

Etude de l'organisation transcriptionnelle et de la régulation post-transcriptionnelle du groupe d'ORF codant l'érythrose 4-phosphate déshydrogénase, la phosphoglycérate kinase et la Fructose-1,6-bisphosphate aldolase de classe II chez Escherichia coli: Applications à la production de protéines recombinantes

Corinne Vallet

#### ▶ To cite this version:

Corinne Vallet. Etude de l'organisation transcriptionnelle et de la régulation post-transcriptionnelle du groupe d'ORF codant l'érythrose 4-phosphate déshydrogénase, la phosphoglycérate kinase et la Fructose-1,6-bisphosphate aldolase de classe II chez Escherichia coli : Applications à la production de protéines recombinantes. Biologie moléculaire. Université Henri Poincaré - Nancy 1, 2005. Français. NNT : 2005NAN10191 . tel-01748075

# HAL Id: tel-01748075 https://hal.univ-lorraine.fr/tel-01748075

Submitted on 29 Mar 2018

**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.



#### **AVERTISSEMENT**

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact: ddoc-theses-contact@univ-lorraine.fr

#### LIENS

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10
<a href="http://www.cfcopies.com/V2/leg/leg\_droi.php">http://www.cfcopies.com/V2/leg/leg\_droi.php</a>
<a href="http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm">http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm</a>





#### FACULTE DES SCIENCES & TECHNIQUES

U.F.R Sciences et Techniques Biologiques Ecole Doctorale Biologie, Santé et Environnement

# Thèse

Présentée pour l'obtention du titre de

# Docteur de l'Université Henri Poincaré, Nancy-I

Biologie Moléculaire

## par Corinne VALLET

Etude de l'organisation transcriptionnelle et de la régulation posttranscriptionnelle du groupe d'ORF codant l'érythrose 4-phosphate déshydrogénase, la phosphoglycérate kinase et la fructose-1,6bisphosphate aldolase de classe II chez *Escherichia coli*, application à la production de protéines recombinantes

Soutenue le 19 Juillet 2005 devant la commission d'examen :

Membres du jury:

Rapporteurs: Mr A. Krol (Président du jury) Directeur de Recherche CNRS, Strasbourg

Mr P. Moreau Chargé de Recherche CNRS, Marseille

Examinateurs: Mr Y. Motorin Professeur, U.H.P., Nancy I

Mme E. Hajnsdorf Chargée de Recherche CNRS, Paris

Directeur de thèse : Mme C. Branlant Directeur de Recherche CNRS, Nancy

Je tiens à exprimer ma profonde gratitude à Madame Christiane Branlant pour m'avoir accueillie dans son laboratoire et m'avoir fait découvrir le monde de la recherche. Je la remercie pour la formation scientifique qu'elle m'a apportée. La rédaction de ce mémoire doit beaucoup à ses conseils et sa rigueur scientifique, et le temps important qu'elle y a consacré.

Je suis très sensible à l'honneur que me font Monsieur Alain Krol et Monsieur Patrice Moreau, d'avoir accepté d'être rapporteur de ce mémoire.

Je tiens à remercier très vivement Madame Eliane Hajnsdorf d'avoir accepté de juger ce travail, ainsi que pour chacune de ses marques de soutien et tout le matériel biologique qu'elle m'a fourni tout au long de ce travail. Je tiens également à remercier Monsieur Yuri Motorin pour avoir accepté de faire partie de mon jury de thèse.

Je suis très reconnaissante à Monsieur Bruno Charpentier d'avoir su me donner des conseils pertinents et judicieux, et pour le temps qu'il a consacré tout au long de mon travail de recherche et de la rédaction de ce manuscrit.

Un immense merci à Benoit Thouvenot pour toutes les discussions scientifiques et non scientifiques et pour son soutien pendant une majeure partie de mon travail de thèse.

Je remercie Monsieur Guy Branlant pour m'avoir permis d'avoir un rôle dans l'enseignement à l'UHP, d'abord en tant qu'ATER, puis en tant que vacataire.

Je remercie l'ensemble des personnes du laboratoire pour leur aide, leurs encouragements et l'ambiance de travail qu'ils ont su créer, en particulier Séverine B, Denise, Delphine R, Fabrice N, Sébastien, Nathalie G, Jacquie, Olivier, Florence, François T, Alex O, Alex K, Arnaud, Tof.

Je tiens à remercier vivement mes amis, en particulier Dylan, Marie, Nathalie, Pierre et Christian, ainsi que ma famille, pour leur soutien et leur conseils.

### **Abréviations**

A : adénine

A<sub>600nm</sub> : absorbance à 600 nm aDB : anti-Downstream Box ADN : acide désoxyribonucléique

ADNc : acide désoxyribonucléique complémentaire

ADP : adénosine 5' diphosphate
AMP : adénosine 5' monophosphate
AMPc : adénosine monophosphate cyclique

ARI/II : "Activating Region" I/II ARN : acide ribonucléique

ARNm : acide ribonucléique messager

ARNnc : ARN non codant

ARNr : acide ribonucléique ribosomique ARRBD : Arginine Riche region Binding Domain

aSD : anti-Shine-Dalgarno ATP : adénosine 5' triphosphate

BIME : Bacterial Interspersed Mosaic Elements

BSA : « Serum Albumine Bovine »

C : cytosine Ci : curie

cpm : coups par minute

CLHP : Chromatographie Liquide à Haute Performance

CTP : cytidine 5' triphosphate
D : taux de dilution (h<sup>-1</sup>)
Da/kDa : dalton/kilodalton
DB : Downstream Box

ddNTP : didésoxyribonucléoside 5' triphosphate

DMS : diméthyle sulfate DNase : désoxyribonucléase

dNTP : désoxyribonucléoside triphophate dsRBD : double strand RNA Binding Domain

DTT : dithiothréitol

dUMP: désoxyuridine 5' monophosphatedUTP: désoxyuridine 5' triphosphateECF: Extra Cytoplasmic Function

EDTA : Acide Ethylène Diamine TétraAcétique

EIIA<sup>Gle</sup> : Domaine A de l'enzyme II du système PTS du glucose EIIBC<sup>Gle</sup> : Domaines B et C de l'enzyme II du système PTS du glucose

E4PDH : érythrose 4-phosphate déshydrogénase EDTA : acide éthylène diamine tétra-acétique

EF-Tu : Facteur d'Elongation Tu endoRNase : endoribonucléase

EtOH : éthanol

 $\begin{array}{lll} E\sigma^{24} & : ARN \ polymérase \ holoenzyme \ ayant \ le \ facteur \ \sigma^{24} \\ E\sigma^{28} & : ARN \ polymérase \ holoenzyme \ ayant \ le \ facteur \ \sigma^{28} \\ E\sigma^{32} & : ARN \ polymérase \ holoenzyme \ ayant \ le \ facteur \ \sigma^{32} \\ E\sigma^{38} & : ARN \ polymérase \ holoenzyme \ ayant \ le \ facteur \ \sigma^{38} \\ E\sigma^{54} & : ARN \ polymérase \ holoenzyme \ ayant \ le \ facteur \ \sigma^{54} \\ E\sigma^{70} & : ARN \ polymérase \ holoenzyme \ ayant \ le \ facteur \ \sigma^{70} \\ \end{array}$ 

exoRNase : exoribonucléase

FBA : Fructose 1,6-bisphosphate aldolase

G3P : glycéraldéhyde 3-phosphate

g/mg/µg/ng : gramme/milligramme/microgramme/nanogramme : glycéraldéhyde 3-phosphate déshydrogénase **GAPDH** 

GDP : guanosine 5' diphosphate

: glucose Glc

: guanosine 5' triphosphate GTP

HK : Histidine Kinase HTH : hélice-tour-hélice

**IPTG** : isopropyl-\(\beta\)-thiogalactoside

: constante thermodynamique de formation du complexe fermé  $K_{B}$ 

d'initiation de la transcription

kb

: constantes cinétiques de formation des complexes d'initiation  $k_1/k_2/k_3$ 

de la transcription

l/ml/ul : litre/millilitre/microlitre

LB/LBA : Luria Bertani/Luria Bertani supplémenté en ampicilline

LBGlc/LBPGlc : Luria Bertani supplémenté en Glc/Glc et phosphate minéral à 0.1 M

: taux de croissance (h-1)

: taux de croissance maximal (h<sup>-1</sup>)  $\mu_{max}$ : molaire/millimolaire/micromolaire  $M/mM/\mu M$ 

: milliampère mA mV : milliVolt

MFD : Mutation Frequency Decline

: minute min

mol/mmol/µmol : mole/millimole/micromole

: nucléotide N

NAD : nicotinamide adénine dinucléotide

: nicotinamide adénine dinucléotide phosphate NADP

: nanomètre nm

: ribonucléosides 5' triphosphate NTP

: nucléotides nt

**ONPG** : O-Nitrophényl \( \beta - D-galactopyranoside \)

ORF : "open reading frame" : poly(A) polymérase PAP : paire de bases pb : phagemide bluescript

pBS

: poids/volume p/v : polyéthylène glycol **PEG** : phospho-énol pyruvate PEP **PGK** : phosphoglycérate kinase : poids moléculaire PM

: polynucléotide phosphorylase **PNPase** 

PTS : système phosphotransférase PEP dépendant

: vitesse spécifique de production ou de consommation de composés q

: quantité suffisante pour qsp

: vitesse spécifique de production ou de consommation de composés

: "ribosome binding site" **RBS** 

**RNase** : ribonucléase

: Represseurs, ORFs, kinases ROK

: complexe fermé d'initiation de la transcription RPc : complexe d'élongation de la transcription RPélong

: complexe intermédiaire d'initiation la transcription **RPint** 

: complexe d'initiation de la transcription **RPinit** 

RPo : complexe ouvert d'initiation de la transcription

: Régulateur de réponse RR

S : substrat

SD : Shine-Dalgarno

SDS : sodium dodécylsulfate

SELEX : systematic evolution of ligands by exponential enrichment

TBE : Tris-borate-EDTA TCA : Tricarboxylic acid cycle

TE : Tris-EDTA

TEMED : N-N-N'N' tétraéthylènediamine TIR : "Translation Initiation Region"

Tn : transposon tr/min : tours/minute

Tris : Tris (hydroxyméthyl) aminométhane

U : unité enzymatique
UI : Unité Internationale
UP : "upstream region"
UTP : uridine 5' triphosphate
UTR : "untranslated region"

UV : ultra-violet

V : volt

V/cm : Volt/centimètre v/v : volume/volume

vvm : volume d'air entrant par volume du réacteur et par minute

W : watt X : biomasse

X-gal : 5-bromo-4-chloro-3-indoxyl-β-D-galactopyranoside

Y<sub>A/B</sub> : Rendement de conversion de B en A

Sommaire

# **Avant-propos**

SCH THIR MADEY 1

ELECTION CONTROLS

ENGLISH CONTROLS

EAGET VICTOR LOCAL ELEX

# Introduction

| Partie A / La Nutrition bactérienne   | 1  |
|---|----|
| I / Les différents systèmes d'entrées assurant l'entrée des nutriments dans la cellule  | 2  |
| I.1 / Cas des nutriments carbonés   |    |
| I.2 / Cas du phosphate et des nutriments contenant un ou des phosphates                 | .3 |
| II / Les systèmes de transport PTS des sources carbonées chez E. coli                   | 4  |
| II.1 / Organisation générale  |    |
| II.2 / Les protéines générales du système PTS   | 5  |
| II.2.a / La protéine EI.  |    |
| II.2.b / La protéine HPr  |    |
| II.3 / Les protéines EII.   |    |
| II.4 / Structure des composants du système PTS du glucose                               |    |
| II.5 / Mécanisme de transport et de phosphorylation des substrats                       |    |
| II.6 / Organisation des gènes du système PTS  | 8  |
| III / Le métabolisme général des sources de carbones                                    | 9  |
| Partie B / La transcription chez <i>E. coli</i>   | 11 |
| I / Le mécanisme d'initiation de la transcription                                       | 11 |
| I.1 / L'apoenzyme ARN polymérase  | 11 |
| I.2 / Les différentes formes d'holoenzymes d'E. coli                                    | 12 |
| I.3 / Les promoteurs reconnus par l'holoenzyme Εσ <sup>70</sup>                         | 13 |
| I.4 / Les facteurs σ d'E. coli  |    |
| I.4.a / La famille de facteurs σ <sup>70</sup>  | 14 |
| I.4.a. $\alpha$ / Les différents groupes de la famille de facteurs $\sigma^{70}$        | 14 |
| I.4.a.β / Structure et rôles des différentes régions des facteurs $\sigma^{70}$         |    |
| I.4.a. $\chi$ / Les membres de la famille de facteurs $\sigma^{70}$ chez <i>E. coli</i> | 17 |
| I.4.b / La famille de facteurs σ <sup>54</sup>  | 19 |
| I.5 / Les différentes étapes de l'initiation de la transcription                        |    |
| II / La régulation de l'initiation de la transcription                                  | 22 |
| II.1 / influence de la courbure de l'ADN au niveau du promoteur                         |    |
| II.2 / Influence des éléments en trans du promoteur                                     |    |
| II.2.a / Choix du facteur sigma   | 23 |
| II.2.b / Influence des nucléotides au niveau du complexe ouvert                         | 23 |
| II.2.c / Influence des facteurs de transcription.                                       |    |
| II.2.d / Intégration du signal  | 26 |
| III / Mécanisme d'élongation et de terminaison de la transcription                      | 27 |

| Partie C / L'initiation de la traduction chez E. coli                                 | .30  |
|---|------|
| I / Mécanisme d'initiation de la traduction   | 30   |
| I.1 / Initiation de la traduction des ARNm portant une séquence SD                    |      |
| I.2 / Initiation de la traduction des ARNm sans région en amont du codon initiateur   |      |
| I.3 / Modèle cinétique d'initiation de la traduction.                                 |      |
| II / Efficacité de l'initiation de la traduction                                      | 22   |
| II.1 / La séquence SD : structure primaire  |      |
| II.2 / La protéine S1   |      |
| II.3 / Le codon initiateur.   |      |
| II.4 / La séquence Downstream Box située en aval du codon d'initiation                |      |
| II.5 / Les codons situés au début des ORF et leur influence sur la structure          |      |
| secondaire de l'ARNm  | . 36 |
| II.6 / L'utilisation des codons des ORF et les codons terminateurs                    |      |
| III / Régulation de l'initiation traduction par les éléments en trans de l'ARNm       | 37   |
| III.1 / Structure secondaire et tertiaire de la région 5' non traduite                |      |
| III.2 / La température  |      |
| III.3 / Influence de la traduction du cistron amont sur la traduction du cistron aval |      |
| d'ARNm polycistronique : le couplage traductionnel                                    |      |
| III.4 / Répression de l'initiation de la traduction par fixation de protéines         |      |
| III.4.a / Autorégulation des protéines de liaison à l'ARNm                            |      |
| III.4.b / Régulation par la protéine CsrA   |      |
| III.4.c / Régulation par les métabolites  |      |
| Partie D / La dégradation des ARNm  | . 45 |
| I / Les RNases impliquées dans la dégradation des ARNm et leur régulation             | . 45 |
| I.1 / La RNase E, la RNase G et le dégradosome  | 46   |
| I.1.a / La RNase E  |      |
| I.1.a.α / Rôle de la RNase E  |      |
| I.1.a.β / Structure de la RNase E   |      |
| I.1.a.χ / Mécanisme de clivage par la RNase E   |      |
| I.1.a.δ / Spécificité de coupure de la RNase E  |      |
| I.1.a.ε / Régulation de la synthèse de la RNase E                                     |      |
| I.1.a.φ / Les homologues de la RNase E  |      |
| I.1.b / La RNase G  |      |
| I.1.b.α / Rôle de la RNase G  |      |
| I.1.b.β / Structure de la RNase G   |      |
| I.1.b.γ / Spécificité de coupure de la RNase G  |      |
| I.1.b.δ / Régulation de la RNase G  |      |
| I.1.c / le dégradosome  |      |
| I.2 / La RNase III  |      |
| I.2.b / Spécificité de clivage de la RNase III d' <i>E. coli</i>                      |      |
| I.2.c / Structure de la RNase III de classe I.  |      |
| I.2.d / Mécanisme de clivage de la RNase III d' <i>E. coli</i>                        |      |
| I.2.e / Régulation de l'expression et de l'activité de la RNase III.                  |      |

| I.3 / Les exonucléases 3'-5': la RNase II, la RNase R et la PNPase   |                |
|--|----------------|
| I.3.a / La RNase II et la RNase R  | 62             |
| I.3.a.α / Rôle de la RNase II  | 63             |
| I.3.a.β / Régulation de la RNase II  | 63             |
| I.3.a.γ / Rôle de la RNase R   |                |
| I.3.b / La PNPase.   |                |
| I.3.b.α / Rôle de la PNPase  | 64             |
| I.3.b.β / Structure de la PNPase   |                |
| I.3.b.γ / Régulation de la PNPase  |                |
| I.4 / La poly(A) polymérase et la polyadénylation des ARN  |                |
| I.4.a / Implication de la polyadénylation dans la dégradation des ARN  |                |
| I.4.b / Mécanisme de dégradation des ARN polyadénylés  |                |
| I.4.c / Activité poly(A) polymérase  |                |
| I.4.d / Structure de la poly(A) polymérase I   |                |
| I.4.e / Régulation de l'expression de la poly(A) polymérase I  |                |
| I.5 / La protéine Hfq  |                |
| I.5.a / Rôle de la protéine Hfq  |                |
| I.5.b / Homologie et structure de la protéine Hfq  |                |
| I.5.c / Spécificité de fixation de la protéine Hfq   |                |
| I.5.d / Mécanisme de régulation exercée par la protéine Hfq  |                |
| I.6 / Les autres enzymes impliquées dans la dégradation des ARNm   |                |
| TO SERVICE STATE OF THE SERVIC |                |
| II / Les éléments intervenants dans la stabilité des ARNm  | 74             |
| II.1 / Les éléments situés en cis  | 74             |
| II.1.a / Les éléments localisés à l'extrémité 5' des ARNm  | 74             |
| II.1.b / Les éléments localisés à l'extrémité 3' des ARNm  | 75             |
| II.2 / Les éléments agissant en trans.   |                |
| II.2.a / Les ribosomes.  |                |
| II.2.b / La protéine H-NS.   |                |
| II.2.c / Les ARNnc.  | 77             |
| Partie E / Régulation de l'expression des gènes du métabolisme central   | 78             |
| , J J J J J J J J J J J J J  |                |
| I / Variation de la production des enzymes du métabolisme central en fonction  |                |
| des conditions de croissance   | 78             |
|  |                |
| II / Régulation des voies du métabolisme carboné exercée par le système PTS du glu   | cose 80        |
| II.1 / Régulation du métabolisme carboné par l'enzyme générale HPr   |                |
| II.2 / Régulation du métabolisme carboné par la protéine EIIA Glc  | 81             |
| II.2.a / Régulation par la protéine EIIA Gle déphosphorylée :  |                |
| l'exclusion d'inducteur  | 82             |
| II.2.a.α / Le rôle de l'exclusion d'inducteur  |                |
| II.2.a.β / Spécificité de l'interaction de la protéine   |                |
| EIIA Glc déphosphorylée  | 83             |
| II.2.a.χ / Exclusion d'inducteur par des sources de carbone dites P  |                |
|  | "S             |
|  |                |
| autres que le glucose  |                |
| autres que le glucose  II.2.b / Régulation par la protéine EIIA Gle phosphorylée :   | 83             |
| autres que le glucose  II.2.b / Régulation par la protéine EIIA Gle phosphorylée : la régulation transcriptionnelle par le complexe CRP/AMPc   | 83             |
| autres que le glucose  II.2.b / Régulation par la protéine EIIA <sup>Glc</sup> phosphorylée : la régulation transcriptionnelle par le complexe CRP/AMPc  II.2.b.α / Gènes régulés par le complexe CRP/AMPc   | 83<br>84       |
| autres que le glucose  II.2.b / Régulation par la protéine EIIA Gle phosphorylée : la régulation transcriptionnelle par le complexe CRP/AMPc   | 83<br>84<br>84 |

| par le complexe CRP-AMPc   |
|--|
| via la protéine EIIA <sup>Glc</sup>  |
| du régulateur transcriptionnel Mlc   |
| II.4.a / Régulation de l'expression de l'opéron pts  |
| II.4.b / Régulation de l'expression du gène <i>ptsG</i> codant la protéine EIIBC <sup>Gle</sup> 92 |
| II.4.b. $\alpha$ / Régulation de l'initiation de la transcription du gène <i>ptsG</i> 92           |
| II.4.b.β / Régulation post-transcriptionnelle du gène <i>ptsG</i> 93                               |
| III / Régulations transcriptionnelles des voies du métabolisme carboné                             |
| exercées indépendamment du système PTS du glucose  |
| III.1 / Le régulateur FruR, appelé aussi Cra   |
| III.2 / Les systèmes de transduction du signal à deux composants                                   |
| Présentation du sujet  |
| Partie A / Les enzymes FBA, GAPDH, E4PDH et PGK et leur expression                                 |
| I / La FBA   |
| I.1 / Les deux classes de la FBA   |
| I.2 / Les deux classes de la FBA présentes chez E. coli  |
|  |
| II / La GAPDH et l'E4PDH   |
| II.1 / Rôles de la GAPDH et de l'E4PDH   |
| II.2 / Structure et mécanisme des GAPDH phosphorylantes NAD-dépendantes                            |
| II.3 / Structure et mécanisme de l'E4PDH   |
| III / La PGK 105   |
| III .1 / Rôles de la PGK   |
| III .2 / Structure et mécanisme de la PGK  |
| ×  |
| IV / Transcription de l'opéron epd-pgk-fbaA et du gène gapA chez E. coli                           |
| IV.1 / Transcription du triplet d'ORF epd-pgk-fbaA   |
| IV.2 / Transcription du gène gapA  |
| Partie B / Objectifs de mes travaux107   |
| Résultats - Discussion   |
|  |
| Partie A: Etude de l'organisation des ORF gap, epd, pgk et fbaA                                    |
| dans les chromosomes bactériens et étude de la régulation du triplet                               |
| d'ORF epd-pgk-fbaA chez E. coli  |

| E4PDH/GAPDH-PGK-FBA au sein des génomes bactériens séquencés   |
|--|
|  |
| I.1 / Etude de l'organisation des ORF autour de l'ORF pgk  |
| I.1.a / Les génomes étudiés et la démarche utilisée  |
| I.1.b / Résultats  |
| I.1.b. $\alpha$ / Les ORF situées en amont de l'ORF $pgk$  |
| I.1.b.β / Les ORF situées en aval de l'ORF pgk   |
| I.2 / Distinction entre ORF <i>epd</i> et <i>gap</i> en amont de l'ORF <i>pgk</i>  |
| I.2.a / La démarche utilisée   |
| I.2.b / Les résultats  |
| I.3 / Classe d'aldolase produite par les groupes d'ORF epd/gapA-pgk-ald  |
| II / Etat des travaux concernant l'expression du tandem d'ORF epd-pgk lorsque  |
| j'ai entrepris ce travail  |
|  |
| III / Poursuite de l'étude du triplet d'ORF epd-pgk-fbaA   |
|  |
| Dublication NO1 and a service of the latest termination of the service of the ser |
| Publication $N^{\circ}1$ : Characterization of the molecular mechanisms involved in the  |
| differential production of Erythrose-4-phosphate dehydrogenase, 3-   |
| Phosphoglycerate kinase and class II Fructose-1,6-bisphosphate aldolase in   |
| Escherichia coli   |
| Estrict forma son  |
|  |
|  |
| Travaux complémentaires réalisés sur les régulations de l'expression des   |
| opérons epd-pgk-fbaA   |
| I/Establish ANN AND THE STATE OF THE STATE O |
| I / Etude de la stabilité relative des ARNm monocistroniques pgk et fbaA chez E. coli 119  |
|  |
| II / Poursuite de l'étude phylogénique des opérons end-nok-fha A en rapport  |
| II / Poursuite de l'étude phylogénique des opérons <i>epd-pgk-fbaA</i> en rapport avec les caractéristiques mises en évidences précédemment  |
| avec les caractéristiques mises en évidences précédemment  |
| avec les caractéristiques mises en évidences précédemment  |
| avec les caractéristiques mises en évidences précédemment  |
| avec les caractéristiques mises en évidences précédemment  |
| avec les caractéristiques mises en évidences précédemment.122II.1 / Etude de la conservation des régions promotrices epd P0 et pgk P1122II.2 / Conservation du terminateur potentiel rho indépendant pgk T1123II.3 / Conservation des sites d'entrée interne de la RNase E123II.4 / Sites d'initiation de la traduction124   |
| avec les caractéristiques mises en évidences précédemment  |
| avec les caractéristiques mises en évidences précédemment  |
| II.1 / Etude de la conservation des régions promotrices epd P0 et pgk P1   |
| II.1 / Etude de la conservation des régions promotrices epd P0 et pgk P1   |
| II.1 / Etude de la conservation des régions promotrices epd P0 et pgk P1   |
| III.1 / Etude de la conservation des régions promotrices epd P0 et pgk P1  |
| II.1 / Etude de la conservation des régions promotrices epd P0 et pgk P1   |

| Partie B : Etude de la régulation de l'expression du gène gapA et de l'opéron epd-<br>pgk-fbaA en fermentation continue dans des conditions de limitation en          |
|---|
| nutriments  |
| I / Etude des transcrits gapA dans des bactéries cultivées en fermentation continue en condition de limitation en glucose   |
| avec 2 g/l de glucose   |
| II / Etude des transcrits epd et gapA dans des bactéries cultivées en fermentation  |
| continue en condition de limitation en phosphate  |
| III / Conclusion et Perspectives  |
| Partie C: Etude comparative des souches TP 2503 et TP 2504  |
| supplémenté avec différentes concentrations en glucose  |
| II / Etude des mécanismes expliquant l'obtention d'une forte densité cellulaire en phase stationnaire de croissance pour la souche d'E. coli TP 2504                  |
| de la souche TP 2504  |
| ou muté   |
| III / Comparaison des capacités de production de protéines recombinantes par  |
| des souches exprimant un gène ptsG sauvage ou muté  |
| III.2 / Etude de l'influence de la concentration en glucose dans le milieu de culture sur la croissance des souches TP 2503 et TP 2504 recombinées et leur production |

| III.3 / Etude de l'effet de l'oxygénation du milieu sur la croissance des souches recombinantes et leur production de GAPDH | 55                                     |
|---|--|
| IV / Conclusion   | 57                                     |
| Conclusion finale   | 159                                    |
|   |  |
| Matériels et Méthodes   |  |
| Partie A / Matériels  | 61                                     |
| I / Souches bactériennes utilisées  | 161<br>161                             |
| Partie B / Méthodes   | 62                                     |
| I. / Cultures des Bactéries       1.1 / Cultures d'E. coli en mode batch  | 162<br>162<br>163                      |
| II / Extraction et analyse de l'ADN.  II.1 / Extraction d'ADN cellulaire  | 164<br>164<br>164<br>165<br>165<br>166 |
| III / Extraction et analyse de l'ARN  | 167<br>167<br>167                      |

| III.3 / Analyse des ARN par Northern-blot  |
|--|
| III.3.a / Séparation des ARN suivant leur taille   |
| III.3.b / Transfert électrophorétique sur une membrane Biodyne   |
| III.3.c / Hybridation  |
| III.3.d / Déshybridation et réhybridation  |
| III.4 / Analyse de la structure secondaire d'ARN par sondes enzymatiques et chimiques169                     |
| III.4.a / Production d'ARN froids par transcription in vitro avec l'ARN                                      |
| polymérase du phage T7169  |
| III.4.b / Analyse de la structure secondaire des ARN par sondes  |
| enzymatiques et chimiques  |
| III.5 / Etude d'interaction ARN-protéines par gel retard   |
| III.5.a / Production d'ARN radiomarqués par transcription in vitro avec                                      |
| l'ARN polymérase du phage T7   |
| III.5.b / Etude des complexes [ARN-Hfq]  |
| III.5.b.α / Formation des complexes ARN-protéines  |
| III.5.b.β / Fractionnement des complexes [ARN-Hfq] par la technique  |
| de retard sur gel173   |
| II.5.b.γ / Expériences d'empreintes  |
| IV/ A l  |
| IV / Analyse des protéines cellulaires   |
| IV.1 / Sonication des bactéries  |
| IV.2 / Analyse des protéines par SDS-PAGE  |
| IV.3 / Mesure des activités enzymatiques   |
| IV.3.a / Mesure d'activité GAPDH   |
| IV.3.b / Mesure d'activité PGK   |
| IV.3.c / Mesures d'activité FBA  |
| V / Analyse des surnageants de cultures  |
| V.1 / Dosage enzymatique du glucose  |
| V.2 / Dosage enzymatique de l'acide acétique   |
| V.3 / Dosage par CLHP du glucose et de l'acide acétique  |
| VI / Exploitations des données aunévimentales  |
| VI / Exploitations des données expérimentales  |
| VI.1 / Détermination du poids sec  |
| VI.2 / Détermination des caractéristiques de croissances des souches d' <i>E. coli</i> en cultures batch     |
|  |
| VI.3 / Détermination des caractéristiques de croissances des souches d' <i>E. coli</i> en cultures continues |
| en cultures continues  |
|  |
| Riblicanaphia  |
| Bibliographie  |
| 182  |

Avant-propos

J'ai réalisé un travail de thèse dans l'équipe de C. Branlant, et ce travail a été encadré par B. Charpentier. Il porte sur «l'étude de l'organisation transcriptionnelle et de la régulation post-transcriptionnelle du groupe d'ORF codant l'érythrose-4-phosphate deshydrogénase, la phosphoglycérate kinase et la fructose-1,6-bisphosphate aldolase de classe II chez *Escherichia coli*». Les objectifs de ce travail avaient été définis de façon à ce que je puisse utiliser mes compétences initiales en matière de fermentations bactériennes, acquises lors de mes études à l'ENSAIA, ceci tout en acquérant les compétences en biologie moléculaire et analyses structurales et fonctionnelles de l'équipe de C. Branlant.

Au laboratoire, l'équipe de G. Branlant s'est intéressée pendant de nombreuses années à la glycéraldéhyde-3-phosphate déshydrogénase (GAPDH), en tant que modèle d'étude pour le remodelage d'enzyme et en vue de mieux comprendre les bases moléculaires et la spécificité et de la réactivité de cette enzyme. E. coli présente la particularité de présenter 2 gènes gap fonctionnels ayant une forte homologie avec les gènes gap (Branlant et al., 1983 ; Alefounder et Perham, 1989). L'équipe de C. Branlant s'est intéressée aux mécanismes de régulation de l'expression de ces deux gènes gap. La GAPDH active est codée par le gène gapA, puisqu'une mutation non-sens dans ce gène gapA conduit à un phénotype GAPDH (Hillman et Fraenkel, 1975). Lorsque ce gène gapA est cloné dans un plasmide multicopie, il conduit à une très forte production de la GAPDH, constituant environ 80 % des protéines solubles (Branlant et al., 1983). L'étude des mécanismes de régulation de l'expression de ce gène gapA a été réalisée par B. Thouvenot et B. Charpentier. B. Charpentier a mis en évidence la présence de 4 promoteurs, régulés différemment, assurant une forte expression du gène gapA quelques soient les conditions de l'environnement (Charpentier et Branlant, 1994). B. Thouvenot et B. Charpentier ont réalisé une étude plus approfondie du promoteur gapA P1, le promoteur le plus efficace en phase exponentielle de croissance. Ils ont mis en évidence la présence d'une boîte -10 étendue, avec une boîte -10 sous-optimale, et de façon inattendue, la présence d'une boîte -35 très proche du consensus (Thouvenot et al., 2004). Le second gène gapB code une protéine présentant de très fortes homologies avec les GAPDH de type bactérien. Cette protéine présente une très faible activité GAPDH et une activité érythrose-4-phosphate deshydrogénase (E4PDH) (Boschi-Muller et al., 1997; Zhao et al., 1995). Le gène gapB, appelé aussi epd, est localisé en amont des gènes pgk et fbaA, codant respectivement la phosphoglycérate kinase (PGK) et la fructose-1,6-bisphosphate aldolase de classe II (FBA) intervenant dans la glycolyse et la gluconéogenèse chez E. coli. Le gène gap est souvent retrouvé en amont du gène pgk chez les bactéries. Nous avons alors entrepris une étude systématique de l'organisation des gènes gap/epd-pgk-fbaA chez les bactéries. Cette étude est basée sur la recherche, au sein des gènes homologues des gènes gap, de la signature protéique définissant l'activité E4PDH décrite au laboratoire par Boschi-Muller et al. (1997). Nous avons ainsi pu montrer que pour de nombreux phyla bactériens, l'organisation gap-pgk-tpi est observée alors que chez les familles des Vibrionales et Enterobactériales de la subdivision γ des protéobactéries, c'est l'organisation *epd-pgk-fbaA* qui est observée.

B. Charpentier, V. Bardey et N. Robas avaient étudié la transcription du tandem d'ORF *epd-pgk* chez *E. coli*. Ils avaient ainsi mis en évidence la présence d'un promoteur unique, *epd* P0 en amont du gène *epd* et observé que le taux de transcrits initiés à ce promoteur augmente en présence de glucose lorsque le transporteur EIIBC<sup>Gle</sup> est fonctionnel. Ce promoteur présente la particularité d'avoir un site de fixation au complexe CRP-AMPc permettant une activation de la transcription du promoteur *epd* P0 par ce complexe. Bien que la concentration du complexe CRP-AMPc diminue en présence de glucose, la présence de ce complexe est indispensable à l'activation du promoteur *epd* P0 en présence de glucose, ceci parce qu'il est requis pour l'expression d'EIIBC<sup>Gle</sup> (Charpentier *et al.*, 1998). Ils avaient montré que ce promoteur permet de transcrire les ORF *epd* et *pgk*, mais qu'un autre promoteur de type étendu, *pgk* P1, est aussi impliqué dans l'expression de base de *pgk*, ceci quelles que soient les conditions de culture. Nos travaux montrent que le gène *fbaA* est transcrit à partir des promoteurs *epd* P0 et *pgk* P1. V. Bardey et N. Robas avaient mis en évidence différents mécanismes conduisant à une faible expression de l'ORF *epd* : une faible traduction de l'ORF *epd* et la présence de sites de clivages par la RNase E dans la région *epd-pgk*, et en particulier dans l'ORF *epd*.

Mon travail a permis de montrer que l'ORF fbaA est aussi transcrite à partir des promoteurs epd P0 et pgk P1 et que les transcrits epd-pgk-fbaA et pgk-fbaA sont clivés par la RNase E, au niveau des régions epd-pgk et pgk-fbaA. V. Bardey et A. Mougin avaient étudié la structure secondaire de la région epd-pgk des transcrits clivés par la RNase E. J'ai, pour ma part, étudié celle de la région pgk-fbaA, ceci par l'emploi de sondes enzymatiques et chimiques de la structure des ARN en solution. Nous observons chez les entérobactéries une conservation des structures secondaires proposées pour les sites RNase E du transcrit epd-pgk-fbaA. Ces données apportent des informations sur la structure secondaire des sites d'entrée de la RNase E dans les ARNm. Nous avons ensuite étudié la stabilité des différents transcrits monocistroniques issus de l'opéron epd-pgk-fbaA.

Comme mentionné ci-dessus, B. Charpentier, V. Bardey et N. Robas avaient montré une augmentation du taux de transcrits initiés aux promoteurs gapA P1 et epd P0 en présence de glucose dans le milieu de culture et lorsque le transporteur spécifique du glucose EIIBC<sup>Glc</sup> est actif. Comme l'inactivation du transporteur EIIBC<sup>Glc</sup> conduit à une diminution du taux de croissance en milieu minimum supplémenté en glucose, nous avons cherché à savoir si c'était l'absence de la protéine EIIBC<sup>Glc</sup> ou la diminution du taux de croissance qui limitait le taux de transcrits gapA P1 et epd-pgk-fbaA. Pour cela, par utilisation de techniques de fermentations continues, nous avons pu mettre en évidence une augmentation du taux de transcrits initiés au promoteur gapA P1 parallèle à

l'augmentation du taux de croissance, et ceci même dans la souche n'exprimant pas le transporteur EIIBC Gle

Au cours de ces fermentations continues, nous avions observé une absence de production d'acide acétique en phase exponentielle de croisssance, dans le cas de la souche exprimant un transporteur EIIBC<sup>Glc</sup> inactif. Ceci pouvait être intéressant sur le plan de la production de protéines recombinantes chez *E. coli*, puisque la production d'acide acétique par *E. coli* est une des limitations de cette production. Ceci m'a conduit à tester la capacité de production de protéines recombinantes par la souche EIIBC<sup>Glc</sup>. J'ai ainsi pu mettre au point des conditions permettant d'obtenir une forte production de protéines recombinantes.

Nous commencerons donc par faire une revue bibliographique sur la nutrition bactérienne et les différentes voies métaboliques pour le métabolisme du glucose. Les différents mécanismes impliqués dans la régulation de l'expression des gènes chez *E. coli* seront aussi présentés. Nous décrirons en particulier les mécanismes d'initiation de la transcription des gènes, puis les mécanismes impliqués dans la traduction des ARNm et enfin, les différents systèmes de maturation et de dégradation des ARNm. Les connaissances dans ces domaines sont en effet nécessaires pour la compréhension des données sur l'opéron *epd-pgk-fbaA*. Enfin, nous décrirons les mécanismes de régulation de l'expression des gènes impliqués dans le métabolisme du glucose chez *E. coli*. Ceci est particulièrement important pour comprendre les mécanismes de régulation observés lors de l'inactivation du transporteur EIIBC<sup>Glc</sup>. Les résultats de l'opéron *epd-pgk-fbaA* seront ensuite présentés, en incorporant l'article sous presse dans Molecular Microbiology, avec une présentation des objectifs et une discussion incorporant certaines données non insérées dans l'article. Les deux autres parties des résultats portent sur les deux volets encore non publiés des travaux. Un bilan et une conclusion finale feront suite à cette partie Résultats.

# Introduction

Dans les années 1940, Jacques Monod a observé qu'en présence d'un mélange de certains sucres, la croissance bactérienne se fait en deux phases, séparées par une phase de latence pendant laquelle la croissance est fortement ralentie. Il a baptisé ce phénomène la diauxie (Monod, 1942). A partir de telles observations, Monod a proposé que la cellule bactérienne n'incorpore pas simplement les différents nutriments carbonés en fonction de leur abondance, mais en fait, contrôle de façon active leur entrée et leur métabolisme dans la cellule, en régulant la quantité des enzymes nécessaires. La phase de latence correspond à la synthèse des enzymes spécifiques nécessaires pour métaboliser des nutriments carbonés alternatifs. De nombreuses études se sont portées sur la compréhension des mécanismes de régulation mis en jeu lors de la diauxie. Récemment, Edwards *et al.* ont suggéré que ce contrôle actif a pour but d'ajuster l'ensemble des flux métaboliques de la bactérie afin d'obtenir un taux de croissance maximal (Edwards *et al.*, 2001).

Dans un premier temps, je vais donc décrire les différents systèmes d'entrée des nutriments carbonés dans la cellule, et les différentes voies métaboliques correspondantes. Je m'attacherai ensuite à décrire les différents mécanismes de régulation de l'expression des gènes dont les produits sont nécessaires au métabolisme des sucres, après avoir détaillé les différentes régulations exercées au niveau transcriptionnel et post-transcriptionnel.

# Partie A / La nutrition bactérienne

Les bactéries à gram négatif, comme Escherichia coli (E. coli), sont constituées de deux membranes. La région située entre ces deux membranes constitue le périplasme, la membrane interne entoure le cytoplasme. Le transport des nutriments à travers la membrane externe est réalisé principalement par diffusion facilitée, grâce à des porines, et quelques transporteurs spécifiques. Par contre, différents transporteurs spécifiques sont localisés au niveau de la membrane interne et assurent un contrôle plus fin du transport des nutriments vers le cytoplasme. Nous allons décrire ces systèmes de transport localisés au niveau de la membrane interne de façon succincte. Ensuite seront détaillés les systèmes de perméation des composés carbonés reposant sur un ensemble de phosphotransférases phospho-énol pyruvate (PEP)-dépendantes. Enfin, seront décrites les différentes voies métaboliques empruntées par les sources de carbone.

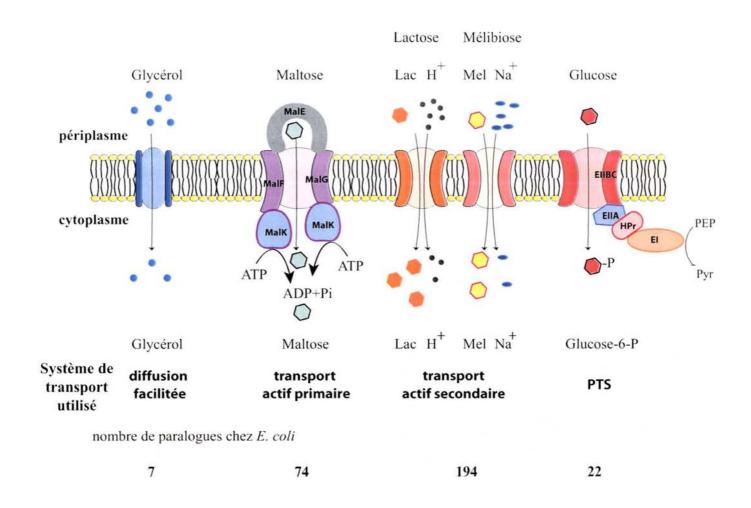


Figure 1 : Représentation des différents modes de transport des solutés chez E. coli (d'après Paulsen et al., 2000 ; Siebold et al., 2001).

# I / Les différents systèmes assurant l'entrée des nutriments chez *E. coli*

#### I.1 / Cas des nutriments carbonés

Afin de subvenir à leurs besoins carbonés dans différents milieux, les bactéries utilisent divers mécanismes de transport des solutés, qui diffèrent suivant le mode de transport et l'énergie utilisée pour effectuer ce transport (Paulsen *et al.*, 2000) (Fig. 1).

Le système le plus simple, mais aussi relativement rare chez les procaryotes, est la diffusion facilitée. Dans ce mode de transport, c'est le gradient de concentration créé entre l'intérieur et l'extérieur de la cellule, qui entraîne le flux de substrat entrant. Ce système ne nécessite aucune dépense d'énergie. La vitesse de transport est cependant plus élevée que par simple diffusion passive. Chez *E. coli*, le système de diffusion facilitée est en particulier utilisé par le glycérol.

Le transport actif primaire est réalisé par les transporteurs ABC (ATP-Binding Cassette) qui forment une des plus larges familles de protéines : E. coli contient 74 transporteurs ABC, où les gènes codant ces protéines représentent environ 5 % du génome (pour revue, Higgins, 2001). Contrairement à la diffusion facilitée, ce transport établit un gradient de concentration au lieu de l'exploiter, en consommant de l'énergie (Paulsen et al., 2000). Un des transporteurs le plus étudié chez E. coli est celui du maltose (Fig. 1). La protéine MalE périplasmique se lie au substrat avec une forte affinité et l'apporte au niveau des domaines transmembranaires MalF et MalG. L'hydrolyse de l'ATP au niveau des domaines cytoplasmiques MalK apporte l'énergie nécessaire à la translocation du substrat par les domaines transmembranaires MalF et MalG.

Le transport actif secondaire représente le mode de transport le plus fréquemment utilisé pour de nombreux solutés (Paulsen et al., 2000). En particulier, E. coli possède 194 transporteurs secondaires. Ce système de transport utilise un gradient ionique préalablement établi. Cette famille de transporteurs est divisée en trois groupes, suivant le gradient ionique utilisé : soit un gradient de protons, soit un gradient d'ions Na<sup>+</sup>, soit un gradient d'autres ions et solutés pouvant être échangés avec les protons. Le gradient de protons, appelé aussi force protomotrice, est généré lors de la respiration cellulaire aérobie, mais il peut aussi être créé par l'hydrolyse d'ATP par une ATPase membranaire. Les autres gradients utilisés sont générés aux dépens de cette force protomotrice. Ainsi, le gradient de Na<sup>+</sup> est réalisé par l'antiport H<sup>+</sup>/Na<sup>+</sup> où le transport des différents ions est réalisé en sens

opposé. Chez *E. coli*, la perméase du lactose, codée par le gène *lacY*, réalise la translocation stoechiométrique d'α ou de β D-galactopyranosides, tels que le lactose et le mélibiose, et de protons H<sup>+</sup> (Sondej *et al.*, 2003). Le mélibiose peut aussi rentrer dans la cellule en utilisant le gradient de Na<sup>+</sup> (Fig. 1). Dans ces deux cas, le transporteur est un symport, transportant le soluté et l'ion dans une même direction, du périplasme vers le cytoplasme.

Le dernier mécanisme de transport repose sur la phosphorylation du substrat. Il est réalisé par un ensemble de **phosphotransférases phospho-énol pyruvate (PEP)-dépendantes**, ces systèmes sont encore appelés **PTS**. Ainsi, ce mécanisme a l'avantage d'économiser de l'énergie par rapport aux deux systèmes de transport actifs décrits ci-dessus. En effet, lorsqu'un substrat est importé dans la cellule par un transporteur actif, une molécule d'ATP est consommée lors de l'entrée du substrat dans la cellule et lors de sa phosphorylation. Cette étape de phosphorylation est généralement la première réaction métabolique intracellulaire réalisée. Au contraire, le système PTS consomme une seule molécule de PEP pour ces deux étapes, soit l'entrée et la phosphorylation du substrat. Cependant, ce mode de transport avantageux n'est utilisé que pour un nombre limité de substrats : seulement 22 systèmes PTS distincts ont été identifiés chez *E. coli* (pour revue, Plumbridge, 2002). La présence de système PTS n'a été observée que dans la moitié des bactéries dont les génomes ont été séquencés actuellement, et il n'a pas été identifié ni chez les Archaea, ni chez les eucaryotes (Paulsen *et al.*, 2000). Ces systèmes PTS réalisent aussi d'autres fonctions cellulaires. Nous décrirons les différentes régulations exercées par le système PTS chez *E. coli* dans la partie E.

La majorité des composés nutritionnels utilise deux types de transport : un de faible affinité et consommant peu d'énergie, utilisé lorsque ce composé est en forte concentration, et un de forte affinité et consommant le plus d'énergie, utilisé lorsque ce composé est en faible concentration. Ceci est le cas pour le glucose : à forte concentration, il entre dans la cellule via le système PTS spécifique du glucose. A faible concentration, le glucose est piégé par la glycoporine LamB et traverse la membrane interne par un système de transport actif primaire, le transporteur ABC Mgl du galactose (pour revue, Ferenci, 1996).

#### I.2 / Cas du phosphate et des nutriments contenant un ou des phosphates

Suivant sa concentration, le phosphate peut être transporté par un mécanisme reposant sur des transporteurs de faible affinité, le système Pit, comprenant les transporteurs PitA et PitB, et un mécanisme reposant sur un transporteur de forte affinité, le système Pst (pour revue, Wanner, 1996). Deux autres systèmes de transports peuvent prendre en charge le phosphate inorganique en tant qu'analogue

de leur substrat mais avec une plus faible affinité. Il s'agit des transporteurs du glycérol 3-phosphate (GlpT) et du glucose 6-phosphate (UhpT). Cependant, ces deux systèmes ne permettent pas la croissance cellulaire sur un milieu contenant du phosphate inorganique, en absence des systèmes Pit et Pst.

La synthèse du transporteur PitA semble constitutive (pour revue, Wanner, 1996), alors que celle de PitB est réprimée ou inactivée par le régulon PHO (Harris et al., 2001). Ce régulon PHO comprend l'ensemble des gènes et opérons co-régulés par le système à deux composants PhoB-PhoR, qui sera présenté plus en détail dans la partie IV. Le système de transport Pst appartient à la famille des transporteurs ABC. Il est constitué d'une protéine de liaison périplasmique PstS, des deux protéines de la membrane interne PstA et PstC, qui constituent un canal et d'une ATPase cytoplasmique PstB (pour revue, Wanner, 1996). Les composants de ce système de transport du phosphate à forte affinité Pst sont codés par l'opéron pstSCAB-phoU. Cet opéron appartient au régulon PHO. La transcription est inhibée par une forte concentration en phosphate seulement en présence du système Pst et de la protéine PhoU (pour revue, Wanner, 1996).

## II / Les systèmes de transport PTS des sources carbonées chez E. coli

#### II.1 / Organisation générale

Les systèmes PTS sont tous composés de deux protéines cytoplasmiques générales, l'enzyme I (EI) et la protéine à histidine (Histidine Phosphocarrier protein HPr), et d'un nombre variable de complexes enzymatiques assurant le transport spécifique des sucres. Ces enzymes sont appelées enzymes II<sup>sucres</sup> (EII<sup>sucres</sup>) (Fig. 2). La protéine EI s'autophosphoryle en utilisant le groupement phosphate du PEP. Elle transfère ensuite son groupement phosphate à la protéine HPr, qui phosphoryle les différentes protéines EII<sup>sucres</sup> des différents complexes de transport PTS. Ces protéines EII<sup>sucres</sup> asssurent ensuite le transport et la phosphorylation de leurs sources carbonées spécifiques. La phosphorylation des protéines EI, HPr et EII est réversible, contrairement à la phosphorylation des sucres transportés au travers de la membrane interne. Ceci est dû au fait de la stabilité décroissante des phosphorylations réalisées sur les protéines successives (pour revue, Postma *et al.*, 1996b).

Lors de la phosphorylation de EI, le PEP, molécule donneuse du phosphate, est transformée en pyruvate. Cette réaction est réalisée de manière irréversible par les pyruvate kinases, et de façon réversible par la protéine EI du système PTS : elle constitue un lien entre le transport des sources carbonées et le métabolisme central.

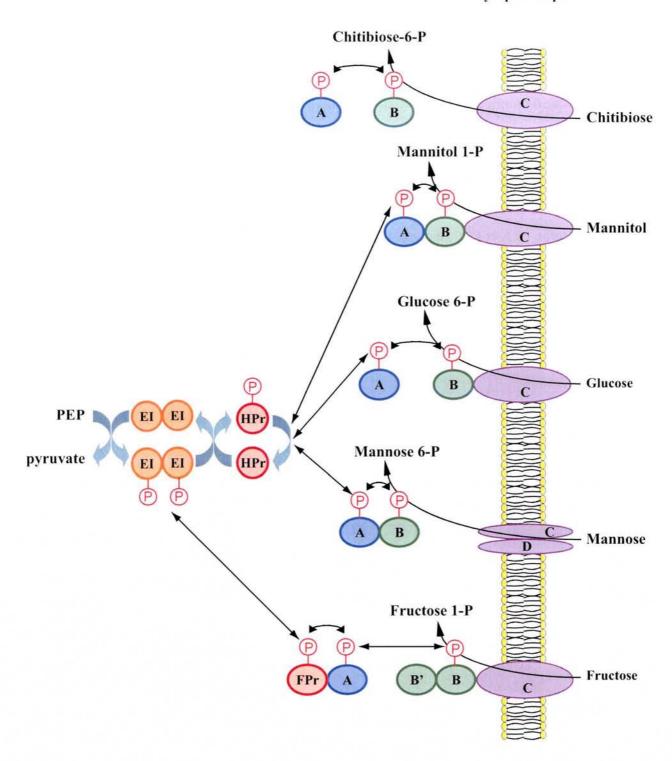
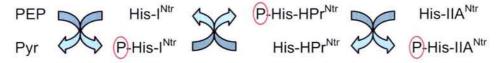


Figure 2: Organisation du système PTS du mannitol, du glucose, du mannose et du fructose chez *E. coli* (Keyhani *et al.*, 2000 ; pour revues, Postma *et al.*, 1993 ; Postma *et al.*, 1996 ; Ginsburg et Peterkofsky, 2002).

La protéine générale EI est en orange, la protéine générale HPr ainsi que le domaine homologue à HPr de FPr du système PTS du fructose, sont représentés en rouge. Les domaines EIIA des protéines EII spécifiques des différentes sources carbonées sont schématisés en bleu, tandis que les domaines EIIB sont représentés en vert, les domaines EIIC et EIID transmembranaires en violet. Le groupement phosphate qui est transféré depuis le PEP jusqu'à la source de carbone est représenté par un P cerclé en rouge.

Chez *E. coli* et les autres protéobactéries à Gram négatif, différents systèmes paralogues au système PTS ont été identifiés (pour revue, Hu et Saier, 2002). Par exemple, les protéines PTS de l'azote catalysent les réactions de transfert de phosphate suivant (d'après Hu et Saier, 2002):



La protéine EI<sup>Ntr</sup> d'*E. coli* rend le métabolisme de l'azote dépendant de celui du carbone. Les conséquences régulatrices de ces voies de transport semblent différentes suivant les bactéries. Ainsi, chez *Legionella pneumophila*, la protéine EI<sup>Ntr</sup> contrôle la virulence (pour revue, Hu et Saier, 2002).

#### II.2 / Les protéines générales du système PTS

#### II.2.a / La protéine EI

La protéine EI, de 63.4 kDa, est constituée d'un domaine amino-terminal (N-terminal) (EI-N) compact contenant l'histidine 189, et d'un domaine C-terminal (EI-C) flexible susceptible à la protéolyse. Elle est en équilibre entre les formes monomériques et dimériques. La forme dimérique uniquement s'autophosphoryle au niveau de l'histidine 189, en présence de PEP et de magnésium (pour revue, Ginsburg et Peterkofsky, 2002). Le domaine EI-N peut être phosphorylé réversiblement, *in vitro*, par la protéine HPr, mais pas par le PEP, le domaine (EI-C) catalyse le transfert du phosphate du PEP à l'histidine 189 (Fomenkov *et al.*, 1998). La phosphorylation de EI-N entraîne un changement conformationnel de la protéine dans lequel l'histidine 189 subit une rotation vers la surface de la protéine. La protéine EI phosphorylée interagit avec la protéine HPr et lui transfert son phosphate. Elle retourne à l'équilibre entre les formes déphosphorylées monomériques et dimériques, jusqu'à ce que le PEP et le Mg<sup>2+</sup> se lient à nouveau et favorisent une dimérisation complète de la protéine EI. Ainsi, les concentrations intracellulaires de PEP et de Mg<sup>2+</sup>, liées aux fluctuations métaboliques, sont des déterminants majeurs pour la régulation de l'autophosphorylation dépendante du Mg<sup>2+</sup> par le PEP (Dimitrova *et al.*, 2002).

Le domaine EI-C est homologue aux PEP synthases et pyruvate : phosphate dikinase, mais le domaine EI-N ainsi que toutes les autres protéines PTS n'ont pas d'homologues à d'autres enzymes de fonction connues (pour revue, Hu et Saier, 2002). Les bactéries à Gram positifs et les spirochètes ne possèdent qu'un seul homologue à la protéine EI tandis que les protéobactéries ont plusieurs paralogues. En particulier, *E. coli* a 5 paralogues à la protéine EI : EI, EI<sup>Dha</sup>, deux EI<sup>Fru</sup> (EI<sup>Fry</sup> et EI<sup>Frw</sup>), et EI<sup>Ntr</sup>, dont les quatre dernières possèdent des domaines supplémentaires. La protéine EI<sup>Ntr</sup> contient un domaine détecteur de type Nif-A; les protéines EI<sup>Frw</sup> et EI<sup>Fry</sup>, identiques à 50 %, contiennent les

domaines HPr et EIIA<sup>Fru</sup> ; la protéine EI<sup>Dha</sup> contient une forme tronquée de la protéine EI, HPr et un domaine EIIA ressemblant à EIIA<sup>Man</sup> (pour revue, Hu et Saier, 2002).

#### II.2.b / La protéine HPr

La protéine HPr, de 9 109 Da, est phosphorylée au niveau de l'histidine 15 par la protéine EI. Comme pour EI, les bactéries à Gram positif n'ont qu'un seul homologue à la protéine HPr. *B. subtilis* fait exception à cette règle : elle contient les protéines homologues HPr et Crh, dans laquelle un résidu glutamine remplace l'histidine du site actif. Au contraire, *E. coli* K12 contient 6 paralogues à HPr : H, H<sup>Fru</sup>, H<sup>Fry</sup>, H<sup>Fry</sup>, H<sup>Dha</sup>, et H<sup>Ntr</sup>, dont trois paralogues à H<sup>Fru</sup>. La protéine H<sup>Fru</sup>, appelée aussi FPr, comporte un domaine EIIA<sup>Fru</sup> en N-terminal, et un domaine homologue à HPr en C-terminal (pour revue, Hu et Saier, 2002). Le domaine C-terminal de FPr phosphoryle son domaine N-terminal EIIA<sup>Fru</sup>, mais semble également capable de phosphoryler le domaine EIIA d'autres système de transport PTS, mais avec une efficacité moindre que la protéine HPr. La protéine FPr peut se substituer à la protéine HPr chez *E. coli* dans les mutants ne produisant pas de protéine HPr (pour revues, Postma *et al.*, 1993; Postma *et al.*, 1996b).

#### II.3 / Les protéines EII

Leur nombre et leur structure varient entre les différentes espèces de bactéries. Chez *E. coli*, l'analyse du génome a identifié 22 protéines EII, dont le substrat spécifique n'est connu que pour 13 d'entres elles (pour revue, Plumbridge, 2002).

Les protéines EII sont des dimères, où chaque monomère est constitué en général de trois domaines structuraux EIIA, EIIB et EIIC, chacun comprenant environ respectivement 100, 100 et 350 acides aminés. Les domaines EIIA et EIIB sont cytoplasmiques et réalisent le transfert séquentiel du phosphate de la protéine HPr au sucre transporté, en utilisant généralement une histidine, ou quelques fois une cystéine pour certains domaines EIIB comme pour la protéine EIIBC<sup>Gle</sup>. Le domaine EIIC est transmembranaire et comprend les déterminants majeurs pour la reconnaissance des sucres et leur transport du périplasme au cytoplasme. Un quatrième domaine EIID transmembranaire est parfois nécessaire au transport des sources de carbone, comme c'est le cas pour le système PTS spécifique du mannose (Fig. 2).

L'organisation de ces domaines varie en fonction du système PTS (Fig. 2). Les différents domaines peuvent être codés par des gènes différents, comme pour le chitibiose, ou au contraire par un seul gène et ne former qu'une seule chaîne peptidique comme pour le mannitol (Keyhani et al., 2000; pour revues, Postma et al., 1993; Postma et al., 1996a). Deux des trois (ou quatre) domaines fonctionnels peuvent aussi constituer une protéine. Ainsi, le système PTS spécifique du glucose comprend un domaine

EIIA<sup>Gle</sup> cytoplasmique indépendant, alors que le domaine EIIB<sup>Gle</sup> cytoplasmique est fusionné au domaine EIIC<sup>Gle</sup> transmembranaire. Dans le système PTS spécifique du mannose, le domaine EIIA<sup>Man</sup> est fusionné au domaine EIIB<sup>Man</sup> cytoplasmique.

Bien que les complexes EII<sup>sucres</sup> soient spécifiques d'un composé carboné, ils peuvent transporter d'autres composés carbonés, avec en général des affinités plus faibles. Ainsi, le complexe EII<sup>Man</sup> spécifique du mannose peut assurer le transport d'autres hexoses, tels que le glucose et le fructose (pour revues, Postma *et al.*, 1993; Postma *et al.*, 1996b).

Six classes de protéines EII<sup>sucres</sup> ont pu être définies, suivant un alignement de séquences et leur spécificité de substrat : les familles glucose / saccharose, lactose, mannitol, mannose, galactitol et glucitol (pour revues, Postma *et al.*, 1996b; Robillard et Broos, 1999). Les protéines EII<sup>sucres</sup> appartenant au même groupe peuvent se complémenter contrairement aux protéines appartenant à des groupes différents où les similarités sont limitées à quelques motifs (pour revues, Postma *et al.*, 1996b; Robillard et Broos, 1999). Ainsi, les protéines EIICBA<sup>nag</sup>, EIICBA<sup>bgl</sup>, appartenant à la même classe que la protéine EIIA<sup>Glc</sup>, peuvent complémenter une souche délétée de la protéine EIIA<sup>Glc</sup>. Ces deux protéines permettent le transport et la phosphorylation de l'α-méthylglucoside par la protéine EIIBC<sup>Glc</sup>, en absence de la protéine EIIA<sup>Glc</sup> (pour revue, Robillard et Broos, 1999). Ces exemples illustrent la complexité du transport des sources de carbone dans les cellules bactériennes.

### II.4 / Structure des composants du système PTS du glucose

La structure des différents complexes impliqués dans le transfert de phosphate par le système PTS spécifique pour le glucose a été obtenue récemment (Fig. 3) (Cai et al., 2003; Wang et al., 2000c). La spécificité des surfaces d'interaction est caractérisée par une complémentarité géométrique et chimique. La protéine EIIA<sup>Gle</sup>, de forme concave, interagit avec la partie convexe des protéines HPr et EIIBC<sup>Gle</sup> via des interactions hydrophobes, au niveau du site actif, et électrostatique à la périphérie (Cai et al., 2003). Chaque enzyme utilise la même surface d'interaction avec ses différents partenaires, où quelques chaînes latérales adoptent une conformation différente, suivant le partenaire à contacter, afin d'optimiser les interactions. Ainsi, la cascade de transfert des phosphates est réalisée par association et dissociation de complexes protéiques bimoléculaires, plutôt que par formation d'un complexe ternaire. L'enzyme générale HPr s'associe avec l'enzyme générale EI, et est phosphorylée. Elle est obligée de se dissocier de la protéine EI pour pouvoir interagir avec la même interface avec la protéine EIIA<sup>Gle</sup>. De même, la protéine EIIA<sup>Gle</sup> phosphorylée se dissocie de HPr pour interagir avec la protéine EIIBC<sup>Gle</sup> avec la même interface d'interaction qu'avec HPr. L'interaction directe de la protéine EI avec la protéine EIIBC<sup>Gle</sup> ne semble pas possible à priori (Wang et al., 2000a; Wang et al., 2000c; Cai et al., 2003). Il est à noter que le domaine N-terminal de la protéine EIIA<sup>Gle</sup> se structure en une hélice amphipathique

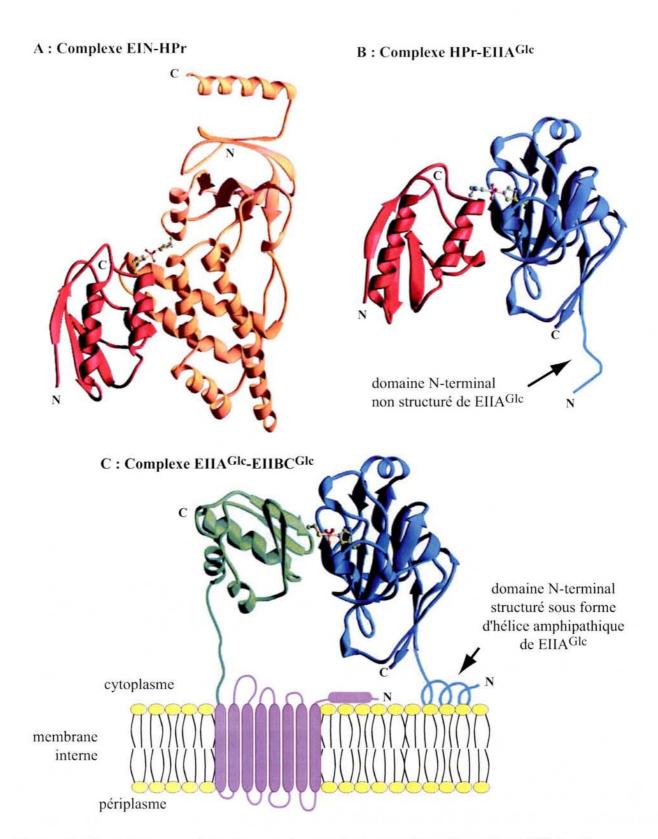


Figure 3 : Structure secondaire des complexes entre les protéines du système PTS du glucose chez *E. coli* (d'après Cai *et al.*, 2003).

- A : Structure secondaire du complexe HPr avec le domaine N-terminal de la protéine EI (EIN).
- B : Structure secondaire du complexe HPr avec la protéine EIIAGlc.
- C : Structure secondaire du complexe de la protéine EIIBGlc avec la protéine EIIAGlc. Le domaine N-terminal de la protéine EIIAGlc se structure en hélice au contact de la membrane.

Les hélices  $\alpha$  et feuillets  $\beta$  sont représentés en orange, rouge, bleu et vert pour les protéines EIN, HPr, EIIA<sup>Glc</sup> et EIIB<sup>Glc</sup>, respectivement. Les résidus réalisant le transfert de phosphate entre deux protéines ainsi que le phosphate sont schématisés dans l'état de transition.

au contact de la membrane (Fig. 3C) (Wang et al., 2000b). Cette hélice stabiliserait le complexe EIIA<sup>Gle</sup>-EIIBC<sup>Gle</sup>, ce qui améliorerait le transfert de phosphate.

#### II.5 / Mécanisme de transport et de phosphorylation des substrats

Les protéines EII<sup>sucres</sup> réalisent les quatre fonctions suivantes (pour revue, Robillard et Broos, 1999) :

- la phosphorylation vectorielle : la source carbonée est transportée du périplasme au cytoplasme, et est phosphorylée,
- la phosphorylation non vectorielle : la source carbonée présente dans le cytoplasme est phosphorylée,
- l'échange de phosphate entre deux sources carbonées du cytoplasme,
- la diffusion passive à travers la membrane interne.

Le modèle de transport et de phosphorylation des sources carbonées par EIIBC<sup>Glc</sup> suivant a été émis par l'équipe de B. Erni (Garcia-Alles et Erni, 2002, Garcia-Alles et al., 2002). Le dimère EIIBC<sup>Glc</sup> possèderait un site de faible et de forte affinité par monomère, où chaque monomère serait en sens opposé par rapport à la membrane. Ainsi, un site de forte affinité et de faible affinité sont situés de chaque côté de la membrane interne. Le substrat se fixerait sur le site de forte affinité côté périplasme, puis serait phosphorylé et transféré au site de faible affinité orienté vers le cytoplasme du même monomère, puis libéré. Le site de forte affinité de l'autre monomère, situé au niveau du cytoplasme, réaliserait la phosphorylation non vectorielle des substrats.

#### II.6 / Organisation des gènes du système PTS

Les gènes codant les protéines assurant le transport et le métabolisme des sources de carbone jusqu'à un intermédiaire de la glycolyse sont souvent regroupés sous la forme d'un opéron (pour revue, Plumbridge, 2002). Ceci est observé pour des sucres transportés par un système PTS, appelés sucres PTS, comme le fructose, où l'opéron fruBKA code les protéines qui assure le transport et le métabolisme du fructose, ou des sucres non PTS, comme le lactose. La transcription de ces différents opérons est régulée en général par un activateur transcriptionnel, le complexe CRP/AMPC (CRP: cAMP Receptor Protein = CAP: Catabolite Activator Protein) / AMPc, et un répresseur: son rôle est d'inhiber l'expression des gènes codant des protéines impliquées dans le transport et le métabolisme d'une source carbonée en absence de celle-ci (pour revue, Postma et al., 1993, Postma et al., 1996b). Par exemple, l'opéron nagE-BACD regroupe les gènes nécessaires au transport et au catabolisme du N-acétylglucosamine et glucosamine. En absence de substrat, l'initiation de la transcription au niveau des promoteurs divergents nagE P1 et nagB P1 est réprimée par la protéine NagC. Le répresseur NagC se fixe au niveau des boîtes B et E, centrées respectivement aux positions -125 et -31 par rapport au site +1 du promoteur nagE P1, et forme une boucle d'ADN. La répression par la protéine NagC est levée

en présence du N-acétylglucosamine-6-phosphate, produit du transport du N-acétyl glucosamine par son transporteur PTS, EIICBA<sup>nag</sup> (Plumbridge, 1991). La transcription de l'opéron *nagE-BACD* est aussi induite par le complexe CRP-AMPc: un site de fixation centré en position -61.5 par rapport au site d'initiation de *nagE* P1 et -71,5 par rapport à celui de *nagB* P1 (Plumbridge et Kolb, 1998). Il est aussi à remarquer que ce complexe CRP/AMPc renforce l'inhibition de la transcription en présence du répresseur NagC. La protéine NagC est produite à partir de deux promoteurs, *nagC* P1 et P2, situés dans la région codante du gène *nagA*, ce qui permet une expression constitutive du répresseur (Plumbridge, 1996).

L'organisation des gènes codant le système PTS spécifique du glucose est différente (pour revues, Postma et al., 1993; Postma et al., 1996b). Le complexe EII est composé des deux protéines EIIA<sup>Glc</sup> et EIICB<sup>Glc</sup>, codées respectivement par les gènes *crr* et *ptsG*. Le gène *crr* est localisé en aval des gènes *ptsH* et *ptsI*, codant les protéines générales HPr et EI. L'ensemble forme ainsi l'opéron *pts*. Le gène *ptsG* est indépendant (Fig. 45). La régulation de l'expression de ces gènes est complexe et sera décrite dans la partie E. Cette organisation particulière pourrait être due à la place fondamentale du métabolisme du glucose chez *E. coli*.

## III / Le métabolisme général des sources de carbone

Après leur transport, les sources de carbone utilisent des voies métaboliques spécifiques et rejoignent les voies métaboliques centrales à partir de différents branchements. Ces voies métaboliques centrales sont la glycolyse, la voie des pentoses phosphate, la voie d'Entner-Doudoroff, et le cycle de Krebs (Fig. 4). Chez *E. coli*, les gènes codant les enzymes intervenant dans les différentes voies métaboliques sont généralement regroupés en opéron, comme c'est le cas pour le cycle de Krebs et la voie d'Entner-Doudoroff. Ainsi, les gènes *sdhCDAB-sucABCD* codant des enzymes du cycle de Krebs sont transcrits à partir du promoteur P<sub>sdh</sub> situé en amont de l'opéron (Park *et al.*, 1997; Cunningham et Guest, 1998). Les gènes codant les enzymes de la voie d'Entner-Doudoroff sont regroupés au sein de l'opéron *edd-eda* (Egan *et al.*, 1992). Par contre, les gènes codant des protéines de la voie des pentoses phosphate et de la glycolyse sont majoritairement monocistroniques (pour revues, Sprenger, 1995; Fraenkel, 1996). Au niveau de la glycolyse, seul les gènes *pgk* et *fbaA* peuvent appartenir au même opéron.

Les sources de carbone transportées dans la cellule par un système PTS sont préférentiellement utilisées. Ceci s'explique par la limitation de l'expression des gènes codant des systèmes de transport et de dégradation de sources de carbone qui n'empruntent pas le système PTS lorqu'une source de carbone utilisant un système PTS est disponible dans le milieu. Ces régulations de l'expression des

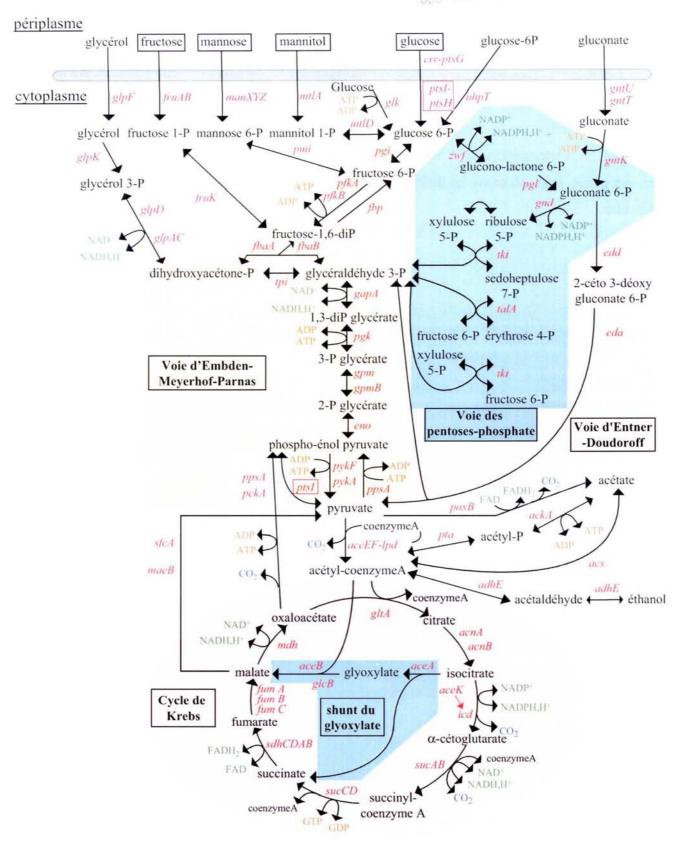


Figure 4 : Principales voies du métabolisme carboné et les gènes codant les différentes enzymes mises en jeu chez *E. coli*.

Les sources de carbone transportées dans la cellule par un système PTS sont encadrées en noir. Les gènes codant les enzymes impliquées dans les voies métaboliques sont indiqués en rouge. Les gènes *ptsI* et *ptsH*, codant les protéines générales EI et HPr du système PTS, respectivement, sont encadrés. Les différentes voies métaboliques (voie d'Embden-Meyerhof-Parnas, voie des pentoses phosphate, voie d'Entner-doudoroff, cycle de Krebs et shunt du glyoxylate) sont indiquées sur un fond de couleur différentes.

gènes sont exercées entre autre par les éléments du système PTS, au niveau de différentes étapes allant de l'initiation de la transcription à la régulation de l'activité des enzymes des voies métaboliques. Nous allons décrire par la suite les différents mécanismes et régulations s'exerçant au niveau de l'initiation de la transcription, de l'initiation de la traduction, de la dégradation des ARNm. Dans une dernière partie, nous décrirons les différentes régulations exercées au niveau des voies métaboliques carbonées.

# Partie B / La transcription chez Escherichia coli

Les bactéries vivent dans des habitats où les conditions changent très fréquemment. Afin de s'adapter aux différentes variations de leur environnement, elles ont développé des réponses très sophistiquées afin d'obtenir les bons produits en quantité et en temps utile (pour revue, Browning et Busby, 2004). Ces réponses passent le plus souvent par l'activation et/ou la répression d'un nombre de gènes pour permettre l'adéquation entre la physiologie cellulaire et/ou le métabolisme avec les nouvelles conditions de croissance rencontrées par la bactérie. Ainsi, les bactéries ont mis au point une large gamme de mécanismes pour réguler finement l'expression des gènes. Cette régulation est exercée par l'action de protéines régulatrices activant ou inhibant l'ARN polymérase au niveau de l'initiation de la transcription, modulant l'élongation et/ou la terminaison à des sites spécifiques. A ceci s'ajoute un autre niveau de régulation au niveau post-transcriptionnel, intervenant au niveau de la stabilité de l'ARN, l'efficacité d'initiation de la traduction, l'activité des protéines et leur dégradation (Rojo, 2001). Seront développées ici les régulations transcriptionnelles effectuées au niveau de l'initiation de la transcription.

La transcription des gènes se décompose en trois étapes : l'initiation, l'élongation et la terminaison, où chacune peut être le siège de différentes régulations. Afin d'éviter une dépense d'énergie inutile, l'étape clé de la régulation de la transcription est l'initiation. Chez les bactéries, la transcription est réalisée par une seule enzyme, l'ARN polymérase ADN dépendante, qui existe sous les formes apoenzyme et holoenzyme. Sous la forme apoenzyme, elle réalise l'élongation, soit la synthèse d'ARN dans le sens 5'-3' en utilisant l'ADN comme matrice, la terminaison correspondant à l'arrêt de la transcription. Elle doit s'associer avec un facteur σ pour former l'holoenzyme afin de se fixer sur l'ADN. Elle reconnaît alors une séquence particulière, appelée promoteur, située en amont de la région codante. Nous décrirons les différentes étapes de la transcription et leurs régulations.

# I / Le mécanisme d'initiation de la transcription

#### I.1 / L'apoenzyme ARN polymérase

Chez les bactéries, l'ARN polymérase est constituée des sous-unités  $\alpha$ ,  $\beta$ ,  $\beta$ ' et  $\omega$ , comprenant respectivement 329, 1342, 1407 et 91 acides aminés. L'assemblage de l'apoenzyme a été décrit comme étant un processus séquentiel ordonné (Ishihama, 1981) :

$$\alpha \longrightarrow \alpha_2 \longrightarrow \alpha_2\beta \longrightarrow \alpha_2\beta\beta'\omega$$

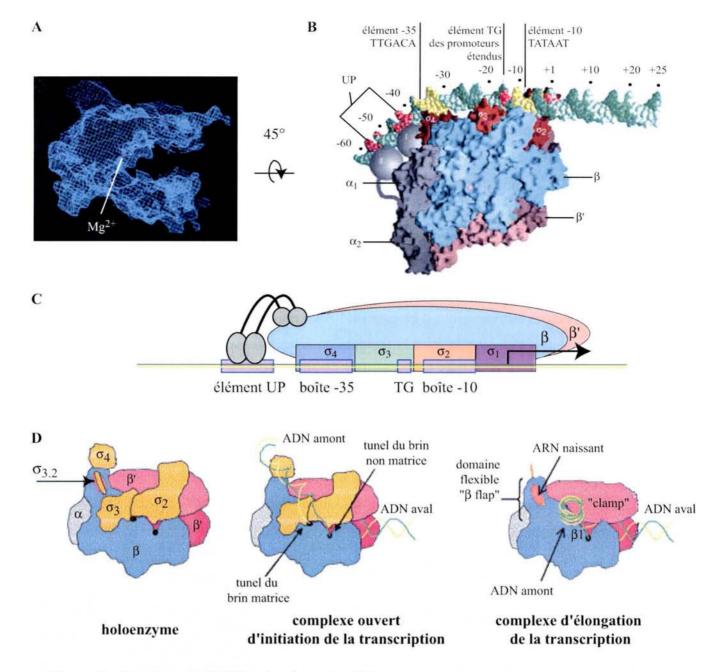


Figure 5 : Structure de l'ARN polymérase bactérienne.

A : Structure de l'apoenzyme d'E. coli à 15 Å, obtenue par cryomicroscopie (d'après Darst et al., 2002). Les sous-unités  $\beta$  et  $\beta$ ' forment une pince de crabe au niveau de l'apoenzyme. Le point rouge indique l'ion magnésium catalytique.

B : Modèle d'appariement entre l'holoenzyme et l'ADN au niveau d'un promoteur (d'après Darst et al., 2002 ; pour revue Browning et Darst, 2004). Les sous-unités  $\alpha$ -NTD sont indiquées en gris au niveau de la structure, tandis que les domaines  $\alpha$ -CTD sont représentés par des boules grises en interaction avec l'ADN. Les domaines  $\beta$  et  $\beta$ ' sont en bleu cyan et rose foncé, respectivement. Les différents domaines  $\sigma_2$ ,  $\sigma_3$  et  $\sigma_4$  du facteur  $\sigma$  sont colorés en rouge et sont en interaction avec l'ADN qui est représenté sous forme de double hélice.

C : Représentation schématique de l'holoenzyme en interaction avec l'ADN. Les couleurs des sousunités  $\beta$  et  $\beta$ ' sont les mêmes que précédemment. Les sous-unités  $\alpha$ -NTD et  $\alpha$ -CTD sont représentées par des boules grises. Les différents domaines de l'ADN pouvant être en interaction avec l'holoenzyme sont schématisés par un rectangle violet entouré d'un cadre bleu.

D: Représentation schématique de l'ARN polymérase lorsqu'elle est associée au facteur  $\sigma$  ou présente dans les complexes d'initiation ou d'élongation de la transcription. Le même code couleur a été utilisé pour l'ARN polymérase. Les différents domaines du facteur  $\sigma$  sont en jaune-orange, le brin d'ARN naissant en orange, les brins d'ADN matrice et non matrice en jaune et vert, respectivement (d'après Young et al., 2002).

Il débute par la dimérisation du domaine N-terminal de la sous-unité  $\alpha$  (domaine  $\alpha$ -NTD constitué des acides aminés 1 à 235), ce qui permet ensuite la fixation des sous-unités  $\beta$  et  $\beta$ '. Les sous-unités  $\beta$  et  $\beta$ ' constituent les deux pinces de la forme en pince de crabe adoptée par l'apoenzyme (Fig. 5A) et contiennent le centre actif de l'enzyme, localisé au niveau de la réunion des deux pinces. Le facteur  $\omega$  se lie au niveau de la partie C-terminale de la sous-unité  $\beta$ ', et stabilise l'association de  $\beta$ ' avec  $\alpha 2\beta$  (Minakhin *et al.*, 2001). Contrairement aux autres éléments de l'apoenzyme, le facteur  $\omega$  n'est pas essentiel pour la survie de la bactérie (Darst *et al.*, 2002).

