAGAAAAUGUUCGCAC**

Région contenant des C

Séquences non clivées

M	***gène 1***	**AAUCUUAAUUAAAUUUGAGGAGAAACACAUGAAACUAAU**
M	**ORF *Y.-2***	**UAAAUAGAUAUAAUUCAGAGGAGACAAUCAUGUCAGAUAAG**
T	***soc***	**AAUCAUGUAAUUUAAAUAAAGGAGAAUUACAUGGCUAGUACU**
T	**ORF *5,1***	**UUCCAAUGAGCGCUGAUUUAGGAGAAUCCAUGGAAGGUUCU**
T	**gene *53***	**UAAUAUAUUUAUAAACAGGAGGGCCCAUGCUCUUUACA**

P : Précoce, M : Moyen, T : Tardif

Figure XI : Comparaison des séquences de quelques ARNm naturels du phage T4 en 3' du motif GGAG. D'après Uzan M., 2001.

La prédiction des structures secondaires de deux ARN issus de la méthode SELEX et possédant deux GGAG dont un seul est clivé par RegB a été réalisée (Jayasena *et al.*, 1996). Le résultat était clair pour seulement un seul des deux ARN : le GGAG clivé appartient à une région prédite non appariée alors que celui qui n'est pas clivé est engagé dans une hélice stable.

Plusieurs ARN cibles dérivés de la méthode SELEX ont été synthétisés dans notre laboratoire. Leurs cinétiques de coupure par RegB ont été corrélées avec leurs structures secondaires. Certaines de ces structures ont été vérifiées par RMN. Il apparaît que le GGAG engagé dans une

double hélice n'est pas coupé par RegB même après l'ajout de la protéine S1. De même, pour un ARN dont le GGAG est située dans une région complètement linéaire, aucune coupure n'a été détectée. En revanche, pour un ARN dont le GGAG est engagé dans une structure intermédiaire, on a constaté que la coupure par RegB est très efficace même en l'absence de la protéine S1 (Lebars *et al.*, 2001).

En conclusion, la spécificité de coupure attribuée à RegB est liée à la présence de la séquence GGAG et à l'existence probablement d'une structure secondaire au voisinage de ce motif. La protéine ribosomique S1 active RegB au moins *in vitro*. Elle pourrait jouer le rôle d'une « présentase » d'ARN.

1.4 Présentation du travail de recherche que nous avons mené sur RegB

Comme nous venons de le voir, un travail de recherche conséquent et rigoureux a permis de définir l'originalité de RegB en terme de spécificité de coupure et a suggéré des signaux de reconnaissance sur l'ARN. Cependant, nous pensons que ce travail a malheureusement manqué de situer de point de vue catalytique la place de RegB dans le paysage ribonucléasique actuel. En d'autres termes, le travail précédent a privilégié l'étude du problème de la reconnaissance du substrat au détriment de l'étude de la réaction de clivage elle même. L'une des explications que nous voyons à ce manque est l'absence de toute homologie de séquence (même très faible) entre RegB et une ribonucléase connue. Cette absence rend délicates et aléatoires les tentatives de situer RegB dans les quelques grandes familles auxquelles appartiennent toutes les ribonucléases connues.

L'objectif de ce travail de recherche est d'identifier le chemin réactionnel catalysé par RegB, de rattacher RegB à la lumière de ce résultat à un groupe de ribonucléases connues, d'identifier dans ce groupe des supports catalytiques communs et de rechercher la présence de leurs équivalents dans RegB.

L'atteinte de cet objectif a nécessité la mise au point de méthodes d'expression hétérologue non-conventionnelles surmontant l'extrême toxicité de RegB. Elle a aussi demandé l'adaptation de techniques physiques, comme la résonance magnétique nucléaire et la résonance plasmonique de surface, à l'étude des activités ribonucléases.

Nous montrerons que la réalisation de l'objectif présenté ci-dessus a eu deux conséquences directes. La première a contribué à la mise en place, pour la première fois, d'une stratégie d'étude de la structure tridimensionnelle de RegB en solution. La deuxième a permis de construire un nouvel outil d'ingénierie génétique exploitant l'originalité de RegB.

Chapitre 1

Expression de la ribonucléase RegB

Chapitre 1
Expression de la ribonucléase RegB

1.1 Le système d'expression pET

Surproduire une protéine recombinante nécessite de mettre le gène d'intérêt sous le contrôle d'un promoteur finement régulé. Beaucoup de ces promoteurs sont d'origine virale .

Le bactériophage T7 est un virus lytique à ADN qui infecte la bactérie *E. coli*. Durant les phases tardives de son cycle de développement, il surexprime fortement une protéine appelée la protéine majeure de la capside (Studier F.W., 1969). Le taux d'expression exceptionnellement élevé de cette protéine, codée par le gène 10, est assuré par trois séquences régulatrices importantes.

Le gène 10 est d'abord sous le contrôle transcriptionnel du promoteur Phi10. Ce promoteur possède une séquence conservée entre la position -17 et la position +6 par rapport au point de démarrage de la synthèse de l'ARNm (Dunn *et al.*, 1983) et se retrouve à 6 reprises dans l'ADN de T7. Parmi les 17 groupes de promoteurs que contient l'ADN de T7, le promoteur Phi10 est connu pour être le plus fort (McAllister *et al.*, 1980 et 1981). Il est reconnu durant les étapes tardives du cycle viral et est spécifiquement et exclusivement transcrit par l'ARN polymérase du phage T7 exprimée dans *E. coli* à partir du génome viral (Chamberlin *et al.*, 1970). Par ailleurs, la transcription

initiée au niveau de ce promoteur est efficacement arrêtée au niveau d'une séquence conservée appelée terminateur Phi. Enfin, l'ARNm transcrit à partir du promoteur Phi10 porte une séquence d'initiation de la traduction très efficace : le site de liaison au ribosome Phi10 (RBS Phi10).

Les trois séquences (promoteur Phi10, terminateur Phi et RBS Phi10) ont été les candidats de choix pour construire une série de vecteurs d'expression, les vecteurs pET, dans lesquels le gène d'intérêt est cloné sous le contrôle transcriptionnel et traductionnel des séquences flanquant le gène 10 du bactériophage T7 (Studier *et al.*, 1986).

Pour exprimer ce gène, il est nécessaire d'apporter à la cellule une source de T7 ARN polymérase. Deux virus, λDE3 et λCE6, ont été construit à cet effet (Studier *et al.*, 1986). Ils portent tous les deux une copie du gène 1 codant pour la polymérase de T7 insérée dans la région *int* du bactériophage λD69 (Figure 1). Dans le cas du virus λCE6, le gène 1 est inséré dans la même orientation que le gène *int* ; il s'exprime à partir des promoteurs constitutifs p_I et p_L et induit donc l'expression de la protéine recombinante dès l'infection d'*E. coli*. Dans le cas du virus λDE3, le gène 1 est inséré dans une orientation inverse de celle du gène *int* ; il est sous le contrôle du promoteur lacUV5.

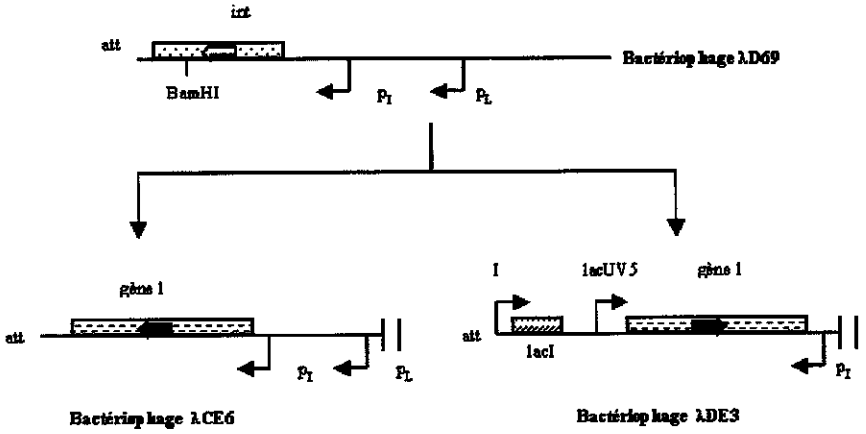


Figure 1. Construction des virus λCE6 et λDE3 portant le gène de la T7 ARN polymérase. Les deux virus sont construits par disruption du gène *int* (nécessaire à l'intégration du phage λD69 dans le chromosome bactérien) suite à l'insertion du gène 1 codant pour la T7 ARN polymérase dans le site BamHI de λD69. att : séquence favorisant l'intégration du génome du virus dans le chromosome bactérien. lacI : gène codant pour le répresseur lactose. I : promoteur du répresseur lactose. D'après Studier *et al.*, 1986.

Grâce au concours d'un virus « helper » apportant en *trans* l'activité « intégrase », le virus λDE3 a pu être intégré dans le chromosome bactérien donnant ainsi naissance à une nouvelle souche

E. coli appelée BL21(λDE3) (Studier *et al.*, 1986). Cette souche, lysogène pour λDE3, est stable mais présente une expression basale relativement élevée. Cette fuite a été minimisée en utilisant non plus le promoteur Phi10 sauvage mais une version composite, appelée T7lac, dans laquelle le promoteur T7 est fusionné à l'opérateur lac. T7lac et lacUV5 sont réprimés en présence du répresseur lactose codé par le gène *lacI* dont une copie se trouve sur le plasmide pET. Cette répression disparaît quand un composé chimique, l'isopropyl-β-D-galactopyranoside (IPTG), est ajouté au milieu de culture. L'IPTG, en effet, lie les molécules de répresseur lactose et les empêchent de reconnaître les séquences promotrices T7lac et lacUV5.

Une fois induite, la T7 ARN polymérase est si sélective et si active vis-à-vis du promoteur T7lac que quasiment toutes les ressources de la cellule sont désormais mobilisées pour exprimer la protéine recombinante. Cette dernière peut alors représenter plus de 50 % des protéines totales de la bactérie au bout de 3 à 4 heures d'induction.

Pour exprimer la ribonucléase RegB, nous avons été amené à mettre au point des stratégies d'expression non-conventionelles inspirées pour la plupart du système pET.

1.2 RegB ne peut être maintenue *in vivo* dans un vecteur pET classique

1.2.1 Preuve expérimentale

Nous avons tenté de cloner le gène *regB* sauvage dans les vecteurs standards pET15b et pET28b (Novagen). Ces vecteurs diffèrent par la composition de leurs sites multiples de clonage et ont conduit au même résultat.

La Figure 2 illustre le schéma de clonage du gène *regB* dans le vecteur pET28b. Ce clonage a été réalisé délibérément en deux étapes (clonage de la partie 5' puis de la partie 3') afin de s'affranchir de tout éventuel problème technique. Bien qu'un vecteur pET recombinant soit obtenu *in vitro* (voir les bandes de PCR du gène entier avec les paires d'amorces A-D et A-E) dans la solution de ligature, celui-ci ne donne aucun clone bactérien viable à l'issue de la transformation d'une souche *E. coli* pourtant dépourvue de T7 ARN polymérase (XL1-Blue).

Nous pensons que l'ARN polymérase de la bactérie est capable de diriger, à partir de promoteurs cryptiques, la synthèse de quelques molécules de messagers du gène *regB* qui une fois traduites en protéines sont suffisantes pour tuer la cellule.

Nous avons ici un cas de toxicité extrême surprenant mais non unique. En effet, un autre gène du phage T4, la gène *ndd* présente un niveau de toxicité comparable à celui de *regB*. Dans leur article d'octobre 1998, J.Y. Bouet et H.M. Krish expliquent : « Toutes les tentatives de clonage du gène *ndd* sauvage dans le vecteur pET11a sous le contrôle d'un promoteur T7 finement régulé ont échoué même dans des souches bactériennes dépourvues de T7 ARN polymérase. Quelques molécules Ndd sont probablement produites à partir de ce plasmide et sont suffisantes pour tuer *E. coli.* » (Bouet *et al.*, 1998).

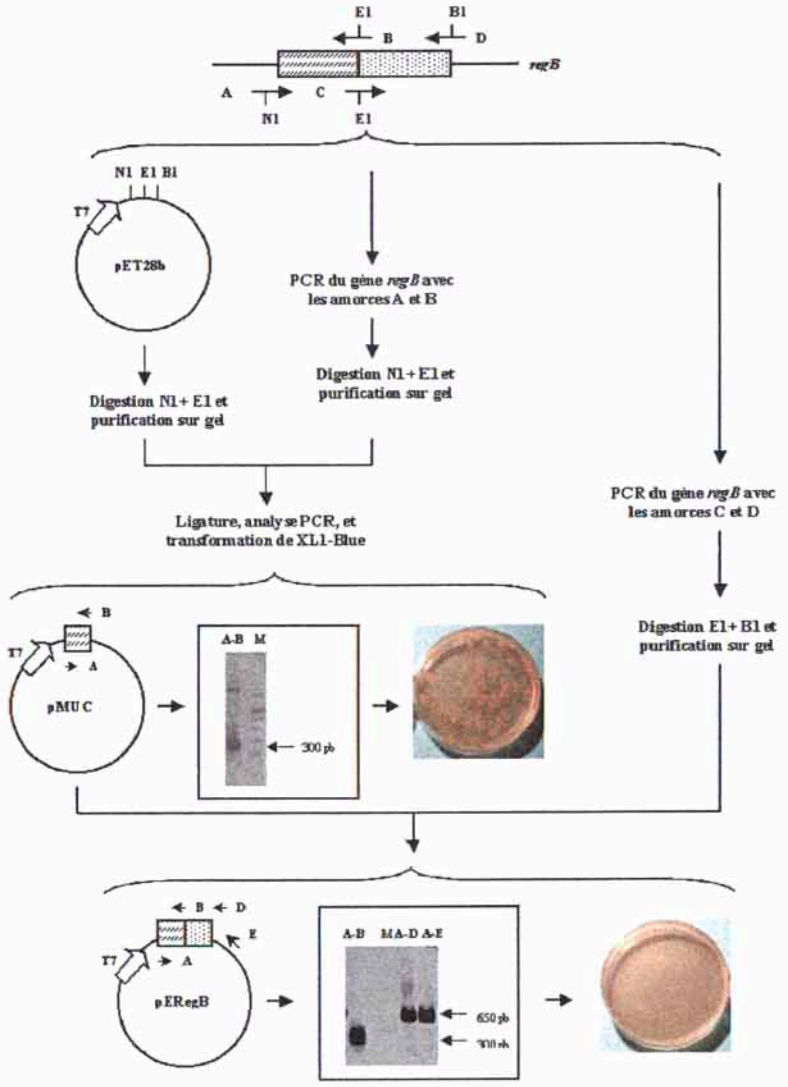


Figure 2. Clonage du gène *regB* dans le vecteur pET28b. La partie 5' du gène a été amplifiée avec les amorces A et B; la partie 3' avec les amorces C et D. N1, E1 et B1 sont les sites NdeI, EcoRI et BamHI respectivement. Chaque solution de ligature est précipitée au n-butanol puis partiellement purifiée sur gel pour éliminer l'excès d'insert. Une fraction est analysée par PCR, le reste sert à transformer la souche XL1-Blue (dépourvue de T7 ARN polymérase). L'amorce E est complémentaire d'une séquence sur le vecteur pET28b juste après le site B1.

Alors que la toxicité du gène *ndd* est relativement bien connue et est attribuée à sa capacité à désorganiser rapidement le nucléoide d'*E. coli*, l'origine de la toxicité de RegB reste imprécise. On pense néanmoins que RegB pourrait cliver spécifiquement un nombre restreint d'ARNm indispensables à la survie d'*E. coli*. Cette hypothèse est soutenue par les travaux de Marc Uzan et de son équipe qui ont démontré que RegB clive *in vivo* le transcrit du gène *ftsZ* (un gène essentiel impliqué dans les étapes précoces de la septation) entraînant la filamentation des cellules bactériennes (Uzan M., 2001).

1.2.2 Des systèmes d'expression plus fermés sont nécessaires

Pour qu'une construction de type pET, contenant le gène *regB*, puisse se maintenir dans *E. coli*, il faut trouver un moyen de réduire fortement l'expression résiduelle du système pET même en absence de source de T7 ARN polymérase.

Trois stratégies peuvent être utilisées pour minimiser l'expression basale d'un vecteur pET : la réduction de la transcription médiée par l'ARN polymérase d'*E. coli*, la réduction de l'efficacité du site d'initiation de la traduction et la manipulation du nombre de copies du plasmide (Figure 3). Nous montrerons dans les deux sous-chapitres qui suivent que les deux dernières stratégies permettent d'exprimer RegB en quantités suffisantes faire l'étude fonctionnelle de l'enzyme.

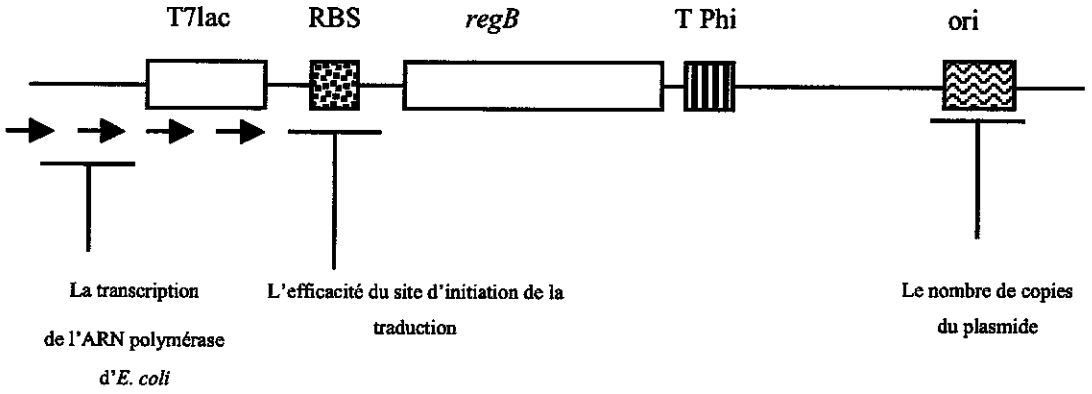


Figure 3. Stratégies de réduction de l'expression basale d'un vecteur pET. Les flèches pleines symbolisent la transcription résiduelle due à l'ARN polymérase d'*E. coli*. T7lac : promoteur composite Phi10 et site opérateur lactose, RBS : site d'initiation de la traduction (RBS Phi10), T Phi : Terminateur de transcription Phi, ori : origine de réplication.

Il est important de noter que la stratégie qui consiste à diminuer la transcription basale liée à l'activité de l'ARN polymérase d'*E. coli*, même si elle est techniquement beaucoup plus fastidieuse à

mettre au point est susceptible de fonctionner. Pour réduire la transcription parasite de la cellule hôte « il suffit » d'introduire un ou plusieurs terminateurs transcriptionnels forts efficacement reconnus par l'ARN polymérase d'*E. coli*. L'idéal est que ces terminateurs juxtaposent la région qui contient le promoteur T7lac et la séquence codante de *regB*. Le tandem des deux terminateurs *rrnB* T1 et *rrnB* T2 (puissants terminateurs naturels du gène *rrnB* codant pour un ARN ribosomique chez *E. coli*) représente un candidat intéressant à intégrer dans un vecteur de type pET (Orosz *et al.*, 1991).

1.3 Expression de RegB à partir d'un vecteur pET contenant un RBS Phi10 modifié

1.3.1 RegB se maintient dans un vecteur pET contenant un RBS peu efficace

Rosenberg et ses collaborateurs ont mis au point dans les années 80 une série de vecteurs, dits de transcription, contenant le promoteur T7 du gène Phi10 mais dépourvus de toute séquence d'initiation de la traduction (Rosenberg *et al.*, 1987).

RegB a été initialement cloné avec succès dans l'un de ces vecteurs, le vecteur pET7. Afin de permettre l'expression de la protéine RegB toute en conservant des clones viables, l'équipe de Marc Uzan a synthétisé *in vitro* une séquence nucléotidique contenant le motif Shine-Dalgarno AGGAG du site RBS Phi10 et l'a introduite dans le vecteur pET7 entre le promoteur T7 et la séquence codante de *regB* (Sanson *et al.*, 2000).

Une construction recombinante (pARNU2) a été obtenue. Cette dernière se maintient dans une souche dépourvue de T7 ARN polymérase. La souche BL21(λDE3) et ses variantes pLysS et pLysE ne sont pas viables en présence de ce plasmide malgré leurs niveaux d'expression basale atténués grâce à l'expression constitutive de molécules de lysozyme (inhibiteurs naturels de la T7 ARN polymérase) (Studier *et al.*, 1991).

Nous avons exprimé la ribonucléase RegB avec le plasmide pARNU2 dans la souche bactérienne BL21 (non lysogène pour le virus λDE3) en utilisant le système de surinfection avec le phage λCE6. Les rendements après purification sont de l'ordre de 200 µg de protéines par litre de culture. Sachant qu'un système pET classique, induit à l'aide du phage λCE6, permet d'obtenir en moyenne 10 mg de protéines par litre, nous concluant à une perte d'efficacité d'expression d'un

facteur 50 (10/0.2). La Figure 4 présente le protocole général d'expression et montre un exemple d'analyse SDS-PAGE d'une purification de RegB avec le système pARNU2/λCE6.

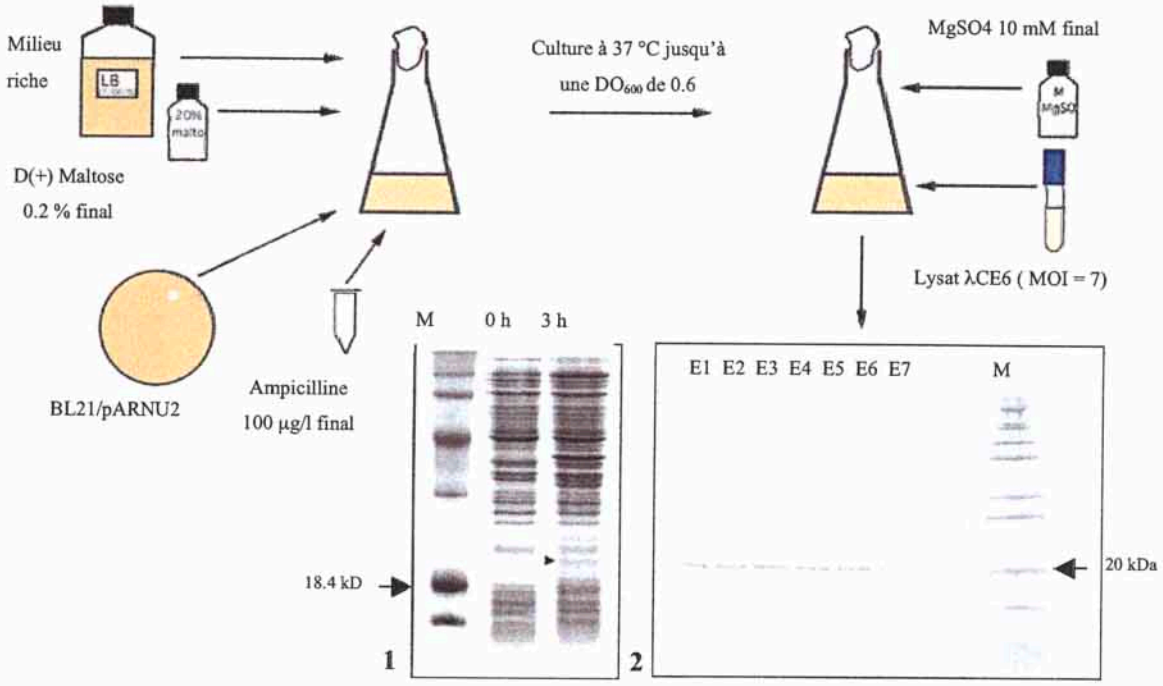


Figure 4. Expression de la Ribonucléase RegB à partir du vecteur pARNU2 contenant un RBS Phi10 atténué en utilisant le système d'induction λCE6. La souche BL21 transformée avec le plasmide pARNU2 est cultivée dans un milieu riche en présence de D(+) Maltose 0.2 % pour induire l'expression de la protéine de surface MalB qui sert à la fois de transporteur de Maltose et de récepteur pour le phage λCE6. A 0.6 unité de DO_{600}, l'expression de RegB est induite en ajoutant un lysat du phage λCE6 à une multiplicité d'infection de 7. Des ions magnésium sont ajoutés à ce stade, à une concentration finale de 10 mM, pour stabiliser les particules phagiques en solution. Les cultures sont récupérées au bout de 3 heures d'induction. Gel 1 : Analyse SDS-PAGE de l'expression de RegB au bout de 0 et de 3 heures d'induction (le culot cellulaire correspond à 0.05 DO_{600}. Gel 2 : Analyse SDS-PAGE de la purification de 1 l de culture (E1 à E7 fractions d'élution, dépôt de 4 µl sur une fraction de 250 µl). M : marqueur de masses moléculaires.

D'après le paragraphe 1.2.2, le faible taux d'expression de *regB* obtenu avec le plasmide pARNU2 ne peut être lié qu'à une altération du site d'initiation de la traduction du gène Phi10. Pour

vérifier cela nous avons séquencé la région RBS Phi10 de pARNU2 et l'avons comparé à la séquence du RBS Phi10 du phage T7. La Figure 5 montre l'alignement de ces deux séquences :

phage T7	T	A	A	G		A	G	G	A	G	A	T	A	T	A	C	A	T	a	t	g
pARNU2	C	C	T	G		A	G	G	A	G	A	T	A	T	A	C	A	T	a	t	g

Figure 5. Alignement du site d'initiation de la traduction du gène Phi10 dans l'ADN du phage T7 et du RBS Phi10 modifié introduit dans le plasmide pARNU2.

La séquence Shine-Dalgarno AGGAG est conservée dans les deux cas et est mise en évidence en caractère gras. La position -14 (par rapport au site de démarrage de la traduction) est en revanche mutée. Cette mutation est probablement à l'origine de la perte d'efficacité traductionnelle car elle correspond à une zone d'appariement avec l'ARN ribosomique 16 S d'*E. coli*. La Figure 5 montre que lorsque la position -14 est un A l'appariement avec l'ARN 16 S est optimal.

L'efficacité de la traduction de l'ARNm du gène Phi10 ne dépend donc pas seulement de la présence du consensus AGGAG mais aussi de la séquence en 5' de ce consensus

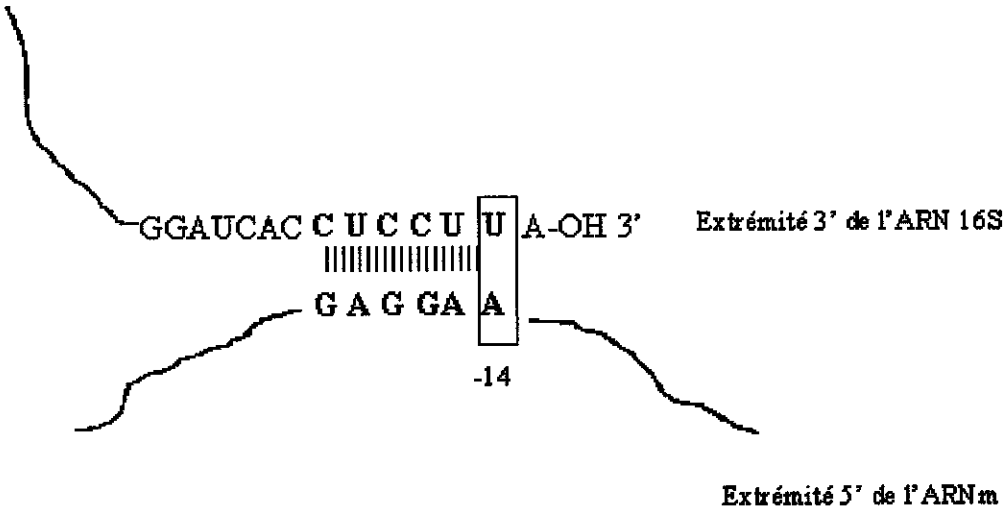


Figure 6. Appariement entre l'extrémité 3' de l'ARN 16 S de la sous-unité 30 S du ribosome d'*E. coli* et le RBS Phi10 illustrant l'importance de la position -14.

Malgré son faible rendement, le système RBS Phi10 modifié/λCE6 nous a permis de préparer des échantillons de RegB de 1 ml à une concentration de 4 à 5 $\mu mol.l^{-1}$ suffisante pour réaliser des tests

d'activité (voir chapitres suivants). L'étude par Résonance Magnétique Nucléaire de l'intermédiaire réactionnel produit durant la coupure d'un ARN par RegB (voir Chapitre 2) nécessite de préparer un échantillon plus concentré, de l'ordre de 50 $\mu mol.l^{-1}$. Nous avons donc adapté le système pARNU2/λCE6 décrit ci-dessus à la culture en fermenteur.

1.3.2 Adaptation du système RBS Phi10 atténué / λCE6 à la culture en fermenteur

La technologie de culture en fermenteur est une méthode efficace qui permet de produire des protéines recombinantes en larges quantités. Comparée à la technique de culture en incubateur, l'utilisation de fermenteurs présente plusieurs avantages.

D'abord, elle permet de manipuler des volumes de cultures un à deux ordres de grandeurs plus importants que ceux des cultures en incubateurs. Par ailleurs, les fermenteurs disposent d'un système de contrôle de la pression partielle en oxygène dans le milieu de culture. Ce contrôle permet, notamment au moment de l'induction de l'expression de la protéine recombinante, de maintenir les cellules dans un métabolisme oxydatif de type « respiration » (oxydation complète de la matière organique) beaucoup plus avantageux de point de vue énergétique (production de 36 moles d'ATP par mole de glucose) que le métabolisme oxydatif de type fermentation (oxydation incomplète de la matière organique produisant seulement 2 moles d'ATP par mole de glucose) qui domine lorsque la pression partielle en oxygène diminue. De plus, un système de contrôle du pH permet d'éviter une forte acidification du milieu de culture susceptible de freiner la vitesse de croissance de la culture bactérienne. Enfin, les fermenteurs disposent d'un système baromécanique qui, tout en préservant la stérilité du milieu de culture, permet d'effectuer des petits prélèvements (pour la mesure de la densité optique et l'observation au microscope) sans interrompre la régulation thermique ou l'oxygénation de la culture.

Nous avons expérimenté l'expression de la ribonucléase RegB avec le système RBS Phi10 modifié/λCE6 dans un fermenteur de 15 L. Le protocole d'expression utilisé diffère de celui présenté au paragraphe 1.3.1 par la technique de préparation du lysat du phage λCE6 et la méthode de cassage des cellules à l'issue de l'étape d'induction.

Le lysat du phage λCE6 est normalement obtenu en cultivant la souche bactérienne LE392 infectée par le phage jusqu'à la lyse complète des cellules et la libération des virions dans le milieu de culture. La centrifugation de ce milieu permet de le débarrasser des débris cellulaires et d'obtenir un lysat de phage dont le titre stable ne dépasse pas en général 10^{+10} particules phagiques par ml. Afin d'éviter d'induire la culture en fermenteur avec un volume important (il faudrait induire les 15 L de

culture avec 1.5 L de lysat à 10^{+10} phages par ml), nous avons préparé un lysat concentré en cultivant la souche LE392 infectée par le phage λCE6 et en récupérant le culot cellulaire juste avant le début de la lyse des cellules. Ce culot est resuspendu dans un petit volume de culture, les virions sont libérés en éclatant les cellules grâce à l'ajout de quelques gouttes de chloroforme. Avec cette méthode, des lysats titrant à 10^{+12} phages/ml ont été obtenus et ont permis d'induire la culture en fermenteur avec un volume limité.

Pour casser les cellules induites, nous avons utilisé une technique de cassage mécanique, le DYNO-Mill, qui permet de disrupter les cellules bactériennes en mélangeant le culot cellulaire avec des billes de verres dans une enceinte mise en rotation et maintenue à 4 °C.

La Figure 7 montre une photo du fermenteur utilisé (capacité maximale 20 L, volume effectif de la culture 15 L, source : Institut de Chimie des Substances Naturelles, CNRS, Gif-sur-Yvette) et le gel d'analyse SDS-PAGE de la purification de RegB. La fraction la plus pure et la plus concentrée (la fraction E2) représentait 500 µL de volume et une concentration de 40 µmol.l^{-1}.

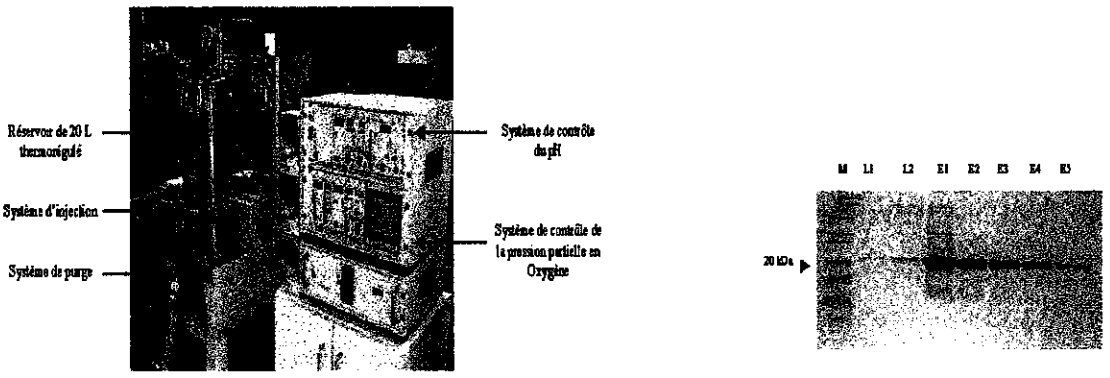


Figure 7. Expression de RegB en fermenteur. A gauche, photo du fermenteur utilisé. A droite, analyse SDS-PAGE de la purification de 15 l de culture. L1, L2 : lavages. E1 à E5 : fractions d'élutions (500 µl).

1.4 Expression de RegB en couplant la stratégie pET à une origine de réplication thermorégulable

La quantité totale d'ARNm de *regB* produite à partir du promoteur T7 dépend, entre autres, du nombre de molécules d'ARNm produites par unité d'ADN matrice (ce nombre est directement lié à la force du promoteur). Elle dépend aussi du nombre total de molécules d'ADN matrice disponibles

dans la cellule pour la machinerie de transcription. Ce nombre est proportionnel au nombre de copies du plasmide pET qui porte le gène *regB*.

Il est raisonnable de penser qu'une augmentation du nombre de copies du plasmide pARNU2 pourrait augmenter la quantité d'ARNm produite donc disponible pour la machinerie de traduction ce qui peut conduire à fortiori à une augmentation de la quantité de protéines exprimées.

Toutefois, en raison de la toxicité de RegB, l'augmentation du nombre de copies du plasmide pARNU2 ne serait réalisable en pratique que si celle ci est conditionnelle, c'est à dire dépendante d'un paramètre de culture, la température par exemple. Dans ce cas, on pourrait maintenir le plasmide pARNU2 à un bas nombre de copies en cultivant les cellules à une température permissive et augmenter ce nombre juste avant l'induction de l'expression de RegB en provoquant un saut vers une valeur restrictive.

Le vecteur pARNU2, un dérivé du plasmide pET7, possède une origine de réplication de type pBR322 (Bolivar *et al.*, 1977) qui maintient le plasmide à 15-20 copies par cellule. Le mécanisme de régulation de ce nombre de copies est un mécanisme de type ColE1 (Brendel *et al.*, 1993). Il est basée sur l'expression constitutive de la protéine Rop codée par le plasmide et qui joue le rôle de senseur du nombre de copies. Quand ce nombre devient trop élevé, la protéine Rop est exprimée de façon proportionnelle et inhibe la formation du complexe d'initiation de la réplication du plasmide au site *ori*. Plutôt que de manipuler ce système de régulation, nous avons extrait du plasmide pARNU2 tous les éléments nécessaires à l'expression (Promoteur T7, RBS Phi modifié, gène *regB* et terminateur Phi) et avons combiné ces éléments avec une origine de réplication régulée par la température dérivée du plasmide R1 (Gustafsson *et al.*, 1979).

1.4.1 Construction du vecteur d'expression pOURegB

1.4.1.1 La réplication thermorégulable du plasmide pOU61

Larsen J.E. et ses collaborateurs ont mis au point, au début des années 1980, une série de plasmides à réplication thermorégulable contenant une origine de réplication dérivée du plasmide R1 (Larsen *et al.*, 1984). Le mécanisme de réplication de ce plasmide est résumé dans la Figure 8. L'initiation de la réplication, au niveau du site ori, nécessite l'expression d'une protéine codée par le plasmide, la protéine RepA dont le rôle est d'ouvrir le duplex d'ADN pour la machinerie d'initiation de la réplication. La concentration de la protéine RepA, activateur de la réplication, est régulée par

deux mécanismes. Dans le premier mécanisme, la protéine RepA est produite de façon constitutive à partir d'un promoteur appelé copBp situé 700 paires de bases en 5' du gène *repA*. La partie 5' de ce transcrit contient l'ARNm de la protéine CopB. RepA est aussi produite à partir d'un promoteur interne (plus fort que le promoteur copBp) situé juste après l'extrémité 3' du gène *copB*. Ce dernier est réprimé par la protéine CopB. Ainsi, lorsque le nombre de copies du plasmide est trop élevé, on assiste à une augmentation de la quantité d'ARNm de RepA transcrits de façon constitutive à partir du promoteur CopBp ce qui entraîne en conséquence l'accumulation de la protéine CopB qui bloque la transcription de RepA à partir du promoteur fort (appelé aussi repAp).

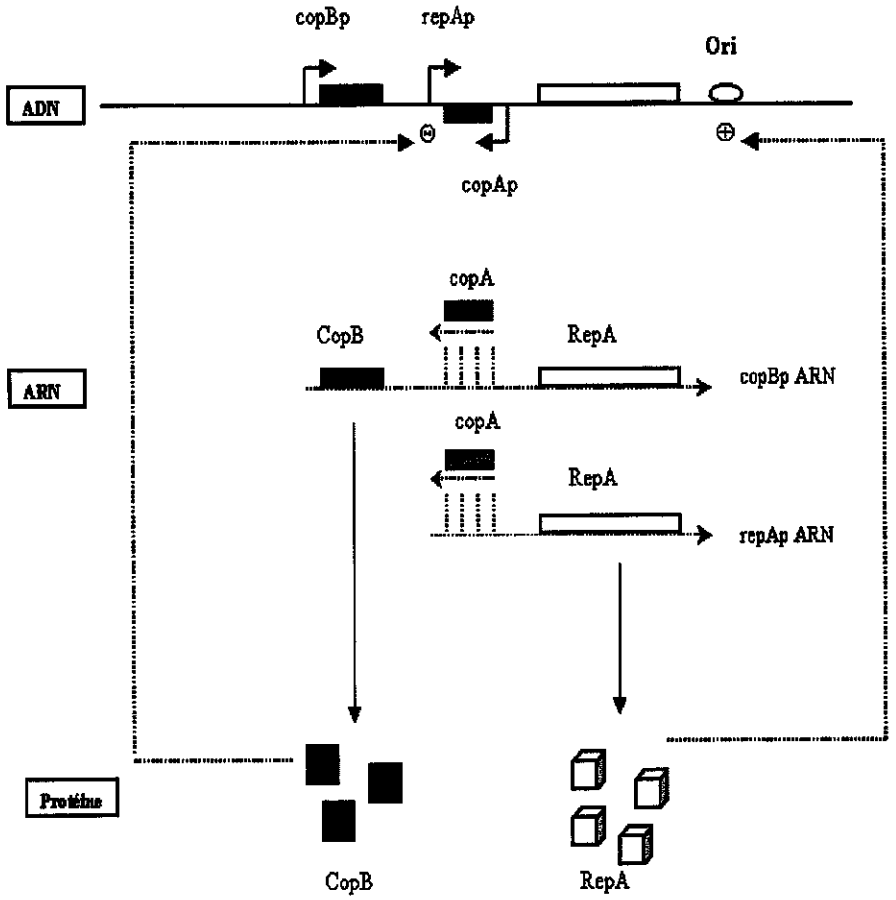


Figure 8. Mécanisme de régulation de la réplication du plasmide R1. La régulation du nombre de copies du plasmide passe par la régulation de l'expression de l'activateur de la réplication : la protéine RepA qui peut s'exprimer à partir du promoteur copBp et repAp (plus fort). La transcription à partir du promoteur repAp est bloquée par la protéine CopB. En cas d'accumulation de l'ARNm de RepA, celui–ci est inactivé suite à la liaison de ARN anti-sens copA.

Le deuxième mécanisme de régulation de la réplication est basé sur un ARN anti-sens appelé copA. Cet ARN est codé par une région située entre les gènes *copB* et *repA* et est exprimé de façon constitutive de telle sorte qu'une augmentation anormale du nombre de copies du plasmide est suivie d'une augmentation de la concentration de l'ARN anti-sens copA qui régule négativement la réplication du plasmide R1 en complexant l'ARNm de RepA.

Larsen et ses collaborateurs ont construit le plasmide thermorégulable pOU61 en remplaçant le promoteur copBp par le promoteur λp_R et en insérant dans le plasmide une séquence d'ADN codant pour l'allèle cI857 du gène du représseur λ. A 32 °C, le représseur cI857 est actif, il lie le promoteur λp_R et réprime l'expression de la protéine repA, maintenant ainsi très bas le nombre de copies du plasmide (1 à 2 copies par cellule). A des températures plus élevées (supérieures à 37 °C), le répresseur cI857 est progressivement inactivé. Le promoteur fort λp_R déréprimé dirige alors la surexpression de la protéine RepA qui augmente le nombre de copies du plasmide. Larsen a observé à 42°C une perte inattendue du contrôle du nombre de copies du plasmide pOU61 entraînant une amplification massive de la quantité d'ADN plasmidique qui peut représenter jusqu'à 1000 copies par cellule. Cette réplication incontrôlée est causée par un phénomène appelé « *la transcription convergente* » qui s'explique par le fait qu'une synthèse importante de l'ARNm de RepA à partir du promoteur λp_R provoque une réduction drastique de la production de l'ARN anti-sens CopA probablement parce que l'onde de changement du degré de super-enroulement qui se propage le long de l'ADN en 3' du promoteur λp_R a un effet drastique sur le fonctionnement du promoteur copAp.

Nous avons testé la réponse du plasmide pOU61 à la température en transformant *E. coli* et en analysant sur gel d'agarose la variation du nombre de copies du plasmide. Les résultats sont consignés dans la Figure 9. Les cellules sont d'abord cultivées à 32 °C puis transférées rapidement dans un deuxième incubateur préalablement mis à 42 °C. Toutes les 15 minutes, une fraction de culture calibrée à 2 unités de densité optique à 600 est prélevée. L'ADN plasmidique de cette fraction est extrait et quantifié sur un gel d'agarose 1%. Le nombre de copies est calculé en rapportant la quantité d'ADN quantifiée sur le gel au nombre de cellules ayant permis d'extraire cette masse de plasmide. La figure 9 montre que le nombre de copies atteint 650 environ au bout de 60 minutes à 42 °C.

Nous avons transféré les séquences contrôlant l'expression de RegB du plasmide pARNU2 au plasmide pOU61. Le plasmide recombinant obtenu, appelé pOURegB, se maintient dans la bactérie lorsque celle-ci est cultivée à 32 °C.

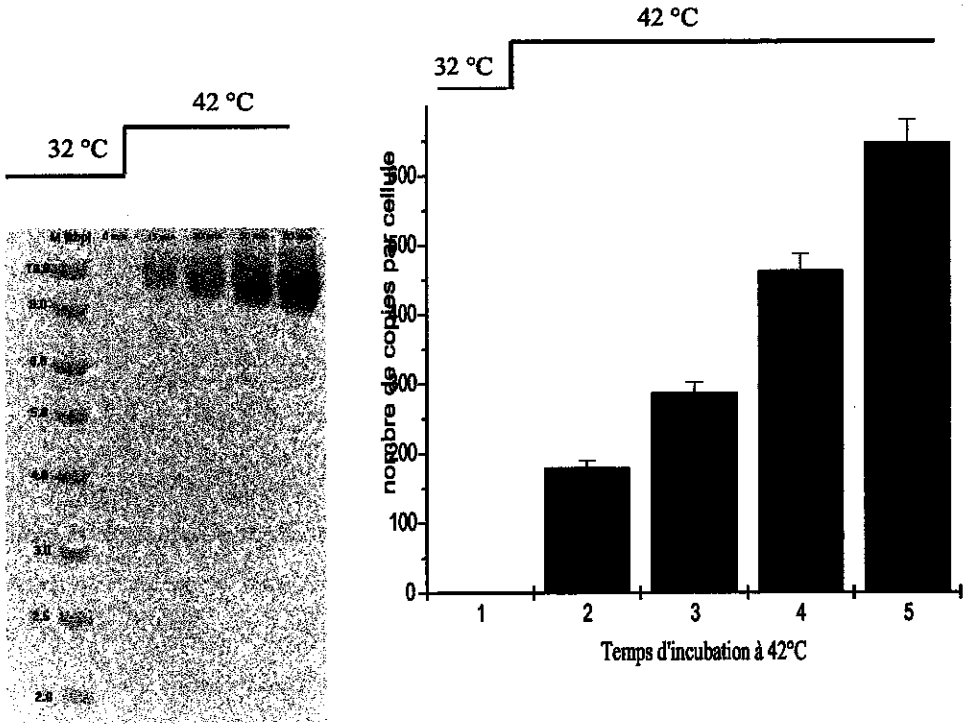


Figure 9 : Amplification du nombre de copies du plasmide pOU61 en réponse à une élévation brutale de la température. A gauche, gel d'agarose montrant le contenu plasmidique d'une fraction cellulaire correspondant à 2 unités DO_{600}. A droite, nombre de copies du plasmide pOU61 par cellule. Les chiffres 1 à 5 désignent des temps d'incubation à 42°C de 0, 15, 30, 50 et 60 min.

1.4.1.2 Caractéristique du plasmide pOURegB

L'organisation générale du plasmide pOURegB est présentée dans la Figure 10. Contrairement au plasmide pARNU2, pOURegB se maintient dans la souche BL21(λDE3) contenant une source de T7 ARN polymérase. Le fait de réduire d'un facteur 20 le nombre de copies du plasmide portant le gène *regB* a réduit la dose de RegB exprimée de façon résiduelle et a autorisé le maintient de la construction dans cette souche.

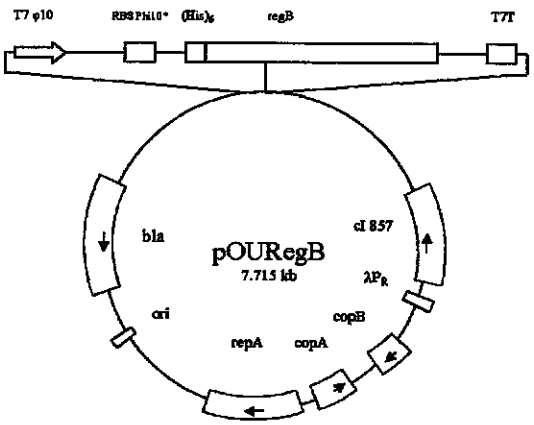


Figure 10 : Caractéristiques du plasmide pOURegB. Le gène *regB* (fusionné à une séquence codant pour une queue polyhistidines et flanqué par les séquences régulatrices dérivées du plasmide pARNU2) a été cloné dans le plasmide à réplication thermorégulable pOU61. T7φ10 : promoteur T7 Phi10 ; RBS Phi10* : RBS Phi10 modifié ; (His)6 : queue polyhistidines ; T7T : terminateur Phi ; bla : gène de résistance à l'ampicilline. Ori : origine de réplication. RepA, copA et copB : gènes de régulation du nombre de copies de pOU61. cI857 : allèle thermosensible du répresseur λ.

1.4.2 Expression de la ribonucléase RegB avec le système pOURegB

La souche BL21(λDE3) transformée avec le plasmide pOURegB est cultivée à 32 °C en milieu riche jusqu'à 0.6 unité DO_{600}. Les cellules subissent ensuite un choc thermique à 42 °C pendant une heure. Les cellules sont alors transférées à 37 °C. L'expression de la protéine est induite pendant 3 heures en ajoutant de l'IPTG (1 mM final). Le résultat est montré dans la Figure 11.

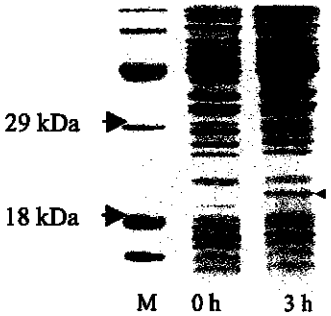


Figure 11 : Expression de la protéine RegB avec le plasmide pOURegB dans la souche BL21(λDE3). Le protocole d'expression est indiqué dans le texte. La bande correspondant à RegB est mise en évidence avec un triangle plein. M : marqueur de masses moléculaires, 0 h et 3 h : heures d'induction de l'expression à 37 °C.

Nous avons estimé le rendement d'expression de RegB avec le système pOURegB à environ 600 µg de protéines par litre de culture, ce qui représente une amélioration du rendement d'expression du système pARNU2 d'un facteur 3. Cette augmentation du rendement d'expression n'est toutefois pas proportionnelle à l'augmentation du nombre de copies du plasmide pOURegB (1-2 à 32 °C, 600 à 42 °C). Ce phénomène pourrait être due à une saturation de la machinerie de traduction avec les ARNm de RegB qui portent un site d'initiation de la traduction très peu efficace (RBS Phi10 modifié). Nous pensons que la quantité d'ARNm produite par unité de temps, avec le système promoteur T7/T7 ARN polymérase, atteint très vite une valeur saturante pour la machinerie de traduction lorsque le nombre de copies du plasmide dépasse une valeur maximale que nous appellerons N_{max}. Lorsque le nombre de copies est inférieur à N_{max}, la quantité d'ARNm effectivement traduite par unité de temps dépend de la quantité totale d'ARNm transcrits qui dépend elle même du nombre de matrices d'ADN disponibles. Lorsque le nombre de copies dépasse N_{max}, la quantité d'ARNm disponible pour la machinerie de traduction devient très vite maximale et quasi-stationnaire et ne dépend donc plus du nombre de copies du plasmide. La vitesse d'apparition de la protéine RegB n'est alors limitée que par la vitesse d'initiation de la traduction qui dépend de la force du RBS. On assiste alors à la disparition du couplage transcription-traduction qui caractérise normalement les systèmes procaryotes. Les ARNm transcrits et non traduits étant très sensibles à la machinerie ribonucléasique de la cellule, la quantité finale de RegB n'augmente pas tellement même si on allonge le temps d'induction. Cette hypothèse est en accord avec les résultats de Marc Dreyfus et de ses collaborateurs qui ont montré que la transcription d'un gène sous le contrôle du promoteur T7 peut être découplée de la traduction lorsque celle-ci est dirigée par un site d'initiation de la traduction peu efficace (Chevrier-Miller *et al.*, 1990).

Cette hypothèse engendre deux conséquences. D'abord, il inutile d'augmenter le nombre de copies du plasmide pOURegB au delà de N_{max}. Pour vérifier cette affirmation, nous avons refait le protocole d'expression de RegB en faisant un choc thermique à 42 °C pendant des intervalles de temps de 15, 30, 45 et 60 min. Après un choc thermique de seulement 15 min, alors que le nombre de copies est estimé à 100 copies par cellule, nous avons remarqué que la quantité de RegB accumulée au bout de 3 heures d'induction à 37 °C était voisine de celle obtenue après un choc thermique de 60 min. Le N_{max} est probablement inférieur ou égal à 100. Par ailleurs, nous pensons que le l'augmentation de l'efficacité du RBS Phi10 (remplacement du RBS Phi10 modifié par le RBS Phi10 sauvage) est l'unique moyen pour améliorer le rendement d'expression de RegB avec ce système thermorégulable. Nous avons pu vérifier cette déclaration en clonant une protéine modèle non toxique, le fragment N-terminal de RegB, dans le plasmide pOU61 sous le contrôle du RBS Phi10 modifié (construction pOURN) et sous le contrôle du RBS Phi10 sauvage (construction pOURNb).

Nous avons testé l'expression du fragment N-terminal de RegB avec ces deux constructions. Le résultat est présenté dans la Figure 12.

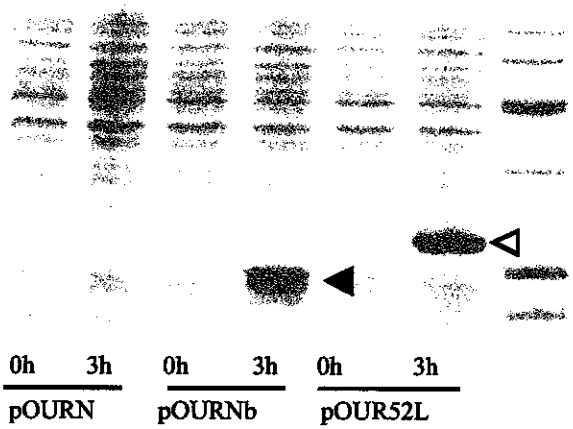


Figure 12 : Expression du fragment N-terminal de RegB avec le système thermorégulable en utilisant deux versions différentes du RBS Phi10. pOURN contient la partie N-terminale de RegB (triangle plein), clonée dans pOU61, avec un RBS Phi10 modifié. pOURNb contient la partie N-terminale de RegB, clonée dans pOU61, avec le RBS Phi10 sauvage. Le gel montre, de plus, le résultat obtenu avec la construction pOUR52L qui contient le mutant ponctuel inactif R52L (triangle vide), cloné dans pOU61, avec le RBS Phi10 sauvage. Le protocole d'expression est identique pour les trois constructions : culture à 32 °C jusqu'à 0.6 unité DO_{600}, choc thermique à 42 °C pendant 30 minutes puis incubation à 37 °C pendant 3 heures en présence de 1 $mmol.l^{-1}$ d'IPTG.

La Figure 12 ci-dessus montre qu'en présence d'un RBS Phi10 modifié le fragment N-terminal est très faiblement exprimé même après l'augmentation du nombre de copies du plasmide. Lorsque ce même fragment est mis sous le contrôle du RBS Phi10 sauvage, sa bande de surexpression représente près de 50 % des protéines totales d'*E. coli*. Le même résultat a été obtenu avec le mutant ponctuel inactif RegB R52L.

Malheureusement, nous avons pu vérifier expérimentalement que l'introduction du RBS Phi10 sauvage dans la construction pOURegB ne donne aucun clone viable même dans une souche dépourvue de T7 ARN polymérase maintenue à une température permissive (32 °C). Ceci s'explique par le fait que le facteur 50 gagné dans la force du RBS Phi10 s'accompagne d'une augmentation d'un facteur identique de la dose de RegB exprimée de façon résiduelle. La diminution du nombre de copies du plasmide d'un facteur 20 n'est pas suffisante pour que la quantité de RegB exprimée de façon résiduelle soit en dessous du seuil de viabilité de la bactérie.

Nous avons résumé dans la Figure 13 l'ensemble des conclusions auxquelles nous sommes parvenues suite à l'étude de l'expression de RegB dans la bactérie *E. coli.*

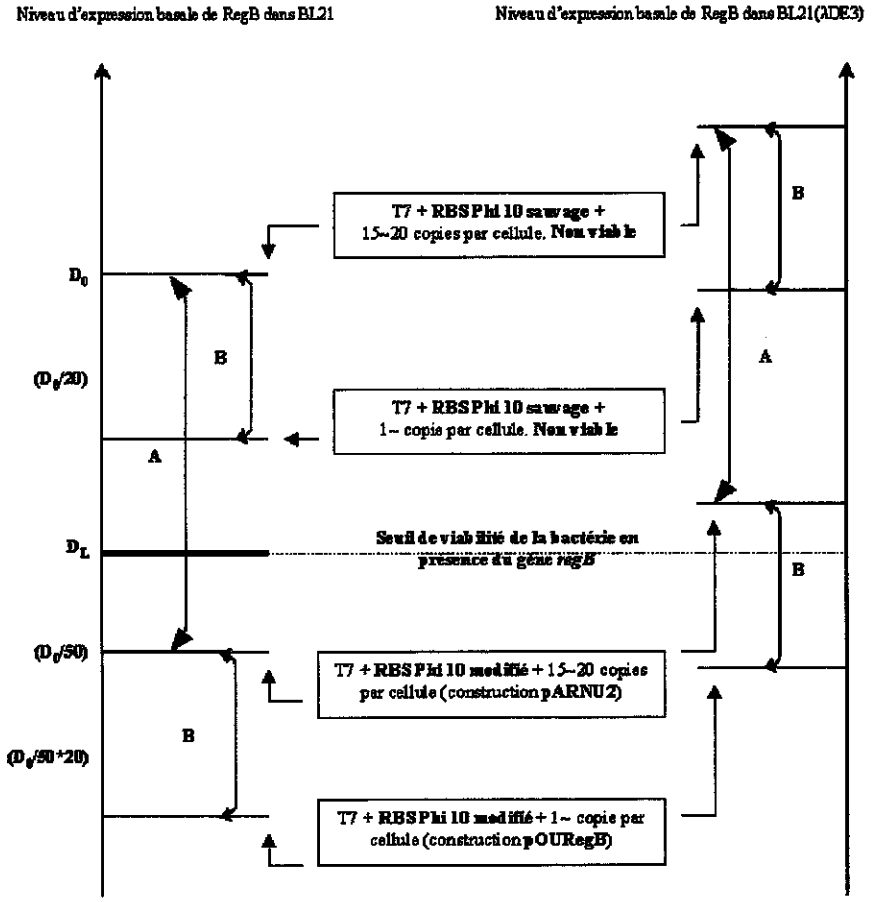


Figure 13 Diagramme de maintient du gène *regB* sous le contrôle du promoteur T7 dans une souche bactérienne dépourvue de T7 ARN polymérase (BL21) et dans une souche contenant cette activité (BL21(λDE3)). Le diagramme n'est pas à l'échelle. La dose de RegB exprimée de façon basale est indiquée sur l'axe vertical. Aucune construction n'est maintenue dans la bactérie lorsque la dose associée est strictement supérieure à une valeur limite D_L. Le système pET classique (RBS Phi10 sauvage + 15 ~ 20 copies par cellule) exprime une dose D_0 largement supérieure à la dose limite D_L et est donc impossible à utiliser avec le gène *regB*. La modification (A) a permis en atténuant l'efficacité du RBS Phi10 de réduire cette dose d'un facteur 50 la plaçant ainsi sous la valeur limite D_L (construction pARNU2). La modification (B) a réduit le nombre de copies du plasmide pARNU2, donc l'expression basale de RegB, d'un facteur 20 environ (soit au total un facteur 50 x 20 = 1000 par rapport au système pET classique). La construction qui en a résulté (pOURegB) est évidemment viable dans une bactérie dépourvue de T7 ARN polymérase ainsi que dans la souche BL21(λDE3).

Les trois systèmes d'expression bactériens que nous venons de voir permettent d'exprimer la ribonucléase RegB avec des rendements de 200 µg/l (système λCE6 en incubateur), 300 µg/l (système λCE6 en fermenteur) et de 600 µg/l (système pOURegB). Il s'agit de systèmes dont les rendements d'expression sont convenables mais dont la technique est fastidieuse. Aussi, nous avons cherché un moyen plus simple d'exprimer RegB sans passer par la préparation de lysat de virus ni par la manipulation de la température de culture. Nous avons pensé à changer complètement de mode d'expression et à utiliser une cellule hôte eucaryote.

1.5 Expression de RegB dans un système eucaryote

1.5.1 Justification du choix

La toxicité de la ribonucléase RegB pour la bactérie *E. coli* est probablement liée à sa capacité à cliver la séquence GGAG qui fait partie du site de liaison au ribosome de plus de 32 % des ARNm de la bactérie. Par ailleurs, il est établi que RegB exploite à travers la protéine ribosomique S1 la machinerie de traduction bactérienne (Ruckman *et al.*, 1994).

Nous avons donc pensé qu'une expression dans un système eucaryote permettrait de s'affranchir de tout problème lié à la toxicité de RegB. Cette idée est motivée par le fait que contrairement aux systèmes procaryotes, l'initiation de la traduction chez les eucaryotes est indépendante de la présence de la séquence GGAG (McKendrick *et al.*, 2003) et ne requiert pas l'intervention d'un homologue de la protéine S1. En l'absence de toute donnée expérimentale relative au comportement de RegB dans un système eucaryote nous avons décidé de tester le clonage et l'expression de RegB dans le système eucaryote le mieux connu : Le système *Saccharomyces cerevisiae (S. cerevisiae).*

1.5.2 Construction des vecteurs d'expression pYRegB1 et pYRegB2

Nous avons choisi de transférer le gène *regB* dans le vecteur navette pYES2 (Invitrogen) pour quatre raisons. D'abord, ce vecteur contient, en plus de ses éléments eucaryotes, une origine de réplication procaryote et un système de sélection à l'ampicilline qui permet de réaliser toutes les étapes de clonage dans la bactérie. Par ailleurs, ce vecteur met le gène *regB* sous le contrôle

transcriptionnel du promoteur GAL1 connu pour sa forte capacité transcriptionnelle inductible au galactose (Stepien *et al.*, 1983). De plus, ce plasmide se maintient dans la levure avec un nombre de copies élevé grâce à son origine de réplication eucaryote de type 2μ (Huberman *et al.*, 1987). Finalement, ce vecteur porte le gène URA3 qui permet de sélectionner très facilement les levures transformées en utilisant l'auxotrophie pour l'uracile (Blanc *et al.*, 1979).

L'initiation de la traduction chez les eucaryotes ne requiert pas la présence de séquences de type Shine-Dalgarno (Kozak M., 1999). Le ribosome s'associe à l'extrémité 5' coiffée puis balaie l'ARNm dans le sens 3' à la recherche du premier codon ATG. Marylin Kozak a montré que l'initiation de la traduction est très efficace lorsque ce codon ATG se trouve dans une séquence consensus : ACC **ATG** G (Kozak M., 1986). Cette séquence porte le nom de séquence Kozak.

Nous avons d'abord cloné le gène *regB* dans le vecteur d'expression pYES2 et avons flanqué l'ATG initiateur avec la séquence consensus Kozak. Le vecteur résultant, appelé pYRegB1 (voir Figure 14) se maintient sans problème dans la levure. Les tests d'expression ont montré, en revanche, que RegB est produite en très faible quantité, difficilement détectable même par une analyse anticorps (western-blot). Plusieurs travaux ont suggéré qu'en plus de la séquence consensus entourant l'ATG initiateur, l'efficacité de la traduction des ARNm dans *S. cerevisiae* dépend fortement des structures secondaires qui peuvent se former autour de cette région (Pompon *et al.*, 1996 ; Pompon D., communication personnelle). Ces structures peuvent, en effet, séquestrer le codon initiateur empêchant ainsi le démarrage de la traduction. Convaincus de l'existence d'un problème traductionnel dans l'expression de RegB à partir du plasmide pYRegB1, nous avons calculé avec le programme Mfold (Zucker M., 2003) les structures possibles que peut adopter la région 5' de l'ARNm de RegB. Le résultat est présenté dans la Figure 14.

La forte structuration de l'ARNm autour de l'ATG initiateur est probablement à l'origine du très faible rendement d'expression obtenu avec pYRegB1. Pour remédier à ce problème nous avons construit un nouveau vecteur, appelé pYRegB2, dans lequel on a remplacé la séquence en 5' de l'ATG initiateur par une séquence prédite par le programme Mfold comme étant peu structurée (voir Figure 15).

Pour exprimer RegB, nous avons transformé la levure *S. cerevisiae* avec le plasmide pYRegB2. Nous avons aussi transformé la même souche avec le vecteur parental pYES2 (témoin négatif).

Nous n'avons pas constaté de différence ni dans la vitesse de croissance ni dans la morphologie des cellules transformées par pYRegB2 ou par pYES2. Lorsqu'on déréprime le système d'expression en ajoutant le Galactose, on observe un léger retard de croissance des cellules portant

pYRegB2 attribuable plutôt à la mobilisation du métabolisme de la levure pour surexprimer la protéine recombinante qu'à un effet toxique de la ribonucléase RegB.

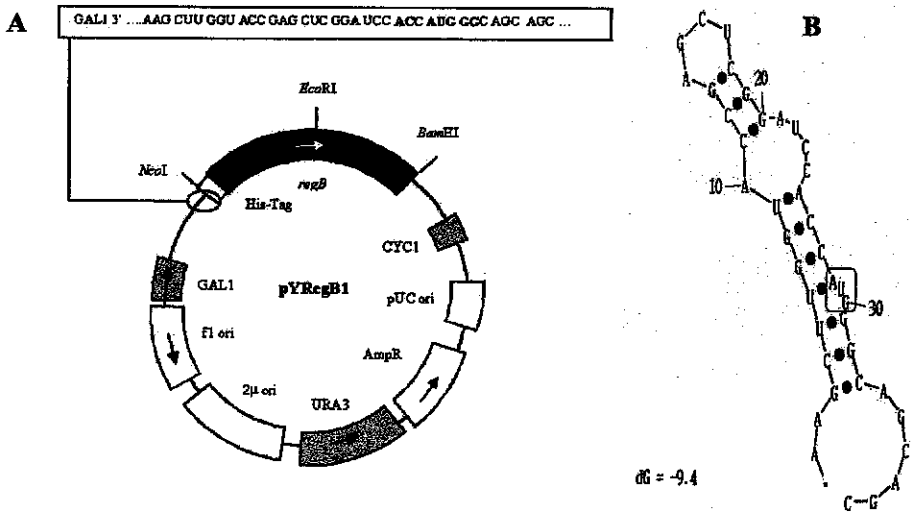


Figure 14 : Expression de RegB dans la levure avec le système pYRegB1. (A) Organisation du plasmide pYRegB1. Le gène *regB* a été cloné, en plusieurs étapes, avec une queue polyhistidines entre les sites PvuII et XhoI du vecteur pYES2. Le codon de démarrage de la traduction de l'ARNm de *regB* correspond à une séquence d'ADN contenant le site de restriction NcoI. La séquence d'ADN autour de ce site est montrée dans l'encadré. GAL1 : promoteur de la galactokinase, CYC1 : terminateur transcriptionnel du gène de iso-1-cytochrome c, URA 3 : gène de l'orotidine-5'-monophosphate décarboxylase, AmpR : gène de la β-lactamase, pUC ori et f1 ori : origines de réplication procaryotes, 2μ ori : origine de réplication de levure. (B) Structure secondaire autour de l'ATG initiateur adoptée par l'ARNm de RegB obtenu à partir de pYRegB1. La structure, calculée avec le programme Mfold, est très stable (sa variation d'enthalpie libre standard G vaut -9.4 kcal/ mole soit -39.3 kj/mole).