## I.2 / Les différentes holoenzymes chez E. coli

La fixation du facteur  $\sigma$  à l'apoenzyme lui confère une affinité pour toute séquence d'ADN, ce qui lui permettrait de s'accrocher à une séquence d'ADN, puis de glisser le long de la double hélice jusqu'à la séquence d'un promoteur, reconnu avec une plus forte affinité. Le facteur  $\sigma$  réalise les trois fonctions suivantes (pour revue, Browning et Busby, 2004) :

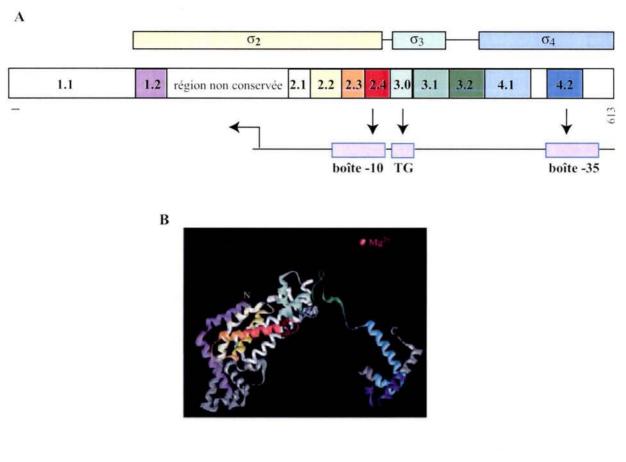
- la reconnaissance de séquences spécifiques de promoteurs,
- le positionnement correct de l'ARN polymérase au niveau du promoteur cible,
- l'ouverture de la double hélice d'ADN à proximité du site d'initiation de la transcription.

Il pourrait aussi jouer un rôle dans l'ouverture et la fermeture des « pinces de crabe » de l'apoenzyme pendant les différentes étapes du processus d'initiation de la transcription (Darst et al., 2002).

Les structures des holoenzymes de *Thermophilus aquaticus* et *Thermus thermophilus* montrent que le facteur  $\sigma$  est localisé sur une face de l'enzyme (Murakami *et al.*, 2002 ; Vassylyev *et al.*, 2002) (Fig. 5, B et C), et interagit avec des élément flexibles de l'apoenzyme situés au niveau des sous-unités  $\beta$  et  $\beta$ ' : le « clamp » pour  $\beta$ ', le domaine  $\beta$ 1 et le «  $\beta$  flap » pour  $\beta$  (Fig. 5D).

Différents paramètres évitent la liaison des sous-unités  $\sigma$  libres à l'ADN qui pourrait rendre les promoteurs inaccessibles aux holoenzymes par inhibition stérique. En particulier, lorsque le facteur  $\sigma$  est à l'état libre, la distance séparant les régions qui interagissent avec les boîtes -10 et -35 ne permet pas de réaliser une interaction simultanée avec ces deux boîtes. Lorsque le facteur  $\sigma$  s'associe avec l'apoenzyme, il subit un réarrangement structural ce qui conduit au positionnement optimal des régions d'interactions du facteur  $\sigma$  aux boîtes -10 et -35.

La majorité des bactéries possèdent, en général, de multiples facteurs  $\sigma$ , de 1 pour *Mycoplasma* genitalium à 63 chez Streptomyces coelicolor (pour revue, Browning et Busby, 2004). Des alignements de séquences protéiques ont permis de différencier deux types de familles de facteurs  $\sigma$  bactériens, les familles  $\sigma^{70}$  et  $\sigma^{54}$  (pour revue, Wosten, 1998). Ces familles comprennent respectivement 4 et 3 régions fortement conservées (Fig. 6, A et C) (pour revue, Wosten, 1998). *E. coli* contient 7 facteurs  $\sigma$ , où 6



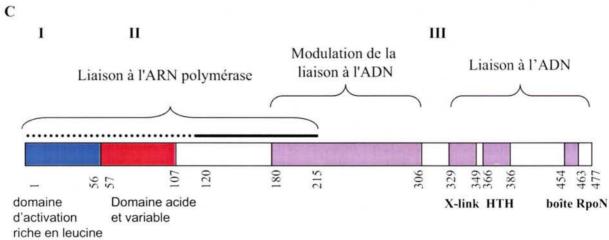


Figure 6 : Organisation des facteurs  $\sigma^{70}$  et  $\sigma^{54}$  et structure de  $\sigma^{A}$  chez *T. thermophilus* (d'après Buck *et al.*, 2000 ; Borukhov et Severinov, 2002).

A : Organisation du facteur  $\sigma^{70}$  et ses domaines d'interactions avec l'ADN. Les trois domaines identifiés par radiocristallographie sont indiqués par les domaines  $\sigma_2$ ,  $\sigma_3$  et  $\sigma_4$ , en jaune, vert et bleu, respectivement. Les différents sous-domaines identifiés par les alignements de séquences et caractérisant la famille de facteur  $\sigma^{70}$  sont représentés par différentes couleurs.

B : Structure du facteur  $\sigma^A$  de *T. thermophilus* obtenue par radiocristallographie. Les différentes couleurs correspondent à celles des différents domaines en A.

C : Organisation du facteur  $\sigma^{54}$  chez *E. coli* et domaine d'interactions avec l'apoenzyme et l'ADN. Les régions I (bleu), II (rouge) et III (violet) conservées et leurs fonctions sont indiquées.

appartiennent à la famille  $\sigma^{70}$  et 1 à la famille  $\sigma^{54}$  (Tableau 1). Les membres de la famille  $\sigma^{70}$ , comprenant les facteurs  $\sigma^{70}$ ,  $\sigma^{S}$ ,  $\sigma^{F}$ ,  $\sigma^{32}$ ,  $\sigma^{E}$  et  $\sigma^{19}$ , peuvent reconnaître et former des complexes transcriptionnellement compétents en absence de source d'énergie. Par contre, les membres de la famille de facteur  $\sigma^{54}$  conduisent à la formation d'une holoenzyme qui a besoin de source d'énergie sous forme d'ATP ou de GTP, pour la formation d'un complexe d'initiation de la transcription actif.

Les différents facteurs sigma reconnaissent des séquences de promoteurs différents, ce qui permet la transcription de classes distinctes de gènes, ou la transcription d'un même gène à partir de promoteurs multiples. Le taux des différentes holoenzymes dépend de la concentration intracellulaire en facteur  $\sigma$ , qui varie en fonction des conditions de croissance, et de son affinité pour l'ARN polymérase sous sa forme apoenzyme. La mesure de ces deux paramètres permet d'estimer le nombre de molécules par cellule de chaque holoenzyme présente, en phase exponentielle de croissance : 550 pour  $E\sigma^{70}$ , 95 pour  $E\sigma^{F}$ , 55 pour  $E\sigma^{N}$ . En phase stationnaire de croissance, la quantité de facteur  $\sigma^{S}$  augmente, et peut représenter jusqu'à 30 % de la quantité de molécule du facteur  $\sigma^{70}$  (pour revue, Ishihama, 2000). Le facteur  $\sigma^{70}$  intervient dans la reconnaissance d'une large majorité des promoteurs transcriptionnels, ce qui assure la transcription de base chez *E. coli*.

# I.3 / Les promoteurs reconnus par l'holoenzyme $E\sigma^{70}$

Le facteur  $\sigma^{70}$  permet la reconnaissance spécifique d'une séquence d'ADN, appelée promoteur de type  $\sigma^{70}$ . En effet, l'alignement des séquences nucléotidiques situées en amont des sites d'initiation de la transcription des gènes d'*E. coli* transcrits par l'holoenzyme  $E\sigma^{70}$  a conduit à l'identification de deux hexamères conservés, possédant le motif consensus suivant :

Le site d'initiation de la transcription est séparé de la boîte -10 par 4 à 8 bp. Les deux hexamères conservés sont centrés approximativement 10 et 35 pb en amont du site d'initiation de la transcription, et sont appelés boîtes -10 et -35, respectivement. Un promoteur typique comprend au moins quatre bases sur six au niveau de chacun des 2 hexamères (Mitchell *et al.*, 2003). Ces 2 hexmères sont séparés par une séquence non conservée de 16 à 18 pb, 17 bp étant la distance la plus fréquente. Un espacement de 16 pb peut être compensé par une diminution du nombre de supertours négatifs de l'ADN et favoriser ainsi la transcription. Au contraire, une augmentation du surenroulement facilite l'initiation de la transcription des promoteurs ayant une séquence interboîte de 18 pb (pour revue, Wang et Syvanen, 1992).

| Facteur               | Gène | PM<br>(Da) | Spécificité de reconnaissance   | Fonction  | Références   |
|-----------------------|------|------------|---|---|--|
| $\sigma^{70}$         | rpoD | 70 236     | TTGACA N <sub>16-18</sub> TATAAT  | Majorité des gènes<br>d'E. coli                                 | Harley et Reynolds,<br>1987  |
| σ <sup>38 (S)</sup>   | rpoS | 38 000     | CTGCAA N <sub>14-16</sub> CGCCAAGT  | Réponses à de<br>nombreux stress                                | Lange et Hengge-<br>Aronis, 1991; Kolb et<br>al., 1995; Tanaka et<br>al., 1995 |
| $\sigma^{32}$         | RpoH | 32 000     | CTTGAA N <sub>13-15</sub> CCCCATNTA   | Réponse à la<br>dénaturation des<br>protéines<br>cytoplasmiques | Cowing <i>et al.</i> , 1985  |
| σ <sup>24(E)</sup>    | rpoE | 24 000     | GAACTT N <sub>16</sub> TCTGA  | Réponse à la<br>dénaturation des<br>protéines<br>périplasmiques | Erikson et Gross, 1989   |
| σ <sup>28(F)</sup>    | rpoF | 28 000     | TAAAGTT N <sub>11</sub> GCCGATAA  | Gènes de synthèse<br>des flagelles                              | Helmann et<br>Chamberlin, 1987   |
| σ <sup>19(Fecl)</sup> | fecI | 19 000     | GGAAAAT N <sub>17</sub> TCCTT   | Opéron de<br>transport du citrate<br>ferrique                   | Angerer et al., 1995;<br>Enz et al., 1995                                      |
| σ <sup>54(N)</sup>    | rpoN | 54 000     | aaN3TGGCAc N <sub>6</sub> TGCNNt<br>petites lettres : 2 à 5 mismatches<br>lettres capitales : 0 ou 1<br>mismatches<br>bases soulignées : 0 mismatches | Gène du<br>métabolisme de<br>l'azote                            | Reitzer et Schneider,<br>2001  |

Tableau 1 : Les différents facteurs sigma mis en évidence chez E. coli.

Pour chaque facteur sigma sont indiqués le nom de son gène, son poids moléculaire, la séquence reconnue par l'holoenzyme correspondante ainsi que le type de gènes transcrits par cette holoenzyme.

Deux autres éléments de séquences peuvent aussi faciliter la reconnaissance du promoteur par l'holoenzyme  $E\sigma^{70}$ : l'élément -10 étendu et l'élément UP. L'élément -10 étendu est constitué du motif 5'-TRTG-3', avec R = A ou G, situé en position -17 à -14 par rapport au site d'initiation de la transcription (il est séparé d'un nucléotide de l'hexamère -10) (Mitchell *et al.*, 2003). Le dinucléotide 5'-TG-3' (position -15 à -14) interagit avec le domaine  $\sigma$ 3.0 du facteur  $\sigma$ 70, tandis que le dinucléotide 5'-TR-3' situé en amont semble affecter la structure et/ou la flexibilité de l'ADN pour faciliter l'interaction du facteur  $\sigma$  avec la région aval. Ce dinucléotide 5'-TR-3' est présent dans 20 % des promoteurs possèdant un élément étendu. L'élément UP est une séquence d'environ 20 nucléotides, située en amont de la boîte -35, et contactée par le domaine  $\alpha$ -CTD de l'ARN polymérase (Fig. 5B) (Estrem *et al.*, 1999 , Gourse *et al.*, 2000). Cet élément UP, favorise la fixation de l'ARN polymérase au niveau du promoteur et l'ouverture de l'ADN, mais pourrait ralentir les étapes postérieures à l'ouverture de l'ADN au sein de l'ARN polymérase (Ellinger *et al.*, 1994).

L'holoenzyme peut ainsi reconnaître un promoteur, par l'intermédiaire de ces quatre éléments, mais aucun promoteur naturel ne comporte les séquences consensus de ces quatre éléments.

## I.4 / Les facteurs o d'E. coli

#### I.4.a / La famille de facteurs $\sigma^{70}$

# I.4.a. $\alpha$ / Les différents groupes de la famille de facteurs $\sigma^{70}$

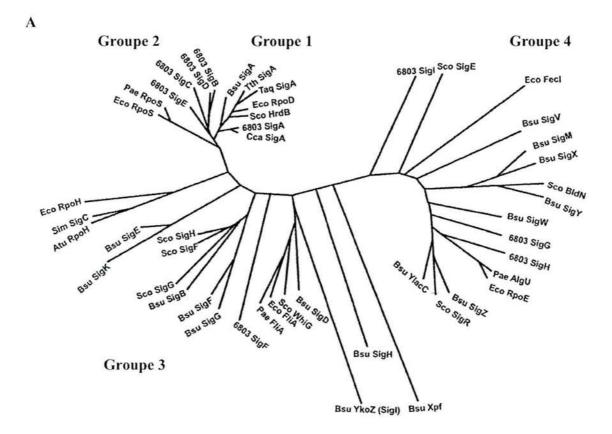
De multiples membres de cette famille ont été découverts dans la plupart des bactéries séquencées. Ils ont été classés en tant que facteur primaire ou alternatif suivant leur fonction. Les facteurs primaires, essentiels, sont responsables de la transcription des gènes de ménages exprimés pendant la phase exponentielle de croissance, tandis que les facteurs alternatifs, non essentiels, sont responsables de la transcription des gènes en réponse à des stimuli environnementaux. La famille de facteur  $\sigma^{70}$  a été divisée en 4 groupes phylogénétiques (Fig. 7).

Le groupe 1 est constitué des facteurs primaires essentiels, tel que  $\sigma^{70}$  (rpoD) d'E. coli.

Les facteurs du groupe 2 sont proches des facteurs du groupe 1, par contre ils ne sont pas essentiels pour la croissance cellulaire, comme  $\sigma^{S}$  (rpoS) chez E. coli.

Les facteurs du **groupe 3**, plus éloignés que le groupe 2, activent les régulons à une réponse à un signal spécifique comme un choc thermique tel que le facteur  $\sigma^{32}$  (rpoH) chez E. coli.

Le **groupe 4** réunit des facteurs fort divergents en séquence des autres facteurs de la famille  $\sigma^{70}$  et, en général, ils permettent la transcription des gènes dont les protéines ont des fonctions extracytoplasmiques (ils ont été appelés facteurs ECF pour ExtraCytoplasmique Function), comme



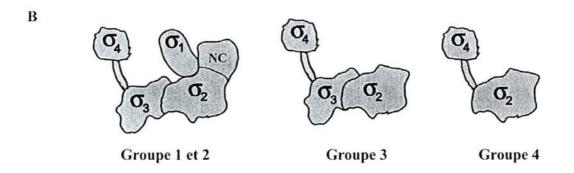


Figure 7 : Les différents groupes de la famille de facteurs  $\sigma^{70}$  (pour revue, Gruber et Gross, 2003).

A: Cet arbre a été établi suivant les alignements de séquences des régions 1.2 à 2.4 des facteurs σ, pour les bactéries suivantes : Eco : E. coli ; Bsu : B. subtilis ; 6803 : Synechocystis sp. PCC6803 ; Sco S. coelicolor ; Pae P.aeruginosa ; Tth : T. themophilus ; Taq : T. aquaticus, Cca : Cyanidium caldarium ; Sim : Sinorhizobium meliloti ; Atu : Agrobacterium tumefaciens.

B : Représentation des différents domaines des facteurs sigma présents dans les groupes 1, 2 , 3 et 4 de la famille de facteur  $\sigma^{70}$ . NC indique la région non conservée, retrouvée dans quelques facteurs de la transcription des gènes de ménages.

 $\sigma^{24(E)}$  (rpoE) et  $\sigma^{19}$  (fecI) chez E. coli. Chez de nombreuses bactéries, le nombre de facteurs différents appartenant à ce groupe peut dépasser la somme des autres facteurs sigma, tels que S. coelicolor qui possède 49 facteurs  $\sigma$  ECF parmi un total de 63 facteurs  $\sigma$  (pour revue, Helmann, 2002).

# I.4.a. $\beta$ / Structure et rôles des différentes régions des facteurs $\sigma^{70}$

Les alignements de séquences protéiques de la famille des facteurs  $\sigma^{70}$  a permis de différencier 4 régions conservées, désignées 1 à 4. Les régions 2 et 4 sont retrouvées dans tous les membres de la famille  $\sigma^{70}$ : elles sont impliquées dans la liaison à l'ARN polymérase, la reconnaissance des promoteurs, et la séparation de l'ADN double brin (ADNdb) (Fig. 6A et 7B) (pour revue, Wosten, 1998). Ces régions conservées correspondent grossièrement aux domaines structuraux du facteur sigma (Campbell *et al.*, 2002a). Les 3 domaines identifiés par radiocristallographie  $\sigma_2$ ,  $\sigma_3$  et  $\sigma_4$ , sont tous constitués d'hélice  $\sigma_3$  (Fig. 6, A et B). Ils correspondent, respectivement, aux régions 1.2 à 2.4, 3.0 (région appelée anciennement 2.5 renommée 3.0 (Campbell *et al.*, 2002a) à 3.1 et 4.1 à 4.2. Ces domaines sont reliés entre eux par des liens flexibles. Une telle organisation structurale permet alors différentes conformations du facteur  $\sigma_3$  par réarrangement des positions relatives des domaines structuraux entre eux, plutôt qu'à des changements dans les domaines structuraux eux-mêmes. Ces différences de conformation expliqueraient les différences de propriétés biochimiques des facteurs  $\sigma_3$  à l'état libres ou associés à l'apoenzyme (pour revue, Borukhov et Severinov, 2002).

Les différents groupes de la famille de facteurs  $\sigma^{70}$  se différencient aussi par le nombre de régions conservées qu'ils possèdent (Fig. 7B). Ainsi, les 4 régions conservées sont retrouvées dans les facteurs sigma des groupes 1 et 2, le groupe 2 ne comprenant pas le domaine 1.1 non conservé (domaine NC, Fig. 7B). La région 1 n'est pas conservée dans les facteurs sigma du groupe 3. Les facteurs sigma du groupe 4 ne possèdent que les deux régions 2 et 4 essentielles (pour revue, Gruber et Gross, 2003).

Ces différentes régions du facteur  $\sigma^{70}$  ont été subdivisées en sous-régions (Fig. 6, A et B).

#### La région 1 est décomposée en deux sous-régions : 1.1 et 1.2.

La sous-région 1.1, possédant de fortes charges négatives, inhibe de façon indirecte la reconnaissance de la boîte -35 par le facteur  $\sigma^{70}$  libre (Camarero *et al.*, 2002). Dans les structures d'holoenzymes de *T. aquaticus* en absence d'ADN, cette sous-région 1.1 est localisée à proximité du site actif de l'ARN polymérase, et cacherait le canal de liaison à l'ADN double brin en aval du site d'initiation de la transcription. Cette région mimerait donc l'ADN et serait déplacée suite à la formation du complexe avec le promoteur (Fig. 8, A et B) (pour revues, Borukhov et Severinov, 2002; Murakami et Darst, 2003).

Cependant, la délétion de cette région 1.1 réduit la vitesse de transition entre les étapes qui conduisent à l'ouverture de la double hélice d'ADN, ce qui suggère que cette région 1.1 intervient dans l'interaction de l'holoenzyme avec le promoteur (pour revue, Murakami et Darst, 2003).

La sous-région 1.2 semble être nécessaire aux changements conformationnels réalisés lors de l'association du facteur  $\sigma^{70}$  avec l'apoenzyme, ou lors de la liaison de l'holoenzyme à l'ADN (Baldwin et Dombroski, 2001). La double délétion des regions 1.1 et 1.2 permet la fixation de l'holoenzyme à l'ADN, mais bloque les étapes ultérieures (Baldwin et Dombroski, 2001).

La région 2 a été subdivisée en 4 régions 2.1, 2.2, 2.3 et 2.4 (pour revue (Helmann et Chamberlin, 1988). Les sous-régions 2.1 à 2.4 avec la sous-région 1.2 constituent le domaine σ<sub>2</sub> cristallisé, constitué de 3 hélices α (Fig. 6, A et B) (Malhotra *et al.*, 1996). L'hélice 2 est le point de contact majeur avec le domaine en fagot de la sous-unité β' de l'apoenzyme (Campbell *et al.*, 2002a). Cette interaction lève une interaction autoinhibitrice du facteur σ à l'état libre qui interfère avec la reconnaissance de la boîte -10 (pour revue, Borukhov et Severinov, 2002). L'hélice 3 comprend la région 2.3, riche en résidus aromatiques et basiques, qui contacte l'ADN et permet l'ouverture de la double hélice (Panaghie *et al.*, 2000; Young *et al.*, 2004; pour revue, Borukhov et Severinov, 2002). En particulier, les acides aminés Tyr 425, Tyr 430, Trp 433 et Trp 434 seraient impliqués dans la stabilisation de l'ouverture de la double hélice d'ADN (Fenton *et al.*, 2000; Panaghie *et al.*, 2000). La sous-région 2.4 serait responsable de l'interaction avec la boîte -10 des promoteurs. En effet, des fragments de σ contenant cette région 2 se lient de façon spécifique au brin non codant de l'élément -10 du promoteur, en présence de l'ARN polymérase (Marr et Roberts, 1997; Murakami *et al.*, 2002).

La région 3 est décomposée en trois régions 3.0, 3.1 et 3.2. Le domaine structural  $\sigma_3$  de T. thermophilus comprend les régions 3.0 et 3.1 et est constitué de 3 hélices  $\alpha$ . La sous-région 3.0, située au niveau de l'hélice 1, interagit avec le dinucléotide TG des promoteurs étendus (Barne et al., 1997). La région 3.1 est constituée d'un domaine hélice-tour-hélice (HTH) commune aux nombreuses protéines de liaison à l'ADN double brin. Cette région 3.1 semble interagir non pas avec l'ADN double-brin comme le laisserait supposer la présence de ce domaine, mais avec l'apoenzyme (Campbell et al., 2002a). Dans les structures d'holoenzymes de T. aquaticus et T. thermophilus, la région 3.2 constitue la partie flexible entre les domaines structuraux  $\sigma_3$  et  $\sigma_4$  (Murakami et al., 2002; Vassylyev et al., 2002, pour revue, Paget et Helmann, 2003). Elle est située vers le centre actif de l'ARN polymérase, et occuperait le canal de sortie de l'ARN naissant (Fig. 5D).

La région 4 contient 2 sous-régions 4.1 et 4.2 et constitue le domaine structural  $\sigma_4$ . Ce domaine est constitué de 2 paires d'hélices, en forme de C. Les résidus hydrophobes recouvrant

l'interieur de la structure en forme de C interagissent avec le domaine flexible « $\beta$  flap » de la sousunité  $\beta$  (pour revue, Murakami et Darst, 2003). Cette interaction conduit au positionnement correct des régions 2.4 et 4.2 requis pour la reconnaissance simultanée des éléments -10 et -35 du promoteur (Kuznedelov et al., 2002). La structure aux rayons X du facteur  $\sigma^A$  de T. aquaticus avec l'ADN montre que la région 4.2 a cette conformation typique HTH : elle se lie à l'élément -35 par des interactions spécifiques avec les groupements phosphates et certaines bases au niveau du grand sillon (Campbell et al., 2002a). La sous-région 4.2 semble aussi réaliser des contacts avec des activateurs transcriptionnels, qui se substituent aux contacts avec l'ADN, lorsque ceux-ci se fixent en position -35, tels que le complexe CRP-AMPc ou le régulateur de réponse PhoB (Campbell et al., 2002a).

La fixation du domaine  $\sigma_4$  à l'ADN introduit une torsion de l'ADN de 36°, ce qui facilite les interactions entre le domaine  $\alpha$ -NTD de l'ARN polymérase et l'ADN amont (Fig. 5, B et D) (Campbell et al., 2002a; pour revue, Murakami et Darst, 2003).

## I.4.a. $\chi$ / Les membres de la famille des facteurs $\sigma^{70}$ chez E. coli

# Représentant du groupe 1 chez E. coli: le facteur o<sup>70</sup>

Le facteur  $\sigma^{70}$  est codé par le gène rpoD. Il est responsable de la transcription de la majorité des gènes exprimés en phase exponentielle de croissance (pour revue, Wosten, 1998).

La disponibilité du facteur  $\sigma^{70}$  est modulée par l'anti-facteur  $\sigma^{70}$  Rsd (Regulator of sigma D). La production de cette protéine est induite pendant la phase stationnaire (Jishage et Ishihama, 1999). La protéine Rsd se fixe au facteur  $\sigma^{70}$ , ce qui empêche sa fixation à l'apoenzyme (Jishage et Ishihama, 1998).

# Représentant du groupe 2 chez E. coli: le facteur $\sigma^{S(38)}$

Ce facteur est codé par le gène rpoS, et est très similaire à  $\sigma^{70}$ . La majorité des régulations de la production du facteur  $\sigma^S$  ont lieu au niveau post-transcriptionnel, soit au niveau de la traduction, soit au niveau de la stabilité du facteur  $\sigma^S$  (pour revue, Hengge-Aronis, 2002a).

Ce facteur  $\sigma^S$  est induit pendant la phase stationnaire et peut représenter jusqu'à 30 % de la quantité de  $\sigma^{70}$  (pour revue, Ishihama, 2000). Les holoenzymes  $E\sigma^S$  et  $E\sigma^{70}$ , bien que pouvant conduire à l'expression de gènes distincts, reconnaissent des séquences promoteurs très similaires. La technique de SELEX (Systematic Evolution of Ligands by EXponential enrichment), qui permet d'extraire, grâce à leur affinité, des ligands potentiels d'une banque d'oligonucléotides composée initialement de séquences aléatoires, a été utilisée pour optimiser la séquence du promoteur *in vitro*. Cette méthode montrent que les holoenzymes  $E\sigma^S$  et  $E\sigma^{70}$  reconnaîssent les mêmes boîtes -10 et -35 (Gaal *et al.*, 2001; pour revue, Hengge-Aronis, 2002b) :

Elément UP boîte -35 région boîte -10 +1
-10 étendue

 $E\sigma^{70}$  AAAWWTWTTTT NNNAAA -5bp - TTGACA -12 bp - TGTG N TATA A T - 4-9 pb - N  $E\sigma^{S}$  AWWWWTTTTT -11 bp - TTGACA -12 bp - TGTG C TATAC/AT - AT-riche - N

Les éléments importants sont indiqués en gras, et ceux qui contribuent à la sélectivité de  $E\sigma^S$  sont soulignés.

Une cytosine en amont de la boîte -10 est plus fréquemment retrouvée dans les promoteurs reconnus par  $E\sigma^S$ . Elle interagirait avec la lysine 173 de la région 3.0 de  $\sigma^S$  (pour revue, Hengge-Aronis, 2002b). En amont de la boîte -35, un élément UP entier ou semi-proximal favorise la reconnaissance du promoteur par l'holoenzyme  $E\sigma^{70}$  alors qu'un élément UP semi-distal favorise une reconnaissance du promoteur par l'holoenzyme  $E\sigma^S$  (Typas et Hengge, 2005).

L'holoenzyme  $E\sigma^S$  peut reconnaître des boîtes -35 dégénérées plus efficacement que l'holoenzyme  $E\sigma^{70}$  et être employée dans des conditions où les interactions entre l'ARN polymérase et le promoteur sont diminuées, telles qu'une forte concentration en sels (Gaal *et al.*, 2001, pour revue, Hengge-Aronis, 2002b). Cette tolérance pourrait s'expliquer par la plus forte interaction entre le domaine flexible «  $\beta$  flap » de la sous-unité  $\beta$  et la région 4 de  $\sigma^S$ , ce qui conduit au positionnement correct de la région 4 au niveau de la boîte -35 (pour revue, Borukhov et Severinov, 2002).

# Représentant du groupe 3 chez *E. coli* : les facteurs $\sigma^{28}$ et $\sigma^{32}$

Le facteur  $\sigma^{28(F)}$  est codé par le gène rpoF (appelé aussi fliA). Il permet la transcription des gènes codant des protéines impliquées dans la mobilité des bactéries, dont les gènes nécessaires à la formation des flagelles. Son activité est régulée par la fixation d'un facteur anti-sigma FlgM sur ce facteur  $\sigma^F$ .

Le facteur  $\sigma^{32}$  est codé par le gène rpoH (Grossman et al., 1984). Il est impliqué dans la réponse à différents stress tels qu'un choc thermique, qui conduisent à la dénaturation de protéines cytoplasmiques, mais aussi en réponse à une carence carbonée (Cowing et al., 1985; Jenkins et al., 1991). Il permet l'activation transitoire de la production de la majorité des protéines de choc thermique, comprenant les protéines chaperones moléculaires comme DnaK, DnaJ, GrpE, GroEL et des protéases (pour revue, Mager et De Kruijff, 1995).

La transcription de  $\sigma^{32}$  est initiée à partir de 4 promoteurs, le promoteur P3 de type  $\sigma^E$  qui est activé lors d'un choc thermique, et 3 promoteurs type  $\sigma^{70}$ : P1, P4 et P5.

# Représentant du groupe 4 chez E. coli: les facteurs $\sigma^{E(24)}$ et $\sigma^{19}$ (fec1)

Le facteur  $\sigma^{E(24)}$  est codé par le gène rpoE (Raina et al., 1995 ; Rouviere et al., 1995), et appartient à l'opéron rpoE-rseABC. Ce facteur  $\sigma^E$  est crucial pour maintenir l'intégrité structurale de la cellule : il

est activé à la fois par des inducteurs de stress généraux, comme un choc thermique, et par des inductions spécifiques, comme la surexpression de porines, telle que OmpC (Ades *et al.*, 1999). Ces différentes inductions conduisent à l'accumulation de protéines mal repliées dans l'enveloppe cellulaire (Ades *et al.*, 2003). En absence de stress, le facteur  $\sigma^E$  est lié au domaine N-terminal de la protéine RseA, avec une affinité de 300 fois supérieure à celle pour l'apoenzyme (pour revue, Ades, 2004). Le domaine C-terminal de RseA périplasmique se lie à RseB, ce qui le stabilise et augmente l'affinité de RseA pour  $\sigma^E$  (Collinet *et al.*, 2000). Par contre, la protéine RseC localisée au niveau de la membrane interne se lierait à RseA, inhibant sa fonction anti-facteur  $\sigma^E$ . L'appartenance à un même opéron pour  $\sigma^E$ , RseA, RseB et RseC permet une régulation fine de l'activité de  $\sigma^E$ .

Lors d'un stress, l'accumulation de protéines mal repliées dans l'enveloppe conduit à la dégradation protéolytique séquentielle du facteur anti-sigma RseA par les protéases membranaires DegS et YeaL et à la libération du facteur  $\sigma^E$  de RseA. Le facteur  $\sigma^E$  se lie à l'ARN polymérase et initie la transcription d'une vingtaine de gènes regroupant des protéases périplasmiques, des protéines facilitant le repliement de protéines, et des enzymes impliquées dans les voies métaboliques de biosynthèse de constituants de l'enveloppe cellulaire (Dartigalongue *et al.*, 2001 ; pour revue, Ades, 2004).

Le facteur σ<sup>19</sup> ou FecI, permet la transcription des gènes *fecABCDE* codant les protéines impliquées dans le transport du citrate ferrique chez *E. coli* (pour revue, Braun *et al.*, 2003). Le gène *fecA* code un transporteur du citrate de la membrane externe, *fecBCDE* un transporteur ABC de la membrane interne. L'activation de la transcription de l'opéron *fecABCDE* a lieu uniquement en présence de citrate ferrique et lors d'une limitation en fer. Les gènes de régulation de cet opéron, *fecIR*, sont situés en amont et forment une unité transcriptionnelle différente. Le facteur transcriptionnel Fur est activé en présence de fer, et réprime la transcription de *fecIR*. Lors de conditions limitées en fer, la répression de la transcription de l'opéron *fecIR* est levée. La présence de citrate ferrique induit ensuite la libération du facteur FecI de la protéine FecR, qui joue un rôle de facteur anti-sigma FecI et un rôle de chaperone pour éviter sa dégradation, ce qui conduit à la transcription de l'operon *fecABCDE*.

#### I.4.b / La famille de facteurs $\sigma^{54}$

Le facteur  $\sigma^{54}$ , appelé aussi  $\sigma^N$ , est codé par le gène rpoN (Hunt et Magasanik, 1985). Il serait exprimé de manière constitutive, et son taux apparaît constant dans différentes conditions de croissance (pour revue, Reitzer et Schneider, 2001). Il permet la transcription d'une vingtaine de gènes dont les protéines sont impliquées dans le métabolisme de l'azote, mais aussi dans l'utilisation de sources de carbone, la synthèse de flagelles et la virulence bactérienne. Il se différencie des facteurs de la

famille  $\sigma^{70}$  par sa séquence en acides aminés (Fig. 6) et par le mécanisme d'initiation de la transcription. Le facteur  $\sigma^{54}$  reconnaît deux régions centrées aux positions -12 et -24 par rapport au site +1 d'initiation de la transcription, seul ou en association avec l'apoenzyme (Hunt et Magasanik, 1985; Buck et Cannon, 1992). L'holoenzyme Eσ<sup>54</sup> forme un complexe inactif avec l'ADN (complexe fermé) (Sasse-Dwight et Gralla, 1988). Elle requiert la présence d'activateurs pour catalyser la séparation des deux brins (Buck et al., 2000). Ceux-ci se lient à l'ADN aux sites qui sont fonctionnellement indépendants de la distance et de l'orientation de l'ARN polymérase. Leur action peut être facilitée par la fixation sur l'ADN des protéines IHF et ArgR qui modifient la courbure de l'ADN. L'activité GTPase de l'activateur permet un changement conformationnel de  $E\sigma^{54}$  et l'initiation de la transcription. Cette activité GTPase est contrôlée par des interactions avec des ligands de faibles poids moléculaires, ou avec différentes protéines de régulations (Shingler, 1996). L'initiation de la transcription par le facteur  $\sigma^{54}$  dépend aussi de la concentration en azote via le système à deux composant Ntr (Nitrogen regulated response). Lorsque l'azote est abondant, une cascade de réactions conduit à stimuler la déphosphorylation par la protéine détectrice NRII (ou NtrB) du régulateur de réponse NRI (appelé aussi NtrC). NRI ne peut alors plus activer  $E\sigma^{54}$ . Lorsque l'azote est limitant, NRII s'autophosphoryle et transfère son phosphate à NRI, qui va ensuite activer l'expression des gènes correspondants.

## I.5 / Les différentes étapes de l'initiation de la transcription

Les différentes étapes correspondant au processus d'initiation de la transcription peuvent être représentées ainsi (pour revue, Record *et al.*, 1996 ; Buckle *et al.*, 1999 ; Rojo, 2001) :

$$R+P \xrightarrow{k1} RPc1 \xrightarrow{k2} RPc2 \xrightarrow{k3} RPo \xrightarrow{NTP} RPinit \xrightarrow{NTP}$$

(avec R = RNA polymérase, P = Promoteur, RPc = complexe fermé (closed), RPo = complexe ouvert (open), RPinit = complexe d'initiation et RPelong = complexe d'élongation)

Dans une première étape, l'apoenzyme s'associe avec l'un des 7 facteurs  $\sigma$  d'E. coli décrits précédemment. L'holoenzyme comprenant un facteur de la famille  $\sigma^{70}$ , reconnaît alors la région promoteur de l'ADN, par une interaction des domaines  $\sigma_4$  et  $\sigma_2$  avec, respectivement, les boîtes -35 et -10 (Fig. 5, 6 et 8) et forme un complexe fermé RPc (RNA polymerase / promoter closed complex), où les deux brins d'ADN sont toujours appariés (Fig. 5, B et C et 8A). Les acides aminés aromatiques des

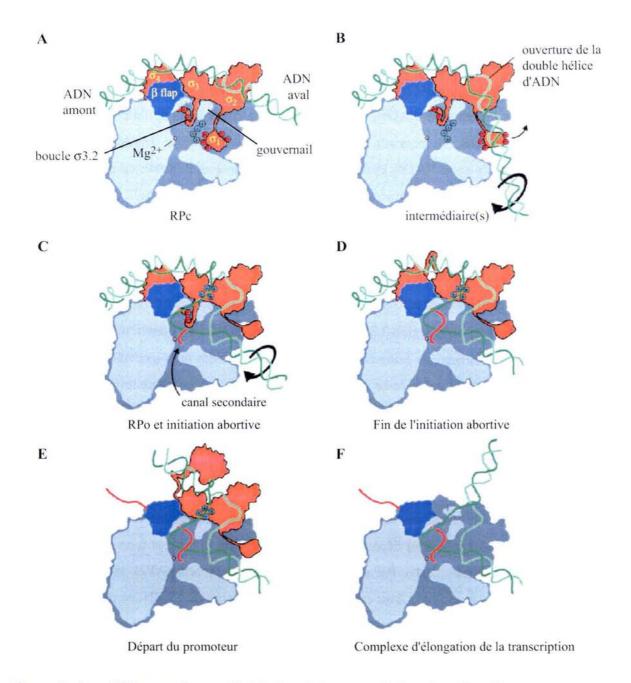


Figure 8 : Les différentes étapes d'initiation de la transcription chez E. coli (d'après Murakami et Darst, 2003).

La réprésentation montre une vue coupée de l'ARN polymérase, délétée d'une partie de la sous-unité  $\beta$ . L'ARN polymérase est indiquée en gris clair et gris foncé, son domaine flexible " $\beta$  flap" contactant  $\sigma_4$  est coloré en bleu. Les quatre domaines du facteur  $\sigma$  sont en orange.

A et B: L'ARN polymérase, sous forme holoenzyme, se fixe au niveau d'un promoteur de la double hélice d'ADN, les brins matrices et non matrices sont en vert foncé et vert clair, respectivement. Le promoteur est contacté par les domaines  $\sigma_2$  et  $\sigma_4$ . L'ouverture de l'ADN est réalisée par la sous-région 2.3. L'ouverture de la double hélice est associée à une torsion de l'ADN, au niveau du site d'entrée de l'ADN, ce qui déplace ainsi la région 1.1 du facteur  $\sigma$ .

C et D : Le brin matrice est localisé à proximité du site actif contenant l'ion magnésium catalytique indiqué en violet. Dans le complexe RPo, l'ADN est ouvert jusqu'au site d'initiation de la transcription, ce qui permet le début de la synthèse d'ARN, représenté en rouge. Les nucléotides empruntent le canal secondaire pour atteindre le site actif de l'ARN polymérase. La présence de la région 3.2 au niveau du canal de sortie provoque l'arrêt et l'éjection de courts transcrits, dits avortés. Dès que les interactions ARN naissant/ADN sont suffisamment importantes, l'ARN naissant déplace cette région 3.2.

E : La région  $\sigma_4$ , liée à la région 3.2, est décrochée de la région flexible " $\beta$  flap" de  $\beta$ .

F : Le facteur σ est dissocié dans la majorité des cas, au niveau du complexe d'élongation.

sous-régions 2.3 et 2.4 contactent l'ADN au niveau de la boîte -10 et initient l'ouverture de la double hélice d'ADN pendant l'état intermédiaire (Fig. 8B). Cette ouverture est facilitée par l'appariement plus ou moins lâche de la séquence d'ADN, riche en A/T. L'ouverture de la double hélice d'ADN de la région -11 à +2 environ, se poursuit jusqu'au site d'initiation de la transcription, grâce à l'établissement de contacts entre l'ADN aval et l'apoenzyme (Young et al., 2004), ce qui conduit à la formation du complexe RPo (RNA polymerase / promoter open complex) (Fig. 8C). L'équilibre de ou des états intermédiaires vers le complexe RPo dépend de nombreux facteurs, tels que la température et la concentration en ions Mg<sup>2+</sup> (pour revue, deHaseth et al., 1998). Les deux brins de la double hélice d'ADN sont alors localisés à deux endroits différents au niveau de l'ARN polymérase (Fig. 5D et 8C). Le brin matrice est situé à proximité du site actif de l'ARN polymérase, alors que le brin non matrice interagit avec les résidus aromatiques de la région 2.3 du facteur σ. Cette interaction maintient l'ouverture de la double hélice d'ADN. L'holoenzyme initie ensuite la polymérisation des premiers ribonucléotides, qui accèdent au site actif de l'ARN polymérase via le canal secondaire. Le complexe tertiaire comprenant l'ARN polymérase, l'ADN et l'ARN naissant est appelé complexe d'initiation (RPinit) (Fig. 8D). L'ARN polymérase synthétise de nombreux transcripts de 2 à 9 nucléotides, appelés transcrits abortifs, avant de pouvoir quitter le promoteur (Fig. 8E). La formation de ces transcrits abortifs semble résulter de la présence de la région 3.2 dans le canal de sortie de l'ARN naissant. Ainsi, si l'interaction ARN/ADN/ARN polymérase n'est pas assez forte, l'ARN naissant ne peut pas déplacer la région 3.2 du canal de sortie, ce qui conduit au départ de l'ARN naissant du complexe (pour revue, Borukhov et Severinov, 2002; Murakami et Darst, 2003). Par contre, dès que les interactions ARN naissant/ADN sont suffisamment importantes, la région 3.2 est expulsée du canal de sortie de l'ARN, ce qui entraîne aussi le décrochement de la région σ<sub>4</sub> du domaine flexible «β flap» de l'ARN polymérase (pour revue, Murakami et Darst, 2003). L'ARN polymérase quitte le promoteur et entre dans la phase d'élongation (Fig. 8, E et F). Dans la majorité des cas, l'ARN polymérase se détache du facteur σ (Bar-Nahum et Nudler, 2001).

L'efficacité de la transition d'une étape de l'initiation à une autre est fonction de chaque promoteur, et peut être définie par une constante cinétique. Cependant, la première étape de fixation de l'ARN polymérase à l'ADN pour former le complexe RPc est souvent réversible, tandis que les étapes suivantes le sont moins. L'efficacité des différentes étapes est fonction de chaque promoteur. L'efficacité globale de chaque promoteur résulte de la combinaison de chacune de ces étapes, la plus lente devenant l'étape limitante de l'efficacité de l'initiation de la transcription pour un promoteur donné (pour revue, Rojo, 2001).

# II / La régulation de l'initiation de la transcription

E. coli contient environ 2 000 molécules d'ARN polymérase sous forme d'apoenzyme, dont 1 300 sont impliquées dans la transcription, pour exprimer environ 4 000 gènes (pour revue, Ishihama, 2000). L'ARN polymérase est donc en quantité limitée dans la cellule. De plus, elle synthétise autant d'ARNm que d'ARNr et d'ARNt utilisés par la machinerie de traduction pour la production de protéines (pour revue, Kennell, 2002). Ainsi, différents mécanismes moléculaires sont mis en jeu pour une utilisation optimale des ARN polymérases en favorisant ou inhibant l'initiation de la transcription (pour revue, Browning et Busby, 2004). Chaque étape de l'initiation de la transcription peut aussi être activée ou inhibée. En considérant chaque étape d'un point de vue thermodynamique, les activateurs ou les répresseurs diminueront ou augmenteront, respectivement, l'énergie d'activation nécessaire à chaque étape (pour revue, Rojo, 2001).

Nous allons décrire les différents paramètres permettant une régulation fine de l'initiation de la transcription des gènes, comprenant l'influence d'une courbure de l'ADN, le choix du facteur σ intervenant dans la reconnaissance des éléments -35 et -10 du promoteur, les ligands influençant la stabilité du complexe ouvert, les facteurs de transcription jouant un rôle dans le recrutement ou l'inhibition de l'ARN polymérase.

#### II.1 / Influence d'une courbure de l'ADN au niveau d'un promoteur

Au niveau d'un promoteur, la présence d'une courbure de l'ADN influence la formation des complexes RPc et RPo. En effet, l'initiation de la transcription nécessite une torsion de l'ADN assez importante qui peut être facilitée ou défavorisée par une courbure. La formation de cette dernière résulte soit d'une séquence particulière de l'ADN, soit de l'interaction de l'ADN avec une protéine régulatrice. Ainsi, l'ajout d'une région courbée en amont d'un promoteur artificiel peut modifier jusqu'à un facteur 10 l'efficacité d'initiation de la transcription (Collis et al., 1989). Les protéines HU et IHF, de structuration du nucléoïde, s'autorégulent en modifiant la courbure au niveau du promoteur de leurs propre gènes (Jauregui et al., 2003).

La recherche bioinformatique d'éléments de courbures au niveau de différents génomes bactériens montre que l'activation transcriptionnelle par ce mécanisme pour différents gènes est un trait commun pour les bactéries même phylogénétiquement éloignées (Jauregui et al., 2003). Cette recherche bioinformatique a permis de mettre en évidence des séquences courbées au niveau de promoteurs de gènes de fonctions complètement différentes, à priori, tels que les gènes de transposase, de l'ADN gyrase, de la glucose-6-phosphate isomérase et de la biosynthèse des flagelles (Jauregui et al., 2003).

Au niveau des gènes fortement exprimés tels que les ARNr et les aminoacyl-ARNt synthétases, la courbure semble faciliter la formation du complexe ouvert. Cette courbure est centrée au niveau d'une série d'adénines située en amont, qui correspond aussi à l'élément UP. Ainsi, l'élément UP constitue à la fois un site d'interaction avec le domaine  $\alpha$ -CTD de l'ARN polymérase et un élément assurant la formation d'une courbure de l'ADN.

#### II.2 / Influence des éléments en trans du promoteur

#### II.2.a / Choix du facteur sigma

La quantité des sept facteurs  $\sigma$  est supérieure par rapport à la quantité d'ARN polymérase libre (1 200 facteurs  $\sigma$  pour 600 à 700 molécules d'apoenzymes libres par bactérie) (pour revue, Ishihama, 2000). Le taux des différentes formes d'holoenzymes dépend de la concentration intracellulaire en facteur  $\sigma$  et de son affinité pour l'ARN polymérase sous sa forme apoenzyme. Différents paramètres modifient l'affinité des facteurs  $\sigma$  pour l'apoenzyme. Par exemple, Jishage *et al.* (2002) ont montré que la compétition pour l'apoenzyme entre les différents facteurs  $\sigma$  est régulée par le nucléotide ppGpp que nous allons décrire plus en détails dans la partie suivante (Jishage *et al.*, 2002). Le ppGpp favorise alors la fixation des facteurs  $\sigma$  et  $\sigma$  au détriment de  $\sigma$  . A ce modèle se rajoute d'autres régulations via les anti-facteurs sigma.

#### II.2.b / Influence des nucléotides au niveau du complexe ouvert

Lorsque les promoteurs possèdent une séquence riche en GC entre la boîte -10 et le site d'initiation de la transcription, région appelée discriminateur, tel que les promoteurs des ARNr, les complexes RPo présentent une faible stabilité. La stabilité des complexes RPo peut être influencée par le taux de deux types de nucléotides, le nucléotide initiateur et le ppGpp. Ces molécules de petite taille permettent une réponse rapide et efficace aux variations de l'environnement (Gaal et al., 1997; pour revue, Browning et Busby, 2004).

Le premier nucléotide utilisé pour la synthèse de l'ARN, correspondant au nucléotide +1, appelé nucléotide initiateur (iNTP), entraîne une stabilisation du complexe ouvert. Ainsi, il a pu être montré in vitro et in vivo, que l'augmentation de la concentration de iNTP augmente l'efficacité d'initiation de la transcription au niveau des promoteurs rrnB avec l'ATP, et rrnD avec le GTP (Gaal et al., 1997, Schneider et al., 2002). In vivo, la variation du pool de nucléotides initiateurs a préalablement été corrélée avec la vitesse de croissance jusqu'à ce que la méthode de mesure du pool de nucléotide

initiateur soit remise en question (Petersen et Moller, 2000, Schneider et Gourse, 2004). Ainsi, la variation de la concentrations de iNTP ne permettrait pas une variation de l'expression des opérons ribosomiques en fonction du taux de croissance. Par contre, l'augmentation de la concentration de iNTP ne serait effective qu'au début de la phase de croissance exponentielle et permettrait une activation de l'expression des ARNr et ARNt (Murray et al., 2003).

La déstabilisation du complexe ouvert est entraîné par le (p)ppGpp, synthétisé entre autre lors d'une carence en acides aminés (pour revue, Cashel et al., 1996). Différents changements physiologiques, appelés réponse stringente, sont alors mis en place afin de répondre à la demande de biosynthèse des protéines, malgré un déficit du pool d'acides aminés, tels qu'une diminution de la synthèse de la majorité des protéines et d'ARN et une augmentation de la production de protéines essentielles à la survie cellulaire comme σ<sup>S</sup> (pour revue, Cashel et al., 1996). Ces variations résultent de l'accumulation de dérivés de GTP ou GDP, portant un groupe pyrophosphate sur le ribose, en 3' hydroxyle, formant ainsi le pppGpp et ppGpp, nucléotides appelés collectivement ppGpp (pour revue, Condon et al., 1995).

Le taux de base de ppGpp varie inversement avec le taux de croissance (pour revue, Gralla, 2005). Il est synthétisé principalement par la protéine RelA, de 84 kDa, associée au ribosome. Son activité est stimulée par l'arrêt d'un ribosome en cours d'élongation, due à la présence au site A du ribosome d'un ARNt ne portant pas d'acides aminés (voir partie C) (Hogg et al., 2004). Il est aussi synthétisé par l'enzyme cytosolique bifonctionnelle SpoT de 79 kDa. En effet, SpoT possède une faible activité synthase mais majoritairement une activité hydrolase du ppGpp (Mechold et al., 2002). Ainsi, l'augmentation de la concentration en ppGpp résulterait de la diminution de l'activité hydrolase, lors d'une carence carbonée brusque (pour revue, Condon et al., 1995), ou d'une augmentation de l'activité synthase, lors d'une carence en phosphate inorganique ou en phospholipides (Spira et Yagil, 1998). La concentration en ppGpp augmente aussi lorsque la fructose-1.6-bisphosphate aldolase est inactivée (Schneider et Gourse, 2003), ou lors de l'inhibition de l'entrée du glucose par l'α-methyl-glucoside (Cashel et al., 1996; Murray et al., 2003). Cette augmentation pourrait être due à l'interaction possible entre la protéine EIIA<sup>Gle</sup> déphosphorylée et la protéine RelA (Sondej et al., 1999).

La molécule de ppGpp diminue l'efficacité d'initiation de la transcription en se liant au niveau d'un site adjacent au site catalytique de l'ARN polymérase (Artsimovitch et al., 2004). Ceci conduit à la diminution de stabilité des complexes ouverts, et donc à l'inhibition de la transcription lorsque celle-ci est limitée au niveau de l'étape de la formation des complexes ouverts. Ainsi, les gènes codant les produits nécessaires pour la traduction possèdent un discriminateur riche en GC, tels que les ARNr, et voient leur expression diminuer en présence de ppGpp.

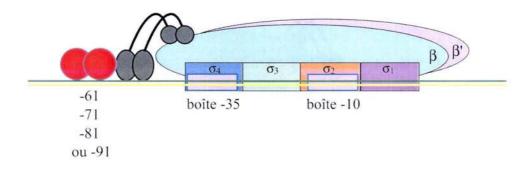
#### II.2.c / Influence des facteurs de transcription

Le rôle des facteurs de transcription est de coupler l'expression des gènes en fonction des conditions de l'environnement (pour revue, Browning et Busby, 2004). Leur activité est régulée via la fixation de ligands ou une modification covalente telle qu'une phosphorylation pour les régulateur de réponse des systèmes à deux composants, soit par une variation de leur concentration, obtenue par un contrôle de leur expression ou de leur dégradation.

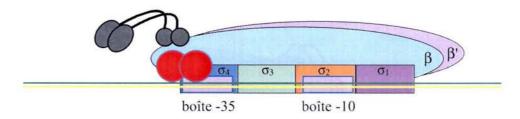
E. coli possède environ 300 gènes qui codant des facteurs de transcription prédits ou connus, sur les 4405 ORFs identifiées (pour revue, Martinez-Antonio et Collado-Vides, 2003). Ces facteurs de transcription possèdent différents niveaux de contrôle. En effet, sur les 150 facteurs de transcription dont l'activité a été confirmée expérimentalement, environ 60 facteurs contrôlent l'expression d'un seul promoteur tandis que 7 facteurs modulent l'expression de 51% des gènes (pour revues, Martinez-Antonio et Collado-Vides, 2003, Browning et Busby, 2004). Ces 7 facteurs, comprenant les facteurs CRP, IHF, FNR, FIS, ArcA, Lrp et HNS, sont appelés des régulateurs globaux, puisque leur inactivation entraîne des phénotypes pléiotropes, et régulent des opérons appartenant à différentes voies métaboliques. La protéine CRP s'associe à l'AMPc, dont la concentration dépend de l'état énergétique de la cellule, pour réguler l'expression des gènes. Les protéines FNR et ArcA sont liées à la production d'énergie, en fonction de l'état redox de la cellule. Le facteur Lrp ajuste le métabolisme général en fonction de la concentration en L-leucine cellulaire. Les facteurs FIS, IHF et HNS sont des protéines de liaison à l'ADN et modulent sa topologie en fonction de l'état énergétique de la cellule (pour revue, Martinez-Antonio et Collado-Vides, 2003).

Un facteur de transcription bactérien peut avoir soit un rôle d'activateur, soit de répresseur, soit exercer les deux fonctions, suivant le promoteur régulé. L'activité du facteur de transcription dépend de la localisation du site de fixation sur la région du promoteur. Lors d'une activation, le facteur de transcription contacte l'holoenzyme seule ou au niveau de la région promotrice. Trois mécanismes généraux d'activation sont possibles (Fig. 9). Dans le mécanisme d'activation de classe I, les activateurs se lient en amont du promoteur, à une position centrée au niveau des positions -61, -71, -81 ou -91 par rapport au site +1 d'initiation de la transcription (Fig. 9A). Ils interagissent avec un domaine  $\alpha$ -CTD de l'ARN polymérase, ce qui permet de recruter l'ARN polymérase au niveau du promoteur (pour revue, Busby et Ebright, 1994). La variation de la position du site de fixation du facteur de transcription est liée à la flexibilité de la région reliant les domaines  $\alpha$ -CTD aux domaines  $\alpha$ -NTD. Dans le mécanisme d'activation de classe II, le facteur de transcription se lie au niveau de la boîte -35 du promoteur, et interagit directement avec l'ARN polymérase, en général, le domaine 4 du facteur  $\sigma$ , quelques fois, le domaine  $\alpha$ -CTD (Fig. 9B) (pour revues, Rhodius et Busby, 1998; Dove *et al.*, 2003). Dans le dernier mécanisme d'activation, le facteur de transcription module la conformation de l'ADN afin

#### A: Mécanisme d'activation de classe I



#### B: Mécanisme d'activation de classe II



#### C: Mécanisme d'activation par modification de la conformation de l'ADN

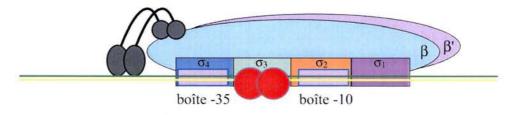


Figure 9 : Mécanisme d'activation de la transcription par un seul facteur de transcription (d'après Browning et Busby, 2004).

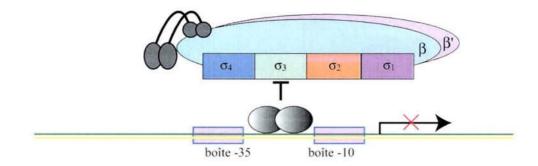
A : L'activateur de transcription de classe I recrute l'ARN polymérase en interagissant avec un domaine  $\alpha$ -CTD. Il est fixé à l'ADN au niveau des sites centrées aux positions -61, -71, -81 ou -91 par rapport au site d'initiation de la transcription.

B : L'activateur de transcription de classe II, localisé au niveau de la boîte -35 de l'ADN, interagit avec la région 4 du facteur  $\sigma$ , en général.

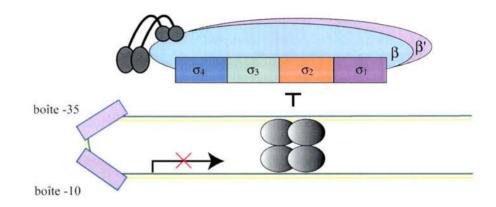
C : La fixation de l'activateur à proximité ou au niveau du promoteur, modifie la conformation de l'ADN, ce qui repositionne correctement les boîtes -35 et -10 l'une par rapport à l'autre, et facilite l'interaction avec l'ARN polymérase.

Les mêmes représentations de l'holoenzyme et de l'ADN que la figure 5C sont utilisées ici. L'activateur transcriptionnel est représenté par deux ronds rouges.

#### A: Mécanisme de répression par empêchement stérique



#### B: Mécanisme de répression par limitation de l'accès du promoteur



#### C: Mécanisme de répression par modification de la conformation de l'ADN

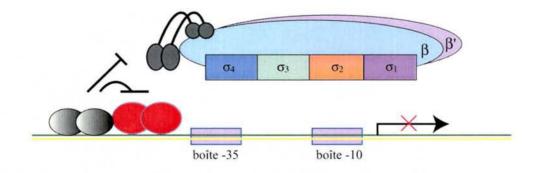


Figure 10 : Mécanisme de répression de la transcription par un seul facteur de transcription (d'après Browning et Busby, 2004).

A : Le répresseur est fixé au niveau du promoteur, ce qui inhibe la fixation de l'ARN polymérase.

B : Le répresseur se lie à des sites distaux du promoteur, et forment une boucle, ce qui limite l'accès de l'ARN polymérase au promoteur.

C : Le répresseur inhibe l'effet d'un activateur de la transcription.

Les mêmes représentations pour l'holoenzyme et l'ADN que la figure 5C sont utilisées ici. L'activateur transcriptionnel est représenté par deux ronds rouges, alors que le répresseur transcriptionnels est indiqué par deux ronds gris.

de faciliter le recrutement de l'ARN polymérase (pour revue, Browning et Busby, 2004). Par exemple, les promoteurs activés par les facteurs de type MerR possèdent des boîtes -35 et -10 séparés par une région interboîte dont la longueur n'est pas optimale. La fixation de l'activateur en présence de mercure modifie la conformation de l'ADN, ce qui permet un positionnement correct des boîtes -10 et -35, et donc le recrutement de l'ARN polymérase (Fig. 9C) (pour revue, Brown et al., 2003).

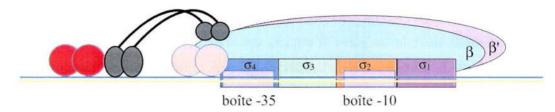
Les facteurs de transcription inhibent la transcription soit de manière stérique, en se fixant directement au niveau du promoteur, ce qui inhibe le recrutement de l'ARN polymérase, soit en limitant l'accessibilité du promoteur par la formation d'une boucle au niveau de l'ADN, soit en inhibant l'activation d'un autre facteur de transcription (Fig. 10).

#### II.2.d / Intégration du signal

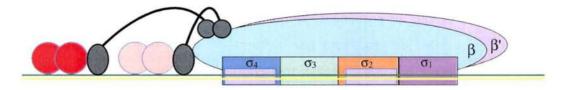
L'activité de la plupart des promoteurs bactériens dépend de plusieurs signaux. Ainsi, différents facteurs de transcription répondant à différents signaux de l'environnement, modulent l'utilisation des promoteurs, ce qui permet d'obtenir une régulation fine de l'expression des gènes. L'intégration des signaux est réalisée par la combinaison d'activateurs et de répresseurs, qui agissent indépendamment ou interagissent directement entre eux.

Quatre types d'intégration conduisant à une activation ont pu être différenciés (Fig. 11) (pour revue, Browning et Busby, 2004). Dans un premier cas, deux activateurs peuvent entrer en contacts avec l'ARN polymérase indépendamment ou de façon coopératifs, ce qui améliore l'affinité de l'ARN polymérase pour le promoteur (Fig. 11, A, B et C). Ainsi, dans le cas du promoteur proP P2 d'E. coli, l'activation de la transcription est réalisée de façon indépendante, par un mécanisme de classe I et de classe II, par la protéine Fis et le complexe CRP/AMPc situés en position -121 et -41, respectivement (Fig. 11A) (McLeod et al., 2002). Les activateurs peuvent interagir avec l'ARN polymérase par deux mécanismes de classe I, comme dans le cas du promoteur promoteur acsP2 activé par un tandem complexe CRP/AMPc (Tebbutt et al., 2002). Dans un deuxième cas, les activateurs peuvent agir en succession, la présence d'un second activateur permet le repositionnement correct du premier activateur fixé à l'ADN, afin que ce dernier puisse recruter l'ARN polymérase (Fig. 11B). Dans le cas du promoteur malK, le complexe CRP/AMPc repositionne l'activateur MalT (Richet et al., 1991). Enfin, un second activateur peut inhiber l'action répresseur d'une protéine du nucléoïde sur un premier activateur (Fig. 11D). Les promoteurs des gènes nir et nrf d'E. coli sont activés par le facteur de transcription FNR par un mécanisme de classe II. La fixation des protéines du nucléoïde IHF ou Fis supprime cette activation. L'inhibition apparente ainsi engendrée est levée par la présence des facteurs transcriptionnels NarL et NarP, régulateurs de réponse des systèmes à deux composants qui répondent à la concentration en nitrites et en nitrates.

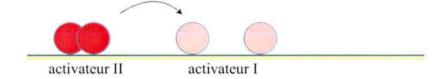
A: Mécanisme d'activation par des activateurs de façon indépendente activateurs de classe I/classe II



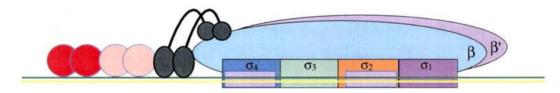
activateurs de classe I/classe I



**B**: Mécanisme d'activation coopérative



C: Mécanisme d'activation par liaison coopérative



D: Mécanisme d'activation par inhibition de la répression

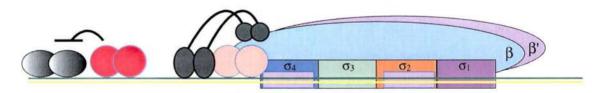


Figure 11 : Régulation de la transcription par plusieurs facteurs transcriptionnels (d'après Browning et Busby, 2004).

A : Deux activateurs peuvent agir de façon indépendante, soit par un mécanisme de classe I et de classe II, soit par deux mécanismes de classe I.

B : L'activateur II repositionne correctement l'activateur I, qui peut alors activer l'initiation de la transcription.

C : Les deux activateurs agissent en synergie.

D : Un des deux activateurs de l'initiation de la transcription inhibe l'effet d'un répresseur.

Les mêmes représentations pour l'holoenzyme et l'ADN de la figure 5C sont utilisées ici. Les activateurs transcriptionnels sont représentés par deux ronds rouges ou oranges, alors que le répresseur transcriptionnel est indiqué par deux ronds gris.

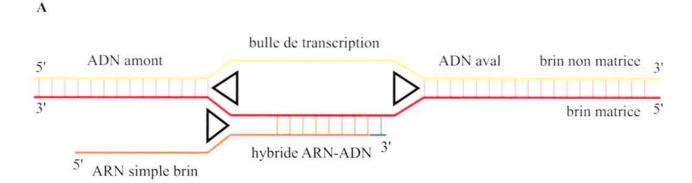
# III / Mécanisme d'élongation et de terminaison de la transcription

Lorsque la taille de l'ARN naissant dépasse 11 à 13 nucléotides, l'ARN polymérase se sépare de la région promotrice et entre en **phase d'élongation**. Cette phase est caractérisée par la formation d'un complexe de la transcription ternaire stable avec l'ADN et l'ARN et l'ARN polymérase jusqu'à la terminaison. Ceci permet la transcription d'opérons entiers, dont la taille peut atteindre 20 000 bp chez *E. coli* (pour revue, Richardson et Greenblatt, 1996). La modélisation du complexe d'élongation montre que l'ADN est ouvert sur 12 à 14 nucléotides, formant ainsi la bulle de transcription (Fig. 12) (pour revue, Darst, 2001; von Hippel et Pasman, 2002). Trois sites de liaison de l'ARN polymérase aux acides nucléiques assurent la stabilité de ce complexe d'élongation: un premier site contacte l'ADN amont sur environ 9 paires de bases, un deuxième site contacte l'hybride ARN/ADN d'environ 8 bp, et le troisième site se lie à l'ARN lorsqu'il est séparé de l'ADN (pour revue, Nudler et Gottesman, 2002).

Chaque nouveau nucléotide s'apparie avec le nucléotide du brin matrice, puis est lié au niveau de l'extrémité 3' de l'ARN naissant, situé en position -1 par rapport au nouveau nucléotide (+1) (pour revue, Nudler, 1999). Cette réaction est accompagnée du départ du groupement pyrophosphate. Ainsi, la réaction réversible d'élongation est :

Deux mécanismes de déplacement de l'ARN polymérase sont proposés. L'ARN polymérase se déplace soit pour chaque nucléotide ajouté, soit par une succession de sauts jusqu'au site de terminaison.

L'arrêt de l'élongation est un facteur essentiel pour permettre une régulation correcte de l'expression des gènes. Chez *E. coli*, deux mécanismes de terminaison de la transcription ont été identifiés, dits rho-indépendant ou rho-dépendant, réalisant à part égale la terminaison dans la cellule (pour revue, Richardson et Greenblatt, 1996). Dans le mécanisme de terminaison rho-indépendante ou terminaison intrinsèque, la séquence de l'ADN suffit pour décrocher l'ARN polymérase. Cette séquence conduit à la synthèse d'un ARN formant une structure tige-boucle stable contenant environ 20 paires de bases riches en G-C, suivie par une séquence d'envivon 8 nucléotides, riche en U, simple brin (Fig. 13) (d'Aubenton Carafa *et al.*, 1990; pour revue, Richardson, 2002). La formation de l'hybride rU-dA correspond à l'hybride ayant une stabilité thermodynamique la moins stable, et entraîne une pause de l'ARN polymérase (pour revue, Nudler et Gottesman, 2002). Cette pause facilite la formation de la structure tige-boucle de l'ARN, en amont, à proximité de l'hybride ADN/ARN, ce qui conduit à une déstabilisation du complexe d'élongation. La déstabilisation des interactions au sein de l'hybride ADN/ARN et entre l'ARN naissant et l'ARN polymérase, mais aussi le réappariement des deux brins d'ADN, conduisent au décrochage de l'ARN naissant (Ryder et Roberts, 2003). La terminaison intrinsèque



B

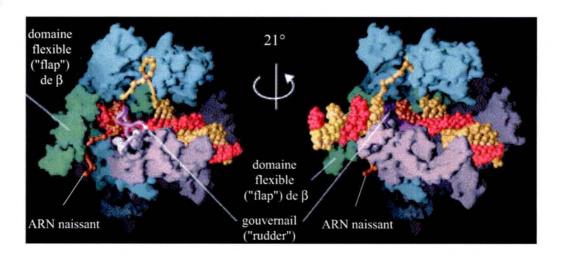


Figure 12 : Etape d'élongation de la transcription (d'après Darst, 2001).

A : Représentation schématique du complexe d'élongation. Les brins d'ADN matrice et non matrice sont indiqués sous forme de traits rouge et jaune, respectivement. L'ARN naissant est indiqué en orange, le nouveau nucléotide entrant en vert. Les triangles noirs symbolisent le contrôle de la longueur de la bulle de transcription, et celle de l'hybride ADN-ARN par l'ARN polymérase.

B : Modèle du complexe d'élongation. Ce complexe comprend la structure de l'apoenzyme de *Thermus aquaticus*, constitué des sous-unités  $\beta$  (cyan),  $\beta$ ' (rose),  $\alpha$  (blanc) et  $\omega$  (blanc). Les acides nucléiques ont le même code couleur qu'en A. L'ARN naissant quitte l'ARN polymérase à proximité du domaine flexible " $\beta$  flap" de  $\beta$ , et serait séparé de l'ADN par le gouvernail ("rudder").

Site de liaison
faible à l'ARN
simple brin

site de liaison
à l'ARN simple brin

site de liaison
l'Hybride ARN/ADN
l'Hybride ARN/ADN
l'Hybride ARN/ADN
l'Hybride ARN/ADN

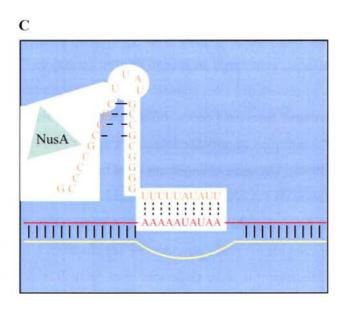


Figure 13 : Mécanisme de terminaison intrinsèque (pour revue, Nudler et Gottesman, 2002).

A: La terminaison débute par une pause induite par la séquence de U.

B : La pause favorise la formation de la structure tige-boucle située en amont de la séquence de U, et diminue les interactions à la fois au niveau de l'hybride ARN/ADN et au niveau de l'ARN naissant simple brin avec l'ARN polymérase, ce qui facilite le décrochage de l'ARN naissant.

C : La terminaison intrinsèque est favorisée par la protéine NusA. Cette protéine semble affaiblir les interactions de l'ARN avec les sites de faible affinité et stabiliser la pause nécessaire à la formation de la structure tige-boucle.

est aussi régulée par des facteurs additionnels. Par exemple, la protéine NusA stimule la terminaison en favorisant la formation de la structure tige-boucle de l'ARN (pour revue, Nudler et Gottesman, 2002).

La dissociation de l'ARN naissant de l'ARN polymérase peut être réalisée par l'intermédiaire de l'hélicase Rho, au niveau de terminateurs dits Rho-dépendants (pour revue, Richardson, 2002). Ce facteur Rho provoque la terminaison à la fin de la transcription de gènes cibles ou au niveau des atténuateurs situés en amont des régions codantes et à l'intérieur d'un transcrit si celui-ci n'est pas normalement traduit (pour revue, Richardson, 2003). Le facteur Rho est un hexamère de six sous-unités identiques (Fig. 14). Il fixe l'ARN au niveau de deux types de sites de liaison, appelés sites primaires et secondaires, puis consomme l'ATP pour se déplacer le long du transcrit à la poursuite de l'ARN polymérase, et désapparier le transcrit de l'ADN et de l'ARN polymérase. Les sites primaires sont localisés au niveau des domaines N-terminaux des monomères (Fig. 14, B et C). Ils se lient avec une forte affinité, avec l'ARN ou l'ADN simple brin au niveau des séquences rut (Rho utilization), riches en pyrimidine, via une interaction par groupe de dinucléotide. L'ARN s'enroule sur 70 à 80 bases à la périphérie de l'anneau autour des sites primaires (Fig. 14C) (Skordalakes et Berger, 2003). Les sites secondaires, de plus faible affinité, sont situés au niveau du trou central de l'hexamère (Fig. 14B). La fixation d'acides nucléiques au niveau des sites primaires entraîne un réarrangement conformationnel du facteur Rho. Ce réarrangement favorise l'accessibilité des sites de liaisons aux nucléotides pour l'activité ATPase, qui est d'autant plus importante lorsqu'il s'agit d'ARN fixé sur les sites primaires au lieu d'ADN (Jeong et al., 2004b). Un autre changement conformationnel se produit après la fixation d'ATP, ce qui permet la fixation de l'ARN au niveau des sites secondaires (Jeong et al., 2004b), puis la fermeture de l'anneau et la formation du site catalytique ATPase (Fig. 14C) (pour revue, Kaplan et O'Donnell, 2003). Deux modèles ont été proposés pour décrire ces réactions (pour revue, Nudler et Gottesman, 2002). Dans le premier, la séquence rut reste fixée aux sites primaires du facteur Rho pendant le déplacement du facteur Rho le long du transcrit, tandis que dans le deuxième cas, le facteur Rho est détaché de la séquence rut. La dissociation du complexe d'élongation est réalisée au niveau des pauses de l'ARN polymérase. Ainsi, les différents sites de terminaison Rho-dependant s'étendent sur environ 100 paires de bases en aval du site rut, au niveau de séries de cinq à huit nucléotides espacés de 20 à 30 nucléotides.

La fixation du facteur Rho au transcrit est inhibée par la traduction : les ribosomes limitent l'accessibilité des séquences rut par inhibition stérique (pour revue, Nudler et Gottesman, 2002). Ceci explique le principe de polarité, dans lequel une mutation non-sens dans un gène diminue la transcription des gènes situés en aval, appartenant au même opéron. Dans le cas des ARNr et ARNt non traduits, la fixation du facteur Rho est limitée par la structuration des ARN, mais aussi par des facteurs dits d'antiterminaison. D'autres facteurs facilitent la fixation de Rho à l'ARN, tel que le

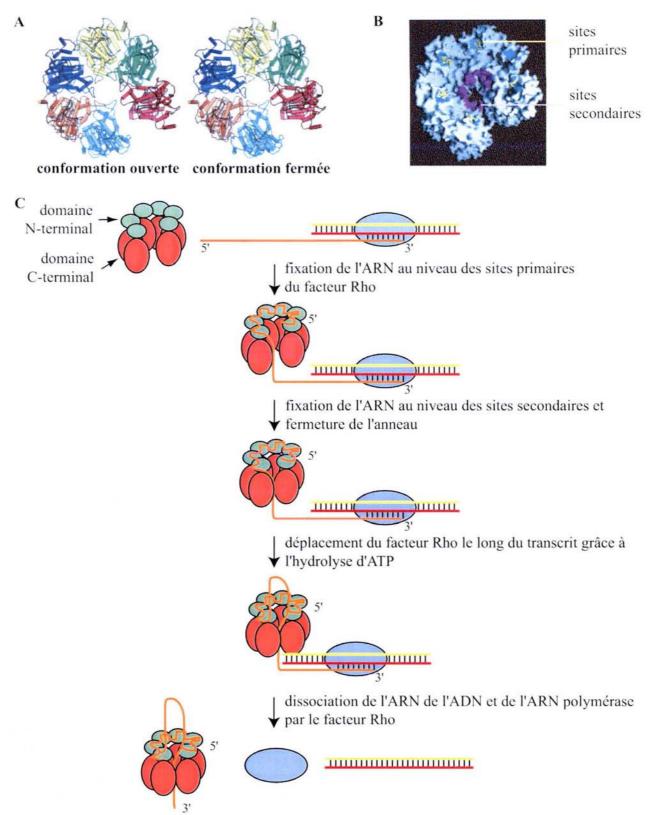
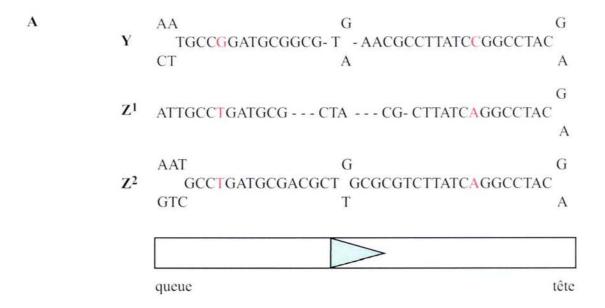


Figure 14 : Structure et mécanisme d'action du facteur Rho

(d'après Skordalakes et al., 2003 ; Kaplan et O'donnell, 2003).

- A : Représentation schématique du facteur Rho dans les conformations ouvertes et fermées, respectivement. Les six monomères sont indiqués par différentes couleurs.
- B : Représentation de la surface moléculaire du facteur Rho. Les acides nucléiques, interagissant avec les sites primaires indiqués en bleu, sont représentés par des batons jaunes . Les sites secondaires sont colorés en violet.
- C : Mécanisme d'action du facteur Rho. Les domaines N-terminaux et C-terminaux du facteur Rho sont schématisés en vert et rouge, respectivement. L'ARN polymérase est représentée par un disque bleu, les brins matrices et non matrices de l'ADN et l'ARN sont symbolisés par un trait rouge, jaune et orange, respectivement.

facteur NusG, qui contacte à la fois l'ARN polymérase et le facteur Rho (pour revue, Nudler et Gottesman, 2002). Récemment, la terminaison Rho-dépendante a été associée aux répétitions palindromiques de type BIME (Bacterial Interspersed Mosaic Elements) (Espeli et al., 2001). Chez les bactéries, les répétitions palindromiques sont constituées soit d'une simple répétition, comme pour les éléments RSA (repeat sequence) et ERIC (Enterobacterial Repetitive Extragenic Consensus), soit de plusieurs répétitions identiques, comme pour les répétitions de 29-bp « iap », ou différentes, comme pour les éléments boxC et BIMEs (pour revue, Bachellier et al., 1999). Les BIMEs comprennent les palindromes imparfaits de 40 nucléotides Y, Z<sup>1</sup> et Z<sup>2</sup> (Fig. 15A). Les palindromes de type Y diffèrent des palindromes Z<sup>1</sup> et Z<sup>2</sup> par la présence de résidus G et C ou T et A en position 7 et 32, les palindromes Z<sup>2</sup> sont plus grands que les Z<sup>1</sup>. Une orientation a été établie pour les palindromes Y, Z<sup>1</sup> et Z<sup>2</sup>, ce qui permet de définir alors une « tête » et une « queue ». Les BIMEs sont constituées de un à six binômes de palindromes orientés « tête à tête », avec les éléments suivants : L, site de fixation de la protéine IHF, ou S sont localisés entre deux «têtes», l, s et r au niveaux des «queues». Les éléments A et B peuvent être associés aux éléments «queue» de chaque extrémité d'un BIME. Suivant les différents palindromes présents dans les BIMEs, deux familles ont pu être définies : les familles de type BIME-1 contiennent les palindromes Y et Z<sup>1</sup> séparés de l'élément L, avec le motif A à proximité de Z<sup>1</sup> et B de Y (Fig. 15C). Les familles de type BIME-2 comprend les palindromes Y et Z<sup>2</sup>, entourant l'élément S, les éléments l, s ou r sont localisés au niveau des « queues ». Différents mécanismes permettent d'expliquer la terminaison au niveau de ces séquences BIME. Ils peuvent constituer des sites rut, bloquer la progression de l'ARN polymérase, ce qui permet au facteur Rho de rejoindre l'ARN polymérase. Enfin, la présence d'une pause pourrait permettre l'accès du facteur Rho à un site rut non identifié (Espeli et al., 2001).



Nom Taille Consensus L 32 à 34 bp YGRTTYCTGCAATATATTGAATTTGCAYGMTYT S 12 à 14 bp 1 8 à 9 bp YGGTGCA C S 1 bp 18 à 31 bp r 20 bp TTTGCGTTTGTCATCAGTCY A B 17 bp AACAAMGCGCACTTTGTCARCAATCTG

В

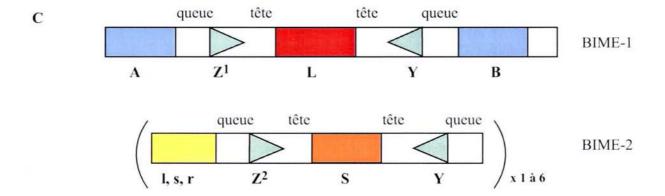


Figure 15 : Structure schématique des BIMES (d'après Bachellier *et al.*, 1999 et www.pasteur.fr/recherche/unites/pmtg/repet/intro.BIME.html).

A : Séquences et représentation schématique des palindromes retrouvés dans les BIMEs.

B : Séquences des éléments autres que les palindromes retrouvés dans les BIMEs. La séquence indiquée en rouge dans l'élément L est le site de fixation de la protéine IHF.

C: Représentation schématique des deux grandes familles de BIMEs retrouvées chez les bactéries. La famille de type BIME-1 contient les palindromes  $Z^1$  et Y entourant l'élément L, les éléments A et B sont à l'extrémité des palindromes  $Z^1$  et Y, respectivement. La famille de type BIME-2 est constituée de deux à douze palindromes inversés, où les palindromes Y et  $Z^2$ , entourent l'élément S, tandis que les éléments I, s et r sont à proximité des queues des palindromes.

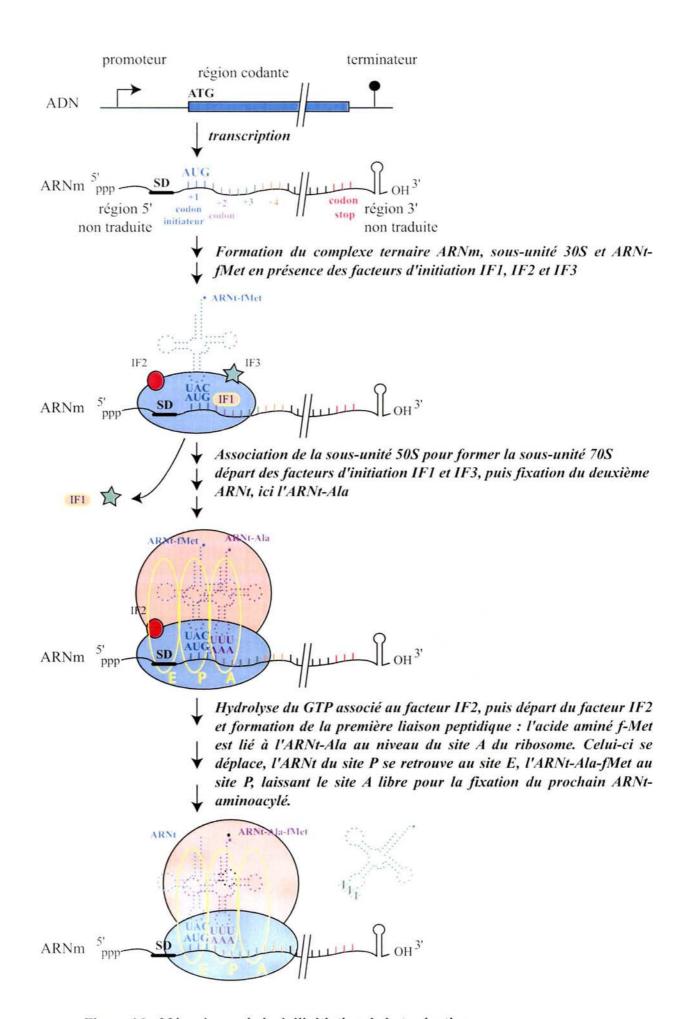


Figure 16 : Mécanisme général d'initiation de la traduction.

# Partie C / L'initiation de la traduction chez E. coli

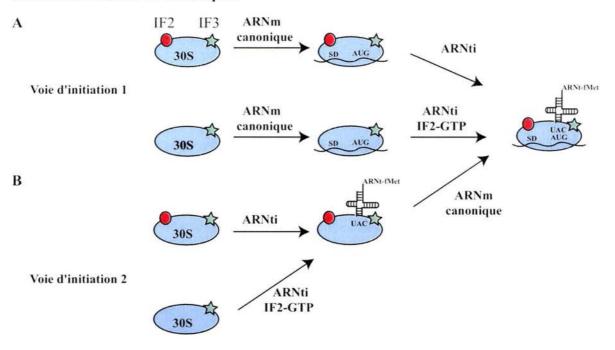
L'ARNm synthétisé au cours de la transcription est traduit ce qui permet la synthèse de la protéine correspondante. Nous allons décrire dans cette partie, les mécanismes d'initiation de la traduction et leur régulation. Dans la prochaine partie, nous étudierons la maturation et dégradation des ARNm et les régulations des activités de dégradation.

Chez les procaryotes, la traduction, réalisée par les ribosomes 70S, débute dès que les sites de fixation des ribosomes sur l'ARNm naissant sont synthétisés. La traduction, comme la transcription, peut se diviser en 3 étapes distinctes : l'initiation, l'élongation et la terminaison, pouvant chacune être le siège de contrôles régulant l'expression génique. Nous limiterons notre étude aux mécanismes de l'initiation de la traduction. Cette étape, généralement limitante, est le siège des principaux contrôles de la traduction (pour revue, Draper, 1996). Elle conduit à la formation d'un complexe, comprenant le ribosome, l'ARNm servant de matrice à la traduction, l'ARNt initiateur formyl-méthionine (ARNt-fMet), et trois protéines auxilliaires appelées facteurs d'initiation (IF1, IF2 et IF3). Le ribosome contient trois sites de fixation des ARNt : les sites A (aminoacyl), P (peptidyl) et E (exit) (Fig. 16). L'ARNt-fMet se fixe au niveau du site P via les interactions codon +1/anticodon avec le codon initiateur AUG. L'ARNt suivant, reconnu via les interactions codon +2/anticodon, est apporté par le facteur EF-Tu au niveau du site A. La formation de la première liaison peptidique entre les acides aminés portés par les deux premiers ARNt permet au ribosome d'entrer en phase d'élongation (Fig. 16).

#### I / Mécanisme d'initiation de la traduction

Le ribosome 70S résulte de l'assemblage de deux sous-unités, 30S et 50S, composées chacunes de protéines ribosomales et d'ARN ribosomiques, ARNr 16S pour la petite sous-unité, ARNr 23S et 5S pour la grande sous-unité. L'initiation de la traduction est principalement réalisée via le recrutement de la sous-unité 30S au niveau de la région 5' non traduite de l'ARNm, puis de la sous-unité 50S pour former le ribosome. La région de l'ARNm protégée par le ribosome, s'étendant des positions –35 à +20, est appelée TIR (Translation Initiation Region). En amont du codon initiateur, elle contient une séquence riche en purine, la séquence Shine-Dalgarno (SD) appelée aussi RBS (Ribosome Binding Site). Cette séquence SD s'associe au niveau de l'extrémité 3' terminale de l'ARNr 16S de la sous-unité 30S, séquence appelée anti-Shine-Dalgarno (aSD). Cette interaction a été démontrée entre autre par des expériences de mutagenèse, soulignant les rôles joués à la fois par l'ARNm et l'ARNr (Jacob et al., 1987, Brink et al., 1995). Cependant, les ribosomes sont aussi capables

#### Traduction des ARNm canoniques



#### Traduction des ARNm sans séquence 5' non traduite

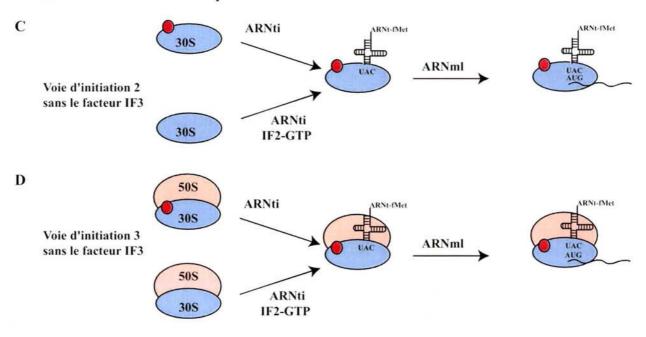


Figure 17 : Les différents mécanismes de formation du complexe pré-ternaire (d'après Moll et al., 2004).

A et B: Initiation de la traduction des ARNm canoniques. Le recrutement de la sous-unité 30S est réalisé via des interactions entre les séquences SD et aSD. L'ARNt initiateur (ARNti) interagit avec le codon initiateur avant ou après la fixation de la sous-unité 30S.

C et D : initiation de la traduction des ARNm sans séquence 5' non traduite (ARNml pour "ARNm leaderless"). La sous-unité 30S, dépourvue du facteur IF3, ou le ribosome 70S, se lie au niveau du codon d'initiation de l'ARNm. L'association avec l'ARNm est réalisée uniquement via les interactions codon-anticodon avec l'ARNti.

La sous-unité 30S est indiquée par un disque bleu, la sous-unité 50S par un disque orange, les facteurs d'initiation de la traduction IF2 et IF3 par un disque rouge et une étoile verte respectivement. Le facteur IF1 est absent de cette figure. Dans chacune des voies d'initiation de la traduction, le facteur IF2 et l'ARNti peuvent se lier indépendemment à la sous-unité 30S, et au ribosome 70S, ou former le complexe ternaire IF2-GTP-ARNti avant la fixation au ribosome.

d'initier la traduction d'ARNm sans séquence SD canonique, tel que l'ARNm *rpsA* codant la protéine S1 (Boni *et al.*, 2001) ou les ARNm issus de chloroplastes (Fargo *et al.*, 1999), ou possèdant peu ou pas de nucléotides situés en amont du codon initiateur (pour revue, Moll *et al.*, 2002). Nous détaillerons ici les différents mécanismes d'initiation de la traduction pour les ARNm comportant une séquence SD canonique, et ceux avec peu ou pas de nucléotides en amont du codon initiateur.

#### I.1 / Initiation de la traduction des ARNm portant une séquence SD

En plus de la particule ribosomique 30S, l'ARNm et l'ARNt-fMet, interviennent les trois facteurs d'initiation de la traduction IF1, IF2 et IF3 (pour revue, Kozak, 1999). Les facteurs IF1 et IF3 maintiennent une portion de particule ribosomique 30S libre : IF1 favorise la dissociation du ribosome 70S alors que IF3 se lie à la sous-unité 30S et permet la dissociation des éléments restants du cycle de traduction précédent, c'est-à-dire l'ARNt non chargé situé au niveau du site P et l'ARNm (Karimi et al., 1999; pour revue, Kozak, 1999). Le facteur IF2 reconnaît spécifiquement l'ARNt-fMet au niveau de son bras accepteur (Guenneugues et al., 2000), et l'apporte au site P de la particule ribosomique 30S, contrairement au facteur d'élongation Tu (EF-Tu), qui apporte les ARNt chargés seulement au niveau du site A du ribosome entier. Le facteur IF1 stimule l'action des facteurs IF2 et IF3 en se liant au site A de la sous-unité 30S (pour revue, Ganoza et al., 2002). Il modulerait l'association des sous-unités et la fidélité des interactions de l'ARNt lors du changement conformationnel du complexe pré-ternaire.

Deux mécanismes d'initiation de la traduction ont été proposés pour les ARNm portant une séquence SD (pour revue, Ganoza et al., 2002). Dans le premier cas (Fig. 17A), la particule ribosomique 30S se fixerait de façon aléatoire à l'ARNm et plus fortement au niveau d'un site d'initiation de la traduction via les interactions SD/aSD, puis à l'ARNt initiateur (de Smit et van Duin, 2003). Dans le deuxième cas (Fig. 17B), l'ARNt initiateur est déjà fixé à la sous-unité 30S avant l'interaction avec l'ARNm (Wu et al., 1996). La présence de l'ARNt-fMet stabilise ensuite la forme GTP-IF2 liée à la sous-unité 30S. Cette forme GTP-IF2 accélère le recrutement de la particule ribosomique 50S, aboutissant au complexe 70S (Antoun et al., 2003). La fixation de la sous-unité 50S est un processus irréversible, qui s'accompagne de la dissociation successive des facteurs d'initiation IF1 et IF3 (Fig. 16). Ceci permet ensuite la liaison rapide de l'ARNt-aminoacylé au site A, apporté par une molécule EF-Tu en présence d'une molécule de GTP. L'hydrolyse du GTP lié au facteur IF2 favorise le départ du facteur IF2 du ribosome et permet la formation de la première liaison peptidique (Antoun et al., 2003).

Le complexe pré-ternaire d'initiation de la traduction, composé d'ARNm, de l'ARNt initiateur et de la sous-unité 30S réalise un changement conformationnel permettant l'interaction du codon d'initiation avec l'anticodon de l'ARNt-fMet, ce qui aboutit à la formation d'un complexe ternaire stable (pour revue, Gualerzi et Pon, 1990; Schlax et Worhunsky, 2003). Ceci se traduit par la variation des

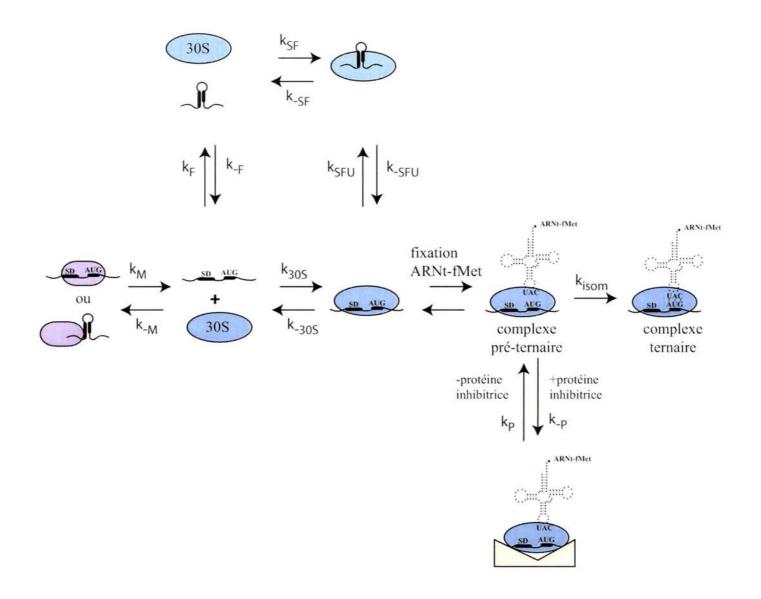


Figure 18 : Modèle cinétique décrivant l'initiation de la traduction et les différentes régulations possibles (d'après Draper, 1996, Schlax et Worhunsky, 2003, de Smit et van Duin, 2003).

La régulation de l'initiation est réalisée à différents niveaux : l'accessibilité de la région TIR au niveau d'ARN fortement structurés dépend de la température mais aussi de la cinétique de repliement : elle est caractérisée par les constantes cinétiques  $k_F$  et  $k_{-F}$ . L'accessibilité de la région TIR peut aussi dépendre d'une protéine ou d'un petit ARN : la fixation de protéine ou d'ARN à l'ARNm est fonction de la constante cinétique  $K_M$ .

L'ARNt se fixerait avant ou après l'étape d'isomérisation. Cependant, l'inhibition de la régulation au niveau du complexe pré-ternaire est réalisée après la fixation de l'ARNt-fMet, avant l'étape d'isomérisation.

La sous-unité ribosomale 30S est représentée par un ovale bleu, la molécule se fixant à l'ARN nu pour inhiber sa traduction en rose, la protéine inhibant la formation du complexe ternaire en jaune.

protections de l'ARNm vis-à-vis de l'attaque des radicaux hydroxyles s'étendant des positions -35 à +5 pendant la liaison de la sous-unité 30S, puis évoluant des positions -35(+/-2) à +19 lorsque le complexe ternaire est formé, puis des positions -9(+/-2) à +19 lorsque le ribosome est formé (Huttenhofer et Noller, 1994). Cette protection est similaire à celle observée au niveau de la structure établie par diffraction aux rayons X du cristal du complexe obtenu entre le ribosome de *Thermus thermophilus* et l'ARNm, où l'ARNm est protégé des positions -15 à +15 (Yusupova et al., 2001).

# I.2 / Initiation de la traduction des ARNm sans région en amont du codon initiateur

Lorsque l'ARNm comprend peu ou pas de séquence en amont du codon initiateur, il recrute le ribosome uniquement, à priori, via le codon initiateur (pour revue, Moll et al., 2002). L'initiation de la traduction de ces ARNm sans région amont est réalisée soit par les ribosomes 70S sans dissociation préalable (Moll et al., 2004), soit par la sous-unité 30S ne comportant pas le facteur IF3 (Fig. 17, C et D) (Tedin et al., 1999). En effet, le facteur IF3 lié à la sous-unité 30S inhibe la formation du complexe d'initiation au niveau d'ARNm sans région amont in vitro, et l'efficacité de la traduction de ces ARNm est inversement proportionnelle à la concentration de IF3 in vivo (Tedin et al., 1999, O'Donnell et Janssen, 2002). L'initiation de la traduction de ces ARNm est alors modulée par le ratio des facteurs d'initiation IF2 et IF3. Le codon d'initiation de la traduction est reconnu par le complexe ribosome-IF2-ARNt-fMet. Cet intermédiaire de la traduction est équivalent à celui formé pendant la traduction des ARNm eucaryotes. Ainsi, la traduction des ARNm sans région amont, présents dans les bactéries, eucaryotes et Archaea, serait une fonction conservée au cours de l'évolution par l'appareil de traduction (pour revue, Moll et al., 2002).

#### I.3 / Modèle cinétique d'initiation de la traduction

Lors de l'initiation de la traduction, la sous-unité 30S des ribosomes se lie à l'ARNm suivant la constante d'équilibre  $K_{30S}$  ( $K_{30S}$ =  $k_{30S}$ /  $k_{.30S}$ ), et l'ARNt-fMet se joint au complexe pour former le complexe pré-ternaire. Ce complexe réalise un changement conformationnel, appelé aussi isomérisation, pour donner lieu au complexe ternaire, suivant la constante  $k_{isom}$  (Fig. 18) (pour revue, Draper, 1996). Lorsque l'isomérisation est rapide, la traduction est immédiatement engagée après la liaison de la sous-unité 30S à l'ARNm, et est proportionnelle à  $k_{30S}$ \*[30S]. Par contre, lorsque l'isomérisation est l'étape limitante, la sous-unité 30S peut quitter l'ARNm suivant  $k_{.30S}$ . La traduction est alors proportionnelle à  $K_{30S}$  \*  $k_{isom}$ \*[30S].

Différents paramètres peuvent affecter la régulation de l'initiation de la traduction. Le repliement de l'ARNm, avec ou sans la sous-unité 30S, fonction de k<sub>SFU</sub> et k<sub>F</sub>, respectivement, inhibe la formation du complexe pré-ternaire. Le repliement inhibiteur de l'ARN peut être dépendant du temps, comme par exemple dans le cas de l'ARNm de la protéine de maturation du phage MS2 (voir paragraphe III.1), de la température, tel que l'ARNm du facteur σ<sup>32</sup> (pour revue, Schlax et Worhunsky, 2003). Différents facteurs protéiques ou nucléiques inhibent l'accès de la sous-unité 30S soit par compétition stérique soit en favorisant une structure inhibitrice de l'ARNm. Ces facteurs peuvent agir au niveau du complexe pré-ternaire en inhibant l'isomérisation. Cette dernière étape d'initiation de la traduction est aussi influencée par la longueur de la séquence SD, l'espacement entre la séquence SD et le codon initiateur, et la présence d'un codon initiateur autre que le codon AUG (Ringquist et al., 1992).

#### II / Efficacité de l'initiation de la traduction

Différents traits de séquence influencent la traduction au niveau de ses trois étapes :

- les structures primaires et secondaires de l'ARNm au niveau de la région TIR pour l'initiation,
- l'utilisation des codons au niveau de l'élongation (Kanaya et al., 1999, Karlin et al., 2001),
- la nature du codon stop pour la terminaison, les codons en amont, et la base suivante (Mottagui-Tabar *et al.*, 1994, Bjornsson *et al.*, 1996; Poole *et al.*, 1995).

Nous détaillerons ici l'influence de la structure primaire et secondaire de la séquence SD, puis de la nature du codon initiateur, et des codons situés en aval et enfin, les codons terminateurs.

# II.1 / La séquence SD : structure primaire

La sous-unité 30S interagit avec l'ARNm au niveau de la séquence SD 5'UAAGGAGGUGA3' située en amont du codon initiateur. Cette séquence SD s'apparie via 3 à 9 résidus à la séquence aSD (3'AUUCCUCCACU5'), située aux positions 1532 à 1542 de la région 3' de l'ARNr 16S d'E. coli (Fig. 19A) (pour revue, Kozak, 1999). La séquence SD favorise la fixation de l'ARNm sur le ribosome, et le positionnement du codon d'initiation à proximité du site P ribosomal (pour revue, Ganoza et al., 2002). Son efficacité est déterminée par son potentiel d'interaction avec la séquence aSD et l'espacement avec le codon d'initiation (Ringquist et al., 1992, de Smit et van Duin, 1994b). En général, les séquences SD de l'ARNm sont constituées de 3 à 4 bases, avec de faibles variations autour du consensus GGAGG (Schurr et al., 1993; pour revue, Kozak, 1999). La distance séparant la séquence SD du codon initiateur a été déterminée au niveau de deux séquences SD différentes, appelées « SD » (GAGGU) et « sd » (UAAGG) (Fig. 19) par Chen et al. (1994). Ceci a conduit à la définition de l'écartement corrigé,

comme étant l'écartement entre le codon d'initiation de la traduction et le nucléotide en face de la position 1536 de l'ARNr 16S (Fig. 19A) (Chen et al., 1994; Ma et al., 2002). L'espacement aligné varie entre 5 et 13 paires de bases, l'espacement optimal étant entre 8 et 10 paires de base chez E. coli (Ringquist et al., 1992; Chen et al., 1994). Ceci permet un positionnement correct du codon d'initiation AUG par rapport à l'ARNr 16S. Ainsi, dans le complexe ternaire d'initiation de la traduction, l'ARNt initiateur peut interagir de façon optimale avec le codon d'initiation, et la traduction est efficace (Fig. 19C, situations 1 et 2). En revanche, si l'écartement est trop faible (Fig. 19C, situation 3) ou trop élevé, le codon d'initiation n'est plus correctement positionné dans le site P du ribosome, l'interaction efficace entre l'ARNt et le codon d'initiation n'est pas efficace (Chen et al., 1994).

L'espacement optimal dépend aussi de la longueur de la séquence SD. Plus la séquence SD est longue, plus l'espacement entre la séquence SD et le codon initiateur doit être élevé (Nivinskas et al., 1999). D'autre part, l'augmentation de longueur de la séquence SD a été corrélée à une augmentation de la traduction pour certains transcrits (Ringquist et al., 1992). Une étude réalisée au niveau de 30 génomes procaryotes montre que les gènes fortement exprimés ont une plus longue séquence SD que les autres gènes (Ma et al., 2002). Cet effet positif sur la traduction a été observé seulement si l'ARNm contient des structures secondaires limitant l'accès des ribosomes au codon initiateur (de Smit et van Duin, 1994b) ou si un codon initiateur alternatif se substitue au codon initiateur généralement retrouvé, l'AUG (Weyens et al., 1988). La plus forte interaction entre les séquences SD et aSD peut aussi ralentir le départ du ribosome au niveau de l'initiation de la traduction. La protéine S1 semble alors interagir pour contrer les effets inhibiteurs d'une trop longue séquence SD et faciliter l'initiation de la traduction (Komarova et al., 2002).

# II.2 / La protéine 51

Par expérience SELEX, il a été montré que la protéine S1 est le composant majeur du ribosome impliqué dans la sélection de l'ARNm chez les bactéries à gram négatif (Ringquist et al., 1995). Cette protéine ribosomale de la sous-unité 30S, de 61 kDa, comprend un domaine N-terminal responsable des interactions protéines-protéines, et un domaine central et C-terminal de liaison à l'ARN simple brin via les répétitions du motif S1 (Bycroft et al., 1997). Les ribosomes dépourvus de protéine S1 peuvent interagir avec les régions TIR des ARNm possédant une forte séquence SD (Farwell et al., 1992), cependant la formation du complexe ternaire n'est pas toujours réalisée (Komarova et al., 2002). En effet, la protéine S1 facilite à la fois la liaison du ribosome à l'ARNm et l'initiation de la traduction, indépendamment de la longueur de la séquence SD, mais aussi la formation du complexe ternaire lorsque la séquence SD est grande (Komarova et al., 2002).

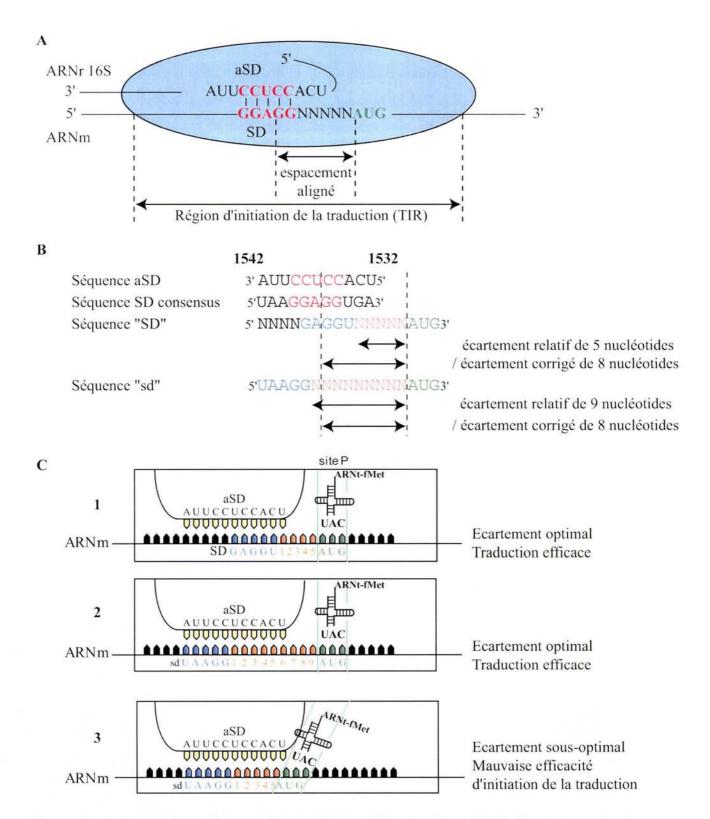


Figure 19 : Influence de la distance séparant le motif SD du codon d'initiation de la traduction.

A : Définition de l'écartement corrigé (Ma et al., 2002). L'écartement corrigé est défini par la distance entre le codon d'initiation de la traduction et le nucléotide complémentaire au résidu 1537 dans la séquence aSD de l'ARNr 16S.

B : Ecartement relatif et corrigé pour les séquences "SD" (GAGGU) et "sd" (UAAGG) (Chen et al., 1994).

C : Modélisation de l'influence de l'écartement sur l'efficacité de la traduction. Les nucléotides importants sont indiqués par des flèches de couleurs : jaune pour la séquence aSD, bleue pour les séquences SD et sd, orange pour les nucléotides séparant la séquence SD, ou sd, du codon initiateur, et vert pour le codon initiateur. Dans les conditions 1 et 2, l'ARNt initiateur interagit de façon optimale avec le codon initiateur. En revanche, dans la situation 3, l'écartement ne permet pas un positionnement correct de l'ARNt initiateur dans le site P.

#### II.3 / Le codon initiateur

En plus du codon classique d'initiation de la traduction AUG, le codon GUG est quelques fois utilisé, les codons UUG, AUU et AUA de façon plus exceptionnelle, (pour revue, Draper, 1996; Ganoza et al., 2002). L'initiation au codon AUA a été observée lors de la production d'une protéine d'Archaea chez E. coli (Kopke et Leggatt, 1991), tandis que le codon UUG initie la synthèse de la protéine ribosomale S20 (pour revue, Draper, 1996). Le codon AUU est utilisé lors de la traduction des gènes infC codant le facteur d'initiation IF3 et pcnB, codant la poly (A) polymérase I (PAP). L'étude de l'efficacité traductionnelle in vivo des différents codons d'initiation possibles variant d'un nucléotide par rapport au codon classique AUG a permis de définir trois classes de codon initiateurs ({Sussman, 1996 #39). La classe I comprend les codons AUG, GUG et UUG permettant une traduction efficace, supérieure à 10 % à celle initiée par le codon AUG. La classe IIA, composée des codons CUG, AUU, AUC, AUA et ACG initie la traduction à un taux de 1 à 3 % par rapport à une traduction initiée au codon AUG. Les deux codons AGG et AAG, de la classe IIB, montrent une très faible initiation de la traduction, non quantifiable. L'initiation de la traduction au niveau des codons de la classe IIA est régulée par le facteur IF3, qui déstabilise les complexes d'initiation de la traduction au niveau de codon non canonique (pour revue, Kozak, 1999). Son inactivation conduit à une augmentation de l'initiation de la traduction au niveau des codons de la classe IIA d'un facteur 3 à 5. Comme le codon initiateur AUU du gène infC codant le facteur IF3 appartient à cette classe IIA, le facteur IF3 s'autorégule négativement (Sacerdot et al., 1996). Dans le cas du gène pcnB, où la surproduction de la protéine correspondante est toxique pour la cellule, la régulation par IF3 diminue d'un facteur 5 la production de cette protéine (Binns et Masters, 2002).

Bien que la nature du codon initiateur varie, il est toujours reconnu par l'ARNt-fMet (pour revue, Ganoza et al., 2002). En effet, seul cet ARNt peut initier la traduction. Ainsi, une formyl-méthionine est incorporée en N-ter de toutes les protéines synthétisées. Cependant, le groupement formyl ou la formyl-méthionine peuvent être enlevés au cours de la synthèse protéique.

#### II.4 / La séquence Downstream Box située en aval du codon d'initiation

En plus de la séquence SD et du codon initiateur, d'autres éléments de l'ARNm favorisent l'efficacité d'initiation de la traduction. Différentes hypothèses d'appariement entre l'ARNm et l'ARNr 16S ont été émises, basées sur des données statistiques et l'observation de l'effet de mutations augmentant l'expression des gènes. Ainsi, la complémentarité de type Watson Crick, similaire à l'appariement SD-aSD, entre la séquence dite <u>D</u>ownstream <u>Box</u> (DB) située en aval du codon d'intiation et la séquence dite anti-Downstream Box (aDB) des bases 1469 à 1483 l'ARNr 16S (5'AGAUACUUAGUGUUUC3') (Sprengart et al., 1990), a été proposée comme élément favorisant la

traduction, en particulier pour les ARNm sans région 5'UTR (Shean et Gottesman, 1992). Bien que cette complémentarité ait été retrouvée chez différents ARNm bactériens et phagiques, aucune évidence biochimique, ni génétique n'a pu confirmer cette interaction (pour revue, Moll et al., 2002). Au contraire, la région aDB de l'ARNr 16S est sous forme double-brin, ce qui l'empêcherait de s'apparier avec l'ARNm (La Teana et al., 2000). De même, la structure de la sous-unité 30S de Thermus thermophilus ne permet pas de soutenir les modèles d'appariement DB-aDB (Moll et al., 2001; pour revue, Moll et al., 2002). L'effet activateur de la traduction par cette séquence DB semble plutôt lié à une absence de structure secondaire de l'ARNm (pour revue, Moll et al., 2002). Ainsi, une vérification expérimentale, via des mutations compensatoires, est nécessaire pour l'identification de séquence activatrice de la traduction par complémentarité avec l'ARNr 16S (pour revue, Draper, 1996).

# II.5 / Les codons situés au début des ORF et leur influence sur la structure secondaire de l'ARNm

Bien que la région située en 5' d'ORF fortement exprimées ne semble pas interagir avec l'ARNr 16S, elle se comporte comme un activateur de la traduction (O'Connor et al., 1999). Cette activation semble être corrélée à l'utilisation des codons dans cette région. En effet, l'utilisation de codons synonymes peu fréquents par rapport à l'utilisation de codons synonymes fréquents portant le même acide aminés localisé dans la région codante de gènes fortement exprimés dans les gènes d'E. coli, montre une diminution des taux de traduction d'un facteur 6 (Sorensen et al., 1989). De même, la différence d'efficacité d'initiation de la traduction dans les fusions trp-lacZ est de l'ordre d'un facteur 20 lorsque le codon en position +2 est un codon rare par rapport à un codon fréquent (Stenstrom et al., 2001). Dans le génome d'E. coli, l'analyse des codons situés en position + 2 montre que les codons riches en A, en particulier le codon AAA, sauf les codons ACU, ACG et AGG, sont plus fortement utilisés (Looman et al., 1987; Stenstrom et al., 2001). Cet effet activateur semble être lié à la faible structuration de la région 5' codante de l'ARNm (Stenstrom et al., 2001; Voges et al., 2004). Elle faciliterait le départ du ribosome du RBS, et donc l'initiation de la traduction suivante. Une utilisation différentielle des codons a pu être retrouvée au niveau de la région 5' des régions codantes, comprenant les 50 premiers codons (Karlin et al., 1998). Ceci pourrait aussi contribuer à faciliter le départ du ribosome. Par contre, la structuration de l'ARN en aval de la région 5' influence faiblement la traduction. Ainsi, une fois que le ribosome en cours de traduction a quitté la région d'initiation de la traduction, il pourrait plus facilement désapparier les structures secondaires potentielles tout en poursuivant la traduction (Voges et al., 2004). La présence d'une structure tige-boucle 15 nts en aval du codon d'initiation peut aussi augmenter l'expression des gènes (Paulus et al., 2004). Cette structure pourrait alors éviter les interactions à longue distance avec la région TIR de l'ARNm, ce qui augmente l'accessibilité du site d'initiation au ribosome.

#### II.6 / L'utilisation des codons des ORF et les codons terminateurs

Parmi les différents éléments impliqués dans la régulation de la traduction, l'élément le plus important pour la production de protéines chez *E. coli* est l'utilisation des codons (Lithwick et Margalit, 2003). En effet, lorsqu'un codon est fréquemment utilisé, la cellule bactérienne comporte une quantité élevée d'ARNt correspondant. Ceci évite au ribosome d'entrer en pause et augmente la vitesse d'élongation du ribosome.

Enfin, la nature du codon stop et les bases situées de part et d'autre influencent la terminaison. En effet, chez *E. coli*, le codon stop UAA est préférentiellement utilisé par les protéines fortement exprimées chez *E. coli* par rapport aux deux autres codon stop UAG et UGA, qui sont moins efficaces (Lithwick et Margalit, 2003). De plus, la présence d'un codon stop inefficace conduit à une faible quantité de protéines (Jin et al., 2002). L'analyse statistique d'environ 2000 gènes du génome d'*E. coli* a mis en évidence un rôle potentiel de la base suivant le codon stop. *In vivo*, l'efficacité de terminaison varie de 80% (UAAU) à 7% (UGAC), avec la hiérarchie U>G>A#C pour les codons UAA et UGA, et G>U#A>C pour UGA (Poole et al., 1995). Enfin, les acides aminés situés en amont du codon de terminaison influencent l'efficacité de la terminaison (Mottagui-Tabar et Isaksson, 1997; Sunohara et al., 2002). En particulier, lorsque le codon rare arginine AGG est situé juste en amont du codon stop UGA peu efficace, le système de *trans*-traduction ARNtm est recruté, ce qui conduit à un clivage de l'ARNm correspondant (Hayes et al., 2002).

# III / Régulation de l'initiation de la traduction par les éléments en trans de l'ARNm

L'efficacité d'initiation de la traduction est aussi modulée par l'accessibilité des ribosomes à la région TIR de l'ARNm. Différents facteurs agissent directement au niveau de la région TIR ou à proximité (Fig. 18).

# III.1 / Structure secondaire et tertiaire de la région 5' non traduite

Les structures secondaires et tertiaires de l'ARNm qui impliquent les régions flanquant la région d'initiation et/ou le codon d'initiation affectent l'expression des gènes (de Smit et van Duin, 1994a; Boni et al., 2001, pour revue, Schlax et Worhunsky, 2003). Elles peuvent diminuer l'efficacité de l'initiation de la traduction en limitant la reconnaissance par la sous-unité 30S. Ainsi, la séquence SD de l'ARNm codant la protéine de surface du bactériophage MS2 participe à la formation d'une hélice d'ARN (Fig. 20A). Une corrélation quantitative entre la stabilité thermodynamique de cette structure et l'efficacité de la traduction a pu être démontrée in vivo (de Smit et van Duin, 1993; de Smit et van Duin, 1994b). Ainsi, l'aide lorsque  $\Delta G$ de la structure, obtenu à du logiciel Mfold [(http://www.bioinfo.rpi.edu/applications/mfold/old/rna/form1.cgi (Zuker, 2003)] n'est pas inférieur à -6 kcal.mol<sup>-1</sup>, la structure secondaire n'inhibe pas l'initiation de la traduction. Par contre, chaque diminution de 1,4 kcal.mol<sup>-1</sup> au-delà de cette valeur de -6 kcal.mol<sup>-1</sup> conduit à une diminution de l'initiation de la traduction d'un facteur 10 (de Smit et van Duin, 1994a). Cette corrélation a pu être aussi démontrée pour d'autres variants de séquence, au niveau du gène lamB, du gène mom du phage Pu, du gène marR (Martin et Rosner, 2004; pour revue, Schlax et Worhunsky, 2003). Lorsque le repliement de l'ARN inhibe la traduction, la limitation de l'initiation de la traduction va dépendre de la cinétique de formation de la structure inhibitrice (pour revue, Schlax et Worhunsky, 2003). Ainsi, la protéine de maturation du bactériophage MS2 contient une région 5' non traduite de 130 nts, qui se replie en forme de feuille de trèfle, dans laquelle la séquence SD interagit avec une séquence complémentaire située à une longue distance (Fig. 20B) (Groeneveld et al., 1995). La synthèse de la protéine a lieu entre le moment de la transcription de cet ARNm et la formation de la structure inhibitrice (van Meerten et al., 2001). Une corrélation a pu ensuite être établie entre la cinétique de repliement de cette région non traduite et l'efficacité de la traduction.

La structure secondaire et tertiaire de l'ARNm au niveau de la région TIR peut aussi jouer un rôle d'activateur de la traduction. En effet, la sous-unité 30S peut se lier à l'ARNm lorsque celui-ci comporte de longues structures en tige-boucle, ces structures peuvent alors optimiser la position des éléments requis pour la formation du complexe d'initiation de la traduction. Ainsi, au niveau des ARNm 25 et 38 du phage T4, la séquence séparant la séquence SD du codon d'initiation de la traduction est de 27 et 22 nts respectivement (Nivinskas et al., 1999). La formation d'une structure tige boucle réduit la distance entre la séquence SD et le codon d'initiation à 11 et 5 nts respectivement (Fig. 20, C et D) (Nivinskas et al., 1999). Cette structure facilite la fixation de la sous-unité 30S (Fig. 21B), mais ne semble pas interférer avec la fixation de la sous-unité 50S (Gold, 1988). La présence d'une structure tige-boucle peut aussi être nécessaire en amont de la séquence SD, afin de rapprocher

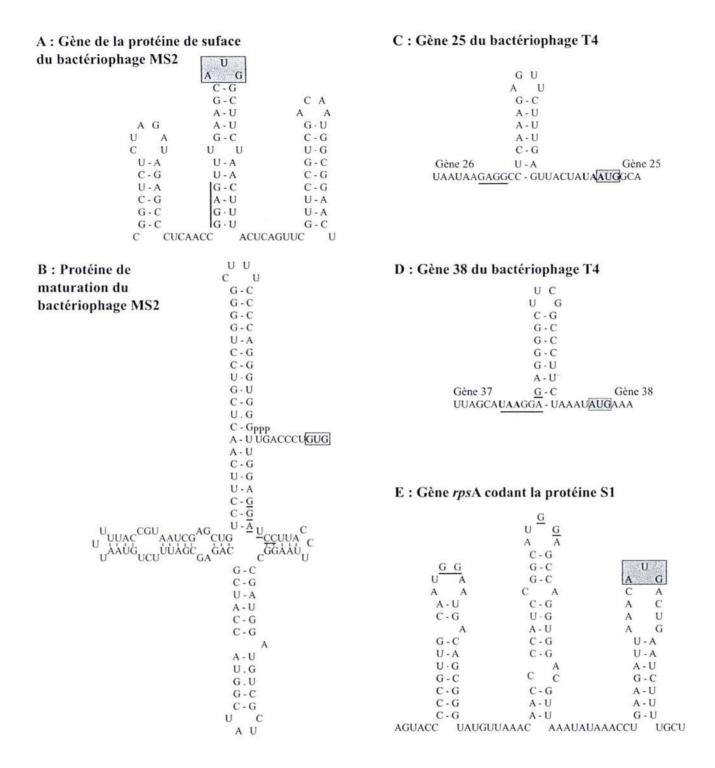


Figure 20 : Exemples de structures tige-boucle affectant l'efficacité de la traduction.

A: Structure secondaire de l'élément TIR de l'ARNm de la protéine de surface du bactériophage MS2 (De Smit et van Duin, 1990). La structure tige-boucle séquestre la séquence SD et diminue l'efficacité d'initiation de la traduction.

B : Structure secondaire en forme de feuille de trèfle de l'ARNm de la protéine de maturation du bactériophage MS2 (d'après van Meerten et al., 2001).

C et D : Structure secondaire de l'élément TIR des ARNm 25(A) et 38(B) du bactériophage T4 (d'après Nivinskas *et al.*, 1999). La structure tige-boucle rapproche la séquence SD du codon initiateur.

E : Repliement de la région TIR de l'ARNm *rpsA* (d'après Boni *et al.*, 2001). La séquence SD pourrait être formée par les séquences GG en amont des boucles apicales.

Le codon terminateur du gène amont, lorsqu'il est présent, est indiqué en gras, le codon initiateur est encadré et grisé, la séquence SD est soulignée.

une séquence activatrice, comme dans le cas de l'ARNm *thrS* de la thréonyl-synthétase (Fig. 21C et 23A) (Sacerdot *et al.*, 1998).

La structure secondaire et tertiaire semblent être aussi requises pour reconstituer la séquence SD, comme cela est proposé pour l'ARNm *rpsA* codant la protéine ribosomale S1 (Boni *et al.*, 2001). En effet, l'ARNm *rpsA* est fortement et efficacement traduit, mais ne possède pas de séquence SD canonique (Boni *et al.*, 2001). Cependant, sa région TIR se replie en trois structures tiges boucles I, II et III (Fig. 20E). Ce mode de repliement est conservé chez les protéobactéries en général, en particulier dans la subdivision γ mis à part le groupe des *Pseudomonas* (Boni *et al.*, 2001; Tchufistova *et al.*, 2003). Il pourrait représenter une région de liaison au ribosome via les boucles apicales I et II. Celles-ci contiennent le motif conservé GG, nécessaire pour la forte traduction de l'ARNm *rpsA*, qui jouerait ici le rôle de séquence SD (Fig. 21D). Le recrutement de la sous-unité 30S pourrait aussi être réalisé via la protéine ribosomale S1, qui interagit via les régions simple brin. Dans ce cas, les motifs GG des boucles apicales I et II seraient nécessaires pour maintenir une structure tertiaire optimale de la région TIR. Ce mode de recrutement des ribosomes via les structures secondaires, voire tertiaires, pourrait être aussi utilisé par les chloroplastes (Boni *et al.*, 2001). En effet, les grandes régions 5' non traduites, fortement structurées, semblent initier la traduction en absence de séquence SD classique (Fargo *et al.*, 1999).

## III.2 / La température

La structure secondaire de la région TIR peut aussi être modulée par la température. En effet, l'expression du gène rpoH, codant le facteur  $\sigma^{32}$ , est régulée au niveau de sa traduction. Ce facteur voit sa synthèse augmenter lors d'un choc thermique. Ceci s'explique en partie par une déstabilisation de la structure secondaire de l'ARNm, à température élevée, qui entraîne une libération du site RBS (pour revue, Schlax et Worhunsky, 2003).

# III.3 / Influence de la traduction du cistron amont sur la traduction du cistron aval d'ARNm polycistronique : le couplage traductionnel

Dans les paires de gènes adjacents cotranscrits, la traduction du cistron amont favorise le plus souvent la traduction du cistron aval. Ce mécanisme, appelé couplage traductionnel, est lié soit à la proximité du site d'initiation du cistron aval du codon terminateur du cistron amont (pour revue, Draper, 1996), soit à la modification de la structure au niveau de la région TIR du gène aval. Ce mécanisme de régulation est utilisé par les opérons codant les protéines ribosomales, afin d'obtenir une quantité

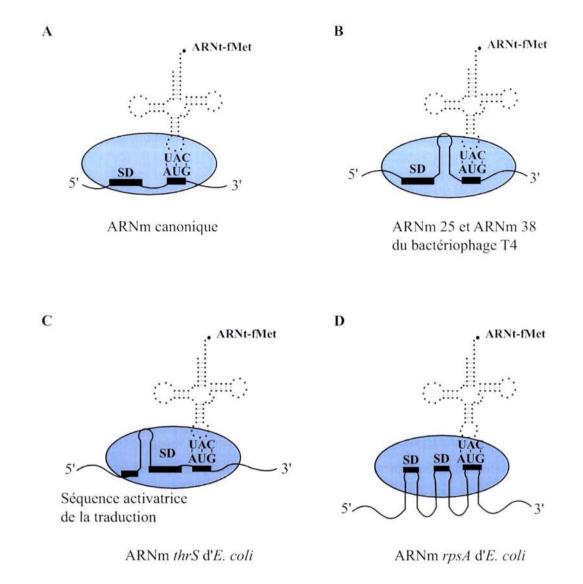


Figure 21 : Modèles de complexes pré-ternaires au niveau d'ARNm avec une séquence SD canonique, ou en présence de structure secondaire activatrice de la traduction (d'après Boni et al., 2001).

A : Les ARNm possèdant une SD classique recrutent la sous-unité 30S via les interactions entre les séquences SD et aSD.

B : La présence d'une structure tige-boucle située entre la séquence SD et le codon initiateur rapproche ces deux éléments et favorise l'initiation de la traduction.

C : La structure tige-boucle est ici située en amont de la séquence SD afin de rapprocher l'élément activateur de la séquence SD.

D : L'ARNm *rpsA* recruterait la sous-unité 30S par les boucles apicales de la région amont. Cet ARNm peut aussi recruter le ribosome par une interaction avec la protéine S1. Ce cas n'est pas représenté ici.

stoechiométrique de protéines (pour revue, Keener et Nomura, 1996). Ainsi, au niveau de l'opéron L35 constitué des gènes *infC-rpmI-rplT* qui codent respectivement le facteur d'initiation IF3 et les deux protéines ribosomales L35 et L20 (Fig. 22), la traduction du cistron *rplT* dépend de la traduction du cistron *rpmI* qui est aussi influencée par la traduction du cistron *infC*. La traduction du cistron *rpmI* libère le site d'initiation de la traduction du cistron *rplT* d'une structure inhibitrice (Lesage *et al.*, 1992). La traduction du gène *infC*, diminue la répression de la traduction exercée par la protéine L20 au niveau du gène *rpmI*, codant la protéine L35, en dénaturant la structure reconnue par la protéine L20 (Guillier *et al.*, 2002).

Dans quelques cas, la traduction du cistron amont inhibe celle du cistron aval. En effet, lorsque les cistrons se recouvrent partiellement, les ribosomes en cours d'élongation sur l'ARNm bloquent l'accès physique au site d'initiation distal. Les ribosomes peuvent aussi limiter l'initiation de la traduction du cistron aval, sans que les cistrons se recouvrent. Ainsi, dans le cas des gènes V et VII du phage filamenteux IKe, la forte traduction du gène V limite celle du gène VII (Yu et al., 2001).

# III.4 / Répression de l'initiation de la traduction par fixation de protéines

Le contrôle de la traduction réalisé par les protéines, est souvent négatif, bien qu'un cas de régulation positive a été caractérisé dans le cas du bactériophage μ (Wulczyn et Kahmann, 1991). Ce mécanisme de répression de la traduction est souvent utilisé par les protéines de liaison à l'ARN. Un exemple particulièrement étudié est celui de la régulation exercé par les protéines ribosomales, pour autoréguler leur synthèse par un mécanisme de rétrocontrôle. Dans ce mécanisme, une compétition s'exerce entre le site naturel de fixation des protéines sur l'ARNr et leur site de liaison au niveau de leur ARNm. Ainsi, une ou plusieurs protéines ribosomales codées par l'opéron se fixent sur leur propre ARNm et bloquent la traduction de leur cistron ou d'un cistron amont lorsque les protéines ribosomales répresseurs sont en excès par rapport à l'ARNr disponible (Fig. 22) (pour revue, Keener et Nomura, 1996). Le couplage traductionnel entre les protéines ribosomales inhibe ensuite la traduction des cistrons situés en aval. Deux types d'inhibition de la traduction au niveau d'un cistron sont possibles (pour revue, Schlax et Worhunsky, 2003) (Fig. 18):

- répression par compétition du site de fixation RBS entre la sous-unité 30S et la protéine agissant en trans,
- répression par blocage du processus d'initiation de la traduction après fixation de la particule ribosomique 30S.

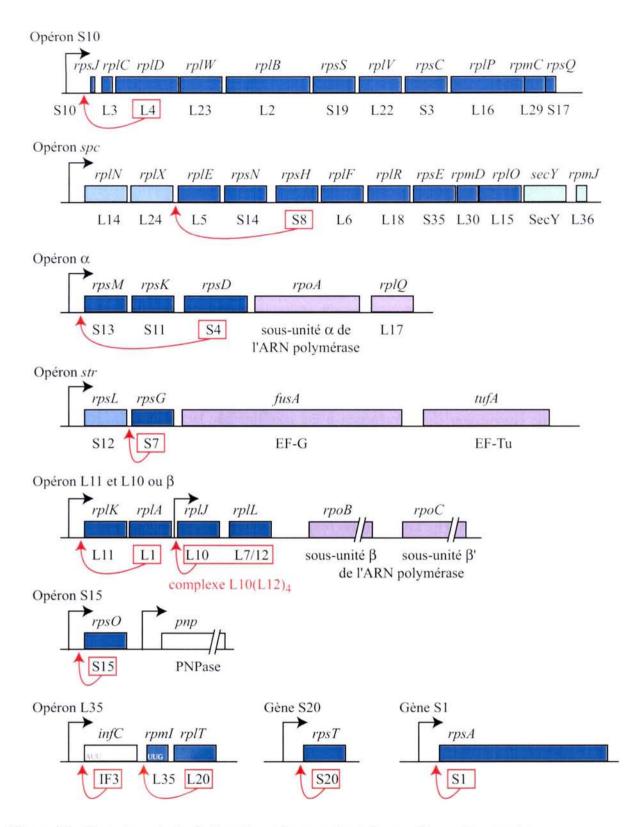


Figure 22 : Structure et régulation des opérons codant des protéines ribosomales (d'après Keener et Nomura, 1996, Draper, 1996).

Les effets de rétrocontrôle de l'initiation de la traduction sont représentés par des flèches et les protéines agissant en répresseur sont encadrées en rouge. Les promoteurs transcriptionnels sont indiqués par des flèches brisées. Les ORF indiquées en bleu foncé sont rétrocontrolées par la protéine répresseur directement, alors que l'expression des ORFs représentées en bleu clair est soumise au rétrocontrôle de façon indirecte. Un contrôle d'expression faible ou nul est observé pour les ORF en violet et blanc, respectivement, et aucun niveau de régulation n'est connu pour les ORF en vert. Les codons initiateurs différents de AUG sont indiqués.

#### III.4.a / Autorégulation des protéines de liaison à l'ARNm

Lors de la répression par compétition au site de fixation RBS, la protéine reconnaît la séquence cible par sa séquence primaire ou par la structure secondaire, qui présentent des similitudes avec l'ARN cible de cette protéine. Cette similitude entre l'ARNm et l'ARN cible de la protéine a été décrite en particulier pour la Thréonyl-tRNA synthétase (ThrRS) (pour revue, Romby et Springer, 2003). En effet, la ThrRS autorégule sa synthèse par inhibition de la traduction de son ARNm, suivant les affinités relatives de ThrRS pour l'ARNt<sup>Thr</sup> et pour son ARNm (Fig. 23) (Romby et al., 1996). La ThrRS se fixe avec une plus forte affinité sur l'ARNt<sup>Thr</sup> que sur la région TIR de son ARNm, au niveau des domaines 2 et 4 de la région TIR de son ARNm, qui miment le bras anticodon de l'ARNt<sup>Thr</sup>. Elle interagit avec les bases de la boucle contenant la séquence mimant l'anticodon du domaine 2 (Torres-Larios et al., 2002; Caillet et al., 2003). Le domaine 2 serait le site principal de liaison de ThrRS, alors que le domaine 4 améliorerait l'affinité de la ThrRS avec son ARNm. De cette manière, la ThrRS empêcherait la fixation de la sous-unité 30S par inhibition stérique (pour revue, Romby et Springer, 2003). Le modèle de compétition de la traduction a été confirmé par le fait que l'augmentation de la production de la ThrRS à partir de l'ARNm thrS résulte soit d'une augmentation du nombre de ribosomes, soit d'une diminution de la quantité d'ARNt<sup>Thr</sup> (Comer et al., 1996). L'appariement de la ThrRS au niveau de son ARNm donne lieu à la formation d'un complexe très stable, ce qui conduit à la dégradation de l'ARNm (Fig. 23C) (Nogueira et al., 2001).

La répression de la traduction peut aussi avoir lieu après l'appariement de la sous-unité 30S à l'ARNm (pour revue, Schlax et Worhunsky, 2003). Deux exemples de ce mécanisme ont été étudiés : la régulation de l'opéron α et celle de la production de la protéine ribosomale S15. La protéine S15, lorsqu'elle est en excès, se lie à son ARNm et réprime sa traduction en inhibant la formation du complexe ternaire (pour revue, Schlax et Worhunsky, 2003). En effet, l'ARNm *rpsO*, codant la protéine S15, est en équilibre entre deux conformations (Fig. 24). Dans la première, la séquence SD est localisée dans une structure tige-boucle peu stable. Dans la deuxième, la séquence SD est simple brin, à proximité d'un pseudonoeud. Cette structure en pseudonoeud présente une meilleure affinité pour la sous-unité 30S et doit être formée pour pouvoir se lier à la protéine S15. La protéine S15 reconnaît le motif de séquence conservé G-U/G-C et la jonction de trois tiges-boucles au niveau de l'ARNr 16S et de l'ARNm *rpsO* (Fig. 24) (Serganov *et al.*, 2002). Ainsi, la formation préférentielle du complexe ternaire avec l'ARNm sous forme de pseudonoeud, la sous-unité 30S et la protéine S15 inhibe la formation du complexe ternaire actif (pour revue, Schlax et Worhunsky, 2003). Un mécanisme de régulation similaire est observé au niveau de l'opéron α.

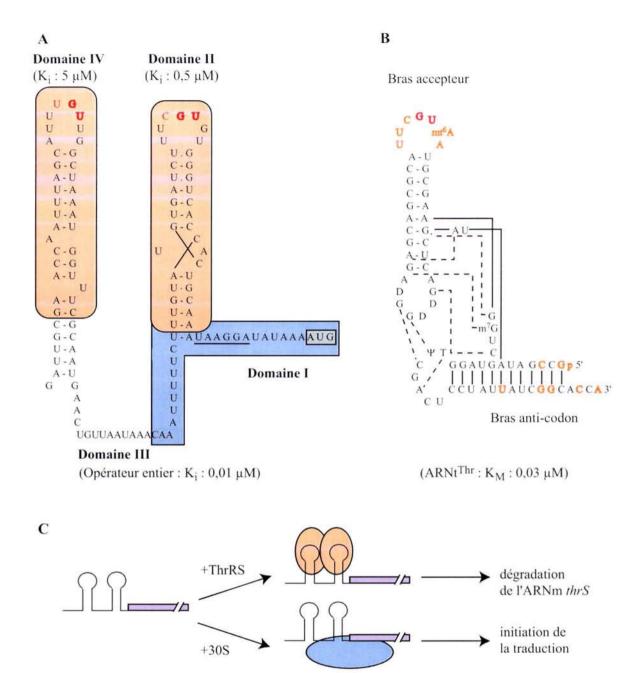


Figure 23 : Autorégulation de la synthèse de la thréonyl-tRNA synthétase (ThrRS).

A: Séquence et structure secondaire de la région 5' non traduite de l'ARNm *thrS* d'*E. coli* (d'après Romby *et al.*, 1996 ; Torres-Larios *et al.*, 2002 ; pour revue, Romby et Springer, 2003). Les domaines I, comprenant le codon d'initiation et la séquence SD, et le domaine III sont en simple brin. Les régions qui interagissent avec la sous-unité 30S du ribosome, sont représentées par un rectangle bleu. L'enzyme ThrRS reconnait spéciquement les domaines II et IV, constitués d'une structure tige-boucle, les zones d'interactions de ces domaines sont inclues dans un cadre orangé. Les bases en rouge ressemblent aux bases de l'anticodon, les bases en gras sont reconnues par ThrRS.

B : Structure des bras anticodons et accepteur de l'ARNt<sup>Thr</sup> (d'après Caillet *et al.*, 2003). Les bases en rouge sont les bases de l'anticodon essentielles pour la réaction catalysée par ThrRS, alors que les bases en orange interagissent avec la ThrRS.

C: Modèle de régulation de la traduction de l'ARNm thrS (pour revue, Schlax et al., 2003). La synthèse de la ThrRS est autorégulée par compétition entre l'ARNt<sup>Thr</sup> et l'ARNm pour la fixation de ThrRS. La structure de la région non traduite forme deux domaines reconnus par la ThrRS. La fixation de la ThrRS conduit ensuite à la dégradation de l'ARNm correspondant. La région codante est indiquée par un cadre violet.

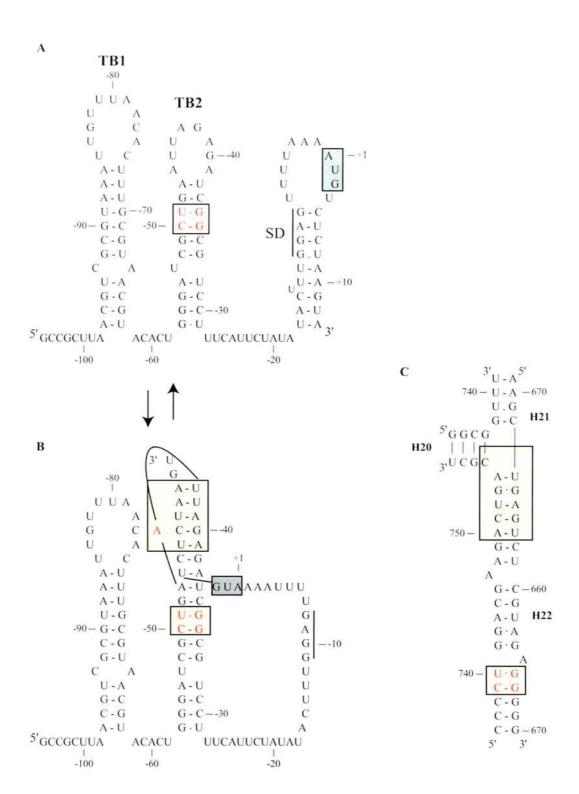


Figure 24 : Sites de fixation de la protéine ribosomale S15 sur l'ARNm rpsO et l'ARNr 16S (d'après Philippe et al., 1990 ; 1993; Serganov et al., 2002).

A et B: Equilibre entre les deux structures de l'ARNm *rpsO*. La structure pseudonoeud (B), sur laquelle se fixe préférentiellement la sous-unité 30S, est requise pour la fixation de la protéine S15. La fixation de la protéine S15 inhibe ensuite la formation du complexe d'initiation de la traduction actif. La séquence SD est soulignée, le codon initiateur est à l'intérieur du rectangle gris.

C : Structure du site de fixation de la protéine S15 sur l'ARNr 16S.

Les deux sites de fixation de la protéine S15 sont indiqués par des rectangles jaunes. Les motifs de séquences reconnus par la sous-unité 30S sont en rouges.

#### III.4.b / Régulation par la protéine CsrA

Pour survivre dans son environnement, la bactérie *E. coli* possède différents systèmes de régulations pour reconnaître et s'adapter aux nouvelles conditions de croissance. Ainsi, la transition de la phase exponentielle à la phase stationnaire est accompagnée de nombreux changements physiologiques importants. Différents systèmes de régulation globaux interviennent, dont le système Csr (Carbon storage regulator). Ce système comprend entre autre la protéine CsrA de 61 acides aminés, protéine de liaison à l'ARN simple brin via son domaine KH. Cette protéine régule l'expression post-transcriptionnelle de certains gènes, positivement ou négativement suivant l'ARNm cible. Ainsi, la protéine CsrA réprime l'expression des gènes de biosynthèse du glycogène, portés par l'opéron glgCAP et cstA, codant un transporteur de peptide. Elle se lie à la région non traduite du transcrit glgCAP, en recouvrant la séquence Shine-Dalgarno de la première ORF, glgC, et la séquence SD du transcrit cstA (Baker et al., 2002; Dubey et al., 2003). La déstabilisation engendrée par la protéine CsrA est liée à l'absence de traduction, par inhibition stérique de la fixation des ribosomes à la région TIR de l'ARNm, et/ou à une modification de la structure de l'ARNm, ce qui le rend alors plus accessible pour une attaque aux endoRNases (pour revue, Romeo, 1998).

#### III.4.c / Régulation par les métabolites

Les métabolites peuvent moduler directement la transcription et/ou la traduction en se fixant au niveau de la région 5' non traduite de certains ARNm. Ces éléments de l'ARN, appelés « riboswitch », sont généralement constitués de deux domaines : ils fixent un ligand au niveau d'un domaine aptamère naturel et modulent en fonction de la concentration en métabolites l'expression des gènes aval via un autre domaine appelé « plateforme d'expression », (pour revue, Winkler et Breaker, 2003). Celle-ci comporte des éléments impliqués dans la régulation de l'expression des gènes au niveau de la terminaison ou de l'anti-terminaison de la transcription, ou au niveau de l'initiation de la traduction (Fig. 25) (pour revue, Nudler et Mironov, 2004). Le domaine aptamère est fortement conservé au niveau de différents organismes, alors que les « plateformes d'expression » varient au niveau de leur séquence, structure et mécanisme d'action. Ces mécanismes de régulation ont été mis en évidence au niveau du métabolisme de vitamines chez B. subtilis, E. coli et Rhizobium etli, mais aussi d'acides aminés comprenant la méthionine, la cystéine et la lysine et de nucléotides, comme l'adénine et la guanine chez B. subtilis (pour revue, Nudler et Mironov, 2004). La majorité des régulations observées est une répression de l'expression des gènes aval, exercée par les « plateformes d'expression » au niveau de la terminaison de la transcription chez les bactéries à gram positif, et au niveau de l'initiation de la traduction chez les bactéries à gram négatif (pour revue, Nudler et Mironov, 2004). Ainsi, chez B. subtilis et E. coli, les ARNm codant les enzymes impliquées dans la voie de biosynthèse de la thiamine (vitamine B1), fixent la

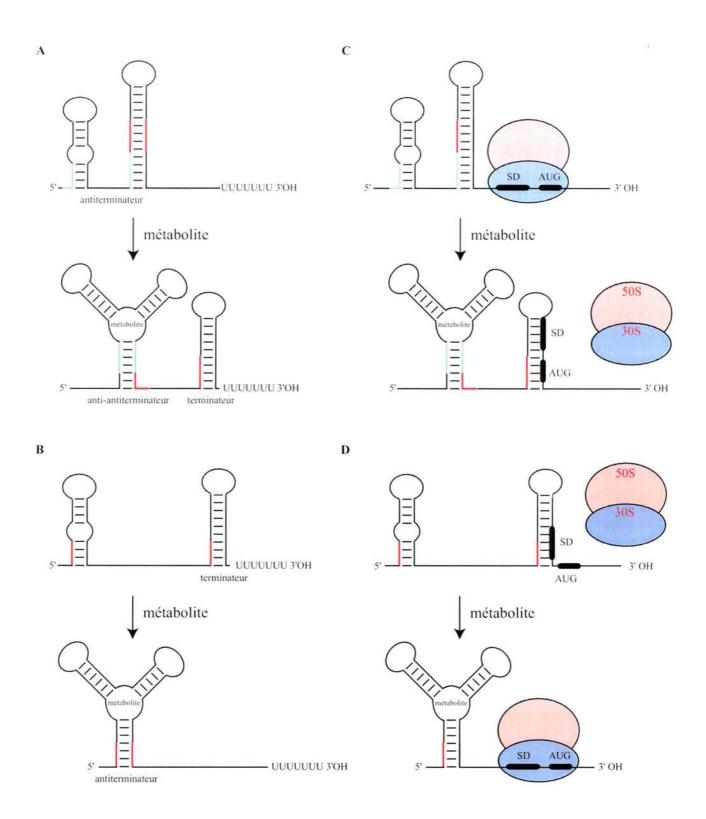


Figure 25 : Représentation schématique du contrôle de l'expression des gènes par les métabolites (d'après Nudler et Mironov, 2004).

La fixation du métabolite au niveau de l'aptamère conduit à un changement de structure de l'ARNm au niveau du domaine appelé "plateforme d'expression". Ce domaine régule l'expression des gènes au niveau de la terminaison de la transcription qui est alors favorisée (A) ou inhibée (B), ou au niveau de la traduction, par modification de l'accessibilité de la région RBS par les ribosomes qui est soit séquestrée (C) ou disponible (D). Il est à noter qu'un seul mécanisme d'activation de la transcription par un métabolite a été observé. Le mécanisme d'activation de la traduction, établi par analogie, n'a pas encore été observé. Les sous-unités 30S et 50S du ribosome sont représentées par des disques bleu et orange, respectivement.

thiamine ou son dérivé pyrophosphate au niveau de l'aptamère, domaine conservé chez les deux organismes. La fixation de l'effecteur conduit à une terminaison à un site précoce de la transcription des ARNm chez B. subtilis (modèle Fig. 25A), tel que pour l'ARNm tenA, tandis que la fixation de l'effecteur conduit à une inhibition de l'initiation de la traduction chez E. coli via le changement de structure des ARNm au niveau de leur région TIR, séquestrant alors le site RBS (modèle Fig. 25C). Ce mécanisme de « riboswitch » conduit à une diminution de l'expression des gènes en fonction de la concentration du métabolite final de la voie de biosynthèse (Winkler et al., 2002).

# III.5 / Régulation de l'initiation de la traduction par les ARN non codants

De nombreux systèmes de mesure et de réponse aux variations de l'environnement utilisent des petits ARN non codants (ARNnc), de 40 à 400 nucléotides. L'avantage d'utiliser ce moyen de régulation par les ARNnc est de permettre une réponse rapide, et donc une adaptation rapide, dû fait de la petite taille de ces molécules d'ARN à synthétiser et de l'absence de traduction. Différentes approches bioinformatiques ont mis en évidence 1001 ARNnc au total chez *E. coli*, chaque approche détectant entre 180 à 250 ARNnc, mais seulement une cinquantaine d'ARNnc ont été identifiés par des approches biochimiques (Hershberg *et al.*, 2003; pour revue, Storz *et al.*, 2004). Ceux-ci peuvent être classés en trois groupes suivant leur mode d'action (pour revue, Storz *et al.*, 2004):

- association avec une protéine pour former un complexe actif tel que l'ARN M1 de la RNase P,
- association avec une protéine pour moduler l'activité de la protéine. L'ARNnc dont la structure mime une séquence cible d'une protéine, interagit avec cette protéine. Ainsi, l'ARNnc 6S ressemblant à un complexe ouvert est reconnu par l'holoenzyme Eσ<sup>70</sup>.
- appariement des ARNnc avec leur ARNm cibles pour modifier leur stabilité ou leur efficacité d'initiation de la traduction.

Le rôle des ARNnc sur la stabilité des ARNm sera décrit dans la prochaine partie.

La régulation au niveau de l'initiation de la traduction est réalisée soit par compétition avec le ribosome, soit par modification de la structure secondaire au niveau de la région TIR. En particulier, les ARNnc DsrA et RprA modifient la structure secondaire de la région d'initiation de la traduction de l'ARNm rpoS par l'intermédiaire de la protéine Hfq, ce qui facilite la traduction de l'ARNm rpoS (pour revue, Repoila et al., 2003). En effet, dans les conditions normales de croissance, la prédiction de structure secondaire au niveau du site RBS de l'ARNm rpoS montre que celui-ci serait inclus dans une tigeboucle, ce qui inhiberait la traduction (Fig. 26D). Les ARNnc RprA et DsrA stimulent la traduction de l'ARNm rpoS en réponse respectivement à une faible température et un stress au niveau des membranes de la cellule et un choc osmotique dans les deux cas (Fig. 26A). Ils se lieraient au niveau

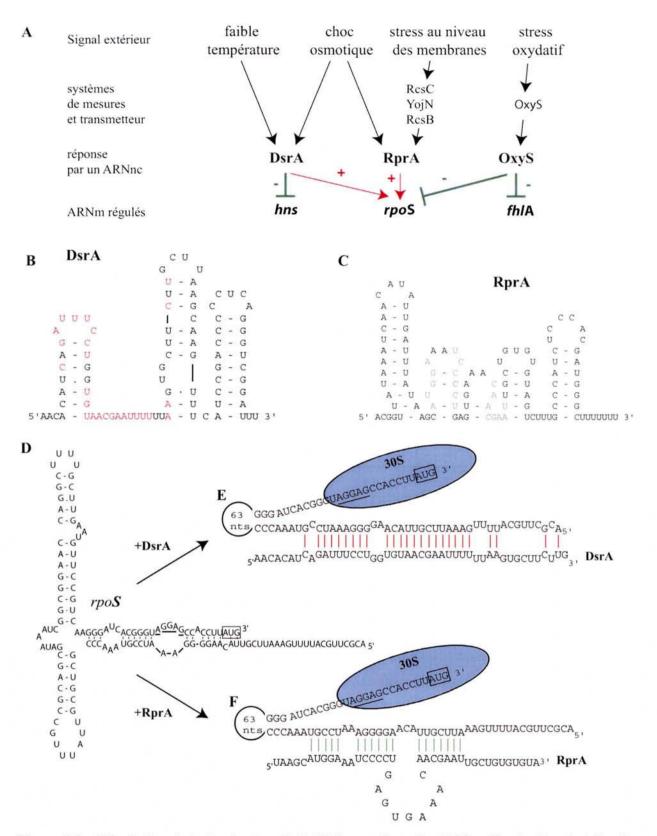


Figure 26: Régulation de la traduction de l'ARNm rpoS par les ARNnc DsrA, RprA et OxyS

A : Régulation de la transcription des ARNnc DsrA, RprA et OxyS en fonction des conditions de l'environnement et quelques gènes cibles correspondants (pour revue, Repoila et al., 2003).

B, C et D : Structure secondaire prédite des ARNnc DsrA et RprA, où les bases interagissant avec la région 5' non traduite de l'ARNm *rpoS* sont en couleur (d'après Lease et Belfort, 2000 et Majdalani *et al.*, 2002).

D : Structure secondaire prédite de l'ARNm rpoS (d'après Worhunsky et al., 2003).

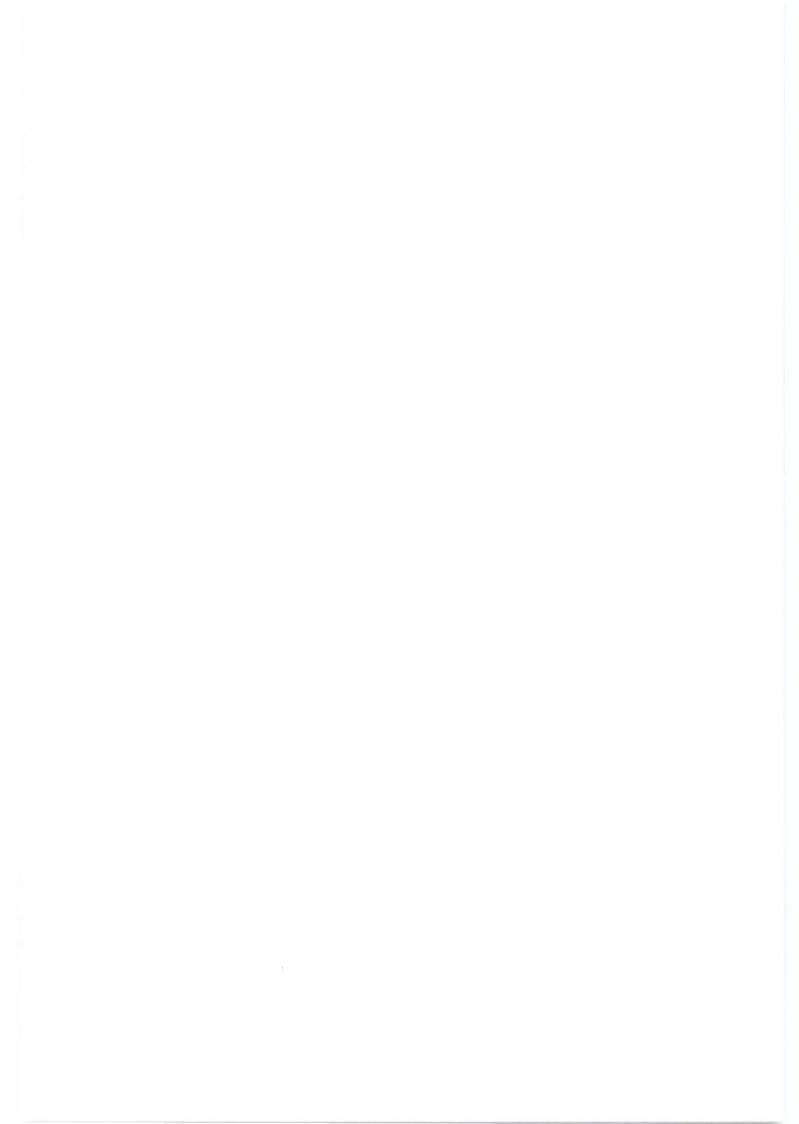
E: Appariement DsrA-rpoS (pour revue, Repoila et al., 2003).

F: Appariement RprA-rpoS (pour revue, Repoila et al., 2003). Les interactions entre les ARNnc DsrA et RprA avec l'ARNm rpoS sont symbolisées respectivement par des traits rouges et des traits verts.

de la région amont qui interagit avec la séquence SD de *rpoS*, levant ainsi l'inhibition de la traduction (Fig. 26, E et F).

L'ARNnc DsrA, est présent dans la cellule sous les formes F et T. Sous sa forme F de 87 nts, il interagit par son extrémité 5' à l'ARNm *rpoS* et se lie aussi avec la sous-unité 30S (Worhunsky *et al.*, 2003). Par contre, sa forme T (de 61 à 65 nts) est délétée des bases interagissant avec l'ARNm *rpoS*, mais elle peut toujours interagir et réguler *hns*, son autre ARNm cible. La forme F est stabilisée à 25 °C, contrairement à la forme T. Ainsi la variation des ratio F/T permettrait une régulation dépendante de la température pour *rpoS*, et indépendante pour *hns* (pour revue, Repoila *et al.*, 2003). L'ARNnc RprA, de 105 nt, est activé au niveau transcriptionnel, via le système phosphorelai RcsC/YojN/RcsB, en réponse à un stress localisé au niveau des membranes de la cellule (pour revue, Repoila *et al.*, 2003).

La traduction de l'ARNm *rpoS* est inhibée par l'ARNnc OxyS, de 109 nts, qui s'accumule après un stress oxydatif (Zhang *et al.*, 2002). L'inhibition de la traduction de l'ARNm *rpoS* pourrait être alors réalisée par une modulation de la quantité de la protéine Hfq disponible, ou bloquer l'accès des ARNnc activateurs au niveau de l'ARNm *rpoS* (pour revue, Repoila *et al.*, 2003).



# Partie D / La dégradation des ARNm

Après avoir été synthétisés, les ARNm sont traduits en protéines, puis dégradés. La stabilité des ARNm est un moyen supplémentaire de contrôle de l'expression des gènes. Pendant la croissance, la dégradation des ARNm bactériens est une activité métabolique majeure puisque 50 % des ARN synthétisés sont des ARNm, alors qu'ils représentent seulement 2 à 3 % de la totalité des ARN (pour revue, Kennell, 2002). La variation de stabilité des ARN permet un ajustement aux changements des conditions de croissance, comme un choc au froid (Yamanaka et Inouye, 2001), le développement d'un programme d'expression génétiquement contrôlé tel que le développement de l'infection des phages (pour revue, Carpousis et al., 1999; Grunberg-Manago, 1999), et une expression différentielle des gènes codés par les ARNm polycistroniques, (pour revues, Grunberg-Manago, 1999; Regnier et Arraiano, 2000).

La stabilité des ARNm est déterminée par leur temps de demi-vie (pour revue, Nierlich et Murakawa, 1996). La demi-vie moyenne des ARNm d'*E. coli* est de 7 minutes en phase exponentielle de croissance, à 37°C, mais peut varier de 30 secondes à 20 minutes selon les transcrits (Bernstein *et al.*, 2002; Selinger *et al.*, 2003, pour revue, Grunberg-Manago, 1999). Ces différences de stabilité des ARNm semblent fortement dépendantes de la structuration de leurs extrémités 5' et 3' et de la polyadénylation 3' terminale (pour revues, Nierlich et Murakawa, 1996; Rauhut et Klug, 1999).

La dégradation des ARNm est un processus complexe variant d'un transcrit à l'autre, dans lequel un nombre limité de ribonucléases (RNases) intervient. Ainsi, chez *E. coli*, parmi les dix-neufs RNases identifiées, la majorité est impliquée dans les mécanismes de maturation des ARN de transfert ou des ARN ribosomiques, seules 7 RNases sont impliquées dans la maturation et dégradation des ARNm (Tableau 2).

# I / Les RNases impliquées dans la dégradation des ARNm et leur régulation

Chez *E. coli*, la dégradation de la plupart sinon de tous les ARNm est initiée par un clivage endonucléolytique, souvent réalisé dans la région 5' des ARN, catalysé par la RNase E, la RNase G ou la RNase III (Cohen et McDowall, 1997; pour revues, Grunberg-Manago, 1999; Coburn et Mackie, 1999; Carpousis *et al.*, 1999; Steege, 2000; Kushner, 2002). Les produits de ce clivage deviennent ensuite des substrats pour d'autres clivages endoribonucléiques par la RNase E, aussi bien que pour la digestion par les exoribonucléases (exoRNases) 3'-5', la polynucléotide phosphorylase (PNPase), la RNase II et la RNase R (Donovan et Kushner, 1986; Cheng et Deutscher, 2005; pour revue Coburn et Mackie, 1999). Une fois le

|                                   | super-<br>famille | ribonucléase            | gène               | Mw<br>(kDa)           | délétion | Rôle   | Distribution dans d'autres organismes   | Références   |
|-----------------------------------|-------------------|-------------------------|--------------------|-----------------------|----------|--|---|--|
| E N D O R I B O N U C L E A S E S | RNase E/G         | RNase E                 | rne (ams,<br>hmp1) | 118 (dimère)          | létale   | Dégradation des ARNm, maturation de l'ARNr 16S et 23S ;<br>maturation de l'ARN M1 de la RNase P ; endoribonucléase spécifique<br>des régions simple brin riches en A et U. | subdivisions β et γ des <i>Proteobacteria</i> : un homologue<br>RNase E et RNase G; un seul homologue de la famille<br>RNase E/G présent dans environ la moitié des bactéries<br>séquencées | pour revue, Condon et<br>Putzer, 2002                        |
|                                   |                   | RNase G                 | rng (cafA)         |                       | viable   | Maturation extrémité 5' de l'ARNr16S; maturation des ARNm; endoribonucléase spécifique des régions simple brin riches en A et U.   |   |  |
|                                   | RNase III         | RNase III               | rnc                | 25 (dimère)           | viable   | Maturation et dégradation des ARNm et des ARNr ; endoribonucléase spécifique des régions doubles-brins   | eucaryotes ; bactéries sauf Deinococcus radiodurans   | pour revue, Condon et<br>Putzer, 2002                        |
|                                   | Rnase P#          | RNaseP/M1<br>RNase P/C5 | rnpA<br>rnpB       | 14                    | létale   | Maturation de l'extrémité 5° des ARNt ; maturation des ARNm  | eucaryotes, bacteries, Archeae à l'exception d'Aquifex<br>aeolicus  | Li et Altman, 2003 ; pou<br>revue, Condon et Putzer,<br>2002 |
|                                   | RNase M           | RNase M                 | rna                | 27                    | viable   | Ribonucléase périplasmique   | subdivisions α et β des Proteobacteria  | pour revue, Condon et<br>Putzer, 2002                        |
| E X O R I B O N U C L E A S E S   | RNase R           | RNase R                 | rnr (VacB)         | 95                    | Viable * | Dégradation de l'extrémité 3' des ARNr, ARNt et ARNm   | Superfamille retrouvée chez les eucaryotes, bactéries mais absente chez les Archeae à l'exception de Halobacterium  | Cheng et Deutscher,  |
|                                   |                   | RNase II                | rnb                | 72.3                  | Viable * | Dégradation extrémité 3' des ARNm et des ARNt (hydrolyse)  |   | 2002, 2005; pour revue,<br>Zuo et Deutscher, 2001            |
|                                   | PDX               | PNPase                  | pnp                | 77.1 kDa<br>(trimère) | Viable * | Dégradation extrémité 3' des ARNm et des ARNt (phosphorolyse) ;<br>activité poly(A)polymérase  | bacteries, à l'exception de Mycobacteria  |  |
|                                   |                   | RNase PH                | rph                | 50                    | Viable * | Maturation de l'extrémité 3' des ARNt (phosphorolyse)  | bactéries, à l'exception de Mycoplasma, Sprirochetes,<br>Chlamydia, et Helicobacter pylori; eucaryotes; Archaea à<br>l'exception de M. jannashii et Halobacterium NRC-1                     | pour revue, Zuo et<br>Deutscher, 2001                        |
|                                   | DEDD              | RNase D                 | rnd                | 43                    | Viable   | Maturation et dégradation des ARNt   | eucaryotes ; Mycobacteria et divisions α et γ des<br>Proteobacteria   |  |
|                                   |                   | RNase T                 | rnt                | 23                    | Viable   | Maturation de l'extrémité 3' des ARNr 5S et 23S ; dégradation du CCA des ARNt  | division γ des Proteobacteria   |  |
|                                   |                   | Oligo- ribonucléase     | orn                | 21                    | létale   | Dégradation 3'-5' des oligoribonucléotides   | eucaryotes ; Actinomycetes et divisions $\beta$ et $\gamma$ des<br>Proteobacteria   | Ghosh et al., 1999; pour<br>revue, Zuo et Deutscher,<br>2001 |
|                                   | RBN               | RNase BN                | rbn                | 33                    | Viable   | Maturation de l'extrémité 3' des ARNt  | bactéries   | pour revue, Zuo et<br>Deutscher, 2001                        |

Tableau 2 : Les différentes superfamilles de ribonucléases impliquées dans la dégradation des ARNm chez E. coli

<sup>\*</sup> Lorsque les mutations rnb, rph et rnr sont associées à une mutation pnp, le double mutant n'est pas viable. # La Rnase P est constituée d'un ARN catalytique M1 et de la protéine C5.

premier clivage initié, l'ARNm est souvent rapidement dégradé. Ce phénomène de tout ou rien évite la présence d'intermédiaires de dégradation qui pourraient conduire à la production de protéines tronquées (Spickler et al., 2001; pour revue, Coburn et Mackie, 1999).

En général, la dégradation apparente a lieu via une polarité 5' vers 3' (Selinger et al., 2003). A première vue, ce résultat est surprenant puisque aucune exoRNase dégradant l'ARN dans la direction 5' vers 3' n'a été trouvée chez les procaryotes (pour revue, Zuo et Deutscher, 2001). Cette polarité 5' vers 3' pourrait être due à la spécificité de la RNase E, qui requiert une extrémité 5' libre pour réaliser le premier clivage endonucléolytique aval. Le premier clivage réalisé dans la région 5' libère un fragment amont qui est dégradé par les exoRNases 3'-5' et un fragment aval qui possède une nouvelle extrémité 5' monophosphate reconnue préférentiellement par la RNase E et qui est clivé. Ce clivage endonucléolytique processif suivi par la dégradation par les exoRNases 3'-5' pourrait expliquer la polarité apparente de la dégradation des ARNm de 5' vers 3' (pour revue, Grunberg-Manago, 1999).