La Figure 16 présente le résultat d'une analyse western blot de l'expression de RegB dans la levure. Alors que l'expression est indétectable avec le plasmide pYRegB1, une bande de surexpression est visible avec le plasmide pYRegB2. Cette surexpression est évidemment plus faible que celle obtenue dans la bactérie avec un système pET/BL21(λDE3) classique. Les systèmes d'expression dans la levure sont, en effet, connus pour être environ 20 fois moins efficaces que les systèmes d'expression bactériens. Nous avons estimé le rendement d'expression de RegB avec le système pYRegB2 à environ 300 μg de protéines par litre de culture.

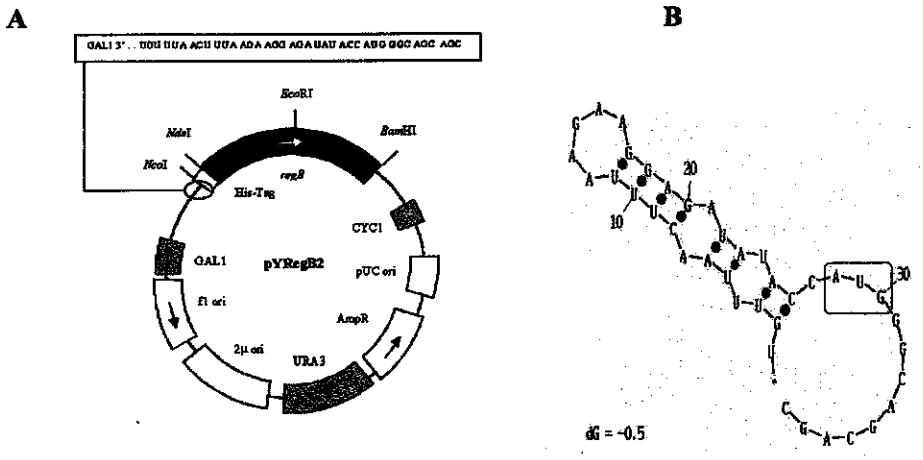


Figure 15 : Expression de RegB dans la levure avec le système pYRegB2. (A) Organisation du plasmide pYRegB2. (B) Structure secondaire autour de l'ATG initiateur adopté par l'ARNm de RegB obtenu à partir de pYRegB2. Les notations sont identiques à celles de la Figure 14.

Figure 16 : Analyse western blot de l'expression de la ribonucléase RegB dans la levure. *S. cerevisiae* est transformée avec pYRegB1 (A) ou pYRegB2 (B), mise en culture dans un milieu synthétique minimal sans uracile puis induite pendant 8 heures avec du galactose 2 %. Un dixième de l'extrait cellulaire brut correspondant à 20 unités de DO_{600} est séparé par SDS-PAGE. La protéine RegB est révélée en utilisant un anticorps primaire spécifique de la queue polyhistidines et un anticorps secondaire couplé à la phosphatase alcaline. (C) Piste témoin : la bactérie BL21(λDE3) est transformée avec un plasmide surexprimant le mutant ponctuel inactif RegB R52L et induite avec l'IPTG pendant 3 heures (cette piste a été grossièrement calibrée à la même quantité totale de biomasse que les cultures de levures).

Nous avons remarqué que les cellules de levures transformées avec le plasmide pYRegB2 conservent leur capacité à exprimer RegB pendant plusieurs générations. Cette stabilité renforce l'hypothèse de l'absence de toxicité de RegB dans un système eucaryote. A titre de comparaison, les cellules BL21(λDE3) transformées avec le plasmide pOURegB (maintenu à 1 copie par cellule)

perdent leur capacité à exprimer RegB lorsque les cellules sont repiquées plus de deux fois de suites sur un milieu de sélection. La toxicité résiduelle de RegB dans cette souche sélectionne la descendance contenant des constructions réarrangées ayant perdu leur capacité à exprimer l'enzyme.

1.6 Avantages et inconvénients des différents systèmes d'expression de RegB

Les caractéristiques principales, les avantages et les inconvénients des 4 systèmes d'expression de RegB que nous avons testés sont résumés dans le Tableau 1.

Désignation	**pARNU2/λCE6 en incubateur**	**pARNU2/λCE6 en fermenteur**	**pOURegB**	**pYRegB2**
Transcription	Promoteur Phi10 de T7	Promoteur Phi10 de T7	Promoteur Phi10 de T7	Promoteur GAL1
Traduction	RBS Phi10 atténué	RBS Phi10 atténué	RBS Phi10 atténué	Consensus Kozak + absence structures secondaires
Réplication	Type pBR322 (15 ~ 20 copies)	Type pBR322 (15 ~20 copies)	Type R1 (1~2 copies à 32 °C, 600 copies à 42 °C)	Type 2μ
Hôte	*E. coli* BL21 ou XL1-Blue	*E. coli* BL21 ou XL1-Blue	*E. coli* BL21(λDE3)	Levure *S. cerevisiae*
(μg/litre)	200	300	600	300
Avantages	1- Transformants stables 2- Culture et expression à la même température (37 °C) 3- Vitesse de croissance bactérienne normale	1- Manipulation de grands volumes de culture 2- Contrôle des conditions de culture (pO_2, pH ...) 3- Rendement d'expression meilleur qu'en incubateur	1- Système d'induction chimique simple (IPTG) 2- Rendement d'expression meilleur que le système pARNU2/λCE6	1- Transformants très stables 2- - Système d'induction chimique simple et peu onéreux (galactose)
Inconvénients	1- Système d'induction fastidieux (λCE6) 2- Faible rendement d'expression	1- Procédure nécessitant un équipement lourd ne pouvant être de routine dans un laboratoire	1- Le saut de température nécessite l'utilisation simultanée de deux incubateurs 2- Les transformants sont instables	1- Temps de culture et d'expression deux fois plus lents que dans la bactérie 2- Cassage des cellules difficile 3- Susceptibilité aux protéases et aux modifications post-traductionnelles inutiles.

Tableau 1 : Les caractéristiques génétiques, les avantages et les inconvénients des 4 systèmes d'expression de RegB.

1.7 Mise au point d'un nouvel outil d'ingénierie génétique exploitant le système RegB

1.7.1 Principe

Nous avons montré dans les sous-chapitres précédents que la ribonucléase RegB présente un profil de toxicité différentiel à l'égard du système procaryote *E. coli* et du système eucaryote *S. cerevisiae*. Nous avons voulu exploiter cette propriété originale de RegB pour construire, dans la levure, un vecteur contenant le gène *regB* sous le contrôle du promoteur Phi10 de T7, du RBS Phi10 sauvage et du terminateur de transcription Phi. Un tel vecteur se maintient très facilement dans la levure et peut être produit et purifié à partir de cet organisme. En revanche, il ne peut se maintenir dans la bactérie *E. coli* que si un insert vient disrupter la séquence codante de RegB annulant ainsi l'effet toxique de la protéine. Ce vecteur peut donc être utilisé comme un vecteur de clonage à sélection positive. Par ailleurs, nous avons pensé qu'en emboîtant les signaux d'expression eucaryote et procaryote dans un tel plasmide, on pourrait en faire un vecteur d'expression duale dans la bactérie et dans la levure.

1.7.2 Particularité du système de sélection positive utilisant RegB

Plusieurs systèmes de sélection positive ont été décrits dans la littérature et sont fréquemment utilisés dans les laboratoires pour forcer l'étape de clonage à sélectionner uniquement les clones bactériens qui portent le plasmide recombinant c'est à dire le plasmide contenant l'insert étudié (Choi *et al.*, 2002). Malheureusement, ces systèmes présentent tous des limitations liées à la nature de la souche bactérienne utilisée et à la nécessité d'étaler les bactéries sur des boites de pétri contenant un inducteur chimique (IPTG).

Le système de sélection coloré blanc/bleu, par exemple, est basé sur la disruption de la séquence codante du peptide α de la β-galactosidase (Vieira *et al.*, 2002). Bien que très pratique, l'usage de ce système est limité aux souches bactériennes portant la délétion Δ(lacZ)M15. De plus, il requiert l'induction de l'expression du peptide α en ajoutant de l'IPTG sur les boites de pétri. Par conséquent, la protéine qui remplace le peptide α se trouve elle même induite ce qui rend le système de sélection inutilisable dans le cas où l'insert code pour une protéine toxique.

D'autres systèmes de sélection positive sont basés sur l'inactivation d'un agent toxique. La ribonucléase Barnase (Deyev *et al.*, 1998), la protéine CcdB (Control cell death ; Bernard *et al.*, 1994)

et le domaine de liaison à l'ADN du facteur de transcription GATA-1 (Trudel *et al.*, 1996) sont utilisés dans plusieurs vecteurs. Comme dans le système blanc/bleu, ces systèmes nécessitent l'ajout d'IPTG sur la boite de pétri. Ceci rend le travail de sélection fastidieux puisqu'à cause de la sensibilité différentielle des souches d'*E. coli* pour l'IPTG, sa concentration doit être optimisée avant d'entreprendre le processus de clonage. Par ailleurs, dans le cas de la protéine CcdB, il n'est pas possible d'utiliser des souches bactériennes contenant le plasmide F car ce dernier code pour la protéine CcdA, un inhibiteur naturel de la protéine CcdB.

Nous avons montré dans le chapitre précédent que la ribonucléase RegB est létale pour la bactérie *E. coli* quand elle est couplée à la fois au promoteur Phi10 de T7 et au RBS Phi10 sauvage. Cette létalité est intéressante parce qu'elle ne demande pas de « forcer » l'expression de RegB en ajoutant de l'IPTG. Les doses de RegB exprimées de façon basale suffisent pour tuer la bactérie. Nous avons pu, par ailleurs, vérifier que cette létalité ne dépend pas de la souche bactérienne utilisée. Les souches XL1-Blue et BL21(λDE3) connues pour leurs niveaux d'expression basale très distincts ont été testées pour leur capacité à maintenir le gène *regB* couplé au promoteur Phi10 de T7 et au RBS Phi10 sauvage. Aucun clone viable n'a été obtenu suite à la transformation par la construction contenant le gène *regB*.

1.7.3 Construction du vecteur pTOXR-1

Le gène *regB* sauvage a été cloné dans le vecteur navette eucaryote/procaryote pYES2 sous la forme d'une séquence d'ADN encadrée par le promoteur T7 Phi10, l'opérateur lac, le RBS Phi10 sauvage et le terminateur Phi. Le gène *regB* a été inséré de telle façon qu'il soit aussi le sous contrôle transcriptionnel eucaryote du promoteur GAL1 et du terminateur CYC1. Le plasmide recombinant, appelé pTOXR-1, a été construit en deux étapes (voir Figure 17).

Nous avons construit dans un premier temps deux fragments d'ADN A et B. Dans le fragment A, la partie 5' du gène *regB* est fusionnée à une séquence codant pour une queue polyhistidines et est mise sous le contrôle du promoteur T7lac et du RBS Phi10 extraits du vecteur d'expression pET28b. Le fragment B contient la partie 3' du gène *regB* incluant son codon stop et le terminateur de transcription Phi extrait du vecteur pET15b. Ce fragment a été obtenu à partir du vecteur pERL52 qui contient le mutant ponctuel inactif RegB R52L. Le fragment A a été cloné dans le vecteur d'expression de levure pYES2 dont le site unique NdeI a été préalablement supprimé en utilisant la technique de remplissage avec la T4 ADN polymérase. Dans le plasmide recombinant ainsi obtenu (pYTOX), le fragment A est localisé en aval du promoteur GAL1 et en amont du

terminateur CYC1. Nous avons inséré dans une deuxième étape le fragment d'ADN B dans le plasmide pYTOX et avons utilisé la solution de ligature pour transformer non pas *E. coli* mais *S. cerevisiae.* Les cellules de levure recombinantes ont été criblées par PCR. Le plasmide recombinant final pTOXR-1 (voir Figure 18) a été extrait de la levure en utilisant un protocole très efficace inspiré de celui mis au point par Robzyk K. et Kassir Y. (Robzyk *et al.*, 1992).

Nous avons vérifié que ce plasmide ne donne aucun clone bactérien viable quand on l'introduit dans les souches bactériennes les plus courantes (BL21, DH5α, JM101, XL1-Blue et BL21(λDE3).

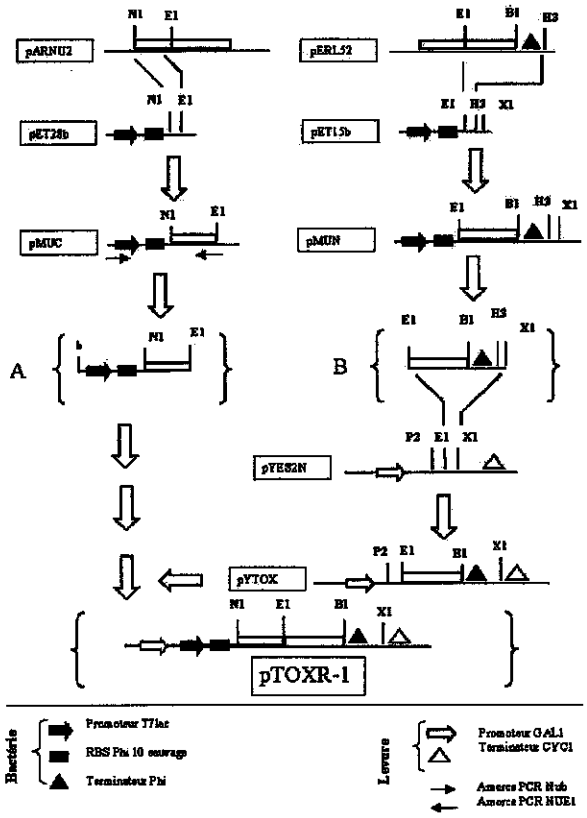


Figure 17 : Construction du vecteur pTOXR-1. Les fragments d'ADN ne sont pas à l'échelle. Le clonage des inserts NdeI-EcoRI (partie 5' de *regB*) de pARNU2 et EcoRI-HindIII (partie 3' de *regB*) de pERL52 dans le vecteur d'expression pET28b a donné les plasmides pMUC et pMUN respectivement. Le fragment A est le produit PCR de l'amplification de pMUC avec les amorces NUE1 et NUb. Le fragment B est l'insert EcoRI-XhoI excisé du plasmide pMUN. pYES2N est le plasmide pYES2 débarrassé de son unique site NdeI. L'insertion du fragment B dans pYES2N a engendré le plasmide pYTOX. L'insertion du fragment A entre les sites PvuII et EcoRI de pYTOX a engendré pTOXR-1. B1 : BamHI, E1 : EcoRI, H3 : HindIII, N1 : NdeI, P2 : PvuII, X1 : XhoI, b : extrémité franche d'un produit de PCR.

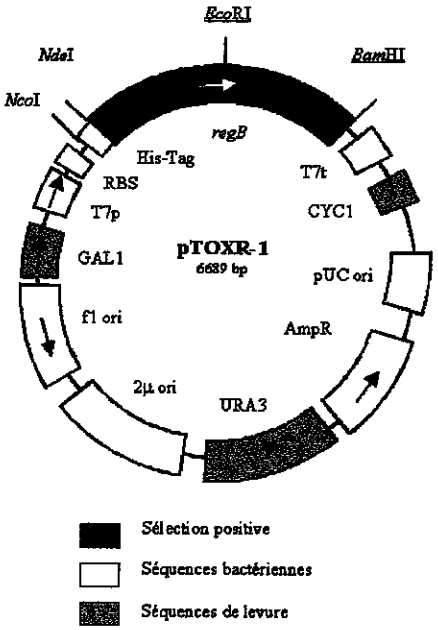


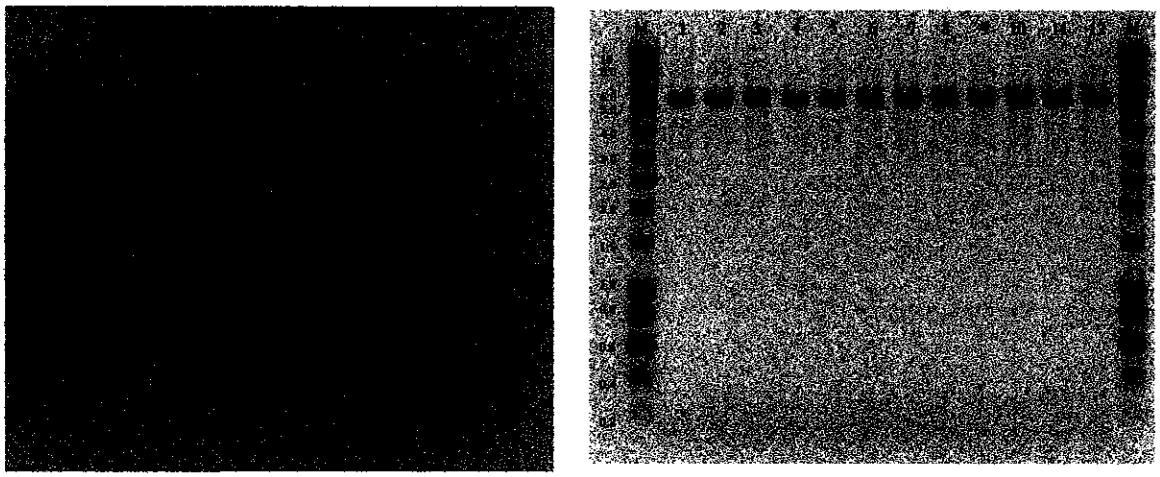
Figure 18 : Caractéristiques génétiques du plasmide pTOXR-1. Les séquences d'ADN ne sont pas à l'échelle. Les flèches indiquent l'orientation transcriptionnelle des gènes. RBS : RBS Phi10 de T7, AmpR : déterminant de résistance à l'ampicilline, f1 ori : origine de réplication produisant un ADN simple brin, T7p : promoteur T7 + opérateur lactose, T7t : terminateur Phi, 2µ : origine de réplication de levure, GAL1 : promoteur de la galactokinase, CYC1 : terminateur transcriptionnel du gène de iso-1-cytochrome c, URA 3 : gène de l'orotidine-5'-monophosphate décarboxylase. Les sites de restriction uniques sont soulignés.

1.7.4 Le système de sélection positive de pTOXR-1 : Application au clonage du bimodule 4-5 de la protéine ribosomique S1

Nous avons choisi de tester le système de sélection positive de pTOXR-1 sur un fragment d'ADN codant pour les deux modules 4 et 5 de la partie C-terminale de la protéine ribosomique S1 d'*E. coli* en cours d'étude dans le laboratoire (Bisaglia *et al.*, 2003). Ce fragment sera désigné par la suite sous le nom de « bimodule 45 ». Il contient un site NdeI dans son extrémité 5' et un codon stop associé à un site BamHI dans son extrémité 3'. Nous avons donc choisi de l'insérer entre les sites NdeI et BamHI de du vecteur pTOXR-1 sans prêter attention à la digestion complète du vecteur et sans lui faire subir un traitement particulier pour éviter la re-circularisation (purification sur gel d'agarose pour éliminer l'insert parental ou déphosphorylation des extrémités cohésives). Le détail du

protocole « clonage en trois heures » est présenté dans la publication n°1. Nous avons par ailleurs appliqué le même protocole au vecteur standard pET15b qui a servi de témoin.

A l'issue de l'étape de transformation de la souche bactérienne XL1-Blue avec les produits de ligature nous avons obtenu 53 clones viables avec pTOXR-1 et 245 avec pET15b. Le résultat de l'analyse de restriction avec les enzymes NdeI et BamHI de 12 clones choisis aléatoirement est présenté dans la Figure 19.



A B

Figure 19 : Analyse de restriction NdeI-BamHI de 12 clones bactériens issus du clonage du bimodule 45 en utilisant le système de sélection positive de pTOXR-1 (A) ou le vecteur classique pET15b (B). Les pistes pS145 et pERL52 de la Figure (A) sont des pistes témoins montrant la taille de l'insert (en paires de bases) correspondant au bimodule 45 et à la protéine RegB (mutant R52L) respectivement. M : marqueur de masses moléculaires (en milliers de paires de bases).

Le système pTOXR-1 a permis d'obtenir 11 clones recombinants sur 12. En revanche, en utilisant le système pET15b, tous les clones sélectionnés sont des plasmides parentaux dépourvus d'insert. Ce résultat montre l'efficacité du système de sélection positive basé sur la ribonucléase RegB. Nous avons choisi arbitrairement un plasmide recombinant obtenu à partir de pTOXR-1, nous l'avons appelé pXS145 et l'avons utilisé pour transformer la bactérie *E. coli* et la levure *S. cerevisiae* afin de tester sa capacité à exprimer le bimodule 45 dans deux hôtes différents.

1.7.5 Le système d'expression duale de pTOXR-1

Dans la bactérie, la traduction commence au niveau du premier codon ATG rencontré en aval du site de liaison au ribosome (Shine *et al.*, 2001 ; Kozak *et al.*, 1999). Dans pTOXR-1, cet ATG initiateur se trouve dans le site NcoI (CC **ATG** G) (voir la Figure 18). Dans la levure, comme nous l'avons déjà décrit dans le paragraphe 1.5.2, la traduction démarre au premier ATG rencontré à partir de l'extrémité 5' de l'ARNm. Nous avons émis l'hypothèse que cet ATG serait confondu avec l'ATG initiateur bactérien puisqu'il il n'existe aucun autre codon entre l'extrémité 3' du promoteur GAL1 et le site NcoI. Par ailleurs, nous avons été agréablement surpris de voir que ce site est flanqué par une séquence qui respecte exactement le consensus Kozak ACC ATG G. Si ce raisonnement est correct, le plasmide pXS145 du paragraphe précédent devrait exprimer la même protéine, fusionnée à une queue polyhistidines, dans la bactérie et dans la levure.

Nous avons utilisé 50 ml de culture pour exprimer le bimodule 45 dans la levure *S. cerevisiae* et 5 ml pour l'exprimer dans la bactérie BL21(λDE3). Les protéines exprimées dans les deux cas ont été purifiées sur une colonne de cobalt, précipitées à l'acide trichloracétique et analysées par Western-blot. Le résultat est montré sur dans la Figure 20.

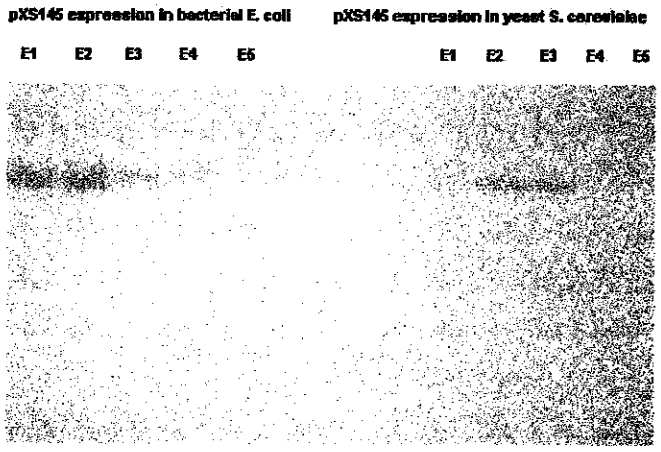


Figure 20 : Analyse Western-blot de l'expression du bimodule 45 avec la construction pXS145 dans *S. cerevisiae (à droite)* et dans *E. coli (à gauche).* E1 à E5 : fractions d'élution obtenues à l'issue de l'étape de purification sur une colonne de cobalt. L'analyse Western-blot est basée sur un anticorps qui reconnaît la queue polyhistidines.

La Figure 20 montre que les deux protéines fusionnées à une queue polyhistidines exprimées dans la levure et dans la bactérie ont la même masse moléculaire. Ce résultat montre que les traductions procaryote et eucaryote ont débuté au même ATG. Nous avons, par ailleurs, vérifié que la

protéine produite avec le plasmide pXS145 possède la même taille que celle produite à partir de la construction pS145 issue du vecteur standard pET15b (voir Figure 21). Ce résultat montre que la protéine produite avec le système pXS145 se termine au codon stop prévu.

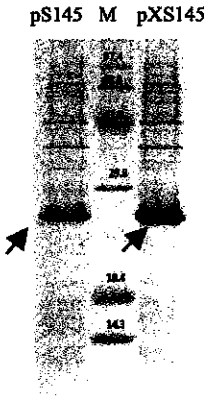


Figure 21 : Expression du bimodule 45 dans la bactérie BL21(λDE3) transformée avec le plasmide pS145 à gauche (issu de pET15b) ou pXS145 à droite (issu de pTOXR-1). Les bandes de surexpression sont indiquées avec des flèches. M : marqueur de masse moléculaire en kDa.

1.7.6 Perspectives d'utilisation du plasmide pTOXR-1

Le système pTOXR-1 présenté ci-dessus permet de cloner rapidement un fragment d'ADN, de sélectionner les plasmides recombinants avec une efficacité avoisinant les 92 % et d'exprimer la protéine d'intérêt dans la bactérie et dans la levure sans étapes supplémentaires de sous-clonage.

Pour le moment, ce système est limité à des inserts de type NdeI-EcoRI ou NdeI-BamHI. La versatilité de ce système peut être étendue en incluant dans le plasmide des sites multiples de clonage dans le site BamHI situé à l'extrémité 3' de gène *regB*. Ceci permettrait de cloner des inserts contenant un site NdeI à leur extrémité 5' et une variété de sites de restrictions dans leur extrémité 3'. Poly1 (GGATCC GGTACC TTAATTAA GTCGAC AGATCT GGATCC) et Poly2 (GGATCC CAGCTG GAGCTC CAATTG GGCGCC GGATCC) sont deux exemples de sites multiples de clonage introduisant respectivement les sites (*Bam*HI, *Kpn*I, *Pac*I, *Sal*I et *Bgl*II) et (*Bam*HI, *Pvu*II, *Sac*I, *Mun*I et *Nar*I).

Chapitre 2

Caractérisation du chemin réactionnel catalysé par RegB

Chapitre 2

Caractérisation du chemin réactionnel catalysé par RegB

La ribonucléase RegB catalyse le clivage de la liaison phosphodiester se trouvant au milieu de certains motifs GGAG, entre le deuxième G et le A (Uzan *et al.*, 1988). Comprendre la base moléculaire de la coupure spécifique catalysée par RegB nécessite de définir le site actif de l'enzyme. En absence de toute homologie significative avec une enzyme connue et ne disposant pas de la structure tridimensionnelle d'un complexe enzyme-substrat, nous avons entrepris l'identification du site actif de RegB en utilisant une approche mécanistique. Afin de limiter le champ d'investigation, nous avons exploité le fait que les ribonucléases catalysent le clivage de la liaison phosphodiester en adoptant des stratégies similaires que l'on peut classer en fonction du chemin réactionnel c'est à dire de l'ensemble des actes élémentaires et des intermédiaires réactionnels qui apparaissent lors de coupure de l'ARN substrat. Identifier le chemin réactionnel de RegB permet de l'associer à un groupe particulier de ribonucléases dont on connaît les résidus du site actif.

Après avoir décrit les caractéristiques de la liaison phosphodiester, nous montrerons d'abord que les ribonucléases utilisent un nombre restreint de mécanismes pour cliver cette liaison. Nous indiquerons ensuite ceux susceptibles d'être suivis par RegB. Nous exposerons enfin le modèle expérimental et la stratégie que nous avons suivis pour identifier le chemin réactionnel de RegB ainsi que les conséquences qui découlent de cette identification.

2.1 La liaison phosphodiester

La liaison phosphodiester est par définition la liaison chimique qui se forme lors de l'estérification de l'acide phosphorique par deux fonctions alcools. Elle désigne indistinctement les deux liaisons covalentes O-P-O notées A et B sur la Figure 22. Toutefois, il est d'usage dans la littérature de considérer la liaison C-O (liaison C sur la Figure 22) comment faisant partie de la liaison phosphodiester.

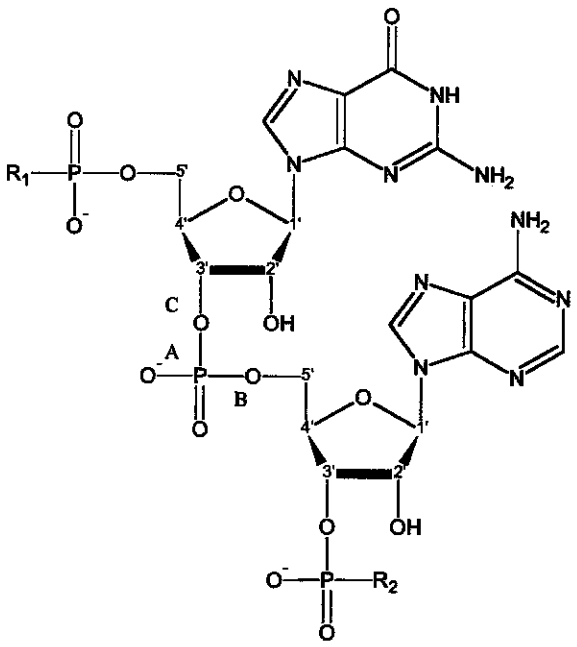


Figure 22 : Une liaison phosphodiester de type G-A. 3 liaisons covalentes différentes, notées A B C, sont désignées sous le terme générique de liaison phosphodiester.

Pour orienter précisément la liaison covalente nous adopterons une nomenclature unique 5'-3' relative à la molécule entière d'ARN. Ainsi, la liaison A sera appelée liaison O3'-P, la liaison B P-O5' et la liaison C C3'-O3'. Cliver une liaison phosphodiester c'est rompre l'une des 3 liaisons covalentes A, B ou C. Dans ce qui suit nous nous limiterons à l'étude du clivage des liaisons A et B.

Le clivage de la liaison phosphodiester par rupture de la liaison covalente carbone-oxygène C n'a jamais été décrit pour des ribonucléases. Il est utilisé par des désoxyribonucléases (enzymes dont le substrat est l'ADN) comme l'Endonucléase V (Fuxreiter *et al.*, 1999) et l'Endonucléase III (Fromme *et al.*, 2003) pour cliver la liaison phosphodiester au niveau d'un site pyrimidique endommagé.

Il est important de noter que la liaison phosphodiester est naturellement très résistante au clivage. A titre d'exemple, le temps de demi-vie du phosphodiester le plus simple (le diméthylphosphate) en présence de 1 molaire d'hydroxyde de sodium à 35 °C est de 15 ans (Chin *et al.*, 1989).

Cette résistance au clivage par attaque nucléophile est attribuée en partie à l'état d'ionisation (monoanion) de la liaison phosphodiester dans une large gamme de pH (la constante d'acidité du dernier groupe OH du phosphore est de 2.4, le groupe se trouve donc majoritairement sous la forme monoanionique pour tout pH supérieur ou égal à 3.4). Les répulsions électrostatiques entre cet anion phosphodiester et le nucléophile qui approche qu'il soit neutre (H_2O) ou chargé négativement (ions RO-) fait que l'efficacité du clivage est très faible (Westheimer F.H., 1968).

Cette propriété confère à la liaison phosphodiester une stabilité importante pour maintenir l'intégrité et la fidélité du transfert de l'information génétique mais soulève un problème catalytique important pour les enzymes impliquées dans la régulation de la quantité d'ARNm de la cellule. Beaucoup de ribonucléases utilisent des mécanismes très sophistiqués qui permettent d'accélérer d'un facteur pouvant atteindre 10^{+16} la vitesse de clivage de la liaison phosphodiester (Gerlt, J.A., 1993). Cette accélération extraordinaire n'est pas totalement comprise mais est attribuée en partie à la capacité de l'enzyme à réduire les répulsions électrostatiques entre les oxygènes du phosphate et les nucléophiles, à favoriser les réactions intramoléculaires dans lesquelles le nucléophile est parfois le groupement hydroxyle en 2' du cycle du ribose et à stabiliser certains intermédiaires réactionnels.

2.2 Clivage de la liaison O3'-P (liaison covalente A)

Le clivage de la liaison O3'-P permet de libérer des ARN possédant une extrémité 3' OH. Le Tableau 2 présente une liste non exhaustive de ribonucléases d'*E. coli* coupant la liaison A. La caractéristique commune des ribonucléases de ce groupe est la présence d'un ion métallique divalent dans le site actif de l'enzyme (très souvent l'ion Mg^{2+}). Cet ion est directement impliqué dans la catalyse enzymatique selon un mécanisme conservé. La Figure 23 montre le mécanisme catalytique proposé pour la RNase HI d'*E. coli* (Uchiyama *et al.*, 1994).

Nom	Substrat	Cofacteur	Site de coupure	Fonction
RNase P	pré-ARNt	Mg^{2+}	Endoribonucléase	Maturation ARNt
RNase III	ARNdb	Mg^{2+}	Endoribonucléase	Maturation ARNm, ARNr
RNase E	ARNm, ARNr 9S	Mg^{2+}	Endoribonucléase	Maturation ARNr 5S, dégradation ARNm
RNase HI	Hybride ADN/ARN	Mg^{2+}	Endoribonucléase	Réplication, réparation ADN
RNase HII	Hybride ADN/ARN	Mg^{2+}, Mn^{2+}	Endoribonucléase	inconnue
RNase N	ARNsb et ARNdb	–	Endoribonucléase	inconnue
RNase II	ARNsb, poly(A)	Mg^{2+}	Exoribonucléase	Dégradation ARNm, maturation ARNt
RNase D	pré-ARNt	Mg^{2+}	Exoribonucléase	Maturation ARNt
RNase BN	pré-ARNt	Co^{2+}	Exoribonucléase	Maturation ARNt
RNase T	pré-ARNt	Mg^{2+}	Exoribonucléase	Maturation ARNt
OligoRNase	oligoribonucléotides	Mg^{2+}	Exoribonucléase	Dégradation ARN

Tableau 2. Liste non exhaustive des ribonucléases libérant une extrémité 3'OH. ARNsb : ARN simple brin, ARNdb : ARN double-brins, ARNr : ARN ribosomique, ARNt : ARN de transfert. D'après D'Alessio *et al.*, 1997.

Dans le cas de la RNase HI, une molécule d'eau activée par le groupe carboxyle monoionisé de l'aspartate 70 forme le nucléophile qui attaque l'atome de phosphore. Un ion Mg^{2+} appartenant au site actif fait une liaison de coordination avec la liaison scissile P-O par l'intermédiaire d'une deuxième molécule d'eau. Cette dernière sert de donneur de proton à la liaison P-O. L'attaque nucléophile du phosphore par le groupe hydroxyle passe par un état transitoire d'intermédiaire phosphoré pentacoordiné. Le clivage de la liaison P-O est accompagné d'une inversion de la configuration de l'atome de phosphore.

Il est important de noter que l'extrémité 3'OH ainsi libérée est fonctionnellement primordiale. Elle sert d'amorce pour l'allongement d'une nouvelle chaîne nucléotidique lors de la réplication ou de la réparation de l'ADN.

La mutagenèse dirigée la RNase HI a révélé l'existence de 4 chaînes latérales de type carboxyles (Asp 10, Asp 134, Asp 10 et Glu 48) et une histidine (H124) importantes pour la catalyse. On pense que l'histidine H124 joue le rôle de base qui régénère l'état monoanionique de l'aspartate 70 (Oda *et al.*, 1993).

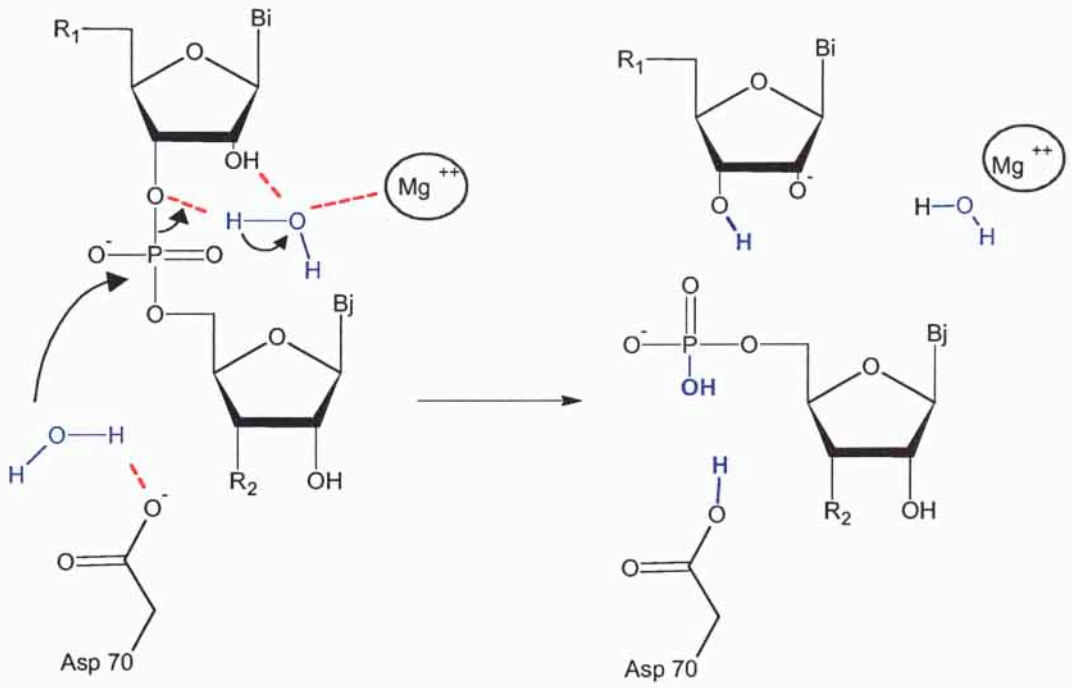


Figure 23 Mécanisme catalytique de clivage de la liaison O3'-P par la RNase HI d'*E. coli*.

Au vu du mécanisme présenté ci-dessus, la RNase HI ainsi que les enzymes du Tableau 2 constituent des hydrolases au sens strict car elles clivent directement les liaisons phosphodiesters en utilisant des molécules d'eau. Le terme de phosphodiestérase employé dans la littérature pour désigner ce groupe de ribonucléases est imprécis. Nous lui préférerons le terme « d'hydrolases libérant une extrémité 3'OH ».

Signalons enfin que la liaison O3'-P peut être clivée selon un mécanisme similaire à celui de la Figure 23 mais dans lequel l'agent nucléophile (la molécule d'eau) est remplacé par une molécule de phosphate inorganique qui attaque la liaison phosphodiester. On parle dans ce cas de « phosphorolyse » et non d'hydrolyse. La polynucléotide phosphorylase PNPase (Jarrige *et al.*, 2002) et la RNase PH (Ishii *et al.*, 2003) sont deux exemples d'enzymes qualifiées de « phosphorylases libérant une extrémité 3'OH ».

2.3 Clivage de la liaison P-O5' (liaison covalente B)

Les ribonucléases qui clivent la liaison P-O5' libèrent une extrémité 3' phosphate. Deux mécanismes réactionnels distincts ont été proposés pour expliquer cette coupure. Le premier est un mécanisme d'hydrolyse, le deuxième est un mécanisme de transfert de liaison phosphate.

2.3.1 Hydrolyse de la liaison P-O5' : Exemple de la nucléase staphylococcale

La nucléase staphylococcale est une enzyme qui clive aussi bien l'ADN que l'ARN en libérant une extrémité 3' phosphate (Loll *et al.*, 1989). Il s'agit d'une protéine de 149 résidus aminés qui utilise un ion calcium Ca^{2+} dans son site actif. Cet ion est maintenu grâce à des liaisons de coordination avec les chaînes latérales des résidus aspartate 21 et aspartate 40, le carbonyle de la thréonine 41, 3 molécules d'eau et l'oxygène non estérifié du groupement phosphate de la molécule d'ARN substrat. Le tout forme autour de l'ion calcium une structure de coordination octaédrique (Hale *et al.*, 1993). Le mécanisme de catalyse est présenté dans la Figure 24.

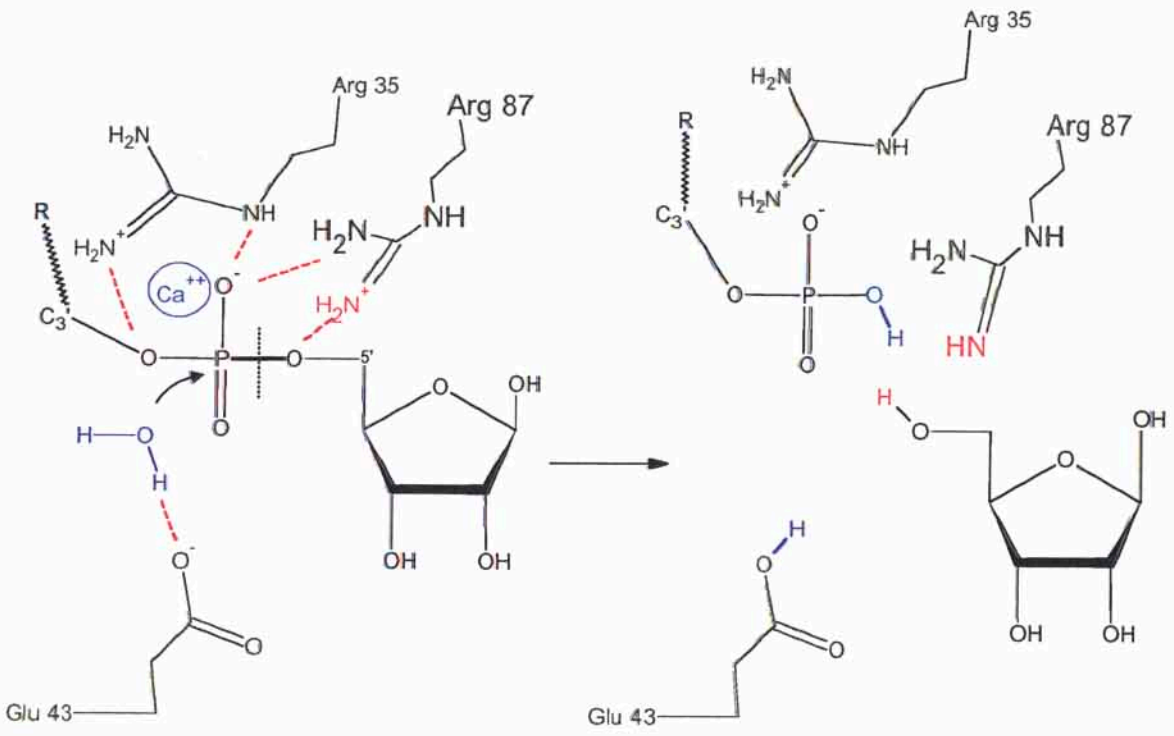


Figure 24 : Mécanisme de clivage de la liaison P-O5' par la Nucléase Staphylococcale. Le structure octaédrique complexe entourant l'ion calcium n'est pas représentée. La liaison scissile P-O5' est dessinée en gras.

Dans ce mécanisme le glutamate 43 active une molécule d'eau par soustraction d'un proton. L'attaque du nucléophile ainsi produit (groupe hydroxyle) sur l'atome de phosphore est facilitée d'une part par l'ion Ca^{2+} positionné dans le site actif et d'autre part par les chaînes latérales des arginines 35 et 87 qui concourent à la neutralisation de la charge négative de la liaison phosphodiester. Comme dans le cas de RNase HI, un intermédiaire pentacoordiné serait produit transitoirement avant que la liaison P-O5' ne soit rompue. La nucléase staphylococcale est, selon notre définition du paragraphe 2.2, une « hydrolase libérant une extrémité 3' phosphate ».

D'autres ribonucléases sont capables de cliver la même liaison P-O5' mais en utilisant un mécanisme de coupure radicalement différent.

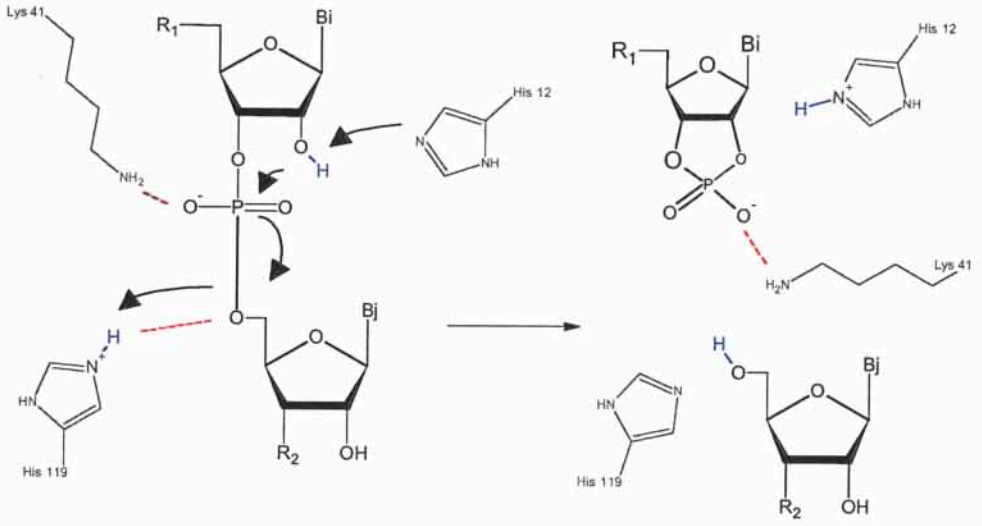
2.3.2 Clivage de la liaison P-O5' par transfert du groupe phosphate : Exemple de la RNase A

Contrairement aux enzymes décrites jusqu'à présent, la ribonucléase A extraite du pancréas bovin (Beintema J.J., 1998 ; Beintema *et al.*, 1997 ; Thompson *et al.*, 1994) clive la liaison covalente P-O5' en deux étapes.

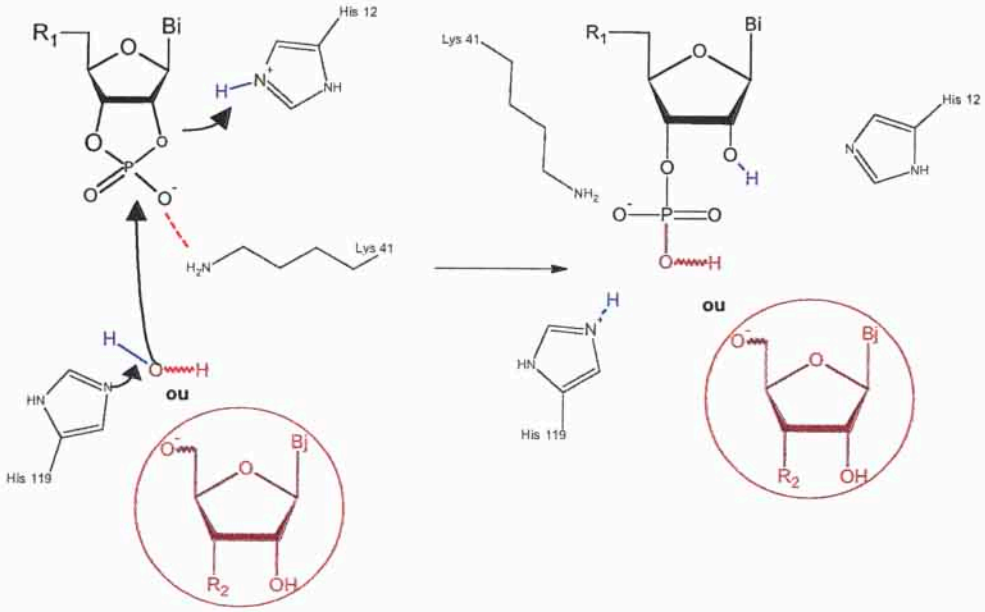
La première étape est une réaction de transphosphorylation dans laquelle la liaison phosphate P-O5' est « échangée » contre une liaison phosphate P-O2' mettant en jeu le groupe hydroxyle en position 2' sur le cycle du ribose. Le produit de cette première étape est un composé qui possède une liaison phosphodiester cyclique atypique 3'O-P-2'O.

Dans une deuxième étape, le composé cyclique est hydrolysé en présence de molécules d'eau pour libérer un ARN possédant une extrémité 3' phosphate.

Le mécanisme de la RNase A est représenté dans la Figure 25. Ce schéma mécanistique a été pendant longtemps l'objet de confusions et de controverses liées à la nature de la molécule cyclique qui apparaît à l'issue de l'étape de transphosphorylation. Le 2',3'-phosphodiester cyclique était considéré au départ comme « un intermédiaire réactionnel » qui restait lié à l'enzyme avant de subir l'étape d'hydrolyse finale. Cuchillo et ses collaborateurs ont montré que le composé cyclique n'était point un intermédiaire réactionnel mais constituait un premier produit libéré dans la solution (Cuchillo *et al.*, 1993). L'hydrolyse de ce premier produit a lieu uniquement en milieu aqueux, une fois que toutes les molécules de substrats ont été transphosphorylées. Cuchillo a interprété cette observation en considérant que l'enzyme est capable de catalyser la réaction inverse en utilisant des molécules d'eau à la place du poly-nucléotide libéré lors de l'étape de transphosphorylation (voir Figure 25).



Réaction n° 1 : Formation du produit cyclique



Réaction n° 2 : Réaction inverse de la Réaction n °1

Figure 25 : Mécanisme catalytique de la RNase A. L'enzyme clive la liaison P-O5' en produisant un phosphodiester 2',3'-cyclique (Réaction 1). En solution aqueuse (forte concentration en molécules d'eau) le produit cyclique est hydrolysé selon la réaction 2. En solution non aqueuse, la réaction 2 reforme le phosphodiester linéaire de départ. Les deux réactions 1 et 2 sont inverses l'une de l'autre et témoignent de la capacité de l'enzyme à catalyser une réaction donnée dans les deux sens.

Dans ce mécanisme l'histidine 12 joue le rôle de base durant l'étape de transphosphorylation et d'acide durant l'étape d'hydrolyse. L'histidine 119 joue un rôle symétrique. La lysine 41 joue un rôle important dans la catalyse en neutralisant la charge négative qui se trouve au voisinage de la liaison phosphodiester qu'elle soit linéaire ou cyclique.

Le Tableau 3 résume les 4 mécanismes connus permettant aux ribonucléases de rompre la liaison phosphodiester.

Mécanisme	Extrémité produite	Nucléophile	« Intermédiaire » réactionnel	Exemple
Hydrolyse	3'OH	H_2O	Phosphorane penta-coordiné[1]	RNase HI
Phosphorolyse	3'OH	Phosphate	Probablement un phosphorane[1]	RNase PH
Hydrolyse	3' Phosphate	H_2O	Phosphorane penta-coordiné[1]	Nuc. Staphy.
Transphosphorylation	3' Phosphate	2'OH	Phosphodiester 2',3'-cyclique[2]	RNase A

Tableau 3 : Les 4 mécanismes possibles permettant de cliver la liaison phosphodiester dans un ARN. [1] Il s'agit d'un intermédiaire réactionnel vrai qui ne peut être libéré en solution. [2] Il s'agit d'un produit de réaction libéré dans la solution. Nuc. Staphy. : Nucléase Staphylococcale.

Afin d'identifier le mécanisme que catalyse RegB il est d'abord nécessaire de connaître la nature de l'extrémité (3'OH ou 3' Phosphate) libérée lors de la coupure du motif GGAG.

2.4 RegB libère une extrémité 3' Phosphate

L'identification de la nature de l'extrémité libérée par RegB est due à Marc Uzan (Uzan *et al.*,1988). La stratégie utilisée repose sur l'utilisation de la T4 polynucléotide kinase. Cette enzyme catalyse le transfert et l'échange du groupe γ-phosphate de l'ATP vers l'extrémité 5'OH (s'il y a lieu) d'un polynucléotide. En présence d'une concentration d'ATP de 1 $\mu mol.l^{-1}$, en absence d'ADP, en se plaçant à un pH de 7.6 et en utilisant un excès d'ARN substrat, l'enzyme fonctionne essentiellement comme une kinase en phosphorylant une éventuelle extrémité 5'OH libre. Dans ces conditions et en utilisant de l'ATP marqué au ^{32}P en position Gamma, un fort marquage radioactif du produit de digestion par RegB de l'ARN substrat MotA a été observé (Sanson *et al.*, 1995). Ce résultat montre

que l'un des fragments d'ARN produits de digestion possède une extrémité 5'OH libre, c'est à dire que l'autre fragment possède une extrémité 3' phosphate.

Le fait que RegB libère une extrémité 3' phosphate permet d'envisager, d'après le Tableau 3, deux mécanismes différents de coupure : une hydrolyse sans intermédiaire détectable expérimentalement ou une transphosphorylation produisant un phosphodiester cyclique. Afin de discriminer entre ces deux mécanismes, nous avons choisi d'analyser la nature des produits de la réaction de coupure par résonance magnétique du noyau phosphore (^{31}P).

2.5 Suivi cinétique par RMN du ^{31}P de la réaction catalysée par RegB

2.5.1 Principe

Les noyaux atomiques possèdent une propriété, dérivant de la théorie de la mécanique quantique, appelée spin. A cette propriété est associé un nombre entier ou demi-entier appelé nombre de spin caractéristique du noyau étudié.

Le noyau ^{31}P (abondance naturelle : 100 %) possède un nombre de spin égal à 1/2. Ce nombre lui confère deux états énergétiques α et β naturellement confondus mais qui deviennent distincts lorsque le noyau est plongé dans un champ magnétique externe uniforme. Moyennant l'apport, sous la forme d'une onde électromagnétique, d'un quanta d'énergie correspondant à la différence entre les états α et β, un signal de résonance du noyau ^{31}P est observé, mesuré par un spectromètre RMN et quantifié sous la forme d'une grandeur appelé déplacement chimique δ exprimé en partie par million (ppm). Le déplacement chimique δ du ^{31}P est très sensible à l'environnement atomique dans lequel il se trouve et constitue donc une excellente sonde pour observer les changements qui ont lieu à l'occasion de la rupture de la liaison phosphodiester.

Gorenstein a montré que le déplacement chimique du ^{31}P est extrêmement sensible à l'angle O-P-O (Gorenstein D.G., 1984). Une diminution de 3 degrés seulement dans la valeur de l'angle O-P-O s'accompagne d'une augmentation de la valeur du déplacement chimique de 4 unités.

La Figure 26 montre la variation de la valeur du δ (^{31}P) en fonction de l'angle O-P-O. Il est intéressant de voir sur cette figure qu'un phosphodiester linéaire résonne dans l'intervalle [0, 5 ppm] alors qu'un phosphodiester cyclique résonne vers 20 ppm.

Nous avons donc ici un moyen de distinguer de façon non ambiguë un phosphodiester linéaire d'un phosphodiester cyclique. Autrement dit, en observant l'apparition éventuelle d'un signal dans la région des 20 ppm, il est possible de savoir si la réaction de coupure catalysée par RegB produit un intermédiaire de type phosphodiester 2',3'-cyclique ou non.

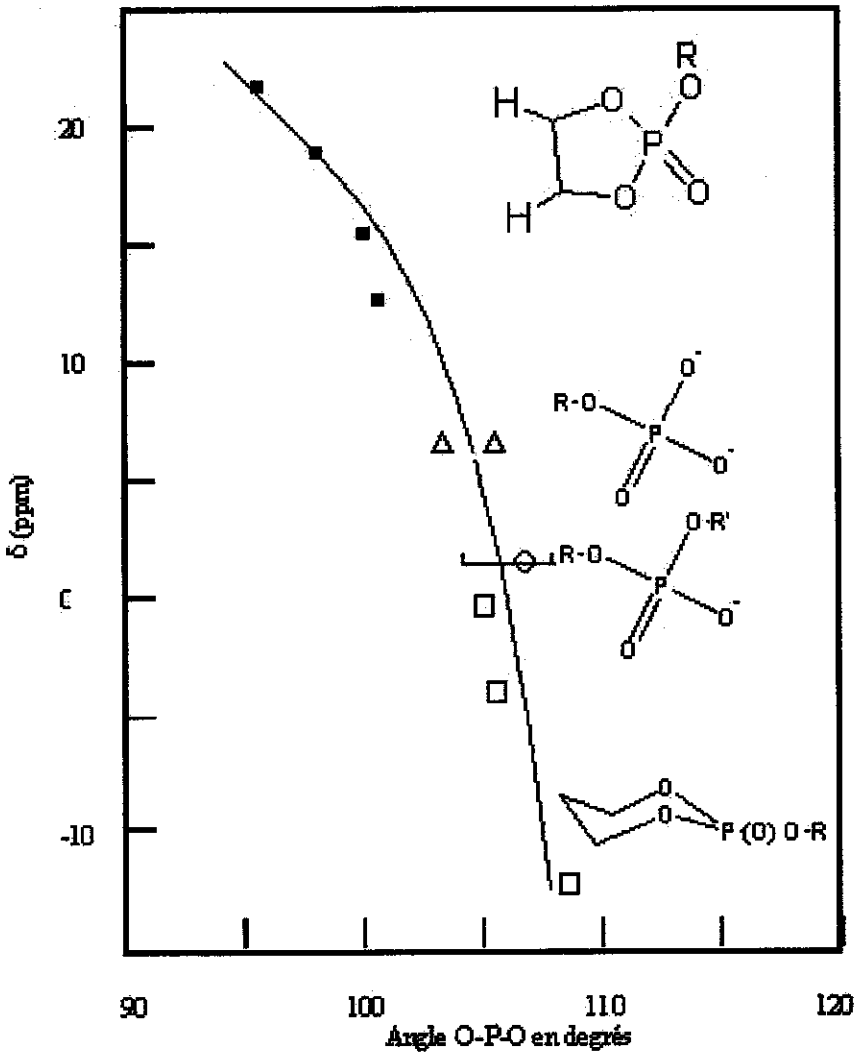


Figure 26 : Déplacement chimique du ^{31}P dans les esters phosphates en fonction de la valeur de l'angle O-P-O. Carré plein : esters pentacycliques, triangle vide : monoesters dianioniques, cercle vide : monoesters monoanioniques, carré vide : ester hexacycliques. D'après Gorenstein D.G., 1984.

2.5.2 Choix d'un modèle expérimental de coupure par RegB

A cause de la faible sensibilité de la RMN, le suivi au cours du temps de la variation du déplacement chimique ^{31}P d'un ARN substrat de RegB n'est possible que si la concentration de la solution en liaisons phosphodiesters scissiles est voisine de 0.5-1.0 mmol.l^{-1}. Dans un tube RMN d'environ 400 µl de volume, cette concentration correspond à 0.2-0.4 micromoles d'ARN ne contenant qu'un seul motif clivable GGAG. Il s'agit d'une énorme quantité d'ARN qui demande le recours à la synthèse chimique (la transcription *in vitro* est difficilement adaptable à la synthèse d'ARN à l'échelle de la micromole). La synthèse chimique est, par ailleurs, incapable de produire efficacement des ARN dont la longueur dépasse la trentaine de nucléotides. Sachant que toutes les cibles ARN naturelles de RegB font plus d'une quarantaine de nucléotides, il nous faut disposer d'un ARN court et efficacement clivé par RegB.

Pour atteindre ce but, nous avons exploité d'une part les travaux de Jayasena et de ses collaborateurs (Jayasena *et al.*, 19996) et d'autre part les travaux de Lebars (Lebars *et al.*, 2001). Jayasena a utilisé la technique SELEX (Systematic Evolution of Ligands by Exponential enrichment) pour sélectionner *in vitro* des ARN clivés par RegB en présence de son activateur : la protéine ribosomique S1. 42 molécules ont été obtenues. Elles contenaient toutes le motif GGAG. La 22éme molécule sélectionnée est présentée dans la Figure 27. Cette molécule qui comprenait au départ 80 nucléotides a été raccourcie en une molécule de 30 nucléotides en éliminant les séquences fixes en 5' et en 3' et ne conservant ainsi que la région centrale qui a subi le processus de sélection à partir d'une bibliothèque aléatoire. Le 30-mer obtenu a été appelé Selex22, il contient deux motifs GGAG dont un seul est clivé par RegB. Lors de ses tentatives d'élucider le mécanisme d'activation de RegB par la protéine S1, Lebars a dessiné un variant de la molécule Selex22, le Selex22tb, qui a été amputé du site GGAG non clivé par RegB.

La molécule Selex22tb nous a paru comme étant la molécule idéale pour étudier sa coupure par RegB. Cette molécule est en effet très efficacement clivée par RegB en absence de la protéine S1 ce qui permet d'étudier la coupure par RegB sans les interférences éventuelles de la protéine ribosomique (Lebars *et al.*, 2001). Elle comprend par ailleurs uniquement 21 nucléotides et peut donc être obtenue très facilement par voie de synthèse chimique.

Nous avons produit chimiquement l'ARN Selex22tb en utilisant des phosphoramides selon un protocole inspiré de celui de la synthèse chimique d'ADN et adapté à la synthèse d'ARN par Snoussi et ses collaborateurs (Snoussi *et al.*, 2001).

Nous avons vérifié que la molécule Selex22tb est efficacement coupée par RegB (voir Figure 28 montrant l'analyse sur gel de polyacrylamide d'une cinétique de coupure sur 60 minutes). Le mélange réactionnel introduit dans le tube RMN contenait au final 30 $\mu mol.l^{-1}$ de RegB et un excès de l'ARN substrat Selex22tb de 300 $\mu mol.l^{-1}$.

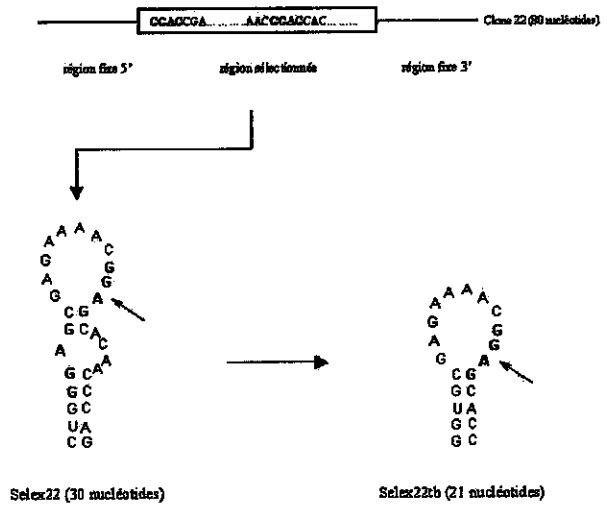


Figure 27 : Conception de la molécule d'ARN Selex22tb. Les deux sites GGAG du Selex22 sont montrés en gras, l'unique site coupé est indiqué par une flèche.

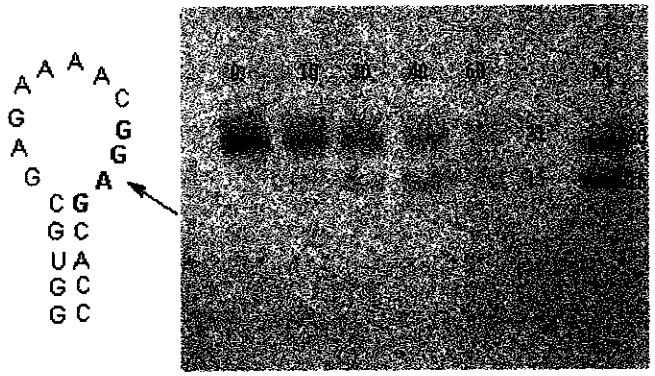


Figure 28 : Coupure de l'ARN Selex22tb par RegB. Suivi sur gel de polyacrylamide de la cinétique de coupure de l'ARN Selex22tb par RegB au niveau du motif GGAG. L'ARN coloré au bromure d'éthidium (BET) est révélé par des ultra-violets. La bande à 21 nucléotides correspond au substrat Selex22tb, la bande à 15 correspond au produit de coupure le plus long. Le 6-mer produit n'est pas observé sur le gel car il ne fixe pas suffisamment de BET pour être détecté sous lumière UV.

2.5.3 Identification du produit de coupure

Nous avons suivi la cinétique de coupure de l'ARN Selex22tb par RegB à 25°C. RegB est en effet stable à cette température pendant plusieurs jours alors qu'elle est rapidement inactivée à 37 °C.

A l'aide d'une sonde large bande, nous avons enregistrés 175 spectres ^{31}P unidimensionnels successifs. Chacun de ces spectres donne, à un instant donné, les signaux de résonance des noyaux phosphores qui se trouvent dans le mélange réactionnel.

Nous avons enregistré pour chaque spectre unidimensionnel 1024 accumulations afin d'augmenter le rapport signal sur bruit. Nous avons en effet anticipé le fait que si un phosphodiester cyclique venait à apparaître, son signal serait très faible. On a estimé que le signal d'un phosphodiester cyclique éventuel serait 40 fois moins intense que le signal du phosphodiester linéaire du Selex22tb. En effet, parmi les 20 noyaux phosphore du Selex22tb un seul est susceptible de former un ester cyclique. Par ailleurs on sait que la coupure de ce même ARN par RegB n'est pas quantitative et atteint un taux de coupure ne dépassant pas les 50-60 % (Lebars *et al.*, 2001).

Nous avons par ailleurs réalisé une réaction témoin en utilisant le mutant inactif RegB R52L (Ruckman *et al.*, 1989). Ce mutant a été exprimé et purifié selon un protocole identique à celui utilisé pour la ribonucléase sauvage. Le spectre ^{31}P de ce mutant a été enregistré dans les mêmes conditions expérimentales et avec les mêmes paramètres RMN.

La Figure 29 montre le résultat obtenu. Sur ce spectre, l'axe des abscisses correspond aux déplacements chimiques ^{31}P. L'axe des ordonnés correspond au temps de la réaction de coupure (le temps augmente de bas en haut). L'aspect à deux dimensions de ce spectre est du au stockage de signaux de précessions des noyaux phosphates dans des blocs correspondant normalement à des expériences RMN à deux dimensions qui ont subit par la suite une transformée de Fourier dans la dimension phosphore. La partie supérieure du spectre représente la somme des 10 derniers spectres unidimensionnels (Spectres 165 à 175) ce qui correspond à l'état de la solution à la fin de la réaction de clivage.