La dégradation des ARNm est liée à la traduction puisque les ribosomes peuvent inhiber directement la dégradation par les endoribonucléases (endoRNases), en occluant les sites de clivages. Ainsi, les ARNm sont stabilisés si leur traduction s'opère de façon active. La dégradation des ARNm est alors fonction de la fréquence d'initiation de la traduction (pour revue, Nicholson, 1999). Comme l'étape initiale implique un clivage spécifique d'un site situé dans la portion 5' proximal de l'ARNm, à proximité du site d'initiation de la traduction, la polarité 5'-3' des clivages endonucléolytiques ultérieurs pourrait refléter en partie le mouvement du dernier ribosome en cours de traduction.

Nous allons décrire tout d'abord, le rôle et le mécanisme d'action des différentes RNases impliquées dans la dégradation, puis les différents éléments intervenant dans la stabilité des ARNm.

### I.1 / La RNase E, la RNase G et le dégradosome

La dégradation des ARNm est généralement initiée par un clivage par la RNase E (pour revue, Grunberg-Manago, 1999). La purification de cette endoRNase a permis de montrer qu'elle interagit spécifiquement avec d'autres protéines pour former une macrostructure spécialisée dans la dégradation des ARNm: le dégradosome (pour revue, Carpousis et al., 1999). Nous décrirons tout d'abord la RNase E et son homologue, la RNase G, puis nous décrirons le dégradosome. Les différentes nucléases appartenant au dégradosome, autre que la RNase E, seront étudiées dans les parties suivantes.

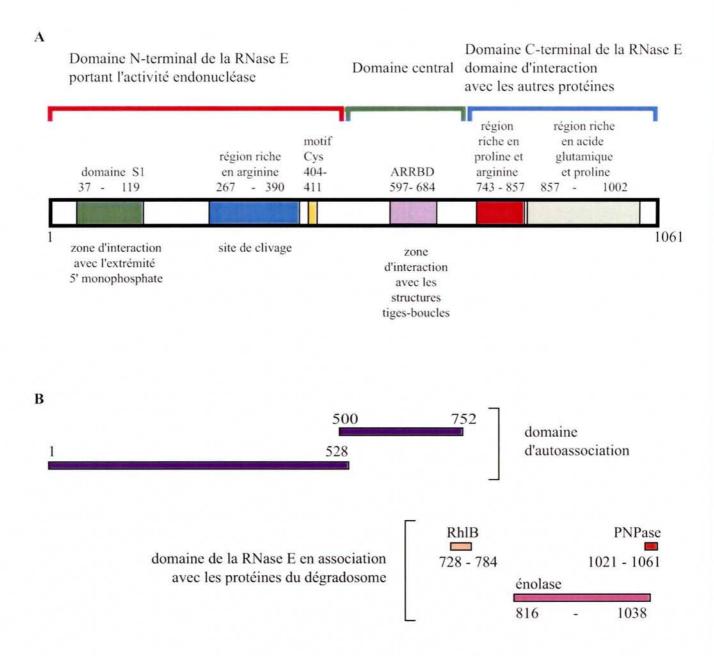


Figure 27: Les domaines structuraux de la RNase E et les régions d'interaction avec les protéines du dégradosome (d'après Grunberg-Manago, 1999, Spickler et al., 2001; Feng et al., 2002, Kaberdin et al., 2000, Callaghan et al., 2004)

A : Représentation schématique de la RNase E : les différentes régions des trois domaines sont indiquées ainsi que leurs positions et leurs fonctions potentielles.

B : Les régions impliquées dans l'autoassociation, ainsi que celles interagissant avec les protéines du dégradosome sont schématisées.

#### 1.1.a / La RNase E

#### I.1.a.α / Rôle de la RNase E

La RNase E, codée par le gène *rne*, est spécifique des régions simple brin. Elle a été identifiée initialement pour son intervention dans la maturation de l'ARNr 9S, précurseur de l'ARNr 5S (Fig. 29) (Ghora et Apirion, 1978; Misra et Apirion, 1979; Coburn et Mackie, 1999; Ehretsmann et al., 1992; Lin-Chao et al., 1994). Elle est aussi impliquée dans la maturation des ARNr 16S et 23 S (Fig. 29A) (Bessarab et al., 1998; Li et al., 1999). D'autres travaux ont montré que la RNase E a un rôle plus général dans le métabolisme des ARN, et est considérée comme la principale endoRNase intervenant dans la maturation et la dégradation des ARN messagers (pour revues, Coburn et Mackie, 1999 ; Grunberg-Manago, 1999 ; Regnier et Arraiano, 2000). Elle intervient dans le contrôle de la réplication des plasmides portant le réplicon ColE1 qui s'effectue via l'ARN régulateur RNA I (Fig. 38A) (Lin-Chao et Cohen, 1991), dans le clivage des queues poly (A) de transcrits (Huang et al., 1998; Walsh et al., 2001), dans la maturation de l'ARN M1, sous-unité catalytique ARN de la RNase P, (Lundberg et Altman, 1995), et dans la maturation des ARNt et de l'ARNtm (Li et Deutscher, 2002; Lin-Chao et al., 1999; Ow et Kushner, 2002; Ray et Apirion, 1981). Le caractère essentiel de cette enzyme pour la viabilité de la cellule a été attribué à son intervention dans la maturation des ARNt (Ow et Kushner, 2002), mais ceci n'a pas été confirmé (Deana et Belasco, 2004). Il est à remarquer que l'inactivation de la RNase E prolongerait la demi-vie chimique des ARNm totaux, sans pour autant avoir un effet sur leur dégradation fonctionnelle (Ono et Kuwano, 1979). E. coli nécessite au moins 10 % du taux de RNase E cellulaire pour être viable (Jain et al., 2002).

#### I.1.a.β / Structure de la RNase E

La RNase E semble être organisée sous forme d'un dimère de dimères, de 1061 acides aminés par sous-unité : les domaines N-terminaux s'associent en dimère de dimères autour de 2 ions Zn<sup>2+</sup>, alors que les domaines C-terminaux s'associent en dimères (Casaregola *et al.*, 1992, Casaregola *et al.*, 1994 ; Callaghan *et al.*, 2004 ; Callaghan *et al.*, 2005). Ainsi, les domaines N-terminaux se situeraient au centre du dégradosome, tandis que les domaines C-terminaux, situés à l'extérieur, recruteraient autour d'eux les autres composants du dégradosome (Fig. 32B). La formation du tétramère est requise pour l'activité catalytique de la RNase E mais chaque dimère peut fixer l'ARN simple brin.

La RNase E a un poids moléculaire de 118 kDa, mais a une migration électrophorétique apparente de 180 kDa sur gel de polyacrylamide SDS (Casaregola *et al.*, 1992, Casaregola *et al.*, 1994). Un monomère comprend trois domaines, le domaine N-terminal (acides aminés 1-528), le domaine central (500-752), et le domaine C-terminal (734-1061) (Fig. 27) (pour revue, Grunberg-Manago, 1999).

Le domaine N-terminal porte l'activité catalytique de l'enzyme au niveau des acides aminés 1 à 498 et les éléments structuraux nécessaires à sa fixation sur l'ARN et à la catalyse (McDowall et Cohen, 1996 ; Taraseviciene et al., 1995). Ce domaine N-terminal est lui-même décomposé en 3 domaines. Le domaine S1 (résidus 37-119), présentant une homologie avec chacun des 6 domaines répétés de la protéine ribosomale S1 (Bycroft et al., 1997), est retrouvé dans un grand nombre de protéines impliquées dans la traduction et dégradation des ARN, qui interagissent avec l'ARN en simple brin (Bycroft et al., 1997). La modélisation et la caractérisation structurale de ce domaine a permis d'identifier les résidus Phe 57, Phe 67 et Lys 112 comme importants pour la liaison à l'ARN simple brin et pour son autorégulation, résidus localisés en périphérie (Diwa et al., 2002 ; Schubert et al., 2004). Les résidus Lys 37 et Tyr 60, requis plus particulièrement pour l'autorégulation, sont localisés au niveau de la surface de dimérisation du domaine S1 (Schubert et al., 2004). Les deux mutations de la RNase E sensibles à la température G66S (rne-1=ams-1) et L68F (rne-3071), localisées dans ce domaine S1, entraînent une déstabilisation du domaine S1 seul, qui peut être compensée, seulement à température permissive, par des interactions additionnelles au niveau de la RNase E entière, (Schubert et al., 2004). Ce domaine S1 de liaison à l'ARN est suivi par une région riche en arginine (acides aminés 267-390), qui pourrait aussi interagir avec l'ARN (Kaberdin et al., 1998; Feng et al., 2002). Enfin, un motif cystéine conservé est localisé en aval (404-411; CPxCxGxG). Il forme une interface de dimérisation autour d'un ion zinc, et est nécessaire pour l'activité catalytique de l'enzyme (Callaghan et al., 2005).

Le domaine central de la RNase E comprend une région riche en arginine (ARRBD pour arginine-riche RNA binding domaine) (acides aminés 597-684), pouvant lier l'ARN in vitro (Kaberdin et al., 1996; McDowall et Cohen, 1996; Taraseviciene et al., 1995; Callaghan et al., 2004). Au contact de l'ARN, cette région se structurerait sous forme de fagots d'hélices (Callaghan et al., 2004). Elle interagirait spécifiquement avec l'ARN structuré, recrutant le domaine catalytique N-terminal. Ceci faciliterait le clivage à des séquences sous-optimales (Kaberdin et al., 2000; Lopez et al., 1999; Ow et Kushner, 2002). Ceci est confirmé par le fait que cette région n'est pas nécessaire ni pour l'autorégulation, ni pour la croissance cellulaire et ni à l'activité RNase (Diwa et al., 2002).

La partie C-terminale (acides aminés 734-1061), riche en proline et résidus acides, peu structurée, sert d'échafaudage, entre autre, pour la formation du dégradosome (Vanzo et al., 1998; Callaghan et al., 2004). Seule la région N-terminale semble être conservée dans une variété d'espèces bactériennes (Kaberdin et al., 1998; Lee et Cohen, 2003), ce qui pourrait suggérer que l'assemblage du dégradosome n'est pas essentiel à la dégradation des ARNm (Ow et al., 2000). Or, l'analyse par puces à ADN de la stabilité des transcrits montre une variation d'expression lorsque cette région C-terminale est délétée (Bernstein et al., 2004). Ainsi, la dégradation effective des ARNm requiert la partie C-terminale de la RNase E, nécessaire à la formation du dégradosome (Lee et al., 2003).

#### I.1.a.χ / Mécanisme de clivage par la RNase E

La RNase E reconnait ses sites de clivages par deux mécanismes (Fig. 28) (Marchand et al., 2001; Marujo et al., 2003; Baker et Mackie, 2003). Dans le premier cas, l'enzyme reconnaît l'extrémité 5' monophosphate du substrat, ce qui facilite la reconnaissance des sites de clivages en aval. Dans le deuxième cas, l'enzyme reconnaît un site de liaison interne de l'ARN. Ce mécanisme nécessite la partie C-terminale de la RNase E, contrairement au premier mécanisme, et peut se réaliser avec une RNase thermosensible. Dans ce cas, la dégradation est fonction des queues poly (A) situées en 3' des ARNm, elle est dite poly (A) dépendante (Marujo et al., 2003).

Le domaine catalytique de la RNase E est en contact avec son substrat via deux sites différents, pouvant être le domaine S1 (aa 36-118) et la région riche en arginine mineure (aa 267-390) de la RNase E, ces deux régions ayant des traits structuraux connus pour réaliser les interactions ARN/protéines. Le domaine S1 pourrait être le point de contact avec l'extrémité 5', la région riche en arginine mineure pourrait interagir avec la séquence cible du substrat (Fig. 28A) (Feng et al., 2002).

Une fois fixée sur l'ARN, la RNase E procéderait à une lecture de l'ARN du côté 3' vers le côté 5' pour trouver les sites de clivages. En effet, la présence de nucléotides 2'-O méthylés à un site de clivage permet la liaison du domaine catalytique, mais empêche le clivage au niveau de ce site et des sites d'hydrolyse amont (Feng et al., 2002). De même, l'ARNm lacZ, codant la β-galactosidase, muté de façon à n'être plus traduit mais possédant une séquence SD qui interagit avec la sous-unité 30S du ribosome, est dégradé en 3' de la séquence SD. Par contre, la région 5' en amont est fortement stabilisée (Joyce et Dreyfus, 1998). La présence de la sous-unité 30S inhibe alors la progression de la RNase E et des exoRNases.

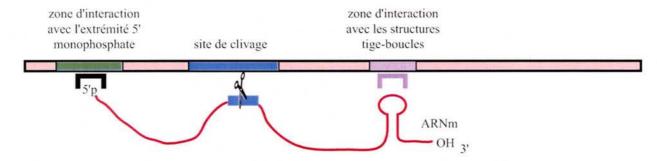
#### I.1.a.δ / Spécificité de coupure de la RNase E

Différents éléments interviennent dans la reconnaissance des sites de clivages par la RNase E, le clivage ayant lieu au moins à plus de 3 nucléotides de l'extrémité 3' de l'ARN (Walsh et al., 2001).

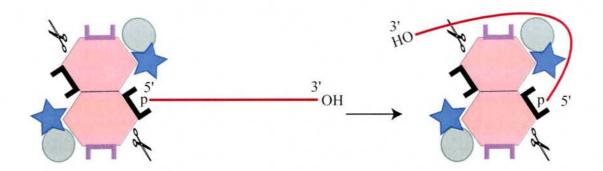
#### Influence de l'extrémité 5' sur le site de clivage

La région 5' des ARN influence l'efficacité d'hydrolyse par la RNase E, montrant une préférence pour les extrémités 5' monophosphorylées, puis triphosphorylées, sous forme simple brin, double brin et enfin circulaire (Baker et Mackie, 2003). En effet, la circularisation de l'ARNm rpsT, codant la protéine ribosomale S20, et l'ARNr 9S inhibe leur dégradation par la RNase E, qu'elle soit purifiée ou intégrée dans le dégradosome (Mackie, 1998). De même, la présence d'appariement en 5' de l'ARNm rpsT, due à la formation artificielle d'une structure tige-boucle ou la présence

#### A: Représentation schématique des fonctions de différents domaines de la RNase E



#### B : Mécanisme d'entrée au niveau de l'extrémité 5' monophosphate



### C : Mécanisme d'entrée interne

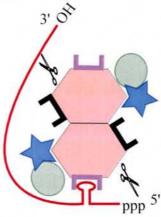


Figure 28 : Mécanisme de reconnaissance du substrat par la RNase E (d'après Marchand et al., 2001).

A : Représentation schématique de la RNase E avec les différents domaines d'interaction avec les régions 5' monophosphate et les structures tiges-boucles et la région catalytique.

B : Mécanisme de reconnaissance de l'ARNm par l'extrémité 5' monophosphate.

C : Mécanisme d'entrée de la RNase E au niveau d'un site interne de l'ARN.

L'ARN est représenté en rouge, la RNase E en orange, l'hélicase en vert et la PNPase en bleu. Les domaines d'interactions avec les régions 5' monophosphate et les structure tige-boucles sont représentés par un rectangle noir et violet, respectivement.

l'oligodésoxynucléotide complémentaire à la région 5' linéaire, inhibe d'autant la dégradation qu'un faible nombre de nucléotides reste non appariés en 5' (Mackie, 1998). Enfin, l'introduction d'un groupement triphosphate à l'extrémité 5' de l'ARNm *rpsT* et de l'ARNr 9S ralentit d'un facteur 20 à 30 l'hydrolyse *in vitro* par la RNase E, qu'elle soit ou non intégrée dans le dégradosome (Mackie, 1998).

Récemment, Jiang et Belasco ont montré que l'extrémité 5' monophosphorylée active l'efficacité catalytique de la RNase E, en particulier lorsque celle-ci est sous forme de multimère (Jiang et Belasco, 2004). Dans ce cas, le site actif d'une sous-unité du multimère accélère le clivage de l'ARN par le site actif d'une autre sous-unité. Ainsi, le clivage catalysé par une sous-unité génèrerait une nouvelle extrémité 5' monophosphate déjà liée à la RNase E, ce qui faciliterait ensuite le clivage en 3' de cet ARN : ceci augmenterait la processivité de la dégradation des ARNm (Jiang et Belasco, 2004).

La RNase E est aussi capable de cliver les queues poly (A) des ARN, par un clivage endonucléolytique (Walsh et al., 2001). Dans ce cas, la présence d'une extrémité 5' triphosphate ralentit au moins d'un facteur 30 la dégradation par le domaine N-terminal de la RNase E. Par contre, aucune différence significative n'est observée dans le cas d'une dégradation par le dégradosome (Walsh et al., 2001).

La reconnaissance de l'extrémité 5' par les différentes nucléases, comme la RNase E et la PAP pour agir en 3' de l'ARN peut être liée à la localisation rapprochée de ces deux extrémités. En effet, les programmes de prédiction de repliement de l'ARN ont tendance à prédire une structure secondaire de l'ARN de façon à ce que de longues interactions entre les extrémités 5' et 3' aient lieu (Schulz et Reznikoff, 1990).

#### Influence de la séquence sur le choix du site de coupure

La RNase E est une RNase spécifique des régions simple brin riches en A et U, précédées ou suivies de structure tige-boucles stables (Lin-Chao et al., 1994; McDowall et al., 1994; pour revues, Coburn et Mackie, 1999; Cohen et McDowall, 1997; Grunberg-Manago, 1999). La RNase E agirait aussi en tant qu'endoRNase des queues poly (A) et poly (U) in vitro (Huang et al., 1998) et in vivo (Walsh et al., 2001).

Des alignements de séquences hydrolysées par la RNase E révélaient l'existence d'une séquence consensus 5'-(A/G)AUU(A/U)-3' (Ehretsmann et al., 1992), où le nucléotide G situé un à deux nucléotides en amont du site de clivage, améliore l'efficacité de clivage (Kaberdin et al., 2000; Redko et al., 2003). L'utilisation d'oligomères a permis d'identifier un motif partiellement dégénéré reconnu par le domaine catalytique de la RNase E (1-498), où la flèche représente le site de clivage (Kaberdin, 2003) :

5'(G,A)(C,A)N(G)(G,U,A) 1 (A,U)(C,U)N(C,A)(C,A) 3'.

L'efficacité d'hydrolyse est alors proportionnelle à la similarité de la séquence testée à ce motif. La RNase E nécessite aussi des régions simple brin de part et d'autre du site de clivage : au moins 5-6 nucléotides en 5', et 3-4 nucléotides en 3'. L'augmentation de la longueur de la région en 5' du site de clivage sur un oligoribonucléotide améliore l'efficacité de clivage lorsque la séquence du site de

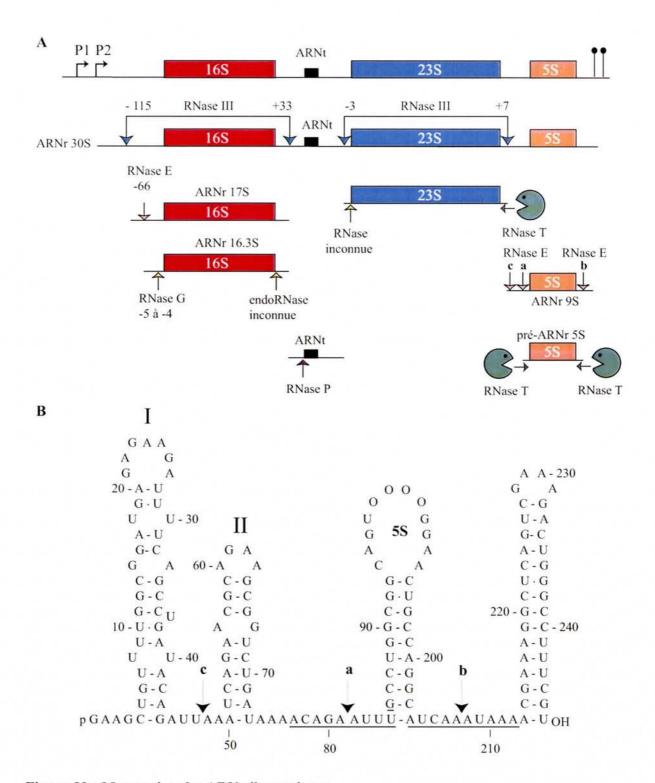


Figure 29: Maturation des ARN ribosomiques.

A: Modèle de maturation des ARN ribosomiques (d'après Li et al., 1999, a et b). Les ARNr transcrits à partir des deux promoteurs P1 et P2 sont clivés par la RNase III, puis par différentes RNases pour conduire à la formation des ARNr 16S, 23S et 5S matures. Un seul clivage par la RNase III permet d'individualiser les pré-ARNr 16S et 23S. Les endoribonucléases sont indiquées par des flèches, les exoribonucléases par des "pacman". Les positions indiquent le clivage par rapport aux positions 5' et 3' de l'ARNr 16S et l'ARNr 23S.

B : Structure secondaire du pré-ARNr 5S (d'après Cormack et Mackie, 1992). Seules, les structures tigeboucles I et II ont été démontrées expérimentalement. Les flèches indiquent les sites principaux de clivage par la RNase E. Les clivages aux sites **a** et **b** sont à l'origine du pré-ARNr 5S. Les nucléotides sont numérotés par rapport au site +1 d'hydrolyse par la RNase III. La séquence de l'ARNr 5S n'est pas indiquée. Les séquences soulignées, contenant les sites de clivages **a** et **b**, correspondent aux oligoribonucléotides 9Sa et 9Sb, respectivement (Kaberdin *et al.*, 2000).

clivage est dégénérée par rapport à la séquence consensus, telle que pour des séquences de poly (A) ou poly (U) (Kaberdin, 2003).

#### Influence des structures tiges boucles adjacentes au site de clivage

Les sites de clivages de la RNase E sont situés à proximité de structures en tige-boucle adjacentes. Celles-ci peuvent améliorer l'efficacité de clivage soit en maintenant la conformation simple brin de la région à cliver, soit en recrutant le domaine ARRBD de la RNase E, qui réalise une lecture de l'ARN jusqu'au site de clivage, comme pour la maturation de l'ARNr 5S (Kaberdin et al., 2000; pour revue, Carpousis et al., 1999). En effet, les ARN ribosomiques sont transcrits à partir de deux promoteurs P1 et P2, pour chacun des 7 opérons que contient *E. coli* (Fig. 29A). Les premiers clivages à la RNase III individualisent les pré-ARNr 5S, 16S et 23S. La synthèse de l'ARNr 5S requiert un clivage de l'ARNr 9S par la RNase E aux sites a et b (Fig. 29). *In vitro*, le clivage des oligoribonucléotides correspondant aux régions a et b ne sont pas équivalents, bien que ces deux régions soient simple-brin. L'ajout du nucléotide G, 1 à 2 nucléotide en 5' du site de clivage b, rapprochant le site b du motif consensus, augmente l'efficacité de clivage par la RNase E, en absence des structures adjacentes. Les structures tige-boucles situées de part et d'autre du site b recrutent la région centrale ARRBD de la RNase E. Ces structures adjacentes améliorent donc le clivage au niveau du site b sous-optimal de l'ARNr 9S.

Les structures tiges-boucles adjacentes au site de clivage peuvent inhiber le clivage par la RNase E. Par exemple, les oligoribonucléotides de 10 et 13 mers, contenant le site de clivage à la RNase E de RNA I, voient leur efficacité de clivage multipliée par 20 en absence des structures tiges boucles adjacentes.

#### I.1.a.ε / Régulation de la synthèse de la RNase E

#### Régulation de l'expression de la RNase E

La RNase E étant une enzyme essentielle, des variations de sa production peuvent empêcher la croissance cellulaire (Diwa et al., 2000). La RNase E autorégule sa synthèse afin de maintenir une concentration optimale. Ce mécanisme d'autorégulation dépend de façon générale du taux de transcrits polyadénylés intracellulaires (Mohanty et Kushner, 2002) et nécessite que le dégradosome soit intact pour avoir une autorégulation efficace (Jiang et al., 2000; Mohanty et Kushner, 2002).

La RNase E est exprimée majoritairement à partir du promoteur p1 situé 361 nucléotides en amont du codon initiateur, et par les promoteurs p2 et p3 situés 34 et 145 nucléotides en amont de p1 (Ow et al., 2002). La région 5' non traduite est requise pour une autorégulation correcte (Diwa et Belasco, 2002). L'ARNm issu du promoteur p1 comporte 3 structure tige-boucles, hp1, hp2 et hp3, et trois

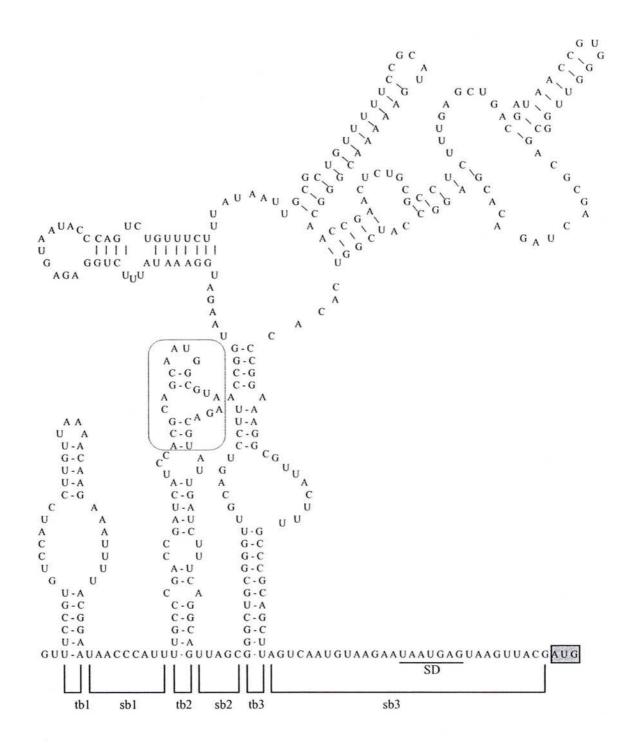


Figure 30 : Structure secondaire de la région 5' terminale non traduite de l'ARNm *rne* transcrit à partir du promoteur p1 (d'après Diwa *et al.*, 2002).

Les crochets délimitent les six domaines structuraux tb1, sb1, tb2, sb2, tb3, sb3. La séquence SD est soulignée, le codon initiateur entouré d'un rectangle grisé. La partie de la tige-boucle tb2 requise pour l'autorégulation est encadrée.

séquences simple brin, en aval de chaque structure tige-boucle (Fig. 30), où les structures tiges-boucles hp2 et hp3 sont très importantes pour l'autorégulation, en particulier hp2. La fusion de hp2 seule en amont d'un site de liaison au ribosome du transcrit *rne* est suffisante pour lui conférer une autorégulation. La boucle interne, située en amont de hp2 recruterait soit la RNase E soit un cofacteur protéique associé qui facilite l'accès de la RNase E à un ou plusieurs sites de clivages situés sur le transcrit *rne* (Diwa et Belasco, 2002).

# Régulation de l'activité de la RNase E

Différents effecteurs modulent l'activité de la RNase E. Lors d'une infection par le bactériophage T7, la protéine kinase à sérine et thréonine, du domaine N-terminal du produit du gène 0.7, phosphoryle le domaine C-terminal de la RNase E, ce qui stabilise un ensemble de substrat à la RNase E (Marchand et al., 2001). Récemment, il a été montré que l'activité RNase E est modulée par l'intermédiaire de l'expression des gènes rraA et rraB. La surexpression du gène rraA (regulator of ribonuclease activity A) inhibe la RNase E, via une interaction avec le domaine N-terminal de l'enzyme, et semble favoriser sa dégradation protéolytique, ce qui diminue la concentration de RNase E (Lee et al., 2003). L'inhibition de la RNase E engendrée par RraA dépend aussi des substrats ARN. Ainsi, la protéine RraA entraîne une stabilisation de certains substrats à la RNase E, ce qui conduit à une modification de l'expression des gènes en fonction d'un stimuli de l'environnement (Lee et al., 2003).

La température et le milieu de culture affectent aussi la dégradation par la RNase E : la stabilité du transcrit *rpsO* est accrue dans un milieu minimum à 44°C après un shift de 15 min, par rapport à un milieu riche à la même température, ou un milieu riche ou minimum à 30°C (Le Derout *et al.*, 2002). Cette modulation d'activité de la RNase E pourrait résulter de la modulation de l'expression du facteur RraA. La physiologie de la bactérie semble aussi influencer l'activité de la RNase E. En particulier, l'anaérobiose affecte la maturation des ARN par la RNase E (Georgellis *et al.*, 1993).

#### I.1.a. / Les homologues de la RNase E

Des homologues de séquences à la RNase E sont présents dans une variété d'espèces bactériennes, d'Archaea et de plantes, et dans les cellules humaines, la plus forte homologie résidant dans la partie N-terminale de la RNase E portant l'activité catalytique (Condon et al., 2001 ; Lee et Cohen, 2003). Chez les plantes, les protéines homologues seraient situées dans le chloroplaste, organelle d'origine bactérienne. La protéine ARD-1 des cellules humaines, partageant des traits structuraux avec la RNase E, peut permettre la survie de souches d'E. coli déficientes en RNase E (Wang et Cohen, 1994).

Les protéines homologues peuvent être divisées en 4 groupes, suivant la position de la possible région catalytique, et la présence ou l'absence de régions contenant des séquences riches en proline,

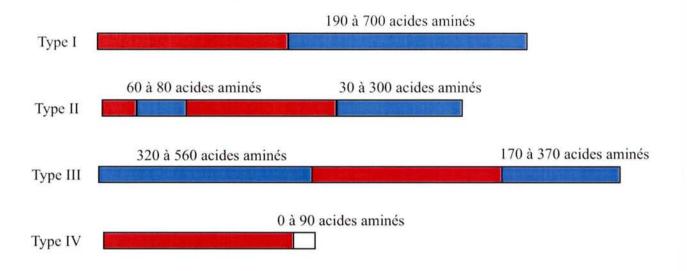


Figure 31 : Classification des homologues de la RNase E/G (d'après Lee et Cohen, 2003).

Les domaines catalytiques potentiels des homologues de la RNase E/G sont représentés en rouges, les domaines pouvant servir d'échafaudage en bleu. Les différentes tailles possibles de ces domaines sont indiquées.

arginine et résidus acides, qui peuvent servir d'échafaudage au dégradosome (Fig. 31) (Lee et Cohen, 2003). Dans les enzymes de type I telles que la RNase E d'*E. coli*, le domaine catalytique est localisé en N-terminal, et le domaine d'échafaudage en C-terminal,. Dans les enzymes de type II, présentes chez les protéobactéries α, la partie d'échafaudage en C-terminale est raccourcie, et le segment riche en proline, arginine et résidu acide interrompt le domaine S1 de la région catalytique (Lee et Cohen, 2003). Les enzymes de type III, identifiées principalement chez les Actinobacter, contiennent la région catalytique au centre de la protéine, comme la RNase ES de *Streptomyces coelicolor*. Les enzymes de type IV n'ont pas de région pouvant servir d'échafaudage, comme la RNase G d'*E. coli*, et ont une taille de l'ordre de 375 à 515 acides aminés qui est inférieure aux autres protéines des autres groupes, de 575 à 1340 acides aminés. Elles sont présentes dans les subdivisions β et γ des protéobactéries, dont les entérobactéries. Aucune protéine ressemblant à la RNase G n'a été retrouvée dans les génomes de bactéries séquencées contenant la RNase E de type II et III. Ainsi, les types I, II et III des homologues de RNase E seraient orthologues, alors que la RNase G serait un paralogue de la RNase E (Lee et Cohen, 2003).

Lorsque les bactéries ne présentent pas d'homologues à la RNase E, elles semblent posséder une RNase M5. C'est le cas du groupe des bactéries à Gram positif à faible G+C, telles que B. subtilis et la spirochète Borrelia burgdorferi (Condon et al., 2001). Cette endoRNase intervient dans la maturation de l'ARNr 5S, au niveau de région double-brin. Dans ce cas, la maturation de l'ARNr 5S est réalisée via un clivage unique réalisé par la RNase M5, alors que chez E. coli, elle comprend un clivage à la RNase E suivi d'une dégradation exoribonucléique par la RNase T (Condon et al., 2001). La présence de l'un ou de l'autre système de maturation de l'ARNr 5S est corrélée à la nature simple brin ou double brin du site de clivage de la molécule précurseur. Quelques bactéries possèdent à la fois un homologue à la RNase E et une RNase M5, comme Clostridium difficile ou Clostridium acetobutylicum, où les différents précurseurs d'ARNr 5S peuvent être clivés par les deux types d'enzymes, alors que d'autres organismes, en particulier le groupe des Archaea, n'ont ni l'une ni l'autre.

#### I.1.b / La RNase G

#### I.1.b.α / Rôle de la RNase G

Originellement découverte comme le produit du gène cafA (cytoplasmic axial filament), dont la surexpression induit la formation de filaments axiaux cytoplasmique, cette endoRNase présente une forte similarité de séquence aux 498 premiers acides aminés de la RNase E (49.5% de similarité de séquence et 34.1% d'identité de séquence en acides aminés (McDowall et al., 1993; Li et al., 1999; Okada et al., 1994; Wachi et al., 1997). Elle présente des mécanismes catalytiques similaires à la RNase E (Tock et

al., 2000; Jiang et al., 2000), mais sa délétion n'est pas létale contrairement à la RNase E (Li et al., 1999; Wachi et al., 1999). Elle n'a qu'un rôle limité dans la maturation des ARN, tel que la maturation en 5' de l'ARNr 16S (Li et al., 1999; Wachi et al., 1999), et dans la dégradation des ARNm, tel que les ARNm eno (Kaga et al., 2002), adhE (Umitsuki et al., 2001). La maturation de 11 transcrits d'E. coli sur 4405 transcrits analysés par puces à ADN est affectée par l'inactivation de la RNase G (Lee et al., 2002). La majorité de ces transcrits sont préalablement clivés par la RNase E ou la RNase III, comme les transcrits polycistroniques ndh, panB, ahpF et les transcrits monocistroniques eno et adhE. Ainsi, la RNase G ne pourrait dégrader les ARNm, que s'ils sont précédemment clivés (Ow et al., 2003). Ceci pourrait s'expliquer par l'absence de la partie C-terminale de la RNase E qui ne permet pas à la RNase G d'utiliser le mécanisme d'entrée interne de la RNase E. Le rôle secondaire dans la dégradation des ARNm attribué à la RNase G est aussi lié à sa faible concentration intracellulaire [inférieure à 5 % de la RNase E (Lee et al., 2002)], mais majoritairement à sa spécificité de séquence. En effet, la surproduction de RNase G plus de 100 fois, ne peut compenser l'absence de RNase E in vivo. Par contre, la complémentation est réalisée par la surproduction d'un variant de la RNase G ayant une extension soit N-terminal, telle que MRKGINM, soit C-terminale, telle que GHHHHHH. L'extension GHis<sub>6</sub> déstabiliserait la structure tertiaire de l'enzyme, puisque la vitesse de sédimentation est plus lente (Briant et al., 2003). La variation de la structure tertiaire de la RNase G lui permettrait alors de cliver des ARN cellulaires importants pour la viabilité de la bactérie, dont l'efficacité de maturation ou de dégradation requiert la RNase E (Deana et Belasco, 2004).

# I.1.b.β / Structure de la RNase G

La RNase G existe principalement sous forme de dimère en solution, en équilibre avec un petit pool de monomère et tétramère et multimère plus important, le poids moléculaire de son monomère est de 55 kDa (489 aa) (Briant *et al.*, 2003 ; Deana et Belasco, 2004). Elle est active sous forme oligomérique.

#### I.1.b.γ / Spécificité de coupure de la RNase G

La RNase G clive les régions riches en A-U simple brin seulement si l'extrémité 5' de l'ARN est monophosphorylée (Jiang et al., 2000 ; Tock et al., 2000). Cette préférence de clivage des substrats monophosphorylés serait une propriété conservée au cours de l'évolution par la partie N-terminale de la RNase E, puisqu'elle est retrouvée aussi bien pour la RNase G que pour SynRne, l'homologue de la RNase E de Synechocystis (Jiang et al., 2000). Sa fonction se superpose à la RNase E, puisque la RNase G clive l'ARN RNA I et l'ARNm ompA aux sites reconnus par la RNase E. Par contre, la RNase G ne clive pas le pré-ARNr 5S au site a, sauf si la séquence est isolée sur un décanucléotide. Ainsi, la

RNase G est inhibée par la structure secondaire des ARNm (Tock *et al.*, 2000). Elle ne montre pas de directionalité de clivages, pour les mêmes substrats que la RNase E (Feng *et al.*, 2002).

# I.1.b.δ / Régulation de la RNase G

La RNase G est régulée au niveau de la quantité de transcrit par la RNase E : la délétion de la RNase E multiplie par 4 la quantité de transcrits *rng* codant la RNase G (Lee *et al.*, 2002) et de protéines RNase G. L'activité de la RNase G, contrairement à la RNase E, n'est pas régulée par la protéine RraA (Lee *et al.*, 2003).

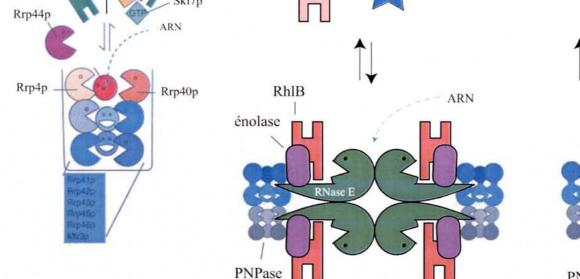
#### I.1.c / le dégradosome

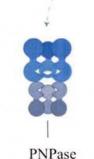
La purification de la RNase E a permis de montrer qu'elle interagit spécifiquement avec d'autres protéines pour former une macrostructure spécialisée dans la dégradation des ARN: le dégradosome (Fig. 32B). Ce complexe contient les composants majeurs suivants : la RNase E, la PNPase, l'hélicase RhlB et l'énolase. L'hélicase RhlB, de la famille des protéines à boîte DEAD, désapparie les structures tige-boucle en consommant de l'ATP, ce qui facilite la dégradation des substrats structurés par la PNPase (pour revue, Rocak et Linder, 2004). L'hélicase RhlB semble faciliter la dégradation des ARNm sans dégradosome par la RNase E (Khemici et al., 2005). Elle nécessite d'interagir avec la RNase E pour présenter une activité ATPase (Py et al., 1996; Vanzo et al., 1998). La PAP est aussi capable d'interagir avec la RNase E (Raynal et Carpousis, 1999). La PNPase, la PAP et la RNase E sont toutes les trois affectées positivement par la phosphorylation en 5' des ARNm, bien que leur action se situent à l'extrémité 3' des ARNm. Ainsi, l'extrémité 3' du transcrit pourrait être apportée à proximité des sites de clivages par les protéines du dégradosome (Feng et al., 2002). L'énolase, enzyme de la glycolyse générant de l'ATP, serait fixée sous formes de dimères au niveau de la région 816-1038 de la RNase E (Callaghan et al., 2004). Elle intervient aussi dans la dégradation des ARN (Bernstein et al., 2004; Morita et al., 2004b). En particulier, sous des conditions de stress liées à l'accumulation d'un sucre phosphaté, l'ARNm ptsG est rapidement dégradé. Cette régulation posttranscriptionnelle nécessite la partie C-terminale de la RNase E et l'énolase. Le rôle de l'énolase n'est cependant pas encore connu.

D'autres protéines sont retrouvées en quantité sous-stoechiométrique dans le dégradosome telles que les protéines chaperones DnaK et GroEL, la polyphosphate kinase (PPK), l'hélicase à boîte DEAD CsdA (pour revues, Carpousis et al., 1999; Symmons et al., 2002). Les protéines chaperones DnaK et GroEL, connues pour faciliter le repliement des protéines, pourraient avoir un rôle dans l'assemblage du dégradosome (Miczak et al., 1996; Blum et al., 1997). La PPK, qui catalyse réversiblement la formation de polyphosphate à partir d'ATP, pourrait réguler l'activité du dégradosome par l'intermédiaire de la

B

A





ARN

C

Figure 32 : Comparaison de l'exosome, du dégradosome et du complexe RNP100 du chloroplaste (Callaghan et al., 2004 ; pour revue, Symmons et al., 2002).

Les sous-unités exoRNase sont indiquées par des cercles, avec des bouches de "pacman", les activités phosphorolytiques sont indiquées en bleu, les RNases hydrolases par les autres couleurs. La position de l'ARN suggérée dans le complexe est indiquée par une ligne en pointillés. Les hélicases sont indiquées par une figure en forme de H, les autres composants sont indiqués par des boîtes.

A : L'exosome de *S. cerevisiae* : son noyau comprend 6 exoRNases phosphorolytiques (Rrp441p, Rrp42p, Rrp43p, Rrp45p, Rrp46p et Mtr3p), deux exoRNases hydrolytiques (Rrp4p et Rrp43p) et la protéine Csl4p d'activité inconnue. L'exonucléase hydrolytique Rrp44p s'associe à l'exosome à la fois du noyau et du cytoplasme, soit directement à l'exosome, soit par l'intermédiaire de l'ARN. La partie supérieure du dessin indique les composants spécifiques de l'exosome du noyau et du cytoplasme.

B : Le dégradosome d'*E. coli* : la partie C-terminale de la RNase E permet l'association avec les autres composants du dégradosome. La copie des composants formant le dimère est indiquée en grisé. Il comprend les composants additionnels tels que la polyphosphate kinase (PPK), la poly (A) polymérase, et la protéine CsdA.

C : Le complexe exoRNase RNP100 de chloroplaste : le noyau est composé d'un dimère de trimères de PNPase. Une activité endoRNase associée avec l'activité exoRNase a été décrite, mais ne proviendrait pas de la protéine p67 isolée au départ.

concentration de polyphosphates (Blum et al., 1997). L'hélicase CsdA, induite lors d'un choc au froid, interagit avec la RNase E au niveau d'un site différent de l'hélicase RhlB, avec une plus faible affinité, et peut se substituer à l'hélicase RhlB au niveau du dégradosome in vitro (Khemici et Carpousis, 2004; Prud'homme-Généreux et al., 2004). In vivo, le dégradosome pourrait contenir soit l'une ou l'autre hélicase, soit les deux types d'hélicases comme pour le dégradosome de Rhodobacter capsulatus (Jager et al., 2001). La RNase E est aussi capable d'interagir avec la protéine ribosomale S1 (Feng et al., 2001). Le dégradosome contient aussi, des ARNm en cours de dégradation, de courts fragments d'ARN stables, fortement structurés. Ceci pourrait impliquer un contrôle qualité du dégradosome pour la biogenèse d'ARN stables.

Par différentes approches de microscopie électronique, il a été montré que le dégradosome existe *in vivo*, et est associé à la membrane, via la partie N-terminal de la RNase E (Liou *et al.*, 2001). La PNPase et l'énolase étant en large excès, vis-à-vis de la RNase E, ces protéines sont retrouvées majoritairement sous forme non liées à la RNase E. Par contre, l'hélicase RhlB est présente en quantité quasi équimolaire à la RNase E, à tous les stades de croissance (Liou *et al.*, 2001). L'hélicase RhlB se lie plus fortement à la RNase E et à elle-même pour former des homodimères qu'à la PNPase (Liou *et al.*, 2002). Ceci est en faveur de la formation du dégradosome par rapport aux mini-machines composées de l'association RhlB/PNPase capables de dégrader les intermédiaires de dégradation qui ont des structures secondaires très stables. De plus, dans un mutant *pcnB*, les mutations de la RNase E interrompant l'interaction avec soit l'hélicase RhlB, soit la PNPase, a les mêmes effets que les mutants délétés de l'hélicase RhlB ou de la PNPase sur la dégradation de la séquence REP de *malE*. Ceci conforte l'idée que l'hélicase RhlB et la PNPase agissent ensemble dans le dégradosome, *in vivo* (Khemici et Carpousis, 2004).

L'analyse de la stabilité des ARN totaux d'E. coli par puces à ADN montre que les parties C-terminales de la RNase E, la PNPase, l'énolase et l'hélicase RhlB sont requises pour une dégradation normale des ARN, puisque la mutation de chacun de ces éléments affecte différemment la stabilité des ARN individuels (Bernstein et al., 2004). La PNPase et la partie C-terminale semblent avoir des actions combinées dans la dégradation des ARNm impliqués dans la glycolyse. Ainsi, la délétion de la PNPase conduit à une stabilisation d'un facteur 2 des ARNm pgk et fbaA, et la délétion de la partie C-terminale conduit à une stabilisation, moins importante que celle observée lors de la délétion de la PNPase, des ARNm gapA, epd et fbaA. Ces deux éléments affectent aussi la stabilité de l'ARNm eno codant l'énolase, enzyme qui appartient aussi au dégradosome. Ainsi, la stabilité de l'ARNm eno intègre les activités de la RNase E et du dégradosome d'un côté, et celle de la RNase G de l'autre côté. La délétion de l'énolase conduit à une stabilisation des transcrits impliqués dans l'entrée et le métabolisme de nombreuses sources de carbone, comme le glycérol, le galactitol, le mannose, le maltose et le fructose (Bernstein et al., 2004). De façon surprenante, ces transcrits sont déstabilisés par la

délétion de la partie C-terminale de la RNase E. Ainsi, la dégradation de ces transcrits ne serait pas réalisée de façon coordonnée par les différents éléments du dégradosome.

Une structure ressemblant au dégradosome, sans PNPase, a été mise en évidence chez Rhodobacter capsulatus (Jager et al., 2001). Chez Streptomyces coelicolor, l'homologue RNase ES possède des segments riches en proline et arginine, ressemblant à la partie C-terminale de la RNase E qui interagissent avec l'homologue de la PNPase de S. coelicolor. Ceci suggère l'existence de complexes ressemblant au dégradosome dans les bactéries à Gram positif (Lee et Cohen, 2003).

Différents complexes protéiques, appelés exosomes, spécialisés dans la dégradation des ARN de 3' vers 5', ont également été mis en évidence dans les cellules eucaryotes (pour revues, Carpousis et al., 1999; Symmons et al., 2002). Les exosomes du cytoplasme et du noyau, isolés de *S. cerevisiae*, comprennent principalement des exoRNases (Fig. 32A), de même que le complexe mitochondrial mtEXO (pour revue, Carpousis et al., 1999; Symmons et al., 2002). Des composants équivalents à l'exosome de *S. cerevisiae* ont été découverts dans les cellules de mammifères (Chen et al., 2001).

Le complexe protéique mis en évidence dans le chloroplaste d'épinard serait constitué d'un dimère de trimère de PNPase avec peut-être une association réversible avec une endoRNase (Fig. 32C). Différentes expériences de modélisation suggèrent que ces exosomes utilisent la PNPase comme structure d'échafaudage au lieu de la RNase E (Aloy et al., 2002).

# I.2 / La RNase III

La RNase III réalise aussi l'initiation de la dégradation des ARN chez E. coli.

#### I.2.a / Rôle de la RNase III

Les protéines de la famille RNase III sont des endoRNases qui clivent les régions d'ARN double brin (ARNdb). Elles interviennent dans la maturation des ARN et le mécanisme d'ARN interférence, entre autre (Lamontagne et Elela, 2004; pour revues, Zamore, 2001; Carmell et Hannon, 2004). Elles sont largement conservées chez les bactéries et chez les eucaryotes mais aucun homologue n'a été retrouvé chez les Archaea (Tableau 2) (pour revues, Conrad et Rauhut, 2002; Condon et Putzer, 2002). Les membres de la famille RNase III sont divisés en trois classes par le nombre de domaines endoRNases qu'ils possèdent, et par la présence de domaines accessoires qui modulent leur activité (pour revue, Carmell et Hannon, 2004). Ainsi, les RNases III procaryotes ne contiendraient que les enzymes de classe I, c'est-à-dire un seul domaine endoRNase lié à un domaine de liaison à l'ARN double brin (Fig. 33). Les enzymes de classes II, telle que Drosha, présentes chez les vertébrés, possèdent deux domaines

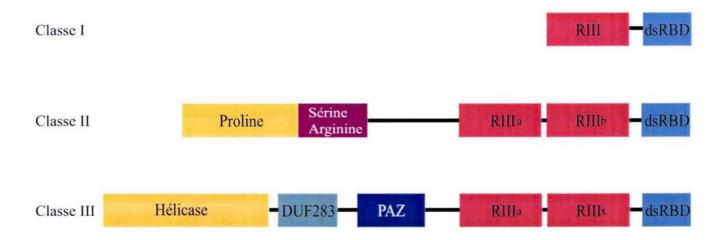


Figure 33 : Les différentes classes d'homologues de la RNase III et leurs différents domaines (d'après Carmell et Hannon, 2004).

Les domaines dsRBD sont les domaines de fixation à l'ARNdb, Sérine Arginine : le domaine riche en arginine et sérine, Proline : le domaine riche en proline, et RIII : les domaines endonucléases. Le site catalytique inactif est symbolisé par x. Le domaine DUF283 est un domaine de fonction inconnue, tandis que le domaine PAZ est impliqué dans le mécanisme d'ARN interférence.

endoRNases, un domaine de liaison à l'ARNdb en C-terminal, et une extension en N-terminale. Chez l'homme et la souris, cette extension de l'enzyme contient deux domaines pouvant être impliqués dans des interactions protéines-protéines: le domaine riche en proline et celui riche en sérine et arginine. Les enzymes de classe II et III sont impliquées dans le mécanisme d'ARN interférence. Les enzymes de classe III, d'environ 200 kDa, se distinguent des enzymes de classes II, par la présence supplémentaire d'un domaine hélicase en N-terminal à boîte DEXH, un domaine de fonction inconnue DUF283, et un domaine PAZ. Ce domaine PAZ doit son nom aux 3 protéines Piwi, Argonaute et Zwille, de la famille de protéines Argonautes impliquées dans les complexes d'ARN interférence (pour revue, Carmell et Hannon, 2004). Il est à noté qu'un seul des deux domaines endoRNases est actif dans cette classe d'enzyme.

Chez *E. coli*, la RNase III (EC 3.1.24) est codée par le gène *rnc*. C'est un homodimère composé de 2 sous-unités de 226 acides aminés et 25.6 kDa (pour revue, Conrad et Rauhut, 2002). Sa principale fonction est la maturation des ARNr. Elle intervient aussi dans la maturation des ARNt, et des ARNm, d'ARN phagiques, et dans les mécanismes de régulation du nombre de copies de certains plasmides (pour revue, Conrad et Rauhut, 2002). Son inactivation n'affecte pas la demi-vie des ARNm totaux (Babitzke *et al.*, 1993), mais modifie le taux d'environ 10 % des protéines détectées par gel d'électrophorèse (Gitelman et Apirion, 1980). Sa délétion entraîne néanmoins un ralentissement de la vitesse de croissance. Ceci serait dû à une diminution de la capacité de synthèse protéique, liée à la présence d'ARNr 23S incomplètement maturés dans les sous-unité 50S des ribosomes (pour revue, Nicholson, 1999). Il a aussi été montré que la RNase III exerce une fonction régulatrice dans la cellule, à travers un substrat particulier, sans clivage de l'ARN comme cela l'a été observé dans le cas du gène codant la protéine cIII du phage λ (Altuvia *et al.*, 1987). La bactérie *E. coli* contient environ une centaine de dimère de RNase III par cellule, qui s'associeraient aux ribosomes (pour revue, Nicholson, 1999).

# I.2.b / Spécificité de clivage de la RNase III d'E. coli

La RNase III est spécifique des régions d'ARN formant des doubles hélices, situées principalement en amont du site d'initiation de la traduction. La longueur des hélices, de forme A, de l'ARNdb doit être au moins d'un tour complet (> 11bp), la longueur minimale est une tige-boucle de 12 bp (pour revue, Nicholson, 1999). Ainsi, un dérivé de l'ARN R1.1 du phage T7, de 12 bp, peut être clivé efficacement *in vitro* (Fig. 34, A et B). La longueur de l'hélice ne change pas le site de clivage : la RNase III identifie son substrat à travers des traits de séquences et de structures spécifiques, appelées épitopes de réactivité. Par alignement de séquences, il a été montré que la présence de liaisons Watson-Crick au niveau des boîtes proximales, de 3 bp de long, et des boîtes distales, de 2 bp de long, inhibent la reconnaissance et donc le clivage de l'ARN (Fig. 34B) (Zhang et Nicholson, 1997). Elles

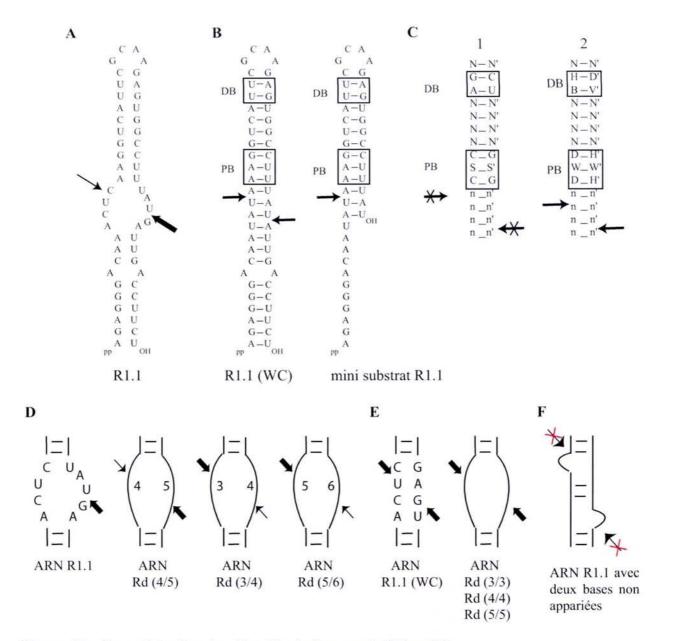


Figure 34 : Caractérisation des sites d'hydrolyse par la RNase III.

Les sites de clivages sont indiqués par des flèches, la taille des flèches reflètant l'efficacité d'hydrolyse. A : Structure de l'ARN R1.1 du phage T7 et ses sites de clivages à la RNase III (pour revue, Nicholson, 1999).

B : Influence de la longueur de la double hélice d'ARN. Un variant de l'ARN R1.1 possédant uniquement des liaisons Watson-Crick au niveau du site de clivage a été réalisé (R1.1(WC)), ce qui a permis de déterminer la taille minimal d'un substrat pour la RNase III (mini substrat R1.1) (pour revue, Nicholson, 1999).

C: Influence de la séquence sur l'efficacité de coupure par la RNase III (d'après Zhang et Nicholson, 1997). 1: séquence antidéterminante, limitant l'hydrolyse symbolisée par les flèches barrées. Les séquences qui limitent l'efficacité de coupure sont encadrées et appelées PB (boîte proximale) ou DB (boîte distale). S correspond à un résidu C ou G, S' son nucléotide complémentaire. N et N' sont des nucléotides strictement complémentaires, et n et n' sont des nucléotides non nécéssairement complémentaires. 2: dégénérescence des séquences hydrolysées par la RNase III. H et H' correspondent à des résidus A, G ou U avec D' et D leur nucléotide complémentaire (U, C ou A). B est un résidu C, G ou U avec son nucléotide complémentaire V' (G, C ou A). W est un résidu A ou U, avec W' son nucléotide complémentaire.

D : Influence de l'asymétrie sur le site de clivage. L'utilisation de séquences aléatoires montrent que le site de clivage est lié à la nature de l'asymétrie de la boucle (Calin-Jageman et Nicholson, 2003a).

auraient un rôle d'antidéterminant : elles participeraient à la sélection du site de clivage en protégeant les ARN de clivages non désirés (pour revue, Nicholson, 1999). La sensibilité aux positions W-C est à rapprocher à la capacité de la RNase III à cliver des ARNdb avec une large variété de séquences. Quelques substrats issus du bactériophage T7, comme l'ARN R1.1, contiennent une boucle interne, avec une asymétrie (4/5), où le clivage est réalisé dans le brin de 5 nucléotides (Fig. 34A) (Calin-Jageman et Nicholson, 2003a). Le clivage ainsi réalisé engendre une structure tige-boucle en 3' qui stabilise la séquence amont. C'est l'asymétrie (4/5) de la boucle interne qui dirige le clivage de la liaison phosphodiester spécifique, dans le fragment de 5 nucléotides de la boucle interne, et non la séquence de nucléotides (Fig. 34D), puisque les structures tiges-boucles (3/4) et (5/6), de séquences aléatoires, favorisent au contraire, un clivage en 5' de la structure tige-boucle. Par contre, lorsque les variants de séquences aléatoires contiennent des boucles internes symétriques, ils sont clivés sur les deux brins (Fig. 34E) (Calin-Jageman et Nicholson, 2003a). La présence d'une structure « base non appariée-hélicebase non appariée » sur un variant de l'ARN de R1.1 permet de fixer la RNase III avec la même affinité, mais conduit à une inhibition de l'activité de clivage (Fig. 34F) (Calin-Jageman et Nicholson, 2003b). Cette structure est un antidéterminant catalytique, inhibant le domaine N-terminal de la RNase III (Blaszczyk et al., 2004).

Ainsi, les sites de clivages de la RNase III sont déterminés par la position du site de fixation de l'enzyme, qui est gouverné par les éléments de séquences de la double-hélice. La présence d'une boucle interne, pouvant être considérée comme un épitope de réactivité secondaire, dirigerait le motif de clivage plutôt que l'activité elle-même (Calin-Jageman et Nicholson, 2003a).

#### 1.2.c / Structure de la RNase III de classe I

Chaque sous-unité de la RNase III est décomposée en deux domaines chez *E. coli*. Le domaine N-terminal, correspondant aux deux tiers de l'enzyme, porte l'activité catalytique, et le domaine C-terminal porte le motif de liaison à l'ARN double brin (ARNdb) (dsRBD : double strand RNA Binding Domain). La délétion de ce motif réduit l'affinité pour l'ARNdb *in vitro*, et exerce une inhibition de clivage du substrat à la fois *in vitro* et *in vivo* (Sun *et al.*, 2001). La structure, obtenue par analyse de la diffraction aux rayons X, du domaine dsRBD d'une protéine de liaison à l'ARN de *Xenopus* et de l'ARNdb, montre que ce domaine interagit avec les deux sillons mineurs et le sillon majeur intermédiaire (Fig. 35A). Dans ce cas, de nombreux contacts ARN-protéine impliquent des groupements 2'OH sur les deux brins de l'ARN : ceci expliquerait la reconnaissance sélective de l'ARNdb, et l'insensibilité générale à des séquences spécifiques.

La structure du domaine endoRNase de Aquifex aeolicus montre que le dimère est formé par de nombreuses interactions hydrophobes, et maintenu par une jonction « balle et panier », jonction ressemblant à celle réalisée entre deux morceaux de puzzle (Fig. 35D). La balle représente la chaîne

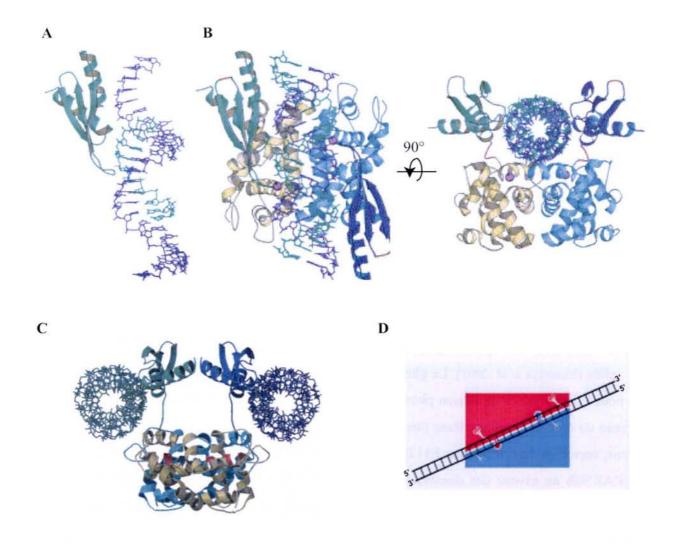


Figure 35 : Structure du domaine endonucléase de la RNase III de classe I d'Aquifex aeolicus et modèle de liaison avec l'ARN (d'après Blaszczyk et al., 2001; Blaszczyk et al., 2004; pour revue, Zamore, 2001).

- A : Structure du complexe du domaine ds-RBD de Xenopus laevis avec l'ARN à 1.9Å.
- B : Modèle hypothétique du complexe de la RNase III d'*A. aeolicus* avec l'ARNdb, à partir de la structure du domaine endonucléase d'*A. aeolicus*. Les ions Mg<sup>2+</sup> catalytiques sont représentés par des sphères violettes.
- C : Structure du complexe Aa-RNaseIII-E110K/ARNdb. La RNase III est constituée des hélices α et les feuillets β. Les deux brins de l'ARN, en vert et bleu, respectivement, sont sous forme de balle et batonnet. Le domaine endonucléase est en jaune et cyan pour la sous-unité A et B, respectivement, alors que le domaine dsRBD est en vert et bleu, pour les sous-unités A et B, respectivement. La double hélice d'ARN associée au dimère A est en vert, celle associée au dimère B est en bleu.
- D : Modèle de clivage endonucléolytique de l'ARNdb par les RNases III de classe I. Les deux sousunités de la RNase III, en bleu et rouge respectivement, sont liées entre elles comme un puzzle, via deux liaisons de type "balle et panier". La balle représente la protubérance, alors que le panier indique le creux fixant la balle. L'ARNdb est indiqué par un double-trait, les sites de clivages sont représentés par une paire de ciseaux.

latérale hydrophobe de la phénylalanine 41, le panier étant formé par les chaînes latérales des résidus V52,V56, L67, S68 et K71 (Blaszczyk *et al.*, 2001). Une vallée est alors formée par la réunion des deux sous-unités, dont la taille correspond avec la taille du substrat double brin (Fig. 35, B et D).

Le domaine catalytique de la RNase III peut hydrolyser une liaison phosphodiester en absence du domaine dsRBD, avec la même spécificité que l'holoenzyme (Sun et al., 2001). L'activité catalytique requiert un ion métallique divalent, de préférence Mg<sup>2+</sup>, et un ou plusieurs groupements acides. L'ion divalent stabiliserait le complexe enzyme substrat, et est requis pour la catalyse : cet ion pourrait générer l'hydroxyle nucléophile nécessaire à la réaction d'hydrolyse (Campbell et al., 2002b). Les alignements d'acides aminés de RNase III montrent 7 résidus acides aspartiques ou glutamiques conservés: E38, E41, D45, E65, E100, D114, E117. La structure du domaine N-terminal d'Aquifex aeolicus montre que les résidus acides correspondant aux résidus E41, D114 et E117 d'E. coli interagissent avec l'ion métallique. Cet ion métallique entraîne un réarrangement des chaînes latérales des résidus acides du site actif, ce qui conduit à l'obtention de deux sites de clivages de part et d'autre de la vallée (Blaszczyk et al., 2001). Le glutamate 117 est essentiel pour la fixation de l'ion magnésium, donc pour l'hydrolyse de la liaison phosphodiester mais aussi pour la fixation de l'ARN double-brin au niveau du domaine endonucléase (Blaszczyk et al., 2004). En effet, la structure du mutant E110K d'A. aeolicus, correspondant au mutant E117K d'E. coli (mutation rnc70), montre que ce mutant s'associe avec l'ARNdb au niveau des domaines dsRBD, mais elle ne peut pas interagir avec l'ARNdb au niveau de ses domaines endoRNases (Fig. 35C). Les résidus E41 et D114 semblent avoir une fonction redondante et beaucoup moins importante pour la fixation du métal (Blaszczyk et al., 2001; Sun et Nicholson, 2001; Sun et al., 2004). Le résidu D45, essentiel à la catalyse, semble déprotoner la molécule d'eau liée au métal pour créer l'espèce réactive. Les résidus E38 et E65 participent à créer une forte charge négative à proximité des sites actifs, et amélioreraient l'affinité de l'ion métallique pour son site (Sun et al., 2004). Le résidu E100 appartient à une région stabilisée par la fixation de l'ion métallique. Son rôle n'est pas encore connu, mais la mutation E100A conduit à une augmentation de la concentration en ion métallique requis, et une variation de l'activité catalytique.

# I.2.d / Mécanisme de clivage de la RNase III d'E. coli

Le modèle de reconnaissance suivant a été proposé : la RNase III reconnaît l'ARNdb par un de ses deux domaines de liaison à l'ARNdb. Lorsque la longueur de l'ARNdb est suffisamment importante et ne contient pas de séquences inhibitrices (antidéterminants), la RNase III semble s'isomériser grâce à son domaine flexible reliant les domaines N et C-terminaux, et l'affinité entre l'ARNdb et le domaine N-terminal. L'ARNdb se fixe alors au niveau de la vallée formée par les deux sous-unités. La RNase III coupe chaque brin de la double hélice de façon asymétrique par l'intermédiaire de ses 4 sites catalytiques, pour conduire à la production de fragments d'ARN

présentant une extrémité 5' monophosphate et une extrémité 3'OH (pour revue, Nicholson, 1999). Le clivage génère un fragment d'ARNdb de 9 pb, avec à chaque extrémité 3'OH deux bases simple brin (Blaszczyk et al., 2001). La liaison et la catalyse de la réaction sont associées à un changement structural de la protéine, l'étape limitante de la réaction serait le changement structural qui précède ou suit la dissociation des produits de clivage (Campbell et al., 2002b). En faveur de ce mécanisme, la structure du mutant E110K d'Aquifex aeolicus avec l'ARNdb montre une différence d'orientation de 90° des domaines de liaison à l'ARNdb par rapport à la structure de la RNase III sauvage avec l'ARN d'A. aeolicus (Fig. 35, B et C) (Blaszczyk et al., 2004).

#### I.2.e / Régulation de l'expression et de l'activité de la RNase III

Chez *E. coli*, le gène *rnc est* localisé en amont de l'opéron *rnc-era-recO* (Fig. 36A). Le gène *era* code une protéine de liaison au GTP avec une activité GTPase, et est essentiel pour la croissance d'*E. coli*. Le produit du gène *recO* participe dans la recombinaison RecF et les voies de réparation (pour revue, Grunberg-Manago, 1999).

La RNase III autorégule son expression via un clivage dans la région 5' non traduite (Matsunaga et al., 1996). En effet, l'ARNm est stabilisé par un facteur 10 dans un mutant rnc par rapport à la souche sauvage. La région 5' non traduite de l'ARNm rnc comporte trois structures tige-boucle TB1, TB2 et TB3 où TB2 et TB3 sont séparées par une séquence simple-brin contenant la séquence nécessaire à l'initiation de la traduction du gène rnc (Fig. 36B). La RNase III clive au niveau de TB2, ce qui élimine la structure tige-boucle TB1 et génère une extrémité 5'monophosphate. Ceci conduit à une dégradation rapide de l'ARNm rnc.

La RNase III est aussi régulée au niveau de son activité via une phophorylation par la protéine kinase à sérine et thréonine, activité portée par le domaine N-terminal de la protéine codée par le gène 0.7 du bactériophage T7. Cette phosphorylation conduit à une augmentation d'un facteur 4 de l'activité de clivage des ARN double brins (pour revues, Grunberg-Manago, 1999; Conrad et Rauhut, 2002), mais aussi des ARN possédant une boucle interne (4/5), fréquemment retrouvés dans les ARN du bactériophage. Ce clivage permet la formation d'une structure tige-boucle en 3' de l'ARN, ce qui protége les séquences amont et favorise la traduction des ARN matures, et donc les ARN du bactériophage (Calin-Jageman et Nicholson, 2003a).

#### I.3 / Les exoRNases 3'-5' : la RNase II, la RNase R et la PNPase

Trois exoRNases 3'-5' sont impliquées dans la dégradation totale des ARN (Donovan et Kushner, 1986; Cheng et Deutscher, 2005), la RNase II et la RNase R, possèdant une activité hydrolytique, et la

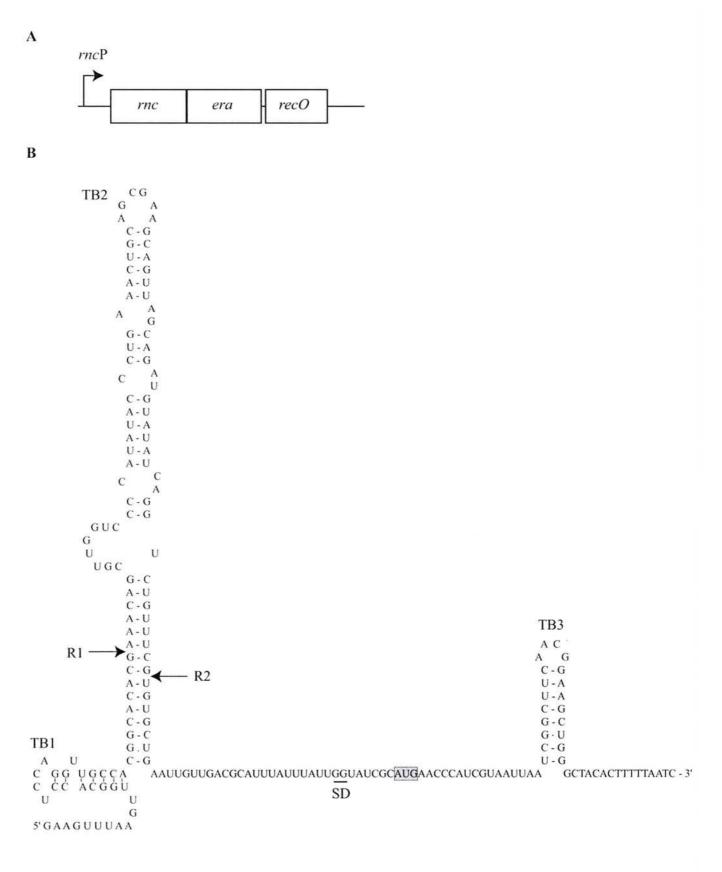


Figure 36 : Autorégulation de la production de RNase III (d'après Matsunaga et al., 1996).

A : Représentation schématique de l'opéron *rnc*. Les différentes ORF sont représentées par des rectangles, le promoteur *rnc*P est indiqué par une flèche brisée. Le gène *era* chevauche le gène *rnc* de quatre nucléotides.

B : Structure secondaire de la région 5' terminale de l'ARNm *rnc*. Les tiges-boucles sont notées TB1, TB2 et TB3. La séquence SD est soulignée, le codon initiateur AUG encadré. Les sites de coupures par la RNase III sont indiqués par R1 et R2.

polynucléotide phophorylase (PNPase), ayant une activité phosphorolytique. Comme 90 % de l'activité exonucléolytique d'extraits d'*E. coli* serait hydrolytique (Deutscher et Reuven, 1991), la PNPase serait responsable de 10 % de la dégradation exonucléolytique des ARNm dans une culture d'*E. coli* à 37°C. La PNPase peut se substituer à la RNase II ou la RNase R, ce qui explique pourquoi une souche inactivée dans l'une ou l'autre exoRNase est viable, alors que l'inactivation de la PNPase avec la RNase II ou la RNase R est létale pour la cellule (Cheng *et al.*, 1998; pour revue, Grunberg-Manago, 1999).

Ces 3 exoRNases se fixent sur l'extrémité 3' simple brin plus ou moins longue : au moins 10 nucléotides pour la PNPase, et 10 à 15 nucléotides pour la RNase II, et 5 à 17 nucléotides pour la RNase R, puis dégradent l'ARN cible jusqu'à quelques nucléotides de la tige-boucle pour la RNase II et la PNPase, en fonction de la séquence et de la structure de l'ARN cible (Cheng et Deutscher, 2005; Spickler et Mackie, 2000). La RNase II semble plus sensible aux structures tige-boucles que la PNPase la RNase R. Cette différence peut s'expliquer par la capacité de la PNPase à s'associer avec l'hélicase RhlB au niveau du dégradosome. Ce complexe peut alors dégrader des ARN fortement structurés, tels que les terminateurs Rho-indépendants ou les structure REP. La RNase II ne peut ni se lier aux composants du dégradosome, ni à l'hélicase RhlB, elle joue donc un rôle limité dans la dégradation de telles structures tiges-boucles (Spickler et Mackie, 2000). La RNase II peut aussi avoir un rôle stabilisateur : au niveau du transcrit *rpsO*, la RNase II génère des transcrits contenant seulement deux nucléotides non appariés en aval de la structure tige-boucle du terminateur. La présence uniquement de deux nucléotides ne permet pas de recruter la machinerie de dégradation exonucléolytique, ce qui conduit à une stabilisation du transcrit *rpsO* par la RNase II (Marujo *et al.*, 2000).

#### I.3.a / La RNase II et la RNase R

La RNase II et la RNase R appartienent à la famille d'exoRNase RNR (Tableau 2) (pour revue, Zuo et Deutscher, 2001). Les protéines de cette famille RNR d'E. coli sont caractérisées par trois domaines : un domaine de choc au froid/S1-like en N-terminal, un domaine central RNase II-like, un domaine S1 de liaison à l'ARN simple brin en C-terminal. Les membres de cette famille RNR sont largement retrouvés parmi les bactéries et les eucaryotes. Aucun homologue n'a été obtenu chez les Archaea, sauf pour Halobacterium NRC-1. Seul, les protéobactéries de la subdivision γ possèdent une RNase II et une RNase R. Par contre, les eucaryotes contiennent au moins trois membres de cette famille RNR. Les homologues chez la levure et chez l'homme sont liés à la GTPase RAN, impliquée dans les mécanismes d'échanges entre le noyau et le cytoplasme, et sont retrouvés dans l'exosome de la levure (Mitchell et al., 1997; Anderson et Parker, 1998). Ces données suggèrent que la famille RNR comprendrait des protéines régulatrices clés chez les eucaryotes (pour revue, Zuo et Deutscher, 2001).

#### I.3.a.α / Rôle de la RNase II

La RNase II, codée par le gène *rnb*, est un monomère de 644 aa et de 72.3 kDa (pour revue, Grunberg-Manago, 1999). Elle présente une activité 3'-5' ARN hydrolase, renforcée en présence d'ions monovalents et d'ions Mg<sup>2+</sup>, produisant des ribonucléotides 5' monophosphates à partir de transcrits portant une extrémité 3'OH (pour revue, Zuo et Deutscher, 2001). Cette activité est indépendante de la séquence d'ARN, mais sensible aux structures secondaires. Une souche délétée de la RNase II montre un plus faible taux de croissance, en milieu riche ou minimum (Cairrao et al., 2001).

#### I.3.a.β / Régulation de la RNase II

Les taux de RNase II et PNPase sont inversement corrélées *in vivo*. Ainsi, dans une souche déficiente en PNPase, l'expression du gène *rnb* est augmentée de 2 à 2,5 fois par un mécanisme de régulation post-transcriptionnelle. Inversement, une souche surproduisant la PNPase contient une quantité réduite de messager *rnb* et de protéine RNase II. Ceci suggère que ces deux exoRNases sont régulées de façon compensatoire (pour revue, Grunberg-Manago, 1999).

Un site de maturation à la RNase E a été identifié dans l'ARNm rnb. Le modèle le plus simple de régulation est que la dégradation de l'ARNm rnb est accomplie par la PNPase dans la direction 3' vers 5', avec ou sans clivage interne de ces ARNm par les endoRNases. La stabilité de la RNase II est aussi contrôlée en fonction des conditions de croissance, via le gène gmr (gene modulating RNase II) localisé en aval du gène rnb. La délétion de ce gène stabilise le transcrit rnb, passant d'une demi-vie inférieure à 1h à plus de deux heures et multiplie par trois la quantité de la RNase II (Cairrao et al., 2001).

#### I.3.a.γ / Rôle de la RNase R

La RNase R, codée par le gène *rnr*, est une protéine de 95 kDa. Elle peut dégrader des ARN faiblement structurés, mais aussi fortement structurés, tels que les ARNt, les ARNr, et les séquences REP d'ARNm (Cheng et Deutscher, 2002 ; Cheng et Deutscher, 2005). Elle nécessite la présence d'une région simple-brin en 3' de l'ARN, comprise entre 5 et 17 nts, pour se fixer à l'ARN et le dégrader (Cheng et Deutscher, 2005).

#### I.3.b / La PNPase

#### I.3.b.α / Rôle de la PNPase

La PNPase est un homotrimère, le poids moléculaire d'un monomère est de 77,1 kDa (711 aa). Elle est codée par le gène *pnp* (Regnier *et al.*, 1987; pour revue, Grunberg-Manago, 1999). Elle catalyse la phosphorolyse réversible de polyribonucléotides en nucléotides diphosphate en présence de phosphate inorganique. Cette réaction permet de dépenser moins d'énergie que l'hydrolyse des liaisons phoshates, puisqu'ici est conservée la liaison phosphate riche en énergie. Etant donné la forte concentration de phosphates inorganiques dans les cellules en croissance exponentielle *in vivo*, l'enzyme est supposée agir en tant qu'exoRNase. Cependant, Mohanty et Kushner ont montré récemment qu'elle permet la polyadénylation résiduelle dans des souches déficiente en PAP, et l'incorporation de résidus C et U dans les queues poly (A) trouvées dans les cultures d'*E. coli* en croissance exponentielle (Mohanty et Kushner, 2000).

#### I.3.b.β / Structure de la PNPase

La structure cristallographique tridimensionnelle de la protéine homologue de *Streptomyces antibioticus* est un dimère de trimères, ce qui permettrait de garder une symétrie par rapport aux composants dimérique du dégradosome (Fig. 32, B et C) (pour revue, Symmons *et al.*, 2002). Cette structure montre que chaque sous-unité de la PNPase est composée d'une structure centrale dupliquée qui contient le domaine catalytique, un domaine composé d'hélice α et deux domaines de liaison à l'ARN simple brin en C-terminal, KH et S1 (pour revue, Symmons *et al.*, 2002). Le domaine dupliqué est issu d'une fusion de produits de duplication de gènes de la RNase PH. Cette enzyme de dégradation phosphorolytique des ARN est impliquée dans la maturation des ARNt chez *E. coli* (Tableau 2) (pour revue, Symmons *et al.*, 2002). Bien que structurellement homologues à la RNase PH, les deux domaines de la PNPase seraient fonctionnellement distincts, l'activité exoRNase serait portée par le deuxième domaine. En plus des domaines S1 et KH, la PNPase comporterait un autre domaine de liaison à l'ARN, requis pour l'autorégulation. En effet, la mutation G454D, située en dehors des domaines S1 et KH, inhibe la liaison à l'ARN, sans avoir de répercussion sur l'activité catalytique de l'enzyme (Regonesi *et al.*, 2004).

La PNPase semble se lier à l'ARN simple brin, sans spécificité de séquence apparente, au niveau de l'extrémité 3', mais aussi au niveau de l'extrémité 5' de l'ARN et/ou à l'intérieur de l'ARN (Bermudez-Cruz et al., 2002 ; Regonesi et al., 2004). Son activité est stimulée d'un facteur 3 lorsque l'ARN possède une extrémité 5' monophosphate (Feng et al., 2002).

Les homologues de la PNPase ont été trouvés chez les mitochondries humaines, les bactéries et les chloroplastes, où la PNPase assure une des principales activités exoRNases, mais aucun orthologue n'a été identifié chez les Archaea (Tableau 2).

## I.3.b.γ / Régulation de la PNPase

Le gène *pnp* est localisé en aval du gène *rpsO*, codant la protéine ribosomale S15. Il est transcrit à partir du promoteur P1 situé en amont de *rpsO* et du promoteur P2 en amont de *pnp* (Fig. 37A), où 80 % des transcrits *pnp* sont issus du promoteur P1 (Regnier et Portier, 1986). Ces deux transcrits sont ensuite maturés par la RNase III, au niveau d'une structure tige-boucle, localisée des sites 82 à 119 nucléotides en amont du site d'initiation de la traduction (Fig. 37B). Ce clivage enlève la partie amont de la tige-boucle, formant un duplex avec une extrémité 3' simple-brin. Une fois ce clivage par la RNase III réalisé, l'expression du gène *pnp* est autorégulée négativement, l'absence de clivage par la RNase III conduit à une augmentation de la stabilité de l'ARNm *pnp* et une surproduction de l'ARNm *pnp* d'un facteur 10. La PNPase se lierait à l'extension 3' du duplex et dégraderait l'ARN amont, ce qui favorise ensuite l'instabilité de l'ARNm *pnp* (Jarrige *et al.*, 2001).

Lors d'un choc au froid, ce duplex serait stabilisé, inhibant l'autorégulation négative par la PNPase (Mathy et al., 2001; Jarrige et al., 2001). Cette augmentation de la synthèse de PNPase est requise pour la phase d'acclimatation lors d'un choc au froid, pour permettre le retour à la croissance de la bactérie. Cette activité ne peut être compensée par la RNase II, puisque les mutants pnp sont incapables de pousser en dessous de 30°C (Yamanaka et Inouye, 2001). Lors de cette phase d'acclimatation à un choc au froid sont synthétisées les protéines spécifiques de choc au froid (CSPs: cold shock proteins). A la fin de la phase d'acclimatation, la quantité de leur transcrit doit être réduite pour que la croissance puisse reprendre. Dans le cas inverse, ces transcrits piègeraient les ribosomes, et empêcheraient la traduction des ARNm totaux. La PNPase, en concert avec la PAP et l'hélicase ARN CsdA, joueraient un rôle important dans la dégradation sélective de ces ARNm CSPs (Yamanaka et Inouye, 2001).