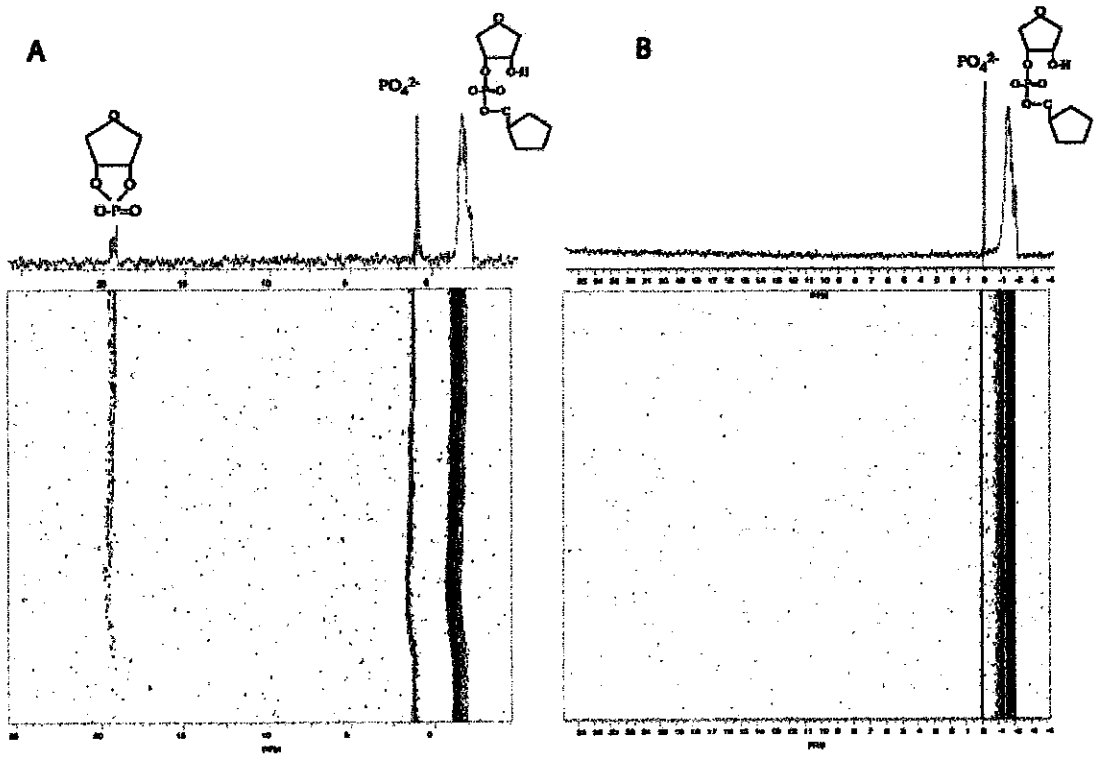


Figure 29 : Suivi par RMN du ^{31}P de la cinétique de coupure de l'ARN Selex22tb par la ribonucléase RegB. (A) Clivage par la protéine RegB sauvage. (B) Expérience témoin utilisant le mutant ponctuel inactif RegB R52L. En abscisses : déplacements chimiques ^{31}P. En ordonnées : le temps.

Le massif large centré à -1.5 ppm correspond à la région où résonnent les phosphores impliqués dans des liaisons diesters linéaires. Le pic fin à 1.5 ppm correspond aux molécules de phosphates résiduelles qui proviennent des tampons de purification de RegB et qui n'ont pas été complètement éliminées par dialyse (on peut noter que la position de ce pic n'est pas exactement la même dans les deux spectres, un léger écart de pH entre les deux expériences explique ce petit décalage de déplacements chimiques).

A 19 ppm, nous identifions clairement un pic qui résonne dans la région des phosphodiesters 2',3'-cycliques. Ce pic est très probablement le produit de coupure de l'ARN Selex22tb par RegB pour deux raisons. D'abord, il est absent au début de la réaction (bas du spectre) et s'accumule progressivement le long de la réaction. Par ailleurs, il est absent dans le spectre témoin enregistré en présence du mutant inactif R52L indiquant que l'apparition de ce composé est liée à l'activité nucléolitique de RegB et non pas à la dégradation de l'ARN Selex22tb par des RNases contaminantes du groupe de la RNase A.

Il est intéressant de remarquer que contrairement à nos attentes, le pic du phosphodiester cyclique ne correspond pas à un pic fin et unique mais à un petit massif où plusieurs pics se chevauchent. Bien qu'il existe dans le Selex22tb un seul noyau phosphate capable de produire le diester cyclique, nous pensons que ce dernier pourrait adopter en solution plusieurs conformations qui donnent au phosphore penta-coordiné des déplacements chimiques disperses autour de la zone de 19 ppm.

Signalons enfin que nous n'observons pas sur le spectre (A) de la Figure 29 de pics RMN correspondant au produit « final » de la coupure du Selex22tb par RegB c'est à dire à la molécule issue de l'hydrolyse du phosphate pentacoordiné (voir mécanisme de la RNase A du paragraphe 2.3.2). Nous pensons que l'observation d'un tel signal nécessiterait le suivi du milieu réactionnel pendant des dizaines d'heures supplémentaires. En effet, comme on l'a expliqué au paragraphe 2.3.2 la réaction de formation du phosphodiester cyclique et la réaction d'hydrolyse obéissent au même mécanisme et se font de manière successive : l'enzyme commence à hyrolyser le phosphore cyclique uniquement quand tout le substrat linéaire de départ a été utilisé.

2.5.4 Conclusion et conséquences

La détection du diester 2',3'-cyclique permet d'affirmer que la ribonucléase **RegB est une transphosphorylase**. RegB utilise donc un mécanisme réactionnel dans lequel le nucléophile est le groupe hydroxyle en position 2' sur le cycle du ribose.

Cette conclusion a deux conséquences. D'abord, elle suggère que RegB n'a pas besoin d'ions métalliques pour catalyser le clivage de son ARN substrat puisque d'après le Tableau 3 seules les ribonucléases de type hydrolases exigent la présence de tels ions. Par ailleurs, elle limite le champ d'investigation du site actif de RegB au groupe des ribonucléases utilisant la stratégie de coupure de la RNase A. Nous verrons dans le chapitre 4 que l'étude des résidus catalytiques de ce groupe de ribonucléases, très documentés dans la littérature scientifique de point de vue structural et fonctionnel, nous permettra de localiser certains des résidus du site actif de RegB.

Avant d'entreprendre ce travail, nous avons voulu auparavant investiguer la structure primaire de RegB par mutagenèse aléatoire à la recherche d'une éventuelle « région hypersensible» sur laquelle nous focaliserons notre étude. C'est l'objectif du chapitre qui suit.

Chapitre 3

Investigation par mutagenèse aléatoire de la structure primaire de RegB

Chapitre 3

Investigation par mutagenèse aléatoire de la structure primaire de RegB

En absence de toute information structurale ou fonctionnelle relative à la structure primaire, la mutagenèse aléatoire constitue souvent une méthode rapide qui permet de comparer l'importance des différents segments peptidiques d'une protéine donnée, voire de détecter les résidus aminés indispensables à l'activité.

Deux éléments concourent à la mise au point d'une stratégie de mutagenèse aléatoire efficace : une technique d'altération de la séquence du gène adaptée au problème étudié et un système de sélection performant permettant d'isoler facilement les clones mutés avant d'entreprendre tout travail de séquençage.

Les techniques d'altération aléatoire de la séquence du gène étudié sont multiples. Elles se distinguent par le nombre et la nature des mutations introduites et peuvent être classées en techniques *in vivo* et techniques *in vitro*.

Les techniques *in vivo* consistent à provoquer une modification de l'information génétique dans la cellule en utilisant, entre autres, des souches bactériennes dites « mutatrices ». Ces souches

ont perdu une ou plusieurs activités assurant la fidélité de la transmission de l'information génétique lors de la réplication et la division cellulaire. Les souches *mutD* (Schaaper *et al.*, 1989), par exemple, sont altérées dans la sous-unité ε de l'ADN polymérase III qui supprime l'activité de correction des faux appariements. Les souches *mutY* sont déficientes en une activité glycosylase qui intervient normalement dans la correction des faux appariements de type G:A (Nghiem *et al.*, 1988 ; Au *et al.*, 1989). Le plasmide contenant le gène d'intérêt est introduit dans la souche bactérienne mutatrice, propagé pendant plusieurs générations puis introduit dans une souche non-mutatrice afin d'y examiner la nature et l'effet des mutations obtenues.

Les techniques de mutagenèse aléatoire *in vitro* sont multiples. L'utilisation d'agents chimiques comme l'hydroxylamine (Busby *et al.*, 1982), l'hydrazine (Myers *et al.*, 1985) le N-nitrosoguanidine (Mitchel *et al.*, 1986), les agents alkylants ainsi que l'irradiation aux ultra-violets et la courte exposition à une source ionisante sont des techniques efficaces mais qui souffrent de deux limitations. Il est d'abord très difficile de contrôler le taux de mutation (nombre de mutation pour une longueur donnée en paires de bases). Par ailleurs, ces méthodes ne génèrent pas toutes les substitutions possibles. Dans le cas de l'hydroxylamine, par exemple, les mutations obtenues sont essentiellement des transitions de paires C:G en paires T:A.

Pour muter aléatoirement RegB, nous avons adopté une technique *in vitro* basée sur l'utilisation de la réaction de polymérisation en chaîne (PCR). Les raisons qui ont motivé ce choix sont exposées dans le paragraphe qui suit.

3.1 La mutagenèse aléatoire par PCR

L'ADN polymérase thermostable Taq de *Thermus aquaticus* catalyse *in vitro* la polymérisation récurrente d'un segment d'ADN en présence d'un mélange des 4 désoxyribonucléotides, de deux amorces d'ADN et d'ions métalliques Mg^{2+} (Tindall *et al.*, 1988). Cette enzyme possède une activité polymérase 5'-3' mais est dépourvue d'activité exonucléase 3'-5' (activité de relecture). La Taq polymérase est donc incapable de corriger les bases incorporées par erreur dans la chaîne nucléotidique naissante. Le taux d'erreur de cette polymérase est relativement élevé. Il est estimé à une base mal appariée sur 3500 alors qu'il est n'est que de 1 base sur 50 000 pour la polymérase Deep Vent de *Pyrococcus GB-D* (*New England Biolabs*). De ce fait la Taq polymérase peut être utilisée pour introduire délibérément des mutations dans la séquence d'ADN au cours du processus d'amplification par PCR.

Le gène *regB* comprend 462 paires de bases. Introduire au minimum une mutation aléatoire sur la séquence d'un gène de cette longueur nécessite d'augmenter le taux d'erreur final de la Taq polymérase d'un facteur 7.6 environ (3500/462). Pour augmenter ce taux d'erreur il est possible d'utiliser deux méthodes mutuellement non-exclusives. La première consiste à utiliser un mélange non-équimolaire des 4 désoxyribonucléosides. Cette méthode se justifie par le fait que la probabilité que la Taq insère la base N à la place de la base M est directement proportionnelle au rapport des concentrations [dNTP]/[dMTP]. En présence d'un excès du désoxyribonucléotide dNTP par rapport au dMTP, la probabilité d'avoir la mutation X :N sera multipliée par un facteur supérieur à 1, égal à [dNTP]/[dMTP]. La deuxième méthode consiste à ajouter un mélange des ions métalliques magnésium Mg^{2+} et manganèse Mn^{2+}. Les ADN polymérases requièrent en général la présence d'un cation divalent pour assurer leur fonction catalytique. Ce cation est très souvent l'ion magnésium Mg^{2+}. Bien que le cation manganèse Mn^{2+} puisse se substituer au magnésium comme cofacteur de la réaction de polymérisation, il a été constaté que le taux d'erreur est significativement augmenté quand le manganèse est utilisé à des concentrations millimolaires (Seal *et al.*, 1979). Des données RMN ont suggéré qu'en présence de manganèse la liaison des désoxyribonucléosides triphosphates au site actif aurait lieu dans des conformations particulières du désoxyribose normalement défavorisées en présence de magnésium (Sloan *et al.*, 1975).

3.2 Mutagenèse aléatoire de RegB et stratégie de sélection

Nous avons choisi de muter indépendamment les parties 5' et 3' du gène *regB*. Deux raisons justifient ce choix. D'abord, les protocoles publiés dans la littérature pour la Taq polymérase en conditions de PCR mutagène font état en moyenne de 1 à 2 modifications pour un fragment de 300 paires de bases ce qui correspond environ à la longueur des parties 5' et 3' du gène *regB* définies par rapport à un site EcoRI central. Par ailleurs, puisqu'il n'existe pas de biais particulier dans la séquence nucléotidique des parties 5' et 3', il est possible de comparer l'importance relative des deux parties du gène *regB* en réalisant en parallèle le même protocole de mutagenèse.

Pour sélectionner les mutants de *regB*, nous avons utilisé un crible biologique simple tirant profit des résultats du chapitre 1. Ce crible utilise la toxicité de RegB comme rapporteur de son activité. Toute mutation qui annule « suffisamment » l'activité de la protéine RegB peut se maintenir dans un vecteur d'expression de type pET standard et donne des clones viables quand on transforme la souche *E. coli* XL1-Blue. Le terme « suffisamment » est imprécis et indique que notre crible biologique ne sélectionnera que les clones dont l'activité est diminuée d'un facteur strictement

supérieur à 20 (voir Figure 13, chapitre 1). Les clones mutés ayant conservé, par exemple, 10 % de leur activité ne seront pas sélectionnés. Notre crible biologique est donc « un filtre aux mailles trop grandes » qui ne sélectionnera que les positions les plus importantes pour la structure et/ou la fonction de RegB. La Figure 30 résume la stratégie de mutagenèse et de sélection utilisée.

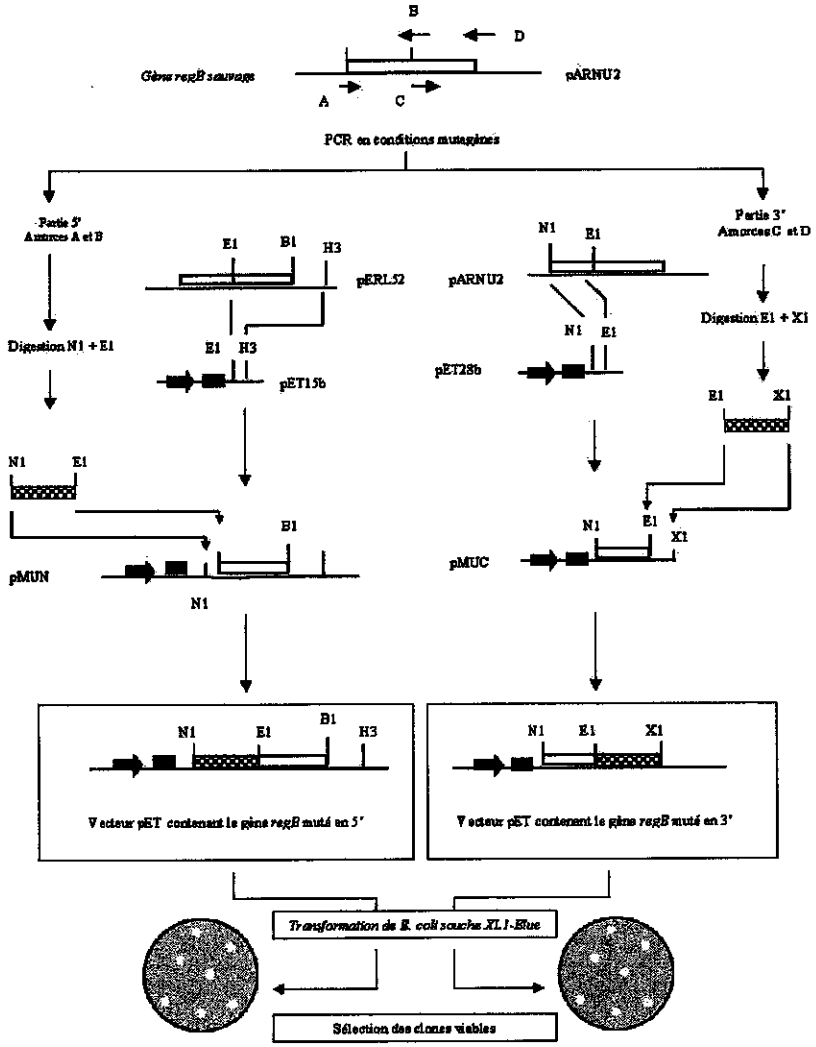


Figure 30 : Mutagenèse aléatoire des parties 5' et 3' du gène *regB* et méthode de sélection des mutants. La partie 5' du gène *regB* sauvage initialement dans le plasmide pARNU2 est amplifiée avec les amorces A et B, digérée par NdeI (N1) et EcoRI (E1) puis clonée dans le vecteur pMUN (provenant du vecteur pET15b standard dans lequel a été clonée la partie 3' du gène *regB*). La partie 3' gène *regB* sauvage est amplifiée avec les amorces C et D, digérée par XhoI (X1) et EcoRI (E1) puis clonée dans le vecteur pMUC (provenant du vecteur pET28b standard dans lequel a été clonée la partie 5' du gène *regB*). Dans les deux cas, la solution de ligature est utilisée pour transformer la bactérie. Les clones viables sont sélectionnés, amplifiés puis éventuellement séquencés. La flèche pleine noire représente le promoteur T7lac, le carré noir le RBS Phi 10 sauvage.

Nous avons testé, dans un premier temps, deux conditions différentes de mutagenèse par PCR sur la partie 5' du gène *regB* (conditions M1 et M2, voir Tableau 4). Dans la condition M1, on a conjugué l'effet d'un mélange magnésium et manganèse au rapport de concentrations Mg^{2+}/Mn^{2+} égal à 1/10 avec l'effet d'un mélange non-équimolaire des 4 désoxyribonucléosides. Des mutants types ont été séquencés et ont révélé que le taux d'erreur est supérieur à celui publié dans la littérature. Il résultait du protocole M1 4 à 6 substitutions sur un fragment ne comptant que 100 résidus. Nous avons donc diminué l'efficacité de mutagenèse en conservant le mélange Mg^{2+}/Mn^{2+} égal à 1/10 mais en utilisant un mélange équimolaire des 4 désoxyribonucléosides (condition M2). Le nombre de substitutions obtenues était de 2 à 3. Nous avons appliqué ce dernier protocole à la partie 5' et à la partie 3' du gène *regB*.

	ADN (pM)	Amorce 5' (µM)	Amorce 3' (pM)	Mg^{2+} (mM)	Mn^{2+} (mM)	Mg^{2+}/Mn^{2+}	Taq (Unités)	Cycles	dATP (mM)	dGTP (mM)	dCTP (mM)	dNTP (mM)
M1	300	0.4	0.4	3	0.3	10	2	22	0.2	0.2	1	1
M2	300	0.4	0.4	3	0.3	10	2	22	0.2	0.2	0.2	0.2

Tableau 4 : Conditions utilisées pour la PCR mutagène. L'ADN matrice est le plasmide pARNU2 contenant la copie sauvage du gène *regB*.

Lorsque des clones viables sont récupérés à l'issue de l'étape de transformation de la bactérie, 4 phénomènes indépendants peuvent expliquer la survie de ces clones : la récupération de clones parentaux ne contenant par exemple que la partie C-terminale de RegB (c'est à dire des clones n'ayant pas intégré la partie 5' mutée du gène *regB*), la récupération de clones réarrangés ayant perdu la capacité d'exprimer la protéine RegB, l'introduction d'un codon stop qui conduit à l'expression d'un petit fragment peptidique inactif et enfin l'introduction d'une mutation ayant un effet structural et/ou fonctionnel. Pour ne sélectionner que ces derniers clones, une deuxième étape de criblage est nécessaire. Cette étape est présentée dans la Figure 31. Les clones parentaux, ne contenant que l'extrémité C-terminale, sont d'abord éliminés en réalisant une PCR directement sur les colonies bactériennes à la recherche d'un fragment d'ADN correspondant au gène *regB* entier. Dans le but d'éliminer les mutants inactivés par apparition prématurée d'un codon stop ainsi que ceux incapables d'exprimer une protéine suite à un réarrangement génétique, nous avons réalisé des tests d'expression protéique en utilisant les clones contenant le gène *regB* reconstitué. Afin d'éviter une étape inutile de

re-transformation d'une souche *E. coli* de type BL21(λDE3), nous avons réalisé les tests d'expression directement dans la souche XL1-Blue en utilisant le système d'expression basé sur l'apport de l'activité T7 ARN polymérase par le bactériophage λCE6 (voir chapitre 1).

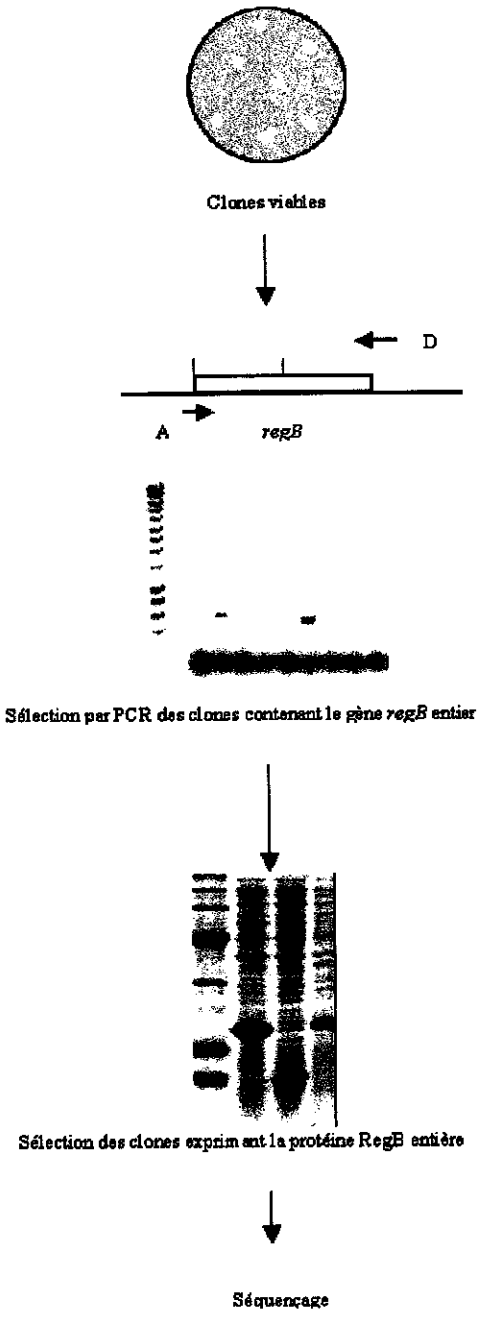


Figure 31 : Criblage des clones viables à la recherche des mutants de RegB affectés dans la structure et/ou la fonction.

3.3 Résultats de la mutagenèse de la partie 5' du gène *regB* (partie N-terminale de RegB)

En appliquant le protocole de mutagenèse M2 à la partie 5' du gène *regB* (correspondant aux résidus M1 à F80), beaucoup de clones viables ont été obtenus à l'issue de la fusion de la partie 3' du gène *regB* sauvage contenue dans le plasmide pMUN avec la partie 5' issue de la PCR mutagène. La Figure 32 montre en exemple le résultat des tests d'expression de 13 clones viables contenant le gène *regB* entier. Les clones M48, M55, M76, M98, M99, M101 et M104 expriment des fragments protéiques de tailles variables inférieures à 14 kDa résultats de l'apparition d'un codon stop dans le fragment N-terminal de RegB. En revanche, les clones M18, M92, M93, M97, M102 et M42 expriment une protéine d'environ 20 kDa qui correspond probablement à la fusion de la protéine RegB à une queue poly-histidines. Ces clones devraient donc contenir des mutations qui ont inactivé la protéine RegB.

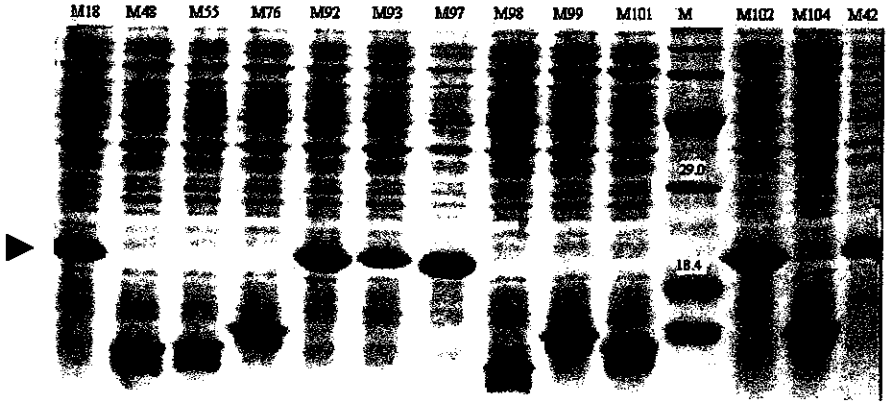


Figure 32 : Analyse SDS-PAGE de l'expression dans la souche XL1-Blue de 13 clones viables issus de la mutagenèse aléatoire de la partie N-terminale de la ribonucléase RegB. Les cellules sont induites par surinfection avec le phage λCE6. M : marqueur de masses moléculaires en kDa. Le triangle noir indique la zone de migration de la protéine RegB entière.

L'ADN plasmidique des clones exprimant une protéine de 20 kDa est extrait puis séquencé. Nous avons analysé au total 156 clones viables. Seuls 12 clones ont franchi la deuxième étape de sélection et ont été séquencés. La Figure 33 montre l'alignement des séquences de ces mutants avec la séquence de la partie N-terminale de la protéine RegB sauvage.

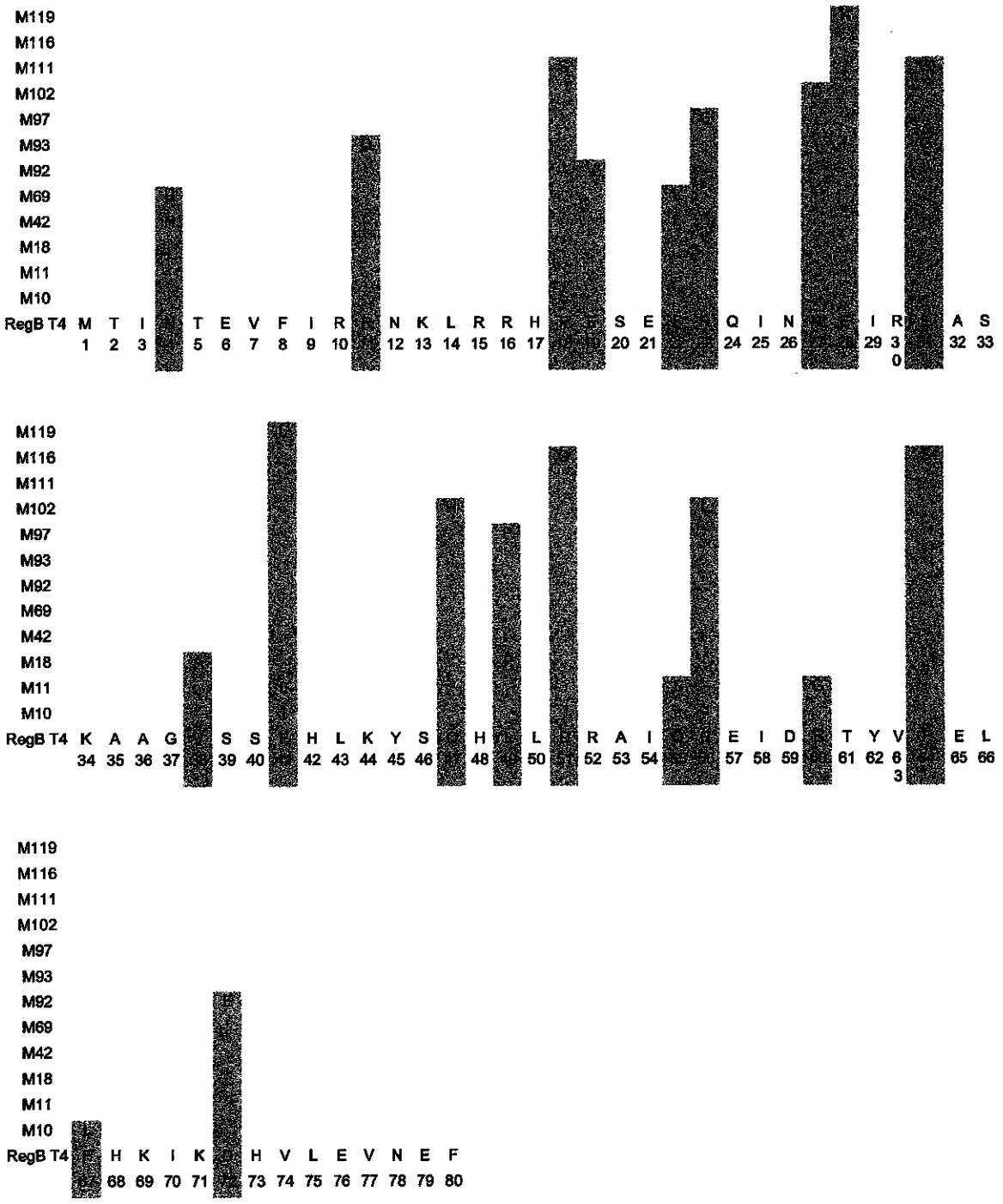


Figure 33 : Alignement des séquences des 12 mutants N-terminaux avec la séquence de la protéine RegB sauvage.

Dans les deux doubles mutants M18 et M92 la leucine 49 a été mutée en proline. Ce résidu introduit des contraintes structurales très fortes liées à l'existence d'une liaison covalente entre le groupe amine NH_2 et un carbone de la Cδ de la chaîne latérale. Ces contraintes obligent en général la chaîne polypeptidique à changer d'orientation. Lorsque ce changement d'orientation survient dans une zone initialement très structurée, il conduit souvent au bouleversement de la structure tridimensionnelle totale de la protéine et à la production d'une protéine mal repliée souvent inactive. Le résidu proline peut ainsi servir de « sonde » structurale qui donne une information sur le degré de structuration d'un fragment peptidique donné. L'effet des mutations M18 et M92 est très probablement un effet structural. Le résidu 49 appartient probablement à une zone structurée.

Les doubles mutants M11, M93 et M111 contiennent les substitutions E60G et E31G. A l'inverse de la proline, le résidu glycine du fait de sa chaîne latérale non-encombrante réduite à un atome d'hydrogène offre à la chaîne polypeptidique des degrés de liberté que ne permettent pas les 19 autres acides aminés. La substitution d'un résidu aminé par une glycine conduit souvent à une protéine mal repliée. Ceci s'expliquerait par le fait que l'univers des possibilités explorées par la chaîne polypeptidique au cours de son processus de repliement devient plus grand. Le processus permettant la formation des éléments de structures secondaires devient limitant et peut conduire à un défaut de repliement lorsque la protéine est surexprimée dans la bactérie. Comme les mutants M18 et M92, nous pensons que les trois mutants glycines ci-dessus sont inactifs à cause d'un problème structural. Les résidus E31 et E60 appartiennent probablement à des zones structurées.

Le mutant M10 contient trois substitutions de résidus phénylalanine dont deux en leucines et une en sérine. L'une au moins des trois phénylalanines F18, F41 et F67 apparaît donc importante pour l'activité de l'enzyme. L'implication du résidu phénylalanine dans l'activité d'un certain nombre de ribonucléases a été déjà décrite. La phénylalanine 120 de la ribonucléase A intervient dans la liaison du substrat grâce non seulement au caractère hydrophobe de sa chaîne latérale mais aussi aux électrons délocalisés des orbitales π du noyau aromatique (Chatani *et al.*, 2001). De la même façon, il a été démontré que la phénylalanine 80 de la ribonucléase MC1 (une ribonucléase qui clive la liaison phosphodiester du coté 5' d'une uridine) intervient aussi bien dans la liaison de l'uridine que dans l'activité hydrolytique dirigée contre un substrat 2',3'-cyclique (Numata *et al.*, 2001).

Les mutants M42 et M69 isolés indépendamment présentent la même séquence. L'asparagine 4 et/ou la phénylalanine 22 sont fonctionnellement importantes pour RegB. Le mutant M92 contient deux substitutions. La substitution du résidu acide aspartate 72 par le résidu acide équivalent glutamate n'est probablement pas à l'origine de la perte complète de l'activité de RegB. En revanche,

le remplacement du résidu glutamate 19 par une chaîne latérale neutre (valine) pourrait expliquer la perte d'activité. Le résidu E19 est donc probablement très important dans l'activité de l'enzyme.

Dans le double mutant M102, deux chaînes latérales polaires non chargées (asparagine 27 et Glutamine 47) ont été mutés en deux résidus polaire chargés (aspartate et histidine respectivement). Dans le mutant M119 on retrouve la mutation F41L et une mutation d'un résidu acide (E28) en un résidu basique (lysine). Il est impossible de conclure quant à l'effet individuel de chaque mutation sans la préparation par mutagenèse dirigée des simples mutants ponctuels.

En conclusion, les différents mutants présentés ci-dessus montrent clairement que l'extrémité N-terminale de RegB est fonctionnellement très sensible aux substitutions des résidus polaires (qu'ils soient chargés ou non) ainsi qu'aux substitutions de certains résidus à chaînes latérales aromatiques. Le glutamate 19 apparaît comme un résidu indispensable à l'activité de RegB. La partie N-terminale apparaît aussi comme sensible à l'introduction de contraintes structurales supplémentaires ou à l'apport de nouveaux degrés de liberté ce qui indique que cette partie est probablement bien repliée.

Plutôt que de préparer les simples mutants ponctuels correspondant à la partie N-terminale, nous nous sommes demandés si le profil de sensibilité à la mutagenèse dirigée de la partie C-terminale de RegB était comparable.

3.4 Résultats de la mutagenèse de la partie 3' du gène *regB* (partie C-terminale de RegB)

En appliquant le protocole de mutagenèse M2 à la partie 3' du gène *regB* (correspondant aux résidus 81 à 153), des clones viables ont été obtenus à l'issue de la fusion de la partie 5' du gène *regB* sauvage contenue dans le plasmide pMUC avec la partie 3' issue de la PCR mutagène. Cependant, pour un même protocole de mutagenèse et pour des quantités d'ADN comparables, le nombre de clones viables était inférieur d'un facteur 10 à celui obtenu lors de la mutagenèse de la partie N-terminale.

Les tests d'expression de 12 clones possédant le gène *regB* reconstitué sont présentés dans la Figure 34. 8 clones n'expriment aucune protéine et pourrait correspondre à des vecteurs réarrangés. Trois clones (C34, C67 et C94) expriment des protéines de tailles inférieures à celle de la protéine

RegB entière. Ils sont le résultat de l'apparition prématurée d'un codon stop. Enfin, seul le clone C72 exprime une protéine de la taille de RegB. Le séquençage de ce clone a révélé l'existence d'une codon stop au niveau de la position 145 qui produit une protéine ayant perdu ses 8 acides aminés C-terminaux. Nous pensons que le clone C72 était viable car la protéine qu'il exprimait était probablement très instable en solution. Nous avons pu vérifié cette hypothèse en préparant par centrifugation un extrait brut des protéines exprimées avec le clone C72. On a retrouvé la protéine tronquée dans le culot cellulaire sous la forme de corps d'inclusions.

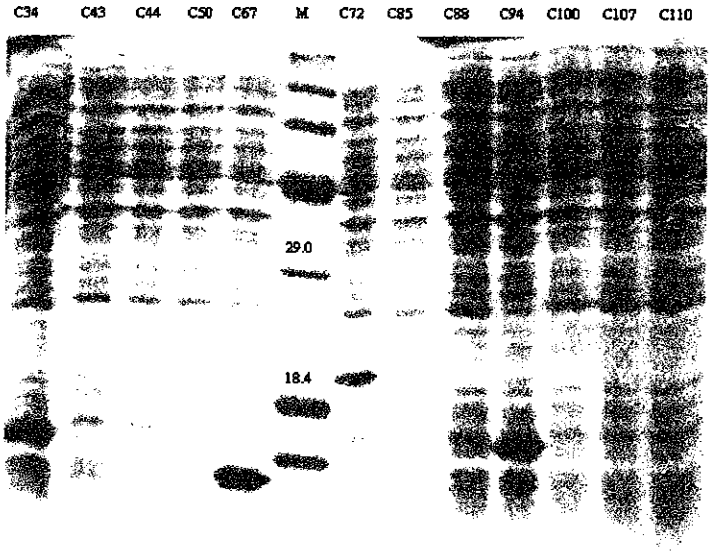


Figure 34 : Analyse de l'expression de 12 clones issus de la mutagenèse de la partie C-terminale de RegB.

Le Tableau 5 présente une analyse statistique comparative des résultats d'expression des clones issus de la mutagenèse des parties N-terminale et C-terminale de RegB. Dans ce tableau, le pourcentage des clones exprimant la protéine RegB entière varie dans le même sens que la probabilité de muter des positions importantes pour la structure et/ou la fonction de RegB *dans les conditions de mutagenèse et de sélection décrites précédemment.* Cette probabilité vaut 0.46 pour la partie N-terminale et est nulle pour la partie C-terminale. Même si ces probabilités sont très grossières car calculées sur un petit échantillon de 12 et 13 clones contenant le gène *regB* reconstitué, elles sont suffisamment disjointes pour affirmer qu'il y a une différence entre les deux parties de RegB. La partie N-terminale semble être plus sensible à la mutagenèse aléatoire que la partie C-terminale.

Il est évident que la probabilité nulle trouvée pour la partie C-terminale ne montre pas que cette partie est dispensable pour la structure et/ou la fonction de RegB. Elle montre tout simplement que la probabilité de trouver au hasard un résidu aminé indispensable à RegB est beaucoup plus faible dans la partie C-terminale que dans la partie N-terminale.

	Partie N-terminale	Partie C-terminale
Nombre de clones analysés	13	12
% clones n'exprimant aucune protéine	0	67
% clones exprimant RegB tronquée	54	33
% clones exprimant RegB entière	46	0

Tableau 5 : Analyse des pourcentages des différentes protéines exprimées à partir de clones recombinants issus de la mutagenèse de la partie N-terminale ou C-terminale de RegB. Ces pourcentages sont basés sur les gels d'expression des Figure 33 et 34.

Dans le chapitre qui suit nous allons conjuguer le résultat de mutagenèse aléatoire ci-dessus avec le résultat du chapitre 2 relatif à la caractérisation du chemin réactionnel catalysé par RegB pour investiguer avec plus de précision la partie N-terminale de RegB. Mutagenèse dirigée, Résonance Magnétique Nucléaire et Résonance Plasmonique de Surface seront utilisées à cet effet.

Chapitre 4

Investigation par mutagenèse dirigée de la partie N-terminale de RegB

Chapitre 4

Investigation par mutagenèse dirigée de la partie N-terminale de RegB

Les expériences RMN du ^{31}P du chapitre 2 ont montré que la ribonucléase RegB est une transphosphorylase qui clive la liaison phosphodiester P-O5' en produisant un composé phosphoré 2',3'-cyclique. Nous montrerons dans ce chapitre que toutes les ribonucléases connues qui produisent un tel composé procèdent par une catalyse acido-basique dont les acteurs sont des résidus aminés bien définis. Nous montrerons qu'une analyse exhaustive de la nature de ces résidus permet de mettre en évidence une stratégie de catalyse commune. En mettant ce résultat en relation avec les conclusions de mutagenèse aléatoire du chapitre précédent, nous consacrerons la suite du chapitre à l'étude du rôle catalytique éventuel de résidus aminés particuliers de RegB. Cette étude comportera deux étapes : la mutagenèse dirigée de ces résidus et la caractérisation des mutants ponctuels obtenus par des tests d'activité *in vivo* et *in vitro*, par des tests d'intégrité structurale (RMN) et par des tests de liaison au substrat (Résonance plasmonique de Surface : RPS).

4.1 Résidus aminés impliqués dans le mécanisme catalytique des ribonucléases de la famille des transphosphorylases

Plusieurs ribonucléases produisant un phosphodiester 2',3' cyclique ont été décrites dans la littérature. Pour la plupart, les résidus aminés impliqués dans la catalyse acido-basique ont été identifié en conjuguant des travaux de mutagenèses dirigées, d'études de courbes de titrage de l'activité de l'enzyme en fonction du pH et d'études de la structure tridimensionnelle de l'enzyme parfois en complexe avec le substrat.

Nous avons identifié 4 groupes de ribonucléases qui se distinguent par leurs structures primaires, leurs spécificités de coupure et la nature des résidus aminés impliqués dans la catalyse acido-basique. Il s'agit des groupes de la RNase A, de la RNase T1, de la RNase T2 et de la Barnase.

4.1.1 Le groupe de la RNase A

Ce groupe comprend toutes les ribonucléases homologues à la ribonucléase pancréatique bovine (RNase A). Ces ribonucléases se retrouvent exclusivement chez les mammifères, les oiseaux, les reptiles et les amphibiens. Elles clivent, avec une spécificité relative, la liaison phosphodiester consécutive à une base pyrimidique en libérant une extrémité 3'Phosphate.

En comparaison avec la RNase A, représentant type, les ribonucléases de ce groupe contiennent entre 3 et 4 ponts disulfures très conservés. Certaines de ces RNases peuvent être secrétées. Le Tableau 6 présente les 6 sous-groupes que comprend le groupe de la RNAse A.

Sous-groupe	**Exemple**
Ribonucléases pancréatiques secrétées	RNase A
Ribonucléases mammifères non secrétées	NDE (Neurotoxine dérivée des éosinophiles)
Ribonucléases 4 mammifères	Ribonucléase 4 humaine
Angiogénines mammifères	Angiogénine bovine
Ribonucléases de volailles	Ribonucléase hépatique du poulet
Ribonucléases de batraciens	Onconase de *Rana pipiens*

Tableau 6 : RNAses du groupe de la RNase A. Cette classification est inspirée de celle de Beintema J.J. (Beintema *et al.*, 1997).

Le mécanisme de catalyse de la RNase A a été extensivement étudié (Beintema *et al.*, 1997). Il en ressort que deux histidines H12 et H119 interviennent directement dans la catalyse acido-basique. Lors de l'étape de transphosphorylation, l'**histidine 12** joue le rôle de catalyseur général basique qui active le groupe hydroxyle en 2' par soustraction d'un proton. L'**histidine 119** joue le rôle de catalyseur général acide qui fournit un proton à l'oxygène 5' partant. Ces rôles sont exactement inversés lors de l'hydrolyse du phosphodiester cyclique. La mutation de l'une ou l'autre de ces histidines annule complètement l'activité de l'enzyme et témoigne du caractère indispensable de ces deux résidus.

Ces deux histidines sont structurellement et fonctionnellement très conservées dans toute la famille de la RNase A. La Figure 35 montre l'alignement des séquences de 10 ribonucléases appartenant au groupe de la RNAse A. Cette figure met en évidence la conservation des deux histidines catalytiques (H12 et H119).

RNase A	K	F	E	R	Q	H	M	D	S	S	P	Y	V	P	C	H	F	D	A	S	V
ECP	W	F	A	I	Q	H	I	S	L	N	P	V	V	P	V	H	L	D	T	T	I
EDN	W	F	E	T	Q	H	I	N	M	T	P	V	V	P	V	H	L	D	R	I	I
hAngiogénine	H	F	L	T	Q	H	Y	D	A	K	P	*	L	P	V	H	L	D	Q	S	I
hRNase 4	R	F	L	R	Q	H	V	H	P	E	P	Q	V	P	V	H	F	D	G	*	*
hRNase panc.	K	F	Q	R	Q	H	M	D	S	D	P	Y	V	P	V	H	F	D	A	S	V
Onconase	T	F	Q	K	K	H	I	T	N	*	P	*	A	P	V	H	F	V	G	*	V
Moel. Oss.	D	F	L	Y	K	H	M	D	F	P	*	*	L	P	V	H	L	D	G	T	S
Tortue	K	F	L	R	Q	H	V	D	Y	P	*	*	L	P	V	H	Y	D	K	S	I
Iguane	S	F	Q	N	K	H	I	D	Y	P	*	*	Q	P	V	H	F	E	L	V	L
						12										119					

Figure 35 : Alignement des séquences de 10 RNases appartenant au groupe de la RNase A. La numérotation est celle de la RNase A. Les délétions sont indiquées par des étoiles. Les histidines conservées sont en gras. ECP : Protéine cationique des éosinophiles, EDN : Neurotoxine dérivée des éosinophiles, hAngiogénine : Angiogénine humaine, hRNAse 4 : RNase 4 humaine, Onconase : Onconase de Rana pipiens, Moel. Oss. : RNase de la moelle osseuse de poulet, Tortue : RNase de Tortue, Iguane : RNase d'Iguane.

4.1.2 Le groupe de la RNase T1

La RNase T1 a été initialement isolée par Sato et Egami à partir d'une mixture enzymatique commerciale provenant du champignon *Aspergillus oryzae* (Sato *et al.*, 1957). Il s'agit d'une endoribonucléase de 11 kDa qui coupe avec une spécificité relative la liaison phosphodiester consécutive à une guanine. Plusieurs ribonucléases séquentiellement homologues à la RNase T1 ont

été isolées à partir de microbes eucaryotes. Ces RNases ont été classées en 5 sous-groupes en fonction de la localisation de leurs ponts disulfures. La RNase T1 contient par exemple deux ponts, l'un entre les cystéines 2 et 10, l'autre entre les cystéines 6 et 103. Ce dernier pont est caractéristique. Il est conservé dans tout le groupe de la RNase T1 (Yoshida *et al.*, 2001 ; Sevcik *et al.*, 1990).

Le mécanisme catalytique et la structure tridimensionnelle ont été investigués pour plusieurs représentants de ce groupe, notamment la RNase T1, la RNase C2 (*Aspergillus clavatus*), la RNase Th1 (*Trichoderma harzianum*) et la RNase Pb1 (*Penicillium brevi-compactum*). Il en ressort une catalyse commune de type acido-basique impliquant lors de la transphosphorylation le résidu **glutamate 58** qui soustrait le proton du groupe 2'OH et **l'histidine 92** qui donne un proton au groupe 5'OH partant (Osterman *et al.*, 1979 ; Steyaert *et al.*, 1990). Ces deux résidus sont très conservées comme le montre l'alignement de séquence de la Figure 36.

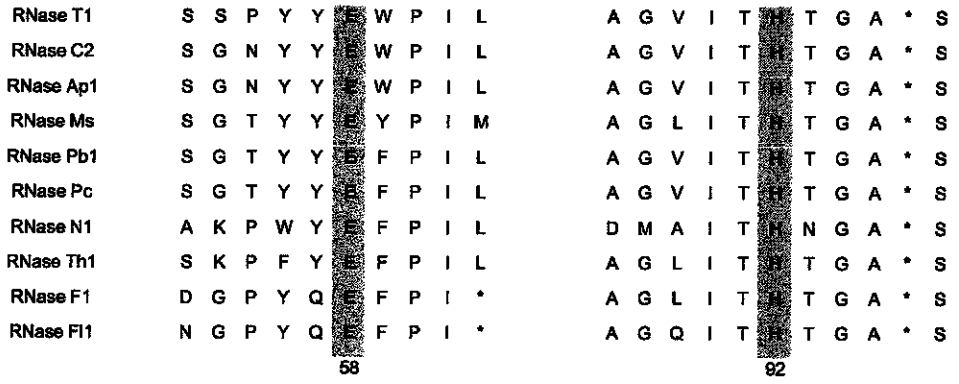


Figure 36 : Alignement des séquences de 10 RNases appartenant au groupe de la RNase T1. La numérotation est celle de la RNase T1. Les délétions sont indiquées par des étoiles. C2 : *Aspergillus clavatus*, Ap1 : *Aspergillus pallidus*, Ms : *Aspergillus saitoi*, Pb1 : *Penicillium brevi-compactum*, Pc : *Penicillium chrysogenum*, N1 : *Neurospora crassa*, Th1 : *Trichoderma harzianum*, F1 : *Fusarium moniliforme*, Fl1 : *Fusarium lateritium*.

Signalons que certaines publications antérieures à l'année 1996 considèrent que c'est **l'histidine 40** qui joue le rôle de catalyseur général basique lors de la transphosphorylation au lieu du glutamate 58. Des travaux de cristallographie classique (Ishikawa *et al.*, 1996) et de cristallographie résolue dans le temps (en utilisant un substrat lentement hydrolysable, l'*exo*-guanosine 2',3'-cyclo-phosphorothioate ; Zegers *et al.*, 1998) ont infirmé ces déclarations. Ils ont montré que le glutamate 58 se positionne dans le site actif selon une direction opposée à celle de l'histidine 92. Il se trouve à une très faible distance du groupe 2'OH et sa position est parfaitement superposable à celle de l'histidine 12 dans le complexe RNase A / 3'-GMP.

4.1.3 Le groupe de la RNase T2

La RNase T2 a été isolée à partir du champignon *Aspergillus oryzae*. Il s'agit d'une glycoprotéine de 36 kDa. Contrairement à la RNase T1, la RNase T2 fonctionne à des pH acides (le pH optimum vaut 4.5) et ne présente aucune spécificité de coupure (Irie *et al*, 2001).

Les RNases du groupe de la RNase T2 se retrouvent chez les virus, les bactéries (exemple de la RNase I d'*E. coli*), les champignons, les amibes, les plantes et les animaux. Elles possèdent toutes deux cassettes hautement conservées de 9 et 12 résidus aminés comprenant la quasi-totalité des résidus du site actif (Figure 37) (Watanabe *et al.*, 1990).

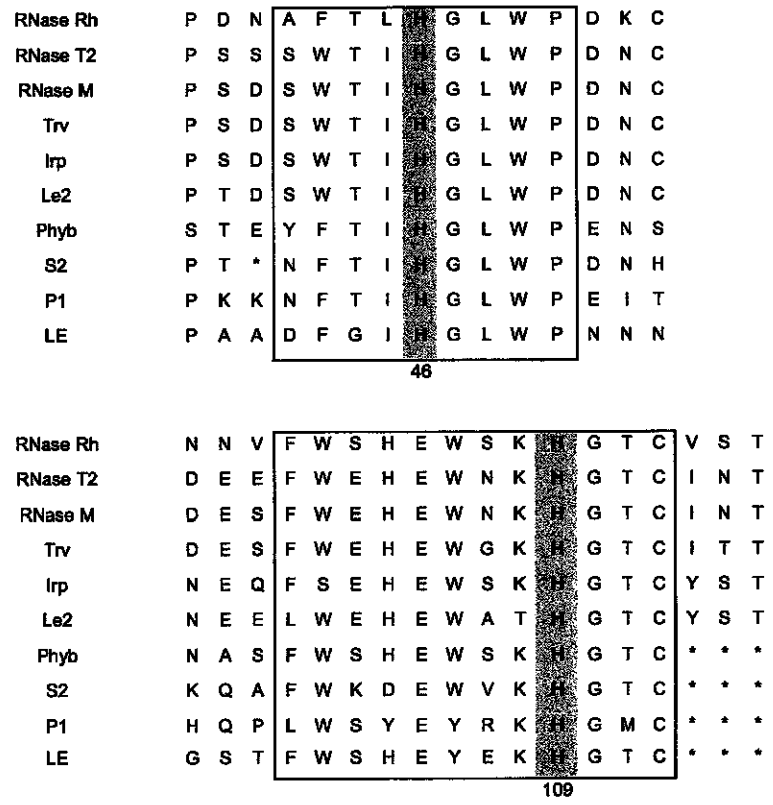


Figure 37 : Alignement des séquences de 10 ribonucléases de la famille de la RNase T2. La numérotation des résidus est celle de la RNase Rh. Les deux cassettes de 9 et 12 résidus caractéristiques du groupe de la RNase T2 sont encadrées. Les délétions sont indiquées par des étoiles. Rh : *Rhizopus niveus*, M : *Aspergillus saitoi*, Trv : *Trichoderma viride*, Le2 : *Lentinus edodes*, Phyb : *Physarum Polycephalum*, S2 : *Nicotiana alata*, P1 : *Petunia Inflata*, LE : *Lycopercison esculentum*. Les histidines catalytiques sont en gras. D'après Irie *et al.*, 2001.

L'étude détaillée de la catalyse acido-basique de la RNase Rh (*Rhizopus niveus*) montre que, lors de la transphosphorylation, l'**histidine 109** joue le rôle de base alors que l'**histidine 46** joue le rôle d'acide. Comme le montre la Figure 37, ces deux histidines sont conservées dans toutes les RNases du groupe (Ohgi *et al.*, 1992).

4.1.4 Le groupe de la Barnase

La Barnase est une enzyme extracellulaire secrétée par *Bacillus amyloliquefaciens* (souche H) (Hartley R.W., 1969). Les homologues de la Barnase se retrouvent chez d'autres souches *Bacillus*, chez certaines souches *Streptomyces* et chez certains champignons.

Bien que structurellement différente de la RNase A, la Barnase catalyse la même réaction de transphosphorylation. Elle utilise les résidus conservés **glutamate 73** et **histidine 102** respectivement comme receveur et donneur de protons lors de la catalyse acido-basique (Figure 38) (Mossakowska *et al.*, 1989 ; Guillet *et al.*, 1993).

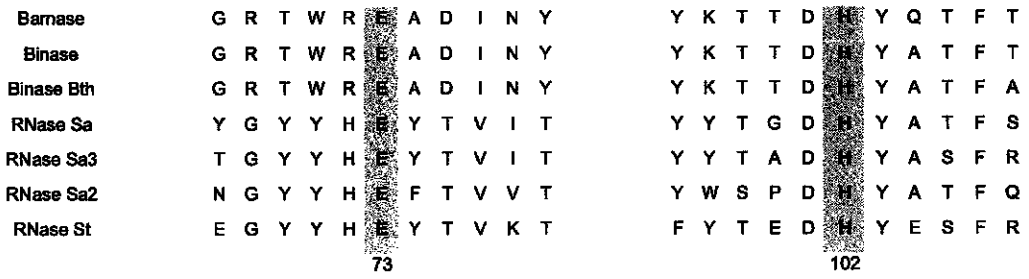
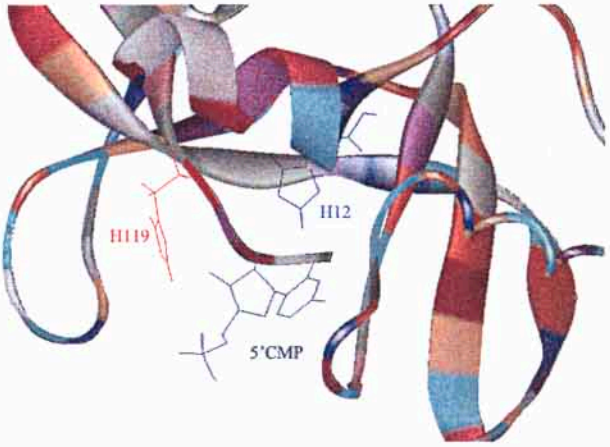
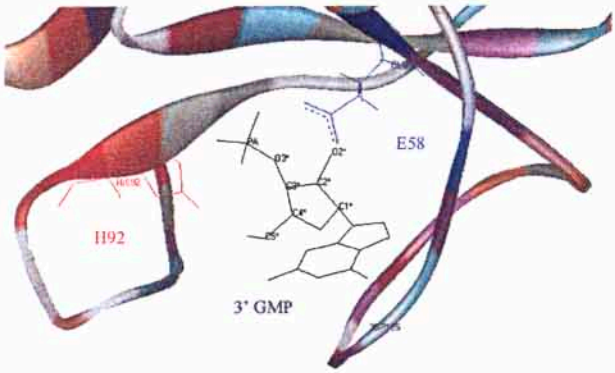


Figure 38 : Alignement des séquences de ribonucléases du groupe de la Barnase. La numérotation se réfère à celle de la séquence de la Barnase. Binase : *Bacillus intermedius* (Souche 7P), Binase Bth : *Bacillus thuringiensis*, Sa : *Streptomyces aureafaciens* (souche BMK), Sa3 : *Streptomyces aureafaciens* (souche CCM 3239), Sa2 : *Streptomyces aureafaciens* (souche R8/26), St : *Streptomyces erythreus* (souche NRRL 2338).

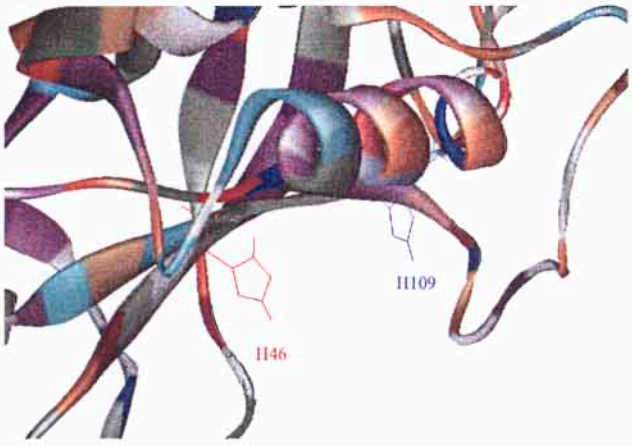
Nous avons représenté sur la Figure 39 ci-après les structures tridimensionnelles des sites actifs de représentants des 4 groupes de ribonucléases que nous venons d'exposer. Nous avons localisé sur cette figure les résidus qui lors de l'étape de transphosphorylation jouent les rôles de catalyseurs acides (couleur rouge) et de catalyseurs basiques (couleur bleue).



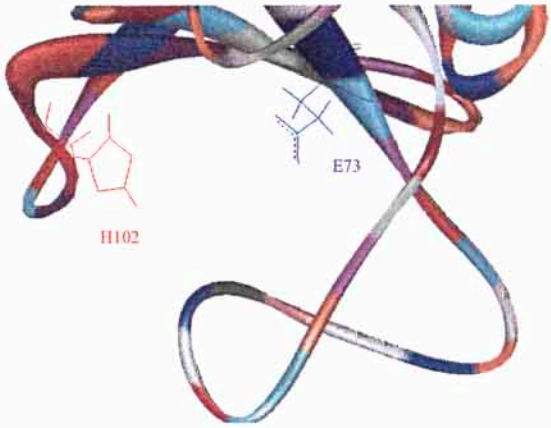
RNase A + 5'CMP



RNase T1 + 3'GMP



RNase Rh (Groupe RNase T2)



Barnase

Figure 39 : Structures tridimensionnelles montrant les résidus aminés impliqués dans la catalyse acido-basique des ribonucléases produisant un phosphodiester 2',3'-cyclique. Les résidus jouant le rôle de base lors de la transphosphorylation sont en bleu, ceux jouant le rôle d'acide sont en rouge. Dans le cas de la RNase A et de la RNase T1, la figure montre le complexe enzyme-mononucléotide. Les codes PDB des structures utilisées pour construire cette figure sont : 1RNN (RNase A), 1RLS (RNase T1), 1BOL (RNase Rh) et 1RNB (Barnase).

Cette figure montre que la réaction de transphosphorylation requiert toujours une **histidine** comme catalyseur général acide et une **histidine** ou un **glutamate** comme catalyseur général basique. Elle montre aussi que des ribonucléases aussi différentes séquentiellement et structurellement que la RNase A, la RNase T1, la RNase T2 et la Barnase ont adopté des stratégies de catalyse communes permettant la formation du phosphodiester cyclique.

L'implication systématique d'au moins une **histidine** dans la catalyse nous a conduit à nous poser la question suivante : « Comment RegB catalyse-t-elle la formation de son phosphodiester cyclique ? Utilise-t-elle des résidus histidines dans une stratégie comparable à celle des ribonucléases précédentes ou a-t-elle mis au point une stratégie radicalement différente ? ».

RegB possède 5 résidus histidines : H17, H42, H48, H68 et H73. Ces cinq résidus sont étonnamment regroupés dans la partie N-terminale de la protéine c'est à dire la partie établie comme étant la plus sensible à l'inactivation par mutagenèse aléatoire.

En absence de toutes homologies significatives entre la séquence RegB et celles des 4 groupes de ribonucléases présentées précédemment, nous avons émis l'hypothèse de l'implication d'au moins une histidine dans la catalyse médiée par RegB. Nous nous sommes proposés de tester la validité de cette hypothèse en mutant en alanines les histidines de RegB en caractérisant les mutants obtenus par des tests fonctionnels et structurels.

4.2 Sélection des résidus histidines à muter

Avant de muter les cinq histidines de RegB, nous avons d'abord cherché si ces résidus étaient conservés dans les versions de RegB codées dans les génomes de virus apparentés au bactériophage T4 ; c'est à dire des virus appartenant à la famille *Myoviridae*, genre *Pseudo-T4*. 16 séquences du gène *regB* appartenant des virus Pseudo-T4 ont été déposées dans la banque *GenBank* par Piesiniene et ses collaborateurs. Malheureusement, parmi ces séquences, 15 avaient entre 99.3% et 100 % d'identité de séquence avec le gène *regB* du phage T4. L'unique séquence significativement éloignée de T4 est celle du phage RB69.

La séquence du gène *regB* du phage RB69 alignée avec celle du phage T4, présentée dans la Figure 40, montre que seules quatre histidines (H42, H48, H68 et H73) sont conservées. Il est légitime de penser que l'histidine H17 du phage T4 (mutée en Isoleucine dans le phage RB69) est dispensable pour l'activité de RegB. Nous avons donc décidé de muter en alanines uniquement les quatre histidines **H42, H48, H68** et **H73**. Le choix du résidu de substitution est un choix standard qui se justifie par le caractère neutre, non réactif et peu encombrant du résidu alanine.

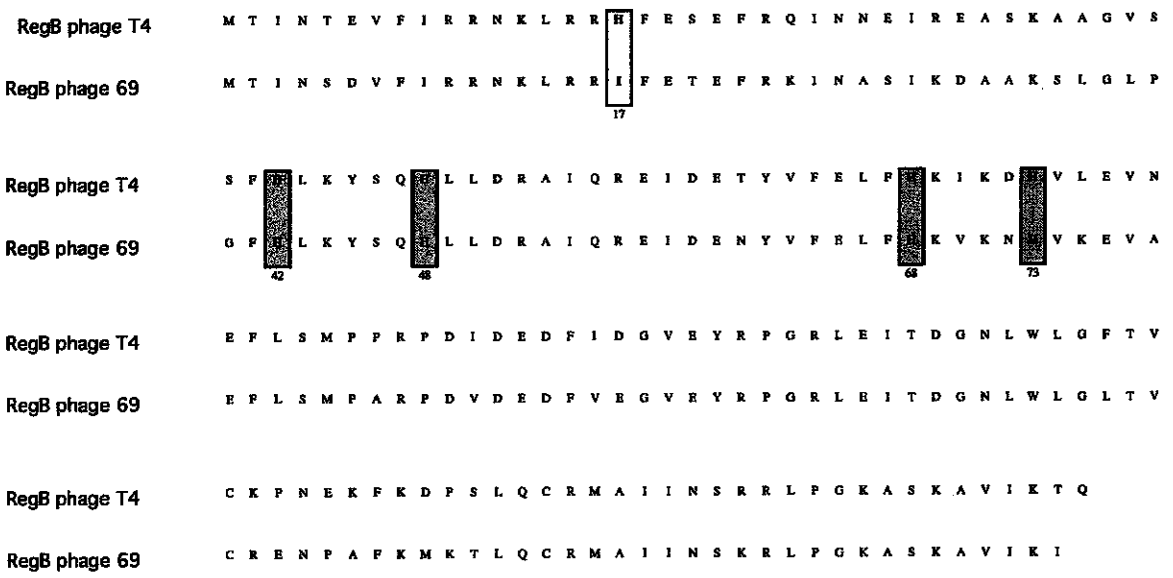


Figure 40 : Alignement des séquences de la ribonucléase RegB du phage T4 et du phage RB69. Seules les histidines H42, H48, H68 et H73 sont conservées.

4.3 Mutagenèse dirigée des histidines 42, 48, 68 et 73

Pour muter les 4 histidines conservées de RegB, nous avons utilisé comme matrice le plasmide pARNU2 contenant la version sauvage du gène *regB* (voir Chapitre 1). Pour chaque mutation ponctuelle nous avons synthétisé un couple d'oligonucléotides d'orientations inverses et dont les extrémités 5' coïncident au niveau du codon contenant la mutation histidine-en-alanine. Une réaction de polymérisation en chaîne en présence de la matrice pARNU2, des deux oligonucléotides mutateurs et de la *Pfu* Polymérase (Cline *et al.*, 1996) permet de synthétiser de nouvelles copies du plasmide pARNU2 ayant incorporé la mutation étudiée. Ces nouvelles copies plasmidiques se distinguent de l'ADN matrice parental (pARNU2) par l'absence de super-enroulements liée à l'absence de liaison phosphodiester entre le premier nucléotide 5' et le dernier nucléotide 3' ainsi que par l'absence de méthylation des sites de restriction se trouvant sur le plasmide.

La fraction d'ADN relaxée est constituée d'un ADN hémi-néo-synthétisé (un brin synthétique et un brin parental) et d'un ADN néo-synthétisé pur. Pour éliminer l'ADN hémi-néo-synthétisé, nous avons réalisé une digestion avec l'enzyme de restriction *DpnI* qui ne clive que la séquence consensus GA|TC contenant des adénines méthylées c'est à dire contenant au moins un brin parental (amplifié *in vivo* dans *E. coli*). Après une étape de purification sur gel d'agarose de l'ADN relaxé résistant à la coupure par *DpnI*, cet ADN est ligaturé avec la T4 ADN ligase et utilisé pour transformer par

électroporation la bactérie XL1-Blue. L'ADN plasmidique de deux clones recombinant est extrait puis séquencé pour vérifier la présence de la mutation recherchée.

L'efficacité de mutagenèse dirigée de ce protocole a été toujours voisine de 100 % grâce aux étapes décisives d'élimination des ADN parentaux par digestion enzymatique et purification sur gel. Les plasmides contenant les mutations H42A, H48A, H68A et H73A ont été appelé respectivement pH42A, pH48A, pH68A et pH73A.

4.4 Caractérisation *in vivo* de l'effet des mutations histidine-en-alanine

L'ADN matrice utilisé pour la mutagenèse des histidines est le plasmide pARNU2 qui contient la copie sauvage du gène *regB* sous le contrôle du promoteur T7lac et d'un RBS Phi 10 modifié. Nous avons vu dans le Chapitre 1 qu'à cause de la toxicité de RegB, pARNU2 ne peut se maintenir dans la souche bactérienne BL21(λDE3) même en absence d'inducteur. Nous avons donc émis l'hypothèse que si l'une des quatre mutations histidine-en-alanine de RegB réduisait significativement l'activité de RegB cela devrait se traduire par une diminution de l'effet toxique de la protéine qui permettra éventuellement le maintien du plasmide correspondant dans la souche BL21(λDE3).

Nous avons transformé la souche BL21(λDE3) avec les plasmides pH42A, pH48A, pH68A et pH73A. Nous avons utilisé le plasmide pARNU2 comme témoin négatif de transformation (aucun clone viable) et le plasmide pERL52L (contenant le mutant ponctuel inactif RegB R52L) comme témoin positif. Le résultat des transformations est présenté dans la Figure 41.

Seuls les plasmides pH48A et pH68A se maintiennent dans la bactérie BL21(λDE3) suggérant que l'activité des mutants correspondants est significativement diminuée. On note par ailleurs que les colonies obtenues avec le mutant H68A sont plus petites que celles obtenues avec le mutant H48A. Ceci qui suggère que la mutation H48A affecte plus l'activité de RegB que la mutation H68A.

Puisqu'il n'est pas possible à ce stade d'exclure des hypothèses comme un défaut de repliement de la protéine mutante ou la disparition d'une interaction cruciale avec une autre molécule

partenaire, nous somme proposés de purifier et d'étudier *in vitro* le comportement des mutants histidines.

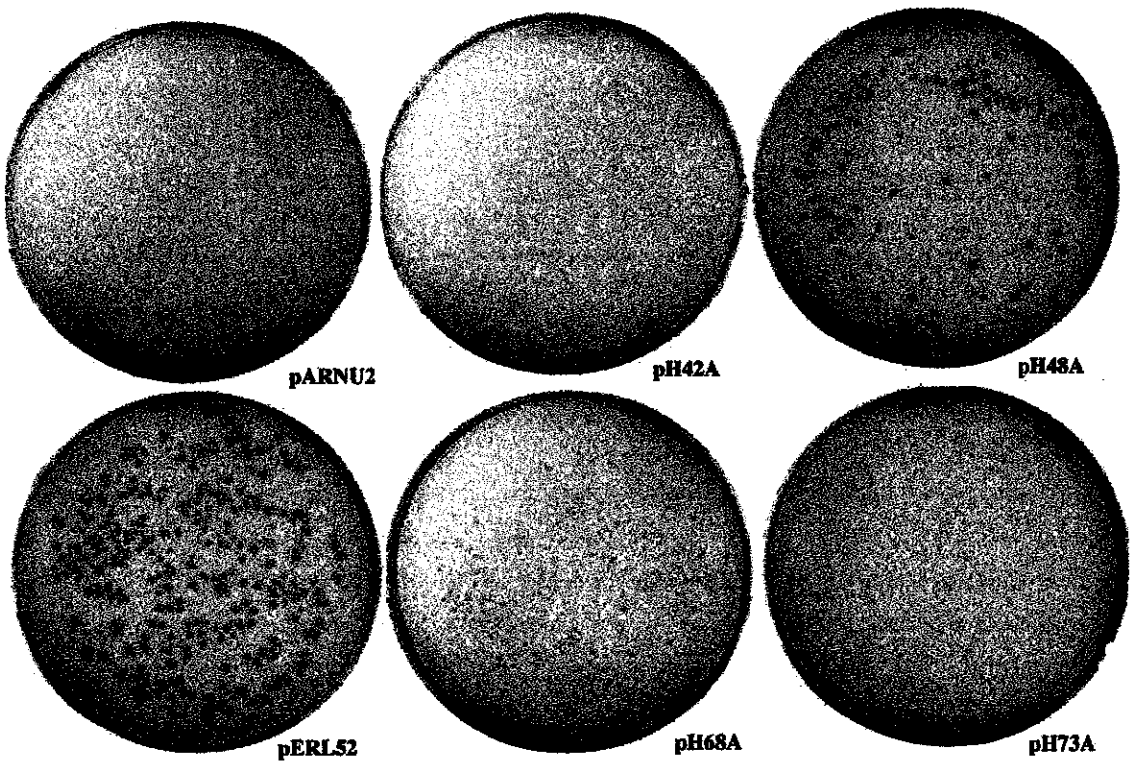


Figure 41 : Etude *in vivo* de la toxicité des 4 mutants ponctuels histidine-en-alanine de RegB. La souche bactérienne BL21(λDE3) a été transformée avec le plasmide pARNU2 qui contient la version sauvage de RegB (témoin négatif), le plasmide pERL52 qui contient la version inactive RegB R52L (témoin positif) et les plasmides pH42A, pH48A, pH68A et pH73A contenant respectivement les mutations H42A, H48A, H68A et H73A.

4.5 Etude des cinétiques de coupure de l'ARN Selex22tb par les différents mutants de RegB

Nous avons exprimé la version sauvage de RegB ainsi que les 4 mutants histidine-en-alanine dans la souche bactérienne XL1-Blue en utilisant le système d'induction avec le bactériophage λCE6. Seul le mutant inactif R52L a été produit avec le système d'expression standard BL21(λDE3)/IPTG.

Afin d'étudier et de comparer les réactivités des différentes versions de RegB, nous avons utilisé le modèle expérimental de coupure de l'ARN Selex22tb (voir Chapitre 2, paragraphe 2.5.2). Les réactions de coupure de l'ARN Selex22tb ont été réalisées dans les conditions explicitées dans la partie matériels et méthodes en présence d'un excès de substrat (10 fois plus d'ARN que de RegB). Afin d'éviter les coupures parasites dues aux ribonucléases contaminantes de type RNase A, nous avons ajouté des unités d'inhibiteurs de RNase A dans la solution de coupure (RegB est insensible à cet inhibiteur). L'apparition de l'ARN 15-mer, produit de coupure de l'ARN Selex22tb par RegB, est suivie pendant 40 min, quantifié puis reportée sur la Figure 42.

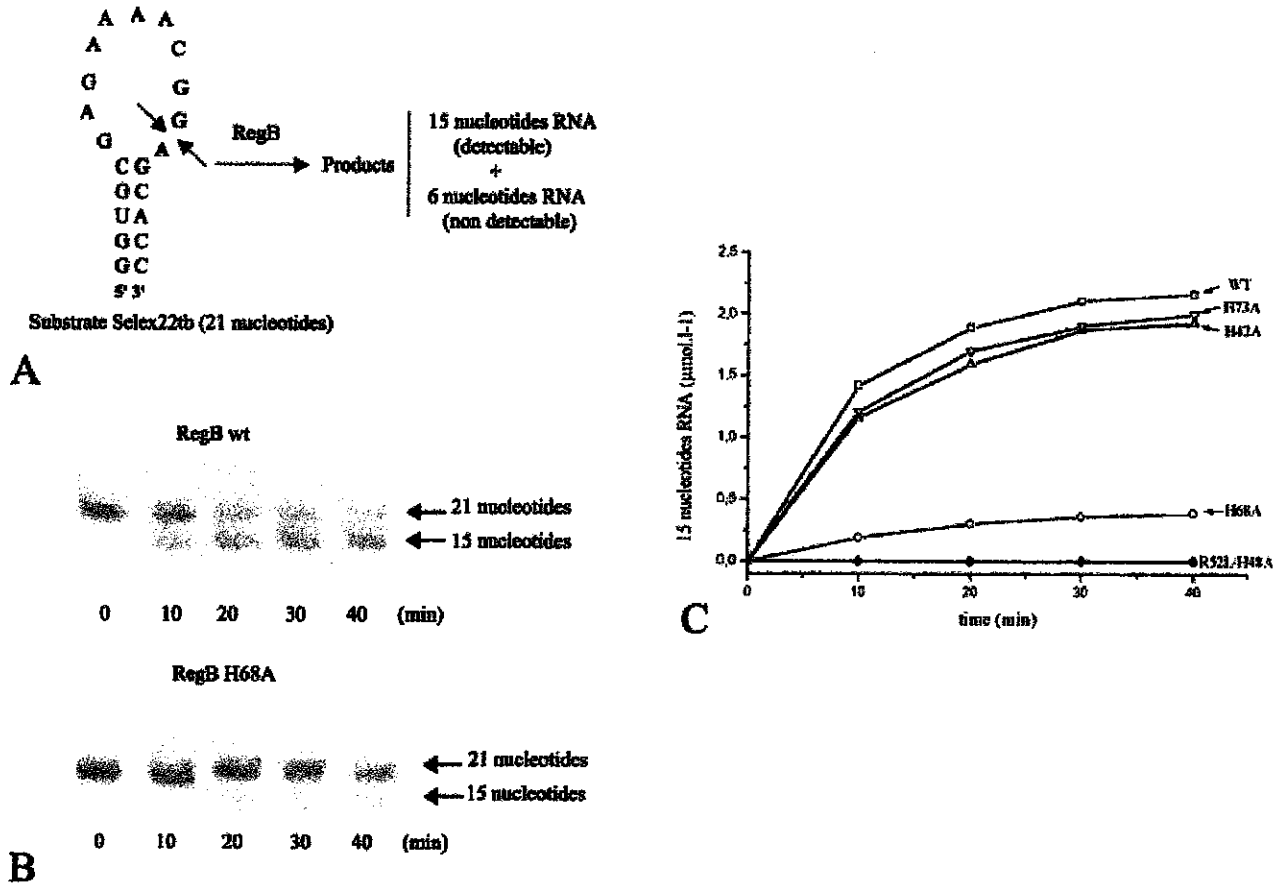


Figure 42 : Cinétiques de clivage de l'ARN Selex22tb par les différentes versions de RegB. (**A**) L'ARN Selex22tb comprend 21 nucléotides. Clivé au milieu de la séquence GGAG, il produit deux fragments de 15 et de 6 nucléotides. Ces deux fragments sont séparés sur un gel d'acrylamide 20 %, colorés au BET, excités aux Ultra-violets et quantifiées par l'appareil VL Imager. (**B**) Exemples de gels obtenus avec RegB sauvage (wt) et RegB H68A. (**C**) Apparition de l'ARN 15-mer en fonction du temps en présence de la version sauvage de RegB (wt) et des versions mutées H42A, H48A, R52L, H68A et H73A.

La Figure 42 montre que la protéine RegB sauvage clive près de 58 % du substrat initial en 40 minutes alors que la protéine RegB R52L est complètement inactive.

Les protéines RegB H42A et H73A clivent l'ARN Selex22tb avec des vitesses comparables à celles de la protéine sauvage. En fait, une légère différence est observée (55% de coupure pour les deux mutants contre 58 % pour le sauvage) mais nous pensons que cette différence est non significative car elle se trouve dans l'intervalle d'erreur expérimentale liée au protocole de quantification d'ARN utilisé.

En revanche, il apparaît clairement sur la Figure 42 que la mutation **H48A** annule complètement l'activité de RegB alors que la mutation **H68A** réduit significativement cette activité sans l'abolir complètement. Ce résultat est en concordance avec l'observation de la Figure 41 qui montre que, bien que les deux mutants H48A et H68A soient viables, le mutant H68A pousse moins vite que le mutant H48A. L'activité résiduelle de RegB H68A serait à l'origine du retard de croissance de ce mutant.

Pour expliquer l'annulation et la diminution de l'activité de RegB pour ces deux mutants nous avons émis trois hypothèses : une perturbation du repliement global de la protéine, une perte de la capacité à la lier le substrat ou la perte d'une chaîne latérale catalytique. Afin de tester la validité de la première hypothèse nous avons sondé par spectroscopie de corrélation nucléaire $^{1}H/^{15}N$ les structures globales des mutants H48A et H68A.

4.6 Etude par RMN de la structure globale des mutants H48A et H68A

L'expérience de résonance magnétique nucléaire $^{1}H/^{15}N$ HSQC (Heteronuclear Single Quantum Correlation) (Mori *et al.*, 1995) permet d'obtenir une carte bidimensionnelle dans laquelle chaque tache représente « un échange d'information magnétique » entre le noyau proton amide du squelette de la protéine et le noyau de l'atome d'azote auquel il est attaché (Figure 43).

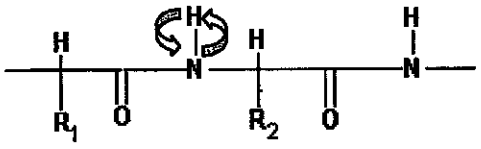


Figure 43 : Noyaux échangeant de l'information magnétique lors d'une expérience hétéronucléaire de type HSQC. Le schéma représente un fragment du squelette peptidique. En bleu : noyau ^{15}N, en rouge : noyau ^{1}H.

Chaque tache sur cette carte représente les corrélations $^{1}H/^{15}N$ d'un résidu particulier de la protéine. La position de cette tache, repérée sur le spectre par un couple de déplacements chimiques dans les dimensions proton et azote, est étroitement liée à l'environnement structural du résidu aminé étudié. Le spectre HSQC global constitue donc une empreinte à deux dimensions de la structure tridimensionnelle de la protéine. Toute modification de la position d'un résidu dans cette structure se traduira par le déplacement de la tache correspondante sur le spectre HSQC.

Nous avons préparés des échantillons de RegB H48A et RegB H68A marqués uniformément à l'isotope ^{15}N en utilisant le système d'expression efficace BL21(λDE3)/IPTG en milieu minimum contenant du chlorure d'ammonium $^{15}NH_4Cl$ comme source d'azote. Puisqu'il n'est pas possible de préparer un échantillon marqué de RegB sauvage à cause de l'incompatibilité des rendements d'expression de cette protéine avec les quantités requises pour la RMN structurale, nous avons préparé un échantillon marqué d'un troisième mutant indépendant : le mutant RegB R52L. La carte HSQC de ce dernier servira d'élément de comparaison pour étudier les cartes des deux mutants histidine-en-alanine. Nous avons enregistré les spectres HSQC des trois versions de RegB sur un spectromètre 600 MHz. Le résultat est présenté dans la Figure 44.

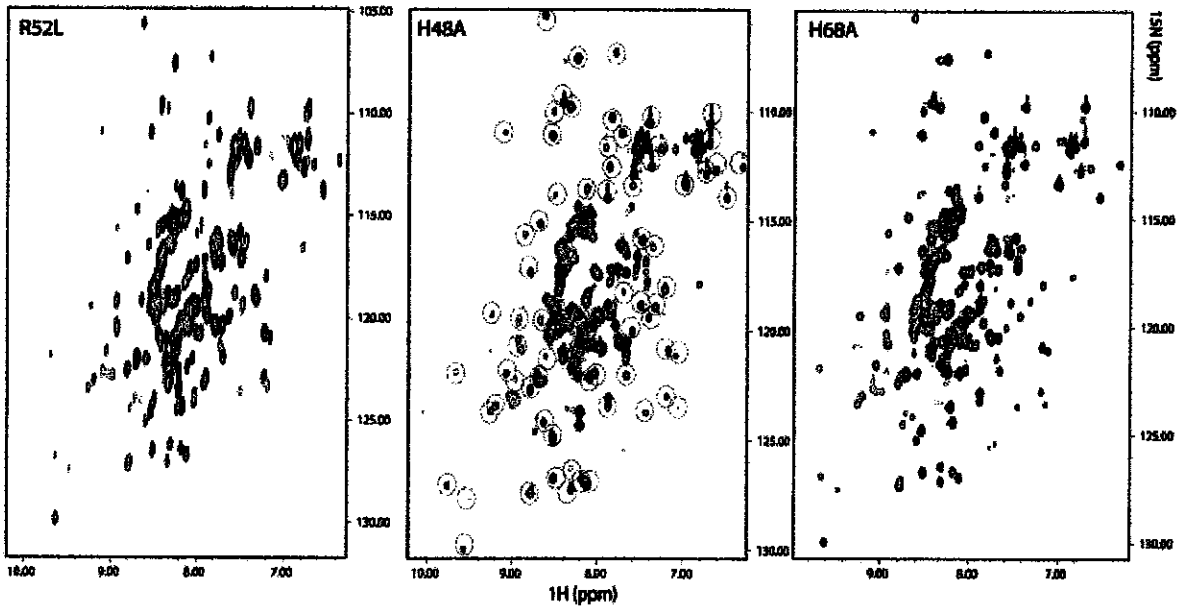


Figure 44 : Spectres $^{1}H/^{15}N$ HSQC des versions R52L, H48A et H68A de RegB. Les spectres ont été enregistrés à 25 °C sur un spectromètre Bruker DRX600. Les pics entourés par un cercle sur le spectre de RegB H48A se retrouvent dans les mêmes positions dans les spectres des mutants indépendants RegB R52L et RegB H68A.

Nous avons entourés par des cercles les pics disperses les mieux résolus du spectre de la protéine RegB H48A. Ces pics correspondent à des résidus aminés se trouvant dans des régions très structurées. Ils constituent une moyen de sonder la structure tridimensionnelle de RegB beaucoup plus précis qu'un simple spectre de dichroïsme circulaire qui ne renseigne que sur la composition globale en structures secondaires d'une protéine donnée. Ces pics se retrouvent exactement aux mêmes endroits dans le spectre du mutant de référence RegB R52L et du mutant RegB H68A. Puisqu'il est très peu probable que trois mutations indépendantes provoquent les mêmes altérations structurales, il est raisonnable de penser que les mutations H48A et H68A ne perturbent pas la structure globale de RegB. Mais qu'en est-il de leurs capacités à lier le substrat ARN ?

4.7 Quantification par Résonance Plasmonique de Surface de la liaison des mutants de RegB à l'ARN substrat

La résonance plasmonique de surface (RPS) est une méthode physique puissante qui permet d'étudier en temps réel les interactions intermoléculaires et de donner des informations à la fois sur l'état d'équilibre et sur les cinétiques de formation du complexe (Katsamba *et al.*, 2002).

La RPS utilise un phénomène optique qui a lieu lorsqu'une lumière incidente est réfléchie à la surface d'une fine couche métallique (en général de l'or). A la surface de ce métal se déplacent des « plasmons de surface » c'est à dire des ondes oscillantes possédant une densité de charge. Ces plasmons ne peuvent être excités par la lumière incidente que si cette dernière arrive avec un angle d'incidence particulier Θ ; c'est à dire lorsque le vecteur d'onde de la lumière incidente coïncide avec celui des plasmons de surface. Dans ce cas, une partie de la lumière incidente est absorbée et se propage dans la couche métallique sous la forme d'une onde perpendiculaire à la surface (onde évanescente). On parle alors de résonance plasmonique de surface. Cette absorption de lumière incidente se répercute sur la lumière réfléchie qui voit son intensité devenir minimale.

La valeur de l'angle Θ, dit de résonance, dépend étroitement de l'indice de réfraction au voisinage de la couche métallique. Le changement de la composition (adsorption ou désorption, fixation ou relargage) de l'interface change cet indice, ce qui conduit à un changement de l'angle de résonance. Des considérations théoriques montrent que Θ est proportionnel à la concentration de macromolécules en surface jusqu'à une valeur élevée de 50 $ng.mm^{-2}$. La Figure 45 résume le phénomène de RPS et montre qu'on peut suivre en temps réel la fixation d'un ligand à son partenaire en échantillonnant au cours du temps la variation de l'angle de résonance Θ.

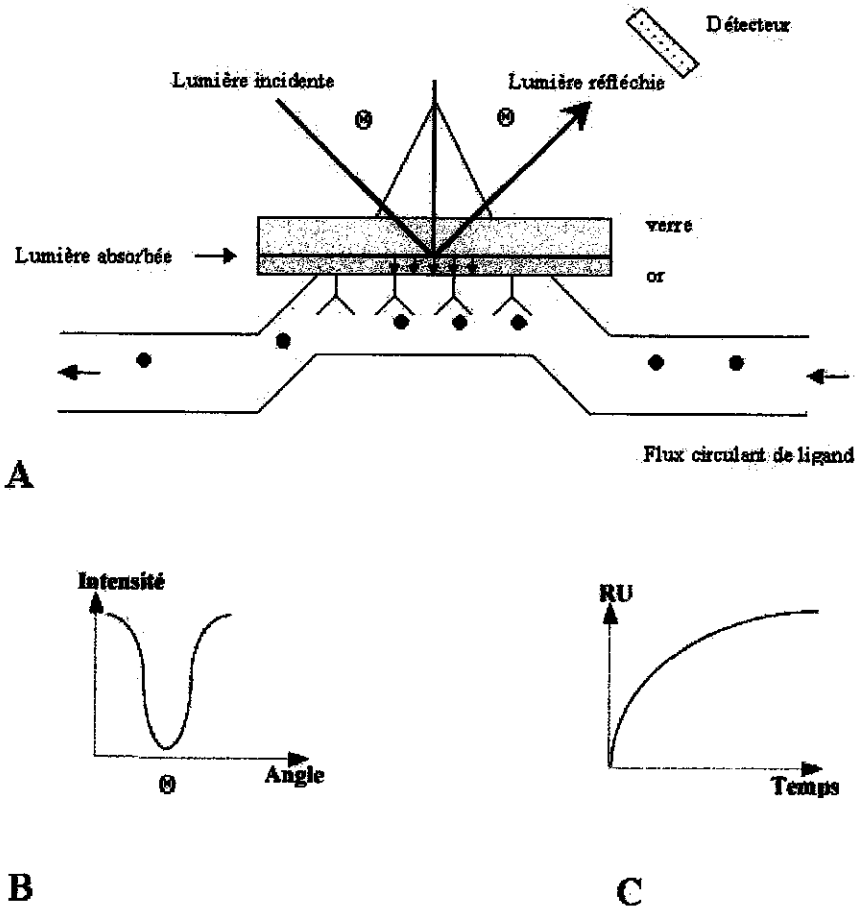
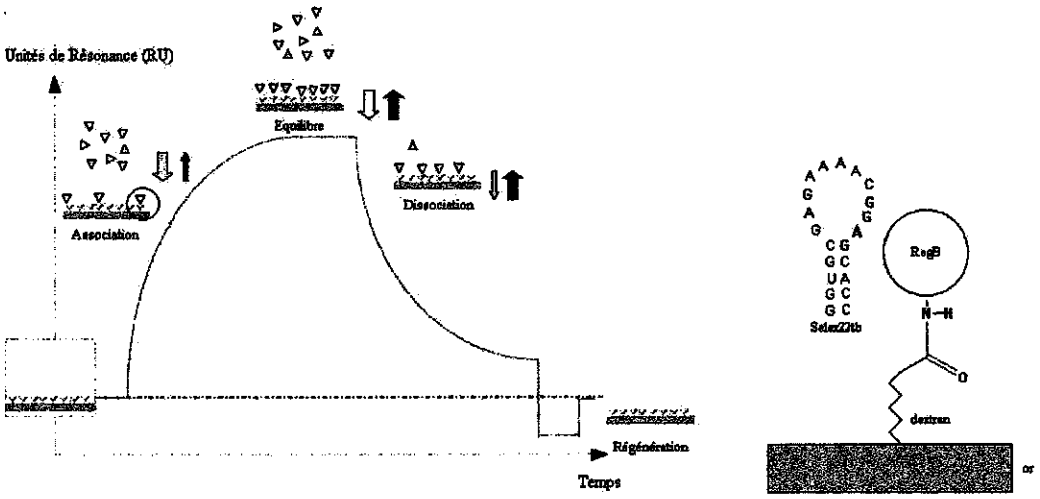


Figure 45 : Principe de la résonance plasmonique de surface. (A) Une lumière incidente d'une longueur d'onde donnée produite par prisme vient frapper une surface de verre sous laquelle se trouve une mince couche d'or. Cette lumière est normalement complètement réfléchie sauf si elle arrive avec un angle d'incidence particulier Θ. Dans ce cas, une partie de la lumière est absorbée par les plasmons de surface de la couche d'or. L'intensité de la lumière réfléchie (quantifiée par un détecteur) devient alors minimale (B). L'angle Θ dépend de la quantité de matière fixée au voisinage de la couche d'or. (C) Lorsqu'un ligand circulant (boule rouge) se fixe sur son partenaire (Y) l'angle Θ varie. On peut suivre cette variation au cours du temps par une grandeur appelée RU (Unité de Résonance, 1 RU vaut 10^{-4} degré).

Nous avons appliqué cette technique pour quantifier l'interaction entre les différentes versions de la protéine RegB et l'ARN substrat modèle que nous avons choisi : l'ARN Selex22tb. Nous avions le choix entre fixer la protéine RegB et faire circuler un flux de l'ARN Selex22tb ou fixer l'ARN et faire circuler RegB. Nous avons opté pour la première solution pour deux raisons. D'abord, il est plus facile de fixer la protéine RegB à la surface de bio-senseurs commerciaux destinés aux mesures par

RPS que de fixer l'ARN. En effet, RegB peut être fixée covalemment sur la couche de dextran juxtaposant le film d'or en utilisant simplement la chimie de couplage $COOH/NH_2$ alors que le couplage de l'ARN est plus fastidieux et demande la synthèse de molécules d'ARN biotinylées et l'utilisation de bio-senseurs à contenant des molécules de streptavidines préalablement immobilisées. Par ailleurs, à cause de la sensibilité des molécules d'ARN aux RNases contaminantes, il est préférable d'injecter un flux d'ARN pendant quelques secondes juste au moment de la mesure RPS que de fixer l'ARN et le laisser de façon continue en contact avec le tampon de circulation.

La Figure 46 montre les caractéristiques principales d'un « sensorgramme » : courbe exprimant la variation de l'angle de résonance (en unité de résonances) en fonction du temps.



A B

Figure 46 : Caractéristiques d'un sensorgramme, exemple de l'interaction RegB/Selex22tb. (A) Différentes partie d'un sensorgramme. L'association est la phase pendant laquelle un flux de l'ARN Selex22tb circule à la surface du bio-sensor contenant la protéine RegB immobilisée. La fixation (flèche bleue) prédomine sur le relargage (flèche rouge). La courbe à pour équation une exponentielle croissante. A l'équilibre, la fixation et le relargage se faisant à la même vitesse, l'angle de résonance ne varie plus. Pendant la dissociation, le flux de l'ARN Selex22tb est remplacé par un flux de tampon. L'ARN se dissocie selon une exponentielle décroissante. Une étape de régénération de la surface permet de dissocier les molécules résiduelles afin de pouvoir réutiliser la surface pour une nouvelle mesure. (B) Agrandissement montrant le couplage de RegB à la surface d'or par l'intermédiaire d'un couplage $COOH/NH_2$.

Nous avons fixé chacune des versions de RegB (RegB sauvage, H42A, H48A, R52L, H68A et H73A) sur un bio-sensor particulier et nous avons enregistré pour chaque protéine deux

sensorgrammes indépendants en utilisant deux concentrations différentes d'ARN. Pour extraire les constantes d'affinités RegB/ARN des sensorgrammes obtenus, nous avons utilisé un modèle de fixation simple uniphasique dans lequel la formation et la disparition du complexe sont dictées respectivement par la constante de vitesse d'association k_a et de dissociation k_d décrites dans l'équation I. La constante de dissociation du complexe RegB:ARN peut alors s'exprimer selon l'équation (II).

$$\text{RegB} + \text{ARN} \underset{k_a}{\overset{k_d}{\leftrightarrow}} \text{RegB:ARN} \qquad \text{(I)}$$

$$K_D = k_d/k_a \qquad \text{(II)}$$

Les valeurs des constantes k_a et k_d sont extraites des parties croissante et décroissante du sensorgramme obéissant aux équations (III) et (IV) respectivement où RU_{max} désigne la valeur RU à l'équilibre et t0 la date de l'instant où l'injection de l'ARN est arrêtée (début de la dissociation).

$$RU(t) = [RU_{max}/(1 + K_D/[ARN])] \,.\, [1 - \exp - [\, [\, k_a \,.\, [ARN] + k_d] \,.\, t\,]\,] \qquad \text{(III)}$$

$$RU(t) = RU_{max} \,.\, \exp - [\; k_d \,.\, [t - t_0\,]\,] \qquad \text{(IV)}$$

Il est important de noter que l'équation (I) ci-dessus suppose qu'on néglige la disparition du complexe RegB:ARN par clivage enzymatique (dans le cas où la version de RegB est totalement ou partiellement active). Cette hypothèse est justifiée par le fait que nous avons enregistré les sensorgrammes sur une courte période (2.8 minutes) et à 25 °C. Dans ces conditions, nous avons estimé que le pourcentage de clivage de l'ARN Selex22tb n'excèderait pas 10 %. En effet, la Figure 42 montre que le pourcentage de clivage de l'ARN est de 10 % après 2.8 min et à 37 °C et il est par ailleurs communément admis d'après la loi d'Arrhenius que la vitesse d'une réaction diminue d'un facteur voisin de 2 lorsque la température passe de 37 °C à 25 °C.

Les résultats des mesures SRP sont consignés dans le Tableau 7 qui regroupe les constantes cinétiques et les constantes d'équilibre. La Figure 47 compare sous forme d'histogrammes les intervalles d'erreur des valeurs du K_D pour RegB sauvage et les quatre mutant histidines.

Protéine	k_a ($(mol^{-1}.l.s^{-1})$)	k_d (10^{-4} s^{-1})	K_D ($\mu mol.l^{-1}$)
RegB sauvage	75 ± 3	1.1 ± 0.3	1.5 ± 0.5
RegB R52L	4.0 ± 0.5	0.60 ± 0.03	15 ± 2
RegB H42A	75 ±5	1.5 ± 0.2	2.0 ± 0.3
RegB H48A	53 ± 2	1.0 ± 0.1	1.9 ± 0.3
RegB H68A	64 ± 10	1.6 ± 0.4	2.5 ± 0.3
RegB H73A	64 ± 10	1.4 ± 0.5	2.2 ± 0.2

Tableau 7 : Constantes cinétiques et constantes d'équilibre à 25 °C de l'interaction des 6 versions de RegB avec l'ARN Selex22tb.

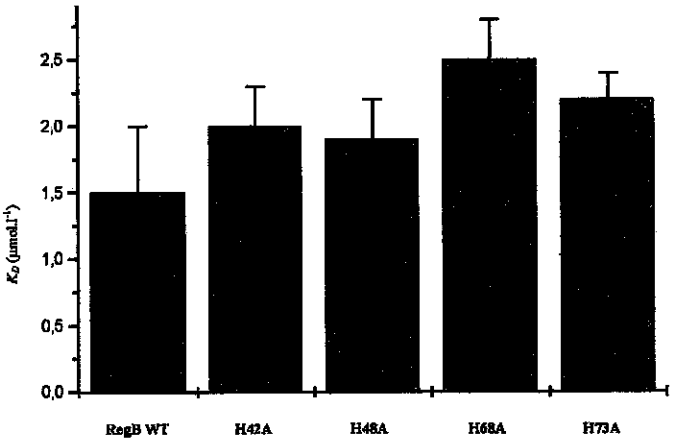


Figure 47 : Les constantes de dissociations K_D (avec leurs intervalles d'erreur) de RegB sauvage et des 4 mutants histidine-en-alanine.

Le tableau ci-dessus montre que les mutants H42A, H48A et H73A ont la même constante de dissociation que la protéine RegB sauvage. Ceci indique que les mutations introduites, en particulier la mutation H48A, ne modifient pas la capacité de l'enzyme à lier son substrat. L'affinité du mutant R52L est en revanche diminuée d'un ordre de grandeur par rapport au sauvage ce qui suggère que la mutation l'arginine 52 serait importante pour la liaison de l'ARN substrat. La constante de

dissociation du mutant H68A semble être légèrement différente de celle du sauvage : 2.5 $\mu mol.l^{-1}$ pour le mutant H68A contre 1.5 $\mu mol.l^{-1}$ pour le sauvage. Toutefois, en regardant la Figure 47 on voit que cette différence est non significative car elle se trouve dans l'intervalle des erreurs de mesures.

4.8 Discussion des rôles des histidines 48 et 68

Nous avons vu au cours de ce chapitre que la mutation des histidines H48 et H68 en alanines permet le maintien de la protéine RegB dans la bactérie *E. coli*. De plus, nous avons vu par des tests d'activité *in vitro* sur le substrat ARN modèle Selex22tb que le mutant H48A est complètement inactif alors que le mutant H68A conserve une activité résiduelle. Les expériences de RMN ont montré que ces modifications de l'activité de RegB ne sont pas dus à une modification radicale de la structure de RegB. Les expériences RPS ont montré par ailleurs que la perte d'activité des mutants H48A et H68A ne s'explique pas par une perte d'affinité de l'enzyme pour son substrat.

Ces résultats suggèrent que les histidines H48 et H68 seraient impliquées directement dans la catalyse médiée par RegB mais de deux façon distinctes. A ce stade, plusieurs hypothèses peuvent être proposées. L'histidine 48 pourrait être un résidu aminé crucial impliqué directement dans la catalyse acide-base (accepteur ou donneur de protons). L'histidine 68 aurait un rôle moins crucial puisque la réaction de catalyse peut se faire, plus lentement certes, en son absence. H68 pourrait par exemple polariser certaines liaisons chimiques pour favoriser le transfert de protons ou intervenir en aval de ce transfert en interagissant avec un intermédiaire particulier. Affirmer l'une ou l'autre de ces hypothèses nécessite non seulement d'étudier en détail la structure tridimensionnelle de RegB mais aussi d'étudier les interactions de RegB avec son ARN.

Pendant que l'on préparait le projet conséquent de la résolution de la structure tridimensionnelle de RegB en solution, des travaux très intéressants ont été publiés. Ils concernent l'évolution naturelle du génome d'un virus apparenté au phage T4 : le phage RB49. Le court chapitre qui suit lui sera consacré.

Chapitre 5

Evolution naturelle du gène *regB* dans un virus apparenté au phage T4 : Etude du phage RB49

Chapitre 5

Evolution naturelle du gène *regB* dans un virus apparenté au phage T4 : Etude du phage RB49

Le bactériophage T4 est un virus du genre *Myoviridae*. Son tropisme bactérien, sa morphologie globale ainsi que son architecture fine ont été utilisées comme critères pour construire une famille virale appelée *pseudo-T4*. Environs 170 virus ont été identifiés dans cette famille (Ackermann *et al.*, 1997). Ces virus ont été isolés chez plusieurs hôtes bactériens et dans des conditions de culture très différentes (Ackermann *et al.*, 1998).

Des expériences d'hybridation génomique et des analyses PCR ont montré l'existence au sein de cette famille de deux sous-familles de virus (Monod *et al.*, 1997). La sous-famille des virus *T-pairs* présente une très grande homologie de séquence avec le génome du phage T4, en général supérieure à 90 % (Cowie *et al.*, 1971). Les protéines RegB des virus T2 et T6 de cette sous-famille, par exemple, ont des identités de séquence de 99 % et de 100 % respectivement. En revanche, les virus formant la sous-famille des *pseudo-T-pairs* contiennent des blocs de séquences qui divergent de façon significative du génome du phage T4. Malheureusement très peu de données sont disponibles pour cette sous-famille et seul le gène *regB* du phage RB69 (voir chapitre 4) a été séquencé.

Très récemment, les travaux de Henry Krish et de ses collaborateurs ont permis d'investiguer par séquençage aléatoire près de 20 % de la séquence du génome du virus RB49 qui fait 170 kpb (Desplats *et al.*, 2002). Ces travaux ont permis d'explorer significativement pour la première fois le génome du bactériophage RB49. L'analyse des séquences obtenues montre que ce virus partage avec le phage T4 un ancêtre commun mais semble avoir évolué différemment. En moyenne, l'homologie de séquence en résidus aminés avec le phage T4 est inférieure ou égale à 60 %. Beaucoup des gènes, identifiés comme non-essentiels dans le phage T4, ont disparu dans le phage RB49 et ont été remplacés par des séquences inconnues. Paradoxalement, le gène *regB* qui a été identifié comme non-essentiel pour le phage T4 est maintenu dans le génome du phage RB49 puisque on a retrouvé parmi les fragments séquencés aléatoirement des petits fragments qui correspondent aux parties extrémales 5' et 3' du gène *regB* (Krish H.M., communication personnelle). Le maintien du gène *regB* dans le génome du phage RB49 pourrait s'expliquer par le fait que le qualificatif « non-essentiel » associé au gène *regB* du phage T4 est un qualificatif relatif aux conditions de cultures spéciales en laboratoire qui ne sont pas forcément les conditions naturelles dans lesquelles le virus a évolué. Quoi qu'il en soit, l'étude de la séquence du gène *regB* du phage RB49 et sa comparaison avec celle du phage T4 sera certainement très informative.

Dans ce chapitre, nous nous proposons de séquencer le gène *regB* du phage RB49, d'analyser sa séquence et de la comparer aux résultats de mutagenèses aléatoire et dirigée des chapitres 3 et 4.

5.1 Séquençage du gène *regB* du phage RB49

Nous nous sommes procuré un lysat du phage RB49 ainsi que la souche *E. coli* B^E dans laquelle ce virus se propage (générosité de Krish H.M.). Nous avons choisi d'extraire d'abord le gène *regB* du génome du virus par PCR et de séquencer par la suite ce produit de PCR sans avoir à le sous-cloner dans un vecteur particulier.

La Figure 48 montre la procédure globale utilisée. Pour la réaction de PCR, nous avons utilisé deux amorces qui couvrent la partie 5' du gène *regB* du phage RB49 contenant l'ATG initiateur et la partie 3' contenant le codon stop. Les séquences de ces deux amorces ont été déduites du séquençage des extrémités 5' et 3' du gène *regB*.

Afin de s'affranchir des erreurs éventuelles pouvant survenir lors du processus d'amplification du gène *regB* par PCR, nous avons pris deux précautions. Nous avons d'abord utilisé l'ADN polymérase hautement fidèle *Pfu* qui ne fait en moyenne qu'une erreur toutes les 50 000 bases (le gène *regB* fait environ 500 bases). Nous avons par ailleurs choisi de faire deux PCR indépendantes

et de séquencer leurs produits d'amplification. Nous pensons, en effet, que s'il y a une erreur lors de l'amplification, il est très peu probable que cette même erreur se retrouve dans deux PCR indépendantes. Les produits de PCR purifiés sont séquencés par la méthode des didésoxynucléotides.

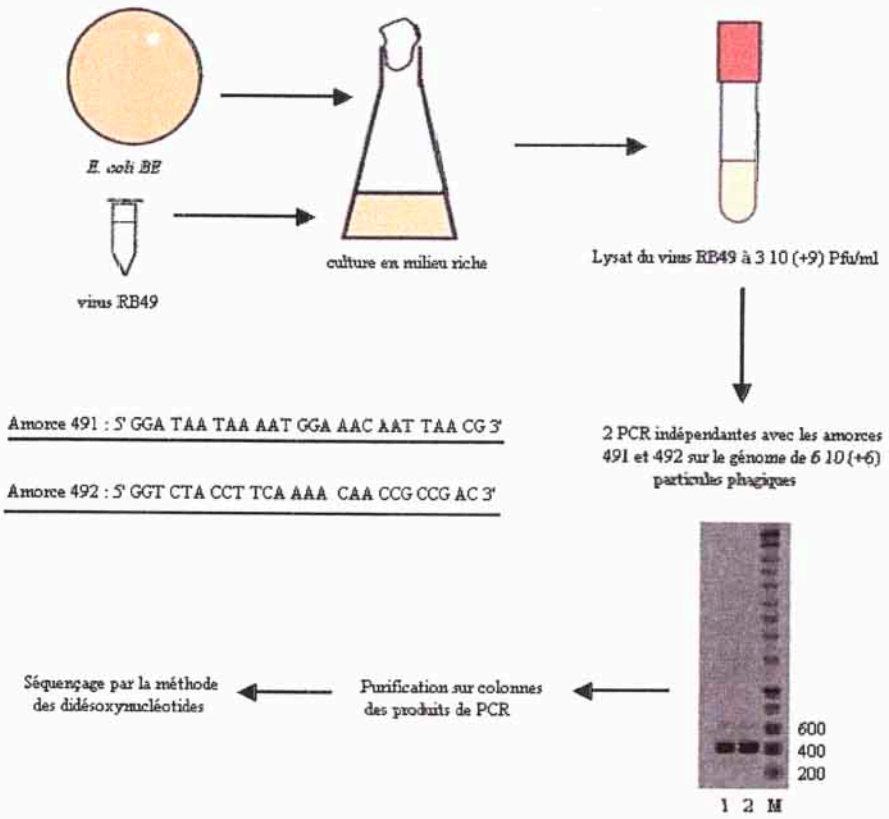


Figure 48 : Séquençage du gène *regB* du phage RB49.

5.2 Comparaison des séquences des protéines RegB du phage T4 et du phage RB49

Le séquençage des deux PCR indépendantes a donné le même résultat. La séquence nucléique de *regB* ainsi obtenue a été ajoutée aux extrémité 5' et 3' déjà disponibles pour reconstituer le gène entier. Ce dernier a été traduit en protéine. Les séquences de la protéine RegB du phage T4 et du phage RB49 sont alignées dans la Figure 49 ci-après.

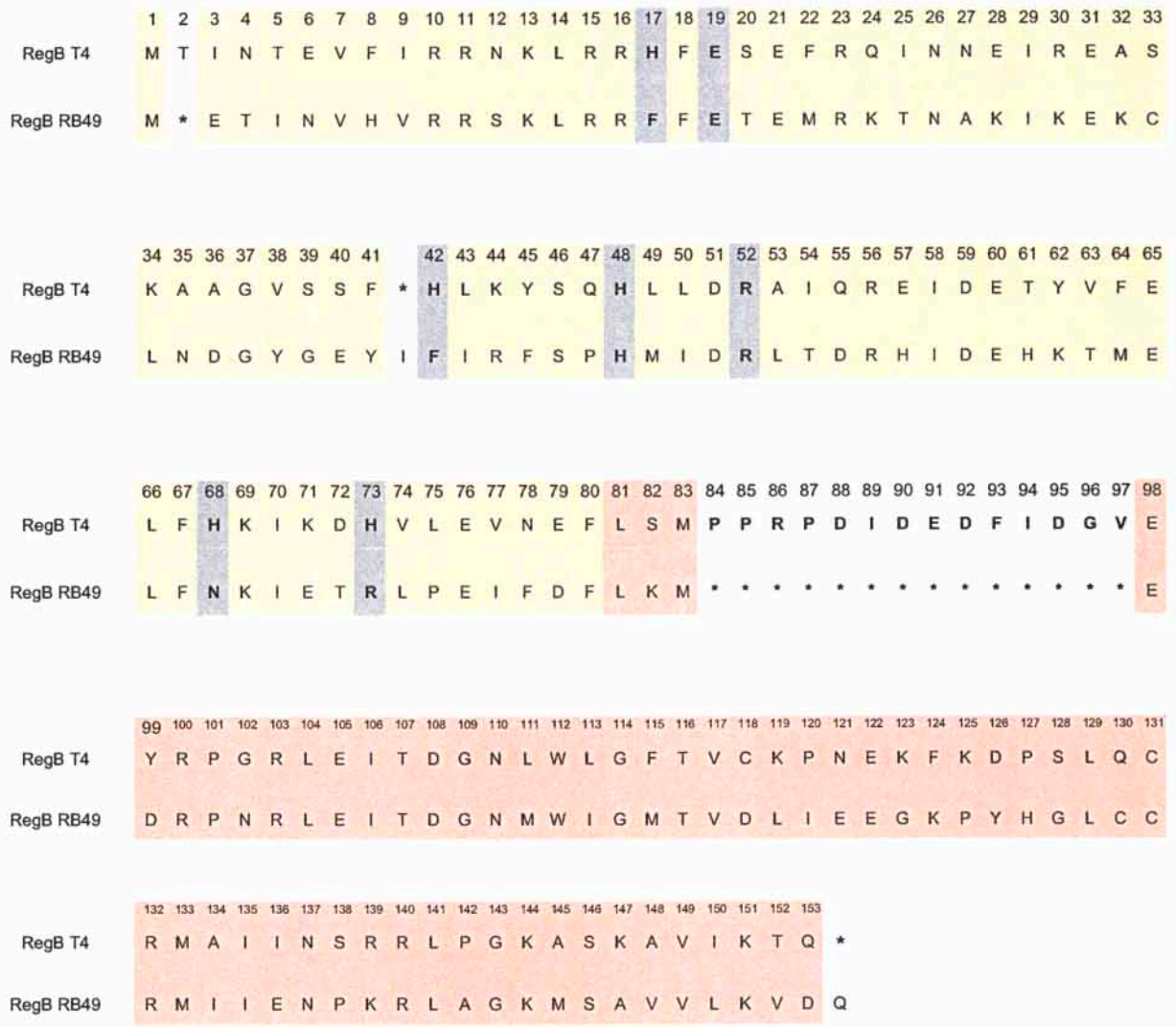


Figure 49 : Alignement des séquences des protéines RegB du phage T4 et du phage RB49.
Jaune : partie N-terminale, orange : partie C-terminale, gris : résidus remarquables, étoile : délétion ou insertion.

La Figure 49 montre que chez le phage RB49 la partie C-terminale de RegB a subi une délétion de 14 résidus aminés (soit près de 20 % de la séquence) et une insertion. La partie N-terminale, en revanche, n'a subi que une seule délétion (1.2 % de la séquence) et une seule insertion. Par ailleurs, on remarque que le pourcentage d'homologie entre le phage T4 et le phage RB49 vaut 60 % dans la partie N-terminale et environ 40 % dans la partie C-terminale. Ces résultats suggèrent que la partie N-terminale de RegB a été sensiblement mieux conservée au cours de l'évolution que la partie C-terminale. Ce résultat, obtenu par mutagenèse naturelle *in vivo*, va dans le même sens que les

résultats de mutagenèse aléatoire du Chapitre 3 qui montrent qu'il est plus probable d'obtenir au hasard des mutants inactifs de la partie C-terminale que de la partie N-terminale.

La Figure 49 montre que les deux histidines H42 et H73 ne sont pas conservées dans la protéine RegB du phage RB49. Ce résultat renforce les conclusions du Chapitre 4 dans lequel on a démontré que la mutation de l'une de ces deux histidines en alanine ne modifie pas l'activité de l'enzyme.

Par ailleurs, nous avons montré dans le chapitre précédent que la position H48 est absolument nécessaire pour la catalyse médiée par RegB. Sa mutation en alanine annule complètement l'activité de la protéine. L'importance de ce résidu est confirmée par la conservation parfaite de ce résidu dans le phage RB49.

Le cas de l'histidine H68 est, en revanche, plus problématique. On sait que la mutation de ce résidu en alanine diminue l'activité de RegB sans l'annuler complètement ce qui suggère que ce résidu a une importance, certes limitée, dans la fonction de RegB. Dans le phage RB49, en revanche, on constate que ce résidu a été substitué par une asparagine, résidu polaire non chargé. Nous ignorons si ce résidu asparagine est capable de jouer le rôle de l'histidine H68 dans la protéine RegB du phage RB49 ou s'il est dépourvu de tout rôle fonctionnel dans cette protéine. Dans le cas de la première hypothèse, il faudrait d'abord étudier l'effet de la mutation ponctuelle H68N sur la protéine RegB du phage T4 ou la mutation N68A sur l'activité de la protéine RegB du phage RB49. Dans le cas de la deuxième hypothèse, le problème devient complexe car il suggère qu'une combinaison particulière de mutations ponctuelles permettrait de s'affranchir du rôle de l'histidine 68. Dans tous les cas, il est clair que l'histidine 68 n'a pas un rôle aussi crucial que celui de l'histidine 48.

Signalons enfin sur la Figure 49 la conservation des résidus Arginine 52 et Aspartique 19. Puisque le mutant ponctuel R52L, qui a été sélectionné *in vivo* par hasard (Ruckman *et al.*, 1989), est complètement inactif, on peut en déduire que l'Arginine 52 joue un rôle important dans l'activité de RegB. Les mesures RPS du Chapitre 4 ont montré que ce mutant a perdu un facteur 10 environ d'affinité pour le substrat ARN. Si on suppose que RegB suit une cinétique Michaelienne, c'est à dire que l'équation de vitesse de la réaction est donnée par l'équation (V), et si on approxime la constante de Michaelis (K_M) par la constante de dissociation K_D, il est aisé de montrer que la perte d'un facteur 10 dans l'affinité pour le substrat ne peut annuler complètement l'activité de l'enzyme puisque le rapport des vitesses initiales (V_i) du mutant R52L et du sauvage (wt) sera donné par l'équation (VI). Dans les conditions expérimentales du paragraphe 4.5 (concentration du substrat S = 3.6 $\mu mol.l^{-1}$, concentration totale de l'enzyme E_T = 0.36 $\mu mol.l^{-1}$) et en prenant 1.5 $\mu mol.l^{-1}$ comme valeur moyenne du K_D, le rapport des vitesses initiales $V_{i\ (R52L)}$ / $Vi_{\ (wt)}$ sera égal à 0.27 c'est à dire que le

mutant R52L conservera 27 % de l'activité de l'enzyme sauvage. Ceci n'est pas vrai expérimentalement. On peut en conclure alors que la mutation R52L doit nécessairement affecter, en plus du K_D, la valeur de la constante de vitesse k_{cat}, ce qui suggère que l'arginine 52 est impliquée dans la catalyse proprement dite.

$$V_i = [k_{cat} \times E_T \times S] / [S + K_M] \quad (V)$$

$$V_{i\,(R52L)} / V_{i\,(wt)} = [S + K_D] / [S + 10 \times K_D] \quad . \quad (VI)$$

La conservation du glutamate 19 chez le phage RB49 coïncide avec le résultat du séquençage du double mutant inactif E19V D72E obtenu par mutagenèse aléatoire (Chapitre 3). Puisque il est fort peu probable que l'annulation de l'activité de ce mutant soit due à la substitution acide-acide D72E et puisque le résidu D72 n'est pas conservé dans le phage RB49 ; on peut raisonnablement suggérer que le glutamate E19 est un résidu indispensable à RegB.

Nous avons résumé dans le Tableau 8 ci-dessous les informations fonctionnelles concernant les 4 résidus aminés remarquables que nous avons étudiés dans les chapitres 3, 4 et 5. Le rôle proposé pour chacun de ces résidus n'est qu'hypothétique et demande à être corrélé avec d'autres données expérimentales notamment structurales afin de construire un modèle du site actif de RegB. A ce titre, les deux chapitres qui suivent seront consacrés à la mise au point de la stratégie d'étude de la structure de RegB par RMN hétéronucléaire.

Résidu	Mutagenèse	Activité	Liaison au substrat	Evolution naturelle (RB49)	Rôle hypothétique
Glutamate 19	E19V D72E	inactif	n.d.	E19 conservé mais D72 non conservé	Catalyseur général acide ou basique
Histidine 48	H48A	inactif	Identique à celle du sauvage	conservée	Catalyseur général acide ou basique
Arginine 52	R52L	inactif	Inférieure à celle du sauvage	conservée	Liaison du substrat et catalyse de la coupure
Histidine 68	H68A	10 à 15 % du sauvage	Identique à celle du sauvage	non conservée	inconnu

Tableau 8 : Résidus remarquables issus de l'étude fonctionnelle de la ribonucléase RegB du phage T4. n.d. : valeur non mesurée.

Chapitre 6

Etude de la structure tridimensionnelle de RegB par RMN :

Mise au point de la stratégie

Chapitre 6

Etude de la structure tridimensionnelle de RegB par RMN :

Mise au point de la stratégie

La compréhension de la base moléculaire de l'interaction entre RegB et son substrat ARN demande non seulement de résoudre la structure tridimensionnelle de la protéine mais aussi d'étudier à l'échelle atomique l'interaction entre RegB et l'ARN. La résonance magnétique nucléaire est, *a priori*, particulièrement adaptée à l'étude de ces deux problèmes. Les acteurs de cette étude ont en effet des tailles moléculaires compatibles avec l'étude RMN. RegB comprend 153 résidus aminés ce qui représente une taille moyenne qui pourrait être investiguée par la RMN hétéronucléaire $^{1}H/^{15}N/^{13}C$. De plus, la molécule d'ARN substrat modèle Selex22tb ne comprenant que 21 nucléotides (environ 6 kDa), il est possible d'étudier son interaction avec RegB sur des cartes bidimensionnelles de type HSQC (voir chapitre précédent) et d'identifier les résidus aminés impliqués dans l'interaction.

Nous présenterons dans ce chapitre le version de RegB que vous avons choisie d'étudier par RMN et les raisons qui ont motivé ce choix. Nous investiguerons ensuite les conditions

spectroscopiques optimales pour étudier cette version et nous discuterons enfin la faisabilité de l'étude structurale de RegB par RMN.

6.1 Choix de la version de RegB

L'un des critères principaux pour étudier la structure d'une protéine par RMN (après la masse moléculaire) est la nécessité de surexprimer cette protéine avec un rendement avoisinant les 20-50 mg par litre de culture. Cette exigence se justifie par le fait que les études RMN requièrent des échantillons protéiques à la fois très concentrés et enrichis aux isotopes ^{13}C et ^{15}N. La faible sensibilité de la RMN comparée aux méthodes spectroscopiques d'absorption (Ultra-Violets par exemple) demande l'utilisation d'échantillons millimolaires. De son coté, l'enrichissement isotopique limite le volumes de culture utilisable à 1 ou 2 litres à cause des prix élevés des métabolites de bases (^{13}C-glucose et $^{15}NH_4Cl$) utilisés comme sources de carbone et d'azote.

Le Tableau 9 présente une estimation des rendements d'expression optimaux des différentes versions de RegB (sauvage et mutants ponctuels) que nous avons étudiées dans les chapitres précédents.

Version de RegB	Plasmide d'expression	Rendement (mg de protéine par litre de culture)
RegB sauvage	pARNU2	0.2
	pOURegB	0.6
	pYRegB2	0.3
RegB H42A	pH42A	0.2
RegB H48A	pH48A	0.2
	pEH48A*	30.0
RegB R52L	pERL52	20.0
RegB H68A	pH68A	0.2
	pEH68A*	6.0
RegB H73A	pH73A	0.2

Tableau 9 : Rendements d'expression des différentes versions de RegB. * : Les plasmides pEH48A et pEH68A sont les deux vecteurs d'expression respectifs des mutants RegB H48A et RegB H68A issus du sous clonage de ces mutants à partir des vecteurs à faible expression pH48A et pH68A (voir chapitre 1).

Il ressort de ce tableau que seuls les mutants RegB H48A et RegB R52L répondent aux exigences de la RMN en terme de rendement d'expression. Nous avons choisi d'étudier le mutant RegB H48A pour deux raisons. D'abord, contrairement au mutant R52L qui a perdu un facteur 10 dans son affinité pour le substrat ARN Selex22, le mutant RegB H48A a exactement la même affinité pour l'ARN que l'enzyme sauvage. Par ailleurs, sur la base d'une comparaison avec les ribonucléases connues produisant un phosphodiester cyclique, RegB H48A est probablement « un mutant catalytique pur » c'est à dire un mutant ne différent de RegB que par la substitution d'une chaîne latérale impliquée dans la catalyse de la rupture de la liaison phosphodiester. Il constitue donc le candidat de choix pour construire un modèle du site actif de RegB.

Nous avons par ailleurs voulu comparer les spectres 1D proton des versions RegB H48A, RegB R52L et RegB H68A avec le spectre 1D proton de la protéine RegB sauvage. Il est en effet possible grâce à la grande sensibilité des noyaux hydrogène (noyaux ayant le plus grand rapport gyromagnétique) d'enregistrer des spectres 1D ^{1}H sur des échantillons protéiques relativement peu concentrés, de l'ordre de 50 $\mu mol.l^{-1}$. De plus puisque ce noyau a une abondance naturelle voisine de 100 %, aucun enrichissement isotopique n'est nécessaire; les échantillons peuvent être donc préparés en milieu riche. Ces deux caractéristiques conviennent parfaitement à la préparation d'un échantillon de la protéine RegB sauvage moyennant le recours, comme on l'a vu au chapitre 1, à la technique de fermentation.

Nous avons préparé à l'aide d'un fermenteur de 15 L un échantillon RMN de la protéine RegB sauvage ayant une concentration de 40 $\mu mol.l^{-1}$. Nous avons enregistré un spectre 1D ^{1}H sur cet échantillon et l'avons comparé aux spectres 1D ^{1}H des trois autres versions de RegB qui se surproduisent avec les méthodes d'expression classiques (H48A, R52L et H68A). Les régions des protons amides des quatre spectres sont présentées dans la Figure 50. Cette figure montre qu'à l'exception des pics fins et intenses qui se trouvent dans la région 7.1-7.5 ppm provenant de tampons phosphates ayant des pH légèrement différents, l'allure globale du spectre 1D proton est conservée pour les quatre versions de RegB ce qui indique que le repliement global des 4 protéines est semblable. En comparant les détails des spectres, on remarque que les spectres du H48A, H68A et R52L sont très voisins mais présentent un certain nombre de pics qui ne résonnent pas à la même fréquence que la protéine sauvage. Même si la qualité du spectre 1D de la protéine sauvage reste discutable (notamment de point de vue de pureté de l'échantillon), les différences observées suggèrent l'existence de variations dans les positions absolues de quelques résidus aminés dans la protéine sauvage d'une part et dans les trois mutants ponctuels d'autre part. Toutefois, dans le cas de RegB H48A et de RegB H68A, ces variations ne semblent pas interférer avec la capacité de l'enzyme à lier le substrat (ces mutants ont la même constante de dissociation pour l'ARN Selex22tb que la

protéine sauvage) ce qui suggère que la conformation du site de liaison du substrat est intact dans ces mutants.

En conjuguant l'ensemble des arguments présentés ci-dessus, on s'aperçoit que la version H48A représente le meilleur candidat pour faire une étude structurale de la ribonucléase RegB.

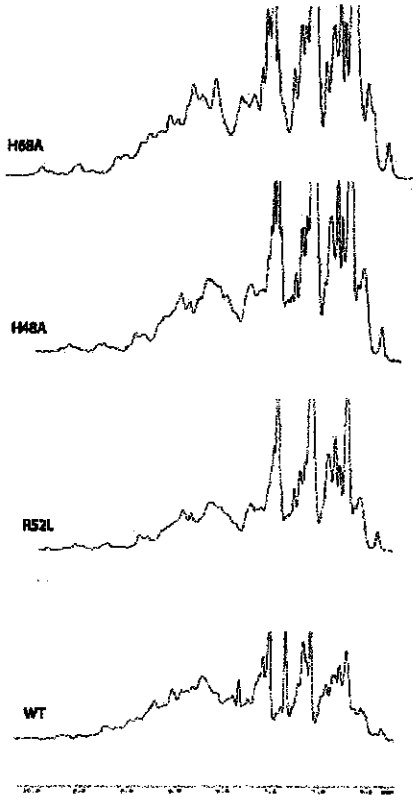


Figure 50 : Régions des protons amides des spectres 1D ^{1}H de RegB sauvage (WT), RegB R52L, RegB H48A et RegB H68A. Les spectres ont été enregistrés à 25 °C, dans un tampon phosphate de sodium à pH 8.0.

6.2 Détermination des conditions spectroscopiques d'étude de RegB H48A

6.2.1 La température

La température influence la largeur des pics RMN donc la résolution globale du spectre. Afin de déterminer la température optimale pour étudier RegB H48A par RMN, nous avons réalisé

une série de spectres 1D proton à différentes températures. En dessous de 15 °C, les spectres sont de très mauvaise qualité. Au dessus de 40 °C, la protéine commence à se dénaturer. Dans l'intervalle 15 °C à 40 °C, nous avons remarqué que le meilleur spectre est obtenu à 32.5 °C. A cette température, les pics sont très fins et laissent apparaître des pics irrésolus à des températures plus basses. Il n'y pas de différence notable entre 32.5 °C et 35 °C mais au dessus de cette température l'effet de la dénaturation thermique devient de plus en important et finit par l'emporter sur la finesse des pics. Nous avons donc choisi de faire les expériences RMN à 32 °C.

Malheureusement, nous avons constaté qu'à cette température un échantillon millimolaire de RegB H48A commence à précipiter au bout de 18 heures. Sachant que les expériences RMN tridimensionnelles s'étalent en général sur plusieurs jours, il est impératif de trouver des conditions de tampon, de force ionique et de pH pour lesquelles l'échantillon de RegB reste stable à 32.5 °C pendant une longue période.

6.2.2 pH, force ionique et molécule tampon

Nous avons mis au point une procédure expérimentale qui permet d'investiguer rapidement plusieurs valeurs, du pH, de la force ionique et de la molécule tampon. La méthode consiste à préparer un échantillon concentré de RegB H48A (2 ml environs d'une solution millimolaire peuvent être obtenus à partir d'un litre de culture en milieu riche) dans un tampon phosphate standard puis de diviser cet échantillon en plusieurs micro-fractions de 50 µl. En utilisant le système des micro-colonnes de chromatographie (*Micro Bio-Spin®, Biorad*) on peut à l'aide de centrifugations courtes débarrasser rapidement la protéine de son tampon originel et l'éluer dans la solution tampon souhaitée.

Nous avons pu ainsi préparer plusieurs micro-fractions avec des pH allant de 4.0 à 9.0. Les pH très acides ont été préparés en utilisant le tampon acétate de sodium. Les autres pH ont été préparés dans le tampon phosphate de sodium. Nous avons criblé aussi plusieurs valeurs de la force ionique en utilisant des concentrations de NaCl allant de 50 mmol.l^{-1} à 600 mmol.l^{-1}. Pour chaque condition, la micro-fraction de 50 µl a été déposée dans un tube de 200 µl à paroi fine (normalement utilisé pour les réactions de PCR), puis incubé pendant trois jours dans un appareil de PCR qui sert de bloc thermique maintenu à 32.5 °C (empêchant de plus, grâce à son couvercle chauffant, les condensations indésirables sur les parois du tube susceptibles de concentrer l'échantillon protéique). La stabilité de l'échantillon protéique a été suivie en examinant le profil de précipitation au fond du tube. Le profil de précipitation minimal est obtenu à pH 6.0 et à une concentration de NaCl de 300

mmol.l^{-1}. Finalement, nous avons examiné l'influence de la molécule de tampon en comparant les stabilités thermiques de RegB H48A en présence de 50 mmol.l^{-1} de phosphate de sodium ou de 50 mmol.l^{-1} de citrate de sodium (tampon couvrant la zone de pH 6.0) à la même concentration et au même pH. Ce choix est motivé par le fait que nous soupçonnons l'existence d'un effet négatif des ions phosphates sur la stabilité de RegB à haute concentration. RegB est en effet une ribonucléase qui interagit avec certains composés phosphatés, en particulier les phosphodiesters. Après 3 jours d'incubation à 32.5 °C, nous avons remarqué que contrairement au tube contenant le tampon phosphate de sodium, le tube contenant un tampon citrate de sodium ne présentait aucun trace de précipitation.

En conclusion, nous avons choisi d'enregistrer tous les spectres RMN de RegB H48A dans les conditions suivantes : *une température de 32.5 °C et un tampon contenant 50 mmol.l^{-1} de citrate de sodium (pH 6.0), 300 mmol.l^{-1} de NaCl et 2 mmol.l^{-1} de dithiothreitol* (sans cet agent réducteur, RegB précipite à toutes les températures et dans tous les tampons à cause de l'oxydation de deux cystéines libres).

Pour une protéine de la taille de RegB, des expériences RMN hétéronucléaires $^{1}H/^{15}N/^{13}C$ sont nécessaires. Nous avons donc décidé de préparer un échantillon de RegB H48A enrichis aux isotopes ^{15}N et ^{13}C.

6.3 Double marquage $^{13}C/^{15}N$ de RegB H48A

La souche bactérienne BL21(λDE3)-pLysS transformée avec le plasmide pEH48A a été cultivée et induite dans un litre de milieu minimum contenant 1.0 g de $^{15}NH_4Cl$ et 2.0 g de ^{13}C-glucose (voir partie matériels et méthodes). Afin d'éviter tout contact avec les ions phosphates, la protéine RegB H48A été purifiée en utilisant d'abord un tampon Tris-HCl qui a été échangé finalement contre le tampon RMN décrit ci-dessus. Il n'est pas possible, en effet, d'utiliser dès le début de la purification un tampon citrate parce que ce dernier ne tamponne que dans des zones de pH acides inférieurs à 6.4. Dans ces zones acides, la chromatographie d'affinité utilisant le cation cobalt est très perturbée parce que les groupes imidazoles de la queue polyhistidines se trouvant chargés positivement sont fortement repoussés à l'approche du cation Co^{2+} sur lesquels ils sont censés se fixer.

La Figure 51 montre le résultat de l'analysé SDS-PAGE de la purification de la protéine RegB H48A doublement marquée $^{13}C/^{15}N$. Le gel montre que la pureté de la protéine est supérieure à 95 %. Il n'est donc pas nécessaire, pour les expériences RMN, de faire appel à une deuxième étape de purification.

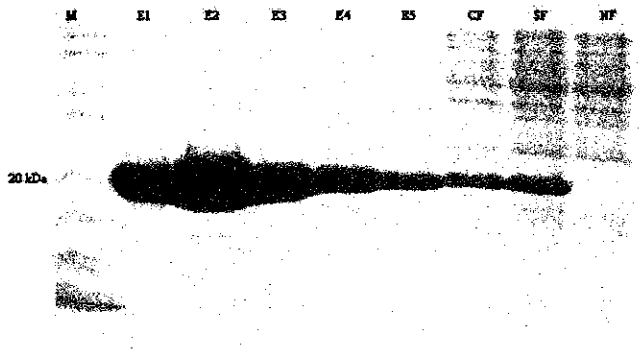


Figure 51 : Analyse SDS-PAGE de la purification de la protéine RegB H48A doublement marquées $^{13}C/^{15}N$. M : marqueur de masses moléculaires. E1 à E5 : fractions d'élution, CF : extrait brut après sonication, SF : fraction soluble, NF : fraction protéique non liée à la résine de cobalt. La masse molaire calculée de la protéine RegB H48A portant une queue polyhistidines vaut 20.073 kDa.