# I.4 / La poly (A) polymérase (PAP) et la polyadénylation des ARN

La PAP réalise l'ajout de queues poly (A) à l'extrémité 3' des ARN. Ce processus de polyadénylation, bien que présupposé unique chez les eucaryotes pour lesquels il joue un rôle fondamental dans la stabilité des ARNm, est impliqué dans la dégradation des transcrits chez les bactéries.



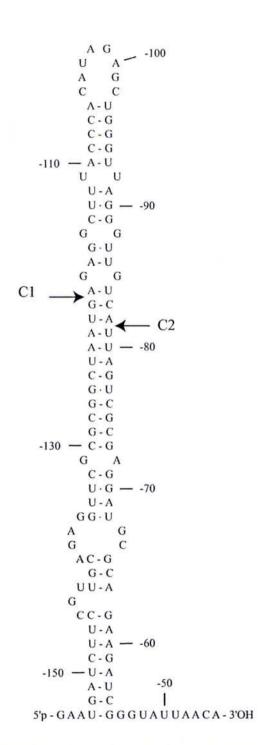


Figure 37 : Représentation schématique de l'opéron *rpsO-pnp* et structure de la région nécessaire à l'autorégulation de l'expression de *pnp* (d'après Jarrige *et al.*, 2001).

A : Représentation schématique de l'opéron *rpsO-pnp*. Les régions codantes sont représentées par des rectangles, les promoteurs P1 et P2 par des flèches brisées, et les sites de terminaison t1 et t2, ainsi que la tige-boucle tb1 reconnue par la RNase III, par des structures tige-boucles.

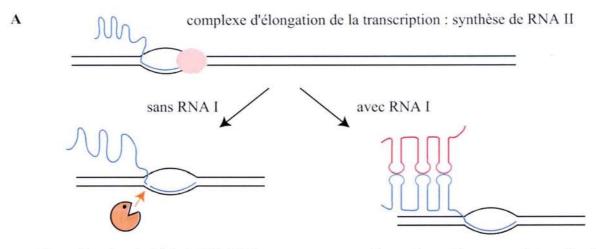
B : Structure secondaire prédite de la tige-boucle tb1 correspondant à la région 5' non traduite du transcrit *pnp*. Les sites de clivage C1 et C2 de la RNase III sont indiqués par des flèches. Les nucléotides sont numérotés par rapport au codon +1 de l'initiation de la traduction.

# I.4.a / Implication de la polyadénylation dans la dégradation des ARN

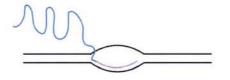
L'intervention de ce processus dans la dégradation des ARN chez les procaryotes a été mise en évidence au travers d'une étude de l'ARN régulateur non traduit RNA I, impliqué dans le contrôle du nombre de copies du plasmide portant le réplicon ColE1 (Fig. 38A). Au cours de la réplication de ce plasmide, l'ARN RNAII est transcrit, ce qui aboutit à la formation d'un hybride stable ARN/ADN au site *ori*. Cet hybride est ensuite clivé par la RNase H. Le produit de clivage en 5' sert d'amorce à la synthèse d'ADN, et donc à la réplication d'ADN. L'ARN régulateur RNA I interagit avec l'ARN RNA II, ce qui inhibe la formation d'un hybride ARN/ADN stable et la réplication du plasmide. La régulation de la stabilité de RNA I contrôle ainsi la régulation du nombre de plasmide portant le réplicon ColE1.

L'ARN RNAI est transcrit à partir d'un promoteur fort, sa demi-vie est de l'ordre de 2 minutes. La dégradation est initiée par un clivage à la RNase E à un site localisé à 5 nucléotides de l'extrémité 5', pour former l'intermédiaire de dégradation RNA L<sub>5</sub> (Fig. 38B) (Lin-Chao et Cohen, 1991). Le transcrit ainsi maturé est rapidement dégradé (Xu et al., 1993). Par contre, en présence de la mutation pcnB codant la poly (A) polymérase I (PAP I), l'ARN RNA I voit sa demi-vie passer de moins d'une minute à plus de 60 minutes (Xu et al., 1993; Cao et Sarkar, 1992a; pour revue, Carpousis et al., 1999). Cette stabilisation de RNA I entraîne une réduction dramatique du nombre de copies de plasmides. In vitro, la polyadénylation favorise la dégradation par la PNPase du fragment RNA I maturé par la RNase E (Xu et Cohen, 1995). D'autre part, la présence de queues poly (A) contenant au moins 3 A au niveau de RNA I inhibe la formation du complexe RNA I / RNA II in vitro, indiquant ainsi un second mécanisme de régulation du nombre de copies de plasmide (Xu et al., 2002). Ainsi, le gène pcnB (plasmide copy number) semble essentiel à la régulation correcte du nombre des plasmides de type ColE1 (Xu et al., 1993). Ce gène est aussi essentiel pour la dégradation des petits ARN Sok et CopA intervenant dans le contrôle du nombre de copies des plasmides R1 (Dam Mikkelsen et Gerdes, 1997 ; Soderbom et Wagner, 1998), et de l'ARN antisens Oop qui régule négativement l'expression du gène cII et affecte l'équilibre entre la lysogénie et l'infection lytique du bactériophage λ (Wrobel et al., 1998).

Bien que la PAP I ne soit pas essentielle (Masters et al., 1993), la surexpression du gène pcnB codant la PAP I inhibe la croissance de la bactérie, engendrant alors des colonies de petites tailles (Binns et Masters, 2002). Cette surexpression pourrait conduire à la polyadénylation des ARN ribosomiques, qui habituellement sont peu polyadénylés, et à leur dégradation (Mohanty et Kushner, 1999a). La PAP I intervient aussi dans la dégradation de nombreux ARNm. En utilisant un plasmide portant le gène pcnB inductible par l'IPTG, la demi-vie d'ARNm spécifiques, tels que rpsO, trxA, lpp et ompA, diminue en fonction de l'augmentation du taux de PAP (Mohanty et Kushner, 1999a).



formation d'un hybride ARN/ADN persitant au site *ori* clivé par la RNase H.



initiation de la synthèse d'ADN conduisant à la réplication du plasmide

formation d'un complexe d'embrassement ("kissing complexe") entre RNA I et RNA II ayant pour conséquences l'absence de formation d'un hybride ARN/ADN persistant. Ceci conduit à une absence d'initiation de la synthèse d'ADN

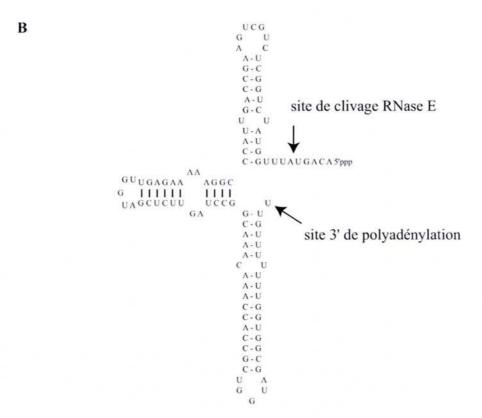


Figure 38 : Contrôle du nombre de copies de plasmides portant le réplicon ColE1 (type pBR322) par RNA I.

A : Représentation schématique du contrôle de la réplication de plasmides par RNA I et RNA II. B : Structure secondaire de l'ARN RNA I (d'après Xu et al., 2002). Le transcrit RNA I clivé par la RNase E, à 5 nucléotides de l'extrémité 5', donnant lieu à RNA I<sub>-5</sub> et polyadénylé en 3' s'apparie plus faiblement avec RNA II. Les queues poly(A) sont clivées par la RNase E et/ou les exoRNases.

# I.4.b / Mécanismes de dégradation des ARN polyadénylés

La polyadénylation réalisée en 3' des transcrits, est en concurrence avec la dégradation par les exoRNases 3'-5'. En effet, l'analyse des séquences poly (A) dans une souche mutée dans les gènes rnb et pnp des exoRNases, montre que ces enzymes participent à la dégradation des séquences poly (A), mais aussi qu'elles contrecarrent leur synthèse (Cao et Sarkar, 1992b; Hajnsdorf et al., 1995; O'Hara et al., 1995). Ceci suggère un mécanisme de compétition entre les 3' exoRNases et les PAP, pour la fixation à l'extrémité 3' des transcrits (Cao et al., 1997). Cependant, une fois la polyadénylation réalisée, celle-ci favorise la dégradation des ARN, en augmentant l'accessibilité des sites de clivages aux RNases. Ainsi, l'ARN RNA I est clivé plus rapidement par les endoRNases RNase A, RNase E, RNase T1 et RNase V1 lorsqu'il est polyadénylé, grâce à une variation de la structure tertiaire, sans observer de variations au niveau de la structure secondaire (Xu et al., 2002). L'ajout de séquences poly (A), contrairement aux séquences poly (G) ou poly (U), favoriserait la fixation des exoRNases et du dégradosome (Blum et al., 1999; Khemici et Carpousis, 2004), permettant ensuite la dégradation de régions peu structurées par la RNase II, moyennement structurées par la PNPase associée ou non avec l'hélicase RhlB, ou très structurées comme les séquences REP par le dégradosome ou la RNase R. L'extrémité 3' des transcrits bactériens est donc le site d'allongements par les PAP et de dégradation. Plusieurs cycles successifs seraient nécessaires à la dégradation totale des transcrits (Cohen, 1995; Coburn et Mackie, 1998).

Des études menées sur l'ARNm *rpsT*, codant la protéine ribosomale S20 (Coburn et Mackie, 1998) et sur les ARNm produits par le phage f1 (Goodrich et Steege, 1999) dans une souche d'*E. coli* où le gène *pcnB* est inactif, suggèrent que la polyadénylation des ARN n'intervient pas ou peu dans les premières étapes du processus de dégradation de ces transcrits. Elle jouerait plutôt un rôle fondamental dans la dégradation des intermédiaires de dégradation stables, en particulier ceux contenant l'extrémité 3' des transcrits, qui sont stabilisés par les structures secondaires de terminateurs de la transcription Rhoindépendant et les structures REP (Coburn et Mackie, 1998; Goodrich et Steege, 1999; Khemici et Carpousis, 2004). Ceci est conforté par le fait que la PAP est stimulée d'un facteur 5 pour un ARN portant une extrémité présentant une extrémité 5' monophosphate, donc déjà clivé, par rapport à une extrémité 5' triphosphate, et par des extrémités 5' et 3' en simple brin (Feng et Cohen, 2000).

La dégradation poly (A) dépendante assurerait un niveau de dégradation de base empêchant l'accumulation d'ARN inactifs, comme démontrés dans le cas d'ARNt mal repliés (Li et al., 2002) et des ARNm faiblement repliés générés par des clivages endonucléolytiques (Coburn et Mackie, 1998; Goodrich et Steege, 1999; Hajnsdorf et Regnier, 1999), ou lors d'une déficience en RNase E (Marujo et al., 2003).

## I.4.c / Activité poly (A) polymérase

Différentes approches ont montré que la PAP I serait responsable de la formation de 90 à 95 % des queues poly (A) dans les cellules d'*E. coli*, mais une autre étude montre qu'elle serait impliquée dans seulement 60 à 65 % des queues de nucléotides ajoutées posttranscriptionnellement (O'Hara et al., 1995; Mohanty et Kushner, 1999a; Mohanty et al., 2004). Ainsi, une autre activité PAP a été mise en évidence, favorisant entre autre la polyadénylation de l'ARN RNA I en milieu minimum et non en milieu LB (Xu et al., 2002). Cette activité, attribuée initialement à l'ORF 310 (Cao et al., 1996), a été identifiée comme issue de la PNPase (Mohanty et Kushner, 1999b; Mohanty et Kushner, 2000). En présence de PAP I, la PNPase semble ajouter préférentiellement des résidus U et C (Mohanty et Kushner, 1999a), alors qu'elle incorpore préférentiellement les résidus A, G, U puis C en absence de PAP I (Mohanty et Kushner, 2000).

# I.4.d / Structure de la poly (A) polymérase I

C'est une enzyme de 473 aa, dont 10 aa sont délétés post-traductionnellement (Binns et Masters, 2002). Ces 10 acides aminés dirigent partiellement la PAP I vers les membranes cellulaires (Jasiecki et Wegrzyn, 2005). Elle appartient à la superfamille des nucléotidyltransférase (Ntr), qui comprend les PAP eucaryotes et les enzymes ajoutant le groupement CCA aux extrémités 3' des ARNt (Raynal et Carpousis, 1999). Leur point commun est la polymérisation d'ARN au niveau des extrémités 3' terminales, sans requérir de brin matrice. Le domaine N-terminal de PAP I d'*E. coli* porte l'activité catalytique, son domaine C-terminal comprend un domaine de liaison à l'ARN et des sites d'interactions avec des facteurs protéiques, tels que la RNase E et les hélicases à boîtes DEAD RhIE, CsdA et SrmB (Raynal et Carpousis, 1999).

#### I.4.e / Régulation de l'expression de la poly (A) polymérase I

La synthèse de PAP I diminue avec le taux de croissance de la bactérie, lorsque les différents taux de croissance sont obtenus en utilisant différents milieux de croissance (Jasiecki et Wegrzyn, 2003). L'utilisation de fusions transcriptionnelles entre la région promotrice du gène *pcnB* et l'ORF *lacZ* montre que la région promotrice du gène *pcnB* est activée inversement avec le taux de croissance. Ainsi, la régulation de la polyadénylation des ARN, en corrélation avec leur stabilité, serait un moyen de réguler l'expression des gènes en fonction du taux de croissance. En particulier, l'ARN RNAI est ainsi stabilisé à de faibles taux de croissance, ce qui entraîne un plus fort nombre de copies de plasmides dans ces conditions de cultures.

L'expression du gène *pcnB* est aussi régulée au niveau traductionnel. La traduction débute au niveau du codon AUU peu fréquemment utilisé et peu efficace, mais aussi sujet à une régulation

négative (Binns et Masters, 2002) (voir Partie C). Ainsi, l'activité PAP I est très finement régulée, ce qui montre le rôle important de la polyadénylation pour obtenir le taux optimal de l'expression d'un gène en fonction des conditions de croissance.

La PAP est activée par la protéine Hfq, régulateur post-transcriptionnel décrit dans le prochain paragraphe. En absence de cette protéine Hfq, la PNPase est la principale polynucléotide polymérase de la cellule (Mohanty *et al.*, 2004).

# I.5 / La protéine Hfg

#### I.5.a / Rôle de la protéine Hfq

La protéine bactérienne Hfq est un régulateur posttranscriptionnel qui interagit avec de nombreux ARN. La protéine Hfq (appelée aussi HF-1) a été identifiée initialement comme un <u>facteur h</u>ôte bactérien requis pour la réplication du bactériophage ARN Qβ (Franze de Fernandez *et al.*, 1968). Cette protéine basique (pI=7.91) de 11,2 kDa, stable à la chaleur est retrouvée sous forme d'hexamère, en solution (Arluison *et al.*, 2002, Vassilieva *et al.*, 2002). Elle est détectée en majorité dans le cytoplasme en association avec les ribosomes via la protéine ribosomale S1, alors qu'une fraction mineure s'associe avec le nucléoïde (Kajitani *et al.*, 1994; Azam *et al.*, 2000; pour revue, Valentin-Hansen *et al.*, 2004). Cette association avec le nucléoïde peut être attribuée à une liaison non spécifique de Hfq à l'ADN (Takada *et al.*, 1997; Azam et Ishihama, 1999), ou être expliqué par la présence de transcrits associés avec le nucléoïde (pour revue, Valentin-Hansen *et al.*, 2004), en encore, par son interaction avec la protéine du nucléoïde H-NS (Kajitani et Ishihama, 1991).

La cellule bactérienne contient de 30 000 à 60 000 molécules de Hfq (Kajitani et al., 1994; Azam et Ishihama, 1999), son contenu varierait avec le taux de croissance. Cependant, d'un côté, la concentration en protéine Hfq semble augmenter avec la vitesse de croissance, sa régulation serait alors similaire à celle des protéines ribosomales (Kajitani et al., 1994; Ali Azam et al., 1999). D'un autre côté, la quantité de protéine Hfq mesurée est inversement proportionnelle au taux de croissance (Tsui et al., 1997; Vytvytska et al., 1998; Vecerek et al., 2003).

L'inactivation du gène hfq affecte la synthèse de plus de 50 protéines, détectées par gel bidimensionnel (Muffler et~al., 1997). Les phénotypes pléiotropes engendrés par cette inactivation comprennent une diminution du taux de croissance, une diminution du surenroulement de plasmides en phase stationnaire, un allongement des cellules et une oxydation plus rapide des sources de carbone (Tsui et~al., 1994). Ces effets pléiotropiques sont à corréler aux cibles de la protéine Hfq: elle interagit directement avec quelques ARNm mais elle régule aussi directement l'expression de régulateurs transcriptionnels globaux, tels que Fur et  $\sigma^S$  (Muffler et~al., 1996; Vecerek et~al., 2003). Ainsi, environ 40 %

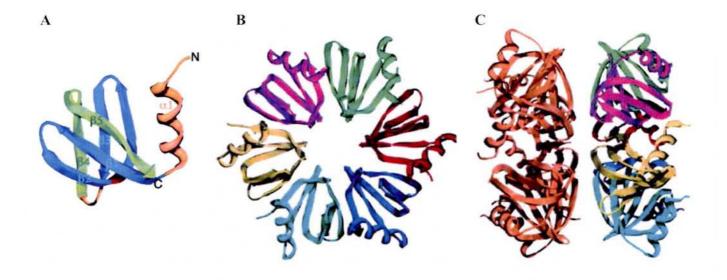
des effets pléiotropiques dus à la mutation de hfq sont superposés aux effets engendrés par la variation d'expression du facteur  $\sigma^S$  (Muffler et al., 1996; Muffler et al., 1997; pour revue, Valentin-Hansen et al., 2004).

### I.5.b / Homologie et structure de la protéine Hfq

Par analyse de séquence, il a été montré une légère similarité entre les domaines N-terminaux de la protéine Hfq et des protéines Sm et Sm-like (LSm) des eucaryotes et des Archaea (Zhang et al., 2002 ; Moller et al., 2002 ; Arluison et al., 2002). Ces protéines forment des hétéro-heptamères chez les eucaryotes, et constituent le composant principal des snRNP. Elles interviennent en particulier dans l'épissage des ARN mais aussi à différents évènements de maturation et de dégradation des ARN (pour revues, Will et Luhrmann, 2001 ; He et Parker, 2000 ; Pannone et Wolin, 2000). La fonction des protéines Lsm d'Archaea, formant un anneau homoheptamérique, est toujours inconnue, mais partage avec ses homologues eucaryotes et Hfq, la capacité de se lier à des séquences riches en U. Les protéines Sm et Lsm sont définies par la conservation d'un motif de séquence bipartite commun, le motif Sm, comprenant deux régions conservées, les segments Sm1 et Sm2, séparées par une boucle variable en taille et en composition (pour revue, Valentin-Hansen et al., 2004). Ce motif Sm est responsable des interactions ARN-protéine et protéine-protéine pour les protéines Sm et LSm d'Archaea et de cellules humaines (pour revue, Vasil'eva Iu et Garber, 2002). L'alignement des protéines homologues de Hfq avec les protéines Sm et Lsm a mis en évidence la conservation de nombreux résidus du motif Sm1, et de quelques résidus hydrophobes du motif Sm2 dans la protéine Hfq (Zhang et al., 2002). La structure cristalline de protéines Sm, comme celle de Hfq d'E. coli et Staphylococcus aureus, montre que le monomère est constitué d'une hélice α en N-terminal et 5 feuillets β antiparallèles (Fig. 39, A et B) (Sauter et al., 2003; Schumacher et al., 2002). Ainsi, le domaine Sm représentant un trait commun entre les familles Hfq et Sm/LSm, la protéine Hfq serait l'homologue structural et fonctionnel des protéines Sm et LSm dans les organismes n'ayant pas ce gène (Moller et al., 2002, Schumacher et al., 2002). En particulier, la protéine Hfq a été retrouvée chez l'archae Methanococcus jannaschii, archae qui ne possède pas de gène hôte Sm/LSm. Les différences de séquences, en particulier au niveau Sm2, suggèrent une évolution divergente à partir d'un ancêtre commun pour conduire aux protéines ayant les traits spécifiques aux Bactéries d'un côté et Archaea et Eucaryotes d'un autre côté (Sauter et al., 2003).

#### I.5.c / Specificité de fixation de la protéine Hfq

Comme les protéines Sm et LSm, la protéine Hfq se lie à l'ARN, au niveau de séquences riches en A-U localisées dans une région simple brin (Senear et Steitz, 1976) et semble requérir la présence de structures tige-boucle soit adjacentes, telles que celles de l'ARNnc Spot42 (Moller et al., 2002) et l'ARNm ompA (Moll et al., 2003b), soit situées de part et d'autre de la séquence simple brin riche en



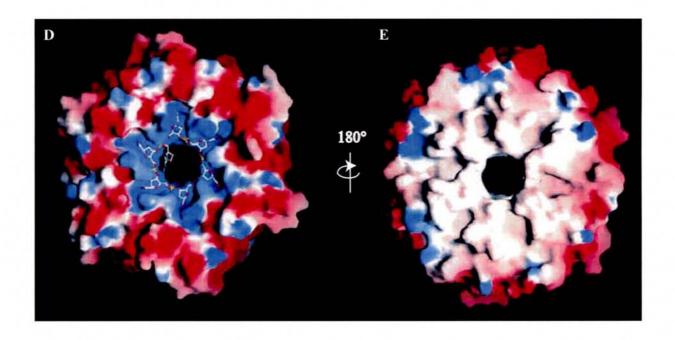


Figure 39 : Structure de la protéine Hfq de Staphylococcus aureus et du complexe Hfq-ARN (d'après Schumacher et al., 2002).

A : Représentation d'une sous-unité de Hfq sous formes d'hélices  $\alpha$  et feuillet  $\beta$ . Le motif Sm1 est coloré en bleu, le motif Sm2 en vert. La région N-terminale et la région variable comprise entre les deux motifs Sm1 et Sm2 sont colorées en orange.

B : Structure de l'hexamère Hfq. Chaque sous-unité est colorée différemment.

C : Structure d'un dimère d'hexamère. Les différentes sous-unités d'un hexamère sont représentées par des couleurs différentes, alors que l'autre hexamère est en orange.

D : Représentation de surface du site de liaison de l'ARN de l'hexamère Hfq. Les régions électropositives sont colorées en bleues, les régions électronégatives sont en rouges. L'ARN est représenté en batonnet.

E : Représentation de surface électrostatique du côté opposé de l'hexamère Hfq.

A/U, telles que celles des ARNnc DsrA (Brescia et al., 2003), OxyS (Zhang et al., 1998a; Moller et al., 2002), RyhB et de l'ARNm SodB (Geissmann et Touati, 2004). Elle se fixe aux queues homopolymériques poly (A) puis poly (U) puis poly (C) et enfin poly (G) (Folichon et al., 2005). La structure de la protéine Hfq de S. aureus avec l'oligonucléotide 5'AUUUUUG 3' montre que l'ARN est lié de manière circulaire autour du canal basique de l'hexamère Hfq, qui est conservé seulement sur un côté de l'hexamère (Fig. 39, D et E) (Schumacher et al., 2002).

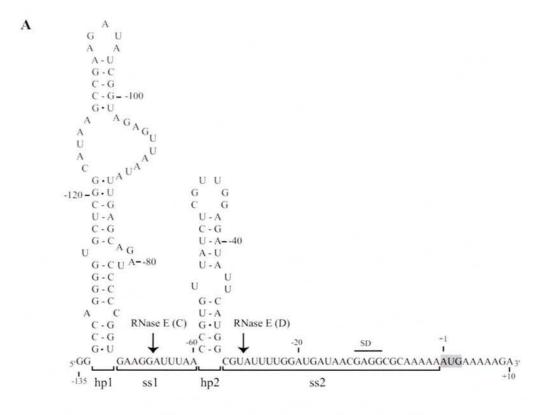
La protéine Hfq peut aussi lier simultanément l'ARNnc DsrA et une séquence poly (A). Ainsi, Hfq pourrait avoir plusieurs sites de liaison aux ARN, ou bien reconnaître la structure secondaire adoptée par la séquence poly (A) (Brescia et al., 2003; Mikulecky et al., 2004).

### 1.5.d / Mécanisme de régulation exercée par la protéine Hfq

### Régulation en fonction du taux de croissance de l'ARNm ompA

La synthèse de la protéine de la membrane externe OmpA (<u>O</u>uter <u>membrane protein A</u>) est corrélée à la stabilité de son ARNm *ompA*, qui varie de 4 à 20 min, de façon inverse à la vitesse de croissance d'E. coli (Nilsson et al., 1984, Georgellis et al., 1992). La dégradation de l'ARNm *ompA* est initiée par un clivage à la RNase E dans la région 5' non traduite, au niveau des sites C et D (Fig. 40A) (Nilsson et al., 1988; Lundberg et al., 1990). Cette région 5' terminale forme deux structures tige-boucle, hp1 et hp2 (Fig. 40A). Les éléments impliqués dans l'initiation de la traduction sont situés après hp2 (Chen et al., 1991). La forte stabilité des transcrits *ompA* est conférée par la structure tige-boucle hp1 et à la forte occupation du site RBS par les ribosomes dans la région simple-brin ss2 (Amold et al., 1998). La stabilisation de cet ARN ne dépend pas de la séquence, mais de la structuration, puisque le même effet stabilisateur est observé après substitution de cette structure par des structures tige-boucle ayant de toutes autres séquences (Emory et al., 1992). En revanche, la localisation de cette structure secondaire à l'extrémité 5' de l'ARNm est fondamentale pour sa stabilité, puisque l'ajout de 5 nucléotides en simple brin, en amont de cette structure, conduit à une forte déstabilisation de l'ARNm *ompA* (Emory et al., 1992).

Comme la quantité de RNase E ne varie pas avec le taux de croissance (Vytvytska et al., 1998), la variation de la stabilité du transcrit ompA en fonction du taux de croissance semble donc être liée à des facteurs agissant en trans, qui modulent l'activité de la RNase E. Ceci a ainsi conduit à l'identification de la protéine Hfq (Vytvytska et al., 1998). En effet, dans une souche inactivée au niveau du gène codant la protéine Hfq, la demi-vie de l'ARNm ompA augmente et n'est plus corrélée au taux de croissance (Vytvytska et al., 1998). La protéine Hfq régulerait la stabilité de l'ARNm ompA par compétition avec la sous-unité 30S pour la liaison avec la région 5' non traduite. Lorsque la sous-unité 30S se lie à cette région du transcrit ompA, elle le protège des clivages à la RNase E au niveau des sites C et D (Fig. 40). Par contre, la protéine Hfq se fixe principalement au niveau de la tige-boucle hp2 et de la région





#### C: Croissance lente

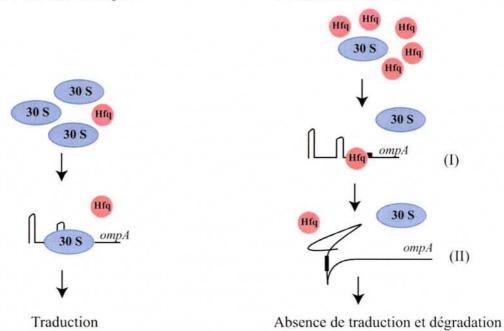


Figure 40 : Structure secondaire de la région 5' non traduite de l'ARNm *ompA*, et modèle de régulation en fonction du taux de croissance de la traduction et de la stabilité de l'ARNm *ompA* par la protéine Hfq.

A: Structure secondaire de la région 5' de l'ARNm ompA (Emory et al., 1992).

B : En conditions de croissance rapide, les sous-unités 30 S des ribosomes sont en excès par rapport à la protéine Hfq. La fixation de la sous-unité 30 S inhibe l'accès des sites de clivages C et D à la RNase E et stabilise le transcrit *ompA* (d'après Vytvyska *et al.*, 2000, Moll *et al.*, 2003, a et b).

C: En conditions de croissence lente, la protéine Hfq se lie préférentiellement à la région 5' de l'ARNm *ompA*. La fixation des ribosomes est inhibée soit de façon stérique par la protéine Hfq (I), soit par le changement structural introduit par la protéine Hfq (II). L'absence de traduction facilite le clivage par la RNase E dans la région 5' non traduite, aussi bien que dans la région codante, ce qui aboutit à une dégradation fonctionnelle rapide du transcrit (d'après Vytvyska *et al.*, 2000, Moll *et al.*, 2003, a et b).

simple brin en aval et déstabilise l'ARNm *ompA*. Elle inhiberait la fixation des ribosomes et la protection de l'ARNm ompA par les ribosomes en provoquant un changement conformationel qui a été mis en évidence par un changement de sensibilité aux nucléases dans la région -70 à + 10 (Vytvytska *et al.*, 2000; Moll *et al.*, 2003a; Moll *et al.*, 2003b). Ce modèle de régulation n'est valable que dans l'hypothèse où la concentration de la protéine Hfq est inversement corrélée avec le taux de croissance (Vytvytska *et al.*, 2000; Moll *et al.*, 2003b; pour revue, Valentin-Hansen *et al.*, 2004).

#### Action concertée avec les ARNnc

La fixation des ARNnc sur leur ARNm cible(s) conduit à une variation de la traduction (voir Introduction, Partie C.III.5) et/ou de la stabilité des ARNm cibles (voir Introduction, Partie D.II.2.d) (pour revue, Gottesman, 2004). La protéine Hfq joue tout d'abord un rôle stabilisateur d'une majeure partie des ARNnc, par inhibition stérique des clivages par la RNase E (pour revue, Gottesman, 2004). Elle favorise l'interaction des ARNnc avec leurs ARNm cibles, en modifiant la structure secondaire de l'un des deux partenaires, ou en augmentant la concentration locale des deux ARN, soit par fixation de deux ARN au niveau d'un hexamère Hfq, soit par interaction de deux hexamères ayant fixé chacun un ARN (Lease et Woodson, 2004).

## Régulation de la longueur des queues poly (A)

Dans le cas de l'ARNm *rpsO*, la protéine Hfq stimule d'une part, la synthèse processive des queues poly (A) par la PAP, favorisant ainsi la dégradation des ARN, et d'autre part, elle protège ces queues poly (A) de la maturation par la RNase E *in vivo* et *in vitro*, et de la dégradation par la PNPase et la RNase II *in vitro* (Hajnsdorf et Regnier, 2000; Folichon *et al.*, 2003; Folichon *et al.*, 2005). Ainsi, la protéine Hfq modulerait la sensibilité d'ARN aux RNases dans la cellule.

La protéine Hfq, copurifiée avec la PAP et la PNPase, favorise la polyadénylation au niveau des terminateurs Rho-indépendants (Mohanty et al., 2004). Elle interagirait au niveau de la structure tigeboucle du terminateur, ce qui la déstabilise, et favorise la polyadénylation par la PAP. Ainsi, la protéine Hfq pourrait alors empêcher la dégradation poly (A) dépendante des transcrits entiers, dont la stabilité est principalement contrôlée par la RNase E, et favoriser la dégradation poly (A) dépendante des ARN structurés stables dépourvus de site RNase E (Folichon et al., 2003).

# I.6 / Les autres enzymes impliquées dans la dégradation des ARNm

La RNase P est une endoRNase spécifique de régions en simple brin, générant des extrémités 5'monophosphate et 3'OH. Elle est composée, chez *E. coli*, d'un ARN M1 catalytique de 377 nucléotides (123 kDa), codé par le gène *rnpA*, et de la protéine C5, de 14 kDa (119 aa), codée par le gène *rnpB* (Tableau 2) (Tsai *et al.*, 2003, pour revues, Hajnsdorf, 1999; Condon et Putzer, 2002).

Elle intervient dans la maturation en 5' des ARNt, mais aussi dans le clivage de structure en pseudonoeud ressemblant aux ARNt, comme les ARN viraux (Mans et al., 1990), la maturation en 5' du précurseur de l'ARN 4.5S (Peck-Miller et Altman, 1991), de l'ARNtm (Komine et al., 1994), et de l'opéron histidine de S. typhimurium chez E. coli (Alifano et al., 1994; pour revue, Hajnsdorf, 1999). Par l'emploi de puces à ADN, il a été montré que la RNase P intervient aussi dans la maturation des gènes impliqués dans différents processus de transport, tels que les gènes codant les protéines impliquées dans le transport du ribose (rbsA-D et rbsK), l'expression de la perméase au tryptophane de faible affinité tnaB, et l'opéron du transport du lactose lacZYA (Li et Altman, 2003). La RNase P serait régulée par le taux de ppGpp (Li et Altman, 2003). La RNase P clive l'ARN dans les régions simple brin, mais l'efficacité est améliorée lorsque la partie en 3' du site de clivage appartient à une structure tigeboucle, et par la présence d'un G juste en aval du site de clivage (Hansen et al., 2001). La présence d'une région RCCA en simple brin en 3' du brin complémentaire permet d'ancrer l'ARN cible à l'ARN catalytique de la RNase P (Hansen et al., 2001).

Les homologues de RNase P ont été retrouvés dans les 3 domaines du vivant. Les sous-unités ARN de la RNase P des Bactéries et de quelques Archaea sont actives catalytiquement *in vitro*, en absence des sous-unités protéiques de la RNase P (Guerrier-Takada *et al.*, 1983; Pannucci *et al.*, 1999). De même, chez les Archaea, comme chez les Eucaryotes, de multiples protéines sont associées à l'ARN de la RNase P, dont aucune ne présente d'homologie aux protéines de RNase P bactérienne (pour revue, Condon et Putzer, 2002).

La RNase M est une endoRNase non spécifique qui hydrolyse les ARN en produisant des extrémités 5'OH et 3'monophosphate (Tableau 2) (Cannistraro et Kennell, 1989). Etant donné que la majorité des extrémités 5' terminales chez *E. coli* présente un groupement hydroxyle (environ 60 % pour une extrémité 5'OH, et 40 % pour une extrémité 5'ppp (Kennell, 2002)) et non phosphomonoester, Cannistraro et Kennell ont supposé que la RNase M serait la principale endoRNase de dégradation des ARNm (Cannistraro et Kennell, 1993). Or, la RNase M apparaît finalement être une forme altérée par une mutation de la RNase I, appelée RNase I\*, retrouvée uniquement dans la souche MRE600. L'endoRNase RNase I non spécifique, codée par le gène *rna*, est trouvée principalement dans l'espace périplasmique de la bactérie. Subbarayan et Deutscher ont pu décrire la mutation du gène *rna* engendrant la RNase M dans la souche MRE600, et montrer que la délétion du locus *rna* conduit à une perte complète des activités RNase I et RNase M (Subbarayan et Deutscher, 2001).

Une oligoRNase, codée par le gène *orn*, a récemment été identifiée chez *E. coli* (Yu et Deutscher, 1995 ; Zhang *et al.*, 1998b). C'est un homodimère thermostable, de 20.7 kDa par monomère, stimulée par les ions Mn<sup>2+</sup> (Zhang *et al.*, 1998b). Cette exoRNase 3'-5', essentielle à la cellule, hydrolyse spécifiquement, de façon processive, les fragments courts d'ARN. Elle serait la protéine impliquée

dans la dégradation des oligonucléotides de 2 à 5 mers issus des produits terminaux de dégradation par la RNase II et la PNPase (Ghosh et Deutscher, 1999).

L'oligoRNase appartient à la superfamille de RNases DEDD, caractérisée par la présence de 4 résidus acides invariants et d'autres résidus acides distribués dans trois motifs séparés (pour revue, Zuo et Deutscher, 2001). Des homologues de l'oligoRNase ont été retrouvés dans tous les génomes eucaryotes examinés, mais seulement chez les Actinomycètes et les Protéobactéries des subdivisions β et γ pour les génomes bactériens séquencés (Tableau 2). Elle n'a pas été mise en évidence chez les Archaea.

## II / Les éléments intervenants dans la stabilité des ARNm

La dégradation des ARN est contrôlée par la structure de l'ARN, facteur agissant en *cis* (Bouvet et Belasco, 1992 ; Joyce et Dreyfus, 1998 ; Vytvytska *et al.*, 2000), et/ou par la liaison de protéines, comme les ribosomes ou les ARN hélicases, ou d'autres ARN non codant, facteurs agissant en *trans*.

#### II.1 / Les éléments situés en cis

Nous avons précédemment décrit que la dégradation des ARNm chez *E. coli* est initiée par un clivage endonucléolytique, principalement par la RNase E, suivie d'une dégradation rapide. Cette vision est relativement simplifiée, puisque de nombreux transcrits procaryotes sont polycistroniques, et les cistrons ont des taux significativement différents de traduction et de dégradation. Pour de grands messagers polycistroniques, la dégradation des cistrons localisés en 5' a souvent lieu avant que les cistrons distaux soient transcrits. Les extrémités 3' des ARN naissants sont protégées des exoRNases, puisqu'elles sont séquestrées dans le complexe ternaire d'élongation contenant la matrice ADN et l'ARN polymérase. Ainsi, chez *E. coli*, la dégradation des ARNm est gouvernée au niveau d'un cistron et non du messager entier. Différentes barrières existent, en 5' et 3' du cistron, pour éviter que la dégradation ne touche les cistrons adjacents.

#### II.1.a / Les éléments localisés à l'extrémité 5' des ARNm

La structure de la région 5' des transcrits joue un rôle fondamental dans la dégradation, en limitant ou favorisant le premier clivage endonucléolytique limitant. Ceci est illustré par le cas de l'opéron *pts*. En effet, les gènes de cet opéron peuvent être transcrits à partir de plusieurs promoteurs alternatifs (Ryu et Garges, 1994a). Les différents ARNm présentent donc des extrémités 5' différentes, qui

leur confèrent des stabilités différentes. Ceci permet de réguler finement l'expression des gènes (Ryu et al., 1995).

#### II.1.b / Les éléments localisés à l'extrémité 3' des ARNm

#### Les terminateurs Rho-indépendants

La transcription des ARN s'achève au niveau de sites de terminaison de la transcription, dont certains dits Rho-indépendants sont constitués d'une longue structure tige-boucle stable, souvent riche en G-C, suivie d'une séquence riche en poly (U). Ces structures 3' terminales des ARNm stabilisent de nombreux ARNm (Abe et Aiba, 1996; Cisneros et al., 1996). La nature de la séquence poly (U) est insuffisante pour permettre la liaison des exoRNases, et la structure tige-boucle renforce cette inhibition.

# La séquence simple-brin

La nature de la séquence simple-brin située en 3' affecte la dégradation des ARNm (Lee et Cohen, 2001). La séquence poly (U) située après le terminateur Rho-indépendant limite l'accessibilité aux exoRNases. Les séquences poly (G) forment des quartets, inhibant l'accès des exoRNases et la dégradation en amont des ARNm. Cette stabilisation, engendrée par une queue poly (G) de longueur optimale de 15 nucléotides, est inversement proportionnelle à la longueur du transcrit. Les séquences poly (A) favorisent au contraire la fixation de la protéine Hfq, et la dégradation de l'ARNm.

# Les séquences REP (Repetitive Extragenic Palindrom)

Les structures tige-boucle, formées par les séquences REP favorisent la stabilité des ARNm. La bactérie *E. coli* possède environ 600 séquences REP localisées dans les régions intercistroniques à l'état de copie unique ou plus souvent multicopies (Bachellier *et al.*, 1999). Ces structures contribuent à l'expression différentielle de gènes au sein d'un même opéron, comme cela a été montré dans le cas de l'opéron maltose *malEFG* d'*E. coli*. Cet opéron code trois protéines requises pour le transport des maltodextrines à travers la membrane cytoplasmique. La protéine codée par le premier gène, MalE, est fortement produite par rapport aux produits des deux gènes en aval, MalF et MalG. La région intergénique *malE-malF* contient deux structures tige-boucle REP, stabilisant l'ARNm *malE*. La délétion d'une séquence REP diminue d'un facteur 9 la production de MalE. Ainsi, la séquence REP située dans l'espace intergénique *malE-malF*, limite la dégradation exonucléolytique du cistron amont *malE* (Newbury *et al.*, 1987a, Newbury *et al.*, 1987b; pour revue, Grunberg-Manago, 1999). *In vitro*, cette structure REP de *malE* inhibe la progression de la RNase II et de la PNPase, conduisant à la production d'intermédiaires de dégradation exonucléolytique stables (McLaren *et al.*, 1991; Py *et al.*, 1996; Blum *et al.*, 1999; pour revue, Cobum et Mackie, 1999).

# II.2 / Les éléments agissant en trans

Tous les facteurs se fixant à l'ARN peuvent intervenir sur sa stabilité, soit en le protégeant, comme les ribosomes, soit en inhibant sa protection par les ribosomes ou en favorisant sa dégradation comme la protéine chaperone H-NS.

#### II.2.a / Les ribosomes

Une étude approfondie de la relation entre traduction et stabilité des transcrits a été réalisée sur l'ARNm *lacZ* d'*E. coli*, dont la stabilité est contrôlée par une hydrolyse à la RNase E (Yarchuk *et al.*, 1992). L'ARN polymérase du phage T7 présente la particularité de transcrire les ARN huit fois plus rapidement que celle d'*E. coli*. Or, les phénomènes de transcription et de traduction sont couplés chez les bactéries. En utilisant cette ARN polymérase, il y a découplage entre transcription et traduction (Iost *et al.*, 1992). Les ribosomes, dont la vitesse d'élongation reste inchangée, ne progressent pas à la même vitesse que l'ARN polymérase, et de larges régions d'ARN ne sont donc plus protégées par les ribosomes. Les ARNm *lacZ* ainsi produits, présentent une stabilité fortement diminuée (Iost et Dreyfus, 1995). Cet effet déstabilisateur, lié au découplage entre transcription et traduction est inversement proportionnel à la vitesse d'élongation de l'ARN polymérase (Makarova *et al.*, 1995), et est réduit par la surproduction de protéines se fixant aux ARN, telles que les hélicases DeaD ou SmrB (Iost et Dreyfus, 1994). Ces expériences montrent de façon claire, que la stabilité du transcrit *lacZ* est influencée par la protection de la région codante qui limite l'accessibilité des sites d'hydrolyse par la RNase E. Lorsque l'ARN n'est pas protégé, sa dégradation par la RNase E est facilitée par l'hélicase RhlB (Khemici *et al.*, 2005).

La limitation de l'accessibilité des sites de coupures aux RNases par les ribosomes a aussi été étudiée en utilisant des antibiotiques. L'utilisation d'antibiotiques qui bloquent les ribosomes en cours d'élongation sur l'ARNm, tels que le chloramphénicol ou la tétracycline, stabilise fortement les transcrits. Au contraire, l'utilisation de la pyromycine, qui favorise le décrochage précoce des ribosomes, ou de la kasugamycine, qui inhibe l'initiation de la traduction, déstabilise fortement les transcrits étudiés (pour revue, Petersen, 1993). La déstabilisation liée à l'absence de traduction est utilisée dans les mécanismes de régulation feed-back, comme pour la ThrRS (Fig. 23C)(Nogueira et al., 2001).

# II.2.b / La protéine H-NS

La protéine H-NS, de 15 kDa, est impliquée dans la régulation de l'expression de nombreux gènes en réponse aux variations de l'environnement et dans l'adaptation à différents types de stress

(pour revue, Schroder et Wagner, 2002). Ce régulateur global de l'expression des gènes chez *E. coli*, fonctionne principalement comme un régulateur transcriptionnel (pour revue, Dorman, 2004). La forme homodimérique ou hétérodimérique avec la protéine paralogue StpA, se lie à l'ADN courbé, et bloque physiquement l'accès de l'ARN polymérase au promoteur. Les différents dimères semblent aussi pouvoir agir en tant qu'activateurs transcriptionels, mais le mécanisme n'est pas connu. La protéine H-NS exerce aussi un rôle au niveau post-transcriptionnel. Entre autre, elle augmente la production de la protéine MalT, activateur transcriptionnel de l'opéron maltose, mais diminue l'efficacité traductionnelle de l'ARNm *rpoS*. Elle intervient aussi dans la dégradation des ARN. En effet, il a été montré que la protéine H-NS augmente la susceptibilité de clivage de l'ARNnc DsrA à la RNase I *in vitro* (Brescia *et al.*, 2004). *In vivo*, l'ARNnc DsrA est dégradé moins rapidement dans un mutant *hns* par rapport à la souche sauvage. Le mécanisme d'action de la protéine H-NS est le suivant : en reconnaissant les régions d'ARN en double brin, le dimère H-NS modifierait la structure tertiaire de l'ARN. Les régions en simple brin non accessibles deviennent alors exposées aux clivages par les RNases. La diminution de stabilité de DsrA, activateur de la traduction de *rpoS* (Fig. 26) pourrait expliquer la diminution de l'efficacité traductionnelle de l'ARNm *rpoS*.

#### II.2.c / Les ARNnc

Les ARNnc sont des régulateurs post-transcriptionnels, intervenant dans la régulation de la traduction des ARN, mais aussi dans leur dégradation. Ils peuvent stabiliser ou favoriser la dégradation des ARNm cibles. Ainsi, l'ARNnc GadY stabilise l'ARNm d'un activateur transcriptionnel GadX en se liant à l'extrémité 3' terminale de l'ARNm cible (Opdyke et al., 2004). L'activateur transcriptionnel GadX active l'expression de deux Glutamate Decarboxylase, GadA et GadB, intervenant dans la résistance à un milieu de croissance acide. D'autre part, l'ARNnc RyhB déstabilise ses ARNm cibles, codant des protéines impliquées dans le métabolisme du fer, comme l'ARNm sodB codant une Superoxyde Dismutase (Masse et Gottesman, 2002; Geissmann et Touati, 2004). L'interaction de l'ARNnc avec son ARNm cible conduit à une dégradation de l'ARNm sodB par la RNase E, mais aussi par la RNase III (Afonyushkin et al., 2005). Il est à noter que la RNase E est plus souvent impliquée dans la dégradation des complexes ARNm/ARNnc. Ceci est à rapprocher de la complémentarité imparfaite entre l'ARNnc et son ARNm cible.

# Partie E / Régulation de l'expression des gènes du métabolisme central

# I / Variation de la production des enzymes du métabolisme central en fonction des conditions de croissance

La bactérie *E. coli* adapte son métabolisme afin de maximiser son taux de croissance dans différentes conditions de l'environnement (Edwards *et al.*, 2001). Ainsi, elle régule la quantité et l'activité des enzymes impliquées dans les différentes voies métaboliques suivant la concentration des différents nutriments et suivant les variations des paramètres physiques de l'environnement, comme la température de croissance, la concentration en oxygène disponible. Les différentes variations engendrées par différentes conditions de croissance ont été quantifiées par l'étude du transcriptome ou du protéome, ou par la mesure d'activités enzymatiques ciblées. Récemment, Peng et Shimizu ont étudié les variations de composition du protéome et de quelques activités enzymatiques du métabolisme carboné, dans différentes conditions de croissance (Peng et Shimizu, 2003).

Ainsi, en milieu LB complémenté en glucose, la diminution de la quantité d'oxygène disponible, où les cellules passent alors de conditions de croissance aérobies à microaérobies, conduit à une faible augmentation des enzymes générales du système PTS (EI, HPr, EIIA<sup>Glc</sup>), mais aussi des enzymes de la glycolyse. Les enzymes du cycle de Krebs sont fortement réprimées, tandis que les enzymes fermentaires, comme celles codées par les gènes pfl et adhE, voient leur concentration fortement augmenter.

Lorsque l'acide acétique est utilisé à la place du glucose, en milieu LB et conditions de croissance aérobie, la production des enzymes spécifiques de la glycolyse, codées par les gènes pfkA, pykF et ppc (Fig. 41), est inhibée alors que celle des enzymes spécifiques de la gluconéogenèse, codées par les gènes fbp, pckA et ppsA (Fig. 4), est activée. La concentration des enzymes du shunt du glyoxylate, et celles du cycle de Krebs augmentent fortement. Cependant, même si la concentration de l'isocitrate désydrogénase (ICDH) codée par le gène icd augmente, son activité est diminuée. En effet, au niveau de l'isocitrate, deux voies métaboliques sont possibles. La poursuite du cycle de Krebs, par la réaction catalysée par l'ICDH, conduit à l'élimination de deux molécules de CO<sub>2</sub> et à la production de coenzymes réduits (pour revue, Cronan et Laporte, 1996). Le shunt du glyoxylate permet d'éviter la formation de ces deux molécules de CO<sub>2</sub> par le cycle de Krebs, et favorise la gluconéogenèse. Le choix entre ces deux voies métaboliques est réalisé au niveau de l'ICDH. Sa phosphorylation, réalisée par l'ICDH kinase / phosphatase codée par le gène aceK, diminue l'affinité pour son substrat. L'isocitrate est alors utilisé préférentiellement par le shunt du glyoxylate. Les enzymes du shunt du

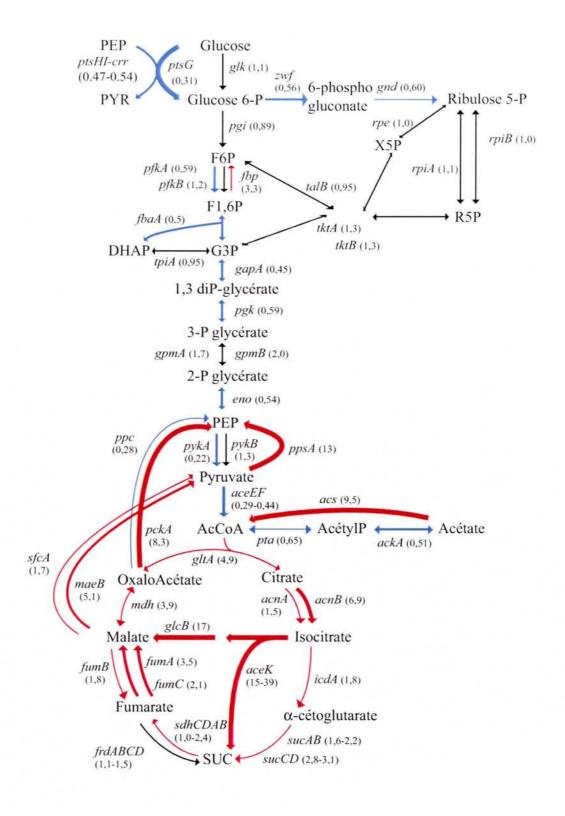


Figure 41 : Variation du taux des transcrits des voies métaboliques carbonées centrales d'une souche d'*E. coli* cultivée sur un milieu minimum contenant de l'acétate par rapport à un milieu minimum contenant du glucose (d'après Oh *et al.*, 2002).

Les réactions catalysées, induites sur acétate, sont représentées par des flèches rouges, tandis que celles inhibées sont représentées en bleu. Le taux d'induction (valeur >1) ou de répression (valeur <1) est indiqué entre parenthèse à proximité du gène correspondant.

glyoxylate sont codées par les gènes aceA et aceB, et appartiennent à l'opéron aceBAK, avec l'ICDH kinase/ phosphorylase (pour revue, Cronan et Laporte, 1996). L'étude du transcriptome de cultures réalisées en milieu minimum complémenté en acétate par rapport à un milieu minimum complémenté en glucose montre des variations similaires à celles observées par l'étude du protéome en milieu LB complémenté en acétate par rapport à un milieu LB complémenté en glucose. En effet, la quantité de transcrit des enzymes générales du système PTS (ptsl, ptsH et crr), des enzymes spécifiques (ptsG) et de la glycolyse (fbaA, gapA, pgk et eno) diminue (Fig. 41) (Oh et al., 2002). Par contre, les transcrits du cycle de Krebs, shunt du glyoxylate et des enzymes spécifiques de la gluconéogenèse voient leur concentration fortement augmenter.

De façon générale, les protéines de la glycolyse et de la néoglucogenèse, qui ne sont pas spécifiquement impliquées dans l'une des deux voies, à savoir les protéines codées par les gènes pgi, fbaA, gapA, pgk et gpmA, montrent un même taux de variation en fonction des conditions de croissance (Peng et Shimizu, 2003).

Les variations du protéome ont été comparées à celle du transcriptome à différents moments d'une culture de forte densité (HCDC : High Cell Density Cultivation) afin de suivre les variations du métabolisme (Yoon et al., 2003). Cette méthode de culture a été développée pour améliorer la productivité des produits d'intérêts (pour revue, Lee, 1996). De façon générale, les variations d'expression des gènes observées par l'analyse du transcriptome et du protéome sont similaires. Là encore, l'expression des gènes eno, fbaA, gapA, gpmA et tpiA de la glycolyse et de la néoglucogenèse varient de la même manière. Leur expression est en particulier accrue pendant la phase stationnaire. Par contre, l'expression des gènes codant les enzymes du système PTS générales et spécifique (EIIBC<sup>Glc</sup>) et celles du cycle de Krebs augmentent pendant la phase de croissance exponentielle, mais diminuent pendant la phase stationnaire.

En général, sous différentes conditions de croissance, les variations d'expression des gènes gapA, pgk et fbaA varient de la même manière, sauf lors de la surproduction de protéines en milieu minimum M9 complémenté en glucose (Oh et Liao, 2000). En effet, lorsque la bactérie synthétise de façon importante un composé d'intérêt non toxique, son métabolisme varie : sa vitesse de croissance diminue, voire s'arrête si la surproduction atteint 30 % des protéines totales synthétisées. L'observation du transcriptome montre une augmentation de la quantité de transcrit des enzymes du cycle de Krebs et du shunt du glyoxylate. Par contre, la quantité de transcrits des gènes du transport du système PTS, à la fois des enzymes générales et celle spécifique du glucose, et des enzymes de la glycolyse diminue. En particulier, la diminution est de l'ordre d'un facteur 1,5 pour ptsG, d'un facteur 2 pour pgk, gapA, ptsHI et crr, mais d'un facteur 5 pour fbaA, dans la souche MC4100 en milieu minimum (Oh et Liao, 2000).

En résumé, nous venons de voir, à travers quelques exemples, que la bactérie modifie son métabolisme sous différentes conditions, telles qu'un milieu de culture complémenté avec différentes Inhibition de l'activité de perméases conduisant à l'exclusion de l'inducteur de certains opérons

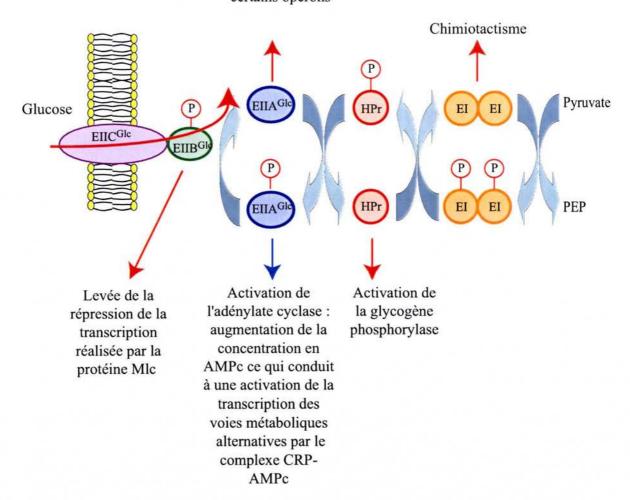


Figure 42 : Les régulations exercées par le système PTS du glucose chez E. coli (d'après Siebold et al., 2001).

Le code couleur pour les composants du système PTS est le même que celui utilisé en Figure 2. Le transport du glucose par la protéine EIIBCGlc entraîne différentes régulations exercées par les composants du système PTS du glucose, régulations représentées par des flèches rouge. En absence de glucose, les régulations indiquées par les flèches bleues sont exercées.

sources de carbone. Ces variations ont été observées au niveau des transcrits, de la quantité et de l'activité des enzymes du métabolisme carboné. Nous allons décrire les différentes régulations exercées au niveau de la production des enzymes du métabolisme carboné, en particulier au niveau de la glycolyse, par le système PTS du glucose, mais aussi la protéine Cra et les systèmes à deux composants.

### II / Régulation des voies du métabolisme carboné exercée par le système PTS du glucose

Le système PTS, décrit précédemment dans la partie A, réalise le transport et la phosphorylation de son substrat. Il régule aussi le métabolisme de la bactérie au niveau de l'activité des enzymes, mais aussi au niveau de la transcription des opérons nécessaires pour le métabolisme d'autres sources carbonées. Ces opérons comprennent les gènes codant les protéines assurant le transport et le métabolisme d'une source carbonée. La régulation de l'expression de ces opérons a pour but de rentabiliser au mieux les dépenses d'énergies et d'équilibrer l'entrée des sources de carbone avec les capacités métaboliques de la cellule (pour revue, Bruckner et Titgemeyer, 2002). L'ensemble de ces régulations appartient au mécanisme de répression catabolique carbonée. Elle est exercée par les différentes enzymes du système PTS, directement, par phosphorylation ou régulation allostérique, ou indirectement, en fonction du rapport entre la forme phosphorylée et la forme non phosphorylée des enzymes du système PTS (Fig. 42). Les régulations exercées par les composants du système PTS du glucose opèrent : 1) au niveau du chimiotactisme, ce qui permet d'orienter la bactérie vers les sources de carbone empruntant le système PTS, régulation exercée via la protéine EI déphosphorylée qui inhibe le senseur kinase CheA de la machinerie du chimiotactisme, 2) au niveau de la régulation de l'expression des gènes codant les enzymes spécifiques du métabolisme d'une source de carbone alternative n'empruntant pas le système PTS et leur transport, régulations exercées directement ou indirectement via les enzyme HPr, IIAGle et IIBCGle (Fig. 42). Nous nous limiterons à l'étude des régulations de l'expression des gènes du métabolisme carboné, exercées par les enzymes HPr, EIIA Gle, EIIBCGlc et Mlc.

#### II.1 / Régulation du métabolisme carboné par l'enzyme générale HPr

Chez de nombreuses bactéries, un des mécanismes de la répression catabolique carbonée repose sur la modulation de la phosphorylation, par un des composants du système PTS, d'un régulateur transcriptionnel, activateur de la transcription ou ayant un rôle d'antiterminateur, (pour revue,

Stulke et Hillen, 1998). Chez *E. coli*, ce mécanisme de régulation a été observé au niveau de l'opéron *bglGFB*, codant respectivement l'antiterminateur BglG de cet opéron, le système de transport de β-glucosides EIIACB<sup>Bgl</sup> et la phospho-β-glucosidase impliquée dans le métabolisme de β-glucosides (pour revue, Plumbridge, 2002). L'antiterminateur BglG requiert une phosphorylation par la protéine HPr pour être actif, et une absence de phosphorylation par le système de transport EIIACB<sup>Bgl</sup>, les sites de phosphorylation étant situés au niveau de deux domaines différents. Ainsi, en absence de β-glucosides, la protéine EIIACB<sup>Bgl</sup> transfert son phosphate à l'antiterminateur BglG, ce qui l'inactive. En présence de β-glucosides, la protéine EIIACB<sup>Bgl</sup> transfert préférentiellement son phosphate au substrat par rapport à l'antiterminateur BglG. Ce dernier doit cependant être phosphorylé par la protéine HPr pour être complètement activé. Cette phosphorylation est fonction du taux de phosphorylation de la protéine HPr, qui dépend de l'utilisation d'autres sucres PTS par la bactérie.

Chez *E. coli*, la protéine HPr peut aussi exercer un contrôle du métabolisme carboné en modifiant l'état oligomérique de l'enzyme cible. Ainsi, la protéine HPr se fixe à la glycogène phosphorylase, codée par le gène *glgP*, dernier gène de l'opéron *glgCAP*, comprenant aussi les gènes codant une ADP-glucose pyrophosphorylase et une glycogène synthase (Seok *et al.*, 1997a). L'interaction de la protéine HPr déphosphorylée avec la glycogène phosphorylase favorise la forme tétramérique de cette enzyme par rapport aux monomères et aux dimères, ce qui conduit à une augmentation de l'activité de l'enzyme de 250 %.

#### II.2 / Régulation du métabolisme carboné par la protéine EIIA<sup>Glc</sup>

La diminution du taux de phosphorylation de la protéine EIIA<sup>Glc</sup>, lié au transport de sucres par le système PTS en général, exerce deux régulations majeures de la répression catabolique carbonée que nous allons détailler. D'une part, elle se lie au système de transport de sucres non PTS et les inactive : il s'agit du mécanisme d'exclusion de l'inducteur. D'autre part, elle conduit à une diminution de la concentration du complexe CRP-AMPc, activateur transcriptionnel d'environ 75 gènes (Gosset *et al.*, 2004). La protéine EIIA<sup>Glc</sup> intervient aussi dans d'autres mécanismes de régulation : elle interagit avec la protéine FrsA, présente chez les γ-protéobactéries anaérobes facultatives, qui régule le flux entre la respiration et la fermentation (Koo *et al.*, 2004), et joue un rôle dans la régulation de l'expression du facteur σ<sup>S</sup>, au niveau transcriptionnel et traductionnel (Ueguchi *et al.*, 2001).

## II.2.a / Régulation par la protéine EIIA Gle déphosphorylée : l'exclusion d'inducteur

#### II.2.a.α / Le rôle de l'exclusion d'inducteur

L'exclusion d'inducteur correspond à l'inhibition de la formation d'un inducteur nécessaire pour l'expression de voies métaboliques alternatives. En effet, la présence d'une source de carbone préférentielle inhibe l'entrée d'une autre source de carbone moins favorable et empêche la formation de l'inducteur nécessaire à l'expression des voies métaboliques alternatives. Ainsi, chez E. coli, la présence de glucose inhibe l'entrée du lactose, son isomérisation en allolactose et donc l'induction de l'opéron lactose par l'allolactose. La présence d'α-méthyl-glucoside, analogue non métabolisable du glucose, inhibe aussi l'entrée d'autres sucres PTS (Saier et al., 1978 ; Kimata et al., 1997). Ce phénomène régulateur est réalisé, chez E. coli, par la protéine EIIA déphosphorylée. En effet, la mutation du gène crr abolit le phénomène d'exclusion d'inducteur. En présence de glucose, la protéine EIIA Gle est principalement sous forme déphosphorylée. Elle se lie aux perméases de sucres non PTS, telles que celles du lactose, maltose, mélibiose et raffinose, et inhibe le transport des sucres correspondants (Saier et al., 1978; Kimata et al., 1997; pour revue, Bruckner et Titgemeyer, 2002). La fixation efficace de la protéine EIIA Glc sur les protéines cibles est favorisée par la présence de leur source de carbone spécifique (Osumi et Saier, 1982; Saier et al., 1983). Ainsi, l'interaction d'une source de carbone avec son transporteur induirait une modification structurale du transporteur, qui favoriserait sa reconnaissance par la protéine EIIA Glc non phosphorylée. Ce mécanisme permet d'éviter la titration possible de la protéine EIIA Glc par l'ensemble de ses protéines cibles.

L'inhibition de la formation de l'inducteur peut aussi être réalisée à une étape ultérieure après le transport dans la bactérie, comme dans le cas du glycérol. Ce composé rentre dans la cellule par diffusion facilitée grâce à un facilitateur (GlpF) (Fig. 1), puis est phosphorylé par la glycérol-kinase pour former le glycérol-phosphate, inducteur des gènes du catabolisme du glycérol (pour revue, Bruckner et Titgemeyer, 2002). La protéine EIIA<sup>Gle</sup> se fixe non pas sur le facilitateur au glycérol, mais interagit avec la glycérol-kinase et inhibe la formation de l'inducteur. Ceci a pour conséquence de limiter le transport du glycérol par diffusion facilitée. La structure tridimensionnelle à haute résolution du complexe glycérol-kinase/EIIA<sup>Gle</sup> montre que la glycérol-kinase sous forme de tétramère fixe une protéine EIIA<sup>Gle</sup> sur chaque monomère, par des interactions essentiellement hydrophobes et électrostatiques, en présence de zinc (Hurley et al., 1993; pour revue, Robillard et Broos, 1999). La fixation de la protéine EIIA<sup>Gle</sup> induit un changement conformationnel de la glycérol kinase qui est renforcé en présence de zinc, ce qui inhibe l'activité de l'enzyme cible (Holtman et al., 2001a). Il est toutefois à remarquer que l'inhibition de la glycérol-kinase est réalisée de façon plus importante par un

intermédiaire de la glycolyse, le fructose-1,6-bisphosphate (Holtman *et al.*, 2001b; pour revue, Bruckner et Titgemeyer, 2002).

#### II.2.a.β / Spécificité de l'interaction de la protéine EIIA Gle déphosphorylée

La perméase du lactose, constituée de 12 hélices transmembranaires, interagit avec la protéine EIIA Gle au niveau de deux régions. Les mutations réalisées dans ces deux régions n'affectent pas la liaison au substrat, puisqu'il y a une activité de transport, mais abolissent la liaison à la protéine EIIA Gle (Seok et al., 1997b; Hoischen et al., 1996). Les acides aminés de ces deux régions ont été alignés avec les séquences correspondantes des perméases du maltose (MalK), raffinose (RafB), mélibiose (MelB) et de la glycérol-kinase. Cet alignement a permis de définir les séquences consensus suivantes pour les régions 1 et 2 (Sondej et al., 1999):

région 1 :  $(V,L)(X)\{0,2\}(S,A,T)(X)\{0,2\}(R)(X)\{0,6\})(G)(K,R)$ 

**région 2**:  $(S,T)(X)\{0,4\}(V,L)(X)\{0,2\}(N,Q,E)(X)\{0,2\}(G)(X)\{0,1\}(N,K)(X)\{0,1\}(S)$ 

(les acides aminés sont indiqués par le code à une lettre ; les deux chiffres entre parenthèses indiquent le nombre d'acides aminés séparant l'acide aminé amont de l'acide aminé aval.)

Il est à noter que la perméase du mélibiose (MelB), régulée par la protéine EIIA<sup>Glc</sup>, contient seulement la région consensus 1, la région 2 étant dégénérée. La recherche de ces régions 1 et 2 sur le génome d'*E. coli* a permis d'identifier 1 456 protéines possédant l'une ou l'autre région, et 37 protéines possédant ces deux régions. Ce groupe de 37 protéines semble constitué de 4 protéines de transport, 2 protéines kinases, 4 protéines de liaison à l'ADN, 10 protéines membranaires, et 17 protéines cytosoliques. Parmi ces 17 protéines se trouve une pyruvate déshydrogénase E2 (p06959) impliquée dans l'utilisation du pyruvate, un des produits de la voie métabolique du PTS, et la GTP pyrophosphorylase (RelA) impliquée dans la synthèse du régulateur global ppGpp (Sondej *et al.*, 1999).

Ces deux régions consensus ont été retrouvées au niveau de la zone d'interaction de la protéine EIIA Gle avec la protéine HPr. Ainsi, les interactions régulatrices et catalytiques impliquant la protéine EIIA Gle sont réalisées par des régions similaires (Sondej et al., 1999).

## II.2.a. $\chi$ / Exclusion d'inducteur par des sources de carbone dites PTS autres que le glucose

En plus du glucose, les autres sources de carbones dites PTS permettent la répression de l'utilisation des sources de carbone dits non PTS (Saier et al., 1976; Saier et al., 1983). Cette répression est réalisée suivant le même mécanisme que celui établit pour le glucose : l'interaction de la forme non phosphorylée de la protéine EIIA<sup>Gle</sup> avec les protéines cibles du mécanisme d'exclusion d'inducteur. En effet, la cascade de phosphorylation du système PTS étant réversible entre les enzymes générales et

les protéines EIIA et EIIB, le passage d'une source de carbone au travers de son système PTS spécifique conduirait à la déphosphorylation successive des protéine EIIB et EIIA spécifiques de cette source de carbone, puis à la déphosphorylation de la protéine générale HPr. Celle-ci est rephosphorylée par la protéine générale EI, mais peut être aussi phosphorylée par des protéine EIIA et EIIB phosphorylées appartenant à d'autres systèmes PTS, comme la protéine EIIA (pour revue, Saier et al., 1996; Crasnier-Mednansky et al., 1997).

## II.2.b / Régulation par la protéine EIIA Gle phosphorylée : la régulation transcriptionnelle par le complexe CRP/AMPc

La répression catabolique carbonée résulte de l'exclusion d'inducteur mais aussi de la répression apparente de la transcription de nombreux gènes impliqués dans le métabolisme carboné. Cette inhibition apparente est due à la diminution de la concentration du complexe CRP/AMPc, effecteur transcriptionnel régulant l'initiation de la transcription de nombreux gènes. Cet effecteur se lie sous forme de dimère au niveau de la séquence consensus symétrique suivante, où les séquences en gras sont les plus importantes : 5'-AAATGTGATCTAGATCACATTT-3' (Zheng et al., 2004). Elles sont situées en amont ou dans les promoteurs, et peuvent activer l'initiation de la transcription par un mécanisme de classe I ou II (Fig. 10) ou de classe III : le complexe CRP-AMPc agit alors en synergie avec d'autres effecteurs (Fig. 11).

#### II.2.b.α / Gènes régulés par le complexe CRP/AMPc

Différentes méthodes ont été utilisées pour déterminer les gènes régulés par le complexe CRP-AMPc. Une approche informatique basée sur la comparaison du génome d'E. coli avec celui de Salmonella typhimurium a identifié 161 sites forts et 285 sites faibles potentiels de fixation du complexe CRP-AMPc (Brown et Callan, 2004). Récemment, une approche de transcription in vitro a montré que 7 % des gènes d'E. coli seraient régulés directement par le complexe CRP-AMPc, permettant l'activation de 280 gènes appartenant à 188 opérons, et réprimant 20 gènes contenus dans 16 opérons (Zheng et al., 2004). In vivo, l'étude du transcriptome de souches mutées dans le gène crp comparée à celui de souche sauvage, cultivées sur un milieu LB avec ou sans glucose, montre que le gène crp affecte la transcription d'au moins soixante quinze gènes, directement ou indirectement (Gosset et al., 2004). Quelques gènes sont activés par le complexe CRP-AMPc en présence de glucose. Ils comprenent les gènes ptsG, codant la protéine EIIBC<sup>Glc</sup>, aceE codant le composant E1 de la pyruvate déshydrogénase, guaB codant l'inosine 5'monophosphate, le gène fis codant la protéine FIS, et celui codant l'ARNnc Spot42. L'ensemble des gènes du métabolisme carboné, réprimés par le glucose et dépendant du complexe CRP-AMPc, comprend de nombreux opérons des voies du métabolisme

carboné central, tel que les enzymes du cycle de Krebs et le shunt du glyoxylate, mais aussi les systèmes de transport et de métabolisme de sources de carbone alternatives ou d'acides aminés. Les gènes réprimés par le glucose, dépendant de la protéine CRP comprend aussi de nombreuses protéines de stress cytoplasmiques, dont les protéines chaperones majeures, 5 complexes protéases ATP-dépendantes, et différentes protéines de choc au froid ou au chaud. L'effet répresseur du glucose dépendant de la protéine CRP n'a pas permis d'identifier à chaque fois un site de liaison à CRP dans les régions de contrôles des opérons. Le rôle exercé par le complexe CRP-AMPc est donc à la fois direct et indirect, utilisant les facteurs de transcription MelR, RpoH (σ³²), BglG, FIS et PdhR (Zheng et al., 2004).

#### II.2.b.β / Régulation de la concentration intracellulaire en AMPc

#### Régulation de la concentration intracellulaire en AMPc par le système PTS

D'une manière générale, le catabolisme des sources de carbone dites PTS conduit à une baisse de la concentration intracellulaire en AMPc. Le mécanisme limitant la synthèse de l'AMPc est lié au transport des composés osidiques par le système PTS et en particulier au taux de phosphorylation de la protéine EIIA Glc (Saier et al., 1975, Saier et al., 1976). En effet, la mutation des enzymes EI, HPr ou la mutation de l'histidine phosphorylable de la protéine EIIAGle conduisent à une production basse et constante d'AMPc (Levy et al., 1990; Reddy et Kamireddi, 1998). L'adénylate cyclase, synthétisant l'AMPc, serait activée, via son domaine C-terminal, par la protéine EIIA Glc phosphorylée. En effet, l'adénylate cyclase est constituée de deux domaines : un domaine catalytique situé en N-terminal, et un domaine C-terminal permettant de limiter la production d'AMPc. Lorsque les adénylates cyclases sont délétées de la partie C-terminale, elles ne sont plus inhibées par la présence de glucose. Dans ce cas, elles synthétisent l'AMPc en forte concentration, ce qui abolit la répression catabolique (Crasnier et al., 1994). La présence de glucose dans le milieu de culture conduit à une déphosphorylation de la protéine EIIA Glc, donc à une absence d'activation de l'adénylate cyclase. La diminution de la concentration en AMPc par les sources de carbones PTS autres que le glucose pourrait impliquer un mécanisme analogue. Le transport de ces sources de carbone par leur système PTS conduirait à une baisse de la phosphorylation des protéines HPr et EI, ce qui induirait la déphosphorylation de la protéine EIIA Gle et limiterait l'activation de l'adénylate cyclase (Crasnier-Mednansky et al., 1997; pour revue, Postma et al., 1996a).

#### Variation de la concentration en AMPc en fonction de la concentration en glucose

La présence d'un sucre PTS inhibe l'entrée et le métabolisme d'autres sucres non PTS, seulement lorsque le sucre PTS est en forte concentration. En effet, les cellules d'*E. coli* cultivées en chémostat, en limitation de glucose, consomment d'autres sucres sans période de latence (Lendenmann et Egli, 1995). Ainsi, à faible concentration de glucose, le glucose n'exerce pas d'effet répresseur sur le

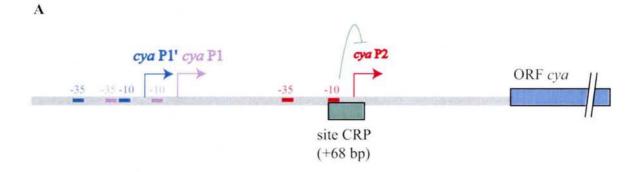
métabolisme de sources de carbone alternatives. Cette absence de répression catabolique carbonée résulte en partie d'une augmentation de la synthèse d'AMPc. Ceci permet à la bactérie de s'adapter à un milieu pauvre en glucose puisque le complexe CRP-AMPc favorise la synthèse des systèmes de transport de forte affinité pour le glucose, tels que le système Mgl et la glycoporine de la membrane externe LamB (pour revue, Ferenci, 1996). La variation de la concentration en AMPc a été alors mesurée par des méthodes directes et indirectes via la mesure de l'expression de gènes dépendant du taux d'AMPc. Il a ainsi été montré que, lorsque la concentration en glucose diminue, la concentration en AMPc intra et extracellulaire augmente brusquement à deux seuils de concentrations différentes : le premier est situé entre 0,3 et 0,4 mM de Glucose alors que le deuxième est à 0,1 mM (Notley-McRobb et al., 1997). De même, la concentration en AMPc diminue lors du catabolisme de sources de carbone PTS lorsque la concentration de ces sources de carbone dépasse 0,4 mM. Ainsi, lorsque du glucose est ajouté à une concentration finale supérieure à 0,4 mM à une culture carencée, la concentration intracellulaire en AMPc chute brusquement, sans que la concentration extracellulaire varie, et n'est pas liée à une augmentation de l'excrétion d'AMPc, comme cela fut supposé (Saier et al., 1975; Notley-McRobb et al., 1997).

#### Régulation de la concentration intracellulaire en AMPc par les sources de carbone non PTS

La présence de certaines sources de carbones dites non PTS, telles que le glucose-6-phosphate, le lactose, le gluconate, conduit également à une forte baisse de la concentration intracellulaire en AMPc, qui semble indépendante de la protéine EIIA<sup>Gle</sup>. En effet, la mutation du gène *crr* n'influence pas cette régulation (Hogema *et al.*, 1997). La surexpression du système de transport UhpT du glucose-6-phosphate dans la cellule s'accompagne d'une forte diminution de la concentration intracellulaire en AMPc, suggérant que le transport de ce composé dans la cellule pourrait être responsable de cette baisse (Dumay *et al.*, 1996). L'hypothèse serait qu'un sucre phosphaté, tel que le glycérol-3-phosphate ou le glucose-6-phosphate inhibe l'interaction de la forme phosphorylée de la protéine EIIA<sup>Gle</sup> avec l'adénylate cyclase et la production d'AMPc (Eppler *et al.*, 2002).

#### Régulation de la production d'AMPc par la protéine CRP

La délétion du gène crp conduit à une forte augmentation de la concentration intracellulaire en AMPc, montrant le rôle fondamental de la protéine CRP dans la régulation de la production d'AMPc (pour revue, Botsford et Harman, 1992). Cette régulation est exercée par cette protéine à la fois au niveau de la transcription du gène cya codant l'adénylate cyclase et de l'activité enzymatique correspondante (Aiba, 1985; Inada et al., 1996; Takahashi et al., 1998; Krin et al., 2002). En effet, la délétion du gène crp entraîne une légère augmentation de l'expression de la fusion d'opéron cya-lacZ, et une forte augmentation de l'activité adénylate cyclase par rapport à la souche sauvage. La répression de la transcription du gène cya est due à la fixation du complexe CRP/AMPc au niveau d'un site



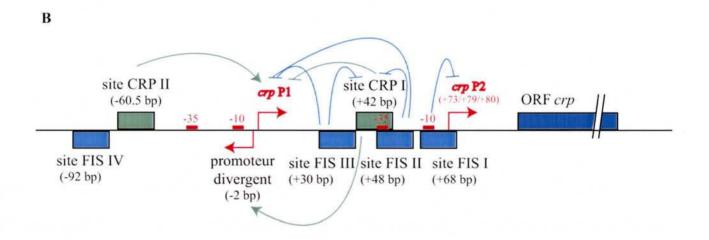


Figure 43 : Régulation de l'initiation de la transcription des gènes cya et crp.

A : Régulation de la transcription du gène *cya* (d'après Aiba, 1985). Le promoteur majoritaire *cya* P2 est indiqué en rouge, ainsi que ses séquences -10 et -35. Les promoteurs mineurs *cya* P1 et *cya* P'1 et leurs séquences -10 et -35 sont représentés respectivement en bleu et en violet. Le site de fixation du complexe CRP-AMPc, chevauchant le promoteur *cya* P2, est représenté en vert. Sa position, indiquant le centre de symétrie est donnée par rapport au site +1 d'initiation de la transcription du site *cya* P1.

B : Régulation de l'initiation de la transcription du gène *crp* (d'après Hanamura et Aiba, 1992 ; Ishizuka *et al.*, 1993 ; Gonzales-Gil *et al.*, 1998). Les promoteurs *crp* P1 et *crp* P2, ainsi que leurs boîtes -10 et -35, sont représentées en rouge. Les sites de fixation du complexe CRP-AMPc sont indiqués en vert, et ceux de la protéine FIS sont indiqués en bleu. Leur position est donnée par rapport au site d'initiation de la transcription du site *crp* P1.

La régulation de la transcription des différents promoteurs par la fixation des protéines effectrices sur leurs sites est schématisée par des flèches de même couleur que le site mis en jeu.

chevauchant le site majeur d'initiation de la transcription *cya* P2 (Fig. 43A). La forte activation de l'activité de l'adénylate cyclase dans les cellules *crp*<sup>-</sup> est liée à l'augmentation de la concentration de la protéine EIIA<sup>Gle</sup> phosphorylée, due à une diminution de la vitesse de déphosphorylation. En effet, la déphosphorylation de la protéine EIIA<sup>Gle</sup> est ralentie, *in vitro*, en présence de surnageant issu de cellules *crp*<sup>-</sup> par rapport à un surnageant issu de cellules isogéniques sauvages (Takahashi *et al.*, 1998). L'ajout de la protéine CRP au surnageant de cellules *crp*<sup>-</sup> ne suffit pas à stimuler la déphosphorylation de la protéine EIIA<sup>Gle</sup> phosphorylée. Ainsi, une ou plusieurs protéines sous le contrôle du complexe CRP-AMPc favoriseraient la déphosphorylation de la protéine EIIA<sup>Gle</sup> (Takahashi *et al.*, 1998). Il pourrait s'agir, entre autre, du transporteur du glucose EIIBC<sup>Gle</sup>, dont la transcription est activée par le complexe CRP-AMPc (Gosset *et al.*, 2004). Deux arguments sont en faveur de cette hypothèse : la protéine EIIA<sup>Gle</sup> phosphorylée interagit préférentiellement avec la protéine EIIBC<sup>Gle</sup> en milieu glucose, pour permettre l'entrée du glucose (Saier *et al.*, 1996) ; la surproduction de la protéine EIIBC<sup>Gle</sup> conduit à la déphosphorylation de la protéine EIIA<sup>Gle</sup> (Kimata *et al.*, 1997).