6.4 Spectres RMN préliminaires

Un échantillon millimolaire de la protéine RegB H48A doublement marquée, contenant 10 % d'oxyde de deutérium (D_20) pour asservir le signal RMN et 0.1 % d'azidure de sodium (NaN_3) pour empêcher la croissance bactérienne, a été utilisé pour enregistrer un spectre bidimensionnel $^1H/^{15}N$ HSQC sur un spectromètre 600 MHz. Un spectre $^1H/^{13}C$ HSQC a été enregistré dans les mêmes conditions à l'exception de l'échange de toutes les molécules d'eau (H_2O) par des molécules d'eau lourde (D_2O) afin de supprimer la résonance aqueuse à 4.8 ppm et ainsi augmenter la qualité du spectre dans cette région.

La Figure 52 présente le spectre $^1H/^{15}N$ HSQC obtenu. Le spectre montre une bonne dispersion des résonances des protons amides qui s'étalent sur environ 3.5 ppm (de 6.5 à 10). Le spectre contient par ailleurs environ 140 taches correspondant à des couples $^1H/^{15}N$ du squelette de la protéine. Ce nombre est en accord avec celui attendu : 145 (RegB contient 153 résidus aminés dont 8 prolines dépourvues de couples $^1H/^{15}N$).

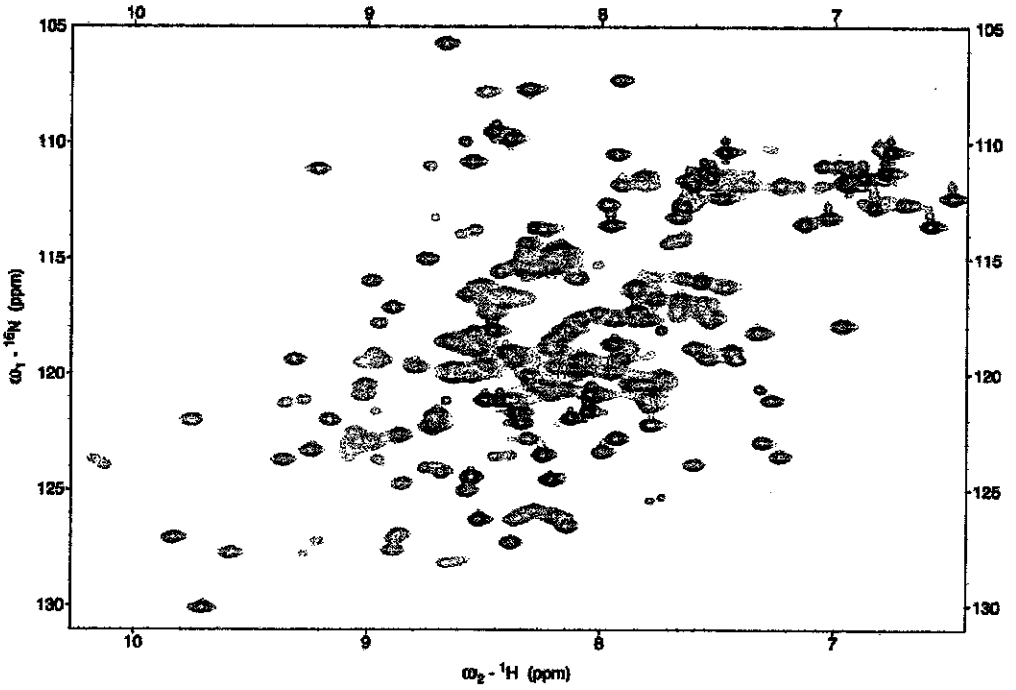


Figure 52 : Spectre bidimensionnel ^{1}H/^{15}N HSQC de la protéine RegB H48A. Le spectre a été enregistré à 32.5 °C sur un spectromètre Bruker DRX600.

La Figure 53 ci-dessous présente l'aspect global du spectre ^{1}H/^{13}C HSQC obtenu.

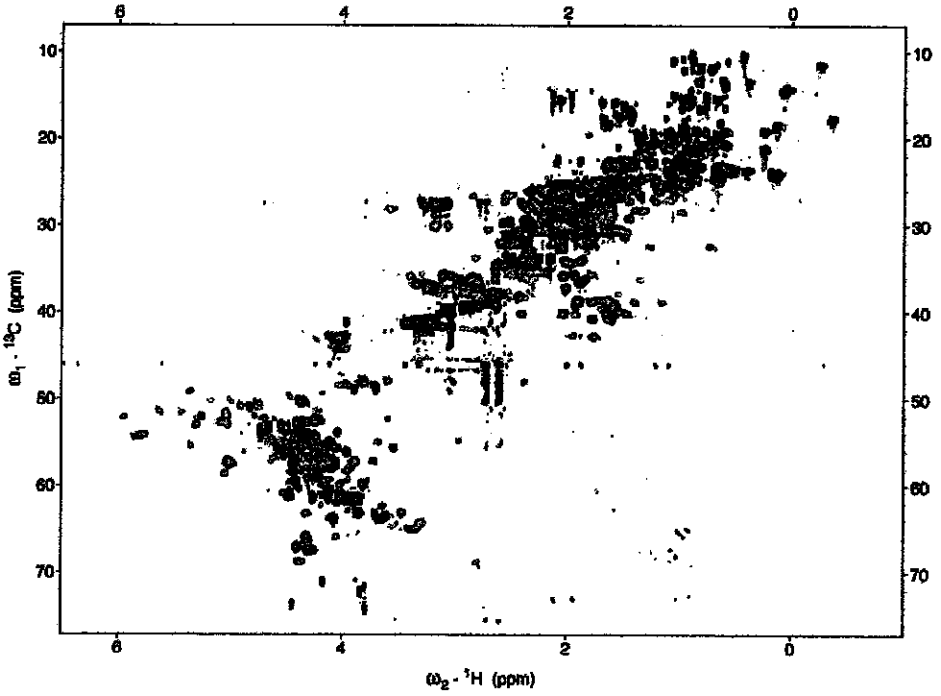


Figure 53 : Spectre bidimensionnel ^{1}H/^{13}C HSQC de la protéine RegB H48A. Le spectre a été enregistré à 32.5 °C sur un spectromètre Bruker DRX600.

La Figure 54 présente deux agrandissements du spectre de la Figure 53. La Figure 54B montre les pics $^{1}H/^{13}C$ correspondant aux groupes méthyles (CH_3) pour lesquels les protons résonnent entre -1 et 2.5 ppm. Ces pics intenses et bien résolus sont très utiles pour la construction d'un modèle de la structure tridimensionnelle de RegB. La Figure 54A montre les pics qui résonnent près du signal de l'eau en dimension proton (4.8 ppm). Des pics peu nombreux mais bien visibles présentent des déplacements chimiques supérieurs à 4.8 ppm. Ceci suggère que le contenu en feuillets β de la protéine RegB H48A est certes faible mais non nul.

Globalement, les spectres $^{1}H/^{15}N$ et $^{1}H/^{13}C$ HSQC de RegB H48A montrent que la protéine est bien structurée. Par ailleurs, la qualité de ces spectres semble à priori suffisante pour entreprendre l'étude de sa structure par RMN.

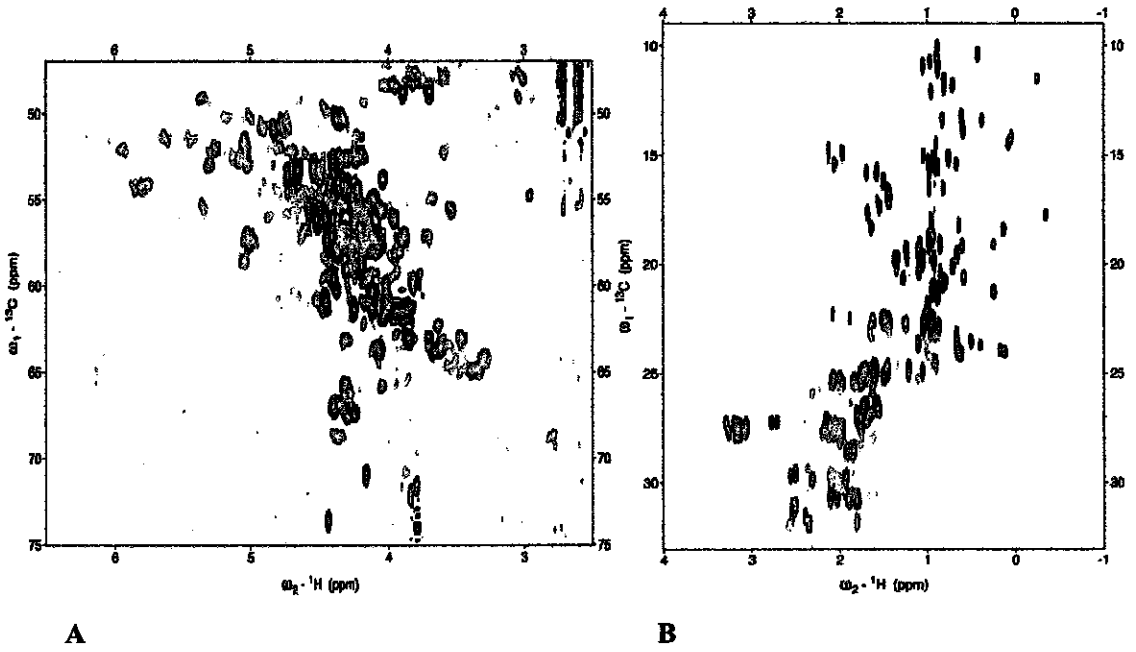


Figure 54 : Détails du spectre bidimensionnel $^{1}H/^{13}C$ HSQC de la protéine RegB H48A. (A) détail montrant les pics qui résonnent au voisinage de 4.8 ppm en dimension proton. (B) région des méthyles.

Chapitre 7

Etude de la structure tridimensionnelle de RegB par RMN :

Début de l'attribution séquentielle de la protéine

Chapitre 7

Etude de la structure tridimensionnelle de RegB par RMN :

Début de l'attribution séquentielle de la protéine

La RMN permet de construire un modèle tridimensionnel de la structure d'une protéine en utilisant un jeu de contraintes sur certaines distances inter-protons (celles inférieures à 5°) et sur certains angles dièdres caractéristiques (voir Figure 55).

Ces contraintes ne peuvent être localisées à l'échelle du résidu aminé que si on attribue une grande partie des atomes de la protéine. On appelle attribution le fait d'affecter à chaque atome une fréquence de résonance particulière. Cette attribution se fait en deux étapes. On réalise dans un premier temps une « attribution séquentielle » c'est à dire on identifie chacun des atomes du squelette de la protéine (proton amide : HN, azote amide : N, carbone α : CA, proton α : HA et carbone carbonyle CO) par des fréquences de résonances particulières (déplacements chimiques particuliers). On utilise ensuite les éléments de cette attribution séquentielle comme points d'ancrage pour « remonter » le long de la chaîne de chaque résidu aminé : c'est l'attribution des chaînes latérales.

Ce chapitre présente et discute la stratégie que nous avons mise au point pour faire l'attribution séquentielle de RegB. Le travail est illustré par quelques exemples de résidus aminés de RegB H48A pour lesquels on a identifié les résonances des noyaux du squelette avec certitude.

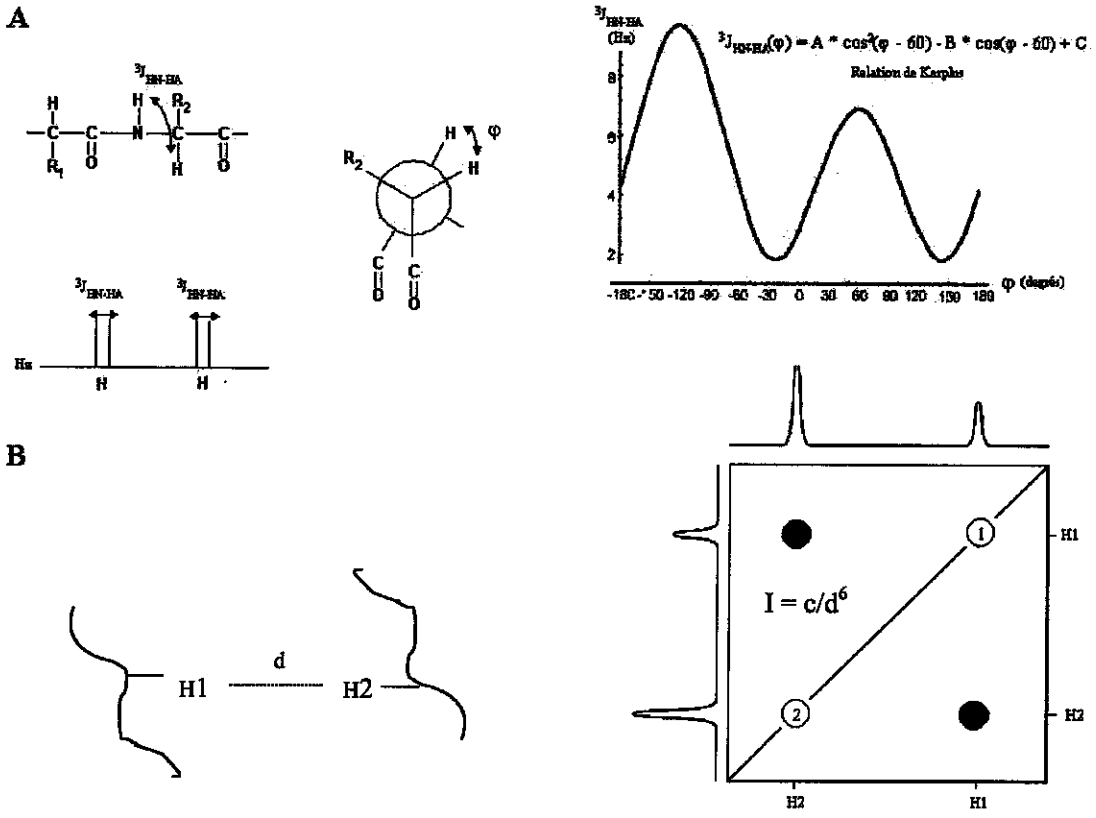


Figure 55 : Les contraintes structurales extraites des données RMN. (A) Des contraintes d'angle de torsion sont extraites des valeurs des constantes de couplage. L'exemple est celui de l'angle dièdre φ qui existe entre le proton amide HN et le proton aliphatique HA du squelette peptidique. La valeur de la constante de couplage entre ces deux protons ($^3J_{HN-HA}$ en Hz) est liée, par la relation empirique de Karplus, à la valeur de l'angle dièdre φ très important dans les définition de la structure secondaire d'une chaîne polypeptidique. La valeur de $^3J_{HN-HA}$ est déterminée en réalisant des expériences RMN dans lesquelles a lieu un échange d'information magnétique à travers les électrons des liaisons covalentes (= couplage scalaire). (B) Des contraintes de distance entre protons séparés d'au plus 5 A° sont extraites des expériences utilisant l'effet Overhauser Nucléaire. Cet effet se traduit par un échange d'information magnétique à travers l'espace (= couplage dipolaire). L'intensité de cet échange I est liée à la distance d qui sépare les deux protons par la relation $I = c/d^6$ (c est une constante).

7.1 La procédure globale

La protéine RegB H48A comprend 173 résidus aminés (en comptant les résidus de la queue polyhistidines). L'attribution d'une protéine de cette taille passe par l'enregistrement d'expériences RMN à trois dimensions dont le but est de lever la dégénérescence des signaux RMN en ajoutant d'abord une dimension ^{15}N puis une dimension ^{13}C.

Les expériences RMN hétéronucléaires ($^{1}H/^{15}N/^{13}C$) que nous avons choisies d'utiliser reposent sur des transferts d'aimantation entre les atomes HN, N, CA, CO et CB (Sattler *et al.*, 1999). Il est possible d'identifier à partir de ces expériences des enchaînements séquentiels puis de projeter ces enchaînements sur la structure primaire de la protéine en utilisant le programme **seq_prob** développé par Grzesiek et Bax (Grzesiek *et al.*, 1993).

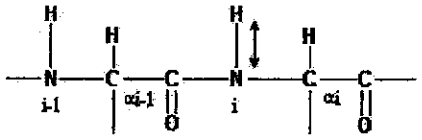
7.2 Les expériences RMN utilisées

La Figure 56 ci-après présente les expériences hétéronucléaires que nous avons enregistrées sur un échantillon doublement marqué de RegB H48A en utilisant les conditions spectroscopiques présentées dans le chapitre précédent.

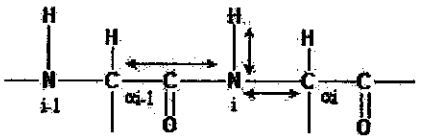
L'expérience de base est l'expérience bidimensionnelle $^{1}H/^{15}N$ HSQC. Cette dernière permet d'identifier pour chaque tache du spectre une paire de déplacements chimiques (dans la dimension proton amide (HN) et la dimension azote (N)).

L'expérience HNCA permet de corréler chacune des paires (HN, N) au déplacement chimique du carbone α du résidu (CA_i) mais aussi au déplacement chimique du carbone α du résidu qui le précède séquentiellement (CA_{i-1}). Ce type de corrélation est très important car il permet de trouver les paires (HN, N) qui correspondent à deux résidus aminés successifs dans la structure primaire de RegB. Le déplacement chimique CA_i d'une première paire (HN, N) coïncidera alors avec le déplacement chimique CA_{i-1} de la paire suivante. Dans le spectre tridimensionnel HNCA, la verticale de chaque tache HSQC contient deux taches l'une correspondant au CA_{i-1} (peu intense) et l'autre correspondant au CA_i (plus intense car le trajet magnétique emprunté est plus court). Cette différence d'intensité sert de critère pour distinguer le carbone α intra-résiduel du carbone α inter-résiduel. Toutefois, pour certaines paires (HN, N), cette différence d'intensité est trop faible pour être significative. On a recourt à une expérience supplémentaire qui sélectionne uniquement le signal de résonance du CA_{i-1} : L'expérience HN(CO)CA. Dans cette expérience, le transfert d'aimantation vers

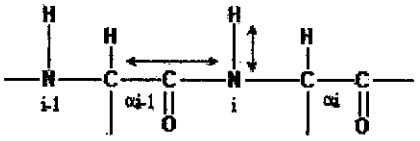
les carbones est unidirectionnel car il passe obligatoirement par le carbone du groupe carbonyle (CO). La superposition des verticales HNCA et HN(CO)CA correspondant à une paire (HN, N) permet alors de d'identifier formellement le CA_{i-1} et le CA_i.



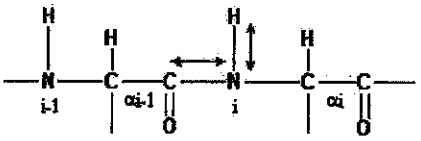
Expérience 2D $^{1}H/^{15}N$ HSQC



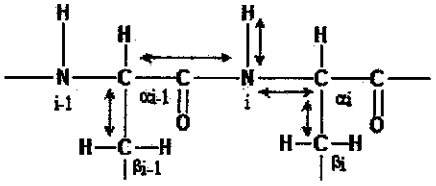
Expérience 3D $^{1}H/^{15}N/^{13}C$ HNCA



Expérience 3D $^{1}H/^{15}N/^{13}C$ HN(CO)CA



Expérience 3D $^{1}H/^{15}N/^{13}C$ HNCO



Expérience 3D $^{1}H/^{15}N/^{13}C$ HNCACB

Figure 56 : Expériences hétéro-nucléaires utilisées pour l'attribution séquentielle de RegB H48A. Les flèches indiquent les trajets magnétiques empruntés.

L'expérience HNCACB emprunte un trajet magnétique comparable à celui de l'expérience HNCA mais contient un transfert supplémentaire entre le carbone α et le carbone β. Dans le spectre

HNCACB, la verticale d'une tache HSQC contient en plus des signaux CA_{i-1} et CA_i une paire correspondant aux signaux CB_{i-1} et CB_i. Au cours du transfert magnétique dans l'expérience HNCACB, l'aimantation est manipulée de façon à ce que les paires CA correspondent à des pics positifs en dimension carbone et les paires CB à des pics négatifs. Cette expérience est très importante car elle permet de lever la dégénérescence de certaines résonances puisque chaque résidu aminé sera désormais défini non plus par 4 coordonnées de déplacements chimiques (δ) (δHN, δN, δCA_i, δCA_{i-1}) mais par 6 coordonnées (δHN, δN, δCA_i, δCA_{i-1}, δCB_i, δCB_{i-1}). La probabilité que deux résidus de RegB aient 6 coordonnées de déplacements chimiques confondues est évidemment beaucoup plus faible que d'avoir seulement 4.

Finalement, l'expérience HNCO permet d'ajouter une coordonnée supplémentaire correspondant au déplacement chimique du carbone du groupe carbonyle (voir Figure 56).

7.3 Recherche des enchaînements séquentiels

7.3.1 Principe

Les expériences RMN décrites dans le paragraphe précédent permettent de construire une matrice de déplacements chimiques δ contenant n lignes (où n est le nombre de taches HSQC correspondant à des paires (HN, N) du squelette) et 8 colonnes. Une ligne correspondant à la tache désignée arbitrairement 'i' sera du type $(i, \delta_{i1}, \delta_{i2}, \delta_{i3}, \delta_{i4}, \delta_{i5}, \delta_{i6}, \delta_{i7})$ où δ_{i1}, δ_{i2}, δ_{i3}, δ_{i4}, δ_{i5}, δ_{i6}, δ_{i7} désignent respectivement les déplacements chimiques δHN, δN, δCA_i, δCA_{i-1}, δCB_i, δCB_{i-1} et δCO (voir Tableau 10).

Trouver deux taches HSQC qui s'enchaînent dans la séquence de RegB, c'est trouver la paire (i, j) écrite dans cet ordre qui vérifie les inéquations $|\delta_{i3} - \delta_{j4}| \leq r$ et $|\delta_{i5} - \delta_{j6}| \leq r$ où r est la résolution en ppm des spectres RMN dans la dimension carbone (c'est à dire le seuil de détection en ppm à partir duquel deux pics dans la dimension carbone sont définis comme distincts, cette valeur est en général voisine de 0.1 ppm). Ces inéquations traduisent simplement le fait que pour que «i » précède « j » dans la séquence de RegB, il faut que le δCA_i coïncide avec le δCA_{j-1} et que le δCB_i coïncide avec le δCB_{j-1}.

n	δ_{HN}	δ_N	δ_{CAi}	δ_{CAi-1}	δ_{CBi}	δ_{CBi-1}	δ_{CO}
1	δ_{11}	δ_{12}	δ_{13}	δ_{14}	δ_{15}	δ_{16}	δ_{17}
.	.	.	.	.	.	.	.
i	δ_{i1}	δ_{i2}	$\boldsymbol{\delta_{i3}}$	δ_{i4}	$\boldsymbol{\delta_{i5}}$	δ_{i6}	δ_{i7}
.	.	.	. ↘	.	. ↘	.	.
j	δ_{j1}	δ_{j2}	δ_{j3}	$\boldsymbol{\delta_{j4}}$	δ_{j5}	$\boldsymbol{\delta_{j6}}$	δ_{j7}
.	.	.	.	.	.	.	.
n	δ_{n1}	δ_{n2}	δ_{n3}	δ_{n4}	δ_{n5}	δ_{n6}	δ_{n7}

Tableau 10 : Matrice des déplacements chimiques. « i » et « j » s'enchaînement dans cet ordre dans la séquence de RegB si $|\delta_{i3} - \delta_{j4}|$ et $|\delta_{i5} - \delta_{j6}|$ sont inférieures ou égales à la résolution spectrale « r » dans la dimension carbone.

Il est possible alors d'écrire un algorithme qui utilise comme entrée cette matrice de déplacements chimiques et qui calcule les paires de résidus qui s'enchaînement séquentiellement.

7.3.2 Exemple d'enchaînement

Nous avons numéroté arbitrairement les taches HSQC de RegB H48A, et avons choisi comme tache de départ la tache portant le numéro 3. L'application de l'algorithme précédent permet de construire sans ambiguïté l'enchaînement 3/113/14/16. Le Tableau 11 présente les déplacements chimiques trouvés.

n	δ_{HN}	δ_N	δ_{CAi}	δ_{CAi-1}	δ_{CBi}	δ_{CBi-1}	δ_{CO}
3	8.320	107.6	45.5 ↘	63.8	0.0 ↘	32.5	175.3
113	8.117	119.7	56.7 ↘	45.5	33.5 ↘	0.0	172.1
14	8.233	123.1	52.8 ↘	56.8	19.3 ↘	33.4	174.5
16	8.312	115.2	58.3	52.9	64.4	19.3	175.8

Tableau 11 : Matrice de déplacements chimiques montrant l'enchaînement 3/113/14/16. La valeur 0.0 indique que le déplacement chimique désigné n'est pas défini.

7.4 Projection d'un enchaînement séquentiel sur la structure primaire de RegB H48A

Pour projeter un enchaînement de taches HSQC sur la séquence de RegB H48A nous avons utilisé le programme **seq_prob** écrit par Stephan Grzesiek (Grzesiek *et al.*, 1993). Ce programme utilise les déplacements chimiques CA et CB pour calculer d'abord la probabilité pour qu'une tache HSQC corresponde à un type donné d'acide aminé puis la probabilité globale pour que l'enchaînement étudié corresponde à un enchaînement dans la séquence de la protéine. Ces calculs de probabilités sont basés sur l'étude de la distribution des déplacements chimiques des carbones CA et CB par rapport aux valeurs des déplacements chimiques en conformation « pelote statistique » (Figure 57).

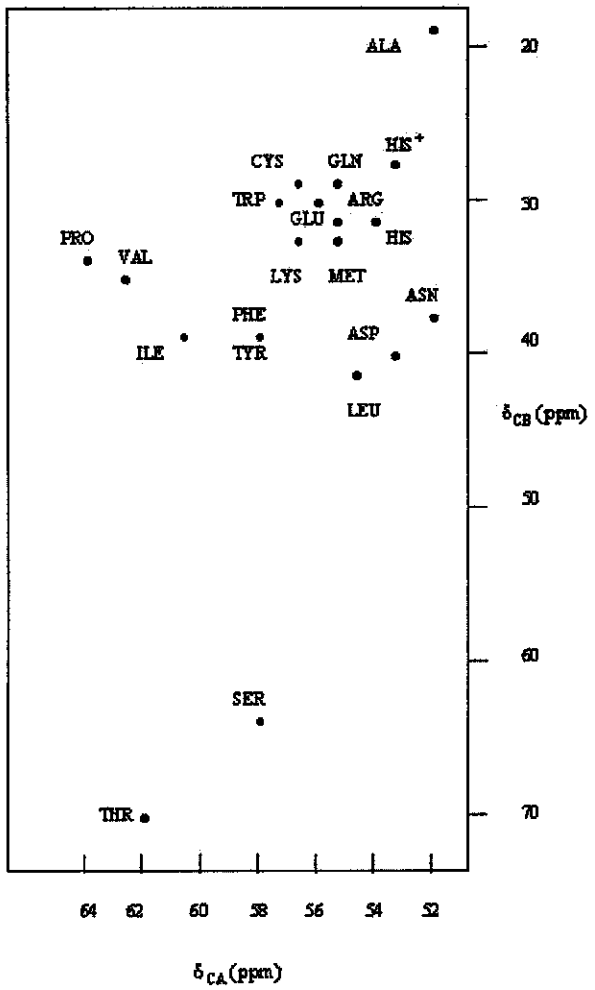


Figure 57 : Déplacements chimiques moyens CA et CB des 19 acides aminés en conformation « pelote statistique ». D'après Grzesiek *et al.*, 1993.

L'application du programme **seq_prob** à l'enchaînement du paragraphe précédent (3-113-14-16) permet de calculer les probabilités suivantes à partir des valeurs des déplacements chimiques des carbones CA et CB :

```
Label    ca     cb      ALA   ARG   ASN   ASP   CYS   GLN   GLU   GLY   HIS
3-1      63.76  32.52   0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0
3        45.53   0.00   0.1   0.0   0.0   0.0   0.0   0.0   0.0  99.9   0.0
14       52.85  19.32 100.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0
16       58.34  64.38   0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0

Label    ca     cb      HIH   ILE   LEU   LYS   MET   PHE   PRO   SER   THR
3-1      63.76  32.52   0.0   0.0   0.0   0.0   0.0   0.0  51.4   0.0   0.0
3        45.53   0.00   0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0
113      56.74  33.46   0.0   0.0   0.0  35.2  23.8   0.0   0.0   0.0   0.0
14       52.85  19.32   0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0
16       58.34  64.38   0.0   0.0   0.0   0.0   0.0   0.0   0.0 100.0   0.0

Label    ca     cb      TRP   TYR   VAL
3-1      63.76  32.52   0.1   0.0  48.5
3        45.53   0.00   0.0   0.0   0.0
113      56.74  33.46  19.5   0.0   0.2
14       52.85  19.32   0.0   0.0   0.0
16       58.34  64.38   0.0   0.0   0.0

prob.: 1.000e+02: PRO142: GLY143: LYS144: ALA145: SER146
prob.: 3.198e-19: ILE29: ARG30: GLU31: ALA32: SER33
prob.: 3.015e-29: HIS-13: HIS-12: HIS-11: HIS-10: SER-9
prob.: 4.486e-36: ARG16: HIS17: PHE18: GLU19: SER20
prob.: 0.000e+00: GLY96: VAL97: GLU98: TYR99: ARG100
```

Ce calcul permet d'attribuer avec certitude les taches 3, 113, 14 et 16 respectivement aux résidus aminés G143, K144, A145 et S146 (Tableau 12). La proline P142 est alors caractérisée par un δ_{CA} = 63.8 ppm et un δ_{CB} = 32.5 ppm. La matrice des déplacements chimiques devient donc celle du Tableau 12 ci-dessous.

n	δ_{HN}	δ_{N}	δ_{CAi}	δ_{CAi-1}	δ_{CBi}	δ_{CBi-1}	δ_{CO}
G143	8.320	107.6	45.5	63.8	0.0	32.5	175.3
K144	8.117	119.7	56.7	45.5	33.5	0.0	172.1
A145	8.233	123.1	52.8	56.8	19.3	33.4	174.5
S146	8.312	115.2	58.3	52.9	64.4	19.3	175.8

Tableau 12 : La projection de l'enchaînement 3/113/14/16 sur la séquence de RegB H48A permet d'attribuer les résidus 142 à 146.

7.5 Limites de la stratégie d'attribution choisie

Plusieurs fragments de RegB H48A ayant des tailles allant de 4 à 8 résidus on pu être attribués avec la stratégie que nous venons d'exposer. Toutefois, plusieurs cas restent très ambigus à cause des problèmes qui suivent. D'abord, on a constaté qu'un nombre non négligeable de taches HSQC présentent dans les expériences 3D des signaux inter-résiduels si faibles qu'il est impossible de les distinguer du bruit de fond de l'expérience. Ces signaux sont pourtant primordiaux pour construire les enchaînements séquentiels. En absence de signaux CB_{i-1}, l'identification de taches HSQC se succédant dans la séquence de RegB se trouve réduite à la comparaison des déplacements chimiques des CA ($|\delta_{i3} - \delta_{j4}| \leq r$) ce qui augmente évidemment les ambiguïtés. Par ailleurs, plusieurs taches HSQC de RegB H48A ont des verticales HNCA qui contiennent des raies anormalement larges et chevauchantes qui rendent l'affectation des déplacements chimiques carbone très difficile et dans certains cas impossible.

La stratégie « classique » d'attribution que nous avons entreprise n'est pas suffisante pour attribuer complètement la séquence de RegB. Atteindre un tel objectif demande de reconstruire une stratégie d'attribution qui conjugue les contributions de plusieurs méthodes dont les plus importantes sont : le marquage ^{15}N sélectif de certains résidus aminés, le découplage des carbones CB lors des expériences HNCA et HNCOCA, l'emprunt de nouveaux trajets magnétiques utilisant des couplages dipolaires et éventuellement la deutération de la protéine (remplacement des protons non échangeables avec des noyaux Deutérium).

Le marquage sélectif d'un résidu aminé avec l'isotope ^{15}N se fait en exprimant la protéine d'intérêt sur un milieu minimum contenant l'acide aminé en question marqué sur son azote amide. Ce marquage permet d'enregistrer un spectre bidimensionnel $^{1}H/^{15}N$ HSQC simplifié ne contenant que les résonances de l'acide aminé marqué. Ce marquage aide à l'attribution puisqu'il attribue à une tache HSQC un type de résidu donné sans avoir recourt aux déplacements chimiques des carbones CB.

Le découplage des carbones CB lors des expériences HNCA et HNCOCA permet de réduire les largeurs de raie dans la dimension carbone, donc de rendre possible l'attribution de certains déplacements chimiques.

L'emprunt de trajets magnétiques utilisant des couplages dipolaires (à travers l'espace) permet de vérifier, de lever certaines ambiguïtés et de compléter des enchaînements séquentiels obtenus à l'aide des expériences utilisant les couplages scalaires. L'expérience 3D $^{1}H/^{15}N/^{1}H$ NOE-HSQC, par

exemple, permet de retrouver des enchaînements séquentiels en étudiant certains des couplages dipolaires présentés sur la Figure 58.

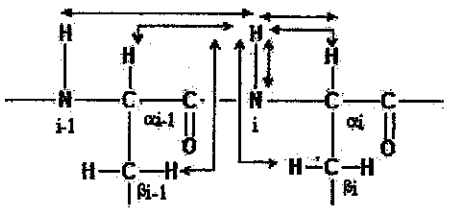


Figure 58 : Couplage scalaire (HN) et certains des couplages dipolaires exploitables à partir d'une expérience 3D $^{1}H/^{15}N/^{1}H$ NOE-HSQC.

Enfin, la deutération de la protéine permet d'augmenter la sensibilité des expériences 3D en diminuant la relaxation (perte de l'excitation magnétique) des noyaux ^{13}C de la protéine. Cette relaxation se fait principalement par l'intermédiaire d'interactions dipolaires entre le carbone et les protons qui lui sont attachés (Browne *et al.*, 1973). Elle varie dans le même sens que le rapport gyromagnétique γ du noyau source de la relaxation. Parce que le noyau deutérium a un γ 6.5 fois plus faible que celui du noyau proton, la deutération s'accompagne d'une réduction significative de la relaxation des carbones et d'une augmentation de la sensibilité des expériences RMN utilisant des transferts d'aimantation à travers les noyaux ^{13}C.

Au moment de la rédaction de ce manuscrit, nous avons entrepris la mise au point et l'application à RegB H48A de plusieurs des solutions d'attribution complémentaires que nous venons de décrire à l'exception de la deutération. Nous sommes raisonnablement optimiste quant à l'attribution sous peu de la plus grande partie de la séquence de RegB et en particulier des positions 17, 48, 52 et 68 dont l'importance fonctionnelle a été soulignée dans les chapitres précédents.

L'attribution séquentielle de RegB H48A permet d'envisager l'étude par RMN de l'interaction de cette protéine avec son substrat modèle (l'ARN Selex22tb) en utilisant des expériences bidimensionnelles de type $^{1}H/^{15}N$ HSQC. Une telle étude peut donner une première localisation du site actif de RegB avant même de construire un modèle de sa structure tridimensionnelle.

Conclusion

Conclusion

La protéine RegB, codée dans le génome du bactériophage T4, est une endoribonucléase séquence spécifique qui clive certains ARNm au milieu du motif GGAG. Il s'agit d'une ribonucléase si originale en terme de spécificité qu'il est légitime de lui associer le qualificatif de « ribonucléase de restriction ». Plusieurs travaux expérimentaux ont permis de mieux comprendre les fonctions physiologiques de cette enzyme dans le cycle de développement du phage T4 et de suggérer par des approches enzymatiques des signaux de reconnaissance au niveau de l'ARN cible. Malheureusement, aucune étude, fonctionnelle ou structurale, n'a été entreprise afin de comprendre à l'échelle moléculaire le comportement de RegB. La difficulté d'exprimer cette protéine très toxique en est certainement la cause majeure.

Au cours de cette thèse, nous avons commencé par caractériser de façon précise le niveau de toxicité de RegB dans *E. coli*. Nous avons établi à cet effet une échelle quantitative qui permet de définir ce niveau de toxicité dans *E. coli* et nous l'avons exploité pour construire un nouvel outil de génie génétique (pTOXR-1) permettant une sélection positive des plasmides recombinants. Par ailleurs, nous avons mis au point trois nouveaux systèmes d'expression de RegB : expression en fermenteur, utilisation d'un vecteur à réplication thermorégulable et expression dans un système eucaryote. La diversité de ces systèmes d'expression permet à présent de mieux répondre aux exigences de certaines méthodes expérimentales en terme de quantités protéiques.

Dans une deuxième partie, nous avons utilisé l'un de ces systèmes d'expression pour démontrer par résonance magnétique nucléaire du noyau ^{31}P que RegB produit un phosphodiester cyclique. L'identification de ce produit a permis de situer RegB dans le monde des ribonucléases connues. RegB occupe désormais le groupe des « transphosphorylases » au même titre que la RNase T1 et la RNase A. Malheureusement, il n'existe aucune homologie séquentielle significative entre RegB et les RNases de ce groupe qui pourrait aider à localiser les segments fonctionnellement important de RegB. Il était donc légitime d'entreprendre un balayage de la structure primaire de RegB par mutagenèse aléatoire.

Nous avons réalisé, indépendamment, la mutagenèse aléatoire de la partie N-terminale et de la partie C-terminale de RegB. Les résultats obtenus étaient difficilement interprétables à l'échelle du résidu aminé à cause de la présence de plusieurs mutants multiples. En revanche, il en ressort que, qualitativement, la partie N-terminale de RegB est beaucoup plus sensible à l'inactivation par substitution de résidus aminés que la partie C-terminale. Ce résultat ne donne qu'une localisation très grossière des résidus importants pour l'activité de RegB, insuffisante pour identifier le site actif de

l'enzyme. Nous avions la possibilité de cribler un plus grand nombre de mutants aléatoires et affiner les conclusions relatives aux mutants multiples en construisant par mutagenèse dirigée les mutants ponctuels correspondants. Il s'agit d'une démarche lourde et fastidieuse. Aussi, avons nous décidé de nous limiter aux résultats qualitatifs de la mutagenèse aléatoire et de les associer à une étude plus rationnelle des résidus aminés impliqués dans la catalyse ribonucléasique.

L'étude comparative des mécanismes réactionnels des ribonucléases de type transphosphorylase a permis de montrer l'intervention récurrente d'un résidu aminé clé : l'histidine. Nous avons réussi à démontrer par des études *in vitro* et *in vivo* que parmi toutes les histidines de RegB, seules les histidines H48 et H68 sont importantes pour certains des aspects de l'activité de l'enzyme à l'exclusion de la capacité à lier l'ARN substrat. Les deux mutants H48A et H68A présentent en effet la même affinité pour l'ARN que l'enzyme sauvage. Nous avons par ailleurs trouvé que l'importance relative de ces deux histidines est très différente. Alors que la substitution de l'histidine 48 annule complètement l'activité de RegB, la substitution de l'histidine H68 ne fait que diminuer l'activité de RegB d'un facteur 8-10. Ce résultat montre que l'histidine H48 est absolument indispensable à l'activité de RegB (elle pourrait jouer le rôle du catalyseur acide ou basique) alors que l'histidine H68 est dispensable au prix d'une perte très partielle de l'activité.

Peu de temps après ces travaux de mutagenèses, la publication de l'article « Snapshot of the genome of the pseudo-T-even bacteriophage RB49 » (Desplats *et al.*, 2002) a motivé notre projet de séquencer le gène *regB* du phage RB49, un virus apparenté au phage T4. Nous exploitons ainsi le résultat de l'évolution naturelle du gène *regB*. La séquence obtenue a montré la protéine RegB du phage RB49 a subi une grande délétion dans sa partie C-terminale soutenant ainsi l'hypothèse que nous avons proposée suite à la mutagenèse aléatoire du gène *regB*. Par ailleurs, la conservation dans le phage RB49 de l'histidine H48 confirme le caractère indispensable de cette histidine que nous avons démontré précédemment. L'histidine H68 n'est pas conservée dans le phage RB49. Parmi les hypothèses que nous avons émises figure celle qui suggère que la fonction de la chaîne latérale 68 pourrait être mimée une autre chaîne que l'histidine. Cette fonction reste à découvrir.

A ce stade nous disposions de plusieurs versions de RegB dont certaines s'exprimaient pour la première fois avec des rendements compatibles avec une étude structurale (unique étude rigoureuse capable à notre point de vue de préciser le rôle des histidines H48 et H68 en particulier et de définir le site actif de RegB en général). Plusieurs arguments ont concouru pour choisir le mutant RegB H48A comme modèle pour étudier la structure de RegB en solution. Nous avons donc mis au point les conditions spectroscopiques nécessaires à l'étude de cette protéine et avons entrepris la première étape de la résolution structurale de cette protéine par RMN : l'attribution séquentielle. Cette attribution devrait être achevée dans peu de temps.

Matériels et méthodes

Matériels et Méthodes

Nous présentons dans cette partie une collection des principaux protocoles utilisés. Ces protocoles, suffisamment détaillés pour être reproductibles, sont classés en trois grandes parties :

- Biologie moléculaire
- Biochimie (expression, purification et tests enzymatiques)
- Méthodes physiques (Résonance Magnétique Nucléaire et Résonance Plasmonique de Surface).

Biologie moléculaire

Souches bactériennes

La souche BL21 [F^- ompT $hsdS_B$ (r_B^- m_B^-) gal dcm], la souche BL21(λDE3) contenant le prophage λDE3 portant le gène de la T7 ARN polymérase et la souche LE392 [F^- el4$^-$ ($McrA^-$) hsdR514(r_K^- m_K^+) supE44 supF58 lacY1 galK2 galT22 metB1 trpR55] proviennent du fournisseur *Novagen*. La souche XL1-Blue [F'::Tn10 $proA^+B^+$ $lacI^q$ Δ(lacZ)M15/recA1 endA1 gyrA96 (Nal^r) thi hsdR17 (r_K^- m_K^+) glnV44 relA1 lac] et la souche JM101 [F' traD36 $proA^+B^+$ $lacI^q$ Δ(lacZ)M15/Δ(lac-proAB) glnV thi] proviennent de *Stratagene*. La souche DH5α [F^- endA1 hsdR17 (r_K^- m_K^+) glnV44 thi-1 recA1 gyrA(Nal^r) relA1 Δ(lacIZYA-argF)U169 deoR (φ80dlac Δ(lacZ)M15)] provient du fournisseur *Life Technologies*.

Enzymes

Les endonucléases de restriction, la T4 ADN polymérase, la T4 ADN ligase et la T4 ADN kinase proviennent de *New England Biolabs*. La Taq polymerase (Promega) a été utilisée pour les analyses qualitatives tel le criblage par PCR des colonies bactériennes ou de levures et pour la mutagenèse aléatoire. La très fidèle Pfu polymerase (Promega) a été utilisée pour les réaction de PCR destinées au clonage.

Calcul des unités d'une enzyme de restriction

La définition et les calculs qui suivent correspondent aux enzymes distribuées par *New England Biolabs.*

Définition d'une unité d'enzyme de restriction : C'est la quantité d'enzyme nécessaire pour digérer en une heure, à la température requise, en présence du tampon requis, dans un volume réactionnel de 50 µl ; 1 µg d'ADN (ce type d'ADN dépend de la nature de l'enzyme, mais est en général l'ADN du phage λ).

Calcul du nombre d'unités : Calculons le nombre d'unités U_E d'une enzyme E à ajouter pour digérer m µg d'ADN de T kpb de long. Soit N_λ le nombre de sites de l'enzyme E dans l'ADN de λ et N_P le nombre de sites de l'enzyme E dans l'ADN étudié.

$$U_E = m_{(\mu g)} * (N_P/N_\lambda) * (48.502/T_{(kb)})$$

Valeurs des N_λ des enzymes usuelles : BglII (6), BamHI (5), ClaI (15), EcoRI (5), HindIII (6), NdeI (7), XhoI (1).

Extraction rapide de fragments d'ADN à partir d'un gel d'agarose

Ce protocole s'applique à l'extraction d'ADN (100 pb à 50 kpb) à partir de gels TAE-agarose en utilisant les solutions Fast 'n' Easy (Tebu).

- Découper la bande d'agarose, la peser et lui ajouter 300 µl par 0.1 g de gel d'une solution d'iodure de sodium conservée à l'abri de la lumière. Bien mélanger.
- Chauffer la solution dans un bain marie à 55 °C jusqu'à la fusion totale du gel (5 min environ) en mélangeant le tube de temps en temps.
- Bien re-suspendre la matrice de silice (vortex 30 s) et l'ajouter au tube d'agarose fondu à raison de 6 µl pour récupérer 0-1 µg d'ADN. Bien mélanger la solution en vortexant doucement. Laisser à RT pendant 5 min en mélangeant de temps en temps puis centrifuger 5 s à 13.000 rpm, enlever le surnageant.

- Re-suspendre, soigneusement en pipetant, le culot dans un 1 ml de la solution de lavage puis centrifuger 5 s à 13.000 rpm, enlever le surnageant. Répéter cette étape de centrifugation puis enlever toutes traces de tampon de lavage en pipetant avec un petit cône.
- Re-suspendre, soigneusement en pipetant, le culot dans 14 µl d'eau bi-distillée (ne pas vortexer), laisser 5 min à RT, centrifuger 1 min à 13.000 rpm puis récupérer 12 µl de surnageant sans toucher le culot.
- Centrifuger de nouveau le surnageant précédent 1 min à 13.000 rpm puis récupérer 10 µl de surnageant sans toucher le culot (les traces mêmes infimes de la matrice de silice sont capables d'inhiber plusieurs enzymes utilisées en Biologie moléculaire).

Ligature efficace de deux fragments d'ADN

La ligation se fait dans un volume final de 10 à 15 µl maximum (de préférence 10 µl).

- Ajouter au mélange insert-vecteur le volume adéquat du tampon de la T4 ADN ligase (*New England Biolabs*). Attention, ce tampon contient de l'ATP, il doit être aliquoté par petits volumes (5 µl) dans des Eppendorfs à usage unique et stocké à -20 °C.
- La masse totale d'ADN (insert + vecteur) doit être entre 0.1 et 0.5 µg et le rapport *molaire* insert/vecteur doit être entre 2 et 5.
- Ajouter 1 µl de T4 ADN ligase de Biolabs en faisant attention à la garder à -20 °C (enzyme très fragile) puis incuber ON à 16 °C (utiliser de préférence l'appareil de PCR réglé à cette température, son couvercle chauffant permet en effet d'éviter l'apparition de condensations sur les parois du tube donc la concentration de la solution de ligature).
- Le protocole suivant (Thomas M.R., 1994) permet d'utiliser l'entièreté d'une ligature tout en évitant l'apparition d'un arc électrique lors de l'électroporation : Ramener le volume de ligature à 50 µl avec de l'eau bi-distillée. Ajouter 500 µl de n-butanol et vortexer 10 sec jusqu'à mélange complet. Centrifuger 10 min à 13.000 rpm. Eliminer le surnageant très soigneusement : un petit culot est généralement visible. Sécher (faire attention à ne laisser aucune trace de n-butanol) et resuspendre soigneusement le culot dans 5 µl d'eau bi-distillée stérile. Stocker à -20 °C jusqu'à utilisation. Pour l'électroporation, utiliser 3 µl de la solution de ligation.

Transformation des bactéries par électroporation

Toute la procédure de préparation des cellules électro-compétentes se déroule à 4 °C.

- Inoculer 1 litre de milieu YENB [7.5 g/l *yeast extract*, 8.0 g/l *nutrient broth*], à l'aide de 5-10 µl d'une pré-culture fraîche de la souche bactérienne étudiée. Laisser pousser à 37°C jusqu'à 0.5-0.9 unités DO_{600}.
- Placer dans la glace puis centrifuger (4000 g, 10 min, 4°C). Eliminer le surnageant, égoutter chaque pot sur papier absorbant, ajouter 100 mL d'eau distillée stérile froide (glace), re-suspendre doucement les culots (par simple rotation du pot, ne pas utiliser de pipetman), centrifuger (4000 g, 10 min, 4°C). Laver une deuxième fois avec de l'eau bi-distillée stérile.
- Reprendre tout le culot dans 20 ml de glycérol 10 % stérile et à 4 °C. Centrifuger (4000 g, 10 min, 4°C). Eliminer le surnageant et reprendre le culot dans 2-3 ml de glycérol 10% stérile froid. Aliquoter par 44 µl, congeler rapidement dans l'azote liquide puis stocker à -80 °C.

Pour l'électroporation, décongeler les cellules dans la glace, ajouter l'ADN transformant (volume inférieur à 5 µl), attendre 1 min, électroporer à 2500 V en utilisant des cuves de 2 mm. La constante de temps doit être entre 4 et 5.5 ms. Ajouter 1 ml de milieu LB, incuber 1 heure à 37 °C sous agitation puis étaler sur le milieu de sélection.

Transformation des bactéries au Chlorure de Calcium

La procédure se déroule à 4 °C.

- Cultiver à 37 °C la souche bactérienne étudiée dans 50 ml de milieu LB jusqu'à 0.5 unités DO_{600} .
- Placer la culture dans la glace, centrifuger (10 min, 6000 rpm, 4 °C) puis reprendre le culot dans un 1.5 ml de $CaCl_2$ 0.1 $mol.l^{-1}$ initialement dans la glace.
- Centrifuger (10 min, 6000 rpm, 4 °C), reprendre le culot dans un 1.0 ml de $CaCl_2$ 0.1 $mol.l^{-1}$, aliquoter par 100 µl puis laisser 30 min dans la glace.
- Ajouter l'ADN transformant (1-2 µl) puis laisser 20 min dans la glace.
- Faire un choc thermique à 42 °C pendant exactement 1 min 30 s, mettre dans la glace pendant 1 min puis ajouter 1 ml de LB.

- Cultiver sous agitation à 37 °C pendant 1 heure puis étaler sur un milieu de sélection.

Extraction ultra-rapide de l'ADN plasmidique bactérien (30 min)

Le protocole est inspiré de celui de Berghammer H (Berghammer *et al.*, 1993). Toute la procédure se déroule dans un seul et unique tube de 1.5 ml.

- Centrifuger 1.5 ml de culture bactérienne en phase stationnaire à 13.000 rpm pendant 2 min. Aspirer le surnageant et *très bien décanter le culot.*
- Re-suspendre le culot dans 60 µl de tampon de lyse (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 15 % w/v sucrose, 2 mg/ml lysozyme, 0.2 mg/ml RNAse pancréatique et 0.1 mg/ml de BSA). Ce tampon de lyse doit être stocké à -20 °C en petits aliquotes à usage unique. Vortexer 15 secondes, laisser à RT pendant 5 min en mélangeant le tube de temps en temps.
- Trouer avec une aiguille de seringue le bouchon du tube, le placer à 95 °C pendant 60 s puis le plonger immédiatement dans la glace pendant 60 s.
- Centrifuger à 13.000 rpm à RT pendant 20 min. Les débris cellulaires et l'ADN chromosomique se retrouvent dans le culot stabilisé par la BSA précipitée par la chaleur.
- Prendre 50 µl de surnageant plasmidique dans un tube propre. L'ADN est prêt pour subir les analyses de restriction pas les enzymes habituelles hormis NdeI (très sensible à la pureté de la préparation).

Pour l'extraction quantitative en moins de 30 min d'un ADN plasmidique bactérien ultra-pur (pour le séquençage par exemple), utiliser le kit onéreux mais extrêmement efficace : Fast 'n' Easy Plasmid Prep (Tebu).

Criblage des colonies bactériennes par PCR

Il est possible de cribler les clones bactériens à la recherche du plasmide recombinant en réalisant une PCR sur des colonies entières en présence des deux amorces encadrant l'insert recherché.

Préparer à 4 °C un mélange réactionnel de 20 µl contenant : le tampon Taq Polymérase (Promega) concentré 1 fois, 3 mM de $MgCl_2$, 0.2 mM dNTP, 100 pmole de chaque amorce et une

unité de Taq Polymérase (Promega). Ajouter ensuite le un tiers d'une colonie bactérienne (18 heures d'incubation à 37 °C) au mélange réactionnel (une trop grande quantité de cellules bactériennes peut inhiber la réaction de PCR). Utiliser enfin le programme de PCR suivant : [(94 °C : 10 min) X 1, (94 °C : 30 s, 40 °C : 30 s, 72 °C : 45 s) X 30, (72 °C : 7 min) X 1]. Notons que ce protocole s'applique aussi au criblage des colonies de levures *S. cerevisiae*.

Préparation et titrage d'un lysat du phage λCE6

Il est recommandé de préparer un lysat du phage λCE6, il choisir une souche hôte ayant le système modification que la souche dans laquelle se fait l'induction avec le phage (exemple, amplifier dans la souche LE392 (rK- mK+) pour induire dans DH5α (r_{K+} m_{K+})).

- Cultiver à 37 °C la souche LE392 dans 10 ml de LB + 0.2 % maltose + 10 mM $MgSO_4$ finaux jusqu'à 0.6 unités DO_{600}.
- Mélanger délicatement 5 ml de cette culture avec 2 10^{+8} particules phagiques puis incuber 15 min à 37 °C sans agitation.
- Ajouter le mélange à 500 ml de LB + 10 mM MgSO4 final préchauffé 5 min à 37 °C puis incuber sous une agitation (250 rpm) jusqu'à observation d'un début de lyse (3 à 5 heures).
- Ajouter 2 ml de chloroforme et agiter à 37 °C (250 rpm, 15 min) puis centrifuger 15 min à 10000 rpm à 4 °C.
- Tout en restant dans la glace, bien décanter le surnageant dans des tubes Falcon de 50 ml (45 ml par tube) et stocker le lysat à 4 °C.

Il est recommandé de titrer le lysat au moins 24 heures après sa préparation afin que sa concentration phagique efficace se stabilise. Le titre obtenu est alors stable pendant 2 à 3 semaines.

- Cultiver à 37 °C la souche LE392 dans 10 ml LB + 0.2 % maltose + 10 mM MgSO4 finaux jusqu'à 0.6 unités DO_{600}. Aliquoter les cellules dans des tubes de 15 ml à raison de 200 µl de cellules par tube. Maintenir dans la glace.
- Préparer dans la glace une solution LB + 0.2 % maltose + 10 mM MgSO4 finaux et l'utiliser pour préparer une série de dilutions du lysat avec une raison égale à 10^{-2}.
- Faire fondre le milieu solide Top-Agar et le laisser en surfusion à 42 °C dans des tubes de 15 ml (3 ml par tube).

- À 200 µl de cellules ajouter 100 µl de dilution du lysat, mélanger délicatement et incuber sans agitation 15 min dans un bain-marie à 37 °C.
- Ajouter le mélange cellules + dilution du lysat au 3 ml de Top-Agar en surfusion, mélanger par inversion puis couler immédiatement et uniformément sur des boîtes LB + 0.2 % maltose + 10 mM MgSO4 finaux préchauffées 5 min à 37 °C puis incuber une nuit à 37 °C.
- Soit *p* le nombre de moyen de plages trouvées pour la dilution la plus forte 10^{-n} (moyenne des résultats des deux dernières dilutions). Le titre du lysat est calculé selon la formule :

$$T = p.10^{n+1} \text{ phage/ml}$$

Extraction du plasmide pTOXR-1 de la levure

Ce protocole est une adaptation de celui publié par Robzyk et Kassir (Robzyk *et al.*, 1992). Des quantités d'ADN voisine du micro-gramme peuvent être obtenues avec ce protocole.

- Re-suspendre les cellules de levures provenant d'une culture, sur la nuit à 28 °C, de 100 ml du milieu SM6 [0.7 % Yeast Nitrogen Base (Difco), 0.01 % L-tryptophane (Sigma), 0.01 % L-leucine (Sigma), 0.005 % L-histidine (Sigma), 0.1 % bacto-casaminoacids (Difco) and 2.0 % D(+)-glucose (Sigma)] dans 6.7 ml du tampon GTET [8.0 % D(+)-glucose, 50.0 mM Tris-HCl pH 8.0, 1.0 mM EDTA, 5.0 % Triton X-100 (Sigma)].
- Ajouter un volume équivalent de billes de verres (0.4-0.6 mm) lavées avec l'acide chlorhydrique et casser les cellules en mélangeant vigoureusement à la main pendant 5 min à température ambiante.
- Ajouter 6.7 ml du tampon GTET puis mettre le mélange à bouillir à 95 °C pendant 4 min.
- Mettre dans la glace pendant 10 min puis centrifuger 10 min à 130000 rpm à 4 °C.
- Ajouter au surnageant un demi-volume équivalent d'acétate ammonium 7.5 M, mélanger, incuber 1 heure à -20 °C puis centrifuger 10 min à 130000 rpm à 4 °C.
- Précipiter le surnageant avec de l'éthanol absolu, laver le culot avec de l'éthanol à 70 %, le sécher et le reprendre dans 200 µl d'eau bi-distillée.
- Charger la préparation sur un gel TAE-agarose 1 % et faire migrer lentement (50V, 5 heures avec un appareil de 15 cm x 15 cm).

- Exciser rapidement sous UV la bande du plasmide pTOXR-1 puis extraire rapidement le plasmide de l'agarose avec le protocole Fast 'n' Easy DNA purification kit (Tebu).

Clonage à sélection positive utilisant le plasmide pTOXR-1

- Digérer 10 ng du plasmide pTOXR-1 avec les enzymes NdeI et BamHI pendant une durée ne dépassant pas 30 min (tenir compte de la définition de l'unité enzymatique). Arrêter la réaction en chauffant 20 min à 80 °C.
- Précipiter la digestion à l'aide d'une solution d'acétate de sodium 3M et d'éthanol absolu, sécher et reprendre dans 5 µl d'eau bi-distillée.
- Ligaturer le vecteur ainsi préparé avec l'insert digéré en utilisant la T4 ADN ligase (1 heure à 16 °C). Précipiter au n-butanol (Thomas *et al.*, 1994) puis transformer la souche bactérienne.

Mutagenèse dirigée des histidines de RegB

Ce protocole est une adaptation du protocole ExSite™ PCR-Based Site-Directed Mutagenesis (Stratagene). Il utilise 2 amorces pour chaque mutagenèse dirigée. Pour la construction des mutants histidine-en-alanine de RegB nous avons utilisé les amorces suivantes : 421 (GCT CTA AAA TAT TCT CAA CAT CTT CTT GAT CGC G) et 422 (AAA CGA TGA GAC TCC TGC TGC TTT TG) pour la substitution H42A; 481 (GCT CTT CTT GAT CGC GCA ATT CAA CGG) et 482 (TTG AGA ATA TTT TAG ATG AAA CGA TGA GAC TCC) pour la substitution H48A; 681(GCT AAA ATA AAA GAC CAT GTT TTA GAA GTT AAT G) et 682 (GAA TAA TTC AAA AAC GTA TGT CTC ATC AAT CTC) pour la substitution H68A et 731 (GCT GTT TTA GAA GTT AAT GAA TTC CTG AGT ATG C) et 732 (GTC TTT TAT TTT ATG GAA TAA TCA AAA ACG TAT G) pour la substitution H73A. Il est primordiale que les amorces soient purifiées sur gel d'acrylamide puis phosphorylées avant de les utiliser. La phosphorylation des amorces se déroule comme suit :

- Dans un volume final de 20 µl, mélanger 1 µg d'amorce et 10 unités de T4 ADN kinase avec son tampon réactionnel.
- Incuber 1 heure à 37 °C.
- Arrêter la réaction en chauffant 10 min à 65 °C. Les amorces ainsi phosphorylées peuvent être stockées à -20 °C pendant plusieurs semaines.

Nous avons utilisé le plasmide pARNU2 contenant le gène *regB* sauvage comme matrice pour la mutagenèse.

- Dans un volume final de 25 µl préparer le mélange suivant : 0.5 pmole du plasmide pARNU2, 15 pmole de chacune des amorces phosphorylées, 1 µl de dNTP 25 mM et 5 unités de l'ADN polymérase ExSite Blend Polymerase dans son tampon.
- Amplifier par PCR en utilisant le programme suivant [(94 °C : 4 min, 42 °C : 2 min, 72 °C : 2 min) X 1 ; (94 °C : 1 min, 45 °C : 2 min, 72 °C : 1 min) X 8 ; (72 °C : 5 min) X 1].
- Mettre les produits de PCR 5 min dans la glace. Ajouter 10 unités de l'enzyme de restriction DpnI et 1.25 unités de l'ADN polymérase recombinante Pfu. Mélanger, incuber 30 min à 37 °C puis inactiver les enzymes en chauffant à 72 °C pendant 30 min.
- Purifier sur gel d'agarose le plasmide néo-synthétisé relaxé et résistant à la coupure par *DpnI*.
- Procéder à la circularisation du plasmide en utilisant la T4 ADN ligase selon le protocole de ligature décrit auparavant puis transformer par électroporation la souche bactérienne XL1-Blue.

Biochimie

Expression/purification de RegB en fermenteur

- Cultiver, à 37 °C jusqu'à la phase stationnaire, la souche bactérienne XL1-Blue transformée avec le plasmide pARNU2 dans 800 mL du milieu de culture Luria Bertani Broth [1.0% Nutriment Broth (Difco), 0.5% Yeast Extract (Difco), 0.8% NaCl, 0.052% Tris] contenant 100 $\mu g.mL^{-1}$ d'ampicilline (Sigma) et 0.2 % D(+) maltose (Sigma)
- Inoculer avec culture le fermenteur contenant 15 L de milieu LB supplémenté de 100 $\mu g.mL^{-1}$ d'ampicilline et 0.2 % D(+) maltose. Cultiver à 37 °C sous agitation.
- A 0.6 unités OD_{600}, ajouter du $MgSO_4$ (10 mM final) et le phage λCE6 (Novagen) à une concentration finale de 2.10^9 $pfu.mL^{-1}$.
- Centrifuger les cellules au bout de 3 heures et les resuspendre dans le tampon A [50 mM phosphate de sodium (pH 8.0), 300 mM NaCl, 0.1 mM AEBSF protease inhibitor (Euromedex) et 5.0 mM β-mercaptoethanol]. Garder les cellules à -80 °C.

- Casser les cellules mécaniquement à 4 °C en ajoutant un volume équivalent de billes de verres (0.4-0.6 mm, Sigma) et en mélangeant deux fois 10 min à la vitesse de 4000 rpm dans l'appareil Dynomill.
- Centrifuger 5 min à 10000 pour récupérer les billes de verres puis centrifuger l'extrait brut 30 min à 15000 rpm à 4 °C pour récupérer l'extrait soluble.
- Laver la résine de cobalt TALON (Clontech) comme suit : prélever 2 ml de suspension homogène/par 5 mg de protéines à purifier, centrifuger 2 min à 800 g et à 4 °C. Jeter le surnageant, resuspendre la résine dans le tampon de lyse A puis centrifuger 2 min à 800 g et à 4 °C. Jeter le surnageant.
- Ajouter l'extrait soluble à la résine et mettre à agiter lentement sur rôtisserie à 4 °C pendant 40 min.
- Récupérer la résine par centrifugation (5 min à 4 °C à 800 g), la laver une fois avec le tampon de lyse A puis une fois avec le tampon de lavage B (tampon de lyse A + 20 mM imidazole).
- Déposer la résine sur la colonne, laisser sédimenter puis laver en continue avec le tampon B (5 fois le volume de la résine).
- Eluer avec 10 ml de tampon C (tampon de lyse A + 150 mM imidazole).
- Dialyser les fractions concentrées contre le tampon 50 mM phosphate de sodium (pH 8.0), 300 mM NaCl et 5.0 mM β-mercaptoethanol.

Expression de RegB dans la levure

- Inoculer, avec une colonie de levure INVSc1 transformée avec le plasmide pTOXR-1, 50 ml de milieu de culture SM6 [0.7 % Yeast Nitrogen Base (Difco), 0.01 % L-tryptophane (Sigma), 0.01 % L-leucine (Sigma), 0.005 % L-histidine (Sigma), 0.1 % bacto-casaminoacides (Difco) et 2.0 % D(+)-glucose (Sigma)].
- Cultiver les cellules ON à 28 °C, puis les centrifuger et les resuspendre dans le milieu d'induction SM6 contenant du galactose 2.0 % de façon à ce que la solution ait initialement 0.4 unités DO_{600}.
- Récupérer les cellules 8 heures plus tard.
- Casser les cellules mécaniquement avec des billes de verre et purifier sur résine de cobalt comme d'usage.

Double marquage $^{13}C/^{15}N$ de RegB H48A

Une pré-culture sur la nuit de 50 ml de la souche BL21(λDE3)-pLysS (Novagen) transformée avec le plasmide pEH48A a été inoculée à 1 l de milieu minimum M9 [12.8 $g.l^{-1}$ (Na_2HPO_4, $2H_2O$); 3.0 $g.l^{-1}$ KH_2PO_4; 0.5 $g.l^{-1}$ NaCl; 1.0 $g.l^{-1}$ $^{15}NH_4Cl$; 0.1 $mmol.l^{-1}$ $CaCl_2$, 1.0 $mmol.l^{-1}$ $MgSO_4$ et 2.0 $g.l^{-1}$ glucose] contenant 100 $\mu g.ml^{-1}$ ampicillin et 34 $\mu g.ml^{-1}$ chlloramphenicol. L'expression de la protéine a été induite à 0.6 unités de OD_{600} pendant 3 heures en utilisant 1.0 $mmol.l^{-1}$ IPTG (Promega). Les cellules ont été récupérées dans le tampon B1 [50 $mmol.l^{-1}$ Tris-HCl (pH 8.0), 300 $mmol.l^{-1}$ NaCl, 0.1 $mmol.l^{-1}$ AEBSF protease inhibitor (Euromedex) et 5.0 $mmol.l^{-1}$ β-mercaptoethanol] et stockées à -80 °C. Toute la procédure de purification se déroule à 4 °C. Les cellules sont cassées par sonication, l'extrait brut est clarifié par centrifugation à 15000 rpm pendant 40 min puis mélangé avec le résine de cobalt (Clontech) pendant 40 min. La résine est ensuite lavée une fois avec le tampon B1 et une fois avec le tampon B2 (tampon B1 + 20 $mmol.l^{-1}$ imidazole). Elle est ensuite déposée dans une colonne de chromatographie et lavée 6 fois son volume avec le tampon B2. La protéine a été éluée avec le tampon B3 (Tampon B1 + 150 $mmol.l^{-1}$ imidazole). Finalement, le tampon B3 a été échangé contre le tampon RMN [50 mM sodium citrate (pH 6.0), 300 $mmol.l^{-1}$ NaCl, 0.1 $mmol.l^{-1}$ AEBSF protease inhibitor (Euromedex) et 2.0 mM dithiothreitol] en utilisant les colonnes de chromatographie Micro Bio-Spin® (Biorad).