#### Régulation de la production de la protéine CRP

La diminution de la concentration du complexe CRP-AMPc reflète une diminution de la concentration en AMPc, mais aussi une baisse de la production de la protéine CRP. En effet, la présence de glucose conduit à une diminution d'un facteur 2 à 3 de la concentration de la protéine CRP (Ishizuka et al., 1993). La transcription du gène crp est à la fois activée et inhibée par le complexe CRP-AMPc (Fig. 43B) (Aiba, 1983; Hanamura et Aiba, 1992; Gonzalez-Gil et al., 1998). La région promotrice du gène crp présente deux sites d'initiation de la transcription : le promoteur fort crp P1 et le promoteur crp P2. L'activation de la transcription est liée à la fixation du complexe CRP-AMPc au niveau du site CRP II situé en position -61.5 par rapport au site d'initiation de la transcription du promoteur crp1, alors que la répression résulte de la fixation du complexe CRP-AMPc au site CRP I de forte affinité, situé en position +42 par rapport au site d'initiation de la transcription du promoteur crp P1. L'occupation du site CRP I activerait la transcription d'un ARN antisens au niveau d'un promoteur divergent, situé en position -2 par rapport au site crp P1 (Okamoto et Freundlich, 1986 ; Okamoto et al., 1988). La transcription du gène crp est aussi régulée par la protéine FIS. Deux sites de fixation de la protéine FIS recouvrent le site CRP I de fixation du complexe CRP-AMPc, un autre est situé au niveau du promoteur d'initiation de la transcription crp P2, et un quatrième est situé en amont du site CRP II. Ainsi, la protéine FIS module l'accessibilité des sites de fixation du complexe CRP-AMPc (Gonzalez-Gil et al., 1998). Le contrôle de la transcription du gène crp serait obtenu par les variations de compositions du complexe nucléoprotéique régulateur formé avec les protéines FIS et le complexe CRP/AMPc en réponse à l'état physiologique de la bactérie. Au début de la phase exponentielle de croissance, la protéine FIS, en forte concentration, se fixe au niveau des sites I à III, la fixation au site IV est inhibée par la fixation du complexe CRP-AMPc au niveau du site de forte affinité CRP II. Ce complexe CRP-AMPc active la transcription au niveau du promoteur *crp* P1. Au cours de la croissance, la concentration de la protéine FIS diminue, et libère les sites de fixation I à III. Le complexe CRP-AMPc peut se fixer au niveau du site CRP I, ce qui inhibe la transcription initiée au promoteur *crp* P2 aval, mais favorise la transcription du promoteur divergent au détriment de l'initiation de la transcription au promoteur *crp* P1. En phase exponentielle de croissance, la concentration de FIS et CRP-AMPc est faible, la transcription du gène *crp* pourrait être réalisée à partir des deux promoteurs.

## II.2.b.γ / Influence de la régulation de l'exclusion d'inducteur par le complexe CRP-AMPc

L'effet glucose, observé en milieu glucose/lactose, ne serait pas lié majoritairement à une diminution de la concentration en complexe CRP/AMPc, mais à l'exclusion d'inducteur (Kimata et al., 1997). En effet, l'ajout d'IPTG sur une souche d'E. coli possédant le promoteur lacL8UV5, promoteur du gène lacZ modifié de façon à être indépendant du taux de complexe CRP/AMPc, abolit l'exclusion d'inducteur. De même, la délétion du gène crr, ou la surproduction de la perméase au lactose LacY, élimine l'effet glucose. Le complexe CRP-AMPc semble plutôt jouer un rôle indirect via l'exclusion d'inducteur. Or, parmi les gènes dont l'expression est diminuée dans des souches délétées des gènes crp ou cya se trouve le gène ptsG. Sa surproduction restaure l'effet glucose dans les souches délétées de l'un ou l'autre des composants du complexe CRP-AMPc. Le complexe CRP-AMPc pourrait exercer un rôle crucial dans l'exclusion d'inducteur, responsable de la diauxie glucose-lactose, en activant l'expression du gène ptsG (Kimata et al., 1997).

#### II.2.c / Régulations exercées par les sources de carbones non PTS via la protéine EIIA Glc

Certaines sources de carbones dites non PTS peuvent induire une forte répression catabolique, bien qu'elles ne soient pas transportées par le PTS. C'est le cas en particulier du lactose, du glucose-6-phosphate, du gluconate et du glycérol-3-phosphate (Hogema et al., 1998a, Hogema et al., 1998b; Eppler et al., 2002). Cet effet dépend en partie de la variation du taux de phosphorylation de la protéine EIIA<sup>Glc</sup>. La déphosphorylation engendrée par le métabolisme du glucose-6-phosphate, du mélibiose, du lactose et de l'arabinose est quasiment aussi important que pour des sucres PTS (Hogema et al., 1998b). Ce phénomène serait lié au métabolisme de ces sources de carbones (Hogema et al., 1998a). En effet, ces sources de carbone non PTS sont métabolisées en pyruvate, par la voie de la glycolyse. Le rapport des concentrations entre le PEP et le pyruvate est alors diminué, ce qui doit limiter le taux de phosphorylation des enzymes EI, HPr et EIIA<sup>Glc</sup> du système PTS (Hogema et al., 1998a). Dans ce cas, la

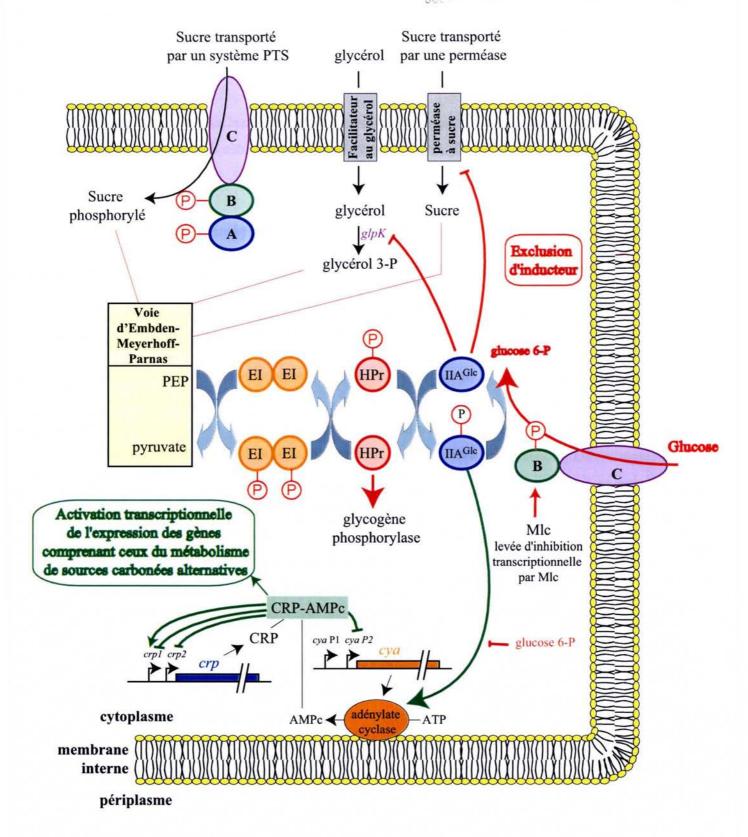


Figure 44 : Représentation schématique de la régulation du métabolisme carboné chez *E. coli* (d'après Gonzalez-Gil *et al.*, 1998; Aiba, 1985; Eppler *et al.*, 2002 ; pour revues, Postma *et al.*, 1993 ; Postma *et al.*, 1996).

Les protéines du système PTS sont représentées comme dans la Figure 2. Les systèmes de transport des sources carbonées autres que celles empruntant un système PTS sont indiquées en gris. L'entrée dans la cellule de substrat PTS induit une déphosphorylation des enzymes du PTS, dont la protéine EIIA<sup>Glc</sup>. Les différentes régulations engendrées par l'entrée d'un substrat PTS, ou d'un substrat non PTS induisant une répression catabolique carbonée sont indiquées en rouge, les régulations réalisées en absence de répression catabolique carbonée sont indiquées en vert.

forme non phosphorylée de la protéine EIIA Gle limiterait l'entrée des sources de carbone non PTS par le phénomène d'exclusion d'inducteur (Fig. 44).

Cependant, d'autres sources de carbone dites non PTS entraînent une répression catabolique en limitant l'interaction de la protéine EIIA<sup>Gle</sup> phosphorylée avec l'adénylate cyclase par un sucre phosphaté, tel que le glycérol-3-phosphate ou le glucose-6-phosphate (Fig. 44) (Eppler et al., 2002).

## II.3 / Régulation du métabolisme carboné par la protéine $EIIBC^{Glc}$ par l'intermédiaire du régulateur transcriptionnel Mlc

L'« effet glucose» correspond à la répression catabolique carbonée exercée par le glucose sur le transport et le métabolisme d'autres sources de carbone. Cet effet nécessite le transport du glucose par son système PTS spécifique. Il comprend les régulations exercées par les enzymes générales et EIIA<sup>Gle</sup>, mais aussi celle exercée par la protéine EIIBC<sup>Gle</sup>. En effet, le transport du glucose par son système PTS conduit à une déphosphorylation de la protéine EIIBC<sup>Gle</sup>, et une séquestration du répresseur transcriptionnel Mlc au niveau de la membrane (Nam et al., 2001). En absence de glucose, ce répresseur se fixe au niveau de l'ADN, et inhibe la transcription des gènes liés au transport et métabolisme en rapport au glucose. Il réprime l'expression de ces gènes ptsHI codant les enzymes générales EI et HPr, le gène ptsG, codant la protéine EIIBC<sup>Gle</sup>, l'opéron manXYZ, codant les enzymes du système PTS du mannose qui transporte le mannose et d'autres hexoses comme le glucose, et la protéine malT, l'activateur transcriptionnel du régulon maltose dont la dégradation conduit à la formation de glucose et glucose-1-phosphate. La protéine Mlc régule aussi l'expression de son propre gène (pour revue, Plumbridge, 2002).

La protéine Mlc a été identifiée lors de la recherche de gènes inhibant la synthèse d'acide acétique dans un milieu LB complémenté en glucose. Sa surproduction conduit à l'obtention de larges colonies sur boîtes LB complémentées en glucose, d'où son nom (make large colonie) (Hosono et al., 1995). Cette protéine, de 44 kDa, appartient à la famille de répresseurs transcriptionnels procaryotiques ROK (Represseurs, ORFs, kinases) (pour revue, Plumbridge, 2002). Elle interagit avec l'ADN par son domaine N-terminal pour réprimer la transcription des gènes cibles, au niveau de la séquence consensus suivante, dans laquelle les bases importantes sont soulignées (Plumbridge, 2001):

#### 5'-TTATTTN<sub>9</sub>AAATAA-3'

Cette séquence consensus, déterminée par la technique SELEX, est très proche de celle reconnue par le régulateur transcriptionnel NagC, qui présente 40 % d'identité avec la protéine Mlc (Hosono *et al.*, 1995; Plumbridge, 2001). Le régulateur NagC coordonne la biosynthèse d'acides aminés sucrés, tels que la glucosamine et la N-acetylglucosamine, avec leur catabolisme. La levée d'inhibition de la protéine Mlc est liée à sa séquestration au niveau de la membrane (Tanaka *et al.*, 2004). La protéine Mlc interagit

avec la protéine EIIBC<sup>Glc</sup> déphosphorylée par son motif de liaison à l'ADN en N-terminal (Seitz et al., 2003; Tanaka et al., 2004). Etant donné que certains régulateurs transcriptionnels exercent leur régulation depuis la membrane, la protéine Mlc semble subir un changement conformationnel lorsqu'elle s'associe au transporteur EIIBC<sup>Glc</sup>. Le domaine C-terminal de Mlc formerait une hélice amphipathique qui interagit avec la membrane, ce qui conduirait à une tétramérisation de la protéine Mlc ou un changement conformationnel de la protéine Mlc (Seitz et al., 2003).

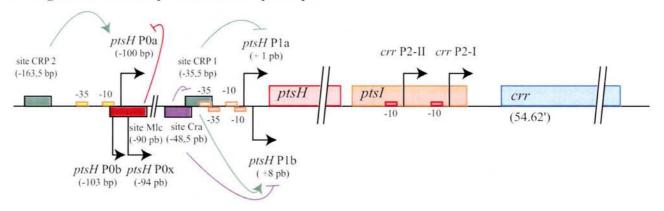
Le gène mlc est exprimé à partir de deux promoteurs P1 et P2, où P1 est reconnu par l'holoenzyme  $E\sigma^{70}$ , et P2 par  $E\sigma^{70}$  et  $E\sigma^{32}$ . Il est soumis à une faible autorégulation, un site de fixation à Mlc recouvre le site d'initiation P2. Un site de fixation du complexe CRP/AMPc permet la répression du promoteur P1, mais l'activation du promoteur P2. La production de Mlc est aussi régulée au niveau post-transcriptionnel. En effet, en présence de glucose, l'ARNm mlc est rapidement dégradé. Il est à noter que la production de Mlc est favorisée sous des conditions de croissance en anaérobie (pour revue, Plumbridge, 2002).

L'effet engendré par le transport de glucose par la protéine EIIBC<sup>Gle</sup> ne serait pas uniquement lié à la protéine Mlc. En effet, la transcription de la fusion transcriptionnelle *sdh-lacZ* est inhibée en présence de glucose, dans un milieu LB tamponné avec du phosphate (Takeda *et al.*, 1999). La surproduction de la protéine Mlc lève la répression par le glucose observée pour ces conditions de croissance. Par contre, la délétion du gène *mlc* n'a aucun effet. Ainsi, l'effet glucose, nécessitant le transporteur EIIBC<sup>Gle</sup>, ne semble pas directement lié à la protéine Mlc (Takeda *et al.*, 1999). De Reuse et Danchin (1991) ont proposé le modèle suivant pour l'activation par le glucose : la protéine EIIBC<sup>Gle</sup> serait la protéine détectrice d'un système à deux composants, associée à un régulateur transcriptionnel (De Reuse et Danchin, 1991). Il a été proposé que ce complexe pourrait activer la transcription des gènes *gapA*, *pgk*, mais inhiberait celle de *sdhCDAB*, codant la succinate déshydrogénase qui appartient au cycle de Krebs (Takeda *et al.*, 1999).

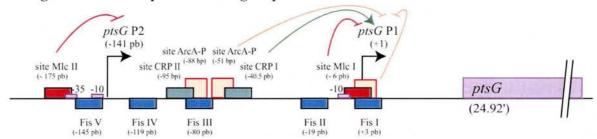
#### II.4 / Régulation de l'expression des gènes du système PTS du glucose

Les gènes codant les enzymes EII spécifiques d'un système PTS d'une source de carbone donnée sont généralement regroupés avec ceux responsables du catabolisme de cette source de carbone, sous la forme d'un opéron fortement régulé. L'organisation des gènes codant le système PTS spécifique du glucose est différente (pour revue, Postma et al., 1993; Postma et al., 1996c). Les protéines EII<sup>Glc</sup> comprennent la protéine EIIA<sup>Glc</sup> et IIBC<sup>Glc</sup>, codées respectivement par les gènes crr et ptsG. Le gène crr est localisé en aval des gènes ptsH et ptsI, codant respectivement les enzymes générales HPr et EI. L'ensemble forme ainsi l'opéron pts, situé à 55 minutes sur le chromosome d'E. coli. Le gène ptsG,

#### A : Régulation transcriptionnelle de l'opéron pts



#### B : Régulation transcriptionnelle du gène ptsG



#### C : Régulation post-transcriptionnelle du gène ptsG

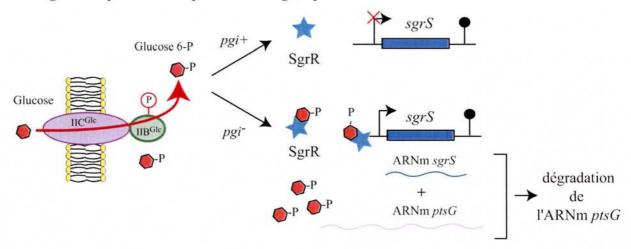


Figure 45 : Régulations transcriptionnelles de l'opéron pts et du gène ptsG chez E. coli (d'après Ryu et al., 1995 ; Shin et al., 2003 ; Jeong et al., 2004a ; Vanderpool et Gottesman, 2004).

A : Les promoteurs contenus dans l'opéron, ainsi que leurs boîtes -35 et -10 potentielles sont indiqués. Les sites de fixation des régulateurs transcriptionnels sont schématisés par des rectangles vert, rouge et violet pour le complexe CRP-AMPc, la protéine Mlc et la protéine Cra, respectivement.

B : Les promoteurs *ptsG* P1 et *ptsG* P2 sont indiqués en bleu, ainsi que leurs séquences -10 et -35 potentielles. Le sites de fixation pour le complexe CRP-AMPc, la protéine Mlc, la protéine FIS et le régulateur de réponse ArcA-P sont respectivement schématisés par des rectangles verts, rouges, bleus et orange clair. L'effet de l'occupation de ces différents sites est symbolisé par des flèches de couleur identique à celle du site de fixation correspondant.

C : Modèle de régulation post-transcriptionnelle de l'ARNm *ptsG*. Le même code couleur de la figure 2 est repris pour représenter le transporteur PTS du glucose. L'activateur transcriptionnel de l'ARNnc est représenté par une étoile, les ARNnc et ARNm par un trait ondulé, bleu et violet, respectivement.

situé à 25 minutes, est indépendant de cet opéron. Cette organisation particulière pourrait être due à la place fondamentale du métabolisme du glucose chez *E. coli*. Nous décrirons ici les régulations de l'expression de l'opéron *pts* et du gène *ptsG*.

#### II.4.a / Régulation de l'expression de l'opéron pts

La transcription des gènes *ptsH*, *ptsI* et *crr* de l'opéron *pts* est initiée à partir de sept promoteurs, qui peuvent être réunis en trois groupes *ptsH* P0, *ptsH* P1 et *crr* P2 (Fig. 45A) (De Reuse et Danchin, 1988; De Reuse *et al.*, 1992; Ryu *et al.*, 1995). Les gènes *ptsH* et *ptsI* sont transcrits à partir des groupes de promoteurs *ptsH* P0 et *ptsH* P1, alors que le gène *crr* est transcrit majoritairement indépendemment des autres gènes de l'opéron, à partir des promoteurs *crr* P2 I et P2 II, localisés dans le gène *ptsI*. Ces deux promoteurs *crr* P2 I et P2 II présentent chacun une boîte -10 étendue, sans boîte -35 apparente (De Reuse et Danchin, 1988). Ils ne semblent pas influencés par le complexe CRP-AMPc, ni par la présence de glucose dans le milieu de culture. Ils assurent ainsi une expression constitutive de la protéine EIIA<sup>Glc</sup> responsable de l'activation de l'adénylate cyclase en absence de glucose. Il est à noter que, dans une souche délétée du gène *crp*, l'expression du gène *crr* est activée par la protéine H-NS. La double mutation *crp hns* diminue la quantité de protéine EIIA<sup>Glc</sup> d'un facteur 2, ce qui a pour conséquences de diminuer la quantité de protéine EIIA<sup>Glc</sup> phosphorylée, donc l'activation de l'adénylate cyclase (Krin *et al.*, 2002).

L'initiation de la transcription au niveau des promoteurs ptsH P0 et ptsH P1 est plus complexe (Fig. 45A). Le promoteur ptsH P1 présente un site de fixation CRP2 pour le complexe CRP-AMPc de faible affinité, centré 35,5 bp en amont du site d'initiation P1a (De Reuse et al., 1992; Fox et al., 1992; Ryu et Garges, 1994a). La fixation du complexe CRP-AMPc sur ce site induit l'utilisation du site d'initiation P1b au détriment du site d'initiation P1a, situé 7 nucléotides en amont. La distance entre le centre du site CRP1 et le site d'initiation de la transcription P1b est alors de 42,5 pb, ce qui est compatible avec le mécanisme d'activation de la transcription de classe I par le complexe CRP-AMPc (Ryu et Garges, 1994a). Cet effet est limité par la présence de la protéine Cra, présentant un site de fixation chevauchant le site CRP 1. Bien que la protéine Cra peut se fixer simultanément avec le complexe CRP-AMPc au niveau de leur site de fixation, l'interaction de Cra avec l'ADN peut déplacer le complexe CRP/AMPc de son site de liaison en présence de l'ARN polymérase et favoriser ainsi le maintien de l'initiation transcriptionnelle au niveau du site P1a (Ryu et al., 1995). Ces régulations antagonistes semblent conduire à un taux constant d'initiation de la transcription du promoteur ptsH P1, quelles que soient les conditions de culture (De Reuse et al., 1992; Ryu et Garges, 1994a; Ryu et al., 1995). De plus, les transcrits initiés au niveau de ce promoteur sont stables, contrairement aux transcrits initiés au niveau du promoteur ptsH P0, permettant sans doute une expression constitutive des gènes de cet opéron (Ryu et al., 1995). Ainsi, les cellules semblent moduler la transcription de l'opéron pts au niveau du promoteur P0. La transcription est réalisée ici à partir de trois sites, P0a, P0b et P0x. Elle débute au site P0b au lieu du site P0a *in vitro* en présence d'un ADN matrice linéaire, ou au site P0x lorsque l'initiation de la transcription est réalisée par l'holenzyme Eσ<sup>32</sup> (Ryu et Garges, 1994a). L'initiation de la transcription au site P0a, qui présente une séquence -10 étendue et une boîte -35, est activée à la fois par le complexe CRP-AMPc et par le glucose (De Reuse *et al.*, 1992; Ryu et Garges, 1994a; Plumbridge, 1999). Le complexe CRP-AMPc possède un site de forte affinité, CRP2 centré 63,5 pb en amont du site d'initiation P0a, et active la transcription par un mécanisme de classe I (De Reuse *et al.*, 1992; Fox *et al.*, 1992). La régulation de la transcription par le glucose est réalisée par la protéine Mlc, qui possède un site de fixation au niveau du site d'initiation de la transcription P0 (Plumbridge, 1999).

#### II.4.b / Régulation de l'expression du gène ptsG codant la protéine EIIBC Glc

#### II.4.b. $\alpha$ / Régulation de l'initiation de la transcription du gène ptsG

La transcription du gène *ptsG* est réalisée à partir des deux promoteurs *ptsG* P1 et *ptsG* P2 (Fig. 45B). Elle est régulée positivement par le complexe CRP-AMPc, et négativement par la protéine Mlc, mais aussi par la protéine FIS et le régulateur de réponse ArcA-P. Le promoteur *ptsG* P1 contribue majoritairement à l'expression de ce gène, le promoteur *ptsG* P2, d'efficacité plus faible, est localisé en amont (Plumbridge, 1998b). Le promoteur *ptsG* P1 ne présente pas de séquence -35 potentielle, mais possède un site de fixation pour le complexe CRP-AMPc, centré 40,5 pb en amont du site d'initiation de la transcription : l'efficacité de ce promoteur est strictement dépendante de ce complexe (Kimata *et al.*, 1997 ; Kimata *et al.*, 1998 ; Plumbridge, 1998b). En présence de glucose, la répression de la transcription du gène *ptsG* par la protéine Mlc est levée, ce qui conduit à une augmentation de l'efficacité transcriptionnelle du promoteur *ptsG* P1, et une légère augmentation de l'activité du promoteur alternatif *ptsG* P2 (Plumbridge, 1998b). La protéine FIS favorise à la fois la transcription par le complexe CRP-AMPc, mais aussi la répression par la protéine Mlc (Shin *et al.*, 2003).

La transcription du gène *ptsG* est aussi régulée en fonction des conditions redox de la cellule (Jeong *et al.*, 2004a). Lorsque le transport d'électrons n'est pas assez important, un transport trop important de sucres conduit à la production de sous-produits acides du métabolisme, tels que l'acide acétique. La bactérie régule alors l'entrée des sucres par une répression de la transcription réalisée par le régulateur de réponse ArcA-P au niveau du promoteur *ptsG* P1 (Fig. 45B).

#### II.4.b.β / Régulation post-transcriptionnelle du gène ptsG

Lorsque la glycolyse est inhibée, soit par une mutation dans les gènes pgi, pfkA ou fbaA, codant les trois premières enzymes de la glycolyse, ou soit par l'utilisation d'α-méthylglucoside, analogue du glucose qui entre et est phosphorylé par le système PTS du glucose mais n'est pas métabolisé, l'ARNm ptsG est rapidement clivé par la RNase E dans la région 5' (Kimata et al., 2001; Morita et al., 2003). Cette dégradation est réalisée lorsque l'ARNm ptsG est traduit et la protéine EIIBC Glc comprenant les deux premiers domaines transmembranaires est insérée au niveau de la membrane interne. Cette localisation de l'ARNm ptsG a proximité de la membrane interne limite la fixation des ribosomes pour de nouveaux cycles de traduction, et favoriserait l'interaction de l'ARNm ptsG avec l'ARNnc SgrS et la protéine Hfq (Vanderpool et Gottesman, 2004; Kawamoto et al., 2005). La transcription de l'ARNnc SgrS est activé en présence de sucres phosphatés : ceux-ci interagiraient avec l'activateur SgrR, produit du gène transcrit de façon divergente par rapport à l'ARNnc SgrS, ce qui active la transcription de l'ARNnc SgrS. D'autres mécanismes de régulations peuvent aussi intervenir dans la régulation post-transcriptionnelle de l'ARNm ptsG. Ainsi, le ppGpp, dont la concentration augmente lors de la mutation du gène fbaA ou de l'utilisation d'α-méthylglucoside (Cashel et al., 1996; Schneider et Gourse, 2003) pourrait jouer un rôle dans cette régulation. En effet, le ppGpp module l'activité de la RNase P, RNase qui intervient dans la dégradation des ARNm dont les produits sont impliqués dans le transport. La RNase P pourrait alors dégrader l'ARNm ptsG lors d'une augmentation de la concentration en ppGpp.

## III / Régulations des voies du métabolisme carboné exercées indépendemment du système PTS du glucose

Les composants du système PTS ne sont pas les seuls éléments contrôlant le métabolisme carboné de la bactérie. Nous décrirons ici les régulations exercées par le régulateur transcriptionnel Cra et les systèmes à deux composants PhoBR et ArcAB.

#### III.1 / Le régulateur FruR, appelé aussi Cra

#### Rôle de la protéine Cra

La protéine Cra (Catabolic repressor/activator) exerce un mécanisme de répression catabolite

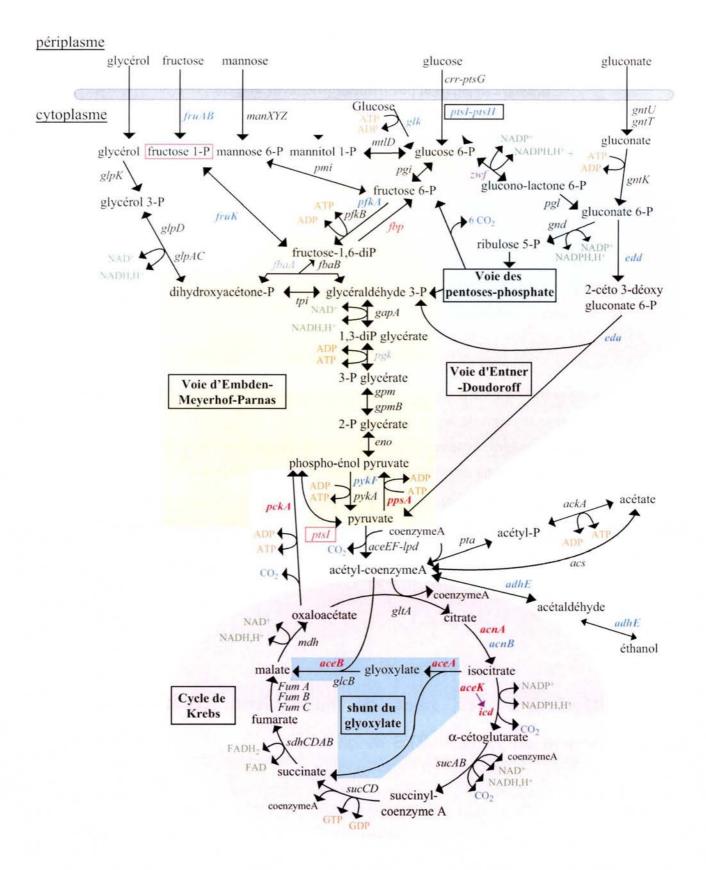


Figure 46 : Régulation de l'initiation de la transcription des gènes du métabolisme carboné exercée par la protéine Cra.

Les gènes dont la transcription est activée par la protéine Cra sont indiqués en rouge, tandis que ceux, dont la transcription est réprimée, sont indiqués en gras et en bleu clair. Les gènes pgk et fbaA sont localisés en aval du gène epd, dont l'expression est réprimée par la protéine Cra, sont indiqués en bleu clair. Un site de fixation à la protéine Cra a été retrouvé en amont du gène zwf, indiqué en violet, où le rôle de Cra n'a pas encore été déterminé.

indépendante du complexe CRP-AMPc. Cette protéine a été identifiée initialement lors de la recherche de mutants suppresseurs de la mutation *ptsH*, capables de se développer sur des milieux contenant des sucres dits PTS (Saier *et al.*, 1970). La protéine Cra, appelée anciennement FruR, réprime la transcription de l'opéron *fruBKA*. Cet opéron code les trois protéines impliquées dans le transport et le métabolisme spécifique du fructose (Reizer *et al.*, 1994). Les protéines FruA (IIB'BC) et FruB (Domaine FPr en Cterminal, et IIA<sup>Fru</sup> en N-terminal) assurent le transport et la phosphorylation du fructose en fructose 1-phosphate (Fig. 46). La fructose 1-phosphate kinase (*fruK*) phosphoryle le substrat en fructose-1,6-diphosphate (Reizer *et al.*, 1994). La mutation du gène *fruR* altère l'expression de nombreux gènes, dont les produits sont impliqués entre autre dans la glycolyse, la gluconéogenèse, le cycle de Krebs, le shunt du glyoxylate (Fig. 46). La protéine FruR, dont l'inactivation montre des effets pléiotropes, a été renommée Cra pour Catabolite repressor/activator (pour revue, Saier, 1996).

La protéine Cra semble limiter l'expression des gènes dont les produits sont impliqués dans la dégradation glycolytique des sources de carbones, tandis qu'elle favoriserait l'expression des gènes intervenant dans la gluconéogenèse (Ramseier et al., 1995; pour revues, Ramseier, 1996; Saier, 1996). Ainsi, en plus de l'opéron fru, la protéine Cra inhibe la transcription des opérons mtl et edd-eda codant les enzymes impliquées respectivement dans le métabolisme spécifique du mannitol et la voie de dégradation d'Entner-doudoroff (Ramseier et al., 1995). Les enzymes générales du système PTS, EI et HPr, sont réprimées au niveau de la transcription par la protéine Cra chez S. typhimurium (Chin et al., 1989). Chez E. coli, la protéine Cra se fixe au niveau de la région régulatrice du promoteur ptsH P1 de l'opéron pts, mais ne jouerait pas un rôle important dans la régulation de l'expression cet opéron chez E. coli (Ryu et al., 1995). De même, l'expression des gènes pfkA et pykF, dont les produits catalysent les deux réactions irréversibles de la glycolyse, est également limitée par la fixation de la protéine Cra au niveau de leur promoteur (Bledig et al., 1996; Negre et al., 1996). Enfin, elle limite l'expression du gène glk, codant la glucokinase (Meyer et al., 1997). L'inactivation de la protéine Cra favorise la croissance des bactéries en présence de sucres glycolytiques, mais permet d'utiliser moins efficacement les acides aminés et les intermédiaires du cycle de Krebs (Ramseier et al., 1995). Ainsi, la transcription des gènes spécifiquement impliqués dans la gluconéogenèse, comme les gènes ppsA ou pckA, codant la PEP synthase et la PEP carboxykinase respectivement, est activée par la protéine Cra (Geerse et al., 1989; Negre et al., 1998; Ramseier et al., 1993; Ramseier et al., 1995). La transcription du gène icd de l'isocitrate déshydrogénase, enzyme du cycle de Krebs, et l'opéron aceBAK, codant les deux enzymes du shunt du glyoxylate et l'isocitrate déshydrogénase kinase/phosphatase semblent être activée par la protéine Cra (Chin et al., 1989; Ramseier et al., 1993).

La protéine Cra active aussi la transcription de l'opéron *cydAB*, codant la cytochrome d oxydase, intervenant dans la chaîne respiratoire (Ramseier *et al.*, 1996), mais réprime l'expression de l'alcool déhydrogénase *adhE*, responsable de la fermentation alcoolique (Mikulskis *et al.*, 1997).

#### Mécanisme d'action

La protéine Cra est composée de quatre sous-unités de 334 acides aminés et 37 kDa, et appartient à la famille des régulateurs transcriptionnels LacI/GalR. Elle possède un domaine Nterminal de liaison à l'ADN HTH, et un domaine C-terminal permettant l'interaction entre les différentes sous-unités et la fixation d'inducteurs (Mikulskis et al., 1997 ; pour revue, Saier, 1996). En effet, la fixation de la protéine Cra aux séquences cibles est inhibée par des concentrations micromolaires de fructose 1-phosphate et millimolaires de fructose-1,6-diphosphate (pour revue, Saier, 1996). La séquence consensus de fixation de la protéine tétramérique Cra est partiellement palindromique (Ramseier et al., 1995) : (G/A)(C/G)TGAA(A/T) \* C(C/G)NT(A/T/C)(A/T/C)(A/T). Cette séquence présente une meilleure conservation de séquence dans le demi-site de gauche par rapport au demi-site de droite (Cortay et al., 1994; pour revue, Saier et Ramseier, 1996). Ceci est en accord avec la fixation asymétrique de Cra au niveau du palindrome imparfait, avec une meilleure interaction réalisée dans le demi-site gauche par rapport au demi-site de droite (pour revue, Saier et Ramseier, 1996). La nature dégénérée de la séquence consensus pourrait refléter sa fixation avec différentes affinités à ses sites cibles et les effets pléiotropes engendrés lors de l'inactivation de la protéine Cra (pour revue, Saier et Ramseier, 1996). Ses séquences cibles sont retrouvées en amont des promoteurs, lorsque ceux-ci sont activés par la protéine Cra, ou au niveau du promoteur ou en aval lorsque les promoteurs sont inhibés par la protéine Cra. Lorsque la protéine Cra active la transcription, sa fixation génère une forte courbure de l'ADN, de l'ordre de 120 degré par exemple au niveau de l'opérateur du gène ppsA. Cette courbure faciliterait l'interaction de l'ADN avec les domaines α-CTD de l'ARN polymérase et donc la formation des complexes ouverts lors de l'initiation de la transcription (Negre et al., 1998; Prost et al., 1999).

De nombreux gènes, controlés par la protéine Cra, sont également régulés par d'autres effecteurs transcriptionnels, comme FNR pour la transcription de l'opéron *cydAB*. Dans ce cas, la fixation de la protéine Cra nécessite la fixation préalable du régulateur transcriptionnel FNR (Ramseier *et al.*, 1996). Certains gènes sont régulés de façon inverse par la protéine Cra et le complexe CRP/AMPc, tel que le gène *pckA* et l'opéron *pts* (pour revue, Saier et Ramseier, 1996). D'autre part, la protéine Cra régule, en milieu minimum M63 fructose, la concentration en AMPc, sa délétion diminuant la concentration en AMPc (Crasnier-Mednansky *et al.*, 1997). L'hypothèse serait que la délétion du gène *cra* déréprime l'expression de l'opéron *fru*, dont la protéine FruB et son domaine FPr. La phosphorylation de FPr est réalisée au détriment de HPr, et donc de la protéine EIIA<sup>Glc</sup>, ce qui diminue le niveau de phosphorylation de la protéine EIIA<sup>Glc</sup>, et donc l'activation de l'adénylate cyclase. Ainsi, la protéine Cra pourrait réguler indirectement l'expression des gènes contrôlés par le complexe CRP-AMPc (Crasnier-Mednansky *et al.*, 1997).

#### III.2 / Les systèmes de transduction du signal à deux composants

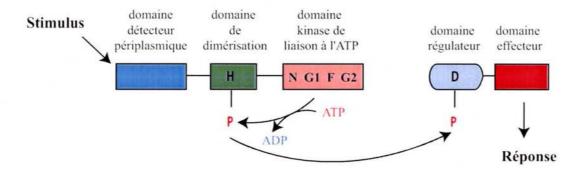
#### Définition

Les bactéries sont des organismes fortement adaptables, pouvant occuper une large variété de niches écologiques. Leur adaptation résulte de leur capacité à exprimer uniquement les gènes codant des enzymes et des voies métaboliques nécessaires pour obtenir une croissance maximale pour l'environnement dans lequel elles se trouvent. Un mécanisme majeur de reconnaissance de l'environnement conduisant à l'expression de gènes spécifiques est assuré par les systèmes à deux composants, et son variant plus complexe, le phosphorelai (pour revue, Hoch et Varughese, 2001).

Le système de transduction du signal à deux composants, ubiquitaires chez les bactéries à l'exception de *Mycoplasma genitalium*, est constitué d'une protéine détectrice (appelée aussi histidine kinase ou HK), sous forme de dimère, qui mesure un paramètre environnemental et d'un régulateur de réponse cytoplasmique (RR), comme le système PhoBR qui répond à la concentration en phosphate (Fig. 47A) (pour revue, Foussard *et al.*, 2001). Lorsque la protéine détectrice détecte le signal environnemental, elle phosphoryle le régulateur de réponse. Celui-ci réalise une réponse adaptative via généralement une variation de l'expression des gènes cibles. Les phosphorelais se différencient de ces systèmes à deux composants par la présence d'un domaine régulateur et phosphotransférase supplémentaires, tels que ArcAB, système phosphorelai fonction de l'état redox de la bactérie (Fig. 47B). En absence de signal, la protéine HK peut quelques fois déphosphoryler le RR.

La protéine détectrice, appelée aussi protéine kinase à histidine, comprend un domaine détecteur variable, pouvant contenir jusqu'à 500 aa, et un domaine transmetteur de 250 résidus, généralement conservé au niveau de la superfamille des protéines à histidine kinase (pour revue, Foussard et al., 2001). Différentes boîtes d'homologies ont pu être définies dans ce domaine transmetteur. La boîte H comprend l'histidine phosphorylable, et est constituée d'un fagot de 4 hélices à partir de la dimérisation de deux domaines ou d'un seul domaine appelé unité de phosphotransfert HPt. Les boîtes N, G1, F et G2 délimitent la cavité de liaison de l'ATP, et sont situées à proximité du domaine H (Fig. 47B). La détection du signal, réalisée au niveau du domaine détecteur de la protéine kinase à histidine, provoque un changement conformationnel dans la portion transmembranaire de la protéine, ce qui active l'activité kinase du domaine transmetteur, et la phosphorylation du résidu histidine. Le domaine receveur du régulateur de réponse autocatalyse ensuite le transfert du phosphate du résidu histidine au résidu aspartate, ce qui engendre un changement conformationnel et la libération du domaine régulateur exerçant la régulation. En absence de stimulus, la plupart des protéines kinase à histidine déphosphorylent leur régulateur de réponse. Ainsi, l'action combinée kinase et phosphatase des protéines kinase à histidine ajuste ainsi la quantité relative des formes phosphorylées du régulateur de réponse, en fonction des différents signaux environnementaux.

Les régulateurs de réponses peuvent aussi être phosphorylés, en absence des protéines à



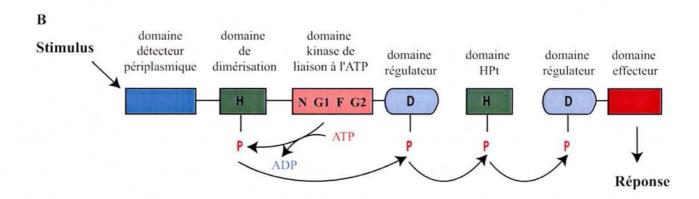


Figure 47 : Schéma de transfert des phosphates au niveau des systèmes à deux composants (pour revue, West et Stock, 2001).

A : Le système de transfert de phoshate à deux composants "classique" : il comprend une protéine détectrice dimérique transmembranaire et un domaine cytoplasmique régulateur de réponse. Un monomère de la protéine détectrice est représenté ici. Les motifs de séquences conservées N, G1, F et G2 sont localisés au niveau du domaine de liaision de l'ATP. Les activités des protéines détectrices sont modulées par les signaux de l'environnement. Le groupe phosphate est transféré au résidu acide aspartique spécifique (D) localisé au niveau du domaine régulateur. La phosphorylation du régulateur de réponse active le domaine effecteur, qui engendre la réponse cellulaire spécifique.

B : Le système de transfert de phoshate phosphorelai. Il comporte une protéine détectrice hybride, qui comporte un domaine regulateur de réponse additionnel au niveau de son domaine C-terminal. Le transfert de phosphate réalisé en trois étapes, His-Asp-His-Asp, implique un intermédiaire de transfert de phosphate (HPt), avec une histidine phosphorylée.

histidine kinase respectives, par des petites molécules, telles que l'acétyl phosphate (Lukat et al., 1992; Danese et Silhavy, 1998). La spécificité de phosphorylation des régulateurs de réponse par ces petites molécules, suggère que ces molécules serviraient de signal pour contrôler les régulateurs de réponses en fonction de l'état métabolique de la bactérie (Lukat et al., 1992).

Les systèmes à deux composants sont impliqués dans différentes régulations en fonction des paramètres de l'environnement. Nous décrirons ici un exemple de système à deux composants, le système PhoBR, puis d'un phosphorelai, le système ArcAB.

#### Le système à deux composants PhoBR : activation lors d'une limitation en phosphate

Lors d'une limitation en phosphate, ce composé est transporté par le système Pst de forte affinité (voir Partie A). Par un mécanisme encore inconnu, l'état fonctionnel de ce système de transport est mesuré par le senseur PhoR. Ainsi, en présence de forte concentration en phosphate, PhoR déphosphoryle et inactive le régulateur de réponse PhoB. Par contre, lors d'une faible concentration en phosphate transportée par le système Pst, ou d'une mutation dans le système de transport Pst, le senseur PhoR phosphoryle PhoB, qui active, ou réprime parfois, la transcription des gènes et opérons du régulon PHO. Ces derniers sont caractérisés par la présence d'une boîte PHO (5' CTGTCATA(A/T)A(A/T)CTGTCA(C/T)3') sur laquelle vient se fixer le régulateur transcriptionnel PhoB phosphorylé. La phosphorylation de l'activateur transcriptionnel PhoB améliore l'affinité de ce régulateur pour ses séquences cibles (Wanner, 1996).

En absence de PhoR, le régulateur de réponse PhoB peut alors être phosphorylé par la protéine détectrice CreC du système à deux composants CreC/CreB, en réponse à la présence de glucose, d'acétate ou de pyruvate en milieu minimum (Avison et al., 2001; pour revue, Bijlsma et Groisman, 2003). En effet, en milieu riche, l'autophosphorylation de CreC semble être inhibée (Avison et al., 2001). En milieu minimum, la protéine CreB phosphorylée se fixe au niveau de la séquence TTCACnnnnnTTCAC, et régule la transcription des gènes cibles : il réprime la transcription de malE de l'opéron transport du maltose malEFG, mais active l'opéron ackA-pta, les gènes talA, radC, trgB, creD, yidS et yieI. Les produits de l'opéron ackA-pta sont impliqués dans la conversion de l'acétyl-CoA en ATP et acétate (Fig. 4). Le gène talA convertit le G3P et sedoheptulose-7-phosphate en fructose-6-phosphate et E4P, dans la voie des pentoses phosphates (Fig. 4). Le gène radC code une protéine possèdant une fonction de réparation/recombinaison de l'ADN comme RecG, le gène trgB, une ADP-ribose phosphorylase, la fonction des protéines issues des trois derniers gènes n'est pas connue.

#### Le phosphorelai ArcAB: activation lors de conditions redox réductrices

Dans le cas des phosphorelais, le transfert du phosphate est réalisé en trois étapes His-Asp-His-Asp (Fig. 47B). Ainsi, le système phosphorelai permet d'intégrer différents signaux tout au long de la cascade de phosphorylation, comme pour le phosphorelai contrôlant la sporulation chez *Bacillus subtilis* (pour revue, Bijlsma et Groisman, 2003).

Le système de transduction du signal ArcAB (Anoxic redox control) d'E. coli, activé lors de conditions anaérobies, régule l'expression, directement ou indirectement, d'au moins 9 % du génome d'E. coli (Liu et De Wulf, 2004; Salmon et al., 2005). En particulier, il régule les gènes et opérons impliqués dans le métabolisme respiratoire et fermentaire, mais aussi dans la synthèse des flagelles, le système de transport du nickel, la dégradation des acides gras à courtes chaines, le cycle TCA, un gène de superoxyde dismutase et un gène impliqué dans la division cellulaire (Fig. 48) (Liu et De Wulf, 2004; Salmon et al., 2005). Il comprend la protéine kinase ArcB et le régulateur de réponse ArcA. La protéine ArcB contient trois domaines cytoplasmiques, comprenant le domaine transmetteur H1 avec l'histidine 292, le domaine receveur D1 contenant l'aspartate 576 et le deuxième domaine transmetteur H2 possédant l'histidine 717. La transmission du signal est réalisée uniquement par chaque intermédiaire du phosphorelai His-Asp-His-Asp (Kwon et al., 2000). L'histidine 292 s'autophosphoryle, puis l'aspartate 576 récupère le phosphate, ce qui accélère le transfert du phosphate à l'histidine 717. L'aspartate 54 du régulateur de réponse ArcA est ensuite phosphorylé par l'histidine 717. Ceci entraîne une multimérisation de ArcA sous forme d'un tétramère de dimère, composé à proportion égale de protéines ArcA phosphorylée et non phosphorylée (Jeon et al., 2001). La protéine ArcA phosphorylée se fixe sur l'ADN au niveau de la séquence consensus 5'-GTTAATTAAATGTTA-3' (Liu et De Wulf, 2004).

Le système ArcAB répond à l'état rédox de la cellule grâce aux quinones oxydées. Les quinones, associées à la membrane, sont des transporteurs d'électrons entre deux complexes enzymatiques donneur et accepteur d'électrons, respectivement. En présence d'oxygène, ou de conditions redox oxydantes, les quinones oxydent deux cystéines de ArcB, localisées sur un domaine cytoplasmique. Cette formation du pont disulfure conduit à une inhibition de l'autophosphorylation de ArcB (Georgellis et al., 2001; Malpica et al., 2004). ArcB agit aussi en tant que phosphatase de ArcA, catalysant sa déphosphorylation (Georgellis et al., 1998). Lors de conditions réductrices, ArcB s'autophosphoryle en présence d'ATP. Cette réaction est favorisée par la présence de certains métabolites anaérobiques tels que le D-lactate, l'acétate, le pyruvate et le NADH (Iuchi, 1993; Georgellis et al., 1999; Rodriguez et al., 2004).

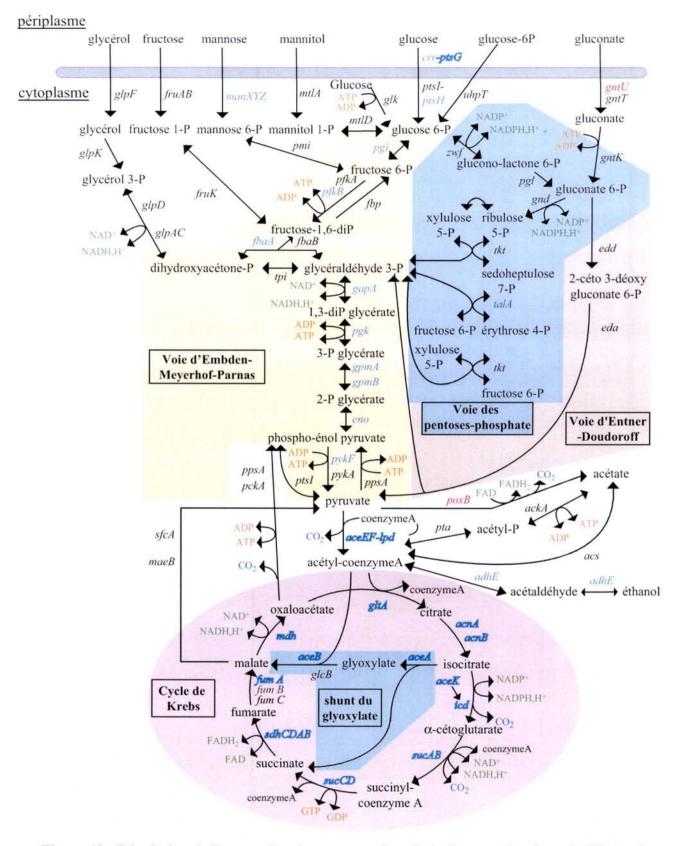
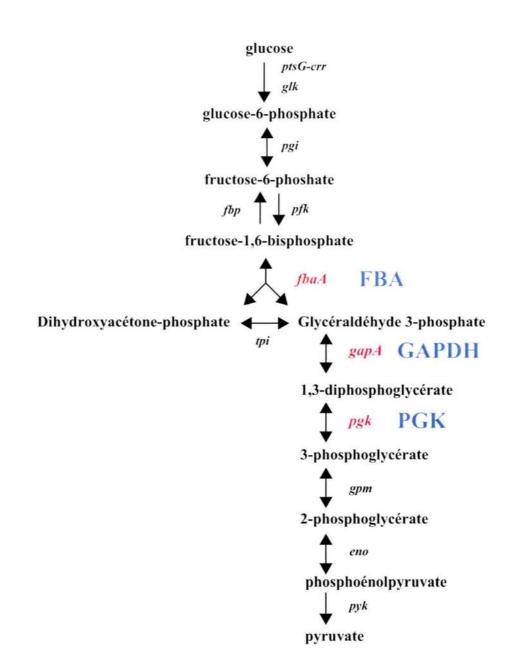


Figure 48: Régulation de l'expression des enzymes du métabolisme carboné par le RR ArcA chez *E. coli* (d'après Liu et De Wulf, 2004; Salmon *et al.*, 2005 et http://www.cifn.unam.mx/Computational\_Genomics/regulondb/).

Les différentes voies métaboliques (voie d'Embden-Meyerhof-Parnas, voie des pentoses phosphates, voie d'Entner-doudoroff, cycle de Krebs et shunt du glyoxylate) sont indiquées sur un fond de couleur différentes. Nous avons ici repertoriés la variation d'expression des gènes d'un facteur supérieur ou égal à 2 lors d'une culture anaérobie en présence ou non du RR. Les gènes dont l'expression est inhibée ou activée par le RR ArcA-P sont indiqués respectivement en bleu ou rouge. Lorsqu'une régulation transcriptionnelle directe a été mise en évidence, les gènes sont indiqués en gras.

# Présentation du Sujet



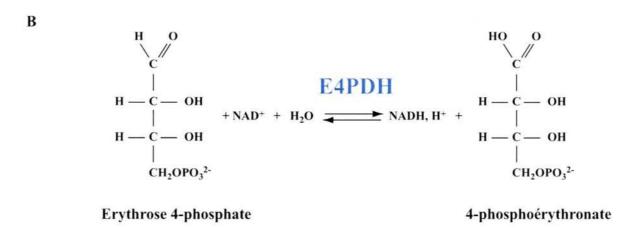


Figure 49 : Représentation de la glycolyse (A) et de la réaction catalysée par l'E4PDH (B).

Les gènes codant les enzymes de la glycolyse ainsi que les enzymes FBA, GAPDH et PGK sont indiquées en A.

#### Partie A / Les enzymes FBA, GAPDH, E4PDH et PGK et leur expression

La voie de la glycolyse, appelée aussi voie d'Embden-Meyerhof-Parnas, est retrouvée chez quasiment tous les organismes vivants. Elle conduit à la formation du pyruvate à partir du glucose (Fig. 49) (pour revue, Fraenkel, 1996). Toutes les réactions de la glycolyse peuvent être réalisées dans le sens de la gluconéogenèse à partir soit d'enzymes catalysant uniquement la réaction inverse, soit des mêmes enzymes, comme pour la FBA, la GAPDH et la PGK. Ces trois enzymes sont fortement produites, et l'organisation des gènes codant ces enzymes est différente suivant les organismes. Chez la majorité des bactéries, le gène gap codant la GAPDH active, est localisé dans un opéron contenant des gènes codant d'autres enzymes de la glycolyse. Ainsi, le gène gap est fréquement retrouvé en amont du gène pgk, comme pour Zymomonas mobilis. Le gène tpi, codant la triose phosphate isomérase (TIM), est retrouvé en aval du tandem de gène gap-pgk chez les bactéries à Gram positif, comme B. megaterium (Schläpfer et Zuber, 1992) ou C. glutamicum (Eikmanns, 1992), chez les spirochètes et les thermotogales (Bardey, 2000). Le gène fbaA, codant une FBA de classe II, est quelque fois cotranscrit avec d'autres gènes codant des enzymes catalysant des réactions réversibles de la glycolyse, tel que la GAPDH et la PGK chez l'algue Euglena gracilis, ou du cycle de Calvin-Benson-Bassham, tel que la ribulose bisphosphate carboxylase/oxygénase et la phosphoribulokinase chez Rhodobacter sphaeroides et Xanthobacter flavus (Plaumann et al., 1997; van den Bergh et al., 1996). Il est à noter que le gène fbaA est retrouvé en aval du gène pgk principalement chez les protéobactéries de la subdivision y (Bardey, 2000).

Dans cette subdivision γ des protéobactéries, deux gènes codent des protéines présentant une forte similitude avec les GADPH, situés à des loci différents sur le chromosome (Bardey, 2000). En particulier, chez *E. coli*, le gène *gapA*, codant la GAPDH active (Hillman et Fraenkel, 1975), est localisé à 40,1 minutes sur le chromosome d'*E. coli*, alors que le gène *gapB* est situé à 66,2 minutes, en amont des gènes *pgk* et *fbaA* (Alefounder et Perham, 1989; Berlyn, 1998). Ce gène *gapB*, renommé *epd*, code une protéine présentant plus de 40 % d'identité en acides aminés avec celle de la GAPDH. Elle porte une activité érythrose-4-phosphate déshydrogénase non phosphorylante (E4PDH), et une faible activité GAPDH (Zhao *et al.*, 1995; Boschi-Muller *et al.*, 1997), tandis que la GAPDH d'*E. coli*, codée par le gène *gapA*, possède une forte activité GAPDH phosphorylante et une faible activité E4PDH phosphorylante (Boschi-Muller *et al.*, 1997).

La GAPDH, codée par le gène *gapA*, présente une plus forte similitude avec les GAPDH d'eucaryotes qu'avec les GAPDH d'origine bactérienne (Branlant et Branlant, 1985). Le gène *gapA* aurait été acquis par un transfert horizontal de matériel génétique d'un ancêtre eucaryote vers un ancêtre des bactéries, alors que le gène *epd* aurait pu être généré par mutation d'un gène *gap* bactérien ancestral (Branlant et Branlant, 1985; Martin *et al.*, 1993).

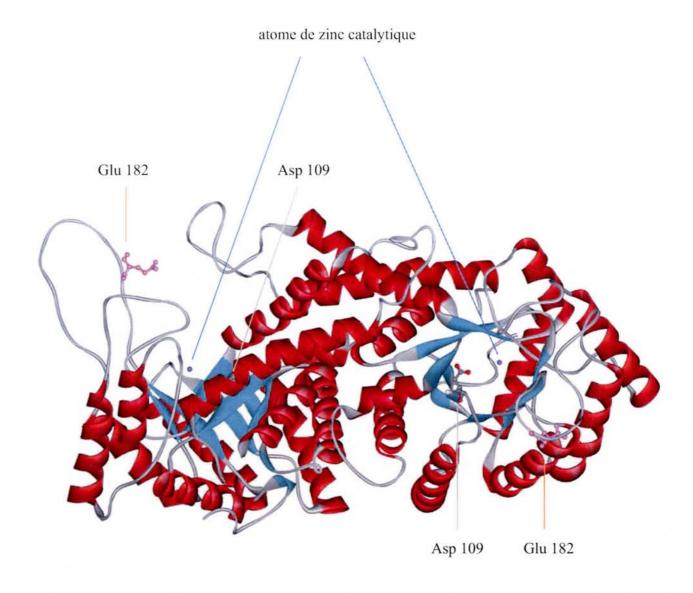


Figure 50 : Structure quaternaire du dimère de la Fructose-1,6-bisphosphate aldolase de classe II de type A chez E. coli (d'après Cooper et al., 1996).

Les hélices  $\alpha$  et les brins  $\beta$  sont indiqués en rouge et vert, respectivement. Les atomes de zinc catalytiques sont représentés par des sphères bleues.

Nous décrirons ici les rôles, structures et activités enzymatiques de la FBA, de la GAPDH, de l'E4PDH et de la PGK, puis l'expression des gènes codant ces enzymes.

#### I / La FBA

Les FBA catalysent le clivage du fructose-1,6-bisphosphate en dihydroxyacétone phosphate (DHAP) et glycéraldéhyde 3-phosphate (G3P). Elles possèdent une architecture en tonneau  $(\alpha/\beta)_8$ . Les FBA ont été subdivisées en deux classes, suivant le mécanisme catalytique.

#### I.1 / Les deux classes de la FBA

Les **FBA** de classe I sont caractérisées par un mécanisme passant par l'intermédiaire d'une base de Schiff, formé entre le carbonyle en C2 du DHAP ou FBP, et l'amine d'un résidu lysine. De ce fait, ces enzymes sont inactivées par une réduction par le borohydrate de sodium (NaBH<sub>4</sub>) en présence de leur substrat (DHAP ou FBP), mais elles ne sont pas inhibées par l'EDTA (pour revue, Horecker *et al.*, 1972).

Deux types de FBA de classe I peuvent être définis. Chez les eucaryotes, les FBA de classe I sont des homotétramères de 160 kDa. Par contre, les FBA de classe I retrouvées chez toutes les Archeae, à l'exception de *T. acidophilum*, et la moitié des bactéries, présentent une faible homologie de séquence avec celles des eucaryotes, de l'ordre de 20 à 30 %, et sont appelées FBA de classe IA (Siebers et al., 2001). Elles ont une structure allant du monomère au décamère, où le poids moléculaire du monomère varie de 27 à 40 kDa (Plaumann et al., 1997; Thomson et al., 1998). De plus, elles peuvent utiliser le fructose-1-phosphate, avec une plus faible affinité que le fructose-1,6-bisphosphate, et leur activité est stimulée par le citrate (Thomson et al., 1998; Siebers et al., 2001).

Les FBA de classe II ont un mécanisme catalytique dépendant d'un métal, généralement l'atome de zinc et sont inhibées par l'EDTA, mais pas par le NaBH<sub>4</sub> contrairement aux FBA de classe I. Les FBA de classe II se différencient aussi par leur structure oligomérique des aldolases de classe I : ce sont des homodimères de 80 kDa. Elles ont été mises en évidence essentiellement chez les procaryotes et les Archeae (Plaumann et al., 1997; Henze et al., 1998). Deux types d'enzymes, A et B, ont été définis dans cette classe II. Ces enzymes utilisent le même mécanisme catalytique, mais présentent une faible identité de séquence (25 à 30 %) (Plaumann et al., 1997). Les enzymes de classe IIA interviennent dans la glycolyse et néoglucogenèse, alors que les enzymes de type IIB peuvent intervenir dans le cycle de Calvin, ou porter une autre activité, comme la tagatose bisphosphate aldolase (Plaumann et al., 1997).

Glu 182

L'Asp 109 joue le rôle de base et déprotone le fructose-1,6-bisphosphate, ce qui conduit à la formation du glycéraldéhyde-3-phosphate.

Le glycéraldéhyde-3-phosphate formé est libéré.

intermédiaire enzyme énediolate

Le Glu 182 joue le rôle de base et donne son proton, ce qui permet la formation et la libération du dihydroxyacétone phosphate.

Figure 51: Mécanisme catalytique de la FBA de classe II chez E. coli (d'après Zgiby et al., 2002).

La majorité des organismes ne possèdent qu'une seule classe de FBA, bien que quelques organismes possèdent les deux classes, comme l'eucaryote *Euglena gracilis* ou la bactérie *E. coli*. Les enzymes FBA de classe I et II d'*E. coli* caractérisées montrent une faible identité de séquence, de l'ordre de 16 %, ce qui suggère une évolution indépendante de ces deux enzymes.

#### I.2 / Les deux classes de la FBA présentes chez E. coli

La FBA de classe IA est codée par le gène *fbaB* (appelé anciennement *dhnA*), localisé à 46.9 minutes sur le chromosome d'*E. coli*. L'enzyme est un oligomère constitué de 8 à 10 monomères, d'environ 340 kDa (Baldwin et Perham, 1978; Thomson *et al.*, 1998). Sa production est induite sur un substrat gluconéogénique. Ainsi, cette enzyme contribue à l'activité FBA de la cellule jusqu'à 60 % lorsque la croissance d'*E. coli* est réalisée sur un milieu contenant du lactate ou du pyruvate, alors qu'elle ne représente que 0 à 5 % de l'activité FBA lorsque la croissance d'*E. coli* est réalisée sur glucose (Baldwin et Perham, 1978; Scamuffa et Caprioli, 1980). *E. coli* possède un autre gène dont la protéine présente une forte homologie aux FBA de classe IA, le gène *yneB* (Galperin *et al.*, 2000).

La FBA de classe II, de type A est codée par le gène *fbaA*, localisé en aval du gène pgk (Alefounder et Perham, 1989). Elle est constituée de deux sous-unités de 359 aa (40 kDa). Sa structure tridimensionnelle, obtenue par cristallographie à une résolution de 2,5 Å, montre une architecture en tonneau  $(\alpha/\beta)_8$ , où les deux dimères sont orientés perpendiculairement l'un par rapport à l'autre (Fig. 50) (Cooper *et al.*, 1996). Dans les familles d'enzymes possèdant cette architecture, le site catalytique est situé dans une poche formée par les côtés C-terminal des brins  $\beta$  et les boucles les reliant aux hélices  $\alpha$  (pour revue, Nagano *et al.*, 2002). L'Asn 35, la Ser 61 et l'Arg 331 sont impliqués dans la liaison au glycéraldéhyde-3-phosphate (Qamar *et al.*, 1996; Zgiby *et al.*, 2000).

La FBA réalise le clivage du fructose-1,6-bisphosphate en deux étapes. L'Asp 109 déprotone le carbone C4 du fructose-1,6-bisphosphate, ce qui conduit à la formation et à la libération du G3P (Fig. 51). Le Glu 182, situé dans la boucle β5-α7 voit son pKa modifié par le mouvement de cette boucle vers son substrat, ce qui lui permet de donner un proton au substrat et former le DHAP (Zgiby et al., 2002).

Les ORF agaY et gatY d'E. coli sont des aldolases de classe IIB. Ces deux gènes présentent de fortes homologies aux FBA de classe II et codent des protéines qui clivent le D-tagatose-1,6-bisphosphate en G3P et DHAP. La D-tagatose-1,6-bisphosphate aldolase 1, codée par le gène agaY, serait impliqué dans le transport et le métabolisme du N-acétylgalactosamine ou N-acétylmannosamine, tandis que la D-tagatose-1,6-bisphosphate aldolase 2, codée par le gène gatY, joue un rôle dans le métabolisme du galactitol (Zgiby et al., 2000).



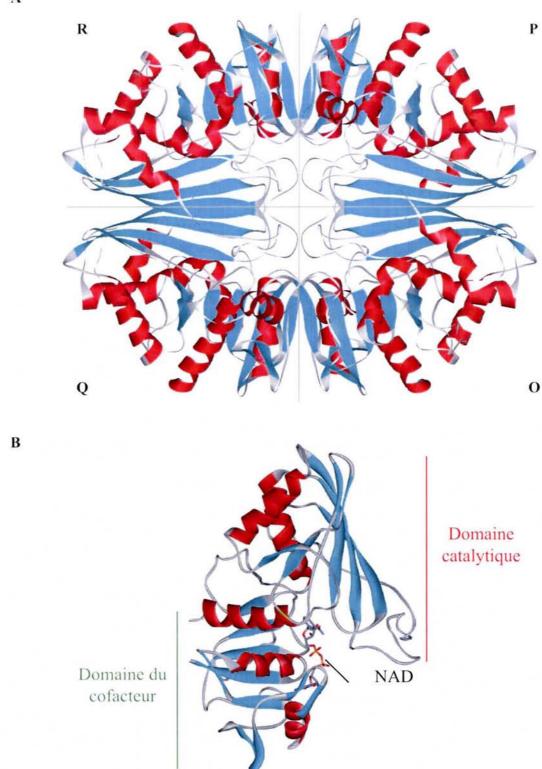


Figure 52 : Représentation schématique de la structure tridimensionnelle de la GAPDH de Bacillus stearothermophilus et E. coli.

A : Représentation schématique de la structure quaternaire de la GAPDH de *Bacillus stearothermophilus* (d'après Skarzynski *et al.*, 1987). Les différentes sous-unités O, P, Q et R sont indiquées. B : Représentation schématique de la structure d'un monomère de la GAPDH d'*E. coli*. Il comprend le domaine de fixation du cofacteur (domaine de Rossman) et le domaine catalytique. Les hélices α sont colorées en rouge et les brins β en bleu. La molécule de NAD<sup>+</sup> est indiquée (d'après Yun *et al.*, 2000).

#### II / La GAPDH et l'E4PDH

#### II.1 / Rôles de la GAPDH et de l'E4PDH

Les GAPDH interviennent au niveau de la glycolyse, gluconéogenèse et au niveau du cycle de Calvin. Trois familles de GAPDH ont pu être distinguées suivant leur rôle physiologique qui peut être corrélé au cofacteur utilisé. En effet, les GAPDH phosphorylantes NAD-dépendantes sont principalement impliquées dans la glycolyse. Elles catalysent la phosphorylation oxydative du D-glyceraldéhyde 3-phosphate (D-G3P) en 1,3-diphosphoglycérate (1,3diPG), en utilisant comme cofacteur le NAD<sup>+</sup> (pour revue, Harris et Waters, 1976). Elles possèdent un résidu Asp en position 32 et Leu-Pro en position 187-188, qui inhibent l'utilisation du cofacteur NADP<sup>+</sup> par encombrement stérique et/ou interactions électrostatiques (Corbier et al., 1990; Clermont et al., 1993).

Les GAPDH NADP-dépendante, à spécificité de cofacteur duale NAD(P) intervenant dans le cycle de Calvin, ou la gluconéogénèse, catalysent préférentiellement la réaction de déphosphorylation réductrice du 1,3diPG en D-G3P, en utilisant préférentiellement le NADP<sup>+</sup> comme cofacteur. Une corrélation entre rôle physiologique et spécificité de cofacteur a été mise en évidence chez B. subtillis: la GAPDH phosphorylante NAD-dépendante, codée par le gène gapA, est impliquée dans la glycolyse, tandis que la GAPDH NADP-dépendante, à spécificité de cofacteur duale NAD(P), codée par le gène gapB, intervient dans la gluconéogenèse (Fillinger et al., 2000).

La troisième famille de GAPDH catalyse l'oxydation non-phosphorylante du D-G3P en acide 3-phosphoglycérate. Chez les eucaryotes, la GAPDH non phosphorylante, retrouvée chez les organismes photosynthétiques, génère du NADPH à partir de NADP, nécessaire pour les réactions de biosynthèse. Chez S. mutans, la GAPDH non phosphorylante permettrait de compenser l'absence de la glucose-6-phosphate déshydrogénase et de la 6-phosphogluconolactonase, les deux premières enzymes du cycle des pentoses phosphates qui conduit à la production de NADPH, (Boyd et al., 1995).

Le rôle de la GAPDH n'est pas limité à son intervention dans la glycolyse et dans le cycle de Calvin. Ainsi, certaines GAPDH de *Streptococcus*, localisées dans les membranes cellulaires, semblent favoriser la virulence de ces bactéries (Gase et al., 1996). Chez les eucaryotes, et en particulier dans les cellules de mammifères, de nombreux rôles ont été attribués à la GAPDH. Elle interviendrait, entre autre, dans la réparation et la réplication de l'ADN, la maturation des ARNt du noyau, l'export d'ARN du noyau, la régulation de la traduction, l'endocytose (pour revue, Sirover, 1999). Elle serait aussi un catalyseur de l'apoptose chez les cellules humaines : une forte concentration de GAPDH nucléaire a été retrouvée dans des échantillons prélevés post mortem chez des patients atteints de la maladie de Parkinson, d'Alzheimer, de la chorée de Huntington. L'utilisation de certains composés pouvant

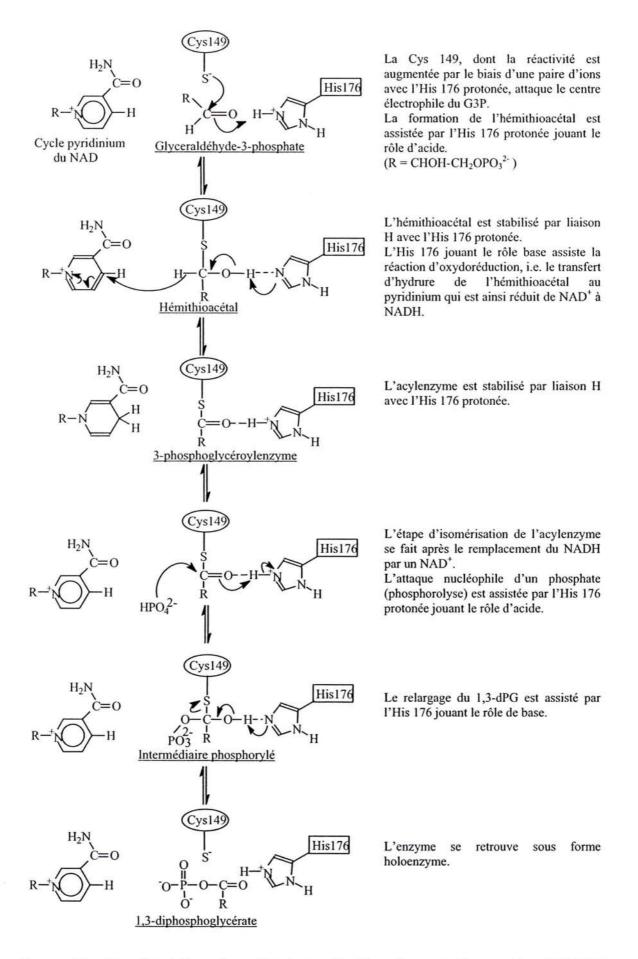


Figure 53: Représentation du mécanisme réactionnel proposé pour les GAPDH phosphorylantes (d'après Segal et Boyer, 1953).

interagir avec la GAPDH et inhiber son accumulation dans le cytoplasme montrent une activité antiapoptotique (berry, 2004).

Chez E. coli, trois gènes gap ont été identifiés. Le gène gapA code la GAPDH active, tandis que le gène gapB code une érythrose 4-phosphate déshydrogénase (E4PDH), et le gène gapC ne semble pas coder une GAPDH active du fait de la présence de différents codons stop dans l'ORF (Hillman et Fraenkel, 1975; Boschi-Muller et al., 1997). L'E4PDH diffère de la famille de GAPDH phosphorylantes NAD-dépendantes par sa spécificité de substrat. Elle réalise l'oxydation de l'érythrose 4-phosphate en 4-phosphoérythronate et montre une plus forte activité E4PDH non phosphorylante qu'une activité GAPDH phosphorylante (Boschi-Muller et al., 1997). Le groupe de Winkler a supposé l'implication de l'E4PDH dans la biosynthèse du pyridoxal 5'-phosphate (Yang et al., 1998; Zhao et al., 1995). Ce coenzyme est essentiel à de nombreuses enzymes impliquées dans le métabolisme des acides aminés et aux glycogènes phosphorylases (pour revue, Helmreich, 1992). Cependant, la croissance de souches délétées au niveau du gène epd, codant l'E4PDH, et gapA, codant la GAPDH, n'est pas arrêtée du fait de l'absence de synthèse du coenzyme (Della-Seta et al., 1997). Le rôle physiologique de l'E4PDH n'a pas encore été déterminé.

#### II.2 / Structure et mécanisme des GAPDH phosphorylantes NAD-dépendantes

Les GAPDH sont constituées de quatre sous-unité de 330 aa (37 kDa) (pour revue, Harris et Waters, 1976). La structure tridimensionnelle de GAPDH phosphorylantes NAD-dépendantes, avec ou sans NAD<sup>+</sup>, a été obtenue à haute résolution pour différents organismes et en particulier pour B. subtilis et E. coli (Fig. 52) (Skarzynski et al., 1987; Duee et al., 1996; Yun et al., 2000). Chaque monomère peut être divisé en deux domaines distincts. Le domaine de fixation du cofacteur est composé des 148 résidus amino-terminaux, ainsi que de l'hélice α<sub>4</sub> carboxy-terminale (résidus 310 à 330). Ce domaine contient le motif βαβαβ appelé domaine de Rossman, caractéristique de la fixation de cofacteurs (Rossmann et al., 1974). Le domaine catalytique est constitué des résidus 149 à 309, qui forment un feuillet β antiparallèle, ainsi que trois hélices α et d'une longue boucle peu structurée, la S-loop, qui intervient dans la tétramérisation de la GAPDH. Ce domaine contient la cystéine 149 et l'histidine 176, ainsi que les sites de fixation des substrats et produits de la réaction. Les sites anioniques Pi et Ps semblent correspondre aux sites de fixation du phosphate inorganique et du groupement phosphate du G3P respectivement (Didierjean et al., 2003). Le site Ps est composé des chaînes latérales des résidus arginine 231 et thréonine 179, et du groupement hydroxyle en 2' du ribose du nicotinamide lorsque le NAD est lié; le site Pi comprend les chaînes latérales des résidus sérine 148 et thréonine 208 et la chaîne principale de la glycine 209.

|              |           | 1                           | 10                                       | 20                                   | 30  | 40                                 |
|--------------|-----------|-----------------------------|--|--------------------------------------|---|------------------------------------|
| E4PDH        | MT\       | VR <b>VAINGE</b>            | FGRIGRNVVR                               | <b>a</b> lyesgr <b>r</b> aei         | TV <b>VAINE</b> LAD                           | <b>a</b> ag <b>ma</b> hl <b>lk</b> |
| <b>GAPDH</b> | MT        | I K <b>v</b> g <b>ing</b> f | GRIGRIVFR                                | <b>A</b> aqK <b>r</b> sdi            | EIVAIN <mark>D</mark> LLD                     | <b>A</b> DY <b>ma</b> ym <b>lk</b> |
|              |           |                             |  |                                      | _   |                                    |
|              | 45        | 50                          | 60                                       | 70                                   | 80  | 90                                 |
| E4PDH        |           |                             | E <b>v</b> rqerdq <b>l</b> f             | <b>V</b> GDDA <b>IRV</b> LH          | ERSLQSLPWRE                                   |                                    |
| GAPDH        | YDS       | THGRFDG                     | T <b>v</b> evkdgh <b>l</b> i             | VNGKKIRVTA                           | ERDPANLK <b>W</b> DE                          | VGVDVVAEA                          |
|              |           |                             |  |                                      |   |                                    |
|              | 0.7       | 100                         | ***                                      | 120                                  | 120   | 1.10                               |
| E4PDH        | 95<br>TGV | 100<br>YGSR <b>E</b> HG     | 110<br>EAHIAAGAKK                        | 120<br>VLFSHPGSND                    | 130<br>LDATV <b>V</b> Y <b>G</b> V <b>N</b> O | 140<br>Dolraehri                   |
| GAPDH        |           |                             |  |                                      | NTPMF <b>V</b> K <b>G</b> A <b>N</b> F        |                                    |
| GAIDII       | 101       | IL DIDEIN                   | RAMILIAGARA                              | V VMIGE-BRD                          | NIFM VIGAN                                    | DRI-AGQD1                          |
|              |           |                             | 1.00                                     | .=0                                  | 100   | 100                                |
| E4PDH        | VCN       | 150                         |  | 170                                  | 180<br>Tihsa <mark>m</mark> hd <b>q</b> qv    | 190                                |
|              |           |                             |  |                                      |   |                                    |
| GAPDH        | VSN       | ASCTTNC                     | LAPLAKVIND                               | NF <b>GI</b> IE <b>G</b> LM <b>T</b> | TVHAT <mark>T</mark> AT <b>Q</b> KT           | VDGPSHKDWR                         |
|              |           |                             |  |                                      |   |                                    |
| EARDII       | 195       | 200                         | 210                                      | 220                                  | 230<br>FEAIAVRVPTI                            | 240                                |
| E4PDH        |           |                             |  |                                      |   |                                    |
| GAPDH        | GGR       | GASQNII                     | P <mark>SS</mark> T <mark>GA</mark> AKAV | gkvl <b>p</b> el <b>n</b> gk         | LTGM <b>A</b> F <b>RVPT</b> F                 | NVSVV <b>DL</b> TV                 |
|              |           |                             |  |                                      |   |                                    |
|              | 245       | 250                         |  | 270                                  |   | 290                                |
| E4PDH        | TVK       | <b>K</b> PVKANE             | VNLLLQK <b>AA</b> Q                      | GAFH <b>G</b> IVD <b>YT</b>          | ELPL <b>VS</b> V <b>DFN</b> H                 | IDPHSAIV <b>D</b> G                |
| <b>GAPDH</b> | RLE       | <b>K</b> AATYEQ             | I KAAVKA <b>AA</b> E                     | EGEMK <b>g</b> vlg <b>yt</b>         | EDDV <b>vs</b> t <b>dfn</b> (                 | GEVCTSVF <b>D</b> A                |
|              |           |                             |  |                                      |   |                                    |
|              | 295       | 300                         | 310                                      | 320                                  | 330   |                                    |
| E4PDH        | TQT       | RVSGAHL                     | IKTLVW DNE                               | WGFA <b>n</b> rm <b>ld</b> t         | TLAMATVAFR                                    |                                    |
| <b>GAPDH</b> | KAG       | SIALNDNF                    | V <b>K</b> LVS <b>WYDNE</b>              | ETGYS <b>n</b> KV <b>ld</b> L        | IAHISK  |                                    |
|              |           |                             |  |                                      |   |                                    |

Figure 54 : Alignement des séquences des protéines GAPDH et E4PDH d'E. coli.

Les résidus en gras sont conservés chez la GAPDH et l'E4PDH d'*E. coli* K12. Les résidus non conservés, spécifiques de la GAPDH ou de l'E4PDH sont encadrés en rouge. La numérotation est indiquée par rapport à la méthionine de l'E4PDH.

$$E_{-S^{-}}^{NAD^{+}} \bigcap_{R}^{OH} Erythrose 4-phosphate}$$

$$E_{-S^{-}}^{NAD^{+}} \bigcap_{R}^{OH} H\acute{e}mithioac\acute{e}tal$$

$$E_{-S^{-}}^{NADH} \bigcap_{R}^{O} Thioacylenzyme$$

$$hydrolyse \downarrow H_{2}O$$

$$E_{-SH}^{NADH} \bigcap_{R}^{O} + NADH, H^{+}$$

$$\downarrow OH$$

$$R = C_{-}^{O} 4-phospho\acute{e}rythronate$$

Figure 55 : Représentation du mécanisme réactionnel proposé pour les E4PDH.

La GAPDH catalyse l'oxydation phosphorylante du G3P en 1,3 diPG via un mécanisme chimique en deux étapes (Fig. 53). La première est une étape d'oxydoréduction qui conduit à la formation de l'intermédiaire thioacylenzyme et du NADH. L'attaque nucléophile par la cystéine 149 sur le carbone aldéhydique du G3P serait rendue possible, à pH physiologique, par la diminution du pK<sub>app</sub> de la Cys 149 de 8,0 à 5,9 par l'His 176. Lors de la deuxième étape de phosphorylation, correspondant à l'étape limitante de la réaction, l'attaque nucléophile du phosphate inorganique envers l'intermédiaire thioacylenzyme conduit à la formation du 1,3diPG.

#### II.3 / Structure et mécanisme de l'E4PDH

L'E4PDH possède une structure très similaire à la GAPDH (Kauffmann et al., résultats non publiés). Elle contient, entre autre, les acides aminés essentiels à la catalyse, la Cys 149 et l'His 176, ainsi que la plupart des résidus impliqués au niveau des sites Pi et Ps (Fig. 54) (La numérotation des résidus est celle préalablement proposée pour les GAPDH par Biesecker et al. (1977). Cependant, certains résidus conservés chez les GAPDH ne sont pas retrouvés chez les E4PDH. Ainsi, l'aspartate 32 impliqué dans l'interaction avec le groupement nII.3 / Structure et cteur est substitué par un glutamate. La thréonine 179 du site Ps corresponda mécanisme de lu phosphate du G3P est remplacée par une méthionine. La séquence SSTGA ent 8, impliquée dans la fixation du phosphate inorganique, est substituée par la séc e substitution pourrait être impliquée dans l'activité non phopshorylante des E t al., 1997). Enfin, le résidu thréonine 311, fortement conservé dans les GAPDH et localisé à proximité du site de fixation du cofacteur est remplacé par une cystéine.

La réaction d'oxydoréduction non phosphorylante de l'E4P est réalisée en deux étapes, comme dans le cas de la GAPDH (Fig. 55) (Boschi-Muller et al., 1997). Lors de la première étape d'acylation, l'attaque nucléophile de la Cys 149 sur le substrat E4P conduit à la formation d'un intermédiaire thioacylenzyme. Contrairement au mécanisme observé pour les GAPDH phosphorylantes, l'étape de déacylation est une hydrolyse par une molécule d'eau. L'étape limitante de la réaction est ici l'étape d'acylation.