Analyse western-blot de RegB

- Séparer les échantillons protéiques à analyser sur un gel SDS-PAGE 15 %. Enlever le gel de concentration et orienter le gel de séparation.
- Laver le gel à l'eau bi-distillée 1 min sous agitation douce.
- Tremper le gel dans le Tampon de Transfert 1 X (voir la partie Tampons usuels) pendant 3 min sous agitation douce.
- Préparer la membrane Hybond - P PVDF (Amersham Pharmacia Biotech) : Utiliser des gants et une pince stérile pour manipuler la membrane, couper un carré de membrane dépassant légèrement les dimensions du gel, imbiber 10 s dans du méthanol 100 %, laver 5 fois à l'eau bi-distillée puis équilibrer la membrane en la trempant dans le tampon de transfert, 10 min sous agitation douce.
- Préparer le « sandwich » comme suit en faisant très attention à l'orientation : Mettre dans l'ordre grille, chiffon, trois papiers wattman 3MM imbibés dans le tampon de transfert, la

membrane, le gel (attention à ne pas emprisonner de bulles d'air), trois papiers wattman 3MM imbibés dans le tampon de transfert, chiffon et grille.

- Fermer le « sandwich », l'introduire dans le bon sens dans l'appareil de transfert (*Hoefer TE Series Electrophoresis Unit, modèle TE 22, Série 96-5071*) puis remplir ce dernier avec du Tampon de Transfert. La protéine doit migrer vers le pole *plus*, il faut donc avoir dans l'ordre : pole *moins*, gel, membrane, pole *plus*.
- Migrer 2 heures à 80 V.
- Récupérer la membrane et marquer la face où se trouve la protéine avec un crayon, tout en orientant la membrane de la même façon que le gel.
- Récupérer le gel puis le colorer au bleu de coomassie afin de vérifier l'efficacité du transfert.
- Rincer la membrane brièvement dans le tampon TBS (voir ci-après), la « bloquer » dans 30 ml de Tampon de Blocage/Incubation (voir ci-après), pendant 1 heure, sous agitation douce puis la laver avec le tampon de Lavage (voir ci-après), 1 fois 5 min.
- Incuber pendant une heure la membrane avec une solution d'anticorps *Anti-polyhistidines tag (Anti-His Antibody, Amersham Pharmacia Biotech, réf. 27-4710)* diluée dans le tampon de Blocage /Incubation au 1/3000 (8.3 µl pour 25 ml final).
- Rincer avec le tampon de Lavage, 3 fois 5 min.
- Ajouter l'anticorps secondaire couplé à la phosphatase alcaline dilué au 1/15000 (1.7 µl dans 25 ml final) dans le tampon de Blocage/Incubation (*Anti-Mouse IgG, Sigma , réf. A-3438*), mélanger pendant une heure puis rincer avec le tampon de Lavage, 3 fois 5 min.
- Préparer la solution de révélation : 33 µl de NBT 5% (conservé à -20 °C) + 16 µl de BCIP 5% (conservé à -20 °C) + 5 ml de tampon TBS. La coloration apparaît au bout de 5-10 min (Ne pas laisser la réaction se poursuivre très longtemps sinon le bruit de fond devient très intense).

Tampon TBS 10X (200 mM Tris-HCl, 1.37 M NaCl, pH 7.6) : 12.1 g de Tris + 40.0 g de NaCl + 400 ml eau bi-distillée, ajuster à pH 7.6 avec de l'acide chlorhydrique, compléter avec de l'eau bi-distillée jusqu'à 500 ml. Stocker à 4°C.

Tampon de Blocage/Incubation (TBS 1X, 0.05 % Tween 20, 5 % w/V lait écrémé) : 10 ml TBS 10 X + 80 ml eau bi-distillée + 5 g lait écrémé en poudre + 50 µl Tween 20 100 %, dissoudre puis compléter à 100 ml avec de l'eau bi-distillée.

Tampon de Lavage (TBS 1X, + 0.05 % Tween 20) : 25 ml TBS 10X + 125 µl Tween 20 100 %, compléter à 250 ml avec l'eau bi-distillée.

Précipitation des protéines à l'acide trichloracétique

- Ajouter à la solution protéique un volume équivalent d'une solution d'acide trichloracétique 20 % conservée à 4°C, mélanger puis incuber dans la glace pendant 30 min.
- Centrifuger à 4 °C, 13.000 rpm pendant 15 min, enlever délicatement le surnageant puis ajouter 300 µl d'acétone froide et centrifuger 5 min à 4 °C.
- Enlever le surnageant, sécher au speed-vacuum puis reprendre le culot dans le *Tampon d'échantillon* 1X (voir partie Tampons usuels).
- Chauffer 5 min à 95 °C avant de charger sur gel SDS-PAGE.

Suivi cinétique du clivage de l'ARN Selex22tb par RegB

Les réactions de coupure ont été réalisées à 37 °C dans un le tampon : 4 mmol.l^{-1} phosphate de sodium (pH 8.0), 24 mmol.l^{-1} NaCl et 0.4 mmol.l^{-1} β-mercaptoethanol. Le mélange réactionnel (50 µl) contient une concentration en ARN Selex2tb de 3.6 µmol.l^{-1}, une concentration de 0.36 µmol.l^{-1} de RegB et 40 unités de l'inhibiteur de RNases *RNA Guard®* (*Pharmacia*).

La réaction est suivie pendant 40 min. Toutes les 10 min, un aliquote est prélevé, la réaction est arrêtée en ajoutant du formamide (40 % final) et en chauffant 5 min à 95 °C. Les échantillons sont stockés à -20 °C jusqu'à utilisation. Les produits de coupure sont séparés et révélés comme suit :

- Préparer un gel TAE-Urée-Acrylamide à 20 % [Mélanger 4.2 g d'urée + 5 ml d'acrylamide 40 % + 1 ml de TAE 10X (400 mmol.l^{-1} Tris-Acétate, 10 mmol.l^{-1}), dissoudre dans un bain-marie à 37 °C, ajuster à 10 ml, ajouter les catalyseurs de polymérisation (60 µl d'ammonium persulfate 10 % + 10 µl de N,N,N',N'-Tetramethyl-1-,2-diaminomethane absolu), couler le gel et laisser polymériser 2 heures.].
- Placer le gel dans l'appareil d'électrophorèse (BioRad Mini Protean II), mettre le tampon de migration et faire migrer 1 heure à 200 V pour préchauffer.
- Rincer soigneusement les puits avec une seringue pour enlever les dépôts d'acrylamide et d'urée.
- Déposer les échantillons au centre du gel afin d'éviter les distorsions dues aux effets de bords. Les échantillons ont été préparés au préalable comme suit : ajouter à 8.5 µl d'échantillon ARN 6 µl de formamide 100 % et 1.6 µl de TAE 10X, dénaturer à 95 °C pendant 5 min puis ajouter 4 µl de tampon de charge ADN 5X.

- Faire migrer le gel à 200 V, pendant 2-3 heures, le temps que le colorant bleu (Bleu de bromophénol) arrive à 1 cm du bas du gel.
- Colorer le gel pendant exactement 7 min dans 50 ml de TAE 1X + 0.25 µg/ml de bromure d'éthidium.
- Rincer avec du TAE 1X pendant 30 secondes.
- Détecter sous lumières Ultra-violettes en utilisant un trans-illuminateur couplé à une carte d'acquisition (*Vilbert Lourmat*).
- Quantifier le produit de coupure (15 nucléotides) en utilisant comme référence la quantité totale d'ARN Selex22tb déposée (temps 0) corrigée pour tenir compte de sa taille (21 nucléotides).

NB : L'ARN Selex22tb a été synthétisé chimiquement sur l'appareil *Pharmacia LKB Gene Assembler Plus* en utilisant des phénoxyacétyl-β-ARN phosphoramidites (*Amersham Pharmacia Biotech*). Le protocole a été adapté à partir de celui utilisé pour la synthèse d'ADN, l'unique différence réside dans l'utilisation de 5-éthyl-thio-1H-tetrazol comme activateur (Snoussi *et al.*, 2001). Le groupe terminal 5'-O-diméthoxytrityl a été enlevé à la fin de la synthèse. Les oligo-ribonucléotides ont été clivés du support et déprotégés selon le protocole décrit dans (Snoussi *et al.*, 2001). Les échantillons d'ARN ainsi produit a été purifiés par HPLC sur des colonnes Q-sépharose. Le pourcentage extrêmement faible des produits abortifs nous a permis d'utiliser les ARN directement après la purification HPLC sans passer par une étape de purification sur gel d'acrylamide.

Tampons usuels

Tampon de charge ADN 5X

Bleu de bromophénol (0.25 %) :	25 mg
Xylène cyanol FF (0.25 %) :	25 mg
Orange G(0.25 %) :	25 mg
Saccharose (40 %) :	4.0 g
Eau bi-distillée :	qsp 10 mL.

Aliquoter par 1 mL, conserver à -20°C.

Tampon de Laemmli 2X (= Tampon Echantillon 2X)

SDS (6 %) :	1.8 g
β-mercaptoethnol (6 %) :	1.8 mL (14.2 M)
Tris-HCl 250 mM pH 6.8 :	15 mL (Tris-HCl 500 mM)
Glycérol (20 %) :	6 mL (100 %)
Bleu de bromophénol (0.6 %) :	0.18 g
Eau bi-distillée :	qsp 30 mL.

Aliquoter par 1 mL, conserver à -20°C.

Tampon de séparation SDS-PAGE 4X

Tris :	18.15 g
SDS :	0.4 g
Ajuster le pH à 8.8	
Eau bi-distillée :	qsp 100 mL

Filtrer, aliquoter, conserver à -20°C.

Tampon de concentration SDS-PAGE 4X

Tris :	3.0 g
SDS :	0.4 g
Ajuster le pH à 6.8	
Eau bi-distillée :	qsp 50 mL.

Filtrer, aliquoter, conserver à -20°C.

Tampon d'électrophorèse 5X

Tris :	15.1 g

Glycine :	72 g
SDS :	5.0 g
Ajuster le pH à 8.3	
Eau bi-distillée :	qsp 1000 mL.

Tampon de transfert

Tampon d'électrophorèse 5X :	200 ml
Méthanol :	200 ml
Eau bi-distillée :	qsp 1000 ml

Tampon de coloration SDS-PAGE

Acide acétique glacial :	200 ml
Ethanol absolu :	400 ml
Eau bi-distillée :	400 ml
Phastgel BlueR :	2.0 g

Tampon de décoloration SDS-PAGE

Acide acétique glacial :	200 ml
Ethanol absolu :	400 ml
Eau bi-distillée :	400 ml

Méthodes physiques

Résonance plasmonique de surface

Toutes les expériences ont été réalisées sur l'appareil BIAcore 2000 (BIAcore) en utilisant un bio-senseur de type CM5 (BIAcore) maintenu à 25 °C.

La protéine a été couplée à la surface du bio-senseur en utilisant la chimie CO_2H/NH_2. La procédure de couplage est automatisée. Le canal N°1 du bio-senseur a été utilisé comme surface témoin (sans RegB). Le tampon de circulation est un tampon : Hepes 10 mmol.l^{-1} (pH 7.4), 150 mmol.l^{-1} NaCl, 3 mmol.l^{-1} EDTA et 0.005% polysorbate 20 (v/v).

Pour chaque expérience, 70 µL d'ARN Selex22tb ont été injecté à la vitesse de 25 µl.min^{-1}. Les concentrations d'ARN explorées vont de 75 à 140 µmol.l^{-1}. Au moins deux injections indépendantes à deux concentrations différentes d'ARN ont été réalisées pour chaque version de la protéine RegB.

La régénération du bio-senseur se fait par l'intermédiaire de deux impulsions courtes de 50 µL d'une solution contenant du NaCl 1 mol.l^{-1} et du NaOH 50 mmol.l^{-1}.

Tous les sensorgrammes ont été corrigés en réalisant la différence du signal observé avec celui obtenu avec la piste témoin.

Pour traiter les données expérimentales obtenues, nous avons utilisé le modèle Langmuir d'interaction 1 :1. Les constantes d'affinités ont été extraites des courbes en utilisant le programme BIA evaluation 3.0 (BIAcore).

Résonance Magnétique Nucléaire

Toutes les expériences RMN ont été réalisées sur un spectromètre Bruker DRX600 muni d'une sonde triple résonance $^{1}H/^{13}C/^{15}N$ de 5 mm de diamètre et d'une unité de gradients. Pour les expériences ^{31}P, une sonde large bande (^{1}H-BB) a été utilisée. Les expériences ont été transformées en utilisant le programme XWINNMR 2.0 (Bruker) puis analysées en utilisant l'excellent programme *Sparky* écrit par Thomas L. Goddard (Université de Californie San Francisco). Les principales expériences réalisées sont :

- 2D $^{1}H/^{15}N$ HSQC.
- 2D $^{1}H/^{13}C$ HSQC.
- 3D $^{1}H/^{13}C/^{15}N$ HNCA.
- 3D $^{1}H/^{13}C/^{15}N$ HN(C0)CA.
- 3D $^{1}H/^{13}C/^{15}N$ HNC0.
- 3D $^{1}H/^{13}C/^{15}N$ HNCACB.
- 3D $^{1}H/^{15}N/^{1}H$ NOEHSQC

La plupart des séquences d'impulsions utilisées ont été écrites par John K. Smith (Université du Michigan, USA) et éditées François Bontems. Elles sont disponibles sur le serveur http://lupin.polytechnique.fr.

Toutes les expériences ont été réalisées à 32.5 °C et à pH 6.0. Par soucis de reproductibilité, les paramètres RMN optimisés pour la protéine RegB H48A sont présentés, dans leur totalité, ci-dessous.

***2D $^{1}H/^{15}N$ HSQC** (Microprogramme FB_fhsqcf3)*

Acquisition ^{1}H

Dummy scans	16
scans	256
receiver gain	2048
td	500
sw	5.007 ppm
sw	3004.81 Hz
sfo	600.1348MHz
nucleus	1H
o1	7.998 ppm
o1	4800 Hz
dw	166.40 usec
aq_mod	dqd
digmod	digital
dspfvs	12
dspfirm	0
decim	64
dwov	2.60 usec

Acquisition ^{15}N

sw	36 ppm
sw	2189.447 Hz
sfo	60.8179 MHz
nucleus	15N
o1	120.045 ppm
o1	7300 Hz
td	300
nd0	2

Impulsions

p0 11.25 usec at pl1 f1
p1 11.25 usec at pl1 f1
p21 32.00 usec at pl3 f3
p25 1.00 msec gradient pulse suppression 15N
p26 2.27 msec gradient pulse suppression RD
p27 600.00 usec gradient pulse purge IzSz
p28 800.00 usec gradient pulse watergate
pcpd2 1.50 msec at pl12 on f2
pcpd3 250.00 usec at pl13 on f3
f3 channel - 90 degree pulse for decoupling sequence

Niveaux de Puissances

pl1 -3 db on f1
pl3 3 db on f3
pl12 20.39 db on f2 power level for CPD/BB decoupling
pl13 20 db on f3 for CPD/BB decoupling
sp15 120 db on f2

Délais

d0 78.28 usec
d1 1.00 sec
d4 2.50 msec
d11 30.00 delay for disk I/O
d12 20.00 usec delay for power switching
d15 200.00 usec delay for gradient recovery
d27 42.00 usec "d27=d28/2-p21"
d28 148.00 usec delay for binomial water suppression.

Autres Paramètres

cpdprg2 t5m4_pcpd_sp15
cpdprg3 garp
in0 0.0002348875
FQ1LIST 2824.51 : 4800

2D $^{1}H/^{13}C$ HSQC (Microprogramme HSQC13Cpr2)

Acquisition ^{1}H

dummy scans	16
scans	128
receiver gain	2048
td	1024
sw	12.977 ppm
sw	7788.16 Hz
sfo	600.132 MHz
nucleus	1H
o1	4.706 ppm
o1	4800 Hz
dw	64.20 usec
aq_mod	dqd
digmod	digital
dspfvs	12
dspfirm	0
decim	24
dwov	2.68 usec

Acquisition ^{13}C

sw	90.003 ppm
sw	13582.343 Hz
sfo	150.909 MHz
nucleus	13C
o1	44 ppm
o1	6639.72 Hz
l3	98
d0	3.00 usec
in0	36.81 usec

Impulsions

p1 12.20 usec at pl1 on f1
p2 11.00 usec at pl2 on f2
p12 680.00 usec using
sec_hyp_1k at sp12 on f2
p16 1.50 msec gradient pulse
pcpd2 140.00 usec at pl12 on f2

Niveaux de Puissance

pl0 120 db on f2
pl1 -3 db on f1
pl2 0 db on f2
pl9 52 db on f1
pl12 18 db on f2 for CPD/BB decoupling
sp12 0 db on f2

Délais

d1 2.00 sec
d4 1.70 msec
d11 30.00 msec delay for disk I/O
d12 20.00 usec
d13 3.00 usec
d16 100.00 usec

Autres Paramètres

cpdprg2 garp

3D $^{1}H/^{13}C/^{15}N$ HNCA (Microprogramme FB_hnca_cbd)

Acquisition ^{1}H

dummy scans	32
scans	16
receiver gain	2048
td	512
sw	5.007 ppm
sw	3004.81 Hz
sfo	600.134 MHz
nucleus	1H
o1	7.998 ppm
o1	4800 Hz
dw	166.40 usec
aq_mod	dqd
digmod	digital
dspfvs	12
dspfirm	0
decim	64
dwov	2.60 usec

Acquisition ^{15}N

sw	27.000 ppm
sw	1642.104 Hz
sfo	60.817 MHz
nucleus	15N
o1	119.5 ppm
o1	7266.87 Hz
l13	25
d10	3.00 usec
d30	7.97 msec
in10	304.49 usec
in30	304.49 usec

Acquisition ^{13}C

sw	30.004 ppm
sw	4527.960 Hz
sfo	150.911 MHz
nucleus	13C
o1	56.9 ppm
o1	8588 Hz
l3	105
d0	92.63 usec
in0	110.42 usec

Impulsions

p1 11.00 usec at pl1 on f1
p5 32.00 usec at pl3 on f3
p6 65.52 usec at pl19 on f1
p11 3.50 msec using
esnob_3500 at sp1 on f1
p13 11.00 usec using
rect_100_100 at sp6 on f2
p14 22.00 usec using
rect_100_100 at sp7 on f2
pcpd1 65.52 usec at pl19 on f1
pcpd2 1.50 msec at pl12 on f2
pcpd3 250.00 usec at pl16 on f3
pcpd7 1.50 msec at pl7 on f2

Niveaux de Puissance

pl0 120 db on f3
pl1 -3 db on f1
pl2 120 db on f2
pl3 3 db on f3
pl7 31.27 db on f2
pl12 31.27 db on f2
pl16 17 db on f3
pl19 12.5 db on f1
sp1 47.05 db on f1
sp6 0 db on f2
sp7 0 db on f2
sp15 120 db on f2

Délais

d1 1.00 sec
d11 30.00 msec
d24 8.00 msec

Autres paramètres

cpdprg1 waltz16a
cpdprg2 t5m4_pcpd_sp15
cpdprg3 garp
cpdprg7 t5m4_pcpd_sp15

FQ1LIST 2824 : 4800 (1H)

3D $^{1}H/^{13}C/^{15}N$ HN(C0)CA (Microprogramme EKhncoca_cbd)

Acquisition ^{1}H

dummy scans	8
scans	16
receiver gain	2048
td	512
sw	5.007 ppm
sw	3004.81 Hz
sfo	600.134 MHz
nucleus	1H
o1	7.998 ppm
o1	4800 Hz
dw	166.40 usec
aq_mod	dqd
digmod	digital
dspfvs	12
dspfirm	0
decim	64
dwov	2.60 usec

Acquisition ^{15}N

sw	27.000 ppm
sw	1642.104 Hz
sfo	60.817 MHz
nucleus	15N
o1	119.5 ppm
o1	7266.87 Hz
l13	25
d10	3.00 usec
d30	12.35 msec
in10	304.49 usec
in30	304.49 usec

Acquisition ^{13}C

sw	30.001 ppm
sw	4527.448 Hz
sfo	150.911 MHz
nucleus	13C
o1	56.911 ppm
o1	8588 Hz
l3	105
d0	67.35 usec
in0	110.44 usec

Impulsions

p1 11.50 usec at pl1 on f1
p5 32.00 usec at pl3 on f3
p6 68.52 usec at pl19 on f1
p11 3.50 msec using esnob_3500 at sp1 on f1
p13 53.00 usec using rect_100_100 at sp6 on f2
p14 53.00 usec using FS_hncoa_recof at sp7 on f2
p15 53.00 usec using rect_100_100 at sp8 on f2
p16 230.00 usec using FS_hncoa_sedon at sp9 on f2
p17 230.00 usec using FS_hncoa_sedof at sp10 on f2
pcpd1 68.52 usec at pl19 on f1
pcpd2 1.50 msec at pl12 on f2
pcpd3 250.00 usec at pl16 on f3

Niveaux de Puissance

pl0 120 db on f3
pl1 -3 db on f1
pl2 120 db on f2
pl3 3 db on f3
pl12 25.29 db on f2
pl16 17 db on f3
pl19 12.5 db on f1
sp1 47.05 db on f1
sp6 15.2 db on f2
sp7 9.2 db on f2
sp8 15.2 db on f2
sp9 13.7 db on f2
sp10 13.7 db on f2
sp15 120 db on f2

Délais

d1 1.00 sec
d11 30.00 msec
d24 12.40 msec

Autres Paramètres

cpdprg1 waltz16a
cpdprg2 t5m4_pcpd_sp15
cpdprg3 garp
td1 210
td2 210
FQ1LIST 2629 : 8588 :26293 (13C)

3D $^{1}H/^{13}C/^{15}N$ HNCO (Microprogramme FB_hnca_cbd)

Acquisition ^{1}H

dummy scans	16
scans	32
receiver gain	2048
td	512
sw	5.007 ppm
sw	3004.81 Hz
sfo	600.134 MHz
nucleus	1H
o1	7.998 ppm
o1	4800 Hz
dw	166.40 usec
aq_mod	dqd
digmod	digital
dspfvs	12
dspfirm	0
decim	64
dwov	2.60 usec

Acquisition ^{15}N

sw	27.000 ppm
sw	1642.104 Hz
sfo	60.817 MHz
nucleus	15N
o1	119.5 ppm
o1	7266.87 Hz
l13	23
d10	3.00 usec
d30	7.97 msec
in10	304.49 usec
in30	304.49 usec

Acquisition ^{13}C

sw	19.997 ppm
sw	3018.185 Hz
sfo	150.929 MHz
nucleus	13C
centerOi	174.94 ppm
o1	26405 Hz
l3	64
d0	175.49 usec
in0	165.66 usec

Impulsions

p1 11.05 usec at pl1 on f1
p5 32.00 usec at pl3 on f3
p6 64.33 usec at pl19 on f1
p11 3.50 msec using
esnob_3500 at sp1 on f1
p13 11.00 usec using
rect_100_100 at sp6 on f2
p14 22.00 usec using
rect_100_100 at sp7 on f2
pcpd1 64.33 usec at pl19 on f1
pcpd2 1.50 msec at pl12 on f2
pcpd3 250.00 usec at pl16 on f3
pcpd7 100.00 usec at pl7 on f2

Niveaux de Puissance

pl0 120 db on f3
pl1 -3 db on f1
pl2 120 db on f2
pl3 3 db on f3
pl7 25.29 db on f2
pl12 25.29 db on f2
pl16 17 db on f3
pl19 12.5 db on f1
sp1 30.8 db on f1
sp6 0 db on f2
sp7 0 db on f2
sp15 120 db on f2

Délais

d1 1.00 sec
d11 30.00 msec
d24 8.00 msec

Autres Paramètres

cpdprg1 waltz16a
cpdprg2 t5m4_pcpd_sp15
cpdprg3 garp
cpdprg7 t5m4_pcpd_sp15
FQ1LIST 2824 : 4800 (1H)

3D $^{1}H/^{13}C/^{15}N$ HNCACB (Microprogramme FB_hnca_cbd)

Acquisition ^{1}H

dummy scans	16
scans	32
receiver gain	2048
td	512
sw	5.510 ppm
sw	3306.88 Hz
sfo	600.134 MHz
nucleus	1H
o1	7.998 ppm
o1	4800 Hz
dw	151.20 usec
aq_mod	dqd
digmod	digital
dspfvs	12
dspfirm	0
decim	48
dwov	3.15 usec

Acquisition ^{15}N

sw	27.000 ppm
sw	1642.104 Hz
sfo	60.817 MHz
nucleus	15N
o1	119.5 ppm
o1	7266.87 Hz
l13	24
d10	3.00 usec
d30	11.97 msec
in10	304.49 usec
in30	304.49 usec

Acquisition ^{13}C

sw	75.457 ppm
sw	11318.619 Hz
sfo	150 MHz
nucleus	13C
o1	41ppm
o1	6187 Hz
l3	64
d0	34.44 usec
in0	44.18 usec

Impulsions

p1 12.60 usec at pl1 on f1
p5 35.00 usec at pl3 on f3
p6 50.20 usec at pl19 on f1
p11 3.50 msec using
esnob_3500 at sp1 on f1
p13 11.00 usec using
rect_100_100 at sp6 on f2
p14 22.00 usec using
rect_100_100 at sp7 on f2
pcpd1 50.20 usec at pl19 on f1
pcpd2 1.50 msec at pl12 on f2
pcpd3 250.00 usec at pl16 on f3

Niveaux de Puissance

pl0 120 db on f3
pl1 -3 db on f1
pl2 120 db on f2
pl3 3 db on f3
pl12 27.35 db on f2
pl16 20 db on f3
pl19 9 db on f1
sp1 45.9 db on f1
sp6 0 db on f2
sp7 0 db on f2
sp15 120 db on f2

Délais

d11 30.00 msec
d24 12.00 msec

Autres Paramètres

cpdprg1 waltz16a
cpdprg2 t5m4_pcpd_sp15
cpdprg3 garp
FQ1LIST 2824.51 : 4800 (1H)

3D $^{1}H/^{15}N/^{1}H$ NOEHSQC (Microprogramme FB_noefhsqcf3)

Acquisition ^{1}H

dummy scans 64

scans	32
receiver gain	2048
td	600
sw	5.007 ppm
sw	3004.81 Hz
sfo	600.134 MHz
nucleus	1H
o1	7.998 ppm
o1	4800 Hz
dw	166.40 usec
aq_mod	dqd
digmod	digital
dspfvs	12
dspfirm	0
decim	64
dwov	2.60 usec

Acquisition ^{15}N

sw	35.001 ppm
sw	2128.679 Hz
sfo	60.817 MHz
nucleus	15N
o1	118.4 ppm
o1	7199.1 Hz
l2	19
d10	85.32 usec
in10	234.89 usec

Acquisition ^{1}H

sw	11.971 ppm
sw	7184.2 Hz
sfo	600.13 MHz
nucleus	1H
o1	8.000 ppm
o1	4800 Hz
td	150

Impulsions

p1 11.75 usec at pl1 on f1
p3 32.00 usec at pl3 on f3
pcpd3 250.00 usec at pl16 on f3

Niveaux de Puissance

pl0 120 db on f3
pl1 -3 db on f1
pl3 3 db on f3
pl16 17 db on f3

Délais

d1 1.00 sec
d11 30.00 msec delay for disk I/O

Autres Paramètres

cpdprg3 garp
td1 38
td2 38
FQ1LIST 2824.2 : 4800 (1H)

Références bibliographiques

Références bibliographiques

Ackermann, H. W. (1998). Tailed bacteriophages: the order Caudovirales. *Adv. Virus Res.*, **51**, 135-201.

Ackermann, H.W. et Krisch, H.M. (1997) A catalogue of T4-type bacteriophages. *Arch. Virol.*, **142**, 2329-2345.

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E. et Tollervey, D. (1999) Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J.*, **18**, 5399-5410.

Allmang, C., Petfalski, E., Podtelejnikov, A., Mann, M., Tollervey, D. et Mitchell, P. (1999) The yeast exosome and human PM-Scl are related complexes of 3' --> 5' exonucleases. *Genes Dev.*, **13**, 2148-2158.

Aloy, P., Ciccarelli, F.D., Leutwein, C., Gavin, A.C., Superti-Furga, G., Bork, P., Bottcher, B. et Russell, R.B. (2002) A complex prediction: three-dimensional model of the yeast exosome. *EMBO Rep.*, **3**, 628-635.

Altman, S. et Kirsebom L.A. (1999) RNA World II. Gesteland, R., Cech, T. et J. Atkins (éditeurs) *Cold Spring Harbor Press.*

Aravind, L. et Koonin, E.V. (2001) A natural classification of ribonucleases. *Methods Enzymol*, **341**, 3-28.

Au, K.G., Clark, S., Miller, J.H. et Modrich, P. (1989) Escherichia coli mutY gene encodes an adenine glycosylase active on G-A mispairs. *Proc. Natl. Acad. Sci. USA*, **86**, 8877-8881.

Beintema, J.J. (1998) Introduction: the ribonuclease A superfamily. *Cell. Mol. Life Sci.*, **54**, 763-765.

Beintema, J.J., Breukelman, H.J., Carsana, A. et Furia, A. (1997) Evolution of vertebrate ribonucleases: Ribonuclease A superfamily. Dans D'Allessio, G. et Riordan,J.F. (ed.), *Ribonucleases structures and functions*,

Academic Press, pp. 245-269.

Bernard, P., Gabant, P., Bahassi, E.M., et Couturier, M. (1994) Positive selection vetors using the F plasmid ccdB killer gene. *Gene*, **148**, 71-74.

Berghammer, H. et Auer, B. (1993) "Easypreps": fast and easy plasmid minipreparation for analysis of recombinant clones in E. coli. *Biotechniques*, **14**, 524-528.

Bernstein, E., Caudy, A.A., Hammond, S.M. et Hannon, G.J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, **409**, 363–366.

Bisaglia, M., Laalami, S., Uzan, M. et Bontems, F. (2003) Activation of the RegB endoribonuclease by the S1 ribosomal protein is due to cooperation between the S1 four C-terminal modules in a substrate-dependant manner. *J. Biol. Chem.*, **278**, 15261-15271.

Blanc, H., Gerbaud, C., Slonimski, P.P. et Guerineau, M. (1979) Stable yeast transformation with chimeric plasmids using a 2 micron-circular DNA-less strain as a recipient. *Mol. Gen. Genet.*, **176**, 335-342.

Blum, E., Py, B., Carpousis, A.J. et Higgins, C.F. (1997) Polyphosphate kinase is a component of the Escherichia coli RNA degradosome. *Mol. Microbiol.*, **26**, 387-398.

Bolivar, F., Rodriguez, R.L., Betlach, M.C. et Boyer, H.W. (1977) Construction and characterization of new cloning vehicles. I. Ampicillin-resistant derivatives of the plasmid pMB9. *Gene*, **2**, 75-93.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L. et Boyer, H.W. (1977) Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene*, **2**, 95-113.

Bouet, J.Y., Krisch, H.M. et Louarn, J.M. (1998) Ndd, the bacteriophage T4 protein that disrupts the Escherichia coli nucleoid, has a DNA binding activity. *J. Bacteriol.*, **180**, 5227-5230.

Bousquet-Antonelli, C., Presutti, C. et Tollervey, D. (2000) Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell*, **102**, 765-775.

Brendel, V. et Perelson, A.S. (1993) Quantitative model of ColE1 plasmid copy number control. *J. Mol. Biol.*, **229**, 860-872.

Browne, D.T., Kenyon, G.L., Packer, E.L., Sternlicht, H., et Wilson, D.M. (1973) Studies of macromolecular structure by carbon-13 nuclear magnetic resonance. Specific labeling approach to the study of histidine residues in proteins. *J. Am. Chem. Soc.*, **95**, 1316-1323.

Burkard, K.T. et Butler, J.S. (2000) A nuclear 3'-5' exonuclease involved in mRNA degradation interacts with Poly(A) polymerase and the hnRNA protein Npl3p. *Mol. Cell. Biol.*, **20**, 604-616.

Busby, S., Irani, M. et Crombrugghe, B. (1982) Isolation of mutant promoters in the Escherichia coli galactose operon using local mutagenesis on cloned DNA fragments. *J. Mol. Biol.*, **154**, 197-209.

Cech, T.R., Zaug, A.J. et Grabowski, P.J. (1981) In vitro splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell*, **27**, 487-496.

Chace, K.V. et Hall, D.H. (1975) Characterization of new regulatory mutants for bacteriophage T4. New class of mutants. *J. Virol.*, **15**, 929-945.

Chamberlin, M., McGrath, J. et Waskell, L. (1970) New RNA polymerase from Escherichia coli infected with

bacteriophage T7. *Nature*, **228**, 227-231.

Chatani, E., Tanimizu, N., Ueno, H. et Hayashi, R. (2001) Structural and functional changes in bovine pancreatic ribonuclease a by the replacement of Phe120 with other hydrophobic residues. *J. Biochem.*, **129**, 917-922.

Chevrier-Miller, M., Jacques, N., Raibaud, O. et Dreyfus, M. (1990) Transcription of single-copy hybrid lacZ genes by T7 RNA polymerase in Escherichia coli: mRNA synthesis and degradation can be uncoupled from translation. *Nucleic Acids Res.*, **18**, 5787-92.

Chin, F., Banaszczyk, F., Jubian, V. et Zou, X. (1989) Co(III) complex promoted hydrolysis of phospho diesters: comparison in reactivity of rigid cis-diaquotetraazacobalt(III) complexes. *J. Am. Chem. Soc.*, **111**, 186-190.

Choi, Y.J., Wang, T.T. et Lee, B.H. (2002) Positive selection vectors. *Crit. Rev. Biotechnol.*, **22**, 225-244.

Cline, J., Braman, J.C. et Hogrefe, H.H. (1996) PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res.*, **24**, 3546-51.

Coburn, G.A. et Mackie G.A. (1998) Reconstitution of the degradation of the mRNA for ribosomal protein S20 with purified enzymes. *J. Mol. Biol.*, **279**, 1061-1074.

Cohen, S.N. et McDowall, K.J. (1997) RNase E: still a wonderfully mysterious enzyme. *Mol. Microbiol.*, **23**, 1099-1106.

Condon, C., et Putzer, H. (2002) The phylogenetic distribution of bacterial ribonucleases. *Nucleic Acids Res*, **30**, 5339-5346.

Court, D. (1993) Control of Messenger RNA Stability. Belasco, J.G. et Brawerman, G. (éditeurs) *Academic Press, New York.*

Cowie, D.B., Avery, R.J., et Champe, S.P. (1971) DNA homology among the T-even bacteriophages. *Virology*, **45**, 30-37.

Crary, S.M., Niranjanakumari, S., et Fierke, C.A. (1998) The protein component of Bacillus subtilis ribonuclease P increases catalytic efficiency by enhancing interactions with the 5' leader sequence of pre-tRNAAsp. *Biochemistry*, **37**, 9409-9416.

Cuchillo, C.M., Pares, X., Guasch, A., Barman, T., Travers, F. et Nogues, M.V. (1993) The role of 2',3'-cyclic phosphodiesters in the bovine pancreatic ribonuclease A catalysed cleavage of RNA: intermediates or products ? *FEBS Lett.*, **333**, 207-210.

D'Alessio, G. et Riordan, J.F. (1997). Ribonucleases structures and functions. *Academic Press.*

Decout J.L. et Maurel, M.C. (1993) N6-substituted adenine derivatives and RNA primitive catalysts. *Orig. Life Evol. Biosph.*, **23**, 299-306.

Desplats, C., Dez, C., Tetart, F., Eleaume, H. et Krisch, H.M. (2002) Snapshot of the genome of the pseudo-T-even bacteriophage RB49. *J. Bacteriol.*, **184**, 2789-2804.

Deyev, S.M., Yazynin, S.A., Kuznetsov, D.A., Jucovic, M., et Hartley, R.W. (1998) Ribonuclease-charged vector for facile direct cloning with positive selection. *Mol. Gen. Genet.*, **259**, 379-382.

Diener, T.O. (1991) Subviral pathogens of plants: viroids and viroidlike satellite RNAs. *FASEB J.*, **5**, 2808-2813.

Dunn, J.J. et Studier, F.W. (1973) T7 early RNAs and Escherichia coli ribosomal RNAs are cut from large precursor RNAs in vivo by ribonuclease 3. *Proc. Natl. Acad. Sci. USA*, **70**, 3296-3300.

Dunn, J.J. et Studier, F.W. (1983) Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.*, **166**, 477-535.

Ehretsmann, C.P., Carpousis, A.J., et Krisch, H.M. (1992) Specificity of Escherichia coli endoribonuclease RNase E: in vivo and in vitro analysis of mutants in a bacteriophage T4 mRNA processing site. *Genes Dev*, **6**, 149-159.

Fedoroff, O.Y., Salazar, M. et Reid, B.R. (1993) Structure of a DNA:RNA hybrid duplex. Why RNase H does not cleave pure RNA. *J. Mol. Biol.*, **233**, 509-523.

Forster, A.C. et Symons, R.H. (1987) Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites. *Cell*, **49**, 211-220.

Fromme, J.C. et Verdine, G.L. (2003) Structure of a trapped endonuclease III-DNA covalent intermediate. *EMBO J.*, **22**, 3461-3471.

Fuxreiter, M., Warshel, A. et Osman, R. (1999) Role of active site residues in the glycosylase step of T4 endonuclease V. Computer simulation studies on ionization states. *Biochemistry*, **38**, 9577-9589.

Gerlt, J.A. (1993) Mechanistic principles of enzyme-catalyzed cleavage of phosphodiester bonds. Nucleases, *Cold Spring Harbor Laboratory Press*, pp. 1-34.

Gorenstein, D.G. (1984) Phosphorus-31 NMR: Principles and Applications. *Academic Press*.

Grzesiek, S. et Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly 13C/15N-enriched proteins. *J. Biomol. NMR*, **3**, 185-204.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. et Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, **35**, 849-57.

Guillet, V., Lapthorn, A. et Mauguen, Y. (1993) Three-dimensional structure of a barnase-3'GMP complex at 2.2A resolution. *FEBS Lett.*, **330**, 137-140.

Gustafsson, P., Dreisig, H., Molin, S., Nordstrom, K., Uhlin, B.E. (1979) DNA replication control: studies of plasmid R1. *Cold Spring Harb. Symp. Quant. Biol.*, **1**, 419-425.

Hale, S.P., Poole, L.B. et Gerlt, J.A. (1993) Mechanism of the reaction catalyzed by staphylococcal nuclease: identification of the rate-determining step. *Biochemistry*, **32**, 7479-7487.

Hartley, R.W. (1969) Some environmental effects on the thermal transition of Bacillus amyloliquefaciens ribonuclease (barnase). *Biochemistry*, **8**, 2929-2932.

Hertel, K.J., Pardi, A., Uhlenbeck, O.C., Koizumi, M., Ohtsuka, E., Uesugi, S., Cedergren, R., Eckstein, F., Gerlach, W.L., Hodgson, R., *et al.* (1992) Numbering system for the hammerhead. *Nucleic Acids Res.*, **20**, 3252.

Hostomsky, Z., Hostomska, Z. et Matthews, D.A. (1993) Ribonucleases H. Linn, S.M., Lloyd, R.S. et Roberts, R.J. (ed.), *Nucleases*, 2ème edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY., pp. 341–376.

Huberman, J.A., Spotila, L.D., Nawotka, K.A., el-Assouli, S.M. et Davis, L.R. (1987) The in vivo replication

origin of the yeast 2 microns plasmid. *Cell*, **51**, 473-481.

Hutchins, C.J., Rathjen, P.D., Forster, A.C. et Symons, R.H. (1986) Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid. *Nucleic Acids Res.*, **14**, 3627-3640.

Irie, M. (1997) Escherichia coli Ribonucleases: Paradigms for understanding cellular RNA metabolism and regulation. D'Allessio,G. and Riordan,J.F. (ed.), *Ribonucleases structures and functions*, Academic Press, p. 101.

Irie, M. et Ohgi, K. (2001) Ribonuclease T2. *Methods Enzymol.*, **341**, 42-55.

Ishii, R., Nureki, O. et Yokoyama, S. (2003) Crystal structure of the tRNA processing enzyme RNase PH from Aquifex aeolicus. J. *Biol. Chem.*, **278**, 32397-32404.

Ishikawa, K., Suzuki, E., Tanokura, M. et Takahashi, K. (1996) Crystal structure of ribonuclease T1 carboxymethylated at Glu58 in complex with 2'-GMP. *Biochemistry*, **35**, 8329-8334.

Jacobs, J.S., Anderson, A.R. et Parker, R.P. (1998) The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *EMBO J.*, **17**, 1497-1506.

Jarrige, A., Brechemier-Baey, D., Mathy, N., Duche, O. et Portier, C. (2002) Mutational analysis of polynucleotide phosphorylase from Escherichia coli. *J. Mol. Biol.*, **321**, 397-409.

Jayasena, V.K., Brown D., Shtatland, T. et Gold, L. (1996) *In vitro* selection of RNA specifically cleaved by bacteriophage T4 RegB endonucléases. *Biochemistry*, **35**, 2347-2356.

Kaberdin, V.R., Miczak, A., Jakobsen, J.S., Lin-Chao, S., McDowall, K.J. et von Gabain, A. (1998) The endoribonucleolytic N-terminal half of Escherichia coli RNase E is evolutionarily conserved in Synechocystis sp. and other bacteria but not the C-terminal half, which is sufficient for degradosome assembly. *Proc. Natl. Acad. Sci. USA*, **95**, 11637-11642.

Karam, J.D., Drake, J.W. et Kreuzer, K.N. (1994) Molecular Biology of Bacteriophage T4. Karam, J. (ed.), *American Society for Microbiology, Washington, D.C.*

Katsamba, P.S., Park, S. et Laird-Offringa, I.A. (2002) Kinetic studies of RNA-protein interactions using surface plasmon resonance. *Methods*, **26**, 95-104.

Kharrat, A., Macias, M.J., Gibson, T.J., Nilges, M. et Pastore, A. (1995) Structure of the dsRNA binding domain of E. coli RNase III. *EMBO J.*, **14**, 3572-3584.

Kirsebom, L.A. (2001) Escherichia coli ribonuclease P. *Methods Enzymol.*, **342**, 77-92.

Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, **44**, 283-292.

Kozak, M. (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene*, **234**, 187-208.

Kufel, J., Allmang, C., Chanfreau, G., Petfalski, E., Lafontaine, D.L., Tollervey, D. (2000) Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding. *Mol.Cell. Biol.*, **20**, 5415-5424.

Kurz, J.C., Niranjanakumari, S. et Fierke, C.A. (1998) Protein component of Bacillus subtilis RNase P specifically enhances the affinity for precursor-tRNAAsp. *Biochemistry*, **37**, 2393-2400.

Langhorst, U., Loris, R., Denisov, V.P., Doumen, J., Roose, P., Maes, D., Halle, B. et Steyaert J. (1999) Dissection of the structural and functional role of a conserved hydration site in RNase T1. *Protein Sci.*, **8**, 722-730.

Larsen, J.E., Gerdes, K., Light, J. et Molin, S. (1984) Low-copy-number plasmid-cloning vectors amplifiable by derepression of an inserted foreign promoter. *Gene*, **28**, 45-54.

Lebars, I., Hu, R.M., Lallemand, J.Y., Uzan, M. et Bontems, F. (2001) Role of the substrate conformation and of the S1 protein in the cleavage efficiency of the T4 endoribonuclease RegB. *J. Biol. Chem.*, **276**, 13264-13272.

Loll, P.J. et Lattman, E.E. (1989) The crystal structure of the ternary complex of staphylococcal nuclease, Ca2+, and the inhibitor pdTp, refined at 1.65 A. *Proteins*, **5**, 183-201.

Masaki, H. et Ogawa, T. (2002) The modes of action of colicins E5 and D, and related cytotoxic tRNases. *Biochimie*, **84**, 433-438.

Maurel, M.C. (1994) Les origines de la vie. *Editions Syros*, Paris.

McAllister, W.T. et Carter, A.D. (1980) Regulation of promoter selection by the bacteriophage T7 RNA polymerase in vitro. *Nucleic Acids Res.*, **8**, 4821-4837.

McAllister, W.T., Morris, C., Rosenberg, A.H. et Studier, F.W. (1981) Utilization of bacteriophage T7 late promoters in recombinant plasmids during infection. *J. Mol. Biol.*, **153**, 527-544.

McKendrick, L. (2003) What's new in translation initiation? The first translation determines the fate of mRNA. *Cell. Mol. Life Sci.*, **60**, 639-647.

Miczak, A., Kaberdin, V.R., Wei, C.L. et Lin-Chao, S. (1996) Proteins associated with RNase E in a multicomponent ribonucleolytic complex. *Proc. Natl. Acad. Sci. USA*, **93**, 3865-2869.

Mitchel, R.E. et Morrison, D.P. (1986) Inducible error-prone repair in yeast. Suppression by heat shock. *Mutat. Res.*, **159**, 31-39.

Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. et Tollervey, D. (1997) The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. *Cell*, **91**, 457-466.

Monod, C., Repoila, F., Kutateladze, M., Tétart, F., et Krisch. H. M. (1997) The genome of the pseudo-T-even bacteriophage, a diverse group that resembles T4. *J. Mol. Biol.*, **267**, 237-249.

Mori, S., Abeygunawardana, C., Johnson, M.O., Berg, J., et van Zijl, P.C.M. (1995) Improved Sensitivity of HSQC Spectra of Exchanging Protons at Short Interscan Delays Using a New Fast HSQC (FHSQC) Detection Scheme that Avoids Water Saturation. *J. Magn. Reson. Series B*, **108**, 94-98.

Mossakowska, D.E., Nyberg, K. et Fersht, A.R. (1989) Kinetic characterization of the recombinant ribonuclease from Bacillus amyloliquefaciens (barnase) and investigation of key residues in catalysis by site-directed mutagenesis. *Biochemistry*, **28**, 3843-3850.

Munishkin, A. et Wool, I.G. (1997) The ribosome-in-pieces: binding of elongation factor EF-G to oligoribonucleotides that mimic the sarcin/ricin and thiostrepton domains of 23S ribosomal RNA. *Proc. Natl. Acad. Sci. USA*, **94**, 12280-12284.

Myers, R.M., Lerman, L.S. et Maniatis, T. (1985) A general method for saturation mutagenesis of cloned DNA fragments. *Science*, **229**, 242-247.

Nghiem, Y., Cabrera, M., Cupples, C.G. et Miller, J.H. (1988) The mutY gene: a mutator locus in Escherichia coli that generates G.C—T.A transversions. *Proc. Natl. Acad. Sci. USA*, **85**, 2709-2713.

Nicholson, A.W. (1999) Function, mechanism and regulation of bacterial ribonucleases. *FEMS Microbiol. Rev.*, **23**, 371-390.

Niranjanakumari, S., Stams, T., Crary, S.M., Christianson, D.W. et Fierke, C.A. (1998) Protein component of the ribozyme ribonuclease P alters substrate recognition by directly contacting precursor tRNA. *Proc. Natl. Acad. Sci. USA*, **95**, 15212-15217.

Numata, T. et Kimura, M. (2001) Contribution of Gln9 and Phe80 to substrate binding in ribonuclease MC1 from bitter gourd seeds. *J. Biochem.*, **130**, 621-626.

Oda, Y., Yoshida, M. et Kanaya, S. (1993) Role of histidine 124 in the catalytic function of ribonuclease HI from Escherichia coli. *J. Biol. Chem.*, **268**, 88-92.

Ogawa, T., Tomita, K., Ueda, T., Watanabe, K., Uozumi, T. et Masaki, H. (1999) A cytotoxic ribonuclease targeting specific transfer RNA anticodons. *Science*, **283**, 2097-2100.

Ohgi, K., Horiuchi, H., Watanabe, H., Iwama, M., Takagi, M. et Irie, M. (1992) Evidence that three histidine residues of a base non-specific and adenylic acid preferential ribonuclease from Rhizopus niveus are involved in the catalytic function. *J. Biochem.*, **112**, 132-138.

Orosz, A., Boros, I. et Venetianer, P. (1991) Analysis of the complex transcription termination region of the Escherichia coli rrnB gene. *Eur. J. Biochem.*, **201**, 653-659.

Osterman, H.L. et Walz, F.G. Jr. (1979) Subsite interactions and ribonuclease T1 catalysis: kinetic studies with

APGpC and ApGpU. *Biochemistry*, **18**, 1984-1988.

Pompon, D., Louerat, B., Bronine, A. et Urban, P. (1996) Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol.*, **272**, 51-64.

Rakesh, C. et al. (1996). *Biotechniques*, **20**, 42-44.

Regnier, P. et Arrainano, C.M. (2000) Degradation of mRNA in bacteria: emergence of ubiquitous factors. *Bioessays*, **22**, 235-244.

Reich, C., Olsen, G.J., Pace, B. et Pace, N.R. (1988) Role of the protein moiety of ribonuclease P, a ribonucleoprotein enzyme. *Science*, **239**, 178-181.

Robzyk, K. et Kassir, Y. (1992) A simple and highly efficient procedure for rescuing autonomous plasmids from yeast. *Nucleic Acids Res.*, **20**, 3790.

Rodriguez, R., Lopez-Otin, C., Barber, D., Fernandez-Luna, J.L., Gonzalez, G. et Mendez, E. (1982) Amino acid sequence homologies in alfa-sarcin, restrictocin and mitogillin. *Biochem. Biophys. Res. Commun.*, **108**, 315-321.

Rosenberg, A.H., Lade, B.N., Chui, D.S., Lin, S.W., Dunn, J.J. et Studier, F.W. (1987) Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene*, **56**, 125-135.

Ruckman, J., Parma, D., Tuerk, C., Hall, DH. et Gold, L. (1989) Identification of a T4 gene required for bacteriophage mRNA processing. *New Biol.*, **1**, 54-65.

Ruckman, J., Ringquist, S., Brody, E. et Gold, L. (1994) The bacteriophage T4 RegB ribonuclease. Stimulation

of the purified enzyme by ribosomal protein S1. *J. Biol. Chem.*, **269**, 26655-26662.

Ryter, J.M. et Schultz, S.C. (1998) Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J.*, **17**, 7505-7513.

Sanson, B. et Uzan, M. (1995) Post-transcriptionnal controls in bacteriophage T4 : roles of the sequence specific endoribonuclease RegB. *FEMS Microbiol. Rev.*, **17**, 141-50.

Sanson, B., Hu, R.M., Troitskaya, E., Mathy, N. et Uzan, M. (2000) Endoribonuclease RegB from bacteriophage T4 is necessary for the degradation of early but not middle or late mRNAs. *J. Mol. Biol.*, **297**, 1063-1074.

Sato, K. et Egami, F. (1957). Studies on ribonuclease in Takadiastase I. *J. Biochem.*, **44**, 753-767.

Sattler, M., Schuleucher, J., et Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **34**, 93-158.

Schaaper, R.M. (1989) Escherichia coli mutator mutD5 is defective in the mutHLS pathway of DNA mismatch repair. *Genetics*, **121**, 205-212.

Scott, W.G., Finch, J.T. et Klug, A. (1995) The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell*, **81**, 991-1002.

Seal, G., Shearman, C.W. et Loeb, L.A. (1979) On the fidelity of DNA replication. Studies with human placenta DNA polymerases. *J. Biol. Chem.*, **254**, 5229-5237.

Sevcik, J., Sanishvili, R.G., Pavlovsky, A.G. et Polyakov, K.M. (1990) Comparison of active sites of some microbial ribonucleases: structural basis for guanylic specificity. *Trends Biochem. Sci.*, **15**, 158-162.

Shine, J. et Dalgarno, L. (1975) The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. USA*, **71**, 1342-1346.

Sloan, D.L., Loeb, L.A. et Mildvan, A.S. (1975) Conformation of deoxynucleoside triphosphate substrates on DNA polymerase I from Escherichia coli as determined by nuclear magnetic relaxation. *J. Biol. Chem.*, **250**, 8913-8920.

Snoussi, K. et Leroy, J.L. (2001) Imino proton exchange and base-pair kinetics in RNA duplexes. *Biochemistry*, **40**, 8898-8904.

St Johnston, D., Brown, N.H., Gall, J.G. et Jantsch, M.A. (1992) conserved double-stranded RNA-binding domain. *Proc. Natl. Acad. Sci. USA*, **89**, 10979-10983.

Stams, T., Niranjanakumari, S., Fierke, C.A. et Christianson, D.W. (1998) Ribonuclease P protein structure: evolutionary origins in the translational apparatus. *Science*, **280**, 752-755.

Stepien, P.P., Brousseau, R., Wu, R., Narang, S. et Thomas, D.Y. (1983) Synthesis of a human insulin gene. VI. Expression of the synthetic proinsulin gene in yeast. *Gene*, **24**, 289-297.

Steyaert, J., Hallenga, K., Wyns, L. et Stanssens, P. (1990) Histidine-40 of ribonuclease T1 acts as base catalyst when the true catalytic base, glutamic acid-58, is replaced by alanine. *Biochemistry*, **29**, 9064-9072.

Studier, F.W. (1991) Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J. Mol. Biol.*, **219**, 37-44.

Studier, F.W. et Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.*, **189**, 113-130.

Studier, FW. (1969) The genetics and physiology of bacteriophage T7. *Virology*, **39**, 562-574.

Subramanian, A.R. (1983) Structure and functions of ribosomal protein S1. Prog. *Nucleic Acid Res. Mol. Biol.*, **28**, 101-142.

Symmons, M.F., Jones, G.H. et Luisi B.F. (2000) A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure Fold. Des.*, **8**, 1215-1226.

Thomas, M.R. (1994) Simple, effective cleanup of DNA ligation reactions prior to electro-transformation of *E. coli*. *Biotechniques*, **16**, 988-990.

Thompson, J.E. et Raines, R.T. (1994) Value of General Acid-Base Catalysis to Ribonuclease A. *J. Am. Chem. Soc.*, **116**, 5467-5468.

Thompson, J.E., Venegas, F.D. et Raines, R.T. (1994) Energetics of catalysis by ribonucleases: fate of the 2',3'-cyclic phosphodiester intermediate. *Biochemistry*, **14**, 7408-7414.

Tindall, K.R. et Kunkel, T.A. (1988) Fidelity of DNA synthesis by the Thermus aquaticus DNA polymerase. *Biochemistry*, **27**, 6008-6013.

Trudel, P., Provost, S., Massie, B., Chartrand, P. et Wall, L. (1996) pGATA: a positive selection vector based on the toxicity of the transcription factor GATA-1 to bacteria. *Biotechniques*, **20**, 684-693.

Uchiyama, Y., Iwai, S., Ueno, Y., Ikehara, M., Ohtsuka, E. (1994) Role of the Mg2+ ion in the Escherichia coli

ribonuclease HI reaction. *J. Biochem.*, **116**, 1322-1329.

Uchiyama, Y., Miura, Y., Inoue, H., Ohtsuka, E., Ueno, Y., Ikehara, M. et Iwai, S. (1994) Studies of the interactions between Escherichia coli ribonuclease HI and its substrate. *J. Mol. Biol.*, **243**, 782-791.

Uzan, M. (2001) Bacteriophage T4 RegB endoribonuclease. *Methods Enzymol.*, **342**, 467-480.

Uzan, M., Favre, R. et Brody, E. (1988) A nuclease that cuts specifically in the ribosome binding site of some T4 mRNAs. *Proc. Natl. Acad. Sci. USA*, **85**, 8895-8899.

Vieira, J., et Messing, J. (1982) The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene*, **19**, 259-269.

Watanabe, H., Naitoh, A., Suyama, Y., Inokuchi, N., Shimada, H., Koyama, T., Ohgi, K. et Irie, M. (1990) Primary structure of a base non-specific and adenylic acid preferential ribonuclease from Aspergillus saitoi. *J. Biochem.*, **108**, 303-310.

White, H.B. 3rd. (1976) Coenzymes as fossils of an earlier metabolic state. *J. Mol. Evol.*, **7**, 101-104.

Westheimer, F.H. (1968) Pseudorotation in the hydrolysis of phosphate esters. *Accts. Chem. Res.*, **1**, 70-78.

Yang, X., Gerczei, T., Glover, L.T. et Correll, C.C. (2001) Crystal structures of restrictocin-inhibitor complexes with implications for RNA recognition and base flipping. *Nat. Struct. Biol.*, **8**, 968-73.

Yoshida, H. (2001) The ribonuclease T1 family. *Methods Enzymol.*, **341**, 28-41.

Zegers, I., Loris, R., Dehollander, G., Fattah Haikal, A., Poortmans, F., Steyaert, J. et Wyns, L. (1998) Hydrolysis of a slow cyclic thiophosphate substrate of RNase T1 analyzed by time-resolved crystallography.

Nat. Struct. Biol., **5**, 280-283.

Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406-3415.

ANNEXES

ANNEXE I

Séquence du gène *regB* du phage RB49

AY271260. Enterobacteria ph...[gi:30575687]

```
LOCUS       AY271260                376 bp    DNA     linear   PHG 12-MAY-2003
DEFINITION  Enterobacteria phage RB49 endoribonuclease RegB (regB) gene,
            partial cds.
ACCESSION   AY271260
VERSION     AY271260.1  GI:30575687
KEYWORDS    .
SOURCE      Enterobacteria phage RB49 (RB49)
  ORGANISM  Enterobacteria phage RB49
            Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
            T4-like viruses.
REFERENCE   1  (bases 1 to 376)
  AUTHORS   Saida,F., Uzan,M. and Bontems,F.
  TITLE     The regB gene from the phage RB49
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 376)
  AUTHORS   Saida,F., Uzan,M. and Bontems,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-APR-2003) ICSN, Ecole Polytechnique, Route de Saclay,
            Palaiseau 91128, France
FEATURES             Location/Qualifiers
     source          1..376
                     /organism="Enterobacteria phage RB49"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:50948"
     gene            <1..>376
                     /gene="regB"
     CDS             <1..>376
                     /gene="regB"
                     /codon_start=3
                     /transl_table=11
                     /product="endoribonuclease RegB"
                     /protein_id="AAP33010.1"
                     /db_xref="GI:30575688"
                     /translation="HVRRSKLRRFFETEMRKTNAKIKEKCLNDGYGEYIFIRFSPHMI
                     DRLTDRHIDEHKTMELFNKIETRLPEIFDFLKMEDRPNRLEITDGNMWIGMTVDLIEE
                     GKPYHGLCCRMIIENPKRLAGK"
BASE COUNT      126 a     70 c     86 g     94 t
ORIGIN
        1 tacatgttcg ccgtagtaag ctgcgtagat tcttcgaaac cgaaatgcgc aaaaccaatg
       61 cgaagatcaa agaaaagtgc ttgaatgatg gttatggtga gtatattttc attcgcttta
      121 gtccgcatat gatcgaccgt ctaaccgatc gtcatatcga tgaacataag acaatggaac
      181 tattcaataa gatagaaact cgtttaccgg agatcttcga cttcctgaaa atggaagacc
      241 gaccgaatcg attagaaatt actgacggta acatgtggat cgggatgaca gtagatctaa
      301 tcgaagaggg taagccatat cacggtctat gctgccgaat gattatagag aatccaaaac
      361 gactggcggg taaaat
//
```

ANNEXE II

L'intervention des histidines dans la catalyse ribonucléolytique est-elle le vestige d'une catalyse pré-biotique ?

Nous avons indiqué dans l'introduction de cette dissertation que les catalyseurs ribonucléolytiques les plus simples sont des chaînes d'acides ribonucléiques : les ribozymes. La découverte de ces ARN catalytiques a participé à la construction de la théorie du « Monde de l'ARN », terme employé pour la première fois par Gilbert Walter en 1986 et qui le décrit ainsi :

> « Durant son premier palier, l'évolution a procédé par des molécules d'ARN réalisant les activités catalytiques nécessaires à leur auto-assemblage à partir d'une soupe de nucléotides. Les molécules d'ARN évoluent selon un schéma autoréplicatif, utilisant recombinaison et mutation pour explorer de nouvelles niches. ... Ces molécules développent alors un arsenal entier d'activités enzymatiques. Au palier suivant, les molécules d'ARN commencèrent à synthétiser des protéines, d'abord en développant des adaptateurs ARN capables de lier un acide aminé activé puis en arrangeant ces derniers en suivant un ARN modèle et en utilisant d'autre molécules d'ARN comme ceux qui se trouvent dans le cœur du ribosome. Ce processus aurait fabriqué les premières protéines, qui allaient devenir tout simplement meilleures enzymes que les ARN de départ ...
>
> Finalement, l'ADN apparut sur la scène, l'ultime support de l'information copié à partir des molécules génétiques d'ARN par transcription inverse. ... l'ARN est alors relégué au rôle intermédiaire qu'il possède de nos jours – il n'est plus désormais le centre du palier, remplacé par l'ADN et les enzymes protéiques plus efficaces. »
>
> (Gilbert Walter, The RNA world, 1986)

Nous avons mis en évidence dans le chapitre 4 l'intervention récurrente des résidus histidines dans le mécanisme catalytique de plusieurs ribonucléases connues actuellement. Pourquoi le résidu histidine en particulier ?

Une des hypothèses plausibles est celle qui considère que des groupes anciennement constitutifs des acides nucléiques ont été incorporés au cours de l'évolution dans les acides aminés, donc les protéines, pour réaliser le même type de catalyses qu'ils exerçaient dans les acides

nucléiques originels (White H.B. III, 1976 ; Decout *et al.*, 1993). Le groupe imidazole de l'histidine en serait un exemple. Le N6 ribozyl-adénine (Figure A ci-dessous) est un composé facile à obtenir dans les conditions présumées de la synthèse pré-biotique. Il possède un groupe imidazole libre et présente des propriétés catalytiques comparables à celle de l'histidine (Maurel M.C., 1994 ; Decout *et al.*, 1993). Il aurait participé à la catalyse réalisée par les acides nucléiques primitifs, puis se serait retrouvé dans les chaînes polypeptidiques sous la forme de la chaîne latérale d'une histidine. A ce titre, il est intéressant de remarquer qu'une purine est aujourd'hui à l'origine de la biosynthèse cellulaire de l'histidine ce qui renforce l'hypothèse d'un lien métabolique primordial entre l'histidine et les catalyseurs nucléiques primitifs (Maurel M.C., 1994).

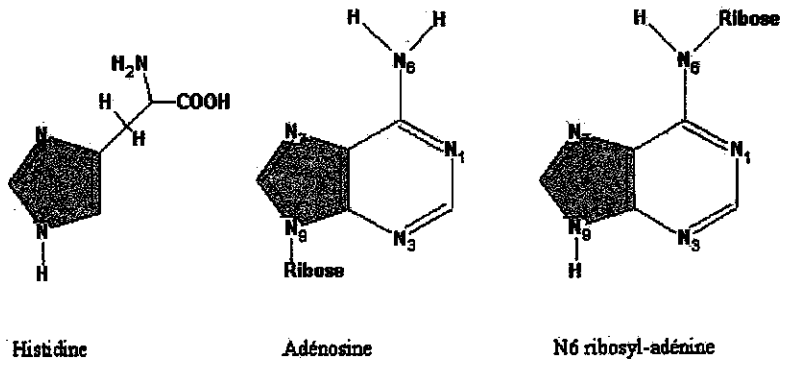


Figure A. Le N6 ribosyl-adénine a des propriétés catalytiques comparables à celles de l'histidine. D'après Maurel M.C., 1994.

PUBLICATIONS

PUBLICATION I

Saida F., Uzan, M., Lallemand, J.Y. et Bontems, F. (2003) New system for positive selection of recombinant plasmids and dual expression in yeast and bacteria based on the restriction ribonuclease RegB. *Biotechnol. Prog.*, **19**, 727-733.

New System for Positive Selection of Recombinant Plasmids and Dual Expression in Yeast and Bacteria Based on the Restriction Ribonuclease *Reg*B

Fakhri Saïda[†], Marc Uzan[‡], Jean-Yves Lallemand[†], and François Bontems*[,†]

Laboratoire ICSN-RMN, Institut de Chimie des Substances Naturelles, Ecole Polytechnique, Route de Saclay, 91128, Palaiseau Cedex, France, and Institut Jacques Monod, UMR 7592, CNRS, Universités Paris 6 et Paris 7, 2 Place Jussieu, Paris, France

By coupling the toxic restriction endoribonuclease *Reg*B, from the bacteriophage T4, to the prokaryotic T7 and the eukaryotic GAL1 promoters, we constructed a two-function plasmid called pTOXR-1. This plasmid is a zero-background cloning vector. It allows an efficient positive selection of recombinant plasmids without the need to completely digest, dephosphorylate, or purify the vector prior to the ligation step. The pTOXR-1 positive selection system requires no special *Escherichia coli* strains, no special culture media, and no addition of inducer to the selective plates. In addition, since this vector carries all signals required for both prokaryotic and eukaryotic expression, it allows the overproduction of heterologous proteins, fused to a polyhistidine tag, in the bacterium *E. coli* and in the yeast *Saccharomyces cerevisiae* from a single plasmid. Hence, this vector may be a useful time-saving tool for one-step cloning and versatile protein expression in bacteria and yeast.

Introduction

The present post-genomic era needs efficient tools for cloning and expressing proteins to meet the requirements for both biochemical and structural studies. Fast cloning of foreign genes needs a system that positively selects the clones containing the recombinant plasmid. Two generations of positive selection systems are in use nowadays. The first one is the well-known blue–white color selection system (*1*), based on the disruption of the coding sequence of the β-galactosidase α-peptide by the cloned insert. Although very practical, this system is limited to special *Escherichia coli* (*E. coli*) strains (those with the Δ(LacZ)M15 deletion) and needs the induction of the α-peptide expression by adding isopropyl-β-D-thio-galactopyranoside (IPTG). Therefore, the protein that replaces the α-peptide is also induced, which make the system useless when the protein of interest is toxic.

The second generation of positive selection systems is based on the inactivation of a toxic agent by the cloned insert. The toxic ribonuclease barnase (*2–4*), the lethal protein CcdB (control cell death) (*5–9*), the toxic DNA binding domain of the mouse eukaryotic transcription factor GATA-1 (*10*), and the rat antimicrobial defensin RatNP-1 (*11*) are used in many selection vectors. Like the blue–white color system, these systems often require the addition of IPTG. Since small differences are observed between different *E. coli* strains, the concentration of IPTG should be optimized before the cloning process. Without IPTG, the positive selection process is limited to *E. coli* strains that are devoid of an endogenous copy of the Lac repressor (to prevent a tight repression of the toxic gene basal expression). Moreover, in the case of the CcdB protein, *E. coli* strains containing the F plasmid must be avoided since this plasmid expresses the CcdA protein, a natural inhibitor of CcdB.

Once cloned in one of these positive selection vectors, the gene of interest often needs to be subcloned into an efficient expression vector. Moreover, it is useful to express the same protein in a prokaryotic system to produce large amounts for structural studies and in an eukaryotic system to obtain correct folding or posttranslational modifications. This process is time-consuming because it requires two additional subcloning steps at the very least (dual expression vectors are rare) and uses commercially available expression vectors often devoid of positive selection systems.

Here we present pTOXR-1, a new system that allows the positive selection of recombinant plasmids in all the common *E. coli* strains, without addition of IPTG, and which also works as a dual expression vector in the bacterium *E. coli* and in the yeast *Saccharomyces cerevisiae* (*S. cerevisiae*). The positive selection system of pTOXR-1 is based on *Reg*B: a toxic restriction endoribonuclease encoded by the genome of the phage T4 (*12*). *Reg*B controls the virus lytic cycle by inactivating some mRNA produced during the early stage of the infection (*13*). It cleaves specifically the GGAG consensus sequence mostly located in the ribosome-binding site that is crucial for the efficient initiation of the translation (*14, 15*). It also has a strong poisoning effect on *E. coli* even at very small amounts (*12*). The exact mechanism of this poisoning effect is unknown, but we think that at least some vital bacterial mRNAs might be targets for *Reg*B cleavage. The pTOXR-1 dual expression system was obtained by coupling the yeast GAL1 promoter with the bacterial T7 promoter and allowing the eukaryotic and the prokary-

* To whom correspondence should be addressed. Phone: 0033169334855. Fax: 0033169333010. E-mail: francois.bontems@polytechnique.fr.
† Ecole Polytechnique.
‡ Universités Paris 6 et Paris 7.

10.1021/bp0257224 CCC: $25.00 © 2003 American Chemical Society and American Institute of Chemical Engineers
Published on Web 05/08/2003

otic translations to meet in the same ATG initiator codon. In this paper, we present the strategy leading to the construction, amplification, and purification of the pTOXR-1 plasmid. We illustrate the positive selection and the dual expression features of pTOXR-1 with a model DNA sequence coding for the C-terminal domains 4 and 5 of the *E. coli* ribosomal protein S1 (hereafter called S1 bimodule 45) (*16*).

Materials and Methods

Strains. *Escherichia coli* strains BL21 strain [F⁻ ompT $hsdS_B$ (r_B^- m_B^-) gal dcm] and BL21(DE3), a BL21 derivative with DE3 being a λ prophage carrying the T7 RNA polymerase gene, were from Novagen. XL1-Blue [F'::Tn10 $proA^+B^+$ $lacI^q$ Δ(lacZ)M15/recA1 endA1 gyrA96 (Nal^r) thi hsdR17 (r_K^- m_K^+) glnV44 relA1 lac] and JM101 [F' traD36 $proA^+B^+$ $lacI^q$ Δ(lacZ)M15/Δ(lac-proAB) glnV thi] were from Stratagene. Strain DH5α [F⁻ endA1 hsdR17 (r_K^- m_K^+) glnV44 thi-1 recA1 gyrA(Nal^r) relA1 Δ(lacIZYA-argF)U169 deoR (φ80dlac Δ(lacZ)M15)] was from Life Technologies. The yeast *S. cerevisiae* strain INVSc1 [MATa his3Δ1 leu2 trp1-289 ura3-52/MATα his3Δ1 leu2 trp1-289 ura3-52] was from Invitrogen.

Plasmids. The DNA sequence coding for the inactive point mutant *Reg*B R52L, initially cloned in the vector M13mp18 (*28*), was PCR amplified using the oligonucleotides NU10 and NU13, digested with *Nde*I and *Bam*HI, and cloned in the pET15b vector (Novagen) yielding to the pERL52 plasmid. The pARNU2 plasmid is a derivative of the transcription vector pET-7 (*17, 18*) in which the wild type *Reg*B gene was cloned under the control of the T7 promoter. A modified gene φ10 ribosome-binding site was added to this vector to express the *Reg*B protein (*19*). This plasmid produces 150 μg of *Reg*B from 1 L of *E. coli* culture. The pS145 plasmid is a pET15b derivative in which the DNA sequence coding for the S1 bimodule 45 was cloned in the *Nde*I and the *Bam*HI sites. The pOU61 runaway replication plasmid (*20*) was a gift from Dr. H. Putzer. The pYES2 shuttle plasmid was from Invitrogen.

Oligonucleotides for PCR. NUb (5' CGA CGT CGA TCT CGA TCC CGC 3') is complementary to the region in plasmid pET28b (Novagen) upstream of the T7 promoter. NUE1 (5' CGG AGC TCG AAT TCA TTA ACT TC 3') is complementary to the sequence in the *Reg*B gene that contains the unique *Eco*RI site. NU10 (5' CCT GCA GGA GAT ATA CAT ATG 3') is complementary to the 5' part of the *Reg*B gene and contains an *Nde*I site within the ATG initiator codon region. NU13 (5' CGG GAT CCT CAT TGA GTT TTA AT 3') and NU15 (5' TTC TCG AGT CAT TGA GTT TTA AT 3') are complementary to the 3' end of the *Reg*B gene and contain a stop codon followed by a *Bam*HI site (NU13) or an *Xho*I site (NU15). 5C4 (5' GAA ATT CCA TAT GTG GGT AGC TAT CGC T 3') contains an *Nde*I site in the ATG initiator codon region. 3C5 (5' CGG GAT CCT TAG TTC AGA GCA ACC CAG 3') contains a stop codon and a *Bam*HI site. 5C4 and 3C5 were used to amplify, from pS145, a DNA fragment coding for the S1 bimodule 45.

Yeast Colony Screening by PCR. A 20 μL PCR mix contained Taq DNA polymerase buffer 1X (Promega), 3.0 mM $MgCl_2$, 0.2 mM dNTP (Boehringer Mannheim), 100 pmol each of oligonucleotide Nub and NU15, and 1 unit of Taq DNA polymerase (Promega). The DNA template originated from the third of a yeast colony. The PCR program, run on a Perkin-Elmer GeneAmp PCR system 2400, was as follows: [(94 °C, 10 min) X 1, (94 °C, 30 s; 42 °C, 30 s; 72 °C, 45 s) X 30, (72 °C, 7 min) X 1].

Extraction of pTOXR-1 Plasmid from Yeast. The protocol is a modification of the one published by Robzyk and Kassir. (*21*). Yeast cells from 100 mL of overnight culture at 28 °C in the SM6 medium [0.7% yeast nitrogen base (Difco), 0.01% L-tryptophan (Sigma), 0.01% L-leucine (Sigma), 0.005% L-histidine (Sigma), 0.1% bacto-casamino acids (Difco), and 2.0% D(+)-glucose (Sigma)] were resuspended in 6.7 mL of the GTET buffer [8.0% D(+)-glucose, 50.0 mM Tris-HCl (pH 8.0), 1.0 mM ethylenediaminetetraacetic acid (EDTA), 5.0% Triton X-100 (Sigma)]. The cells were disrupted by adding an equal volume of acid-washed 0.4−0.6 mm glass beads (Sigma) and shaking vigorously by hand for 5 min at room temperature. A 6.7 mL aliquot of the GTET buffer was added, and the mix was boiled (4 min, 95 °C), chilled in ice (10 min), and centrifuged (10 min, 10 000 g, 4 °C). The supernatant was cleared by adding ammonium acetate (7.5 M, half-volume) and incubating at −20 °C (1 h). After a new centrifugation, the supernatant was precipitated with 100% ethanol, the pellet washed with 70% ethanol, resuspended in 200 μL of double distilled water, migrated in a 1% TAE-agarose gel (size, 15 cm × 15 cm) at 50 V for 5 h, and finally visualized with 0.1 μg/mL ethidium bromide. The pTOXR-1 band was extracted from the agarose with the Fast 'n' Easy DNA Purification kit (Tebu). DNA concentrations obtained with this protocol are up to 20 ng/μL.

pTOXR-1 3 h Cloning Protocol. The pTOXR-1 vector (10 ng) was digested with *Nde*I and *Bam*HI (New England Biolabs) for less than 30 min; the PCR amplified insert (the S1 bimodule 45) was digested with the same enzymes for 2 h. The reactions were stopped by heating (80 °C, 20 min). The digested vector and insert were then precipitated [3 M sodium acetate (0.3 M final) and 100% ethanol (70% final)], resuspended in double distilled water, mixed, and ligated using the T4 DNA ligase (New England Biolabs) for 1 h at 16 °C. The insert-to-vector ratio was equal to 4. The ligation mixture was *n*-butanol precipitated (*22*), resuspended in double distilled water, and used to transform the suitable electrocompetent *E. coli* strain using the Eppendorf Electroporator 2510.

Expression, Purification, and Detection of the S1 Bimodule 45 Using the pXS145 Plasmid. The *E. coli* BL21(DE3) strain transformed with pXS145 was inoculated in 5 mL of a modified Luria Bertani broth medium [1.0% Nutriment Broth (Difco), 0.5% yeast extract (Difco), 0.8% NaCl, 0.052% Tris] containing 100 μg/mL ampicillin (Sigma). When the optical density at 600 nm reached 0.8 (37 °C), the protein expression was induced for 3 h using 1.0 mM IPTG (Promega). pXS145 was expressed in a 50 mL culture of the yeast INVsc1, in the SM6 medium, according to the manufacturer protocol (Invitrogen). The crude extracts were obtained by sonication (*E. coli*) or by shaking with glass beads (yeast). The protein was purified according to the cobalt affinity chromatography protocol (Clontech) with the lysis buffer [50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 0.1 mM AEBSF protease inhibitor (Euromedex), 5.0 mM β-mercaptoethanol], the wash buffer (lysis buffer supplemented with 10 mM imidazole) and the elution buffer (lysis buffer supplemented with 150 mM imidazole). The protein expressed from pXS145 was detected by western blot [primary antibody, anti-polyhistidine tag from mouse

Table 1. Viability of Five *E. Coli* Strains Transformed with Different Constructions Containing the *RegB* Gene[a]

	plasmid name	parental plasmid	transcriptional control element	translational control element	plasmid copy no.	JM101, XL1-Blue	DH5α BL21	BL21 (DE3)
wild type *Reg*B gene	pARNU2	pET7	T7 promoter	weak	15–20	viable	viable	lethal
		pET15b	T7 promoter	efficient	15–20	lethal	lethal	lethal
	pOURegB	pOU61	T7 promoter	weak	1	viable	viable	unstable
		pOU61	T7 promoter	efficient	1	lethal	lethal	lethal
inactive mutant *Reg*B gene	pOURL52	pOU61	T7 promoter	efficient	1	viable	viable	viable
	pERL52	pET15b	T7 promoter	efficient	15–20	viable	viable	viable

[a] The wild type *Reg*B gene is lethal to all the five strains whenever it is under the control of both a strong T7 promoter and efficient translation initiation elements including gene φ10 ribosome-binding site. The same results are obtained when the plasmid copy number is reduced from 15 to 20 in pET15b to one copy per cell in pOU61 (*20*). As a control, the mutant *Reg*B gene (R52L) encoding an inactive protein was maintained successfully in all five strains. pOURegB and pOURL52 are two pOU61 derivatives that differ respectively from pARNU2 and pERL52 only by their low copy number. T7 promoter: phage T7 gene φ10 promoter.

(Amersham Pharmacia); secondary antibody, anti-mouse antibody coupled to the alkaline phosphatase (Sigma)].

Results

Conditions for RegB Lethality. We studied the viability of five *E. coli* strains [BL21(DE3), BL21, DH5α, JM101, and XL1-Blue] when transformed with different constructions containing the *Reg*B gene. The BL21(DE3) strain has an endogenous copy of the T7 RNA polymerase gene. It is suitable for the overproduction of proteins cloned under the control of the strong gene φ10 promoter from the phage T7. However, the leaky expression of the polymerase leads to a relatively high basal expression of the heterologous protein in the absence of the inducer (*23*). The four other strains (BL21, DH5α, JM101, and XL1-Blue) lack the T7 RNA polymerase. The basal expression is thus dramatically reduced since it is limited to the read-through transcription mediated by the *E. coli* RNA polymerase, presumably from cryptic promoters on the plasmid, continuing through the T7 promoter region. In addition, the JM101 and the XL1-Blue strains overexpress the Lac repressor (genotype $lacI^q$). This molecule binds to the lac operator placed just downstream of the T7 promoter and blocks the transcription. These two strains, thus, exhibit an even lower basal expression.

Table 1 shows that the *Reg*B gene cannot be maintained in any of the five *E. coli* strains when it is simultaneously coupled to both a strong transcription signal (T7 promoter upstream of the *lac* operator) and an efficient translation signal including the bacteriophage T7 gene φ10 ribosome-binding site. Such constructions are lethal not only in BL21(DE3) but also in BL21 and DH5α, and more interestingly in JM101 and XL1-Blue, plated without IPTG. Reducing gene dosage by decreasing the vector copy number from 15 to 20 for pET15b to one copy per cell for pOU61 (*20*) does not suppress the lethality of *Reg*B. The only stable constructions containing the *Reg*B gene are plasmids pARNU2 and pOURegB (this plasmid contains the *Reg*B expression region originating from pARNU2 cloned in pOU61). pARNU2 and pOURegB share a modified ribosome-binding site that decreases the translation efficiency (40–50 times) compared to the standard gene φ10 ribosome-binding site (data not shown). The sequence of this modified ribosome-binding site (which may be useful to maintain other toxic genes) is presented in Figure 1.

Taking these initial findings into account, we concluded that *Reg*B was sufficiently toxic to be used as a positive selection system that would require no special *E. coli* strains, no special culture media, and no inducer addition (IPTG). Furthermore, since *Reg*B has been cloned under the control of the eukaryotic GAL1 promoter and maintained successfully in the yeast *S. cerevisiae* using the shuttle vector pYES2, we hypothesized that a plasmid containing the wild type *Reg*B gene under the control of the strong T7/lac promoter and the efficient translation initiation elements of vectors in the pET series, could be constructed and amplified in yeast and then used as a positive selection system in *E. coli*. Besides, we anticipated that by judiciously placing the prokaryotic and eukaryotic expression signals we could use this plasmid as a dual expression vector in yeast and bacteria. To test these assumptions we decided to construct the pTOXR-1 plasmid.

Construction of the pTOXR-1 Plasmid. The structure of pTOXR-1 is presented in Figure 2. The wild type *Reg*B gene was inserted in to the yeast shuttle vector pYES2 as a DNA sequence flanked by the bacterial expression signals, including the T7 promoter with the lac operator, an efficient translation initiation region and the phage T7 φ transcription terminator. The *Reg*B gene was inserted in such a way that it is also under the control of the yeast GAL1 promoter and CYC1 transcription terminator. The construction was built in two steps (Figure 3). First, we constructed two DNA fragments (Figure 3). In fragment A, the 5′ part of the wild type *Reg*B gene is fused in frame to a sequence encoding a polyhistidine tag and put under the control of the T7/lac promoter and the translation initiation region from the pET28b expression vector. Fragment B contains the 3′ part of the *Reg*B gene, including its stop codon, and the phage T7 φ transcription terminator from the pET15b expression vector. This fragment was obtained from the pERL52 plasmid (see Materials and Methods) that contains the inactive mutant *Reg*B R52L (*24*). Fragment A was then cloned into the yeast expression vector pYES2 which had its unique *Nde*I site suppressed using the standard T4 DNA polymerase filling-in protocol. In the recombinant plasmid, called pYTOX, fragment A is localized downstream of the eukaryotic GAL1 promoter and upstream of the eukaryotic CYC1 terminator. In a second step, fragment B was ligated to the pYTOX plasmid and the ligation mixture was used to transform the yeast *S. cerevisiae*. The recombinant yeast colonies were screened by PCR using primers Nub and NU15 (see Materials and Methods). The recombinant pTOXR-1 plasmid was extracted from yeast according to the procedure outlined in Materials and Methods. Because the poor plasmid recovery from yeast prevented the pTOXR-1 direct sequencing, we checked the sequence of the wild type *Reg*B gene by sequencing a PCR amplified fragment.

pTOXR-1 Plasmid Lethal to *E. coli* under Non-Inducing Conditions. We transformed, by electroporation, the five previously described *E. coli* strains with the pTOXR-1 plasmid. Plasmid pERL52 containing the

| | | |

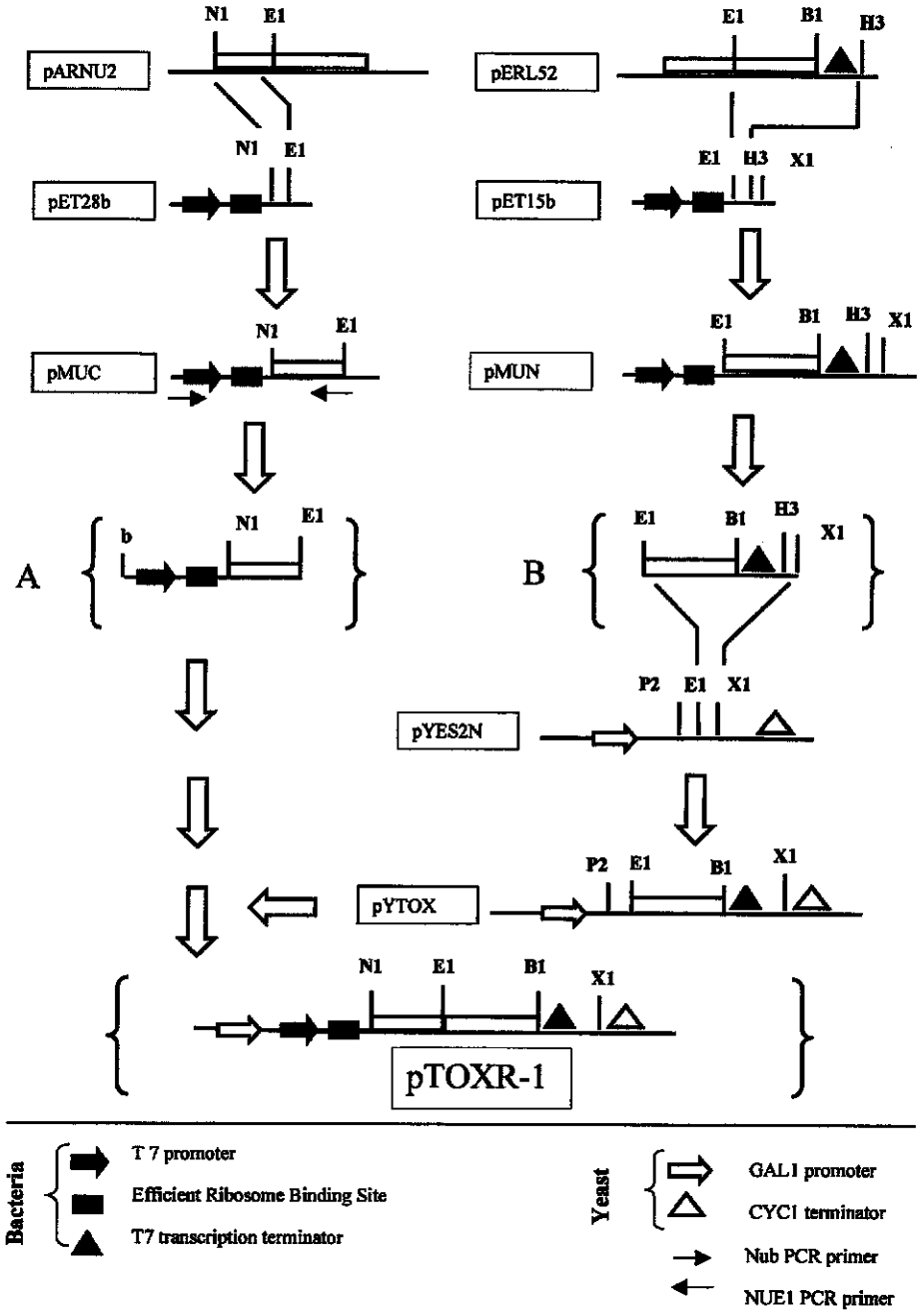


Figure 3. Schematic diagram of the pTOXR-1 plasmid construction. Fragment sizes are not drawn to scale. Cloning the *Nde*I-*Eco*RI insert from pARNU2 and the *Eco*RI-*Hin*dIII insert from pERL52 (see Materials and Methods) in the pET28b vector yielded, respectively, pMUC and pMUN. Fragment A is the PCR product of pMUC amplification with primers NUE1 and NUb. Fragment B is the *Eco*RI-*Xho*I fragment excised from pMUN. pYES2N is pYES2 devoid of its unique *Nde*I site. Inserting fragment B in pYES2N yielded to pYTOX. Inserting fragment A in the *Pvu*II/*Eco*RI sites of pYTOX yielded pTOXR-1. The N1E1 and the E1B1 fragments are respectively the 5′ and the 3′ parts of the wild type *Reg*B gene. Abbreviations: B1, *Bam*HI; E1, *Eco*RI; H3, *Hin*dIII; N1, *Nde*I; P2, *Pvu*II; ×1, *Xho*I; b, blunt end PCR product.

the protein produced from pS145 (the parental plasmid of which is pET15b), we analyzed by SDS-AGE the crude extracts from the induced BL21(DE3) cells transformed with these two plasmids. Figure 6 shows that the two plasmids overexpress a protein with the same molecular weight (between 18.4 and 29 kDa), which is compatible with the predicted molecular weight of 22.6 kDa for the polyhistidine-tagged S1 bimodule 45.

Discussion

The pTOXR-1 plasmid presented in this paper is an overexpressing vector carrying a copy of the wild type *Reg*B gene as a positive selection system. This system overcomes the need to prepare special culture media, to transform special *E. coli* strains or to add IPTG.

In the absence of the T7 RNA polymerase gene, the lethality of pTOXR-1 may be explained by the read-

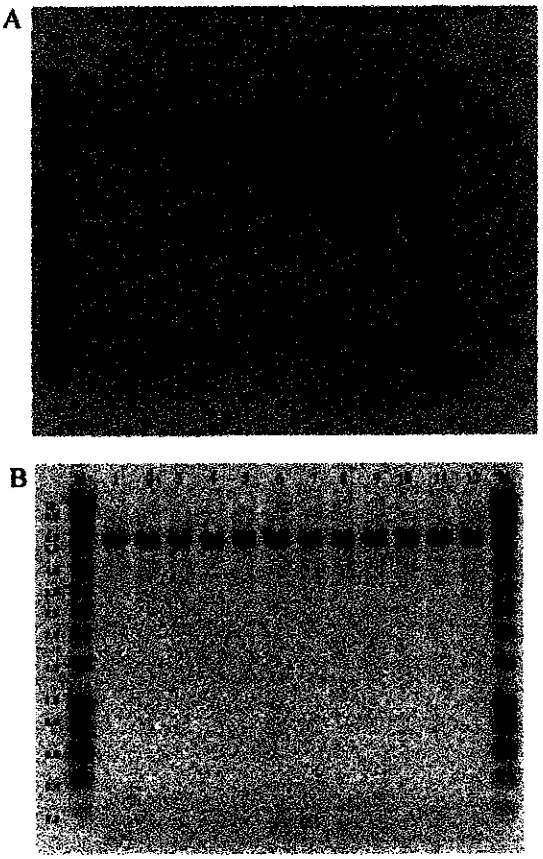


Figure 4. Cloning of the S1 bimodule 45 using the positive selection system of pTOXR-1 (A) or the standard pET15b vector (B). Plasmid DNA of 12 random picked clones was extracted, digested with *Nde*I and *Bam*HI, and loaded on a 1% agarose gel. Two positive controls are used: pS145 and pERL52 (see Materials and Methods), the *NdeI/Bam*HI digestion of which produces, respectively, a 557 bp insert containing the protein S1 fragment and a 465 bp insert containing the coding sequence of the *Reg*B inactive point mutant R52L. M: molecular weight marker; values in kbp.

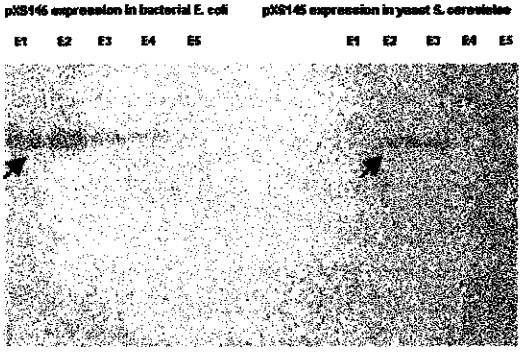


Figure 5. Western-blot analysis of the S1 fragment expression in *E. coli* and *S. cerevisiae* transformed with pXS145. Cells from 5 mL (*E. coli*) or 50 mL (*S. cerevisiae*) cultures were disrupted, purified using cobalt affinity chromatography, trichloroacetic acid-precipitated, separated by SDS–PAGE (18%), transferred to a nitrocellulose membrane (HyBond-P, Pharmacia), and detected using an anti-mouse antibody coupled to the alkaline phosphatase that recognizes the mouse anti-polyhistidine-tag antibody. E1–E5: elution fractions. The arrows indicate the positions of the detected bands.



Figure 6. Comparison of the molecular weight of the S1 fragments produced in *E. coli* after transformation with the pS145 plasmid (left) or the pXS145 plasmid (right). M: molecular weight marker; values in kDa. The calculated molecular weight of the S1 bimodule 45 is 22.6 kDa. The arrows indicate the position of the target proteins.

through transcription mediated by the *E. coli* RNA polymerase. The efficient ribosome-binding site allows then the synthesis of a lethal dose of the *Reg*B protein. This lethal dose is probably very low. Hence, we suggest that *Reg*B may cleave only a few mRNA that are vital for *E. coli*. Uzan and co-workers found, indeed, that *Reg*B cleaves efficiently the *ftsZ* transcript (an essential gene involved in the early steps of the cell septation) causing filamentation, whereas the bulk *E. coli* mRNA is not significantly modified after *Reg*B expression (*12*). Because *Reg*B protein targets the GGAG sequence mainly found in prokaryotic ribosome-binding sites, it is easy to understand why our positive selection system does not depend on *E. coli* strains. It is thus possible that the *Reg*B positive selection could be used in other prokaryotic systems, like gram-positive bacteria (*Bacillus subtilis*). However, upon transformation of the eukaryotic yeast *S. cerevisiae*, we noticed no change in cell phenotype or cell growth, even after the induction of the *Reg*B expression with galactose. Although we cannot exclude that some eukaryotic mRNAs may be targets for *Reg*B cleavage (in their coding sequence for example), we can assert that *Reg*B is not as detrimental to yeast as it is to *E. coli*. This differential effect may be explained not only by the fact that the eukaryotic initiation of translation is independent of the presence of a GGAG sequence but also by the fact that *Reg*B may require some components of the prokaryotic ribosome (ribosomal protein S1) (*15*) to be fully active. The differential effect of *Reg*B toward the prokaryotic and the eukaryotic systems is the key point of our strategy because it allowed us to build the pTOXR-1 plasmid, by reversing the standard methods used for the *E. coli*/yeast shuttle vectors. In general, these plasmids are built in *E. coli* and used to transform yeast because plasmid amplification and purification is easier in bacteria. Here, we have built, amplified, and purified the plasmid in yeast, to avoid toxicity, and used it in bacteria. Although plasmid recovery from yeast is low, our system circumvents this problem by its high efficiency (as few as 10 ng of pTOXR-1 were sufficient to build the pXS145 plasmid with a yield of 92% of recombinant plasmids).

The construction of the pXS145 plasmid, using pTOXR-1, yielded to only one nonrecombinant plasmid (clone 11

in Figure 4A). It contained the *Reg*B gene as an insert in a 2.2 kbp plasmid. The restriction analysis of this clone revealed that it was viable because it underwent a large deletion that suppressed the translational signal directing the *Reg*B expression. We think that such problems might be overcome if pTOXR-1 was digested for a shorter time or with less enzyme units prior to the ligation step. This would avoid plasmid overdigestion and unspecific cut and hence increase the percentage of recombinants.

The pTOXR-1 positive selection system is limited to inserts with an *Nde*I site in their 5′ end and an *Eco*RI or a *Bam*HI site in their 3′ end. These three restriction sites occur rarely in a random DNA sequence (one site every $4^6 = 4096$ bp) and are used successfully for expressing all the proteins studied in our laboratory. However, if needed, the versatility of our plasmid could be enhanced by constructing different pTOXR-1 derivatives containing different tandems of additional cloning sites (polylinkers). These polylinkers could be inserted in the *Bam*HI site of pTOXR-1, allowing the cloning of fragments with an *Nde*I start site but a variety of C-terminal sites. Examples of such polylinkers, which could be obtained by annealing two complementary oligonucleotides flanked by two *Bam*HI sites, are POLY1 (GGATCC GGTACC TTAATTAA GTCGAC AGATCT GGATCC) and POLY2 (GGATCC CAGCTG GAGCTC CAATTG GGCGCC GGATCC). POLY1 and POLY2 contain respectively the following cloning sites (*Bam*HI, *Kpn*I, *Pac*I, *Sal*I, and *Bgl*II) and (*Bam*HI, *Pvu*II, *Sac*I, *Mun*I, and *Nar*I).

We think that the positive selection and the dual expression systems presented in this paper may be useful for repetitive standardized cloning/expression of new genes for structural genomics projects for example.

Acknowledgment

We are grateful to Drs. D. Pompon and P. Urban for yeast genetics recommendations and Dr. M. Bisaglia for his generous gift of the pS145 plasmid. We thank Dr. C. Condon for helpful comments on the manuscript. This work is supported by a national grant from the French Departement of Scientific Research (Ministere de la Recherche).

References and Notes

(1) Vieira, J.; Messing, J. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **1982**, *19*, 259–269.
(2) Yazynin, S. A.; Deyev, S. M.; Jucovic, M.; Hartley, R. W. A plasmid vector with positive selection and directional cloning based on a conditionally lethal gene. *Gene* **1996**, *169*, 131–132.
(3) Deyev, S. M.; Yazynin, S. A.; Kuznetsov, D. A.; Jucovic, M.; Hartley, R. W. Ribonuclease-charged vector for facile direct cloning with positive selection. *Mol. Gen. Genet.* **1998**, *259*, 379–382.
(4) Yazynin, S.; Lange, H.; Makros, T.; Deyev, S.; Lemke, H. A new phagemid vector for positive selection of recombinants based on a conditionally lethal barnase gene. *FEBS Lett.* **1999**, *452*, 351–354.
(5) Bernard, P.; Couturier, M. Cell killing by the F plasmid ccdB protein involves poisoning of the DNA-topoisomerase II complexes. *J. Mol. Biol.* **1992**, *226*, 735–745.
(6) Bernard, P.; Kezdy, K. E.; Van Melderen, L.; Steyaert, J.; Wyns, L.; Pato, M. L.; Higgins, P. N.; Couturier, M. The F plasmid ccdB protein induces efficient ATP-dependent DNA cleavage by gyrase. *J. Mol. Biol.* **1993**, *234*, 534–541.
(7) Bernard, P.; Gabant, P.; Bahassi, E. M.; Couturier, M. Positive selection vetors using the F plasmid ccdB killer gene. *Gene* **1994**, *148*, 71–74.
(8) Mondon, P.; Chang, Y. C.; Varma, A.; Kwon-Chung, K. J. A novel episomal shuttle vector for transformation of Cryptococcus neoformans with the ccdB gene as a positive selection marker in bacteria. *FEMS Microbiol. Lett.* **2000**, *187*, 41–45.
(9) Lew, B. M.; Paulus, H. An in vivo screening system against protein splicing useful for the isolation of nonsplicing mutants or inhibitors of the RecA intein of *Mycobacterium tuberculosis*. *Gene* **2002**, *282*, 169–177.
(10) Trudel, P.; Provost, S.; Massie, B.; Chartrand, P.; Wall, L. pGATA: A positive selection vector based on the toxicity of the transcription factor GATA-1 to bacteria. *Biotechniques* **1996**, *20*, 684–693.
(11) Chan, R. Y.; Palfree, R. G.; Congote, L. F.; Solomon, S. Development of a novel type of cloning vector for suicide selection of recombinants. *DNA Cell. Biol.* **1994**, *13*, 311–319.
(12) Uzan, M. Bacteriophage T4 RegB endoribonuclease. *Methods Enzymol.* **2001**, *342*, 467–480.
(13) Sanson, B.; Uzan, M. Dual role of the sequence-specific bacteriophage T4 endoribonuclease RegB. mRNA inactivation and mRNA destabilization. *J. Mol. Biol.* **1993**, *233*, 429–446.
(14) Sanson, B.; Uzan, M. Post-transcriptional controls in bacteriophage T4: Roles of the sequence-specific endoribonuclease RegB. *FEMS Microbiol. Rev.* **1995**, *17*, 141–150.
(15) Ruckman, J.; Ringquist, S.; Brody, E.; Gold, L. The bacteriophage T4 regB ribonuclease. Stimulation of the purified enzyme by ribosomal protein S1. *J. Biol. Chem.* **1994**, *269*, 26655–26662.
(16) Schnier, J.; Kimura, M.; Foulaki, K.; Subramanian, A. R.; Isono, K.; Wittmann-Liebold, B. Primary structure of Escherichia coli ribosomal protein S1 and of its gene rpsA. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 1008–1011.
(17) Rosenberg, A. H.; Lade, B. N.; Chui, D. S.; Lin, S. W.; Dunn, J. J.; Studier, F. W. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **1987**, *56*, 125–135.
(18) Studier, F. W.; Rosenberg, A. H.; Dunn. J. J.; Dubendorff, J. W. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **1990**, *185*, 60–89.
(19) Sanson, B.; Hu, R. M.; Troitskaya, E.; Mathy, N.; Uzan, M. Endoribonuclease RegB from bacteriophage T4 is necessary for the degradation of early but not middle or late mRNAs. *J. Mol. Biol.* **2000**, *297*, 1063–1074.
(20) Larsen, J. E.; Gerdes, K.; Light, J.; Molin, S. Low-copy-number plasmid-cloning vectors amplifiable by derepression of an inserted foreign promoter. *Gene* **1984**, *28*, 45–54.
(21) Robzyk, K.; Kassir, Y. A simple and highly efficient procedure for rescuing autonomous plasmids from yeast. *Nucleic Acids Res.* **1992**, *20*, 3790.
(22) Thomas, M. R. Simple, effective cleanup of DNA ligation reactions prior to electro-transformation of *E. coli*. *Biotechniques* **1994**, *16*, 988–990.
(23) Studier, F. W.; Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **1986**, *189*, 113–130.
(24) Ruckman, J.; Parma, D.; Tuerk, C.; Hall, D. H.; Gold, L. Identification of a T4 gene required for bacteriophage mRNA processing. *New Biol.* **1989**, *1*, 54–65.
(25) Shine, J.; Dalgarno, L. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome-binding sites. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 1342–1346.
(26) Kozak, M. Initiation of translation in prokaryotes and eukaryotes. *Gene* **1999**, *234*, 187–208.
(27) Kozak, M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **1986**, *44*, 283–292.
(28) Uzan, M. Unpublished data.

Accepted for publication March 12, 2003.

BP0257224

PUBLICATION II

Saida, F., Uzan, M. et Bontems F. (2003) The phage T4 restriction endoribonuclease RegB: a cyclizing enzyme that requires two histidines to be fully active. *Nucleic Acids Res.*, **31**, 2751-2758.

Nucleic Acids Research, 2003, Vol. 31, No. 11 2751–2758
DOI: 10.1093/nar/gkg377

The phage T4 restriction endoribonuclease RegB: a cyclizing enzyme that requires two histidines to be fully active

Fakhri Saïda, Marc Uzan[1] and François Bontems*

Laboratoire ICSN-RMN, Institut de Chimie des Substances Naturelles, Ecole polytechnique, Route de Saclay, 91128 Palaiseau, France and [1]Institut Jacques Monod, UMR 7592, CNRS, Universités Paris 6/Paris 7, France

Received March 6, 2003; Revised and Accepted April 1, 2003

ABSTRACT

The *regB* gene, from the bacteriophage T4, codes for an endoribonuclease that controls the expression of a number of phage early genes. The RegB protein cleaves its mRNA substrates with an almost absolute specificity in the middle of the tertranucleotide GGAG, making it a unique well-defined restriction endoribonuclease. This striking protein has no homology to any known RNase and its catalytic mechanism has never been investigated. Here, we show, using ^{31}P nuclear magnetic resonance (NMR), that RegB produces a cyclic 2′,3′-phosphodiester product. In order to determine the residues crucial for its activity, we prepared all the histidine-to-alanine point mutants of RegB. The activity of these mutants was characterized both *in vivo* and *in vitro*. In addition, their binding capability was quantified by surface plasmon resonance and their structural integrity was probed by $^{1}H/^{15}N$ NMR correlation spectroscopy. The results obtained show that only the H48A and the H68A substitutions significantly reduce RegB activity without changing its ability to bind the substrate or affecting its overall structure. Altogether, our results define RegB as a new cyclizing RNase and present His48 and His68 as potent catalytic residues. The effect of the *in vivo* selected R52L mutation is also described and discussed.

INTRODUCTION

The RegB protein, encoded in the bacteriophage T4 genome, is responsible for the specific inactivation of several of the phage early mRNAs, favoring therefore the transition between the early and the middle phase. RegB is an RNase that cleaves very specifically in the middle of GGAG tetranucleotides (1), with a strong bias toward those found in the region responsible for the initiation of translation (the Shine–Dalgarno region). RegB has, in particular, little if any activity on the GGAG found in the coding regions or in the Shine–Dalgarno regions of late mRNAs. Several studies have established that this effect is due to a different activity of RegB towards different GGAG classes and not to an inactivation of the enzyme in the middle phase or a protection of GGAG in the coding sequence (by the ribosome, for example) (2). In fact, we demonstrated in a previous study that RegB preferentially cleaves the GGAG involved in a particular secondary structure, in which the last G of the tetranucleotide belongs to a short stem separating two non-structured loops (3). Another striking point is that RegB activity, as measured *in vitro*, is very low, but can be enhanced by a factor up to 100-fold in the presence of the ribosomal protein S1 (4), the S1 protein being involved in the recognition of the Shine–Dalgarno region by the *Escherichia coli* ribosome.

From the molecular point of view, RegB is a 153 amino acid protein whose structure is still unknown. Its sequence appears very conserved among the phages related to T4, but presents no homology to other proteins, in particular to the other known RNases. As a first step toward the determination of the molecular basis of RegB activity and specificity, we wanted to characterize the catalytic mechanism of the enzyme.

The RNases can be separated into two classes depending on whether they produce 5′- or 3′-phosphate extremities. The first class regroups many intracellular endo- and exoribonucleases such as those of the RNase H family (RNase H1, H2, HI, HII and HIII) (5,6), the prokaryotic RNases II, III, E and PNPase (7) or the mammalian RNases XRN1 and HKE1 (8). One of their characteristics is that they all possess a metal ion in their active site. The second class is mainly represented by the families of barnase (9), RNase A (10–12), RNase T1 (13) and RNase T2 (14,15). The catalysis is ensured by two side chains, one acting as a general base, the other as a general acid. The base activates the 2′-hydroxyl of the ribose which becomes able to attack the phosphodiester bond, generating a cyclic 2′,3′-phosphodiester. The acid gives its proton to the living 5′-hydroxyl group. It was proposed initially that the cyclic product would be an intermediate, immediately hydrolyzed in a second step necessary to recover the original state of ionization of the active site residues; however, it was shown that the cyclic 2′,3′-phosphodiester is the true reaction product, whose hydrolysis takes place only when all of the polynucleotide substrate has been used (16). In addition, many structural and mutational studies have shown that, in all known cases, the general acid catalyst is always a histidine, while the general base catalyst can be a histidine or a glutamate (17–21).

*To whom correspondence should be addressed. Tel: +33 1 69 33 48 55; Fax: +33 1 69 33 30 10; Email: francois.bontems@polytechnique.fr

RegB produces a 3′-phosphate extremity (1). It thus seems reasonable to suppose that its mechanism is similar to that described previously. However, the RegB sequence is unrelated to those of barnase, RNase A, RNase T1 and RNase T2. In addition, its very low activity and very narrow specificity make it unique among RNases. Therefore, we could not exclude a completely different mechanism. We thus decided to check the production of a cyclic phosphate by RegB and the involvement of histidine residues in its mechanism.

MATERIALS AND METHODS

Site-directed mutagenesis

The ExSite™ PCR-based Site-directed Mutagenesis Kit (Stratagene) was used to construct the four histidine-to-alanine RegB mutants. The oligonucleotides, given hereafter 5′ to 3′, were purchased from Eurogentec and purified on a polyacrylamide gel: 421 (GCT CTA AAA TAT TCT CAA CAT CTT CTT GAT CGC G) and 422 (AAA CGA TGA GAC TCC TGC TGC TTT TG) for the H42A substitution; 481 (GCT CTT CTT GAT CGC GCA ATT CAA CGG) and 482 (TTG AGA ATA TTT TAG ATG AAA CGA TGA GAC TCC) for the H48A mutation; 681 (GCT AAA ATA AAA GAC CAT GTT TTA GAA GTT AAT G) and 682 (GAA TAA TTC AAA AAC GTA TGT CTC ATC AAT CTC) for the H68A mutation, and 731 (GCT GTT TTA GAA GTT AAT GAA TTC CTG AGT ATG C) and 732 (GTC TTT TAT TTT ATG GAA TAA TCA AAA ACG TAT G) for the H73A mutation. The DNA template was the plasmid pARNU2 (2) containing the wild-type copy of the *regB* gene. This plasmid yielded the plasmids pH42A, pH48A, pH68A and pH73A.

Purification of different RegB forms using plasmids pARNU2, pH42A, pH48A, pH68A and pH73A

Transformed *E.coli* strain XL1-Blue (Stratagene) was grown at 37°C in Luria–Bertani broth medium [1.0% nutriment broth (Difco), 0.5% yeast extract (Difco), 0.8% NaCl, 0.052% Tris] containing 100 μg ml^{-1} ampicillin (Sigma) and supplemented with 0.2% (w/v) maltose. The protein expression was induced at OD_{600} = 0.6 by adding $MgSO_4$ (10 mmol l^{-1}) and phage λCE6 (Novagen) to a final concentration of 2×10^9 p.f.u. ml^{-1}. Cells were harvested 3 h later and the RegB protein purified, at 4°C, by cobalt affinity chromatography (Clontech) with lysis buffer [50 mmol l^{-1} sodium phosphate pH 8.0, 300 mmol l^{-1} NaCl, 0.1 mmol l^{-1} aminoethylbenzenesulfonyl fluoride (AEBSF) protease inhibitor (Euromedex) and 5.0 mmol l^{-1} β-mercaptoethanol], wash buffer (lysis buffer supplemented with 10 mmol l^{-1} imidazole) and elution buffer (lysis buffer supplemented with 150 mmol l^{-1} imidazole).

Purification of ^{15}N-labeled RegB samples for NMR analysis

An overnight preculture of transformed *E.coli* strain BL21(DE3) (Novagen) was inoculated to 1 l of the M9 minimal medium [12.8 g l^{-1} ($Na_2HPO_4 \cdot 2H_2O$), 3.0 g l^{-1} KH_2PO_4, 0.5 g l^{-1} NaCl, 1.0 g l^{-1} $^{15}NH_4Cl$, 0.1 mmol l^{-1} $CaCl_2$, 1.0 mmol l^{-1} $MgSO_4$ and 4.0 g l^{-1} glucose] containing 100 μg ml^{-1} ampicillin. Protein expression was induced at OD_{600} = 0.8 for 3 h using 1.0 mmol l^{-1} isopropyl-β-D-thiogalactopyranoside (IPTG) (Promega). The protein was purified as described above.

RNA synthesis

The Selex22tb RNA was synthesized on a Pharmacia LKB Gene Assembler Plus using phenoxyacetyl β-RNA phosphoramidites (Amersham Pharmacia Biotech). The protocol was adapted from that used for DNA synthesis, the main difference being the use of 5-ethylthio-1H-tetrazol as activator (22). The terminal 5′-*O*-dimethoxytrityl group was removed in the synthesizer at the end of the synthesis. The oligoribonucleotides were cleaved from the support and deprotected according to the procedure described in Snoussi and Leroy (22). The fragment purity was checked using a Q-Sepharose HPLC column running on a Beckman system and, due to the very low amount of the abortive forms, was not purified further.

In vitro assay of cleavage of Selex22tb RNA by different RegB forms

Cleavage reactions were performed at 37°C in 4 mmol l^{-1} sodium phosphate pH 8.0, 24 mmol l^{-1} NaCl and 0.4 mmol l^{-1} β-mercaptoethanol. The reaction mixture contained 3.6 μmol l^{-1} of the RNA substrate Selex22tb, 0.36 μmol l^{-1} RegB and 40 U of RNA Guard® RNase inhibitor (Pharmacia) to prevent cleavage by contaminating RNase A-like enzymes. The reaction was followed for 40 min. Aliquots were taken every 10 min and the reaction was stopped by adding formamide (40% final) and heating for 5 min at 95°C. The substrate and the products were separated by gel electrophoresis (20% acrylamide, 7 mol l^{-1} urea), stained with ethidium bromide (0.25 μg l^{-1}) for 7 min, washed for 30 s with sterile water and detected with a UV light transilluminator coupled to a digital acquisition card (Vilbert Lourmat). The band intensities were quantified using Bio1D software (Vilbert Lourmat) and corrected for the RNA length (21 nt for the RNA substrate Selex22tb, and 15 and 6 nt for the products). The 6 nt product is too short to fix sufficient ethidium bromide, and hence we quantified the appearance of the 15 nt product.

NMR experiments

All nuclear magnetic resonance (NMR) experiments were carried out on a Bruker DRX600 spectrometer equipped with a TXI triple resonance probe [$^1H/^{15}N$ heteronuclear single quantum correlation (HSQC) spectra] or a ^1H-BB probe (phosphorus experiments). The spectra were processed off-line using GIFA software (23).

The formation of the 2′,3′-cyclic phosphodiester could be followed conveniently by ^{31}P NMR. Two samples were prepared, containing 300 μmol l^{-1} of the Selex22tb substrate with 30 μmol l^{-1} of either the RegB enzyme or the R52L inactive mutant previously dialyzed against a Tris cleavage buffer (final concentration in the sample: Tris 5 mmol l^{-1} pH 8.0, NaCl 30 mmol l^{-1}, β-mercaptoethanol 0.5 mmol l^{-1}). The reaction was followed at 25°C by recording 175 successive ^{31}P spectra (carrier: 2915.24 Hz, spectral width: 7300 Hz and transients: 1024). The pH of the negative control sample (R52L mutant) was slightly different from that of the reaction sample (wild-type RegB), resulting in a small shift of the peaks corresponding to the residual amount of inorganic phosphorus.

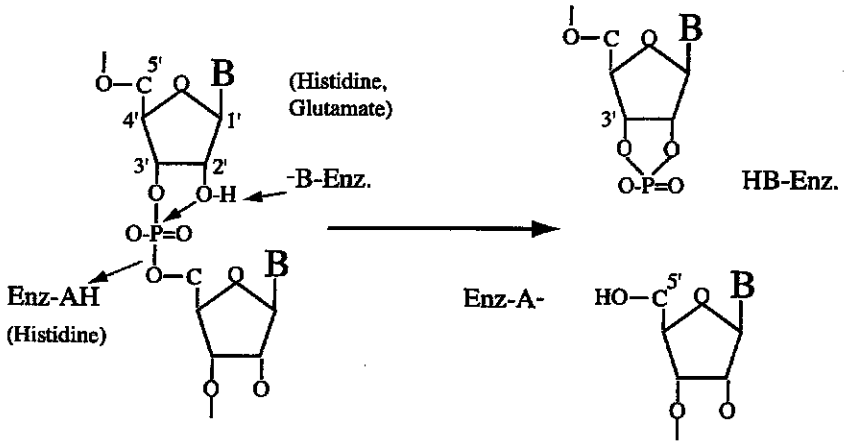


Figure 1. Proposed acid–base catalysis of phosphodiester bond cleavage by barnase, RNase A, RNase T1 and RNase T2. HB-Enz. and $^{-}$B-Enz. (Enz-HA and Enz-A^{-}) denote, respectively, the acidic and the basic forms of the residue acting as a general base catalyst (general acid catalyst).

The ^{15}N-labeled samples of the R52L, H48A and H68A RegB mutants were prepared in a sodium phosphate 50 mmol l^{-1} pH 6.0, 300 mmol l^{-1} NaCl, 5 mmol l^{-1} β-mercaptoethanol buffer at a final concentration of ~500 µmol l^{-1}. The HSQC spectra were recorded at 25°C using the fast HSQC sequence (24). Each experiment consisted of 264 rows of 512 points, each row being the sum of 256 transients. The spectral widths were set to 36 p.p.m. (^{15}N) and 5.5 p.p.m. (^{1}H), and the frequency offsets were to 7300 Hz (^{15}N) and 4800 Hz (^{1}H). The spectra were processed to obtain real matrices of 512 × 512 points.

Surface plasmon resonance experiments

The experiments were carried out using a BIAcore 2000 instrument (BIAcore) with a CM5 sensor chip maintained at 25°C. The protein was coupled to the surface of the sensor chip using CO_2H/NH_2 chemistry. Flow cell 1 was used as a blank surface. The running buffer consisted of 10 mmol l^{-1} HEPES pH 7.4, 150 mmol l^{-1} NaCl, 3 mmol l^{-1} EDTA and 0.005% polysorbate 20 (v/v). A 70 µl aliquot of RNA was injected, at 25 µl min^{-1}, at different concentrations ranging from 75 to 140 µmol l^{-1}. At least two independent injections with two different RNA concentrations were carried out for each RegB form. Regeneration of the sensor chip was achieved by two short 50 µl pulses of NaCl 1 mol l^{-1} and NaOH 50 mmol l^{-1}. All sensorgrams were corrected for background and bulk refractive index by subtracting the reference flow cell signal. The 1:1 Langmuir binding interaction model was used to fit the data with the BIA evaluation 3.0 software (BIAcore).

RESULTS AND DISCUSSION

Characterization of the reaction product

The cleavage of the GGAG tetranucleotide by RegB leads to the appearance of 3′-phosphate and 5′-OH extremities (1). This suggests that the mechanism of RegB is similar to that of barnase, RNase A, RNase T1 and RNase T2-related enzymes. The true product of these enzymes is a cyclic 2′,3′-phosphodiester (Fig. 1). We thus wondered whether RegB was also producing a cyclic phosphodiester or directly releasing a linear 3′-monoester. To check this, we analyzed the product of the reaction by ^{31}P phosphorus NMR. Indeed, the chemical shift of the ^{31}P nuclei involved in a linear 3′,5′-phosphodiester bond and in a 3′-monoester range from –2.0 to –0.5 and from 3.0 to 4.0 p.p.m., respectively while that of a cyclic 2′,3′-phosphodiester resonates near 19 p.p.m. (25,26). The nature of the product of the reaction (linear or cyclic) can thus be determined unambiguously on a simple one-dimensional spectrum.

As substrate, we chose the Selex22tb RNA (GGUGCGA-GAAAACGG↓AGCACC), already studied in this laboratory (3). This substrate derives from the 22nd clone obtained in a study intended to isolate good RegB substrates by the SELEX method (Systematic Evolution of Ligands by EXponential enrichment) (27). We previously showed that Selex22tb is cleaved efficiently by RegB, in the absence of the ribosomal protein S1 (3). In addition, it is sufficiently short to be produced by chemical synthesis on the mg scale. The reaction was carried out at 25°C (instead of 37°C) because RegB is stable at this temperature for several days. We recorded 175 successive one-dimensional experiments. Each experiment corresponded to 1024 transients. We indeed anticipated a very faint signal for the product; about 40 times lower than that of the substrate. RegB indeed cleaves only one phosphodiester bond among the 20 present in the 21 nt Selex22tb. The cleavage, in addition, is not quantitative, the maximal amount of product obtained corresponding to 50–60% of the substrate (3). Finally, the reaction was performed twice, once in the presence of the wild-type RegB enzyme and once in the presence of the inactive R52L mutant (28). The mutant enzyme was produced and purified rigorously in the same conditions as the wild-type and used as a negative control, to ensure that the observed reaction was not due to a contaminant RNase.

Figure 2A represents the results obtained when we mix the wild-type RegB protein with Selex22tb at a ratio of 1:10. In this two-dimensional-like spectrum, the *x*-axis represents the ^{31}P chemical shifts, and the *y*-axis the time course of the

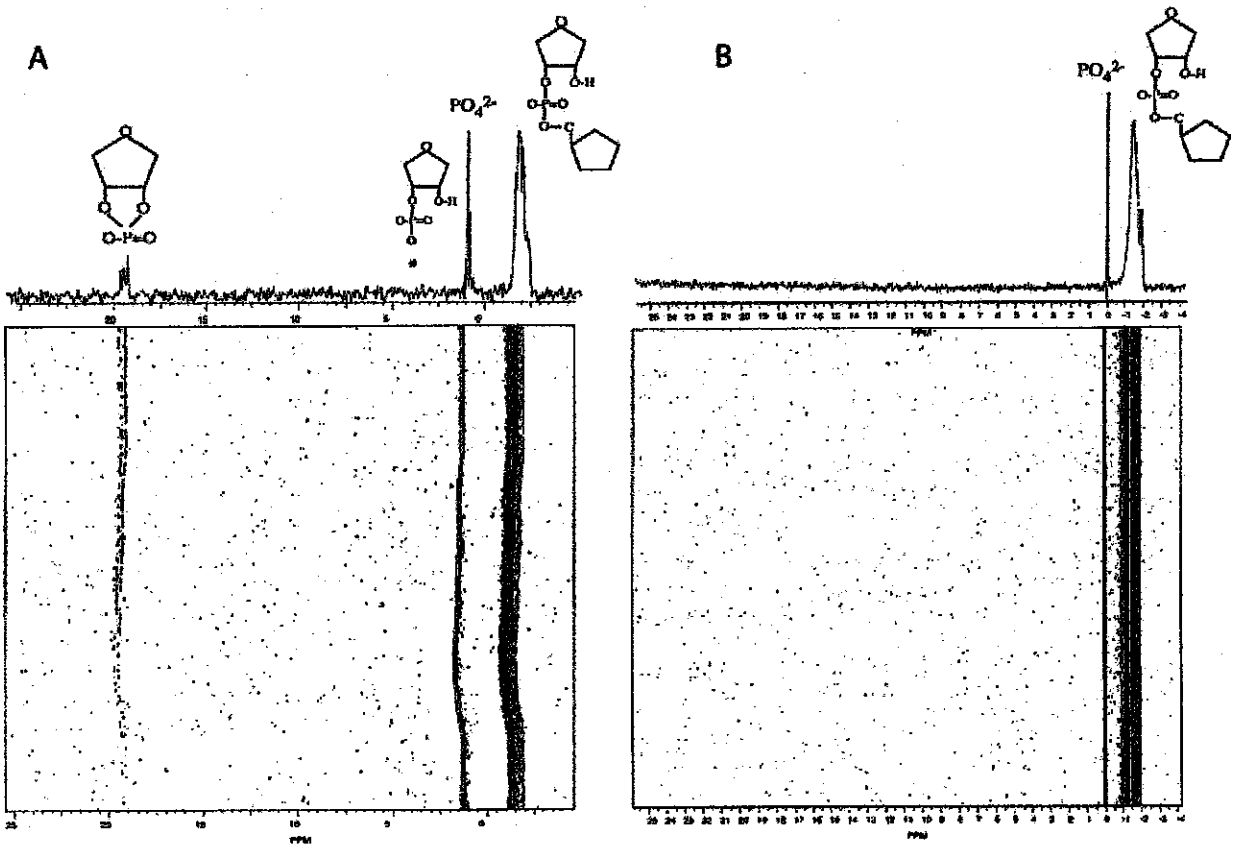


Figure 2. Detection, by ^{31}P NMR, of a 2′,3′-cyclic phosphorus product during the cleavage of the substrate Selex22tb RNA by RegB. (A) Cleavage by the wild-type RegB protein. (B) Negative control experiment using the inactive point mutant RegB R52L. The asterisk indicates the position of the hydrolyzed form of the cyclic product. A peak at this position could be observed provided that the experiment was recorded for a much longer time.

reaction. The one-dimensional spectrum above is the sum of the last 10 rows. The broad region centered at –1.5 p.p.m. represents the different phosphorus nuclei of the substrate Selex22tb. The narrow peak at 1.5 p.p.m. corresponds to residual phosphate molecules originating from the first dialysis buffer. At 19 p.p.m., we clearly identified a peak that resonates in the region of the 2′,3′-cyclic phosphorus. This peak is likely to be the cyclic compound produced by RegB. It is absent at time zero (beginning of the reaction) and gradually accumulates during the time course of the reaction. In addition, it is not seen when Selex22tb is mixed with the RegB inactive mutant R52L (Fig. 2B), indicating that it results from the catalytic activity of the enzyme and is, in particular, not due to an RNase A-like contaminant. We anticipated a sharp single peak corresponding to the cyclic product. This is not the case in Figure 2A where we can see a broad multiplet at 19 p.p.m. Multiple solution conformations of the cyclic 15 nt product could explain this observation.

Functional properties of the RegB histidine-to-alanine mutants

All the RNases described to date which produce a cyclic 2′,3′-phosphodiester possess a histidine in their active site (17–21). We thus decided to check whether this was also the case for RegB. The phage T4 RegB sequence possesses five histidines (H17, H42, H48, H68 and H73). We checked in GenBank all the sequences that have a detectable homology to RegB. We found only 16 sequences. All these sequences encode RegB proteins originating from 16 T4-related phages (L.Piesiniene, unpublished data). Among these 16 sequences, 15 were not of interest because they were almost identical (99.4–100%) to the T4 RegB protein. However, one sequence, originating from phage RB69, was informative (only 77.5% identity). Figure 3 presents the alignment of the RegB proteins from phage T4 and phage RB69. It shows that histidines H42, H48, H68 and H73 are conserved. They were thus mutated into alanines by a PCR-based site-directed protocol (see Materials and Methods). The occurrence of each mutation was controlled by DNA sequencing.

The template for the mutagenesis was the plasmid pARNU2 (2) which contains a copy of the wild-type *regB* gene under the control of the T7 promoter. This plasmid can be propagated in a BL21 *E.coli* strain (devoid of the T7 polymerase) but not in BL21(DE3) (possessing the T7 polymerase) even in the absence of inducer. Indeed, the leaky expression of the T7 polymerase in this latter strain is sufficient to produce enough of the highly toxic RegB protein to prevent bacterial growth (F.Saïda, unpublished data). Since the fully inactive mutant

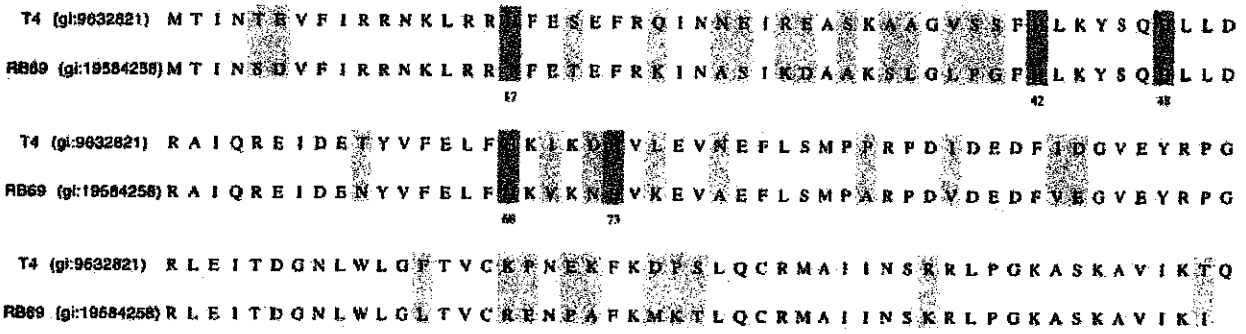


Figure 3. Sequence alignment of the RegB proteins from phage T4 and phage RB69. Four histidines (H42, H48, H68 and H73) out of five are conserved.

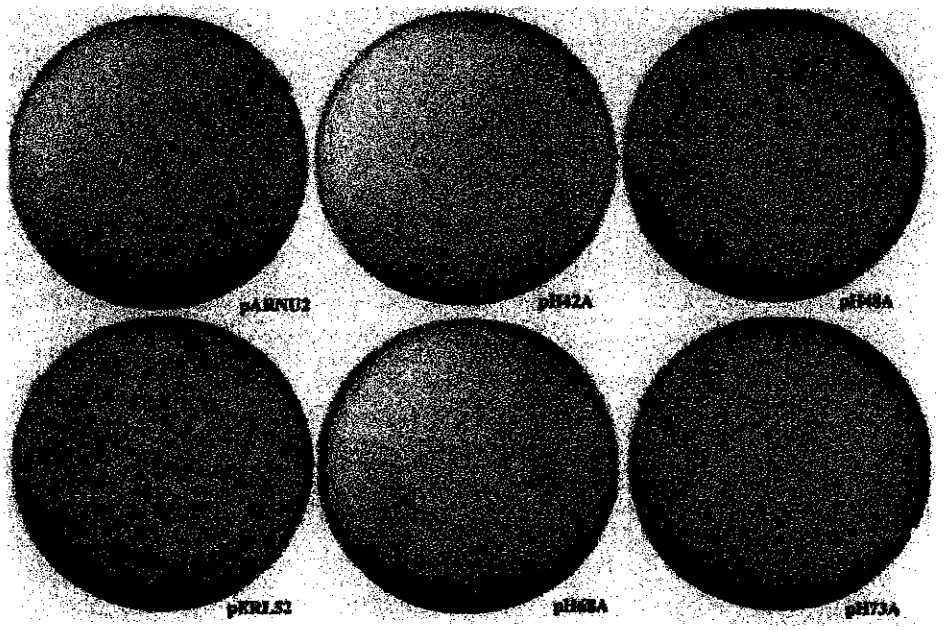


Figure 4. *In vivo* toxicity test of RegB mutants. The viability of *E.coli* strain BL21(DE3) was checked after transformation with plasmids pARNU2, pERL52, pH42A, pH48A, pH68A and pH73A expressing, respectively, the wild-type RegB protein, the inactive point mutant R52L (positive control) and the four histidine-to-alanine mutants. The only mutants that are maintained in BL21(DE3) are H48A and H68A.

R52L is not toxic, we hypothesized that RegB toxicity is linked to its activity and thus speculated that a histidine-to-alanine mutant reducing the enzyme activity would also reduce its toxicity. Such a mutant would thus be easily isolated by checking the viability of a transformed BL21(DE3) strain. Figure 4 shows the result of the electro-transformation of BL21(DE3) cells with the four plasmids resulting from the mutation protocol (pH42A, pH48A, pH68A and pH73A). The plasmid pARNU2 was used as a negative control, and pERL52 (coding for the inactive R52L mutant) as a positive control. Only pH48A and pH68A are maintained in BL21(DE3), suggesting that the activity of the corresponding mutants, RegB H48A and RegB H68A, is reduced. However, other hypotheses, such as a folding default or the disappearance of a critical interaction with another partner, for example, could not be excluded at this stage.