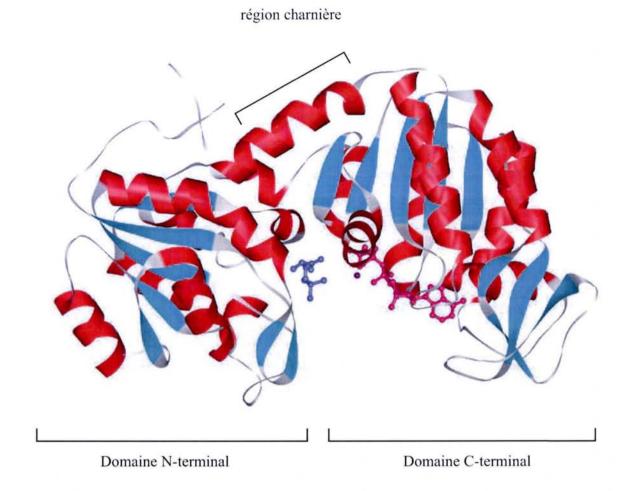


Figure 56: Structure tertiaire de la PGK de Thermotoga maritima (d'après Bernstein et al., 1997).

Les hélices  $\alpha$  et les feuillets  $\beta$  sont indiquées en rouge et bleu, respectivement. La molécule d'ATP est indiquée en rose, le G3P en bleu.

#### III / La PGK

#### III .1 / Rôles de la PGK

La PGK intervient au niveau de la glycolyse, de la gluconéogenèse et de la photosynthèse. Elle réalise le transfert d'une liaison phosphate riche en énergie de l'ATP au 3-phosphoglycérate (3PG), ce qui conduit à la formation de 1,3 di-phosphoglycerate et d'ADP (pour revue, Scopes, 1973).

La PGK joue aussi un rôle au niveau de l'angiogénèse des cellules cancéreuses (Lay et al., 2000). Elle intervient dans la réduction de la plasmine, qui est ensuite clivée par différentes protéases pour aboutir à la formation de l'angiostatine, un inhibiteur de l'angiogénèse. L'administration de PGK à des souris portant une tumeur conduit à une augmentation de la concentration d'angiostatine dans le plasma, à une diminution de la vascularisation de la tumeur et une diminution de la croissance de la tumeur.

#### III .2 / Structure et mécanisme de la PGK

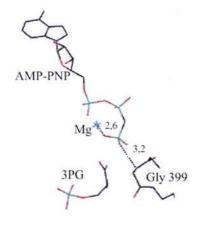
Le monomère de la PGK, dont le poids moléculaire est estimé à 41 kDa (387 aa), comprend un domaine N-terminal et un domaine C-terminal reliés entre eux par une large région charnière, comme pour la PGK de *T. maritima* (Fig. 56). Le G3P se lie au domaine N-terminal, alors que l'ADP-Mg<sup>2+</sup> ou son analogue non hydrolysable comme l'AMP-PNP, interagit avec le domaine C-terminal (Bernstein *et al.*, 1997). La fixation de chacun de ses éléments conduit à une torsion de la région charnière située entre les deux domaines. La torsion finale, de l'ordre de 32°, rapproche les deux substrats ce qui permet ensuite un transfert direct d'un groupement phosphorylé. La PGK catalyse la réaction en stabilisant l'état de transition (Fig. 57B) (Bernstein et Hol, 1998).

# IV / Transcription de l'opéron epd-pgk-fbaA et du gène gapA chez E. coli

#### III.1 / Transcription de l'opéron epd-pgk-fbaA

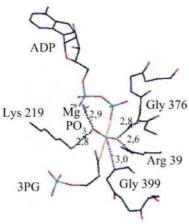
La caractérisation des protéines produites par le triplet d'ORF *epd-pgk-fbaA* montrait que l'E4PDH est faiblement produite, contrairement aux protéines PGK et FBA (Alefounder et Perham, 1989).

A



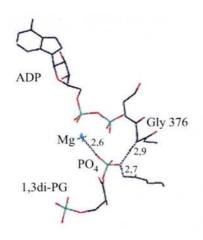
Deux oxygènes du phosphate en position  $\gamma$  de l'analogue non hydrolysable de l'ATP, l'AMP-PNP, est stabilisé par une interaction avec l'ion  $Mg^{2+}$  et le Gly 399.

B



Dans l'état de transition, le phosphate est en conformation bipyramide trigonale. Les trois atomes d'oxygènes chargés négativement du phosphate sont stabilisés un premier par l'Arg 39, le Gly 376, un second par le Gly 399 et le troisième par Lys 219 et l'ion Mg<sup>2+</sup>.

C

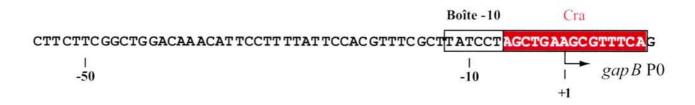


Deux oxygènes du phosphate de C1 est stabilisé par l'ion Mg<sup>2+</sup>, l'Arg 39 et le Gly 376.

Figure 57 : Mécanisme réactionnel proposé pour la PGK de *Trypanosoma brucei* (d'après Bernstein et Hol, 1998).

CRP-AMPc (-70,5)

# A CGCC AGCCT GGAA TAAAG CTTCC CACA AAAGG GGATA AAAG TGTGA TGTGAGTCA GA TAA ATGT -100



# TC GATTA AATGT TCGACAA TTAA CCAAT CAGTC GCAG TTTGC GACA GGTAA GGTTT CCCC GGACG +50

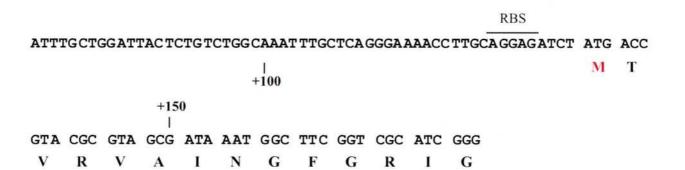


Figure 58 : Région régulant la transcription du gène epd chez E. coli

(d'après Ramseier et al., 1995; Charpentier et al., 1998; Bardey, 2000).

Le site d'initiation de la transcription du gène *epd*, à partir du promoteur appelé *epd* P0 est indiqué par une flèche brisée. La boîte -10 du promoteur est encadrée en noir. La séquence du site de fixation du complexe CRP-AMPc est indiquée en vert et la séquence du site de fixation de la protéine Cra en rouge. Les premiers acides aminés de la protéine E4PDH sont représentés (la méthionine est indiquée en rouge).

cellule (Della-Seta *et al.*, 1997). Par comparaison, 360 000 tétramères de GAPDH sont présents par cellule, dans les mêmes conditions de croissance (Boschi-Muller *et al.*, résultats non publiés).

Au laboratoire, V. Bardey, N. Robas et B. Charpentier ont étudié l'expression du tandem de gènes epd-pgk. Les études préliminaires de Nelleman et al. (1989) suggèraient que le gène pgk était exprimé à partir d'un promoteur situé en amont du gène epd, et d'un promoteur situé juste en amont de l'ORF pgk. Le promoteur epd P0 (appelé anciennement gapB P0) situé en amont de l'ORF epd, a été identifié (Fig. 57) (Charpentier et al., 1998). L'initiation de la transcription à partir du promoteur epd P0 est inhibée par la protéine Cra (Ramseier et al., 1995 ; pour revue, Saier et Ramseier, 1996). Un site de fixation au complexe CRP-AMPc a été identifié en amont du promoteur epd P0, localisé à 70,5 nucléotides en amont du site d'initiation de la transcription (Alefounder et Perham, 1989; Charpentier et al., 1998). Une expérience de gel retard montre que le complexe CRP-AMPc se fixe à cette région d'ADN in vitro (Ramseier et al., 1995). Cependant, in vivo, la délétion de ce site de fixation du complexe CRP-AMPc n'a pas d'effet sur l'initiation de la transcription à ce promoteur (Charpentier et al., 1998). En effet, l'activation par le complexe CRP-AMPc est inhibée par la protéine Cra. Ceci a été montré en utilisant des fusions transcriptionnelle epd-lacZ dans des souches sauvages ou inactivées dans la protéine Cra (Saier et Ramseier, 1996). En milieu minimum supplémenté en pyruvate, l'inactivation de la protéine Cra entraîne une augmentation de l'activité β-galactosidase d'un facteur 3. Par contre, l'activité βgalactosidase reste faible en milieu minimum supplémenté en glucose, avec ou sans la protéine Cra. Ainsi, le promoteur epd P0 est activé par le complexe CRP-AMPc, en présence de pyruvate, seulement si la protéine Cra n'exerce pas son effet répresseur. Le complexe CRP-AMPc aurait aussi un rôle indirect dans l'activation de l'initiation de la transcription au promoteur epd P0 (Charpentier et al., 1998).

Les travaux du laboratoire ont montré que le gène pgk est transcrit à partir des promoteurs epd P0 et pgk P1, identifié comme étant un promoteur étendu (Bardey, 2000). Ainsi, bien que le gène epd soit transcrit à partir d'un promoteur fort, la protéine E4PDH est présente à un taux très faible dans la cellule. Ceci est dû à un clivage de son ARNm, qui est renforcé lorsque l'ARNm pgk est présent en aval, mais aussi à une faible initiation de la traduction (Bardey, 1995; Bardey, 2000). L'ORF fbaA, séparée du gène pgk amont par un terminateur, possède un promoteur potentiel en aval de l'ORF pgk. Aucune donnée experimentale ne permettait de mettre en évidence comment cette ORF est transcrite.

#### IV.2 / Transcription du gène gapA

Le gène *gapA* codant la GAPDH active, est fortement exprimé chez *E. coli*, quelles que soient les conditions de croissance. En particulier, lorsque ce gène est cloné dans un plasmide multicopie, la GAPDH exprimée représente jusqu'à 80 % des protéines solubles (Branlant *et al.*, 1983). La fusion de gènes de la région promotrice du gène *gapA* à la séquence codant l'interleukine 1-β (IL1-β) montre

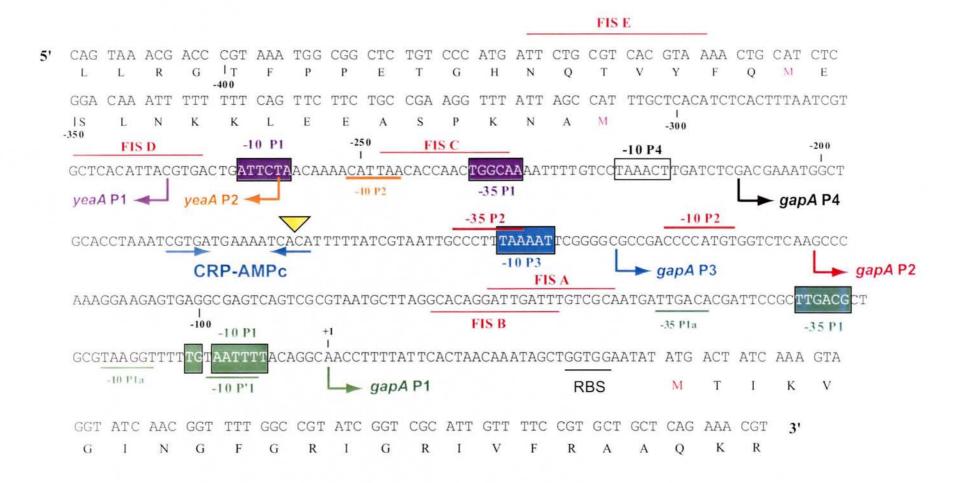


Figure 59: Région régulant la transcription des gènes gapA et yeaA chez E. coli (d'après Charpentier et Branlant, 1994; Charpentier, 1994; Thouvenot, 2005).

Les 4 sites d'initiation de la transcription du gène *gapA* sont indiqués par des flèches de couleurs différentes : P1 en vert, P2 en rouge, P3 en bleu et P4 en noir. Les boîtes -10 et -35 des promoteurs permettant l'initiation aux sites P1, P3 et P4 sont encadrées et les boîtes -10 et -35 dirigeant la transcription au site P2 sont soulignées. Les boîtes -10 et -35 du promoteur P1a et P'1 sont soulignées en vert. Les sites potentiels de fixation de la protéine FIS (A, B, C, D et E) sont indiqués par des traits rouges. Le centre d'une forte courbure de l'ADN est représenté par un triangle jaune, il est localisé au niveau du site de fixation du complexe CRP-AMPc activant l'initiation de la transcription au site P3. La séquence de l'ORF *yeaA* est indiquée, ainsi que les promoteurs *yeaA* P1 et P2 permettant la transcription de cette ORF. Les boîtes -10 P1, -35 P1, -10 P2 et les sites d'initiation de la transcription (représentés par des flèches) sont indiqués en violet et orange. RBS : site de fixation du ribosome.

aussi une forte expression de l'IL1-\( \text{G}\) (Charpentier, 1994). Ainsi, la forte expression de GAPDH résulte principalement de la région promotrice du gène gapA (Charpentier, 1994). Au laboratoire, B. Charpentier a identifié quatre promoteurs en amont du gène gapA, responsables de la production de la GAPDH à un taux quasiment constant dans un grand nombre de conditions de croissance chez E. coli (Fig. 58) (Charpentier, 1994; Charpentier et Branlant, 1994). Le promoteur P2 est transcrit par l'holoenzyme Eσ<sup>32</sup>, en réponse à un choc thermique, alors que les promoteurs P1, P3 et P4 sont transcrits par l'holoenzyme  $E\sigma^{70}$ . Le promoteur P1 est le promoteur majoritaire en phase exponentielle de croissance. Deux couples de boîtes -10 et -35 avaient été proposés, par examen de la séquence et identification du site d'initiation de la transcription (Fig. 58). Ces études ont montré que le couple de boîtes -10P1a et -35 Pla n'est pas actif, aussi bien in vivo, qu'in vitro (Charpentier, 1994). Un troisième couple de boîte, chevauchant la boîte -10P1 et -35 P1 a été proposé, ce couple de boîte correspondant à celles d'un promoteur étendu. B. Thouvenot et B. Charpentier ont montré, par expérience de mutagenèse dirigée, que le promoteur P1 est un promoteur étendu, qui a besoin d'une boîte -35 pour être actif (Thouvenot et al., 2004). Dans ce cas, la boîte -35 compense la boîte -10 sous-optimale. Le promoteur P2 permet de maintenir l'expression de la GAPDH en conditions de choc thermique. Le promoteur P3 est soumis à la répression catabolique (Charpentier et Branlant, 1994). Il est activé par le complexe CRP-AMPC, dont le site de fixation est centré -40.5 pb en amont du site d'initiation de la transcription, par un mécanisme de classe II. Le promoteur P4, très faiblement exprimé en phase exponentielle de croissance, a été très peu étudié.

B. Charpentier et B. Thouvenot ont aussi étudié la région promotrice du gène yeaA, situé sur le brin opposé à celui de gapA. Ce gène code une Méthionine sulfoxide réductase (MsrB) impliquée dans la réduction des méthionines oxydées produites lors d'un stress oxydatif. Deux promoteurs P1 et P2, situés en amont de gapA P4 ont été identifiés pour le gène yeaA. Ainsi, la région inter-ORF yeaA-gapA contient six promoteurs, ainsi que des éléments de régulation, tels que les sites de fixation pour la protéine Fis ou des éléments de courbure de l'ADN, susceptibles de réguler l'expression de l'un et/ou l'autre gène.

### Partie B / Objectifs de mes travaux

Une grande partie de mes travaux portent sur l'expression de l'opéron epd-pgk-fbaA chez E. coli. Afin de pouvoir réaliser une étude comparative de l'expression de cet opéron dans les différents génomes bactériens, j'ai tout d'abord recherché la présence de l'ORF epd et sa localisation au sein des génomes bactériens, ainsi que l'organisation des gènes autour des ORF epd et pgk. A mon arrivée au

laboratoire, V. Bardey, N. Robas et B. Charpentier avaient montré que les gènes epd et pgk d'E. coli sont exprimés à partir des promoteurs epd P0 et pgk P1. J'ai tout d'abord regardé si le gène fbaA, situé en aval du gène pgk chez E. coli, était exprimé uniquement à partir des mêmes promoteurs que le gène pgk. Ceci était assez inattendu, vu qu'un terminateur potentiel rho-indépendant pgk T1 est situé en aval du gène pgk (Alefounder et Perham, 1989) et qu'un promoteur potentiel est situé en amont du gène fbaA (Salgado et al., 2004). J'ai ensuite étudié la maturation des transcrits issus des promoteurs epd P0 et pgk P1. V. Bardey, N. Robas et B. Charpentier avaient mis en évidence le rôle de la RNase E dans la dégradation de l'ORF epd, expliquant en partie sa faible expression. J'ai recherché la (les) RNases impliquées dans la maturation et dégradation des transcrits pgk et fbaA. Enfin, j'ai regardé l'influence de la protéine Hfq, qui possède des sites de reconnaissance similaires à la RNase E, sur la maturation des transcrits issus de cet opéron.

Les gènes gapA, pgk et fbaA codent trois enzymes agissant en succession dans la glycolyse et leur expression est co-régulée (voir partie E.I de l'Introduction). Etant donné que le gène gapA est séparé physiquement sur le chromosome des gènes pgk et fbaA, il était intéressant d'étudier comment ces gènes pouvaient être co-régulés. Les études au laboratoire avaient montré l'existence d'un mécanisme d'activation de l'expression du gène gapA et de l'opéron epd-pgk pour un milieu supplémenté en glucose par rapport à un milieu supplémenté en pyruvate. Ce mécanisme d'activation semble être fonction de la vitesse de croissance obtenue avec une source carbonée, et dépend aussi du transporteur spécifique du glucose, EIIBCGlc (Charpentier et al., 1998). Or, l'inactivation de la protéine EIIBC Gle sur un milieu contenant du glucose conduit aussi à une diminution de la vitesse de croissance (Charpentier, résultats non montrés). Ainsi, l'activation de la transcription des gènes gapA, et epd-pgk par le glucose en présence du transporteur EIIBC<sup>Gle</sup> peut directement résulter de l'augmentation de la vitesse de croissance. J'ai alors mis utilisé la technique de fermentation continue pour étudier l'effet de la variation de la vitesse de croissance sans changement de la source carbonée sur la transcription des gènes gapA et de l'opéron epd-pgk-fbaA. Cette technique permet aussi de différencier le rôle du taux de croissance de celui de l'inactivation du transporteur sur l'expression du gène gapA et de l'opéron epd-pgk-fbaA.

Enfin, j'ai analysé les sous-produits du métabolisme lors des fermentations continues : j'ai alors remarqué l'absence de production d'acide acétique lors d'une limitation en phosphate pour la souche où le transporteur EIIBC<sup>Glc</sup> est inactivé. Ceci m'a conduit à étudier la production de protéines d'intérêt, dans cette souche. En effet, l'acide acétique inhibe la croissance bactérienne, et de façon plus importante la production de protéines d'intérêt. Cette étude a mis en évidence une forte surproduction de GAPDH à partir d'un plasmide multicopie, en milieu riche, oxygéné, en présence de glucose. Ce travail est décrit dans la partie C des Résultats.

# Résultats Discussion

glycolyse, alors que le second gène *gap* code une GAPDH intervenant dans la gluconéogenèse et le cycle de Calvin (Koksharova *et al.*, 1998). De même, chez *B. subtilis*, les gènes *gapA* et *gapB* codent des GAPDH plus particulièrement impliquées, l'une dans la glycolyse et l'autre dans la gluconéogenèse (Fillinger *et al.*, 2000). Par contre, l'étude de V. Bardey et N. Robas n'avait mis en évidence qu'un seul gène *pgk* chez toutes les espèces bactériennes étudiées (Bardey, 2000). De plus, l'analyse phylogénétique des espèces bactériennes faites par comparaison des séquences des protéines PGK des trois domaines du vivants donne des résultats similaires à celle basée sur la comparaison des séquences des ARNr 16S (Pollack *et al.*, 2005). L'ORF *pgk* semblait donc être le meilleur point d'entrée pour la recherche dans les génomes bactériens du triplet d'ORF *epd/gapA-pgk-fbaA*.

#### I.1 / Etude de l'organisation des ORF autour de l'ORF pgk

#### I.1.a / Les génomes étudiés et la démarche utilisée

J'ai utilisé l'ensemble des génomes bactériens partiellement ou entièrement séquencés disponibles à partir du site internet <u>www.ncbi.nlm.nih.gov</u> au début de l'année 2005. Ces espèces recouvrent la plupart des ordres et familles bactériennes répertoriées à ce jour, avec, en particulier, un nombre important d'espèces du phylum Proteobacteria, et des classes  $\alpha$ ,  $\beta$  et  $\gamma$  de ce phylum qui nous intéressait tout particulièrement, puisque jusque-là des gènes epd n'avaient été trouvés que dans les  $\gamma$ -protéobactéries.

Pour rechercher l'ORF pgk chez les différents génomes bactériens étudiés, nous avons utilisé le logiciel BLAST (www.ncbi.nlm.nih.gov) et l'ORF pgk d'E. coli comme base de recherche. Nous avons ensuite recherché les ORF orientées dans le même sens et localisées autour des ORF pgk identifiées. Cette approche a bien confirmé le fait qu'une seule ORF pgk est présente dans toutes les espèces bactériennes examinées. J'ai ensuite recherché si l'ORF située en amont de pgk pouvait être une ORF epd.

#### I.1.b / Résultats

Les résultats obtenus sont représentés dans le Tableau 3. Les bactéries sont classées suivant le dendrogramme obtenus par comparaison de l'ARNr 16S (Cole et al., 2005).

Dans quelques cas assez peu nombreux, aucun gène susceptible de coder une enzyme de la glycolyse est retrouvé au voisinage de l'ORF pgk. C'est le cas pour l'ensemble des bactéries étudiées appartenant aux phyla Aquificae, Cyanobacteria, Chlorobi, Chlamydiae, Planctomycetes et Bacteroidetes.

| Ordre   | Famille   | Espèces  | Nombre de<br>gènes<br>gap/epd | Organisation des ORF autour de l'ORF pgk  | Class  |
|---|---|--|-------------------------------|---|--|
|   |   | Phy  | lum Aquificae Classe          | Aquificae   |  |
| Aquificales   | Aquificaceae  | Aquifex aeolicus VF5   | 1                             | Prot 1 à 3 -PGK- Prot 4 à 5   |  |
|   |   | Phylum   | Thermotogae Classe            | Thermotogae   |  |
| Thermotogales   | Thermotogaceae  |  | 1                             | Prot 1-Prot 2-GAPDH -PGK- Prot 3 à 9  |  |
|   |   | Phylum De  | inococcus-Thermus C           | lasse Deinococci  | The state of the s |
| Deinococcales   | Dinococcaceae   | Deinococcus radiodurans R1   | 1                             | GAPDH -PGK- Prot 1 à 2-TIM  |  |
| Thermales   | Thermaceae  | Thermus thermophilus HB8   | 1                             | GAPDH -PGK- Prot 1 à 4  |  |
|   |   | Phyli  | um Chloroflexi Classe         | Chloroflexi   | •  |
| Chloroflexales  | Chloroflexaceae   | Chloroflexus aurantiacus   |                               | Prot 1 -PGK   |  |
| Dehalococcoidetes   | Dehalococcoides   | Dehalococcoides ethenogenes 195  | 1                             | Prot 1 -PGK PGM-TIM-Prot 2 à 4  |  |
|   |   | Phylum   | Cyanobacteria Classe          | Cvanobacteria   |  |
| Gloeobacterales   | Gloeobacter   | Gloeobacter violaceus PCC 7421   | 2                             | Prot 1 à 3PGK Prot 4  |  |
| Prochlorales  | Prochlorococcaceae  | Prochlorococcus marinus str. MIT 9313  | 2                             | PGK- Prot 1 à 3   |  |
|   |   | Prochlorococcus marinus subsp. pastoris str.   | 2                             | PGK   |  |
|   |   | CCMP1986   | 2                             | PGK   |  |
|   |   | Prochlorococcus marinus subsp. marinus str.  | 2                             | PGK   |  |
| Chroococcales   |   | CCMP1375<br>Crocosphaera watsonii WH 8501  | 2                             | PGK- Prot 1 à 3   |  |
| Chroococcales   | non classée   | Synechocystis sp. PCC 6803   | 2                             | PGK- Prot 1   |  |
| Chroococcales   | non classes   | Synechococcus sp. WH 8102  | 2                             | PGK   |  |
|   |   | Synechococcus elongatus PCC 6301   | 3                             | PGK- Prot 1   |  |
|   |   | Synechococcus elongatus PCC 7942   | 2                             | PGK   |  |
|   | N 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2   | Thermosynechococcus elongatus BP-1   | 3                             | PGK- Prot 1 à 2-ARNt-Prot 3 à 4   |  |
| Oscillatoriales   | Trichodesmium   | Trichodesmium erythraeum IMS101  | 1                             | PGK- Prot 1 à 6   |  |
| Vostocales  | Nostocaceae   | Anabaena variabilis ATCC 29413   |                               | PGK- Prot 1 à 2   |  |
|   | Nostocaceae   | Nostoc punctiforme PCC 73102   | 1                             | PGK- Prot 1   |  |
|   | T-  |  | lum Chlorobi Classe           |   |  |
| Chlorobiales  | Chlorobiaceae   | Chlorobium tepidum TLS   |                               | 7 -PGK- Prot 1 à 4  |  |
|   |   |  | Phylum Proteobact             | eria  |  |
|   |   |  | Olassa - (Alabamanta ab       |   |  |
|   |   |  | Classe a (Alphaproteob        |   |  |
| Rhodospirillales  | Acetobacteraceae  | Gluconobacter oxydans 621H   | 1                             | Prot 1-GAPDH -PGK   |  |
| Rhodospirillales<br>Rhodospirillales  | Acetobacteraceae<br>Rhodospirillaceae   | Gluconobacter oxydans 621H<br>Rhodospirillum rubrum  | 1                             | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1  |  |
| Rhodospirillales  | Rhodospirillaceae   | Gluconobacter oxydans 621H<br>Rhodospirillum rubrum<br>Magnetospirillum magnetotacticum MS-1   | 1 1                           | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7   |  |
|   |   | Gluconobacter oxydans 621H<br>Rhodospirillum rubrum  | 1 1                           | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1  |  |
| Rhodospirillales  | Rhodospirillaceae   | Gluconobacter oxydans 621H<br>Rhodospirillum rubrum<br>Magnetospirillum magnetotacticum MS-1   | 1 1                           | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7   |  |
| Rhodospirillales  | Rhodospirillaceae   | Gluconobacter oxydans 621H<br>Rhodospirillum rubrum<br>Magnetospirillum magnetotacticum MS-1<br>Ehrlichia ruminantium str. Welgevonden<br>Wolbachia endosymbiont of Drosophila melanogaster  | 1 1 1 1 1                     | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7 PGK 40 PGK- Prot 1  |  |
| Rhodospirillales  | Rhodospirillaceae   | Gluconobacter oxydans 621H Rhodospirillum rubrum Magnetospirillum magnetotacticum MS-1 Ehrlichia ruminantium str. Welgevonden Wolbachia endosymbiont of Drosophila melanogaster Wolbachia endosymbiont of Drosophila ananassae   | 1 1 1 1 1                     | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7 PGK 40 PGK- Prot 1 PGK- Prot 1 à 3  | ш  |
| Rhodospirillales  | Rhodospirillaceae   | Gluconobacter oxydans 621H<br>Rhodospirillum rubrum<br>Magnetospirillum magnetotacticum MS-1<br>Ehrlichia ruminantium str. Welgevonden<br>Wolbachia endosymbiont of Drosophila melanogaster  | 1 1 1 1 1                     | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7 PGK 40 PGK- Prot 1  | IIB  |
| Rhodospirillales  | Rhodospirillaceae   | Gluconobacter oxydans 621H Rhodospirillum rubrum Magnetospirillum magnetotacticum MS-1 Ehrlichia ruminantium str. Welgevonden Wolbachia endosymbiont of Drosophila melanogaster Wolbachia endosymbiont of Drosophila ananassae Oligotropha carboxidovorans   | 1 1 1 1 1                     | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7 PGK  40 PGK- Prot 1 PGK- Prot 1 à 3 FBP-Prot 1-TKT-GAPDH -PGK- ALD-Prot 3 à 9   | IIB<br>I   |
| Rhodospirillales  | Rhodospirillaceae<br>Anaplasmataceae  | Gluconobacter oxydans 621H Rhodospirillum rubrum Magnetospirillum magnetotacticum MS-1 Ehrlichia ruminantium str. Welgevonden Wolbachia endosymbiont of Drosophila melanogaster Wolbachia endosymbiont of Drosophila ananassae Oligotropha carboxidovorans Rhodopseudomonas palustris CGA009 Rhodobacter sphaeroides 2.4.1 Silicibacter pomeroyi DSS-3   | 1 1 1 1 1                     | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7 PGK  40 PGK- Prot 1 à 3 FBP-Prot 1-TKT-GAPDH -PGK- ALD-Prot 3 à 9 TKT-GAPDH -PGK- Prot 1 à 2-FBA-Prot 3 à 5 PGK FBA-Prot 1 à 5 PGK  | IIB<br>I   |
| Rhodospirillales<br>Rickettsiales<br>Rhodobacterales                                    | Rhodospirillaceae Anaplasmataceae Rhodobacteraceae Rhodobacteraceae   | Gluconobacter oxydans 621H Rhodospirillum rubrum Magnetospirillum magnetotacticum MS-1 Ehrlichia ruminantium str. Welgevonden Wolbachia endosymbiont of Drosophila melanogaster Wolbachia endosymbiont of Drosophila ananassae Oligotropha carboxidovorans Rhodopseudomonas palustris CGA009 Rhodobacter sphaeroides 2.4.1 Silicibacter pomeroyi DSS-3 Silicibacter sp. TM1040   | 1<br>1<br>1<br>1<br>1         | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7 PGK  40 PGK- Prot 1 PGK- Prot 1 à 3 FBP-Prot 1-TKT-GAPDH -PGK- ALD-Prot 3 à 9 TKT-GAPDH -PGK- Prot 1 à 2-FBA-Prot 3 à 5 PGK PGK PGK- Prot 1 à 6   | IIB<br>I   |
| Rhodospirillales  | Rhodospirillaceae  Anaplasmataceae  Rhodobacteraceae Rhodobacteraceae  Sphingomonadaceae  | Gluconobacter oxydans 621H Rhodospirillum rubrum Magnetospirillum magnetotacticum MS-1 Ehrlichia ruminantium str. Welgevonden Wolbachia endosymbiont of Drosophila melanogaster Wolbachia endosymbiont of Drosophila ananassae Oligotropha carboxidovorans Rhodopseudomonas palustris CGA009 Rhodobacter sphaeroides 2.4.1 Silicibacter pomeroyi DSS-3 Silicibacter sp. TM1040 Erythrobacter litoralis HTCC2594  | 1<br>1<br>1<br>1<br>1         | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7 PGK  40 PGK- Prot 1 à 3 FBP-Prot 1-TKT-GAPDH -PGK- ALD-Prot 3 à 9 TKT-GAPDH -PGK- Prot 1 à 2-FBA-Prot 3 à 5 PGK FBA-Prot 1 à 5 PGK PGK- Prot 1 à 6 TKT-GAPDH -PGK- Prot 1 à 2   | IIB<br>I   |
| Rhodospirillales<br>Rickettsiales<br>Rhodobacterales                                    | Rhodospirillaceae  Anaplasmataceae  Rhodobacteraceae Rhodobacteraceae  Sphingomonadaceae Sphingomonadaceae  | Gluconobacter oxydans 621H Rhodospirillum rubrum Magnetospirillum magnetotacticum MS-1 Ehrlichia ruminantium str. Welgevonden Wolbachia endosymbiont of Drosophila melanogaster Wolbachia endosymbiont of Drosophila ananassae Oligotropha carboxidovorans Rhodopseudomonas palustris CGA009 Rhodobacter sphaeroides 2.4.1 Silicibacter pomeroyi DSS-3 Silicibacter sp. TM1040 Erythrobacter litoralis HTCC2594 Novosphingobium aromaticivorans DSM 12444  | 1<br>1<br>1<br>1<br>1         | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7 PGK  40 PGK- Prot 1 à 3 FBP-Prot 1-TKT-GAPDH -PGK- ALD-Prot 3 à 9 TKT-GAPDH -PGK- Prot 1 à 2-FBA-Prot 3 à 5 PGK PGK PGK PGK TKT-GAPDH -PGK- Prot 1 à 6 TKT-GAPDH -PGK- Prot 1 à 2 TKT-GAPDH-Prot 1 -PGK- FBA  | IIB<br>I   |
| Rhodospirillales Rickettsiales Rhodobacterales Sphingomonadales                         | Rhodospirillaceae  Anaplasmataceae  Rhodobacteraceae Rhodobacteraceae Sphingomonadaceae Sphingomonadaceae Sphingomonadaceae                               | Gluconobacter oxydans 621H Rhodospirillum rubrum Magnetospirillum magnetotacticum MS-1 Ehrlichia ruminantium str. Welgevonden Wolbachia endosymbiont of Drosophila melanogaster Wolbachia endosymbiont of Drosophila ananassae Oligotropha carboxidovorans Rhodopseudomonas palustris CGA009 Rhodobacter sphaeroides 2.4.1 Silicibacter pomeroyi DSS-3 Silicibacter sp. TM1040 Erythrobacter litoralis HTCC2594 Novosphingobium aromaticivorans DSM 12444 Zymomonas mobilis subsp. mobilis ZM4   | 1<br>1<br>1<br>1<br>1         | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7 PGK  40 PGK- Prot 1 à 3 FBP-Prot 1-TKT-GAPDH -PGK- ALD-Prot 3 à 9 TKT-GAPDH -PGK- Prot 1 à 2-FBA-Prot 3 à 5 PGK FBA-Prot 1 à 5 PGK PGK- Prot 1 à 6 TKT-GAPDH -PGK- Prot 1 à 2 TKT-GAPDH -PGK- FBA TKT-GAPDH -PGK- FBA   | IIB  |
| Rhodospirillales Rickettsiales Rhodobacterales Sphingomonadales Rhizobiales             | Rhodospirillaceae  Anaplasmataceae  Rhodobacteraceae Rhodobacteraceae Sphingomonadaceae Sphingomonadaceae Sphingomonadaceae Rhizobiaceae                  | Gluconobacter oxydans 621H Rhodospirillum rubrum Magnetospirillum magnetotacticum MS-1 Ehrlichia ruminantium str. Welgevonden Wolbachia endosymbiont of Drosophila melanogaster Wolbachia endosymbiont of Drosophila ananassae Oligotropha carboxidovorans Rhodopseudomonas palustris CGA009 Rhodobacter sphaeroides 2.4.1 Silicibacter pomeroyi DSS-3 Silicibacter sp. TM1040 Erythrobacter litoralis HTCC2594 Novosphingobium aromaticivorans DSM 12444 Zymomonas mobilis subsp. mobilis ZM4 Sinorhizobium meliloti 1021   | 1<br>1<br>1<br>1<br>1         | Prot 1-GAPDH -PGK GAPDH -PGK - Prot 1 GAPDH -PGK - Prot 1 à 7 PGK  40 PGK - Prot 1 à 3 FBP-Prot 1-TKT-GAPDH -PGK - ALD-Prot 3 à 9 TKT-GAPDH -PGK - Prot 1 à 2-FBA-Prot 3 à 5 PGK FBA-Prot 1 à 5 PGK PGK - Prot 1 à 6 TKT-GAPDH -PGK - Prot 1 à 2 TKT-GAPDH-Prot 1 - PGK - FBA TKT-GAPDH -PGK - FBA TKT-GAPDH -PGK - FBA TKT-GAPDH-Prot 1 - PGK - FBA TKT-GAPDH-Prot 1 - PGK - FBA   | IIB<br>I<br>I  |
| Rhodospirillales Rickettsiales Rhodobacterales Sphingomonadales                         | Rhodospirillaceae  Anaplasmataceae  Rhodobacteraceae Rhodobacteraceae Sphingomonadaceae Sphingomonadaceae Sphingomonadaceae                               | Gluconobacter oxydans 621H Rhodospirillum rubrum Magnetospirillum magnetotacticum MS-1 Ehrlichia ruminantium str. Welgevonden Wolbachia endosymbiont of Drosophila melanogaster Wolbachia endosymbiont of Drosophila ananassae Oligotropha carboxidovorans Rhodopseudomonas palustris CGA009 Rhodobacter sphaeroides 2.4.1 Silicibacter pomeroyi DSS-3 Silicibacter sp. TM1040 Erythrobacter litoralis HTCC2594 Novosphingobium aromaticivorans DSM 12444 Zymomonas mobilis subsp. mobilis ZM4   | 1<br>1<br>1<br>1<br>1         | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7 PGK  40 PGK- Prot 1 à 3 FBP-Prot 1-TKT-GAPDH -PGK- ALD-Prot 3 à 9 TKT-GAPDH -PGK- Prot 1 à 2-FBA-Prot 3 à 5 PGK FBA-Prot 1 à 5 PGK PGK- Prot 1 à 6 TKT-GAPDH -PGK- Prot 1 à 2 TKT-GAPDH -PGK- FBA TKT-GAPDH -PGK- FBA   | IIB<br>I<br>I  |
| Rhodospirillales Rickettsiales Rhodobacterales Sphingomonadales Rhizobiales             | Rhodospirillaceae  Anaplasmataceae  Rhodobacteraceae Rhodobacteraceae Sphingomonadaceae Sphingomonadaceae Sphingomonadaceae Rhizobiaceae                  | Gluconobacter oxydans 621H Rhodospirillum rubrum Magnetospirillum magnetotacticum MS-1 Ehrlichia ruminantium str. Welgevonden Wolbachia endosymbiont of Drosophila melanogaster Wolbachia endosymbiont of Drosophila ananassae Oligotropha carboxidovorans Rhodopseudomonas palustris CGA009 Rhodobacter sphaeroides 2.4.1 Silicibacter pomeroyi DSS-3 Silicibacter pomeroyi DSS-3 Silicibacter sphaeroides DSS-3 S | 1<br>1<br>1<br>1<br>1         | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7 PGK  40 PGK- Prot 1 à 3 FBP-Prot 1-TKT-GAPDH -PGK- ALD-Prot 3 à 9 TKT-GAPDH -PGK- Prot 1 à 2-FBA-Prot 3 à 5 PGK PGK- Prot 1 à 5 PGK PGK- Prot 1 à 6 TKT-GAPDH -PGK- Prot 1 à 2 TKT-GAPDH-Prot 1 -PGK- FBA TKT-GAPDH -PGK- FBA   | IIB<br>I<br>I<br>I   |
| Rhodospirillales Rickettsiales Rhodobacterales Sphingomonadales Rhizobiales             | Rhodospirillaceae  Anaplasmataceae  Rhodobacteraceae Rhodobacteraceae  Sphingomonadaceae Sphingomonadaceae Sphingomonadaceae Rhizobiaceae Bartonellaceae  | Gluconobacter oxydans 621H Rhodospirillum rubrum Magnetospirillum magnetotacticum MS-1 Ehrlichia ruminantium str. Welgevonden Wolbachia endosymbiont of Drosophila melanogaster Wolbachia endosymbiont of Drosophila ananassae Oligotropha carboxidovorans Rhodopseudomonas palustris CGA009 Rhodobacter sphaeroides 2.4.1 Silicibacter pomeroyi DSS-3 Silicibacter sp. TM1040 Erythrobacter litoralis HTCC2594 Novosphingobium aromaticivorans DSM 12444 Zymomonas mobilis subsp. mobilis ZM4 Sinorhizobium meliloti 1021 Bartonella henselae str. Houston-1 Bartonella quintana str. Toulouse  | 1<br>1<br>1<br>1<br>1         | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7 PGK  40 PGK- Prot 1 à 3 FBP-Prot 1-TKT-GAPDH -PGK- ALD-Prot 3 à 9 TKT-GAPDH -PGK- Prot 1 à 2-FBA-Prot 3 à 5 PGK PGK PGK PGK- Prot 1 à 5 PGK PGK- Prot 1 à 6 TKT-GAPDH -PGK- PFOT 1 à 2 TKT-GAPDH-Prot 1 -PGK- FBA TKT-GAPDH -PGK- FBA TKT-GAPDH -PGK- FBA GAPDH -PGK- FBA GAPDH -PGK- FBA GAPDH -PGK- FBA   | IIB  |
| Rhodospirillales Rickettsiales Rhodobacterales Sphingomonadales Rhizobiales Rhizobiales | Rhodospirillaceae  Anaplasmataceae  Rhodobacteraceae Rhodobacteraceae  Sphingomonadaceae Sphingomonadaceae Sphingomonadaceae Bartonellaceae  Brucellaceae | Gluconobacter oxydans 621H Rhodospirillum rubrum Magnetospirillum magnetotacticum MS-1 Ehrlichia ruminantium str. Welgevonden Wolbachia endosymbiont of Drosophila melanogaster Wolbachia endosymbiont of Drosophila ananassae Oligotropha carboxidovorans Rhodopseudomonas palustris CGA009 Rhodobacter sphaeroides 2.4.1 Silicibacter pomeroyi DSS-3 Silicibacter pomeroyi DSS-3 Silicibacter pomeroyi DSS-3 Silicibacter sp. TM1040 Erythrobacter litoralis HTCC2594 Novosphingobium aromaticivorans DSM 12444 Zymomonas mobilis subsp. mobilis ZM4 Sinorhizobium meliloti 1021 Bartonella henselae str. Houston-1 Bartonella quintana str. Toulouse Brucella abortus biovar 1 str. 9-941 Brucella melitensis 16M Brucella suis 1330  | 1<br>1<br>1<br>1<br>1         | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7 PGK  40 PGK- Prot 1 à 3 FBP-Prot 1-TKT-GAPDH -PGK- ALD-Prot 3 à 9 TKT-GAPDH -PGK- Prot 1 à 2-FBA-Prot 3 à 5 PGK PGK- Prot 1 à 5 PGK PGK- Prot 1 à 6 TKT-GAPDH -PGK- Prot 1 à 2 TKT-GAPDH-Prot 1 -PGK- FBA TKT-GAPDH-Prot 1 -PGK- FBA TKT-GAPDH-Prot 1 -PGK- FBA TKT-GAPDH -PGK- FBA GAPDH -PGK- FBA TKT-GAPDH -PGK- FBA | B <br>   |
| Rhodospirillales Rickettsiales Rhodobacterales Sphingomonadales Rhizobiales             | Rhodospirillaceae  Anaplasmataceae  Rhodobacteraceae Rhodobacteraceae  Sphingomonadaceae Sphingomonadaceae Sphingomonadaceae Rhizobiaceae Bartonellaceae  | Gluconobacter oxydans 621H Rhodospirillum rubrum Magnetospirillum magnetotacticum MS-1 Ehrlichia ruminantium str. Welgevonden Wolbachia endosymbiont of Drosophila melanogaster Wolbachia endosymbiont of Drosophila ananassae Oligotropha carboxidovorans Rhodopseudomonas palustris CGA009 Rhodobacter sphaeroides 2.4.1 Silicibacter pomeroyi DSS-3 Silicibacter sp. TM1040 Erythrobacter litoralis HTCC2594 Novosphingobium aromaticivorans DSM 12444 Zymomonas mobilis subsp. mobilis ZM4 Sinorhizobium meliloti 1021 Bartonella henselae str. Houston-1 Bartonella quintana str. Toulouse Brucella abortus biovar 1 str. 9-941 Brucella melitensis 16M   | 1<br>1<br>1<br>1<br>1         | Prot 1-GAPDH -PGK GAPDH -PGK- Prot 1 GAPDH -PGK- Prot 1 à 7 PGK  40 PGK- Prot 1 à 3 FBP-Prot 1-TKT-GAPDH -PGK- Prot 1 à 2 FBA-Prot 3 à 9 TKT-GAPDH -PGK- Prot 1 à 2-FBA-Prot 3 à 5 PGK PGK- Prot 1 à 2-FBA-Prot 3 à 5 PGK- Prot 1 à 2 TKT-GAPDH-PGK- PGK- PTOT 1 à 2 TKT-GAPDH-PGK- PGK- FBA TKT-GAPDH-PGK- FBA TKT-GAPDH-PGK- FBA GAPDH-PGK-FBA TKT-GAPDH-PGK-FBA TKT-GAPDH-PGK-FBA TKT-GAPDH-PGK-FBA TKT-GAPDH-PGK-FBA TKT-GAPDH-PGK-FBA TKT-GAPDH-PGK-FBA TKT-GAPDH-PGK-FBA                              | IIB<br>I<br>I  |

|                                 |  |   | Classe B (beta | proteobacteria)                                       |  |
|---------------------------------|--|---|----------------|---|--|
| Burkholderiales                 | Burkholderiaceae   | Burkholderia cepacia R1808                            |                | Prot 1 à 2 -PGK- PYK-ALD-Prot 3 à 4                   | II-B                                     |
|                                 |  | Burkholderia cepacia R18194                           |                | Prot 1 à 2 -PGK- PYK-ALD-Prot 3 à 6                   | II-B                                     |
|                                 |  | Burkholderia fungorum LB400                           | 1              | Prot 1 à 2 -PGK- PYK-ALD-Prot 3 à 6                   | II-B                                     |
|                                 |  | Burkholderia mallei ATCC2344                          | 1              | Prot 1 à 2 -PGK- PYK-ALD-Prot 3 à 6                   | 0000                                     |
|                                 |  | Burkholderia pseudomallei K96243                      | 1              | Prot 1 à 2 -PGK- PYK-ALD-Prot 3 à 6                   | II-B                                     |
|                                 | Burkholderiaceae   | Cupriavidus necator                                   |                | Prot 1 à 9-GAPDH -PGK- ALD-Prot 10-Prot 11            | II-B                                     |
| lurkholderiales                 | Burkholderiaceae   | Ralstonia eutropha JMP134                             |                | PGK- PYK-ALD-Prot 1-Prot2                             | II-B                                     |
|                                 |  | Ralstonia metallidurans CH34                          |                | PGK- PYK-ALD-Prot 1 à 4                               | II-B                                     |
|                                 |  | Ralstonia solanacearum GMI1000 chromosome             | 1              | PGK- PYK  |  |
| lurkholderiales                 | Rubrivivax   | Rubrivivax gelatinosus PM1                            | 2              | Prot 1à 6 -PGK- PYK-ALD-Prot 7-Prot 8                 | II-B                                     |
| urkholderiales                  | Comamonadaceae   | Polaromonas sp. JS666                                 | 2              | Prot 1 à 4 -PGK - PYK-Prot 5 à 7-ALD-Prot 8 à 12      | II-B                                     |
| urkholderiales                  | Alcaligenaceae   | Bordetella bronchiseptica RB50                        | -              | TKT-GAPDH -PGK  | 11-6                                     |
| untionenalea                    | Alcangenaceae  | Bordetella parapertussis 12822                        | 4              | TKT-GAPDH -PGK  |  |
|                                 |  | Bordetella pertussis Tohama I                         | - 4            | TKT-GAPDH -PGK-                                       |  |
| Hydrogenophilales               | Hydrogenophilaceae   | Thiobacillus denitrificans ATCC 25259                 | 8              | Prot 1 à 3-TKT-GAPDH -PGK- PYK-ALD                    |  |
| Methylophilales                 | Methylophilaceae   | Methylobacillus flagellatus KT                        |                | Prot 1-TKT-GAPDH -PGK- PYK-ALD-Prot 2 à 4             | II-B                                     |
| leisseriales                    | Neisseriaceae  | Chromobacterium violaceum ATCC 12472                  | 2              | TKT-GAPDH -PGK- Prot 1-ALD-Prot 2 à 4                 | II-B                                     |
| leisseriales                    | iveisseriaceae   | Neisseria gonorrhoeae FA 1090                         | 2              | Prot 1 à 4 - PGK                                      | 11-6                                     |
|                                 |  | Neisseria meningitidis MC58                           | 2              | Prot 1 à 5 - PGK                                      |  |
|                                 |  |   | -              |   |  |
|                                 |  | Neisseria meningitidis Z2491                          | _              | Prot 1 à 16 -PGK                                      |  |
| accessor conservation           | Di i   | Nitrosomonas europaea ATCC 19718                      | 2              | TKT-GAPDH -PGK- PYK-ALD                               | II-B                                     |
| thodocyclales                   | Rhodocyclaceae   | Azoarcus sp. EbN1                                     | 1              | Prot 1à 3-TKT-GAPDH -PGK- PYK-ALD-Prot 4 à Prot 19    | II-B                                     |
|                                 | Rhodocyclaceae   | B - 1/2 - 1/2 - DOD                                   |                | FBP-Prot 1-TKT-E4PDH 7-ALD                            | II-B                                     |
|                                 |  | Dechloromonas aromatica RCB                           |                | Prot 1 à 2-FBP-Prot 3-TKT-GAPDH-E4PDH ? -PGK- PYK-ALD | II-B                                     |
|                                 |  | C   | lasse y (gamn  | paproteobacteria)                                     |  |
| hiotrichales                    | Francisellaceae  | Francisella tularensis subsp. tularensis Schu 4       | 1              | TKT-GAPDH -PGK- PYK-FBA                               | II-B                                     |
| Kanthomonadales                 | Xanthomonadaceae   | Xanthomonas axonopodis pv. citri str. 306             | 2              | GAPDH-Prot 1 à 4 ·PGK- Prot 1-PYK-FBA                 | 1  |
|                                 |  | Xanthomonas campestris pv. campestris str. ATCC 33913 | 2              | GAPDH-Prot 1 à 3 PGK- Prot 1-PYK-FBA                  | i  |
|                                 |  | Xanthomonas oryzae pv. oryzae KACC10331               | 2              | Prot 1 à 4 -PGK- Prot 5- PYK-FBA                      | 1  |
|                                 | Xanthomonadaceae   | Xylella fastidiosa 9a5c                               | 1              | PGK- PYK-Prot1-FBA                                    | 1  |
|                                 | Colored Color Superior Color State Colored Col | Xylella fastidiosa Ann-1                              | 1              | PGK- PYK-FBA  | l î                                      |
|                                 |  | Xvlella fastidiosa Dixon                              | 1              | PGK- PYK-FBA  | l i                                      |
|                                 |  | Xylella fastidiosa Temecula1                          | 1              | PGK- PYK-FBA  | i i                                      |
| egionellales                    | Legionellaceae   | Legionella pneumophila str. Lens                      | 1              | TKT-GAPDH -PGK- PYK-SdhB                              |  |
| egionenales                     | Logioniando  | Legionella pneumophila str. Paris                     | 1              | TKT-GAPDH -PGK- PYK-SdhB                              |  |
|                                 |  | Legionella pneumophila str. Philadelphia 1            | 4              | TKT-GAPDH -PGK- PYK-ALD Prot 2 à 4                    | II-B                                     |
| egionellales                    | Coxiellaceae   | Coxiella burnetii RSA 493                             | 4              | TKT-GAPDH -PGK- PYK-SdhB                              | 11-0                                     |
| egiorienales<br>fethylococcales | Methylococcaceae   | Methylococcus capsulatus str. Bath                    | ė              | Prot 1 à 9 -PGK- Prot 10 à 15                         |  |
| seudomonadales                  | Pseudomonadaceae   | Azotobacter vinelandii Avin02 21                      | •              | TKT-GAPDH -PGK- ALD                                   | II-B                                     |
| Seudomonadales                  | rseudomonadaceae   | Azotobacter vinelandii Avin67                         |                | TKT-GAPDH -PGK- ALD                                   | II-B                                     |
| seudomonadales                  | Pseudomonadaceae   | Pseudomonas aeruginosa UCBPP-PA14                     | 4              | GAPDH -PGK- Prot1 à 2-ALD-Prot 3                      | II-B                                     |
| seudomonadales                  | r seddomonadaceae  |   | •              |   | Pre5000000000000000000000000000000000000 |
|                                 |  | Pseudomonas fluorescens PfO-1                         |                | TKT-GAPDH -PGK- Prot 1 à 2-ALD-ALD                    | II-B / II-B                              |
|                                 |  | Pseudomonas putida KT2440                             | 5              | TKT-GAPDH -PGK- Prot 1 à 2-ALD-Prot 3 à 5             | II-B                                     |
|                                 |  | Pseudomonas syringae pv. syringae B728a               |                | TKT-GAPDH -PGK- Prot 1 à 3-ALD                        | II-B                                     |
|                                 |  | Pseudomonas syringae pv. tomato str. DC3000           | 2              | TKT-GAPDH -PGK- Prot 1-Prot 2-ALD                     | II-B                                     |
| seudomonadales                  | Moraxellaceae  | Acinetobacter sp. ADP1 complete genome                | 2              | PGK- Prot 1-ALD-Prot 2                                | II-B                                     |
|                                 | Moraxellaceae  | Psychrobacter sp. 273-4                               |                | Prot 1 à 2 -PGK- Prot 3 ALD-Prot 4 à 5                | II-B                                     |

| Alteromonadales    | Alteromonadaceae                     | Microbulbifer degradans 2-40   | ?                    | TKT  | -PGK  |                                 |         |
|--------------------|--------------------------------------|--|----------------------|--|-------|---------------------------------|---------|
|                    | Idiomarinaceae<br>(Alteromonadaceae) | Idiomarina loihiensis L2TR   | 3                    | TKT-E4PDH?   | -PGK- | ALD-Prot 1 à 6                  | II-B    |
|                    | Shewanellaceae<br>(Alteromonadaceae) | Shewanella oneidensis MR-1   | 4                    | TKT-E4PDH  | -PGK- | ALD-Prot 1 à 4                  | II-B    |
| Vibrionales        | Vibrionaceae                         | Photobacterium profundum SS9   | 4                    | TKT-E4PDH  | PCK   | EBA Prot 1                      | II-A    |
| Violidiaica        | Vibrionaceae                         | Vibrio cholerae O1 biovar eltor str. N16961                              | 3                    | IrgA-E4PDH   |       |                                 | II-A    |
|                    | Vibrioriaceae                        | Vibrio fischeri ES114  | 4                    | TKT-E4PDH  |       |                                 | II-A    |
|                    |                                      | Vibrio parahaemolyticus RIMD 2210633                                     | 3                    | IrgA-E4PDH   |       |                                 | II-A    |
|                    |                                      | Vibrio vulnificus CMCP6  | 3                    | TKT-E4PDH  |       |                                 | II-A    |
|                    |                                      | Vibrio vulnificus YJ016  | 3                    | Prot 1-E4PDH   |       |                                 | II-A    |
| Enterobacteriales  | Enterobacteriaceae                   | Buchnera aphidicola str. APS (Acyrthosiphon pisum)                       |                      |  |       | FBA-YggB                        | II-A    |
|                    | Enterobacteriaceae                   | Buchnera aphidicola str. Bp (Baizongia pistaciae)                        |                      | Prot 1 à 2   | -PGK- | FBA-YggB                        | II-A    |
|                    | Enterobacteriaceae                   | Buchnera aphidicola str. Sg (Schizaphis graminum)                        | 2                    | Prot 1   | -PGK- | FBA                             | II-A    |
|                    | Enterobacteriaceae                   | Candidatus Blochmannia floridanus  |                      |  | PGK-  | FBA                             | II-A    |
|                    | Enterobacteriaceae                   | Edwardsiella ictaluri  |                      |  |       | FBA-YggB-YggE                   | II-A    |
|                    | Enterobacteriaceae                   | Erwinia carotovora subsp. atroseptica SCRI1043                           | 2                    |  |       | FBA-YggB-YggA-YggE              | II-A    |
|                    | Enterobacteriaceae                   | Escherichia coli CFT073  | 3                    |  |       | FBA-MscS(YggB)-YggA-YggE-Ygfl   | II-A    |
|                    |                                      | Escherichia coli K12   | 2                    | TKT-Prot 1 à 6-E4PDH   | -PGK- | FBA-MscS(YggB)-YggA-YggE-Ygfl   | II-A    |
|                    |                                      | Escherichia coli O157:H7   | 3                    |  |       | FBA-MscS(YggB)-YggA-YggE        | II-A    |
|                    | 2000 20 0000                         | Escherichia coli O157:H7 EDL933  | 3                    |  |       | FBA-MscS(YggB)-YggA-YggE-Prot 1 | II-A    |
|                    | Enterobacteriaceae                   | Photorhabdus luminescens subsp. laumondii TTO1                           | 2                    | Prot 1 à -E4PDH  | -PGK- | FBA                             | II-A    |
|                    | Enterobacteriaceae                   | Salmonella enterica subsp. enterica serovar<br>Choleraesuis str. SC-B67  | 2                    | E4PDH  | -PGK- | FBA-YggB-YggA-YggE              | II-A    |
|                    |                                      | Salmonella enterica subsp. enterica serovar Paratypi<br>A str. ATCC 9150 | 2                    | E4PDH  | -PGK- | FBA-YggB-YggA-YggE              | II-A    |
|                    |                                      | Salmonella enterica subsp. enterica serovar Typhi Ty2                    | 2                    | E4PDH  | -PGK- | FBA-Prot 1 à 4                  | II-A    |
|                    |                                      | Salmonella enterica subsp. enterica serovar Typhi str.<br>CT18           | 2                    | E4PDH  | -PGK- | FBA-YggB-YggA-YggE              | II-A    |
|                    |                                      | Salmonella typhimurium LT2   |                      | E4PDH  | -PGK- | FBA-YggB-YggA-YggE              | II-A    |
|                    | Enterobacteriaceae                   | Shigella flexneri 2a str. 301  | 3                    |  |       | FBA-YggB-YggA-YggE-Ygfl         | II-A    |
|                    |                                      | Shigella flexneri 2a str. 2457   | 2                    |  |       | FBA-YggB-YggA-YggE-Ygfl         | II-A    |
|                    | Enterobacteriaceae                   | Yersinia pestis CO92   |                      | TKT-Prot 1 à 3-E4PDH   | -PGK- | FBA-YggB-YggA-YggE              | II-A    |
|                    |                                      | Yersinia pestis KIM  |                      | TKT-E4PDH  | -PGK- | FBA-Prot 1 à 3                  | II-A    |
|                    |                                      | Yersinia pestis biovar Medievalis str. 91001                             | 2                    | TKT-Prot 1-E4PDH   |       |                                 | II-A    |
|                    |                                      | Yersinia pseudotuberculosis IP 32953                                     | 2                    | TKT-E4PDH  | -PGK- | FBA-MscC-YggA-YggE              | II-A    |
|                    | Enterobacteriaceae                   | Wigglesworthia glossinidia endosymbiont of Glossina<br>brevipalpis       | 2                    |  | PGK-  | FBA-Prot 1 à 8                  | II-A    |
| Pasteurellales     | Pasteurellaceae                      | Actinobacillus pleuropneumoniae serovar 1 str. 4074                      |                      | Prot 1 à 2   | -PGK- | FBA-Prot 3 à 6                  | II-A    |
|                    |                                      | Mannheimia succiniciproducens MBEL55E                                    | 1                    | Prot 1   | -PGK- | FBA-Prot 2                      | II-A    |
|                    |                                      | Pasteurella multocida subsp. multocida str. Pm70                         | 2                    | Prot 1   | -PGK- | FBA-Prot 2 à 3                  | II-A    |
|                    |                                      | Haemophilus ducreyi 35000HP  | 1                    | Prot 1   | -PGK- | FBA-Prot 2                      | II-A    |
|                    |                                      | Haemophilus influenzae 86-028NP  |                      | Prot 1   | -PGK- | ?                               | 522,057 |
|                    |                                      | Haemophilus influenzae R2846   | 1                    | Prot 1   | -PGK- | FBA-Prot 2                      | II-A    |
|                    |                                      | Haemophilus influenzae Rd KW20   | 1                    | Prot 1   | -PGK- | FBA-Prot 2                      | II-A    |
|                    |                                      | Haemophilus somnus 2336  | 1                    |  | PGK-  | FBA                             | II-A    |
|                    |                                      | Haemophilus somnus 129PT   |                      |  | PGK-  | FBA                             | II-A    |
|                    |                                      |  | Classe δ (deltaprote | Control of the Contro |       |                                 |         |
| Bdellovibrionales  | Bdellovibrionaceae                   | Bdellovibrio bacteriovorus HD100   | 1                    | Prot 1 à 4-GAPDH   |       | TIM                             |         |
| Desulfovibrionales | Desulfovibrionaceae                  | Desulfovibrio desulfuricans G20  |                      | Prot 1 à 8-TKT   | -PGK- |                                 |         |
|                    |                                      | Desulfovibrio vulgaris subsp. vulgaris str.<br>Hildenborough             | 2                    | ткт  | -PGK- |                                 |         |
| Desulfobacterales  | desulfobacteriaceae                  | Desulfotalea psychrophila LSv54  | 2                    | GAPDH-Prot 1   | -PGK- | TIM-Prot 2-ARNt-ARNt-Prot 3 à 6 |         |
| Desulfuromonales   | Geobacteraceae                       | Geobacter sulfurreducens PCA   | 1                    |  |       | Prot1-ARNt                      |         |
|                    |                                      | Geobacter metallireducens GS-15  | 1                    | GAPDH  |       |                                 |         |
|                    |                                      |  | lasse ε (epsilonpro  |  |       |                                 |         |
| Campylobacterales  | Campylobacteraceae                   | Campylobacter jejuni subsp. jejuni NCTC 11168                            | 1                    | GAPDH  |       | TIM-prot 1 à 2                  |         |
| Campylobacterales  | helicobacteraceae                    | Helicobacter hepaticus ATCC 51449  | 2                    |  |       | Prot 1 à 4                      |         |
|                    |                                      | Helicobacter pylori J99  | 2                    | Prot 1 à 18-GAPDH  |       |                                 |         |
|                    | CONTRACTOR CONTRACTOR                | Helicobacter pylori 26695  | 2                    |  |       | prot 1 à 3                      |         |
|                    | helicobacteraceae                    | Wolinella succinogenes DS1740  | 2                    | 7-GAPDH  | -PGK- | TIM-7                           |         |

|                        |  |   | Phylum Firmi       | cutes                             |         |  |      |
|------------------------|--|---|--------------------|-----------------------------------|---------|--|------|
|                        |  |   | Classe Clostr      | ridia                             |         |  |      |
| lostridiales           | Clostridiaceae   | Clostridium acetobutylicum ATCC824  | 2                  | Prot 1 à 17-TR-GAPDH              | -PGK-   | TIM-PGM-ENO-Prot 18 à 22   |      |
| roomana                | C.Colindiabedo   | Clostridium perfringens str. 13   | 2                  |                                   |         | TIM-PGM-Prot 12-ENO-Prot 13 à 18   |      |
|                        |  | Clostridium tetani E88  | 1                  |                                   |         | TIM-PGM-ENO-Prot 1 à 3   |      |
|                        |  | Clostridium thermocellum ATCC 27405   | 1                  |                                   |         | TIM-PGM-Prot 28  |      |
| lostridiales           | Peptococcaceae   | Desulfitobacterium hafniense  | 62                 | GAPDH                             | -PGK-   | TIM  |      |
|                        | Thermoanaerobacteriaceae   | Moorella thermoacetica ATCC 39073   | 1                  |                                   |         | TIM-PGM-ENO-Prot 1   |      |
| normound or obditional | Thermoanaerobacteriaceae   | Thermoanaerobacter tengcongensis MB4  | 1                  |                                   |         | TIM-ENO-Prot 5 à 8-ARNtMet-Prot 9 à 12   |      |
|                        | Thermoanaerobacteriaceae   | Thermoanaerobacter tengcongensis mo4  | Classe Mollic      |                                   | T GIT   | Tim-End-Flot 3 a 6-ARNumet-Flot 3 a 12   |      |
|                        | 144  | Tit   | A TOTAL CONTRACTOR |                                   | DOW     |  |      |
| ycoplasmatales         | Mycoplasmataceae   | Mycoplasma gallisepticum R  | 2                  | Prot 1 à 18-GAPDH                 |         |  |      |
|                        |  | Mycoplasma genitalium G-37  | 1                  | Prot 1 à 8-GAPDH                  |         |  |      |
|                        |  | Mycoplasma hyopneumoniae 232  | 1                  |                                   |         | Prot 1 à 4   |      |
|                        |  | Mycoplasma mycoides subsp. mycoides SC str. PG1   | 2                  | Prot 1 à 9-GAPDH                  |         |  |      |
|                        |  | Mycoplasma penetrans HF-2   | 2                  | Prot 1-GAPDH                      | -PGK-   | Prot 2 à 16  |      |
|                        |  | Mycoplasma pneumoniae M129  | 1                  | Prot 1 à 10-GAPDH                 | -PGK-   | PTA-Prot 11 à 20   | 1    |
|                        |  | Mycoplasma pulmonis UAB CTIP  | 1                  | Prot 1                            | -PGK    |  |      |
|                        |  | Ureaplasma parvum serovar 3 str. ATCC 700970  | 1                  | Prot 1 à 4                        | -PGK-   | Prot 5 à 15  |      |
| ntomoplasmatales       | Entomoplasmataceae   | Mesoplasma florum L1  | 2                  | Prot 1 à 7-GAPDH                  |         |  |      |
| choleplasmatales       | Acholeplasmataceae   | Onion yellows phytoplasma OY-M  | 1                  | Prot 1 à 8-TIM-Prot 9-ALD-Prot 10 |         | GAPDH-Prot 11 à 21   | II-B |
| on one production of   | rioneropiasmanaces   |   | Classe Bac         |                                   | , , , , |  | 11-0 |
| acillales              | Bacillaceae  | Bacillus anthracis str. 'Ames Ancestor'   | 3                  | ARNtArg-Prot 1 à 2-TR-GAPDH       | DCK     | TIM DOMENO BUILDING  |      |
| aciliales              | Bacillaceae  |   | S 1                |                                   |         |  |      |
|                        |  | Bacillus anthracis str. Stern   | 3                  | ARNtArg-Prot 1 à 2-TR-GAPDH       |         | The state of the s |      |
|                        |  | Bacillus cereus ZK  | 3                  | Prot 1 à 5-TR-GAPDH               |         |  |      |
|                        |  | Bacillus cereus G9241   | 1                  |                                   |         | Prot 1-TIM-PGM-ENO   |      |
|                        |  | Bacillus cereus ATCC 10987  | 4                  | Prot 1 à 2-TR-GAPDH               | -PGK-   | TIM-PGM-ENO  | - 1  |
|                        |  | Bacillus cereus ATCC 10987  | 3                  | ARNt-Prot 1 à 2-TR-GAPDH          | -PGK-   | TIM-PGM-ENO  |      |
|                        |  | Bacillus clausii KSM-K16  | 3                  | TR-GAPDH                          | -PGK-   | TIM-PGM-ENO-Prot 1 à 17  |      |
|                        |  | Bacillus halodurans C-125   | 4                  | Prot 1 à 2-TR-GAPDH               | -PGK-   | TIM-PGM-ENO-Prot 3 à 6   |      |
|                        |  | Bacillus licheniformis ATCC 14580   | 2                  | Prot 1-Prot 2-TR-GAPDH            | PGK-    | TIM-PGM-ENO-Prot 3 à 4   |      |
|                        |  | Bacillus subtilis subsp. subtilis str. 168  | 2                  | AraE-TR-GAPDH                     |         |  |      |
|                        |  |   | 3                  |                                   |         |  |      |
|                        |  | Bacillus thuringiensis serovar konkukian str. 97-27   | 3                  | ARNtArg-Prot 1 à 3-TR-GAPDH       |         |  |      |
|                        | Bacillaceae  | Exiguobacterium sp. 255-15  | 1                  |                                   |         | TIM-PGM-ENO-Prot 2 à 10  |      |
|                        | Bacillaceae  | Geobacillus kaustophilus HTA426   | 2                  | ARNtArg-Prot 1 à 3-TR-GAPDH       |         |  |      |
|                        |  | Geobacillus stearothermophilus  |                    | GAPDH                             |         |  |      |
|                        | Bacillaceae  | Oceanobacillus iheyensis HTE831   | 2                  | TR-GAPDH                          | -PGK-   | TIM-PGM-ENO-Prot 1   |      |
| acillales              | Listeriaceae   | Listeria innocua  |                    | TR-GAPDH                          | -PGK-   | TIM-PGM-ENO-Prot 2 à 5   |      |
|                        |  | Listeria monocytogenes str. 4b H7858  | 1                  | Prot 1-GAPDH                      | -PGK-   | TIM-PGM-ENO-Prot 2 à 3 -?  |      |
|                        |  | Listeria monocytogenes str. 4b F2365  | 1                  | Prot 1 à 5-TR-GAPDH               | -PGK-   | TIM-PGM-ENO-Prot 6 à 13  |      |
|                        |  | Listeria monocytogenes str. 1/2a F6854  | 1                  | ?-Prot 1 à 2-TR-GAPDH             |         |  |      |
| acillales              | Staphylococcaceae  | Staphylococcus aureus   | i l                |                                   |         | TIM-PGM-non ség  |      |
| domaios                | Staphylococcaccac  | Staphylococcus aureus subsp. aureus COL   | 2                  |                                   |         | TIM-PGM-ENO-Prot 1 à 6   |      |
|                        |  | Staphylococcus epidermidis ATCC 12228   | *                  |                                   |         | TIM-PGM-ENO-Prot 1 à 2   | 1    |
|                        |  |   | 2                  |                                   |         |  |      |
| a a taka a Walasa      | 1 t-b ///  | Staphylococcus epidermidis RP62A  | 2                  |                                   |         | TIM-PGM-ENO-Prot 1 à 5   |      |
| actobacillales         | Lactobacillaceae   | Lactobacillus acidophilus NCFM  |                    | ARNtArg-TR-GAPDH                  |         |  |      |
|                        |  | Lactobacillus delbrueckii subsp. lactis   |                    | 7-GAPDH                           |         |  |      |
|                        |  | Lactobacillus gasseri   |                    | non seq-Prot 1-GAPDH              |         |  |      |
|                        |  | Lactobacillus johnsonii NCC 533   | 1                  |                                   |         | TIM-ENO-Prot 1 à 3   |      |
|                        |  | Lactobacillus lactis subsp. lactis II1403   | 2                  | Prot 1 à 6                        |         |  |      |
|                        |  | Lactobacillus plantarum   |                    | 7-TR-GAPDH                        |         |  |      |
|                        |  | Lactobacillus sakei   |                    | ?-TR-GAPDH                        | -PGK-   | TIM-ENO-?  |      |
| actobacillales         | Lactobacillaceae   | Pediococcus pentosaceus ATCC 25745  |                    |                                   |         | TIM-ENO-Prot 1 à 3   |      |
| actobacillales         | Enterococcaceae  | Enterococcus faecalis V583  | 2                  |                                   |         | TIM-ENO-Prot 1 à 2   |      |
| actobacillales         | Leuconostocaceae   | Leuconostoc mesenteroides subsp. mesenteroides  |                    | Prot 1 à 18                       |         |  |      |
| actobacillales         | Streptococcaceae   | Streptococcus pyogenes MGAS8232   | 2                  | Prot 1 à 4 ALD-Prot 5 à 10        |         |  | II-B |
| actobacillales         | Streptococcaceae   | Streptococcus pyogenes M1 GAS   | 2                  | Prot 1 à 4 ALD-Prot 5 à 10        | -PGK-   | Prot 11 à 13   | II-B |
|                        | NAME OF THE PROPERTY OF THE PR | Streptococcus pyogenes SSI-1  | 1                  | Prot 1-ARNtLeu-ALD-Prot 2 à 6     | -PGK-   | Prot 7 à 9   | II-B |
|                        |  | Streptococcus mutans UA159  | 8.0                | ARNt-Prot 1 à 9-GAPDH-            |         |  |      |
|                        |  |   |                    |                                   |         |  | 1    |
|                        |  | Streptococcus suis 89/1591  |                    | GAPDH                             | -PGK-   | Prot 1 à 3   |      |
|                        |  | Streptococcus suis 89/1591 Streptococcus thermophilus CNRZ1066  | 2                  |                                   |         | Prot 1 à 3   |      |
|                        |  | Streptococcus suis 89/1591<br>Streptococcus thermophilus CNRZ1066<br>Streptococcus thermophilus LMG 18311 | 2 2                | GAPDH<br>Prot 1<br>Prot 2         | -PGK    | Prot 1 à 3   |      |

|                  |                      | Phylum Act   | inobacteria Classe | Actinobacteria            |   |
|------------------|----------------------|--|--------------------|---------------------------|---|
|                  | Symbiobacterium      | Symbiobacterium thermophilum IAM 14863                               | 2                  | PGK-                      | Prot 1-TIM-Prot 2   |
| brobacterales    | Rubrobacteraceae     | Rubrobacter xylanophilus DSM 9941                                    | 7                  | ?-Prot 1 à 6 PGK-         | TIM-Prot 7 à 8 -?   |
| inomycetales     | Cellulomonadaceae    | Tropheryma whipplei TW08/27  | 1                  | Prot 1 à 9-GAPDH PGK-     | TIM-Prot 10 à 14  |
| inomycetales     | Corynebacteriaceae   | Corynebacterium diphtheriae NCTC 13129                               | 2                  | Prot 1 à 25 -PGK-         |   |
|                  | Corynebacteriaceae   | Nocardia farcinica IFM 10152   | 1                  | GAPDH -PGK-               | TIM   |
| tinomycetales    | Nocardiopsaceae      | Thermobifida fusca   | - "                | Prot 1 à 4-GAPDH -PGK-    |   |
| tinomycetales    | Mycobacteriaceae     | Mycobacterium tuberculosis H37Rv                                     |                    | GAPDH -PGK-               |   |
|                  |                      | Mycobacterium leprae TN  | 1                  | GAPDH -PGK-               | TIM-Prot 1 à 6  |
|                  |                      | Mycobacterium avium subsp. paratuberculosis str. k10                 | 1                  | Prot 1 à 3-GAPDH -PGK-    | TIM   |
|                  |                      | Mycobacterium bovis AF2122/97  | 1                  | GAPDH -PGK-               | - TIM   |
| tinomycetales    | Propionibacteriaceae | Propionibacterium acnes KPA171202                                    | 4                  | Prot 1 à 5-GAPDH -PGK-    | TIM-Prot 6 à 9  |
| tinomycetales    | Streptomycetaceae    | Streptomyces avermitilis MA-4680                                     | 2                  | Prot 1 à 6-GAPDH -PGK-    |   |
| anioni y cetales | Sir opioniy coldecae | Streptomyces coelicolor A3 (2)                                       | ā I                | Prot 1 à 6-GAPDH -PGK-    |   |
| fidobacteriales  | Bifidobacteriaceae   | Bifidobacterium longum DJO10A  | 3                  | Prot 1 à 10 PGK           | T. 10 T.  |
| nuobacteriales   | Billobacteriaceae    | Bifidobacterium longum NCC2705                                       |                    | Prot 1 à 10 -PGK          |   |
|                  |                      |  | Chlamydiae Classe  |                           | Tim-Flot II a 14  |
| tensi diatas     | Chlamydiaceae        | Chlamydophila abortus S26/3  | dilaniyulae Classi | Prot 1 à 2 -PGK           | Deat 2 à C  |
| nlamydiales      | Chiamydiaceae        | Chlamydophila caviae GPIC  | 2                  | Prot 1 à 2 -PGK-          |   |
|                  |                      |  | 2                  | Prot 1 à 2 -PGK-          |   |
|                  |                      | Chlamydia muridarum Nigg   | 1 1                |                           |   |
|                  |                      | Chlamydophila pneumoniae J138  | 1                  | Prot 1 à 2 -PGK-          |   |
|                  |                      | Chlamydophila pneumoniae TW-183                                      | 1                  | Prot 1 à 2 -PGK           | A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1   |
|                  |                      | Chlamydia trachomatis D/UW-3/CX                                      | 1                  | Prot 1 à 2 -PGK-          | Prot 3 a 6  |
|                  | (a)                  |  | tomycetes Classe   | Planctomycetacia          |   |
| anctomycetales   | Planctomycetaceae    | Rhodopirellula baltica SH 1  | 1                  |                           | Prot 1 à 3  |
|                  |                      |  | Spirochaete Class  |                           |   |
| oirochaetales    | Spirochaetaceae      | Borrelia burgdorferi (Lyme disease spirochete)                       |                    | Prot 1 à 2-GAPDH -PGK     |   |
|                  |                      | Borrelia burgdorferi B31   | 1                  | Prot 1 à 4-GAPDH -PGK-    |   |
|                  |                      | Borrelia garinii PBi   | 1                  | Prot 1 à 4-GAPDH -PGK     | 1.50 (VA) (VA) (VA) (VA)  |
| pirochaetales    | Spirochaetaceae      | Treponema denticola ATCC 35405                                       | 1                  | Prot 1 -PGK               |   |
|                  |                      | Treponema pallidum subsp. pallidum str. Nichols                      | 1                  | PGK-                      | TIM-Prot 1 à 10   |
| pirochaetales    | Leptospiraceae       | Leptospira interrogans serovar Copenhageni str.                      | 1                  | GAPDH -PGK                |   |
|                  | 2 2                  | Fiocruz L1-130 Leptospira interrogans serovar Lai str. 56601         |                    | 2 CARDH BOX               | Prot 1 à 6-TIM-Prot 7   |
|                  |                      |  | cteroidetes Class  |                           | Prot 1 a 0-11m-Prot 7   |
| cteroidales      | Bacteroidaceae       | Bacteroides fragilis YCH46   | 1                  | Prot 1 à 3 -PGK-          | Prot 4 à 9  |
| acterolunies     | Dacteroldaceae       | Bacteroides fragilis NCTC 9343                                       | 4                  | Prot 1 à 3 -PGK           |   |
|                  |                      | Bacteroides tragilis NCTC 9343 Bacteroides thetaiotaomicron VPI-5482 | 4                  | Prot 1 à 3 -PGK           | 100 (0.00)  |
|                  | Pornhummonodoco      | Porphyromonas gingivalis W83   |                    | PGK PGK                   | 11.07 T.17 T.17 T.17  |
|                  | Porphyromonadaceae   |  | usobacteria Classo |                           |   |
|                  |                      | Fusobacterium nucleatum subsp. nucleatum ATCC                        | usopaciena Classi  | r usonacteria             | The appropriate of the second |
| usobacterales    | Fusobacteriaceae     | ir usobacterium nucioatum subsp. nucioatum ATCC                      |                    | Prot 1-GAPDH-Prot 2 -PGK- | D4216   |

#### Tableau 3: Organisation des ORF autour de l'ORF pgk.

Les protéines dont les ORF sont orientées dans le même sens que celle de la PGK sont indiquées par leur nom lorsqu'elles sont impliquées soit par Prot suivi d'un chiffre lorsqu'elles ne sont pas impliquées dans ces voies métaboliques ou lorsque leur fonction n'est pas connue. La protéine TR indique la présence d'un régulateur transcriptionnel, la protéine ALD correspond à une Frutose/Tagatose-1,6-bisphosphate aldolase, la présence d'un point d'intérogation correspond à une séquence inconnue. Les bactéries sont classés selon l'arbre phylogénétique obtenu avec l'ARNr 16S (http://rdp.cme.msu.edu/index.jsp; Cole et al., 2005). Les deux protéines présentant de fortes similitude à une E4PDH ont été appelées ainsi pour Dechloromonas aromatica RCB.

#### I.1.b.α / Les ORF situées en amont de l'ORF pgk

Dans de nombreux phyla, une ORF codant une GAPDH ou une E4PDH est retrouvée en amont de l'ORF pgk. C'est le cas pour les phyla Thermotogae, Deinococcus-Thermus, et aussi chez la majorité des représentants des Firmicutes, Actinobacteria, Spirochaetes, Fusobacteria et Proteobacteria. Dans le phylum des Fusobacteria, dans la classe des α-protéobactéries et dans la famille des Xanthomodaceae des γ-protéobactéries, une protéine de fonction inconnue est quelques fois intercalée entre les ORF de la GAPDH et de la PGK. Dans la majorité des représentants de l'ordre des bacillales, et dans les familles des Clostridiaceae et Lactobaccillaceae du phylum des Firmicutes, l'ORF d'un régulateur transcriptionnel (TR) est présente en amont de l'ORF de la GAPDH.

Notons que dans les classes α à δ appartenant au phylum des protéobactéries, l'ORF d'une transcétolase (TKT) est fréquemment localisée en amont de l'ORF de la GAPDH. La TKT catalyse le transfert d'un groupement cétol de différents substrats donneurs sur différents accepteurs : elle constitue entre autre un lien entre les voies de la glycolyse et la voie des pentoses phosphates (pour revue, Sprenger, 1995). Ainsi, la TKT intervient dans la formation du sédoheptulose 7-phosphate et du glycéraldéhyde 3-phosphate à partir du xylulose 5-phosphate et de ribose 5-phosphate et de fructose 6-phosphate et du glycéraldéhyde 3-phosphate à partir de xylulose 5-phosphate et d'érythrose 4-phosphate (voir Introduction, Fig. 4). L'activité TKT est codée majoritairement par le gène *tktA*, retrouvé en amont des ORF codant la GAPDH et la PGK, et minoritairement par le gène *tktB*.

#### I.1.b.β / Les ORF situées en aval de l'ORF pgk

L'ORF pgk est souvent suivie de l'ORF codant la Triose Phosphate Isomérase (TIM). C'est le cas chez la majorité des phyla Actinobaceria, Spirochaetes, Firmicutes à l'exception de la classe des Mollicutes et de la famille des Streptococcaceae de la classe des Bacilli, et pour une partie des classes δ et ε des protéobactéries. Chez Dehalococcoides ethenogenes 195 du phylum Chloroflexi et Deinococcus radiodurans R1 du phylum Deinococcus-Thermus, une ou deux ORF sont respectivement localisées entre les ORF de la GAPDH et de la TIM. Chez les espèces M. genitalium et M. pneumoniae de la classe des Mollicutes, l'ORF codant la Phosphotransacétylase (PTA) est localisée en aval de pgk. L'ORF codant la phosphoglucomutase (PGM) et l'ORF codant l'énolase (ENO) sont souvent localisées après l'ORF de la TIM, comme c'est le cas pour la classe des Clostridia et pour une partie de la classe des Bacilli.