Kinetics of cleavage of the Selex22tb RNA by RegB mutants

The wild-type RegB protein and the four histidine-to-alanine mutants were expressed in *E.coli* strain XL1-Blue using the phage λCE6 induction system (see Materials and Methods). The inactive point mutant RegB R52L was produced from a pET15b-derivative plasmid. The quick non-radioactive RNA cleavage assay described in Materials and Methods was used to compare the cleavage ability of the wild-type RegB protein and the H42A, H48A, H68A, H73A and R52L RegB mutants. All cleavage reactions were monitored for 40 min in the same conditions in the presence of an RNase A-like RNases inhibitor. The appearance of the 15 nt RNA product plotted as function of time is presented in Figure 5. As expected, the R52L mutant is devoid of any activity, whereas the wild-type

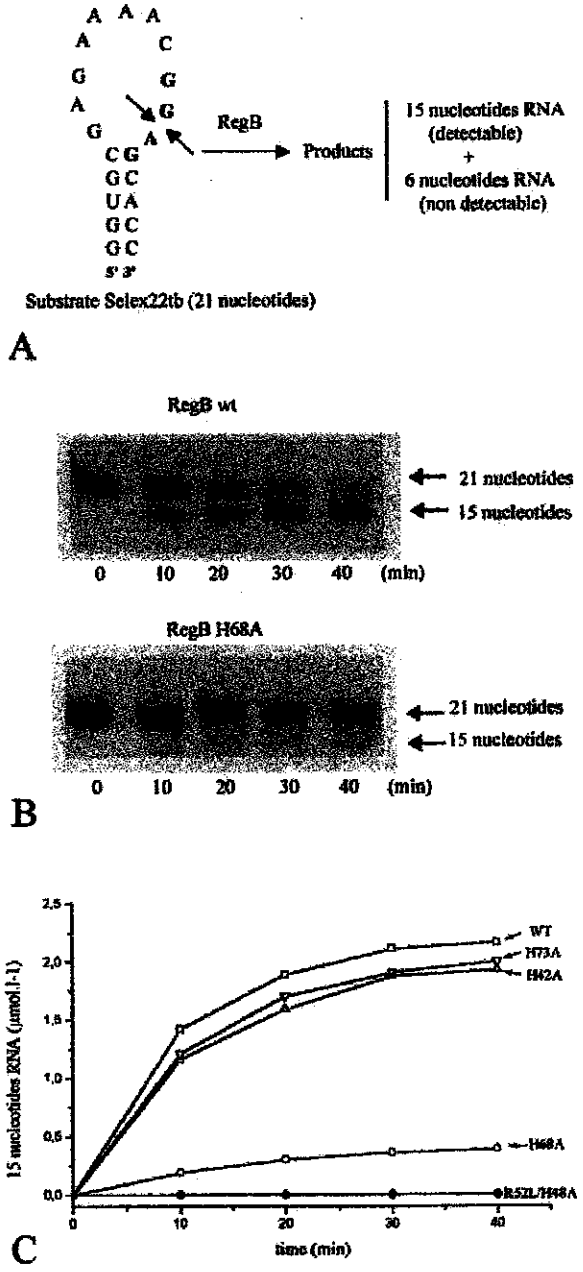


Figure 5. Cleavage reaction kinetics of the Selex22tb RNA by different forms of RegB. (A) The 21 nt Selex22tb RNA is cleaved specifically by RegB in the middle of the GGAG motif, leaving 15 and 6 nt fragments. These fragments are separated on a 20% polyacrylamide gel and stained with ethidium bromide. The light emission of the UV-excited ethidium bromide, in RNA, was recorded using a Vilbert Lourmat Imager and quantified using Biol D V99 software. (B) Example of two UV-excited gels showing the cleavage of the Selex22tb RNA by the wild-type and the H68A RegB forms. (C) Appearance of the 15 nt product during cleavage of Selex22tb RNA by different forms of RegB (wild type, H42A, H48A, R52L, H68A and H73A).

RegB cleaves Selex22tb up to 58% in 40 min. The H42A and H73A mutants cleave the substrate with a rate similar to that of the wild-type. A slight difference is in fact observed (55% for the mutant instead of 58% for the wild-type) but does not seem to be significant because it is within the experimental error. In contrast, RegB activity is completely abolished by the H48A substitution and dramatically reduced by the H68A mutation. This result is in accordance with the colony phenotype observed in Figure 4 where we notice that colonies transformed with the partially active H68A mutant are smaller (the growth rate is lower) than those transformed with the H48A mutant (devoid of any activity). We hypothesized that these substitutions either disrupt the overall structure of RegB or highlight the significance of H48 and H68 in the binding of the substrate or in the direct catalysis of the cleavage reaction. To distinguish between these three hypotheses, we decided to probe by $^{1}H/^{15}N$ correlation spectroscopy the overall structure of RegB H48A and RegB H68A and to quantify the RNA-binding capabilities of these two mutants by surface plasmon resonance (SPR).

$^{1}H/^{15}N$ correlation spectroscopy

^{15}N uniformly labeled samples of RegB H48A and RegB H68A were prepared using the BL21(DE3)/IPTG induction system (see Materials and Methods). The $^{1}H/^{15}N$ HSQC spectra (24) were recorded on a Bruker proton 600 MHz spectrometer. Each peak on the two-dimensional spectrum represents the backbone H/N correlation of one amino acid residue, and its special position reflects the effect of the structural environment. Hence, the HSQC spectrum can be considered as a two-dimensional fingerprint of the three-dimensional structure of the protein. Since the HSQC spectrum of the wild-type RegB protein is unavailable because RegB is too toxic to be produced on the NMR scale, we compared the HSQC spectra of the H48A and the H68A mutants with that of the independent inactive point mutant RegB R52L. The results are presented in Figure 6. We highlighted using circles the highly resolved widespread peaks of the RegB H48A HSQC spectrum. These peaks originate from residues in a well-structured environment and thus represent a significant probe of the protein scaffold (this probe is more relevant than circular dichroism which only describes the secondary structure composition of the protein). All circle-enclosed peaks are found in the same positions in the spectrum of RegB H68A and of the independent mutant RegB R52L. Since it is highly improbable that three independent and different substitutions cause the same structural alteration, we suggest that the amino acid substitutions H48A and H68A (as well as R52L) did not affect the overall structure of the protein.

Binding capabilities of RegB mutants

We used the SPR optical phenomenon to monitor the interactions between RegB mutants and the RNA Selex22tb (29). SPR provides both equilibrium and kinetic information about the RNA–protein interactions. We hypothesized that the formation and breakdown of the complex between RegB and the RNA Selex22tb are dictated by both association (k_a) and dissociation (k_d) rates, as described by

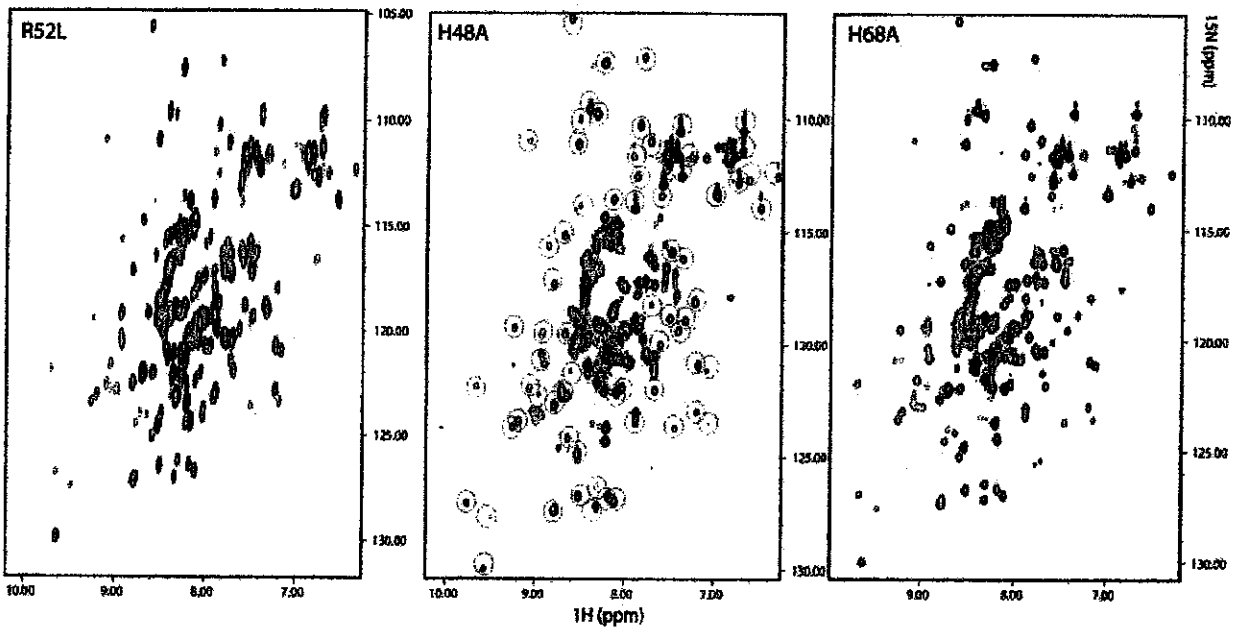


Figure 6. $^{1}H/^{15}N$ HSQC spectra of the R52L, H48A and H68A RegB mutants recorded at 25°C on a Bruker DRX600 spectrometer. The encircled peaks in the spectrum of RegB H48A are found in the same positions in the spectra of the two independent mutants RegB R52L and RegB H68A.

$$\text{RegB} + \text{RNA} \underset{k_d}{\overset{k_a}{\rightleftharpoons}} \text{Regb:RNA} \qquad 1$$

The dissociation constant can then be calculated according to

$$K_D = k_d/k_a \qquad 2$$

Equation 1 assumes that we neglect the breakdown of the RegB:RNA complex by enzymatic cleavage (in the case of the wild type and the mutants that have a residual activity). This assumption is justified by the fact that we recorded SPR sensorgrams over a short time period (2.8 min) and at 25°C. In these conditions, the percentage of cleavage of Selex22tb will be much lower than 10%. Indeed, Figure 5 shows that the percentage of cleavage, after 2.8 min, is 10% at 37°C, and it is known, according to the law of Arrhenius, that the rate of a catalyzed reaction is decreased by a significant factor (about 2) when the temperature drops from 37 to 25°C. The results of SPR measurements are presented in Table 1. We notice that the wild-type RegB protein and the mutants H42A, H48A and H73A have similar dissociation constants within the range of experimental error. This demonstrates that the introduced mutations, in particular H48A, do not change the binding capabilities of RegB. The affinity of the inactive mutant R52L is decreased by an order of magnitude compared with that of the wild type, suggesting that Arg52 may be important for the binding of the substrate RNA. The result for the H68A mutant seems to be slightly different from that of the wild-type: 2.5 μmol l^{-1} for H68A and 1.5 μmol l^{-1} for the wild type. This difference of 1.0 μmol l^{-1} in the K_D may not be significant according to experimental errors (Table 1).

The fact that the *in vitro* cleavage kinetics and the HSQC NMR experiments show that the R52L, H68A and H48A mutations inactivate (completely or partially) the RegB enzyme without altering its structure supports the idea that these mutations result in a perturbation of the global catalytic process (binding of the substrate and/or direct chemical catalysis). The SPR results indicate that the R52L mutation induces a 10-fold decrease in the affinity between the enzyme and its substrate. Arg52 may be important for binding of the substrate, but we cannot exclude that it is also involved in the direct chemical catalysis because the decrease in the affinity for the RNA substrate (10-fold) seems too weak to abolish the activity of the enzyme completely. In contrast, the H48A and H68A mutations have no significant effect on the binding constant. In addition, the H48A mutant has no detectable activity, while H68A conserves a small (~14%) but significant residual activity. These results suggest that H48 and H68 may be directly involved in the catalysis of the reaction mediated by RegB but in a different fashion. At this stage, several hypotheses could be presented. H48 may be a crucial residue involved directly in the acid–base catalysis

Table 1. Kinetics and affinity constants at 25°C for the interactions of wild-type RegB protein and the five RegB point mutants with the Selex22tb RNA substrate

Protein	k_a (M^{-1} s^{-1})	k_d (10^{-4} s^{-1})	K_D (μM)
RegB wild type	75 ± 3	1.1 ± 0.3	1.5 ± 0.5
RegB R52L	4.0 ± 0.5	0.60 ± 0.03	15 ± 2
RegB H42A	75 ±5	1.5 ± 0.2	2.0 ± 0.3
RegB H48A	53 ± 2	1.0 ± 0.1	1.9 ± 0.3
RegB H68A	64 ± 10	1.6 ± 0.4	2.5 ± 0.3
RegB H73A	64 ± 10	1.4 ± 0.5	2.2 ± 0.2

Values were recorded on a BIAcore 2000 SPR apparatus.

(acceptor or donor of protons). H68 may polarize chemical bonds, enhancing the protons transfer, or interact with a particular intermediate.

The fact that the reaction product is a 2′,3′-cyclic phosphodiester and the involvement of at least one histidine in the catalytic mechanism strongly suggest that RegB shares the catalytic mechanism of barnase, RNase A, RNase T1 and RNase T2. However, RegB is structurally and functionally unrelated to all these RNases. First, RegB primary structure has no significant homology to any RNase from the barnase, RNase A and RNase T1 families, which have characteristic conserved residues encompassing their catalytic residues (9–12,15). Also, RegB does not possess the two typical cassettes of 9 and 12 residues that characterize the RNases from the RNase T2 family (13–15). Moreover, RegB is an intracellular enzyme (with two probably reduced cysteines, a reducing agent being required to avoid precipitation), whereas barnase, RNase A, RNase T1 and RNase T2 are secreted extracellular enzymes, the members of the three last families possessing several disulfide bridges. Finally, RegB has absolute sequence specificity, towards the GGAG sequence, which is not found in any described RNase. This suggests that RegB could form a new class of enzymes possessing a cyclizing mechanism.

ACKNOWLEDGEMENTS

We are very grateful to Dr J. L. Leroy for helpful advice on RNA chemical synthesis and to Mrs N. Duvignaud for critical reading of the manuscript. This work was supported by a national grant from the French Department of Scientific Research (Ministere de la recherche).

REFERENCES

1. Uzan,M., Favre,R. and Brody,E. (1988) A nuclease that cuts specifically in the ribosome binding site of some T4 mRNAs. *Proc. Natl Acad. Sci. USA*, **85**, 8895–8899.
2. Sanson,B., Hu,R.M., Troitskaya,E., Mathy,N. and Uzan,M. (2000) Endoribonuclease RegB from bacteriophage T4 is necessary for the degradation of early but not middle or late mRNAs. *J. Mol. Biol.*, **297**, 1063–1074.
3. Lebars,I., Hu,R.M., Lallemand,J.Y., Uzan,M. and Bontems,F. (2001) Role of the substrate conformation and of the S1 protein in the cleavage efficiency of the T4 endoribonuclease RegB. *J. Biol. Chem.*, **276**, 13264–13272.
4. Ruckman,J., Ringquist,S., Brody,E. and Gold,L. (1994) The bacteriophage T4 regB ribonuclease. Stimulation of the purified enzyme by ribosomal protein S1. *J. Biol. Chem.*, **269**, 26655–26662.
5. Lima,W.F., Wu,H. and Crooke,S.T. (2001) Human RNases H. *Methods Enzymol.*, **341**, 430–440.
6. Kanaya,S. (2001) Prokaryotic type 2 RNases H. *Methods Enzymol.*, **341**, 377–394.
7. Nicholson,A.W. (1997) *Escherichia coli* ribonucleases: paradigms for understanding cellular RNA metabolism and regulation. In D'Allessio,G. and Riordan,J.F. (eds), *Ribonucleases: Structures and Functions*. Academic Press, New York, pp. 1–49.
8. Ross,J. (1997) RNA-processing RNases in mammalian cells. In D'Allessio,G. and Riordan,J.F. (eds), *Ribonucleases: Structures and Functions*. Academic Press, New York, pp. 553–587.
9. Mauguen,Y., Hartley,R.W., Dodson,E.J., Dodson,G.G., Bricogne,G., Chothia,C. and Jack,A. (1982) Molecular structure of a new family of ribonucleases. *Nature*, **297**, 162–164.
10. Beintema,J.J., Breukelman,H.J., Carsana,A. and Furia,A. (1997) Evolution of vertebrate ribonucleases: ribonuclease A superfamily. In D'Allessio,G. and Riordan,J.F. (eds), *Ribonucleases: Structures and Function*. Academic Press, New York, pp. 245–269.
11. Zhang,J., Dyer,K.D. and Rosenberg,H.F. (2002) RNase 8, a novel RNase A superfamily ribonuclease expressed uniquely in placenta. *Nucleic Acids Res.*, **30**, 1169–1175.
12. Beintema,J.J. (1998) Introduction: the ribonuclease A superfamily. *Cell. Mol. Life Sci.*, **54**, 763–765.
13. Yoshida,H. (2001) The ribonuclease T1 family. *Methods Enzymol.*, **341**, 28–41.
14. Deshpande,R.A. and Shankar,V. (2002) Ribonucleases from T2 family. *Crit. Rev. Microbiol.*, **28**, 79–122.
15. Irie,M. and Ohgi,K. (2001) Ribonuclease T2. *Methods Enzymol.*, **341**, 42–55.
16. Cuchillo,C.M., Pares,X., Guasch,A., Barman,T., Travers,F. and Nogues,M.V. (1993) The role of 2′,3′-cyclic phosphodiesters in the bovine pancreatic ribonuclease A catalysed cleavage of RNA: intermediates or products? *FEBS Lett.*, **333**, 207–210.
17. Thompson,J.E. and Raines,R.T. (1994) Value of general acid–base catalysis to ribonuclease A. *J. Am. Chem. Soc.*, **116**, 5467–5468.
18. Nishikawa,S., Morioka,H., Kim,H.J., Fuchimura,K., Tanaka,T., Uesugi,S., Hakoshima,T., Tomita,K., Ohtsuka,E. and Ikehara,M. (1987) Two histidine residues are essential for ribonuclease T1 activity as is the case for ribonuclease A. *Biochemistry*, **26**, 8620–8624.
19. Steyaert,J., Hallenga,K., Wyns,L. and Stanssens,P. (1990) Histidine-40 of ribonuclease T1 acts as base catalyst when the true catalytic base, glutamic acid-58, is replaced by alanine. *Biochemistry*, **29**, 9064–9072.
20. Irie,M., Ohgi,K., Iwama,M., Koizumi,M., Sasayama,E., Harada,K., Yano,Y., Udagawa,J. and Kawasaki,M. (1997) Role of histidine 46 in the hydrolysis and the reverse transphosphorylation reaction of RNase Rh from *Rhizopus niveus*. *J. Biochem.*, **121**, 849–853.
21. Irie,M., Ohgi,K., Watanabe,H., Iwama,M., Nakamura,K.T., Kurihara,H., Nonaka,T., Mitsui,Y., Horiuchi,H. and Takagi,M. (1994) pH profile of kinetic constants of RNase Rh from *Rhizopus niveus* and its mutant enzymes towards UpU and possible mechanisms of RNase Rh. *J. Biochem.*, **115**, 1083–1087.
22. Snoussi,K. and Leroy,J.L. (2001) Imino proton exchange and base-pair kinetics in RNA duplexes. *Biochemistry*, **40**, 8898–8904.
23. Pons,J.L., Malliavin,T.E. and Delsuc,M.A. (1996) Gifa V.4: a complete package for NMR data set processing. *J. Biomol. NMR*, **8**, 445–452.
24. Mori,S., Abeygunawardana,C., Johnson,M.O., Berg,J. and van Zijl,P.C.M. (1995) Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. *J. Magn. Reson. Ser. B*, **108**, 94–98.
25. Thompson,J.E., Venegas,F.D. and Raines,R.T. (1994) Energetics of catalysis by ribonucleases: fate of the 2′,3′-cyclic phosphodiester intermediate. *Biochemistry*, **14**, 7408–7414.
26. Palmer,H.R., Bedford,J.J., Leader,J.P. and Smith,R.A. (1990) ^{31}P and ^{1}H NMR studies of the effect of the counteracting osmolyte trimethylamine-*N*-oxide on interactions of urea with ribonuclease A. *J. Biol. Chem.*, **275**, 27708–27711.
27. Jayasena,V.K., Brown,D., Shtatland,T. and Gold,L. (1996) *In vitro* selection of RNA specifically cleaved by bacteriophage T4 RegB endonuclease. *Biochemistry*, **35**, 2349–2356.
28. Ruckman,J., Parma,D., Tuerk,C., Hall,D.H. and Gold,L. (1989) Identification of a T4 gene required for bacteriophage mRNA processing. *New Biol.*, **1**, 54–65.
29. Katsamba,P.S., Park,S. and Laird-Offringa,I.A. (2002) Kinetic studies of RNA–protein interactions using surface plasmon resonance. *Methods*, **26**, 95–104.

PUBLICATION III

Saida, F., Odaert, B., Uzan, M. et Bontems, F. (2003) First structural investigation of the restriction ribonuclease RegB: NMR spectroscopic conditions, $^{13}C/^{15}N$ Double-Isotopic labeling and two-dimensional heteronuclear spectra. *Protein. Expr. Purif.* (Sous presse).

First structural investigation of the restriction ribonuclease RegB: NMR spectroscopic conditions, $^{13}C/^{15}N$ double-isotopic labelling and two-dimensional heteronuclear spectra

Fakhri Saïda[1], Benoît Odaert[1], Marc Uzan[2] and François Bontems[*,1]*

[1] Laboratoire ICSN-RMN, Institut de Chimie des Substances Naturelles, Ecole polytechnique, route de Saclay, 91128, Palaiseau Cedex, France.

[2] Institut Jacques Monod, UMR 7592, CNRS, Universités Paris 6 et Paris 7, 2 place Jussieu, Paris, France.

* To whom correspondence should be addressed. Phone: 0033169334855. Fax: 0033169333010. E-mail: francois.bontems@polytechnique.fr

ABSTRACT

The bacteriophage T4 genome-encoded ribonuclease RegB is the unique well-defined restriction endoribonuclease. This protein cleaves with an almost absolute specificity its RNA substrate in the middle of the GGAG tetranucleotide mainly found in the Shine-Dalgarno sequence (required for the prokaryotic initiation of the translation). This protein has no significant homology to any known ribonuclease and its structure has never been investigated. The extreme toxicity of this ribonuclease prevents the expression of large quantities for structural studies. Here, we show that the toxicity of RegB can be bypassed by using the RegB H48A point mutant and explain why resolving the structure of this mutant is relevant. For nuclear magnetic resonance (NMR) purposes, we report the preparation of highly pure $^{13}C/^{15}N$ double-labelled millimolar samples of RegB H48A using a high yield expression procedure in minimal medium (30 mg/liter). We also present a set of solution conditions that maintain the millimolar samples of this protein stable for long periods at the NMR-required temperature. Finally, we present the first $^{1}H/^{15}N$ and $^{1}H/^{13}C$ two-dimensional NMR spectra of RegB H48A. These spectra show that the protein is folded and that the full structural analysis of RegB by NMR is feasible.

Keywords: RegB; Restriction ribonuclease; Toxicity; Phage T4; Solution NMR.

INTRODUCTION

The endoribonuclease RegB, encoded in the genome of the bacteriophage T4, controls the expression of many viral early genes [1-3]. RegB inactivates its substrate mRNA by cleaving with high specificity the consensus sequence GGAG mainly found in the intergenic regions involved in the efficient initiation of the translation (Shine-Dalgarno sequence) [4]. Previous studies have established that, at least *in vitro*, the activity of RegB can be enhanced by a factor up to 100 in the presence of a particular component of the prokaryotic ribosome: the ribosomal protein S1 [5-6]. The *in vitro* substrate selection method SELEX (Systematic Evolution of Ligands by EXponential enrichment) has established that the GGAG motif is necessary for the cleavage of RegB in presence of the S1 protein [7] and subsequent enzymatic studies have shown that among the GGAG containing molecules, RegB preferentially cleaves those in which the motif is involved in a particular secondary structure [8]. However, the accurate mechanism by which RegB recognizes its substrate remains unknown.

Resolving the three dimensional structure of RegB (153 residues) will provide a valuable mean to understand the molecular basis of the RegB/RNA specific interactions. Structural studies require the production of large quantities of highly pure RegB protein. High-resolution NMR studies require in addition the production of $^{13}C/^{15}N$ isotope-enriched samples that are stable for several days at a concentration near millimolar. Unfortunately, RegB is extremely toxic to *Escherichia coli.* This ribonuclease cannot be maintained in any efficient T7-based expression vector [9]. In this study, we summarize the results of an extensive investigation of the wild type RegB protein expression using non-conventional procedures. We quantitatively compare the yields of these expression procedures with those of five RegB point mutants selected either *in vivo* or *in vitro* (by site-directed mutagenesis). We show that we can bypass

RegB overproduction problems by using the single point mutant RegB H48A in which the residue histidine 48 was mutated into alanine. We discuss the relevance of studying the structure of this point mutant by the light of functional evidences [10].

Overproduction and purification of $^{13}C/^{15}N$ doubly labelled RegB H48A at the yield of 30 mg per liter of minimal medium is reported and the solution conditions that ensure the long-term stability of millimolar RegB samples are presented. The two-dimensional $^{1}H/^{15}N$ and $^{1}H/^{13}C$ heteronuclear single quantum correlation spectra show that the protein is folded and that a full structure determination of RegB is feasible.

MATERIALS AND METHODS

Expression and purification of the wild type RegB protein by fermentation

E. coli strain XL1-Blue was transformed with the pARNU2 plasmid [3] containing the wild type *regB* gene. Cells were grown overnight at 37 °C in 800 mL Luria Bertani Broth medium [1.0% Nutriment Broth (Difco), 0.5% Yeast Extract (Difco), 0.8% NaCl, 0.052% Tris] containing 100 $\mu g.mL^{-1}$ ampicillin (Sigma) supplemented with 0.2 % D(+) maltose (Sigma). This 800 mL preculture was used to inoculate a 20 L Fermentor containing 15 L of the LB medium supplemented with 100 $\mu g.mL^{-1}$ ampicillin and 0.2 % D(+) maltose. The protein expression was induced at OD_{600}= 0.6 by adding $MgSO_4$ (10 mM) and phage λCE6 (Novagen) (final concentration: 2.10^9 $pfu.mL^{-1}$). Cells were harvested 3 hours later and resuspended in the A buffer [50 mM sodium phosphate (pH 8.0), 300 mM NaCl, 0.1 mM AEBSF protease inhibitor (Euromedex) and 5.0 mM β-mercaptoethanol]. Cells were then mechanically disrupted at 4 °C using the DYNO®-Mill KD apparatus. The polyhistidines tagged RegB protein was purified, at 4 °C, by cobalt affinity chromatography (Clontech) using the wash buffer (buffer A supplemented with 10 mM imidazole) and the elution buffer (buffer A supplemented with 150 mM imidazole).

Expression of the wild type RegB protein using the thermo-regulated pOURegB system

The pOURegB plasmid contains the *regB* gene under the control of the T7 promoter and the mutated RBS Phi10 cloned in the runaway-replication plasmid pOU61 [9,11]. *E. coli* BL21(λDE3) cells transformed with pOURegB were grown at 32 °C in the LB medium supplemented with 100 $\mu g.mL^{-1}$ ampicillin until OD_{600}= 0.6. Adding 1 mM IPTG and shifting the temperature from 32 °C to 42 °C for one hour then from 42 °C to 37 °C for 90 min induced protein expression. The protein was purified according to the standard procedure.

Expression of the wild type RegB protein in the yeast *S. cerevisiae*

The shuttle vector pTOXR-1 contains the *regB* gene under the control of the eukaryotic GAL1 promoter and CYC1 terminator [9]. *S. cerevisiae* strain INVSc1 (Invitrogen) transformed with pTOXR-1 was inoculated into the SM6 medium [0.7 % Yeast Nitrogen Base (Difco), 0.01 % L-tryptophan (Sigma), 0.01 % L-leucine (Sigma), 0.005 % L-histidine (Sigma), 0.1 % bacto-casaminoacids (Difco) and 2.0 % D(+)-glucose (Sigma)] and cultured overnight at 28 °C. The culture was then centrifuged and resuspended in the induction medium [SM6 medium without glucose and supplemented with 2.0 % D(+)-galactose (Sigma)] to OD_{600}= 0.4. Cells were harvested 8 hours later and the protein purified according to the standard procedure.

Cloning of the RegB H48A mutant in the pTOXR-1 positive selection and expression vector

An insert containing the coding sequence of RegB H48A was obtained by PCR from the pH48A plasmid [10] using primers NUb (5' CGA CGT CGA TCT CGA TCC CGC 3') and NU13 (5' CGG GAT CCT CAT TGA GTT TTA AT 3') and the high fidelity Pfu DNA Polymerase (Promega). The PCR program, run on a Perkin Elmer GeneAmp PCR system 2400, was: [(94 °C: 3 min) X 1, (94 °C: 30 s, 42 °C: 55 s, 72 °C: 55 s) X 22, (72 °C: 7 min) X 1]. This insert contains an NdeI site within the ATG initiator codon and a BamHI site just upstream of the stop codon of *regB*. 50 ng of the positive selection vector pTOXR-1 [9] initially amplified in yeast were digested for 15 min with NdeI and BamHI, ethanol precipitated and ligated to the digested insert using the T4 DNA ligase at 16 °C for 2 hours. The insert to vector ratio was equal to 4. The ligation mixture was n-butanol precipitated [12] and used to transform electro-competent XL1-Blue cells using the Eppendorf Electroporator 2510. The recombinant plasmid was named pEH48A.

Screening for suitable buffer, ionic strength and pH conditions

The RegB H48A protein, initially in a standard buffer [50 mM sodium phosphate (pH 8.0), 300 mM NaCl, 0.1 mM AEBSF protease inhibitor (Euromedex) and 5.0 mM β-mercaptoethanol], was divided into several 50 μl samples. The pH (from 4 to 10), the ionic strength (NaCl concentrations from 50 mM to 600 mM) and the buffering molecule (sodium phosphate or sodium citrate) of these samples were modified by using the very practical Micro Bio-Spin® Chromatography columns (Biorad).

Overexpression and purification of the $^{13}C/^{15}N$ labelled RegB H48A

A 50 mL overnight preculture of the *E. coli* strain BL21(λDE3)-pLysS (Novagen) transformed with the pEH48A plasmid was inoculated to 1 L of the M9 minimal medium [12.8 g.L^{-1} (Na_2HPO_4, $2H_2O$); 3.0 g.L^{-1} KH_2PO_4; 0.5 g.L^{-1} NaCl; 1.0 g.L^{-1} $^{15}NH_4Cl$; 0.1 mM $CaCl_2$, 1.0 mM $MgSO_4$ and 2.0 g.L^{-1} glucose] containing 100 μg.mL^{-1} ampicillin and 34 μg.mL^{-1} chloramphenicol. The protein expression was induced at OD_{600}=0.6 for 3 hours using 1.0 mM isopropyl-β-D-galactopyranoside IPTG (Promega). Cells were then harvested, resuspended in the B1 buffer [50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 0.1 mM AEBSF protease inhibitor (Euromedex) and 5.0 mM β-mercaptoethanol] and stored at –80 °C until use. The whole purification procedure was made at 4°C. Cells were disrupted by sonication. The crude cell extracts were clarified by centrifugation at 15 000 rpm for 40 min and mixed with the Talon cobalt affinity resin (Clontech) for 40 min. The resin was pelleted by centrifugation for 2 min at 800 g, washed once with the B1 buffer and once with the B2 buffer (B1 buffer supplemented with 20 mM imidazole), put into a chromatography column and washed again with 6 bed volumes of the B2 buffer. The protein was eluted with the B3 buffer (B1 buffer supplemented with 150 mM imidazole). Finally the elution buffer was exchanged with the B4 NMR buffer [50 mM sodium citrate (pH 6.0), 300 mM NaCl, 0.1 mM AEBSF

protease inhibitor (Euromedex) and 2.0 mM dithiothreitol] using the Micro Bio-Spin® Chromatography columns (Biorad).

NMR spectroscopy

All NMR experiments were carried out on a Bruker DRX600 spectrometer equipped with a TXI triple resonance probe. The $^{1}H/^{15}N$ HSQC spectrum was recorded on a millimolar sample of RegB H48A at 32.5 °C using the fHSQC sequence [13]. Each experiment consisted in 264 rows of 512 points, each row being the sum of 256 transients. The spectral widths were set to 36 ppm (^{15}N) and 5.5 ppm (^{1}H) and the frequency offset to 7300 Hz (^{15}N) and 4800 Hz (^{1}H). The spectra were processed to obtain 512 x 512 points real matrices. The $^{1}H/^{13}C$ HSQC spectrum was recorded in 100 % deuterium oxide to completely suppress the water resonances. We used a two-dimensional $^{1}H/X$ correlation sequence via double INEPT transfer [14]. Each experiment consisted in 196 rows of 1024 points, each row being the sum of 128 transients. The spectral widths were set to 90 ppm (^{13}C) and 12.97 ppm (^{1}H) and the frequency offset to 6639 Hz (^{13}C) and 2824 Hz (^{1}H). The spectra were processed to obtain 2048 x 512 points real matrices. The spectra were processed using the XWINNMR 3.0 software (Bruker) and analysed using the Sparky Software (University of California San Francisco, Thomas L. Goddard).

RESULTS AND DISCUSSION

Choice of the RegB H48A point mutant

The wild type ribonuclease RegB is extremely toxic to the bacterium *E. coli* [9]. This toxicity is probably linked to the capability of RegB to inactivate some of the host essential mRNA by directed cleavage within the GGAG motif found in the ribosome-binding sites of about 32 % of *E. coli* mRNA [15-16]. RegB cannot be expressed using the standard efficient T7-based vectors [9]. Extensive investigation of non-conventional expression techniques revealed four low expressing systems. These systems are a T7 promoter coupled to a mutated gene Phi10 ribosome-binding site and induced by infection with the λCE6 phage [3], a thermo-regulated origin of replication (see materials and methods), the expression in the eukaryotic system *S. cerevisiae* (see materials and methods) and the adaptation of the λCE6 induction system with the large scale production in Fermentor (see materials and methods). The expression yields of the wild type RNase RegB using these four systems as well as the optimised expression yields of the five available RegB point mutants are presented in Table 1. The RegB point mutant R52L was recovered by *in vivo* selection [2]. The four others mutants were constructed by site-directed mutagenesis in a previous study that investigated the catalytic mechanism of RegB [10].

Table 1 shows that not only the expression yields of the RegB wild type protein are insufficient for structural studies (in particular for NMR) but also require tedious procedures very difficult to adapt in minimal medium for the isotopic labelling. Table 1 shows that only two RegB versions have their expression yields compatible with a structural study by NMR: RegB H48A and RegB R52L. These two completely inactive mutants can be easily expressed using the standard procedure [transformation of the *E. coli* strain BL21(λDE3) and induction with IPTG]. However, we think that the mutant RegB H48A represents the best candidate to

build a structural model of the ribonuclease RegB for two reasons. First, since RegB cleaves its RNA substrate via a 2', 3'-cyclic phosphodiester like ribonucleases from the RNase A [17], the RNase T1 [18], the RNase T2 [19] and the Barnase [20] groups, it is very likely that histidine 48 side-chain acts as a general acidic or general basic catalyst [10]. Therefore, its displacement by alanine would preserve the structural integrity of the RegB active site. This assumption is confirmed in the case of many RNases such as the RNase T1 where the structural investigation by X-Ray crystallography revealed the perfect matching between the wild type protein and the catalytic point mutant H40A [21-22]. In addition, surface plasmons resonance revealed that, unlike the RegB R52L mutant, RegB H48A binds its RNA substrate with the same affinity as the wild type protein [10]. Therefore, we decided to investigate the solution structure of RegB H48A by NMR.

The expression yield of RegB H48A presented in Table 1 was obtained using the homemade positive selection and expression vector pTOXR-1 [9] (see materials and methods). pTOXR-1 contains a copy of the wild-type *regB* gene flanked by the T7 promoter, the T7 Phi10 ribosome-binding site and the T7 Phi transcription terminator. This plasmid, constructed and recovered in yeast, cannot be maintained in *E. coli* unless a foreign insert disrupts the wild type *regB* gene coding sequence. The structure of the recombinant pEH48A plasmid is presented in Figure 1. As few as 500 mL of the rich Luria Bertani broth medium are needed to prepare an NMR sample of the non-labelled RegB H48A protein.

Screening for the suitable NMR conditions to study RegB H48A by NMR: temperature, pH, ionic strength, and buffering molecule

To check the suitable NMR conditions for studying RegB H48A, we have prepared millimolar samples in a standard phosphate buffer at pH 8.0 containing 5 mM of β-mercapthoethanol (without this reducing agent RegB precipitates because of the oxidation of

its two cysteines). We first looked for the suitable temperature by recording several one-dimensional proton spectra ranging from 15 °C to 37 °C. The spectra showed that the peaks width decreases with the increasing temperature (see the example in Figure 2). The best-resolved spectrum is obtained at 32.5 °C. Unfortunately; we found that millimolar samples of RegB precipitate at this temperature within 18 hours. Because three-dimensional NMR experiments need several days to be recorded, it was important to find conditions that enhance the stability of RegB at 32.5 °C. We first investigated the influence of the pH by preparing different samples ranging from pH 4.0 to pH 9.0. Very acidic pH were prepared using the sodium acetate buffer, the others using the sodium phosphate buffer. We also screened different ionic strengths corresponding to NaCl concentrations ranging from to 600 mM. For each condition, a small millimolar protein sample (50 μl) was pipeted into a 200 μl thin-wall PCR vial and incubated for 3 days in a PCR apparatus programmed to maintain the temperature at 32.5 °C. The stability of the protein sample was monitored by checking the precipitation profile in the vial. The lowest precipitation profile was obtained at pH 6.0 and at 300 mM NaCl ionic strength. Finally, we checked the influence of the buffering molecule by comparing the thermal stability of the protein in the 50 mM sodium phosphate buffer at pH 6.0 and in a sodium citrate buffer at the same concentration and the same pH. This choice was motivated by our suspicion that, at high concentration, phosphate ions may decrease the stability of the ribonuclease RegB known for its capability to interact with phosphate containing compounds, especially phosphodiester bonds. In the presence of sodium citrate molecules, we noticed no visible precipitation after a 3 days-incubation at 32.5 °C.

These findings show that the suitable conditions to investigate the NMR structure of RegB H48A are a temperature of 32.5 °C and a millimolar protein sample in 50 mM sodium citrate buffer at pH 6.0 with 300 mM NaCl. Because the reducing agent β-mercapthoethanol

is not very stable for long-term experiences, we replaced it with 2 mM dithiothreitol in the final protein sample.

Preparation of $^{13}C/^{15}N$ double-isotopic labelled RegB H48A

Resolving by NMR the three-dimensional structure of a 153 residues protein like RegB H48A needs to record triple resonance experiments on $^{13}C/^{15}N$ double-isotopic labelled samples. The *E. coli* strain BL21(λDE3)-pLysS (Novagen), transformed with the plasmid pEH48A and cultured on a minimal medium containing ^{13}C-glucose as a Carbon source of and $^{15}NH_4Cl$ as a Nitrogen source, was used to label the RegB H48A protein. The expression protocol is described in Materials and Methods. Because phosphate buffers decrease the stability of RegB at high concentration, because it is not possible to purify RegB at pH 6.0 using a cobalt-affinity chromatography (at this acidic pH, protonated imidazole groups will prevent the binding of the polyhistidines tag to the positively charged cobalt cation) and because sodium citrate buffering power is limited to pH values ranging from 2.1 to 6.4, we needed to purify RegB first in a Tris-HCl buffer at pH 8.0 and then dialyse the pure protein against the sodium citrate buffer described above. Figure 3 shows the SDS-PAGE analysis of the $^{13}C/^{15}N$ RegB H48A purification. The purity of the sample is greater than 95 %. Therefore there is no need, for NMR experiments, to add a second purification step. The final expression-purification yield is estimated to 30 mg per liter of double-labelled M9 minimal medium.

NMR spectroscopy

A millimolar sample of $^{13}C/^{15}N$ double-labelled RegB H48A containing 10 % of deuterium oxide (to lock the NMR signal) and 0.1 % of NaN_3 to prevent microbial growth was used to record a $^{1}H/^{15}N$ heteronuclear single quantum correlation spectrum (HSQC) on a

600 MHz DRX spectrometer (Bruker). The $^{1}H/^{13}C$ HSQC spectrum was recorded in the same conditions except that H_20 molecules were completely exchanged with D_2O molecules to suppress the water resonances and enhance the quality of the spectrum in this region.

Figure 4 presents the obtained $^{1}H/^{15}N$ HSQC spectrum. This spectrum has a good dispersion of the amide-proton resonances over 3.5 ppm (from 6.5 to 10), which shows that the expressed protein is folded. The spectrum contains about 140 proton-nitrogen backbone resonances. This number is in agreement with the 145 expected resonances (the 153 residues RegB protein has 8 prolines devoid of $^{1}H/^{15}N$ correlation signals).

Figures 5 presents the details of the two major parts of the recorded $^{1}H/^{13}C$ HSQC spectrum. Figure 5A shows the $^{1}H/^{13}C$ cross peaks corresponding to the methyl groups (CH_3) the protons of which resonate between -1 and 2.5 ppm. These intense and well resolved peaks are very helpful for building a three-dimensional model of RegB. Figure 5B shows the $^{1}H/^{13}C$ cross peaks that resonate within the water region (4.8 ppm) in the proton dimension. A few peaks are found downfield of the water region (chemical shift greater than 4.8 ppm), which suggest that the β-sheet content of the RegB H48A secondary structure is probably poor but not nil.

In conclusion, it is now possible to produce easily large quantities of the double-labelled restriction ribonuclease RegB and to maintain the millimolar-concentrated samples stable for several days at high temperature. The quality of the recorded ^{13}C and ^{15}N two-dimensional spectra indicates that building a three-dimensional model of RegB by NMR is feasible. The sequential assignment of the RegB backbone using three-dimensional $^{1}H/^{15}N/^{13}C$ experiments is already under work.

Acknowledgment

We thank Drs. Ouazzani J. and Lopez P. for providing fermentation facilities. This work is supported by a national grant from the French Department of Scientific Research (Ministère de la Recherche) for Saïda F. and an ICSN postdoctoral fellowship for Odaert B.

REFERENCES

[1] M. Uzan, R. Favre, E. Brody, A nuclease that cuts specifically in the ribosome binding site of some T4 mRNAs, Proc. Natl. Acad. Sci. USA 85 (1988) 8895-8899.

[2] J. Ruckman, D. Parma, C. Tuerk, D.H. Hall, L. Gold, Identification of a T4 gene required for bacteriophage mRNA processing, New Biol. 1 (1989) 54-65.

[3] B. Sanson, R.M. Hu, E. Troitskaya, N. Mathy, M. Uzan, Endoribonuclease RegB from bacteriophage T4 is necessary for the degradation of early but not middle or late mRNAs, J. Mol. Biol. 297 (2000) 1063-1074.

[4] J. Shine, L. Dalgarno, The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome-binding sites, Proc. Natl. Acad. Sci. USA 71 (1974) 1342-1346.

[5] J. Ruckman, S. Ringquist, E. Brody, L. Gold, The bacteriophage T4 regB ribonuclease. Stimulation of the purified enzyme by ribosomal protein S1, J. Biol. Chem. 269 (1994) 26655-26662.

[6] M. Bisaglia, S. Laalami, M. Uzan, F. Bontems, Activation of the endoribonuclease RegB by the S1 ribosomal protein is due to cooperation between the S1 four C-terminal modules in a substrate dependent-manner, J. Biol. Chem. 278 (2003) 15261-15271.

[7] V.K. Jayasena, D. Brown, T. Shtatland, L. Gold, In vitro selection of RNA specifically cleaved by bacteriophage T4 RegB endonuclease, Biochemistry 35 (1996) 2349-2356.

[8] I. Lebars, R.M. Hu, J.Y. Lallemand, M. Uzan, F. Bontems, Role of the substrate conformation and of the S1 protein in the cleavage efficiency of the T4 endoribonuclease RegB, J. Biol. Chem. 276 (2001) 13264-13272.

[9] F. Saida, M. Uzan, J.Y. Lallemand, F. Bontems, New system for positive selection of recombinant plasmids and dual expression in yeast and bacteria based on the restriction ribonuclease RegB, Biotechnol. Prog. 19 (2003) 727-733.

[10] F. Saida, M. Uzan, F. Bontems, The phage T4 restriction ribonuclease RegB: A cyclizing enzyme that requires two histidines to be fully active, Nucleic Acids Res. 31 (2003) 2751-2758.

[11] J.E. Larsen, K. Gerdes, J. Light, S. Molin, Low-copy-number plasmid-cloning vectors amplifiable by derepression of an inserted foreign promoter, Gene 28 (1984) 45-54.

[12] M.R. Thomas, Simple, effective cleanup of DNA ligation reactions prior to electro-transformation of *E. coli*, Biotechniques 16 (1994) 988-990.

[13] S. Mori, C. Abeygunawardana, M.O. Johnson, J. Berg, P.C.M. van Zijl, Improved Sensitivity of HSQC Spectra of Exchanging Protons at Short Interscan Delays Using a New Fast HSQC (FHSQC) Detection Scheme that Avoids Water Saturation, J. Magn. Reson. Series B 108 (1995) 94-98.

[14] G. Bodenhausen, D.J. Ruben, Natural abundance Nitrogen-15 NMR by Enhanced Heteronuclear Spectroscopy, Chem. Phys. Lett. 69 (1980) 185-189.

[15] M. Uzan, Bacteriophage T4 RegB Endoribonuclease, Methods Enzymol. 342 (2001) 467-480.

[16] D. Barrick, K. Villanueba, J. Childs, R. Kalil, T.D. Schneider, C.E. Lawrence, L. Gold, G.D. Stormo, Quantitative analysis of ribosome binding sites in *E. coli*, Nucleic Acids Res. 22 (1994) 1287-1295.

[17] J.J. Beintema, Introduction: the ribonuclease A superfamily, Cell. Mol. Life Sci. 54 (1998) 763-765.

[18] H. Yoshida, The ribonuclease T1 family, Methods Enzymol. 341 (2001) 28-41.

[19] M. Irie, K. Ohgi, Ribonuclease T2, Methods Enzymol. 341 (2001) 42-55.

[20] Y. Mauguen, R.W. Hartley, E.J. Dodson, G.G. Dodson, G. Bricogne, C. Chothia, A. Jack, Molecular structure of a new family of ribonucleases, Nature 297 (1982) 162-164.

[21] U. Langhorst, R. Loris, V.P. Denisov, J. Doumen, P. Roose, D. Maes, B. Halle, J. Steyaert, Dissection of the structural and functional role of a conserved hydration site in RNAse T1, Protein Sci. 8 (1999) 722-730.

[22] P. Mignon, J. Steyaert, R. Loris, P. Geerlings, S. Loverix, A nucleophile activation dyad in ribonucleases. A combined X-ray crystallographic/ab initio quantum chemical study, J. Biol. Chem. 277 (2002) 36770-36774.

FIGURE LEGENDS

Figure 1: Characteristics of the pEH48A expression vector. DNA sequences are not drawn to scale. The arrows indicate the transcriptional orientation of the corresponding genes and promoters. The *regB* H48A gene (dots) is fused to a DNA sequence (lines) coding a polyhistidines tag (His-Tag). Abbreviations: RBS, Ribosome-Binding Site; AmpR, Ampicillin-resistance determinant; f1 ori, origin of replication that produces single strand DNA upon super-infection with a helper phage; T7p, T7 promoter; T7t, T7 terminator; 2μ ori, yeast 2 micron origin of replication; URA3, orotidine-5'-monophosphate decarboxylase gene; CYC1, iso-1-cytochrome c transcription terminator; GAL1 promoter, galactokinase promoter. Unique restriction sites are underlined.

Figure 2: Amide proton regions from one-dimensional ^{1}H spectra of RegB H48A at 25 °C and 32.5 °C. Spectra were recorded on a Bruker DRX600 spectrometer. The peaks resolution is better at 32.5 °C than at 25 °C.

Figure 3: SDS-PAGE analysis of the purification of $^{13}C/^{15}N$ labelled RegB H48A. M: molecular weights marker. E1 to E5: Elution fractions, CF: crude cell extract after sonication, SF: Soluble fraction, NF: cobalt resin-unbound protein fraction. The calculated molecular weight of the polyhistidines tagged RegB H48A protein is 20.073 kDa.

Figure 4: Two-dimensional $^{1}H/^{15}N$ spectrum HSQC spectrum of RegB H48A. The spectrum was recorded at 32.5 °C on a Bruker 600 MHz spectrometer.

Figure 5: Two-dimensional $^{1}H/^{13}C$ spectrum HSQC spectrum of RegB H48A. (A) Detail showing the methyl-groups region, (B) detail showing the peaks that resonate in the proton dimension near the suppressed water signal (4.8 ppm), (C) full view. The spectrum was recorded at 32.5 °C on a Bruker 600 MHz spectrometer.

Table 1: The different expression systems of the wild type RegB protein and the five point mutants H42A, H48A, R52L, H68A and H73A. The activity of each RegB version is reported to that of the wild type (100 %). Expression yields are in mg of purified protein per liter of culture in the Luria-Bertani rich medium. "Mat. & met." indicates that the original expression procedure is presented in Materials and methods. Mutated RBS Phi10 designates a low-efficiency version of the phage T7 gene Phi10 ribosome-binding site that permits the expression of the fully active RegB versions under the control of the T7 promoter [9].

TABLES

RegB version	Activity (%)	Vector	Expression strategy	Yield (mg/liter)	References
Wild type	100	pARNU2	Mutated RBS Phi10 + λCE6 infection	0.2	[3]
		pOURegB	Mutated RBS Phi10 + thermo-regulated origin of replication	0.6	Mat. & met.
		pYRegB2	Expression in yeast *S. cerevisiae*	0.3	Mat. & met.
		pARNU2	Expression in Fermentor	0.3	Mat. & met.
RegB H42A	100	pH42A	Mutated RBS Phi10 + λCE6 infection	0.2	[10]
RegB H48A	0	pEH48A	Efficient T7-based strategy using vector pTOXR-1	30.0	Mat. & met.
RegB R52L	0	pERL52	Efficient T7-based strategy using vector pET15b	20.0	[9-10]
RegB H68A	~ 10	pEH68A	Efficient T7-based strategy using vector pTOXR-1	6.0	Mat. & met.
RegB H73A	100	pH73A	Mutated RBS Phi10 + λCE6 infection	0.2	[10]

Table 1

FIGURES

FIGURE 1

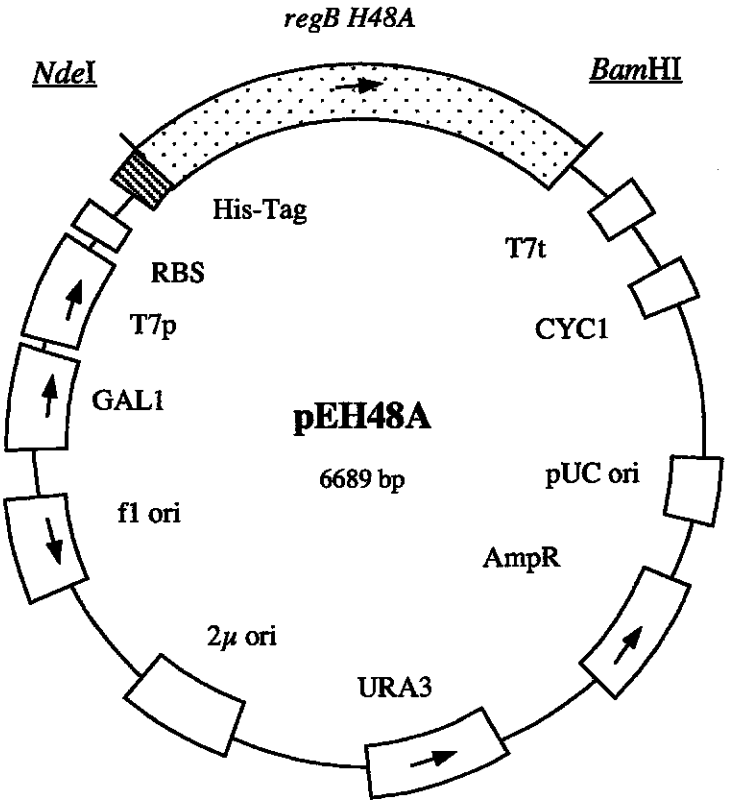


FIGURE 2

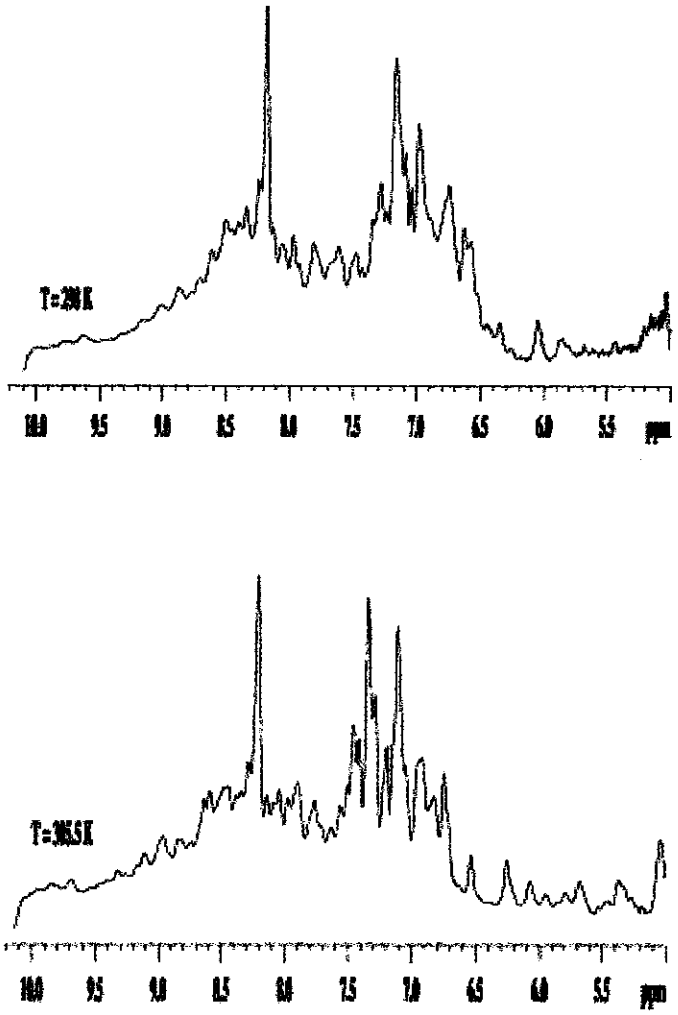

T=200 K
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
ppm
T=265.5 K
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
ppm

FIGURE 3

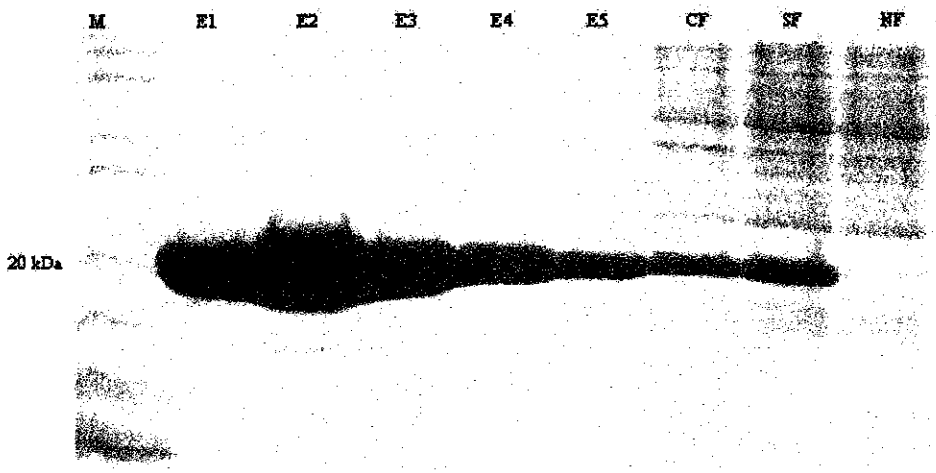


FIGURE 4

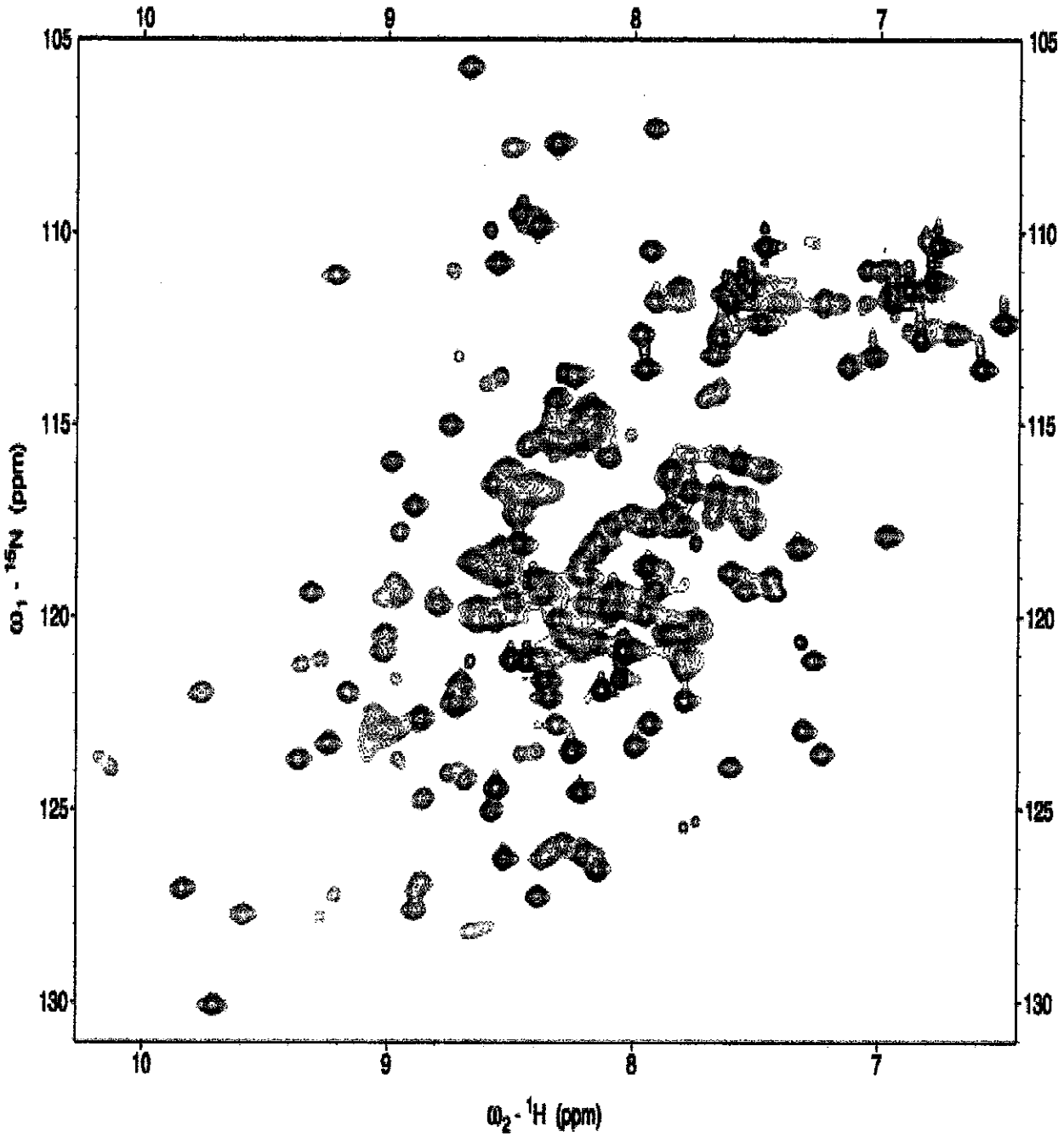


FIGURE 5A

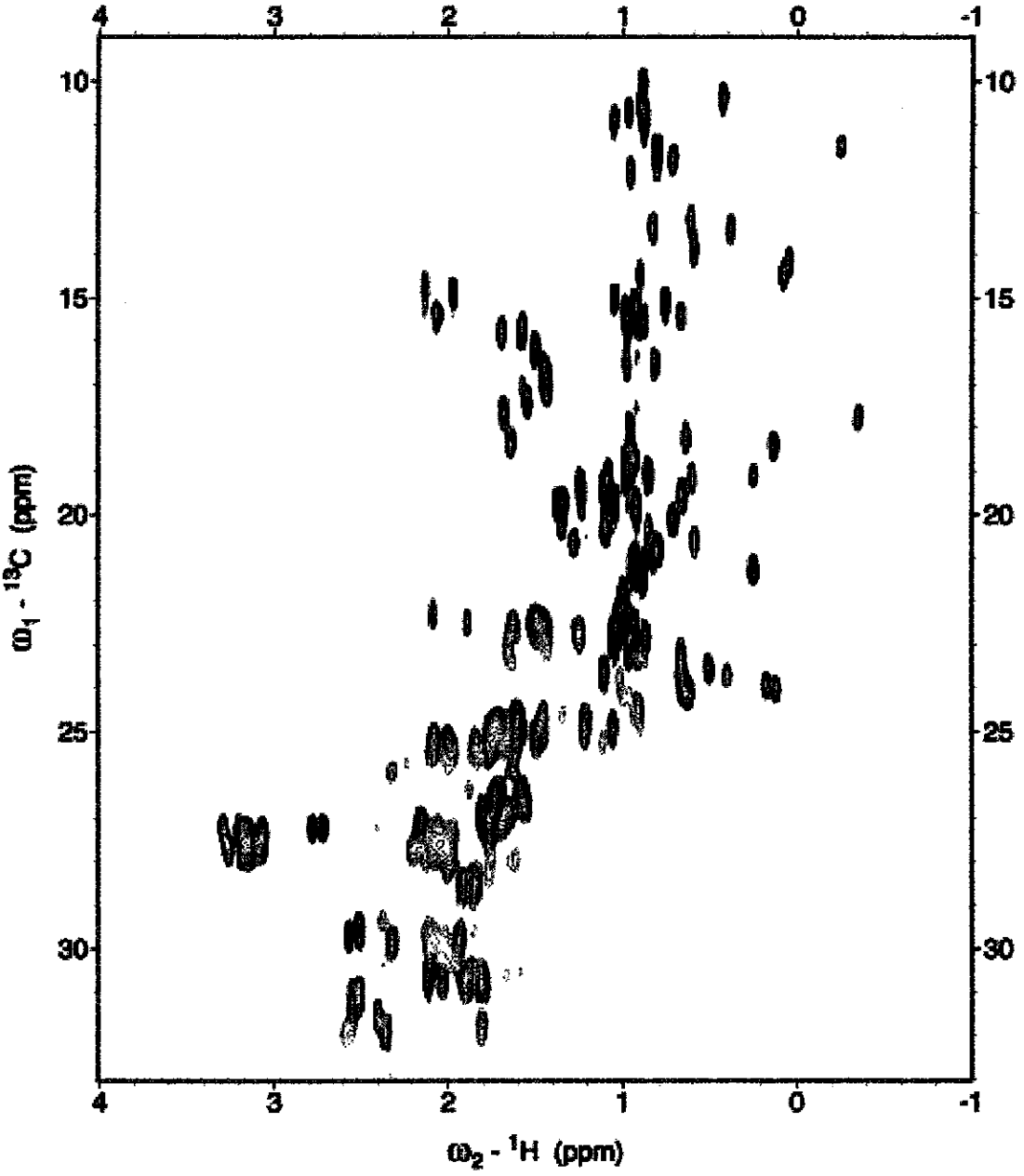


FIGURE 5B

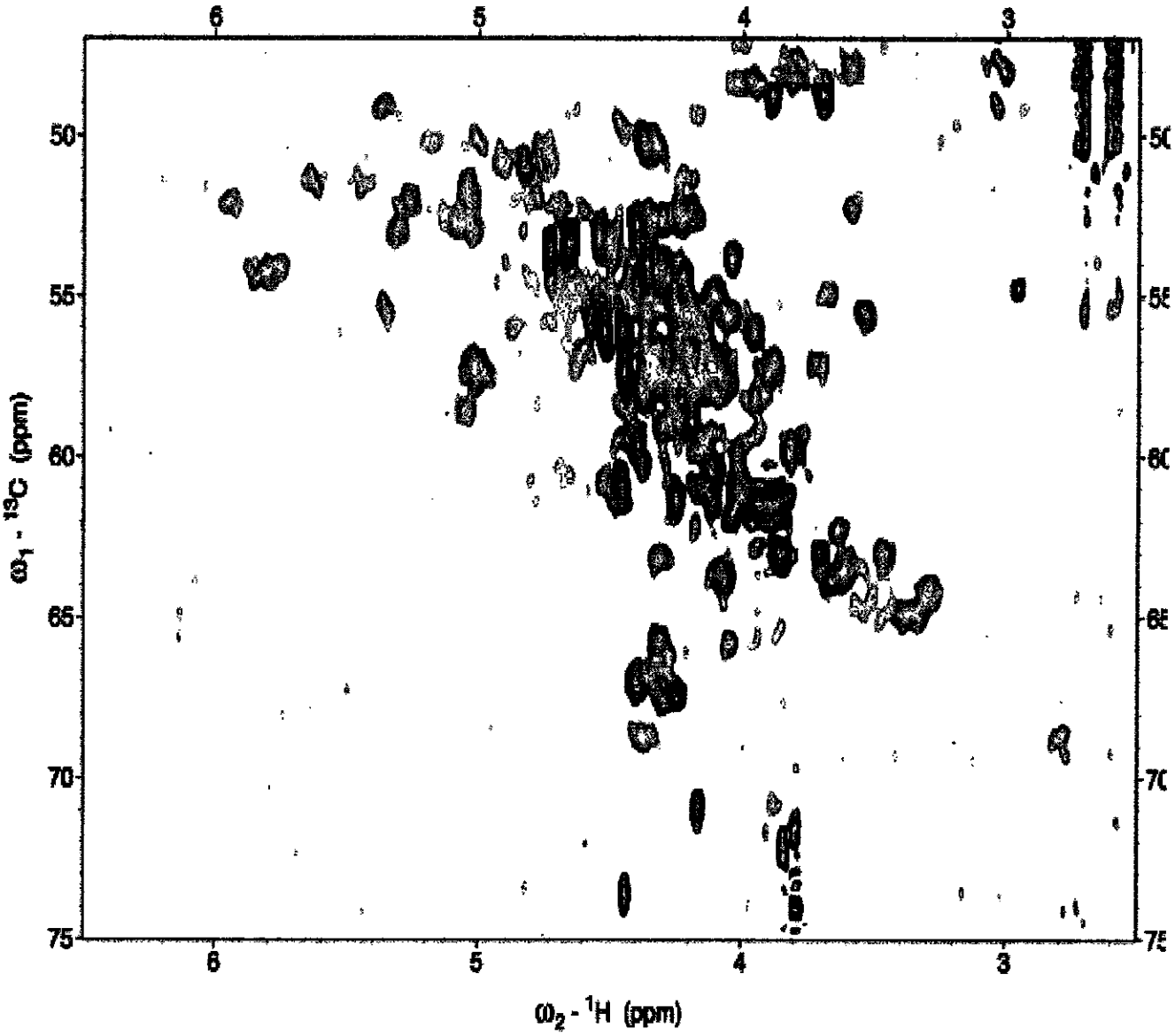


FIGURE 5C