C'est seulement chez des Protéobactéries des subdivisions  $\alpha$ ,  $\beta$  et  $\gamma$  que nous trouvons la présence d'une aldolase codant une Fructose/Tagatose Bisphosphate Aldolase (ALD ou FBA) en aval de l'ORF pgk. Il est à noter que chez les  $\beta$  Protéobactéries, ainsi que chez les Thiotrichales, les

| Acides aminés spécifiques de l'activité E4PDH                        | E32             | M <sub>179</sub> | VDT <sub>208</sub> KL               | C <sub>311</sub> |
|--|-----------------|------------------|-------------------------------------|------------------|
| Acides aminés spécifiques de l'activité GAPDH                        | D <sub>32</sub> | <b>T</b> 179     | SST <sub>208</sub> GA               | Y <sub>311</sub> |
| Classe $\beta$ Rhodocyclales   |                 |                  |                                     |                  |
| Rhodocyclaceae   | -               | -                | 200                                 |                  |
| Dechloromonas aromatica RCB Daro03001453                             | E               | M                | VSTGL                               | F                |
| Dechloromonas aromatica RCB Daro03001484                             | Ε               | M                | <b>VSTG</b> L                       | F                |
| Classe $\gamma$  |                 |                  |                                     |                  |
| Pseudomonadales  |                 |                  |                                     |                  |
| Pseudomonadaceae   | 2               | PFI              | Verse.                              | _                |
| Azotobacter vinelandii Avin_67                                       | D               | <b>*</b>         | VSTGL                               | F                |
| Pseudomonas aeruginosa PAO1  | D               | M                | V <mark>STG</mark> L                | F                |
| Pseudomonas fluorescens PfO-1  | D               | M                | VSTGL                               | F                |
| Pseudomonas putida KT2440  | D               | M                | V <mark>S</mark> T <mark>G</mark> L | F                |
| Pseudomonas syringae pv. syringae                                    | D               | M                | VSTGL                               | F                |
| Pseudomonas syringae pv. tomato Alteromonadales                      | D               | M                | VSTGL                               | F                |
| Idiomarinaceae   |                 |                  |                                     |                  |
| Idiomarina loihiensis L2TR   | 8               | M                | <b>VDT</b> RL                       | 6                |
| Shewanellaceae   | -               |                  |                                     |                  |
| Shewanella oneidensis MR-1   | E               | M                | VDTKL                               | C                |
| Vibrionales  | -               | -                |                                     | -                |
| Vibrionaceae   | _               | _                |                                     | _                |
| Photobacterium profundum SS9   | Ε               | M                | VDTKL                               | C                |
| Vibrio cholerae O1 biovar eltor str. N16961                          | E<br>E          | M                | VDTKL                               | 9 9              |
| Vibrio fischeri  |                 | M                | VDTKL                               |                  |
| Vibrio logei   | E<br>E          | M                | VDTKL                               | _                |
| Vibrio parahaemolyticus  | Ε               | M                | VDTKL                               | C                |
| Vibrio vulnificus CMCP6  | Ε               | M                | VDTKL                               | c<br>c           |
| Vibrio vulnificus YJ016  | Ε               | M                | VDTKL                               | C                |
| Enterobacteriales  |                 |                  |                                     |                  |
| Enterobacteriaceae   | =               | PT .             | VID TIVE                            |                  |
| Erwinia carotovora subsp. atro                                       | E               | M                | VDTKL                               | C                |
| Escherichia coli CFT073  | E               | M                | VDTKL                               | G                |
| Escherichia coli K12   | E               | M                | VDTKL                               | C                |
| Escherichia coli O157:H7   | E               | M                | VDTKL                               | C                |
| Photorhabdus luminescens subsp laumondii TT01                        | E               | M                | VDTKL                               | C                |
| Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67 | E               | M                | VDTKL                               | C                |
| Salmonella enterica subsp. enterica serovar Paratypi A str. ATCC     | =               | M                | VDTKL                               | C                |
| Salmonella enterica subsp. enterica serovar Typhi str. CT18          | E               | M                | <b>VDTKL</b>                        | C                |
| Salmonella enterica subsp. enterica serovar Typhi Ty2                | E               | M                | VDTKL                               | C                |
| Salmonella typhimurium LT2   | E               | M                | VDTKL                               | C                |
| Shigella flexneri 2a str. 2457                                       | E               | M                | VDTKL                               | C                |
| Shigella flexneri 2a str. 301  | E               | M                | VDTKL                               | C                |
| Yersinia pestis biovar Medievalis                                    | E               | M                | VDTKL                               | C                |
| Yersinia pestis CO92   | E               | M                | VDTKL                               | C                |
| Yersinia pestis K1M  | E<br>E          | M                | VDTKL                               | C                |
| Yersinia pseudotuberculosis IP 32953                                 | Ε               | M                | VDTKL                               | C                |

Tableau 4 : Protéines de type GAPDH/E4PDH présentant tous ou certains des acides aminés spécifiques des E4PDH.

Les séquences protéiques correspondant aux protéines GAPDH/E4PDH des bactéries répertoriées dans le tableau 3 ont été obtenues à partir de la base de données NCBI et ont été alignées à l'aide du logiciel ClustalW. Celles comportant au moins deux résidus spécifiques de l'activité E4PDH (Boschi-Muller et al., 1997) ont été répertoriées ici. Les acides aminés responsables de la distinction des activités E4PDH-GAPDH sont indiqués, respectivement, en blanc sur fond rouge (E4PDH) et noir sur fond jaune (GAPDH). Les espèces bactériennes possédant une protéine comprenant la majorité ou tous les acides aminés spécifiques des E4PDH sont respectivement indiquées en rouge (partiel) et bleu (tous). Les bactéries sont classées selon l'arbre phylogénétique obtenu par comparaison de l'ARNr 16S (http://rdp.cme.msu.edu/index.jsp).

Xanthomonadales et *Legionella pneumophila* str *Pneumophila* de la classe γ des Proteobacteria, l'ORF codant la pyruvate kinase (PYK) est intercalée entre celles de la PGK et celle de l'ALD. Comme ces bactéries sont phylogénétiquement proches, on peut imaginer qu'au cours de l'évolution, l'ORF codant la PYK se soit insérée entre les ORF codant la PGK et l'ALD.

Le groupe d'ORF E4PDH/GAPDH-PGK-FBA qui nous intéresse est donc observé uniquement chez des bactéries des classes  $\alpha$ ,  $\beta$  et  $\gamma$  du phylum des Proteobacteria.

#### I.2 / Distinction entre ORF epd et gap en amont de l'ORF pgk

#### I.2.a / La démarche utilisée

Afin d'identifier les ORF susceptibles de coder une E4PDH, nous nous sommes basés sur les données que l'équipe de G. Branlant avait obtenues au laboratoire sur les E4PDH (Boschi-Muller et al., 1997). A savoir, la présence d'une activité E4PDH est liée à la substitution de certains acides aminés conservés des GAPDH par d'autres acides aminés. La signature E4PDH correspond aux substitutions ponctuelles E32D, M179T, C311Y et au remplacement de la séquence VDTKL entre les positions 206 et 210 par la séquence SSTGA (Fig. 54 et Tableau 4). J'ai donc recherché la présence de ces 3 substitutions ponctuelles et de la séquence VDTKL au sein des séquences identifiées comme codant des protéines de type GAPDH/E4PDH.

L'analyse a été réalisée à la fois sur les ORF de type gap/epd localisées en amont de l'ORF pgk et sur celles localisées ailleurs dans le génome. Pour cela, l'ensemble des séquences de types GAPDH/E4PDH codées par les gènomes séquencés ont été alignées par le logiciel ClustalW.

#### I.2.b / Les résultats

La présence de tous les acides aminés spécifiques des E4PDH a été retrouvée dans les protéines codées par l'ORF de type gap/epd située en amont de pgk, chez tous les représentants de l'ordre des Enterobacteriales et des Vibrionales. Chez les Alteromonadales, seule l'espèce Shewanella oneidensis MR-1 présente une ORF en amont de pgk codant une protéine ayant tous les acides aminés caractéristiques des E4PDH (Tableau 4). Dans le cas de l'espèce Idiomarina loihiensis L2TR, le résidu K209 est remplacé par un résidu arginine, l'activité de la protéine codée devra être vérifiée pour définir son identité. Dans l'ordre des Pseudomonadales, l'ORF de type gap/epd située en amont de pgk présente une partie des acides aminés spécifiques de l'activité E4PDH (M179, VST<sub>208</sub>GL et F311) (Tableau 4). Chez les Pasteurellales, aucune ORF de type gap/epd détectée n'était susceptible de correspondre à une E4PDH. Dans les classes α et β des protéobactéries, les ORF de type gap/epd

situées en amont de *pgk* ne contiennent aucun des acides aminés spécifiques des E4PDH, à l'exception de l'espèce *Dechloromonas aromatica*. Cette espèce pourrait coder deux protéines (Daro03001453 et Daro03001484) présentant certains des acides aminés caractéristiques des E4PDH (E32, M179, VST<sub>208</sub>GL, F311). C'est la seule espèce possédant deux ORF de type *gap/epd* susceptibles de coder une protéine ayant à la fois des caractéristiques de GAPDH et des caractéristiques d'E4PDH.

Il est à noter que l'ORF codant la TKT est plus rarement présente en amont de l'ORF codant l'E4PDH, que de l'ORF codant la GAPDH. Une hypothèse est que ce gène *epd* pouvait avoir été généré par quelques substitutions de bases entraînant un nombre limité de changements d'acides aminés dans la protéine codée, ceci après acquisition, vraisemblablement, par un transfert horizontal du gène *gapA* codant la GAPDH active (Branlant et Branlant, 1985; Martin *et al.*, 1993). Sur la base de cette hypothèse, on peut imaginer que ces évènements aient pu ensuite parfois être suivis par une perte de l'ORF *tkt* localisée en amont de l'ORF *epd*.

#### I.3 / Classe d'aldolase produites par les groupes d'ORF epd/gapA-pgk-ald

Comme nous l'avons décrit dans la présentation du sujet, il existe deux classes d'aldolases agissant, dans le cas de la classe I sur le fructose-1,6-bisphosphate, dans le cas de la classe II sur le fructose-1,6-bisphosphate ou le tagatose-1,6-bisphosphate. La distinction entre classe I et classe II repose sur une différence de mécanisme catalytique. Dans les enzymes de classe I, les résidus Lysine 146 et 229 sont essentiels : la Lysine 146 est nécessaire pour la catalyse et intervient dans le positionnement du substrat. La Lysine 229 est impliquée dans la formation d'une base de Schiff avec le substrat. Dans les enzymes de classe II, les histidines 110 et 264 sont des ligands du zinc catalytique. Les résidus Asn35, Ser61, Asp109, Arg 331 et Asp288 interagissent avec le substrat de l'enzyme, et Glu182 est nécessaire à l'activité de l'enzyme (Zgiby et al., 2002). Les enzymes de classe II peuvent être encore subdivisées en enzymes de type A ou de type B sur la base de leur séquence (Plaumann et al., 1997; Sanchez et al., 2002). Chacun de ces deux types d'enzymes de classe II est caractérisé par des similarités de séquences de l'ordre de 40 %. La similarité de séquence n'est que de l'ordre de 25 à 30 % entre les enzymes appartenant à 2 types différents. Les aldolase de type IIA portent une activité fructose-1,6-bisphosphate aldolase (FBA), par contre les aldolases de type IIB peuvent avoir, soit une activité FBA, soit une activité Tagatose-1,6-bisphosphate aldolase (Zgiby et al., 2000). Nous appelerons par la suite Aldolase (ALD) les enzymes de classe IIB et Fructose-1,6-Bisphosphate Aldolase (FBA) celles de classe I et IIA.

L'alignement des séquences protéiques déduites des séquences des génomes montre que chez les α-protéobactéries, lorsqu'une ORF d'aldolase est présente en aval de l'ORF pgk, il s'agit dans la majorité des cas d'une FBA de classe I. Par contre, chez les β-protéobactéries, les ORF d'aldolases

localisées en aval de pgk appartiennent toujours à la classe IIB. Trois cas de figures sont trouvés chez les γ-protéobactéries. Dans l'ordre des Xanthomonadales, l'aldolase appartient à la classe I. Chez les espèces renfermant une ORF gap en amont de pgk, l'aldolase appartient à la classe IIB. Par contre, chez les espèces renfermant une ORF epd en amont de pgk, l'aldolase est de classe IIA, ceci à l'exception de l'ordre des Alteromonadales. Chez les Pasteurellales, où il n'y a ni ORF epd, ni ORF gap en amont de pgk, l'aldolase est de classe IIA. L'ensemble de ces données suggère une évolution concertée des ORF situées en amont et en aval de l'ORF pgk, avec une perte de l'ORF epd dans l'ordre des Pasteurellales.

# II / Etat des travaux concernant l'expression du tandem d'ORF epdpgk lorsque j'ai entrepris ce travail

V. Bardey, N. Robas et B. Charpentier avaient débuté l'étude de l'expression du tandem d'ORF *epd-pgk* au laboratoire. Ils avaient tout d'abord montré que le gène *epd* est transcrit à partir d'un seul promoteur reconnu par l'holoenzyme Eσ<sup>70</sup>, le promoteur *epd* P0, et que ce promoteur contribue à l'expression de la protéine PGK (Charpentier *et al.*, 1998). Ce promoteur *epd* P0 est fortement activé en présence de glucose dans le milieu de culture, et dépend de l'expression du transporteur spécifique du glucose EIIBC<sup>Gle</sup> (Charpentier *et al.*, 1998). Cette activation par le glucose du promoteur *epd* P0 dépend aussi du complexe CRP-AMPc dans les cellules. Un site de fixation potentiel du complexe CRP-AMPc, centré 70,5 pb en amont du site d'initiation de la transcription, est présent en amont du gène (Fig. 60). Ce positionnement était compatible avec une activation par un mécanisme de classe I. En fait, le rôle du complexe CRP-AMPc s'est avéré être indirect : ce complexe est nécessaire pour la production du transporteur EIIBC<sup>Gle</sup>, qui permet l'activation du promoteur *epd* P0 par un mécanisme encore inconnu (Charpentier *et al.*, 1998). D'autre part, l'équipe de G. Branlant au laboratoire avait mis en évidence la faible expression de l'ORF *epd*, en présence de glucose (Boschi-Muller *et al.*, 1997). Ainsi, bien que l'ORF *epd* soit précédée par un promoteur transcriptionnel fort, la protéine E4PDH est très faiblement produite.

V. Bardey et N. Robas avaient poursuivi l'étude de l'expression du tandem d'ORF epd-pgk en utilisant différentes constructions plasmidiques, portant les ORF epd-pgk, avec ou sans le promoteur epd P0 (Fig. 1A de l'article 1). Ils avaient pu mettre en évidence un autre promoteur pgk P1. Ce promoteur localisé dans l'ORF epd ne renferme pas de boîte -35, mais comporte une boîte -10 étendue (Fig. 4A de l'article 1). Il n'est pas activé par la présence de glucose dans le milieu de croissance, contrairement au promoteur epd P0 (Fig. 4C de l'article 1). La réalisation et l'étude de différentes constructions plasmidiques avait aussi permis de mettre en évidence des sites de clivages pgk 2 à 6

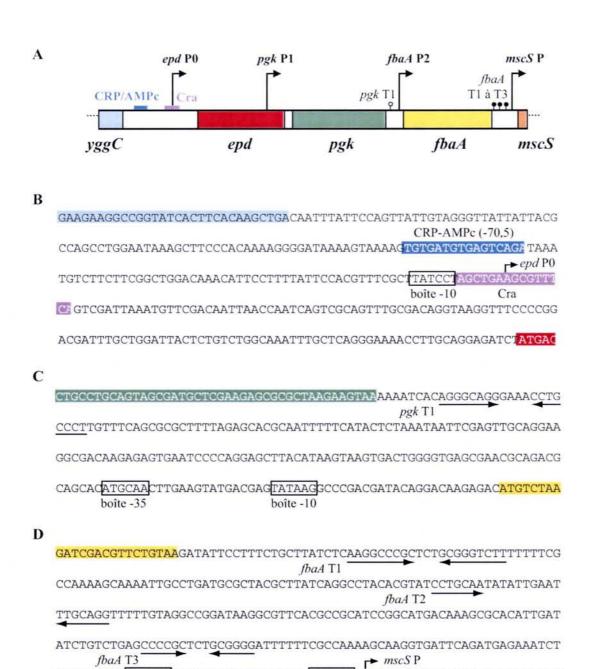


Figure 60 : Représentation schématique du triplet d'ORF epd-pgk-fbaA et séquence de la région promotrice epd P0 et des régions intergénique pgk-fbaA et fbaA-mscS.

boîte -10

GTGATCTATTTGGCAAAATTATGCTTTATTGTTTACCCTTGTCAGACTGCCCGTCATAAGGCGGCG

GAGTGTATTTCTCCATTTTGAGTCAGTTGAAAAGGAATATTGAATGGAAGATTTGAATGTTGTCGA

A : Organisation du triplet d'ORF *epd-pgk-fbaA*. Les ORF sont représentées par des rectangles colorés. Les promoteurs connus ou potentiels sont indiqués par des flèches brisées, les terminateurs par des tiges surmontées d'une boucle. Les sites de fixation au complexe CRP-AMPc et à la protéine Cra sont représentés respectivement par des rectangles bleu et violet.

B, C et D : Séquence de la région promotrice *epd* P0 et des régions intergéniques *pgk-fbaA* et *fbaA-mscS* de la région en aval de l'ORF *fbaA*. Les séquences des différentes ORF sont représentées avec le même code couleur qu'en A. Les boîtes -10 et -35, connues ou proposées sont représentées par des rectangles. Les terminateurs rho-indépendant sont indiqués sur les séquences par des flèches.

dans l'ORF epd et la région interORF epd-pgk se produisant in vivo (Fig. 4B de l'article 1). V. Bardey et N. Robas avaient alors utilisé différentes souches mutées dans différents gènes de RNases, ce qui avait permis de montrer que la RNase E est responsable des clivages localisés dans l'ORF epd et la région interORF epd-pgk (Fig. 6A de l'article 1). Ces clivages par la RNase E avaient aussi été confirmés in vitro par l'utilisation d'une fraction cellulaire contenant le dégradosome (Fig. 6B de l'article 1). En complément de ce travail, F. Torterotot et A. Mougin avaient déterminé la structure secondaire de la région du transcrit epd-pgk contenant les sites de clivages pgk 2 à 6 (Fig. 7A de l'article 1). La structure mise en évidence pour le site pgk 2 (région en simple-brin riche en A et U, entourée de deux structures tiges-boucles) était caractéristique de certains sites forts de clivage par la RNase E, comme celui présent dans l'ARN préribosomique. V. Bardey avait aussi observé une conservation de cette structuration de l'ARN contenant le site de clivage pgk 2, chez Salmonella typhi et Yersinia pestis. L'existence de clivage dans l'ORF epd suggérait quelle était faiblement traduite, les ribosomes en cours de traduction ayant un rôle protecteur du transcrit vis-à-vis des clivages par les endonucléases. V. Bardey avait donc étudié l'efficacité de traduction de l'ORF epd, ceci en utilisant des fusions transcriptionnelles et traductionnelles. Il avait ainsi mis en évidence une faible efficacité de la région TIR (Translation Initiation Region) de l'ORF epd. Il avait aussi montré, par mutagenèse dirigée, que cette faible efficacité résulte, entre autre, d'un faible écartement entre la séquence de liaison au ribosome (RBS) et le codon d'initiation de la traduction.

## III / Poursuite de l'étude du triplet d'ORF epd-pgk-fbaA

Comme nous l'avons montré ci-dessus, l'analyse des génomes de γ-protéobactéries mettait en évidence la présence d'une ORF *fbaA* codant une FBA de classe II et de type A, en aval du tandem *epd-pgk*. Etant donné que la FBA codée par cette ORF intervient dans le métabolisme du glucose, on pouvait se demander si elle pouvait être coexprimée avec l'ORF *pgk* et si oui, il nous paraissait important de reprendre l'étude réalisée par V. Bardey et N. Robas sur les seules ORF *epd-pgk* en utililisant des constructions exprimant les 3 ORF.

Nous pouvions exclure la possibilité que l'ORF mscS, située en aval de fbaA, soit coexprimée avec les ORF epd, pgk et fbaA, car elle est principalement exprimée en phase stationnaire de croissance, contrairement aux ORF pgk et fbaA (Peng et Shimizu, 2003; Stokes et al., 2003). Un autre argument suggérant une absence de coexpression est la présence d'un ensemble de 3 terminateurs transcriptionnels rho-indépendants en aval de l'ORF fbaA, ce qui est en faveur d'un arrêt total de la transcription à la fin de l'ORF fbaA. D'autre part, les ORF situées en amont de l'ORF epd sont peu conservées chez les Enterobacteriaceae et il y avait peu de chances quelles soient coexprimées avec les 3 ORF epd, pgk et fbaA (Tableau 3).

Plusieurs données suggéraient que l'expression des ORF pgk et fbaA puissent être exprimées de manière indépendantes, et d'ailleurs dans les bases de données, ces 2 ORF sont proposées comme étant exprimées de manière indépendante (<a href="http://www.cifn.unam.mx/Computational\_Genomics/regulondb/">http://www.cifn.unam.mx/Computational\_Genomics/regulondb/</a>). Néanmoins une co-expression ne pouvait pas être exclue. En défaveur d'une coexpression de pgk et fbaA, la région interORF pgk-fbaA est grande (213 nts) et elle comporte un terminateur pgk T1 rhoindépendant potentiel, dont l'activité n'avait pas été étudiée (Alefounder et Perham, 1989, Lesnik et al., 2001). Par ailleurs, sur la base d'un alignement de séquence, un promoteur potentiel, reconnu par l'holoenzyme Eσ<sup>70</sup>, situé à environ 17 pb en amont de l'ORF fbaA avait été proposé, cependant, son activité n'avait jamais été testée expérimentalement (Fig. 60) (Salgado et al., 2004). Seules, les données d'une étude du transcriptome d'E. coli par puces à ADN apportaient des données en faveur de la cotranscription de l'ORF fbaA avec l'ORF pgk (Tjaden et al., 2002b; Tjaden et al., 2002a). En effet, la région intergénique pgk-fbaA semblait être transcrite au même taux que les 2 ORF pgk et fbaA. Des expériences devaient donc être réalisées pour savoir si ces 2 ORF étaient coexprimées.

Nous avons donc décidé de cloner l'ensemble des 3 ORF epd-pgk-fbaA avec le promoteur epd P0 situé en amont d'epd et les terminateurs transcriptionnels situés en aval de fbaA. Nous avons aussi produit d'autres constructions comportant les 3 ORF sans le promoteur epd P0, l'une avec le promoteur pgk P1 sauvage, l'autre contenant un promoteur pgk P1 muté (Fig. 1A de l'article 1). L'étude des souches d'E. coli TG1 transformées par ces plasmides nous a permis de montrer, comme il est décrit dans l'article ci-joint, que l'ORF fbaA n'est exprimée qui si au moins l'un des deux promoteurs epd P0 ou pgk P1 est présent (Fig. 3 de l'article 1). J'ai par ailleurs mis au point un test permettant de mesurer l'activité FBA. En effet, cette enzyme n'avait jamais été étudiée au laboratoire. Grâce à la mesure de cette activité enzymatique, j'ai pu montrer que les 2 promoteurs contribuent à peu près à égalité à la production de PGK et FBA, ceci en présence de glucose dans le milieu de culture. Bien que seulement deux promoteurs transcriptionnels soient utilisés pour la transcription des 3 ORF epd, pgk et fbaA, l'analyse par Northern-Blot des ARN produits à partir de ce triplet d'ORF, nous a permis de mettre en évidence une grande complexité des ARNm, avec des ARN tricistroniques, bicistroniques et monocistroniques (Fig. 2 de l'article 1). Cette complexité pouvait être liée à la fois à l'utilisation du terminateur transcriptionnel potentiel pgk T1 situé en aval de pgk, mais aussi à l'existence de sites de clivages par des RNases. En effet, V. Bardey avait mis en évidence des sites de clivage dans l'ORF epd et la région inter-ORF epd-pgk, ceci en utilisant le tandem d'ORF epd-pgk. J'ai vérifié que les coupures par la RNase E aux sites pgk 2 à 6 avaient aussi lieu dans les longs transcrits epd-pgk-fbaA. Par ailleurs, j'ai recherché l'existence de sites de clivages par les RNases dans la région interORF pgk-fbaA. En utilisant différentes souches d'E. coli mutées dans des gènes de RNases, j'ai ainsi pu mettre en évidence 2 sites de clivage par la RNase E dans cette région inter-ORF (Fig. 5 de l'article).

Figure 61 : Localisation des sites de clivages enzymatiques par les RNases, et des modifications chimiques des bases de l'ARN par le DMS, CMCT et kethoxal

(d'après Ehresmann et al., 1987).

#### Sondes enzymatiques:

RNase T1 : G simple brin

RNase T2: A simple brin, puis G simple brin, et enfin C et U

RNase V1 : N double brin ou empilé

#### Sondes chimiques:

DMS (diméthyl sulfate): N3-C, N1-A, N7-G simple brin

CMCT (1-cyclohexyl-3-(2-morpholinoéthyl) carbodiimide métha-p-toluène sulfonate):

N3-U puis N1-G simple brin

kethoxal : N1-G, N2-G simple brin

Il nous a alors paru intéressant d'établir la structure secondaire de la région interORF pgk-fbaA contenant ces deux sites de clivages (Fig. 8A de l'article 1). En effet, une notion nouvelle était apparue en ce qui concerne les sites d'action de la RNase E. Comme nous l'avons dit dans l'introduction, cette enzyme clive plus facilement les ARN portant un résidu monophosphate à leur extrémité 5' du fait de la présence d'un site de fixation du phosphate 5' terminal qui augmente lorsqu'il est simple-brin activité enzymatique. Néanmoins, ceci pose le problème du clivage initial des ARNm qui portent un triphosphate à leur extrémité 5' par cette enzyme. Récemment, la notion d'existence de sites internes présentant une structure particulière permettant de passer outre la nécessité d'un résidu monophosphate en 5' a été émise et nos travaux sur la structure de la région interORF pgk-fbaA et ceux de V. Bardey et A. Mougin sur la structure de la région inter-ORF epd-pgk apportent des informations sur la structure et la séquence des sites internes de clivages par la RNase E (site d'entrée).

Pour l'identification des endonucléases impliquées dans les clivages des transcrits epd-pgk-fbaA, E. Hajnsdorf et P. Régnier nous avaient fourni un ensemble de souches isogéniques d'E. coli dans lesquelles les gènes codant la RNase III, la RNase E ou la RNase G étaient individuellement mutés. Aucun effet des RNases III et G n'a été détecté, seule la RNase E coupe les transcrits epd-pgk-fbaA, epd-pgk et pgk-fbaA. Comme décrit dans les compléments à l'article, j'ai aussi étudié la stabilité des produits de clivages par la RNase E in vivo, ceci en bloquant la transcription par ajout de rifampicine dans le milieu de culture.

Pour l'analyse de la structure secondaire de la région inter-ORF pgk-fbaA, j'ai utilisé l'ensemble des sondes enzymatiques et chimiques utilisées au laboratoire et D. Ropers m'a transmis son expérience de l'utilisation de ces sondes. Trois enzymes ont été utilisées :

- les RNases T1 et T2 qui hydrolysent préférentiellement les régions d'ARN simple brin en conditions limitantes, et qui clivent respectivement les liaisons phosphodiester en 3' des résidus G pour la RNase T1 ou les liaisons phosphodiester en 5' de tous les résidus pour la RNase T2
- la RNase V1 qui clive spécifiquement des régions d'ARN en double-brin de type Watson-Crick ou structure par empilement (Fig. 61).

J'ai aussi utilisé 3 agents chimiques :

- le diméthylsulfate (DMS), qui méthyle les positions N-7 des guanines, N-1 des adénines et plus lentement, les positions N-3 des cytosines,
- le 1-cyclohexyl-3-(2-morpholinoéthyl) carbodiimide métha-p-toluène sulfonate (CMCT) qui ajoute un groupement carbodiimiide en position N-3 des résidus uridines et, à un moindre degré, en position N-1 des résidus guanines,
- le kéthoxal qui méthyle les positions N-1 et N-2 des guanines.
   Les ARN modifiés ont ensuite été analysés par élongation d'amorces par la transcriptase inverse.
   Lorsqu'un résidu est modifié chimiquement sur une position impliquée dans la formation des liaisons

de type Watson-Crick, un arrêt d'élongation d'amorce est observé un nucléotide en amont du résidu modifié. Pour les clivages enzymatiques, l'arrêt a lieu au niveau du dernier résidu en 5' de l'ARN pour les clivages par les RNases T1, T2 et V1. La localisation des clivages enzymatiques et des modifications chimiques a permis de caractériser les segments en simple-brin ou en double-brin. Les données ont ensuite été utilisées pour rechercher une structure secondaire thermodynamiquement stable et compatible avec les données expérimentales par emploi du logiciel M-fold (http://www.bioinfo.rpi.edu/applications/mfold/old/rna/form1.cgi). Enfin cette étude expérimentale et informatique a été complétée par une étude phylogénique: j'ai recherché si les structures possibles étaient compatibles avec la séquence des régions interORF pgk-fbaA d'autres bactéries de la classe des γ-protéobactéries. Nous avons ainsi abouti à un modèle phylogéniquement conservé et en accord avec les données expérimentales.

En résumé, l'ensemble de cette étude du triplet d'ORF epd-pgk-fbaA montre l'existence de nombreux niveaux de régulations au niveau transcriptionnel et post-transcriptionnel :

- l'ORF fbaA appartient à un opéron contenant aussi le tandem d'ORF epd-pgk,
- la transcription de cet opéron dépend d'un promoteur fortement régulé, le promoteur *epd* P0, qui est activé en présence de glucose et d'un promoteur peu régulé, *pgk* P1 de la classe des promoteurs étendus.
- le processus de maturation/dégradation des transcrits est initié par la RNase E. Nous avons mis en évidence deux sites de d'entrée de la RNase E et défini leurs caractéristiques structurales. La structure secondaire de chacun des sites de clivage internes au niveau des régions *epd-pgk* et *pgk-fbaA* est conservée chez les *Enterobacteriaceae*, ce qui suppose la conservation du mécanisme de maturation/dégradation des transcrits issus de l'opéron *epd-pgk-fbaA* chez ces bactéries.
- la faible expression de l'ORF epd par rapport aux ORF pgk et fbaA s'explique par une traduction inefficace, en partie due à un écartement sous-optimal entre la séquence de fixation des ribosomes et le codon initiateur de l'ORF epd, les données suggèrent aussi l'existence d'un facteur inhibiteur saturable agissant en trans.

## Publication n°1

Characterization of the molecular mechanisms involved in the differential production of Erythrose-4-phosphate dehydrogenase,

3-Phosphoglycerate kinase and class II Fructose-1,6bisphosphate aldolase in *Escherichia coli* 

V. Bardey, C. Vallet, N. Robas, B. Charpentier, B. Thouvenot, A. Mougin,
E. Hajnsdorf, P. Régnier, M. Springer, and C. Branlant

## Characterization of the molecular mechanisms involved in the differential production of erythrose-4-phosphate dehydrogenase, 3-phosphoglycerate kinase and class II fructose-1,6-bisphosphate aldolase in *Escherichia coli*

Vincent Bardey, 118 Corinne Vallet, 18 Nathalie Robas, 1 Bruno Charpentier, 1 Benoit Thouvenot, 1 Annie Mougin, 24 Eliane Hajnsdorf, 2 Philippe Régnier, 2 Mathias Springer and Christiane Branlant 1\* 1 Laboratoire de Maturation des ARN et Enzymologie Moléculaire, UMR 7567 CNRS-UHP Nancy I, Faculté des Sciences et Techniques, BP 239, 54506 Vandoeuvre-lès-Nancy, Cedex, France. 2 Institut de Biologie Physico-Chimique CNRS UPR 9073, rue Pierre et Marie Curie, 75005 Paris, France.

#### Summary

A gapA-pgk gene tandem coding the glyceraldehyde 3-phosphate dehydrogenase and 3-phosphoglycerate kinase, is most frequently found in bacteria. However, in Enterobacteriaceae, gapA is replaced by an epd open reading frame (ORF) coding an erythrose-4phosphate dehydrogenase and an fbaA ORF coding the class II fructose-1,6-bisphosphate aldolase follows pgk. Although epd expression is very low in Escherichia coli, we show that, in the presence of glucose, the 3 epd, pgk and fbaA ORFs are efficiently cotranscribed from promoter epd P0. Conservation of promoter epd P0 is likely due to its important role in modulation of the metabolic flux during glycolysis and gluconeogenesis. As a consequence, we found that the epd translation initiation region and ORF have been adapted in order to limit epd translation and to create an efficient RNase E entry site. We also show that fbaA is cotranscribed with pgk, from promoter epd P0 or an internal pgk P1 promoter of the extended -10 class. The differential expression of pgk and fbaA also depends upon an RNase E segmentation process, leading to individual mRNAs with different sta-

Accepted 8 June, 2005. \*For correspondence. E-mail Christiane. Branlant@maem.uhp-nancy.fr; Tel. (+33) 3 83 68 43 03; Fax (+33) 3 83 68 43 07. Present addresses: \*Laboratoires Sérobiologiques-Cognis France, 3 rue de Seichamps, BP40, 54425 Pulnoy, France; \*Laboratoire de Biologie Moléculaire Eucaryote UMR5099-CNRS, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 4, France. \*The two authors contributed equally to this work.

bilities. The secondary structures of the RNA regions containing the RNase E sites were experimentally determined which brings important information on the structural features of RNase E ectopic sites.

#### Introduction

Escherichia coli is capable of using numerous carbohydrates as single source of carbon and energy. However, glucose is preferentially used in the presence of other carbon sources, so that glucose metabolism plays an important role for growth of E. coli. In all living cells, the phosphorylating glyceraldehyde 3-phosphate dehydrogenase (GAPDH), 3-phosphoglycerate kinase (PGK) and class II fructose-1,6-bisphosphate aldolase (FBA) are important enzymes for carbon metabolism (Fraenkel, 1996). In most bacterial genomes, the gene encoding the phosphorylating GAPDH (gap) is located upstream of the PGK gene (pgk) (Conway and Ingram, 1988; Branlant et al., 1989; Eikmanns, 1992; Meijer et al., 1996; Gebbia et al., 1997; Branny et al., 1998). In Gram-positive bacteria, such as Bacillus megaterium (Schlapfer and Zuber, 1992), Bacillus subtilis (Ludwig et al., 2001), Corynebacterium glutamicum (Eikmanns, 1992) and Lactobacillus delbrueckii (Branny et al., 1998), the gap-pgk gene tandem is followed by a gene encoding another glycolytic enzyme, the triose phosphate isomerase (TPI). However, in γ-proteobacteria, that include several highly pathogenic bacteria species such as E. coli, Salmonella typhimurium, Yersinia pestis, Pseudomonas aeruginosa, Shewanella putrefaciens and Vibrio cholerae, the pgk gene is followed by the fbaA gene encoding a class II FBA and instead of being preceded by a gap gene, it is preceded by an epd gene (previously designated as gapB) (Alefounder and Perham, 1989; Zhao et al., 1995; Della Seta et al., 1997; Charpentier et al., 1998; Bardey, 2000; Fillinger et al., 2000) (http://www.tigr.org). The epd gene encodes a protein showing strong sequence (Boschi-Muller et al., 1997; Fillinger et al., 2000) and 3D structure (Kauffman et al. pers. comm.) similarities with GAPDHs. However, it has erythrose-4-phosphate dehydrogenase activity (E4PDH) (Zhao et al., 1995; Boschi-Muller et al., 1997;

© 2005 Blackwell Publishing Ltd

Della Seta et al., 1997) and catalyses the oxidation of erythrose-4-phosphate into 4-phosphoerythronate in the presence of NAD+ (Zhao et al., 1995). It was proposed that the ancestor gap gene encoding a phosphorylating dehydrogenase acting specifically on glyceraldehyde 3phosphate (G3P) was at sometime during evolution converted into the present epd gene encoding a nonphosphorylating dehydrogenase, that acts efficiently on erythrose-4-phosphate (Boschi-Muller et al., 1997). Indeed, conversion of a GAPDH into an E4PDH by site directed-mutagenesis was found to depend upon a limited number of amino acid substitutions: four punctual amino acid replacements (D32E, T179M, P187Y, Y311C) and substitution of the SSTGA sequence around residue T208 in GAPDHs for the VDTKL sequence in E4PDHs (Boschi-Muller et al., 1997; Fillinger et al., 2000). As a relic of the previous GAPDH activity, E4PDH has a residual poorly efficient phosphorylating dehydrogenase activity on G3P (Zhao et al., 1995; Boschi-Muller et al., 1997), but it is not sufficient to ensure growth on glucose in the absence of the gapA gene encoding the fully active phosphorylating GAPDH (Della Seta et al., 1997). The gapA gene is located at a distinct locus on the E. coli chromosome (40.1 min, instead of 66.2 min for the epd-pgk gene tandem) (Hillman and Fraenkel, 1975; Alefounder and Perham, 1989). It may have arisen from a duplication of the ancestor gene or an horizontal transfer of the genetic information (Branlant and Branlant, 1985; Martin et al., 1993).

As the expression of the genes coding for glycolytic enzymes have to be coordinated, it was interesting to know how regulations of the gapA and pgk glycolytic genes have evolved, following their physical separation. In addition, whereas the epd and pgk open reading frames (ORFs) are cotranscribed from a common strong promoter (previously denoted gapB P0, and now denoted epd P0 (Saier and Ramseier, 1996; Charpentier et al., 1998), the E4PDH production is very low in all the growth conditions tested (Zhao et al., 1995; Boschi-Muller et al., 1997). The epd P0 promoter activity is strongly increased in the presence of glucose in the medium and promoter epd P0 is responsible for about half of the cellular PGK activity (Charpentier et al., 1998). Interestingly, although the pgk and fbaA ORFs are found side by side on the E. coli genome, their possible coexpression or their independent expression had not been investigated.

Bacteria use several transcriptional and post-transcriptional mechanisms for the differential expression of proteins from polycistronic mRNAs (for review: Condon et al., 1995; Romby and Springer, 2003). Surprisingly, in spite of the strong importance of glycolytic genes in E. coli, very limited information was available on the regulation of the epd-pgk-fbaA ORF triplet: based on nucleotide sequence inspection, three transcription start-sites were

previously predicted. However, only the presence of promoter epd P0, its activation in the presence of glucose and its downregulation by the Cra protein in the presence of pyruvate were experimentally demonstrated (Ramseier et al., 1995; Saier and Ramseier, 1996; Charpentier et al., 1998). The fact that the epd P0 promoter is responsible for only half of the PGK production (Charpentier et al., 1998) and previous data obtained by Tn5 insertion mutagenesis (Nellemann et al., 1989) suggested the presence of a yet unidentified internal transcriptioninitiation site in the ORF cluster. The activities of a predicted Eo70 initiation site, located 30 nts upstream of the fbaA start codon (Salgado et al., 2004) and of several putative rho-independent terminators, one located between the pgk and fbaA ORFs and three other ones sitting, side by side, downstream from the fbaA ORF (Alefounder and Perham, 1989; Lesnik et al., 2001), had not been tested experimentally. Finally, as polycistronic mRNAs encoding GAPDH and PGK genes were found to be subjected to processing mechanisms in Zymomonas mobilis and B. subtilis (Burchhardt et al., 1993; Ludwig et al., 2001), the E. coli epd-pgk-fbaA transcript may also be subjected to a complex processing mechanism, which also had to be investigated.

In the present study, we tried to bring answers to several of these questions. By Northern blot hybridization, Primer extension analysis, site-directed mutagenesis and S1 mapping, we established the transcriptional organization of the *E. coli epd-pgk-fbaA* genomic region, and studied the maturation and degradation processes of the various transcripts. The RNase involved was identified by using *E. coli* strains defective in the production of various RNase(s). The secondary structure of the cleaved RNA regions was established by the use of chemical and enzymatic probes of the RNA structure and by phylogenetic comparisons. Finally, by construction of *epd-lacZ* transcriptional and translational fusions, we demonstrated a translational limitation of the *epd* expression in *E. coli*.

#### Results

The E. coli epd-pgk-fbaA locus generates a complex array of mRNAs

To study the transcripts and cleavage products of the <code>epd-pgk-fbaA</code> ORF cluster (Fig. 1A), this ORF cluster was cloned into a multicopy plasmid, together with the 243 bp long upstream and 431 bp long downstream DNA regions (plasmid pEPF) (Fig. 1A). A shorter construct, missing promoter <code>epd</code> P0, was built (plasmid p $\Delta$ P0-EPF) (Fig. 1A). The <code>E. coli</code> TG1 strain transformed with plasmid pEPF or p $\Delta$ P0-EPF (designated as pEPF- or p $\Delta$ P0-EPF-transformed cells) or with the empty pBS plasmid

epd-pgk-fbaA expression in E. coli 1267

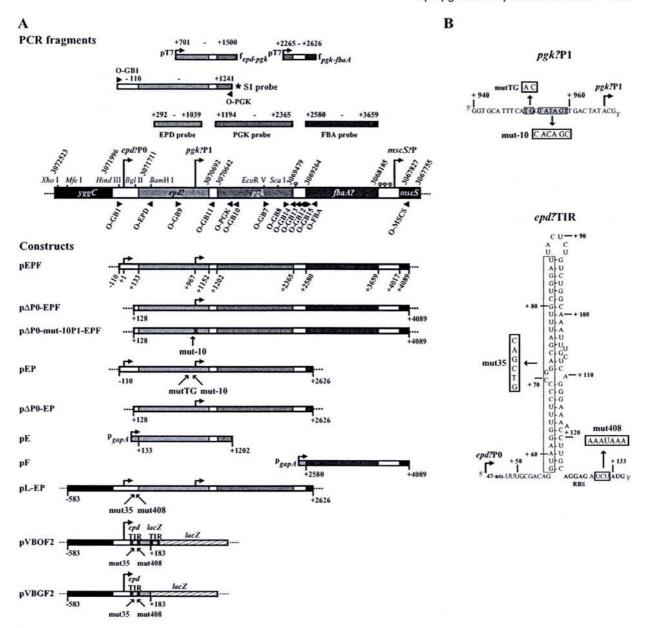


Fig. 1. Schematic representation of the *epd-pgk-fbaA* ORF cluster with the different constructs, probes and PCR fragments used in this study. The *epd-pgk-fbaA* ORF cluster, bordered by the *yggC* and *mscS* ORFs is depicted at the centre of A. White rectangles represent untranslated regions, grey and black rectangles correspond to the ORFs. The *epd* P0 (Charpentier *et al.*, 1998), *pgk* P1 (this study) and *mscS* P (Stokes *et al.*, 2003) transcription-start sites are marked by bent arrows. Potential terminators are symbolized by thin lines surmounted by circles. The positions of the restriction sites used are indicated. The oligonucleotide targeted sequences are shown by black triangles (▶ non-coding, ◄ coding strand). Numbering of the genomic fragment is according to the complete sequence of the *E. coli* K12 chromosome (Blattner *et al.*, 1997). For numbering of fragments inserted in the various constructs (bottom part of A) and PCR amplification products (upper part of A), the *epd* P0 transcription-start site was taken as the reference. The DNA matrices used for *in vitro* transcription of the f<sub>epd-pgh</sub>, f<sub>pgk-tbaA</sub> RNAs and the 5′-end-labelled DNA probe used for S1 mapping are shown in the upper part of A. The star indicates the labelled extremity. The PCR amplification products used as labelled probes for Northern blot analysis of transcripts from untransformed cells (EPD, PGK and FBA probes) are also shown in the upper part of A. The DNA fragments of the *epd-pgk-fbaA* ORF cluster inserted in the plasmids utilized to transform the various *E. coli* strains used in this study (Table 2), are shown in the bottom part of the figure. The plasmid names are given on the left side of the panel. The colour rules used for the fragments are as in the central scheme. The mutations generated in the plasmids (mut35 and mut408 in the *epd* TIR, and mutTG and mutTG and mutTG are proposed 2D structure of the *epd* 5′ UTR region.

(control-transformed cells) were grown until mid-log phase in M63 minimal medium, in the presence of glucose as the sole source of carbon. For each transformant, one part of the culture was used for protein analysis by SDS-PAGE, the other half was used for Northern blot analysis of total RNA (Figs 2A and 3). The oligonucleotides O-EPD, O-PGK and O-FBA, that are, respectively, complementary to each of the epd, pgk and fbaA ORFs, were used as the probes (Fig. 1A). For each culture, total RNA and soluble proteins were always extracted from a similar number of cells (about 1.5 × 109) and the amounts of total RNA or soluble proteins loaded on the gels (Figs 2A and 3) were determined taking into account the plasmid copy numbers in transformed cells (plasmid p∆P0-EPF was present at a higher copy number compared with plasmid pEPF). The experiment was repeated several times, and reproducible results were obtained. The strong PGK and FBA productions and the very low E4PDH production observed with the cloned ORF triplet is illustrated in Fig. 3 [compare lane 2 (ORF triplet) to lanes 7 (PGK), 8 (FBA) and 9 (E4PDH)]. A strong production of PGK and FBA was also observed in the absence of promoter epd P0 (lane 3).

When promoter epd P0 was present in the plasmid, each of the oligonucleotide used for Northern blot analysis reproducibly revealed mRNAs ranging in size from 3700 to 1100 nts (Fig. 2A). The 3700 nt long RNA present in low amount, was expected to be a tricistronic epd-pgkfbaA mRNA (Fig. 2C). Each of the three probes also detected bicistronic RNAs (epd-pgk RNA 2400 nts or pgkfbaA RNA 2700 nts) (Fig. 2A). The epd-pgk RNAs might result from transcription arrest at the putative terminator located downstream from the pgk ORF or endonucleolytic cleavage of epd-pgk-fbaA transcripts (Fig. 2C). Detection of bicistronic pgk-fbaA transcripts, in cells that were transformed with plasmid p∆P0-EPF, was a strong indication for the presence of an internal transcription-start site. In the pEPF-transformed cells, the three probes also enlightened 1100-1500 nt long RNAs, that were expected to be monocistronic. The absence of putative transcriptional terminator between the epd and pgk ORFs suggested that monocistronic epd RNAs resulted from endonucleolytic cleavage of a longer molecule. Altogether, the data suggested the presence of at least one internal transcriptionstart site, located upstream of the pgk ORF and the occurrence of endonucleolytic cleavages. By Northern blot analysis of total RNA extracted from untransformed E. coli TG1 cells, using highly labelled specific DNA probes produced by polymerase chain reaction (PCR) amplification, we verified the cotranscription of the epdpgk and fbaA ORFs in these cells. The level of tricistronic RNA was higher in untransformed cells compared with transformed cells (Fig. 2A and B). However, as in untransformed cells, bi- and monocistronic mRNAs were also present (Fig. 2B).

A promoter of the extended -10 class is located upstream of the pgk ORF

To identify the transcription promoter located upstream of the pgk ORF, E. coli TG1 cells were transformed with plasmids pEP or pAP0-EP, containing the epd and pgk ORFs with or without promoter epd P0 respectively (Fig. 1A). These two plasmids produced large amounts of PGK (Fig. 3, lanes 6 and 7) and had a stronger stability in E. coli compared with the large pEPF plasmid. Total RNAs were extracted from transformed cells grown on glucose. The 5' extremities of the PGK mRNAs were identified by two approaches: (i) reverse transcription using oligonucleotide O-PGK as the primer (Fig. 4A) and (ii) S1 mapping using the 5'-end-labelled O-GB1/O-PGK DNA (Fig. 1A), described in Experimental procedures, as the probe (Fig. 4B). The occurrence of transcription initiation at promoter epd P0 was verified by Primer extension using the O-EPD probe (Fig. 4A). Reproducible results, illustrated in Fig. 4A and B, were obtained in a series of independent experiments.

In the presence of promoter epd P0, several reverse transcriptase (RT) stops were detected between positions +967 (in the epd ORF) and +1184 (in the epd-pgk intergenic region). In contrast, no strong RT stop was detected in the 400 nt region located upstream of position +967 (not shown). Almost all the RT products obtained in high yields had a counterpart among the S1 nuclease digestion products, indicating that they corresponded to RNA extremities (Fig. 4B, lane 2). In addition, the same pgk mRNA 5' extremities were detected by RT analysis of total RNA extracted from untransformed cells (Fig. 4B, lane 1). Hence, at least, six distinct pgk mRNAs, differing by the identity of their 5' extremities, were present in the cells. They were designated as pgk mRNAs 1-6 and the location of their 5' extremities in the epd-pgk sequence is shown in Fig. 4D.

Upon deletion of promoter epd P0, all the cDNA bands had a strongly decreased intensity, except that at position +967 (Fig. 4A, compare lanes 1 and 4). We concluded that the presence of the pgk 1 mRNA was not dependent on the activity of promoter epd P0. Accordingly, a TATAGT sequence, carrying the characteristic features of the -10 elements of promoters recognized by the Eo70 RNApolymerase holoenzyme (Harley and Reynolds, 1987; Lisser and Margalit, 1993), is located 7 bps upstream of position +967 (Fig. 4D). No obvious -35 element is detected. However, as the TATAGT sequence is preceded by a TGG triplet, the corresponding promoter may belong to the special class of E. coli extended -10 promoters, which may function without -35 hexamer (Bown et al., 1997; Mitchell et al., 2003). As a confirmation, the replacement of the putative -10 hexamer segment (TATAGT, position +954 to +959) for a CACAGC sequence

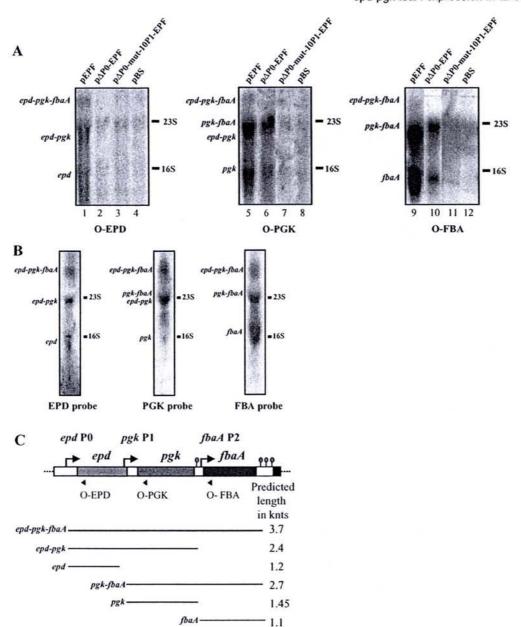


Fig. 2. A complex array of RNAs is produced by the *epd-pgk-fbaA* ORF cluster.

A. Cells transformed with the plasmids (Fig. 1A) indicated above the lanes were grown in M63 medium in the presence of glucose until an A<sub>600</sub> of 0.5. The amount of total RNA, loaded on the agarose gel was adjusted taking into account the plasmid copy number of the transformed cells. In the experiment shown, 8.6, 4.3, 6.4 and 6 μg of total RNA were loaded for plasmids pEPF, pΔP0-EPF, pΔP0-mut-10P1-EPF and pBS respectively. Conditions for Northern blot analysis are given in the *Experimental procedures*. The 5'-end-labelled probes O-EPD, O-PGK and O-FBA used are indicated below the panels. The identity of the hybridization products was deduced taking into account their estimated lengths and their differential hybridization with the probes. Lengths were estimated, by using the *E. coli* 16S and 23S rRNAs as molecular markers.

B. Untransformed *E. coli* TG1 cells were grown in M63 medium supplemented with glucose until an A<sub>600</sub> of 0.5, 30 μg of total RNA extracted from these cells were loaded in each lane of the gel and Northern blot experiments were performed with highly labelled PCR amplification products of the *epd, pgk* and *fbaA* ORF prepared as described in *Experimental procedures*. The 23S and 16S rRNA size markers are shown and the identities of the hybridization products are given on the left side of the panels.

C. The possible transcriptional organization of the epd-pgk-fbaA ORF cluster and the putative transcription and maturation products with their predicted lengths (same representation as in Fig. 1).

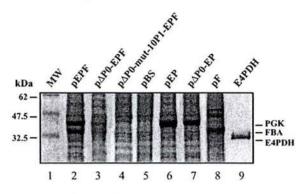


Fig. 3. Only PGK and FBA are overproduced in pEPF-transformed cells. SDS-PAGE analysis of the soluble proteins extracted from cells transformed with the various plasmids indicated above the lanes was performed as indicated in the *Experimental procedures*. Molecular weight markers (MW) were loaded in lane 1. PGK and FBA (lanes 7 and 8) were individually expressed by plasmids p\DPO-EP and pF respectively. Purified E4PDH, produced in pE-transformed *E. coli* cells, was loaded in lane 9. Lane 5 is a control experiment with an empty pBS plasmid. The same batches of transformed cells were used in Figs 2 and 3.

(plasmid pmut-10P1-EP) (Fig. 1B) and the substitution of the TG dinucleotide (positions +951 to +952) for an AC dinucleotide (plasmid pmutTGP1-EP) (Fig. 1B) led to the complete disappearance of the RT stop at position +967 (Fig. 4A, lanes 2 and 3 respectively). This promoter with an extended -10 sequence was designated as *pgk* P1 promoter.

Both PGK and FBA productions depend upon the epd P0 and pgk P1 promoters

To test whether a third transcriptional promoter is present upstream of the fbaA ORF, we transformed E. coli TG1 cells with plasmid p∆P0-mut-10P1-EPF, that carries both the epd P0 deletion and the CACAGC sequence instead of the wild-type (WT) pgk P1-10 sequence (Fig. 1). No bicistronic pgk-fbaA mRNA and no monocistronic pgk and fbaA mRNAs were detected by Northern blot analysis of total RNA (Fig. 2A). Thus, individual fbaA mRNAs are produced from transcripts initiated at the epd P0 and/or pgk P1 promoters. Accordingly, in cells transformed with plasmid p∆P0-mut-10P1-EPF, the PGK and FBA amounts (Fig. 3, lane 4) and activities (Table 1) were both reduced to levels corresponding to the expression of endogenous genes (same levels as with the empty pBS plasmid) (Fig. 3, lane 5 and Table 1). Hence, in spite of the putative rho-independent terminator downstream from the pgk ORF (Fig. 1A) (Alefounder and Perham, 1989), the fbaA ORF is predominantly cotranscribed with the pgk ORF. According to the PGK and FBA activities measured for pEPF and pΔP0-EPF-transformed cells (Table 1), the epd P0 and pgk P1 promoters in plasmid constructs, likely have similar contributions to the PGK and FBA productions when glucose is present in the medium (Table 1).

Reverse transcriptase analysis of the PGK mRNAs in TG1 cells grown in M63 medium supplemented with glucose or pyruvate (Fig. 4C) revealed no significant variation of the steady-state level of the pgk P1 transcripts. Thus, promoter pgk P1 is not activated in the presence of glucose in the medium. Note that the low activity of promoter epd P0 in the presence of pyruvate was accompanied by an almost complete disappearance of the pgk mRNAs 2-6 (Fig. 4C). This suggest that pgk mRNAs 2-6 are generated by endonucleolytic cleavages of the epd-pgk and/or epd-pgk-fbaA transcript(s), but not of transcripts initiated at promoter pgk P1. In support to the occurrence of endonucleolytic cleavages, none of the pgk mRNAs 2-6 5' extremities correspond to a transcription-start site according to the DNA sequence. Whereas cleavage sites 5 and 6 are located within the epd-pgk inter-ORF region, sites 2, 3 and 4 are located in the 3'-terminal part of the epd ORF (Fig. 4D). These three cleavages are, thus, expected to abolish E4PDH production. Remarkably also, whereas sites 2-6 are present in transcripts initiated at promoter pgk P1, these transcripts are inefficiently cleaved compared with the epd P0 transcripts.

The pgk-fbaA interORF region is cleaved on both sides of the pgk putative terminator

The 5' extremities of individual *fbaA* mRNAs were identified by Primer extension analysis using RNA extracted from pEPF- or pΔP0-EPF-transformed cells and the O-FBA probe (Fig. 5). None of the detected RT stops corresponded to the previously predicted transcription-start site (Salgado *et al.*, 2004) (not shown). In contrast, successive RT stops were found in the series of adenosine residues located from position +2364 to +2369 (*fbaA* 5 series of stops) and at four other positions denoted *fbaA* stops 1–4

**Table 1.** The PGK- and FBA-specific activities<sup>a</sup> of *E. coli* TG1-transformed cells<sup>b</sup>.

| PGK°                                     | FBA°   |
|--|--|
| (µmol NADH/                              | (µmol NADH/  |
| min/A <sub>280</sub> /A <sub>600</sub> ) | min/A <sub>280</sub> /A <sub>600</sub> )   |
| 2.5                                      | 0.031  |
| $1.05 \pm 0.11$                          | 0.017 ± 0.002  |
| $0.14 \pm 0.09$                          | $0.0015 \pm 0.001$   |
| $0.10 \pm 0.09$                          | 0.001 ± 0.001  |
|  | ( $\mu$ mol NADH/<br>min/A <sub>280</sub> /A <sub>600</sub> )<br>2.5<br>1.05 ± 0.11<br>0.14 ± 0.09 |

a. Specific activities were measured on cells harvested in the midlog phase of growth as described in the Experimental procedures.
 b. E. coli TG1 cells transformed with plasmids pEPF, pΔP0-EPF, pΔP0-mut-10P1-EPF or pBS were grown in M63 minimal medium supplemented with glucose and Casamino Acids.

c. The reported values take into account the variations of plasmid copy number as explained in the Experimental procedures. They are means and standard deviations of three independent experiments.

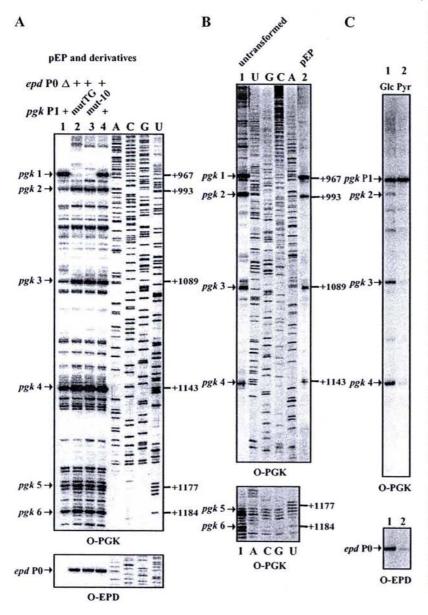
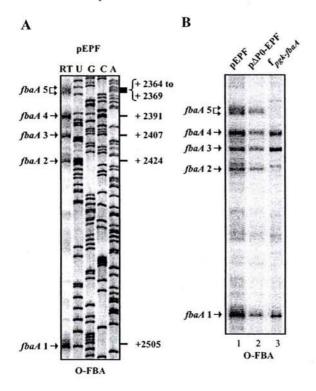
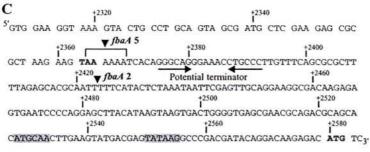


Fig. 4. Identification of six distinct pgk mRNA 5' extremities and a pgk internal promoter. A. Primer extension analysis of total RNAs (5 μg) extracted from pΔP0-EP (lane 1), pmutTGP1-EP (lane 2), pmut-10P1-EP (lane 3), or pEP (lane 4) E. coli TG1-transformed cells using oligonucleotide O-PGK or O-EPD as the primers. Growth was performed in M63 medium supplemented with glucose. The sequencing ladders (lanes A, C, G and U) were obtained with the O-PGK or O-EPD oligonucleotide. The oligonucleotides primers used are indicated below the autoradiograms. The 5' extremities of pgk mRNAs (pgk 1-6) and their positions as referred to the epd P0 start site are indicated on the left and right sides of the autoradiogram respectively.

B. S1 mapping assay on total RNA extracted from untransformed (30 μg) (lane 1) or pEP-transformed (5 μg) (lane 2) *E. coli* TG1 cells grown in M63 medium supplemented with glucose. The 5'-end-labelled O-GB1/O-PGK DNA fragment (Fig. 1A) was used as the probe. Lanes U, G, C and A are sequencing ladders obtained with oligonucleotide O-PGK. The identified *pgk* mRNA 5' extremities and their positions as referred to the *epd* P0 start site are indicated on the left and right sides of the autoradiogram respectively.

C. Effect of the carbon source identity on the pgk mRNA composition. Untransformed E. coli TG1 cells were grown in M63 medium supplemented with glucose (lane 1) or pyruvate (lane 2). Primer extension analyses were performed on 20 µg of total RNA using oligonucleotide O-PGK or O-EPD as indicated below the autoradiograms. The pgk mRNA numbering is as in A. D. Schematic representation of the data shown in A and B on the nucleotide sequence of the epd-pgk region from position +874 to +1243. The extended -10 element of promoter pak P1 is boxed and the corresponding transcription start site is shown by a bent arrow. The 5' extremities of the pgk mRNAs 2-6, corresponding to in vivo RNase E cleavages, are marked by triangles. The epd and pgk ORFs are represented by triplets of nucleotides. The epd stop codon and pgk initiation codon are in bold. The sequence complementary to oligonucleotide O-PGK is underlined.





(Fig. 5). As three of these stops (fbaA 1, 3 and 4) were also detected by RT analysis of an fpok-tbaA transcript (Fig. 1A) produced in vitro (position +2265 to +2626) (Fig. 5B, lane 3), they likely resulted from RT pauses. Only, the fbaA 2 stop and the fbaA 5 series of stops corresponded to fbaA mRNA 5' extremities. In accord with the above data, no putative transcriptional promoter was detected in the corresponding upstream DNA region (Fig. 5C). Interestingly, in the pgk-fbaA inter-ORF region, the fbaA 5 series of cleavage sites overlaps the pgk stop-codon and is located upstream of the potential pak transcription terminator. Site fbaA 2 is located downstream from this motif (Fig. 5C). Identical patterns of cleavages of the pgk-fbaA interORF region were observed in the presence or absence of promoter epd P0 (Fig. 5B, lane 1 and 2), suggesting that similar cleavages of the pgk-fbaA interORF region take place in the epd-pgk-fbaA and pgk-fbaA transcripts.

Fig. 5. Identification of the *fbaA* mRNA 5' extremities.

A. The 5' extremities of *fbaA* mRNAs from

pEPF-transformed E. coli TG1 cells were identified by extension of the O-FBA probe with RT

(lane RT). Lanes U, G, C and A correspond to

the sequencing ladder. Several cleavages were

detected in the run of A residues from position +2364 to +2369 (fbaA 5 series of sites). Positions of the RT stops in the nucleotide sequence downstream from the epd P0 transcription start-site are given on the right side of the autoradiogram. B. RT analysis with oligonucleotide O-FBA of total RNA (5 µg) extracted from pEPF (lane 1) and pΔP0-EPF (lane 2) transformed TG1 cells. Growth was in M63 medium supplemented with glucose. Lane 3 corresponds to the reverse transcription of the in vitro transcribed feet that RNA (Fig. 1A) (30 ng) using the O-FBA primer. The cDNA bands are numbered as in A. C. Schematic representation of the data shown in A and B on the nucleotide sequence of the pgk-fbaA interORF region. The series of fbaA 5 cleavage sites are indicated by a horizontal line surmounted by a triangle and the fbaA 2 cleavage site is shown by a triangle. The pgk and fbaA ORFs are represented by triplets of nucleotides and the pgk stop codon and fbaA initiation codon are in bold. The two horizontal inverted arrows indicate the putative pgk T1 transcription terminator. The -10 and -35 hexamers of the previously proposed fbaA tran-

scriptional promoter (Salgado et al., 2004) are

The epd ORF, the epd-pgk and pgk-fbaA inter-ORF regions are cleaved by RNase E

boxed in grey.

Three endonucleases were found to be involved in mRNA degradation and maturation in *E. coli*: RNase III, RNase E (for review: Grunberg-Manago, 1999) and the more recently identified RNase G (for review: Wachi and Nagai, 2001; Lee *et al.*, 2002). The *E. coli* isogenic strains IBPC602 and IBPC600 (Table 2) were used to test for a possible implication of RNase III in cleavage of the *epd-pgk*, *pgk-fbaA* and *epd-pgk-fbaA* transcripts. The *rnc-105* mutation in strain IBPC600 abolishes RNase III activity. The N3433 strain and its isogenic strain IBPC935 containing a disrupted RNase G gene (Hajnsdorf *et al.* unpubl. data) (Table 2) were used to test for the possible implication of RNase G. The four strains were transformed with plasmid pEPF and total RNAs extracted from cells grown

Table 2. The E. coli bacterial strains used in this study.

| Bacterial strains | Relevant genotypes  | Source or origin           |
|-------------------|---|----------------------------|
| TG1               | supE hsdΔ5 thi Δ(lac proAB), F'(traD36 proAB* lacF lacZΔM15)                          | Stratagene                 |
| IBPC5321          | thi-1 argE3 ΔlacX74 mlt-1 xyl-5 tsx-29 rpsL argG6 his-4                               | Plumbridge and Soll (1989) |
| IBPC6311          | thi-1 argE3 \(\Delta\)lacX74 mlt-1 xyl-5 tsx-29 rpsL argG6 his-4 recA56 srl-300::Tn10 | Plumbridge and Soll (1989) |
| MG1693            | thyA715   | Arraiano et al. (1988)     |
| SK5665            | thyA715 ams-1   | Arraiano et al. (1988)     |
| IBPC602           | thi-1 argE3 ΔlacX74 mtl-1 xyl-5 rpsL argG6 his-4 nadA51::Tn10                         | P. Régnier                 |
| IBPC600           | thi-1 argE3 ΔlacX74 mtl-1 xyl-5 rpsL argG6 his-4 nadA51::Tn10, rnc105                 | P. Régnier                 |
| N3433             | HfrH lacZ43 λ⁻ relA1 spoT thi-1   | D. Apirion                 |
| IBPC935           | HfrH lacZ43 λ- relA1 spoT thi-1 rng::CmR  | E. Hainsdorf               |

on Luria–Bertani (LB) were studied by Primer extension analysis. The patterns of cDNAs obtained with the O-PGK and O-FBA primers were identical for the four strains (data not shown). Thus, RNases III and G are not responsible for cleavages of the *epd-pgk* and *pgk-fbaA* inter-ORF regions.

Then, by using the MG1693 and SK5665 isogenic strains (Arraiano et al., 1988) (Table 2), we tested the possible implication of RNase E. Due to the ams-1 mutation, strain SK5665 (RNase Ets) produces a thermosensitive RNase E. As RNase E is essential for growth of E. coli, the two transformed strains were grown at 30°C (permissive temperature) until mid-log phase. Then, half of each culture was subjected to a temperature shift at 42°C. RT analysis of the pgk mRNAs with the O-PGK probe showed a disappearance of the pgk mRNAs 2-6 after incubation of the RNase Ets strain for 15 min at 42°C (Fig. 6A). Such disappearance was not observed for the WT strain. By using primer O-EPD, we verified that this disappearance was not due to the absence of transcription initiation at promoter epd P0 (not shown). Based on these observations, we concluded that RNase E is responsible for the cleavages that generate the pgk mRNAs 2-6. We confirmed this proposal by Northern blot analysis of total RNA from the RNase Ets strain transformed with the pEP plasmid, that produces an epd-pgk transcript from promoter epd P0. Indeed, an epd-pgk transcript appeared at non-permissive temperature. It was absent at permissive temperature and in extract of the WT strain (Fig. 6B).

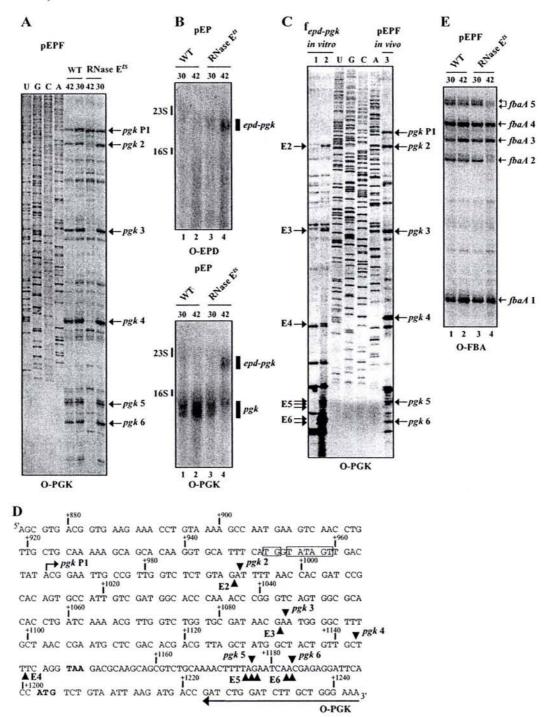
Reverse transcriptase analysis of the *fbaA* mRNAs from the WT and RNase E<sup>ts</sup> strain with the O-FBA probe (Fig. 6E) showed that cleavage in the series of A residue from position +2364 to +2369 (*fbaA* 5 series of sites) is temperature sensitive in the WT strain. However, after the temperature shift, a stronger decrease of the yield of cleavages in the +2364 to +2369 segment and at site *fbaA* 2 was observed in the RNase E<sup>ts</sup> strain compared with the WT strain. We, thus, concluded that RNase E is also responsible for the cleavages detected in the *pgk-fbaA* interORF region.

RNase E cleaves similarly the epd-pgk region in vitro and in vivo

To compare in vitro and in vivo cleavages of the epd-pgk region by RNase E, we submitted an fepd-pak RNA (positions +701 to +1500) (Fig. 1A), produced by in vitro transcription, to the action of a degradosome fraction of the E. coli RNase Ets strain. Indeed, Braun et al. showed that degradosome fractions carrying a mutation in the S1 like binding domain of RNase E have in vitro RNA cleavage properties similar to those of WT degradosomes (Braun et al., 1996). In the fepd-pgk RNA, the cleaved region was flanked by a 292 nt long upstream and a 316 nt long downstream fragment. The possibility to abolish the RNase E activity of the degradosome fraction by a 10 min incubation at 45°C (Arraiano et al., 1988) was used as an internal control. For comparison, the reverse transcripts obtained with the in vitro maturation products were compared with those of total RNA from the WT MG1693 strain transformed with plasmid pEPF. Some of the in vitro cleavage sites were shifted by 1 or 2 nts compared with the in vivo sites and some of the single cleavages found in vivo were replaced by two or three cleavages in succession in vitro (Fig. 6C and D). Nevertheless, a striking similarity of the in vivo and in vitro cleavages was observed. The minor difference observed may be due to the fact that we used a degradosome fraction containing an RNase E mutated in the S1 like RNA-binding domain. The data suggested a very low protection of the epd ORF by translating ribosomes and thus, a low translation efficiency of the epd ORF in vivo.

### Cleavage of the epd ORF is probably due to its low translation efficiency

As translation efficiency largely depends upon the efficiency of the translation initiation region (TIR), we built epd—lacZ transcriptional and translational fusions to test the efficiency of the epd TIR. In plasmid pVBGF2 (epd—lacZ translational fusion), a  $\beta$ -galactosidase truncated of its six N-terminal amino acids, fused to the 18 N-terminal



amino acids of E4PDH, was expressed under the control of the  $\ensuremath{\textit{epd}}$  P0 promoter and  $\ensuremath{\textit{epd}}$  TIR. In plasmid pVBOF2 ( $\ensuremath{\textit{epd}}$ –lacZ transcriptional fusion), the WT  $\beta$ -galactosidase was expressed under the control of promoter  $\ensuremath{\textit{epd}}$  P0 and the  $\ensuremath{\textit{lacZ}}$  TIR (Fig. 1A). The cellular  $\beta$ -galactosidase-specific activity was 12-fold higher for cells transformed with the transcriptional fusion compared with cells

expressing the translational fusion (Table 3). By restoration of an entire  $\beta$ -galactosidase ORF in plasmid pVBGF2, only a limited increase of the  $\beta$ -galactosidase activity was observed (factor of 2) (data not shown). Hence, these results were an indication for a limited efficiency of the epd TIR. To confirm this hypothesis, we used the  $\lambda$ RS45 phage to introduce the transcriptional and translational

Fig. 6. RNase E cleaves both the epd-pgk and pgk-fbaA regions of the epd-pgk-fbaA transcripts.

A. The epd-pgk and epd-pgk-fbaA mRNAs are cleaved by RNase E in vivo. The MG1693 (WT) and SK5665 (RNase E<sup>15</sup>) E. coli strains transformed with plasmid pEPF were grown at 30°C in LB medium, until an Acco of 0.5. One half of each culture was subjected to a temperature shift at 42°C, the other half was maintained at 30°C. Total RNAs extracted from the different batches of cells were analysed by extension of primer O-PGK with RT. Lanes U, G, C and A correspond to the sequencing ladder.

B. E. coli MG1693 (WT) and SK5665 (RNase Etc) cells transformed with plasmid pEP were grown at 30°C in LB medium until an A<sub>600</sub> of 0.5. Then one half of each culture was incubated for 15 min at 42°C (lanes 2 and 4) and the other half was maintained at 30°C (lanes 1 and 3). Equal amounts of total RNA (5 µg) extracted from each cell batches were subjected to Northern blot analysis using the O-EPD and O-PGK oligonucle-

otides as the probes. Transcripts were identified by reference to 16S and 23S rRNAs.

C. The epd-pgk region is cleaved by a degradosome fraction in vitro. The in vitro transcribed fepd-pgk RNA (Fig. 1A) was incubated 10 min with a degradosome fraction, prepared from the E. coli RNase Ets strain. This fraction was preincubated either 10 min at 45°C (lane 1) or 10 min at 30°C (lane 2). The reaction products were analysed by extension of primer O-PGK with RT. For comparison, an RT analysis was performed on total RNA (5 µg) extracted from pEPF-transformed E. coli cells (lane 3). Lanes U, G, C and A correspond to the sequencing ladder. The RNase E cleavage sites obtained in vitro are numbered from E2 to E6.

D. The in vivo and in vitro RNase E cleavage sites detected in the epd-pgk RNA region are shown on the RNA sequence. Triangles above the

sequence correspond to the in vivo cleavage sites, triangle below the sequence to the in vitro cleavage sites.

E. The pgk-fbaA interORF region is cleaved by RNase E in vivo. As in A, the MG1693 (WT) (lanes 1 and 2) and SK5665 (RNase Ett) (lanes 3 and 4) E. coli cells transformed with plasmid pEPF were grown at 30°C in LB medium. One half of each culture was incubated for 15 min at 42°C (lanes 2 and 4), and the other half was maintained at 30°C (lanes 1 and 3). Total RNA extracted from each cell batches (5 μg) were analysed by extension of primer O-FBA with RT. The resulting fbaA cDNAs are numbered as in Fig. 5.

fusions as unique copies in the E. coli TG1 chromosome. The cellular β-galactosidase-specific activity of the lysogenic strain containing the translational fusion was 150-fold lower, compared with that of the lysogenic strain containing the transcriptional fusion (Table 3). This brought strong additional support to the idea that the epd TIR has a low efficiency.

The small distance between the RBS and initiator AUG probably limits the epd TIR efficiency

The epd TIR has two peculiar features: first, a long stemloop structure can be formed immediately upstream of the Ribosome Binding Sequence (RBS) (Fig. 1B) and, second, the distance between the RBS and AUG initiation codon is unusually short (aligned distance of 5, instead of the 8-10 nts optimal distance) (Ma et al., 2002) (Fig. 1B). To test whether one of these two features may be involved in translation limitation, two variant epd TIRs were produced. In variant 35, the 5' strand of the stem (29 nts) was substituted for a GTCGAC sequence, so that formation of the long stem-loop structure was abolished. In variant 408, the WT UCU sequence, that links the RBS to the AUG, was replaced by a 7 nt sequence (AAAUAAA) (aligned distance of the mutated RBS of 9 nts) (Fig. 1B). Each of these two mutations was introduced in the pVBGF2 and pVBOF2 plasmids, described above, and were integrated as unique copies in the E. coli TG1 chromosome. Measurement of the β-galactosidase-specific activities of the transformed cells and monolysogenic cells (Table 3) showed that both TIR mutations had no marked effect on the expression of the transcriptional fusions. Mutation 35 also had a very limited effect on the translational fusion. In contrast, with an optimal spacing between the RBS and initiation codon (variant 408), the β-galactosidase activity was increased by a factor of about 5 for the transformed cells and 15 for the monolysogenic cells. Hence, the small distance between the RBS sequence

Table 3. The β-galactosidase-specific activities produced from E. coli epd-lacZ fusions in E. coli TG1-transformed cells or monolysogenic cells\*.

|                              |                      | β-Galactosidase activit | y (Miller Unit per A <sub>600</sub> ) <sup>b,c</sup> |                        |
|------------------------------|----------------------|-------------------------|--|------------------------|
|                              | Pla                  | asmid                   | Monol  | ysogenic               |
| Tested TIR                   | Translational fusion | Transcriptional fusion  | Translational fusion                                 | Transcriptional fusion |
| E. coli epd WT               | 800 ± 150            | 9150 ± 2000             | 8 ± 3  | 1230 ± 96              |
| E. coli epd 35 <sup>d</sup>  | 1400 ± 400           | 9300 ± 800              | 11 ± 3   | 1100 ± 66              |
| E. coli epd 408 <sup>d</sup> | $4300 \pm 500$       | 8600 ± 1700             | 110 ± 17   | $1380 \pm 110$         |

a. E. coli TG1 cells transformed with derivatives of plasmids pRS414 (translational fusion) or pRS415 (transcriptional fusion) and monolysogenic E. coli TG1 cells containing unique copies of the translational or transcriptional fusions were grown in M63 minimal medium supplemented with glucose and Casa Aminoacids until mid-log phase of growth.

c. The reported values are means and standard deviations of three independent experiments.

b. β-Galactosidase activities were measured using the method of Miller.

d. The E. coli epdmut35-lacZ and epdmut408-lacZ fusions contained the mut35 and mut408 epd TIR sequence respectively.

<sup>© 2005</sup> Blackwell Publishing Ltd, Molecular Microbiology, 57, 1265-1287

Table 4. E4PDH- and PGK-specific activities of E. coli TG1-transformed cells.

| Plasmid      | Identity of epd ORF | E4PDH <sup>c</sup> (μmol<br>NADH min <sup>-1</sup> A <sub>280</sub> <sup>-1</sup> /A <sub>600</sub> ) | PGK <sup>c</sup> (μmol<br>NADH min <sup>-1</sup> A <sub>280</sub> <sup>-1</sup> /A <sub>600</sub> ) |
|--------------|---------------------|---|---|
| pL-EP        | E. coli epd WT      | 0.47 ± 0.15   | 203 ± 10  |
| pmut35-L-EP  | E. coli epd 35      | $0.49 \pm 0.03$   | 170 ± 16  |
| pmut408-L-EP | E. coli epd 408     | $0.98 \pm 0.03$   | 190 ± 11  |

- a. E. coli TG1 cells transformed with plasmids pL-EP, pmut35-L-EP or pmut408-L-EP were grown in M63 minimal medium supplemented with plucose and Casamino Acids
- b. Specific activities were measured on cells harvested in the mid-log phase of growth.
- c. The reported values are means and standard deviations of three independent experiments.

and the AUG initiation codon probably contributes to the low activity of the epd TIR.

Mutations 35 and 408 were also introduced in plasmid pL-EP containing the *epd-pgk* gene tandem (Fig. 1A and Table 4). Alteration of the stem-loop structure had no marked effect on both the PGK and E4PDH activities. In contrast, increasing the distance between the RBS sequence and initiator AUG led to the doubling of the E4PDH activity, whereas the PGK-specific activity was unmodified. This reinforced the idea that the *epd* TIR limits *epd* translation efficiency.

The secondary structures of the upstream epd and pgk-fbaA RNase E target sites are conserved in enterobacteria

RNase E cleaves preferentially RNAs that contain a (G,A)(C,A)N(G)(G,U,A)/(A,U)(C,U)N(C,A)(C,A) sequence (Kaberdin, 2003), and the presence of stem-loop structures flanking this sequence reinforces RNase E efficiency (Kaberdin *et al.*, 2000). By using chemical and enzymatic probes of the RNA secondary structure, we tested whether the *epd-pgk-fbaA* RNase E target sites display

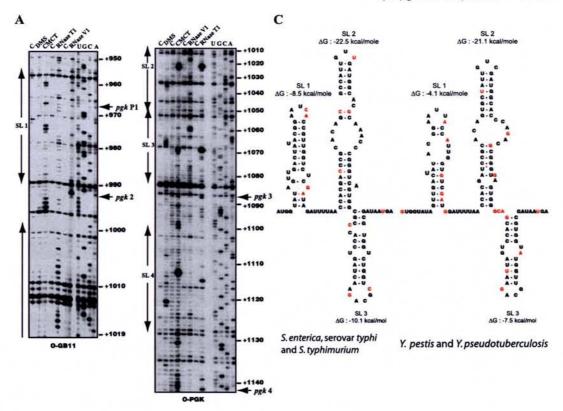
such structural characteristic features. T1 and T2 RNases were used to probe single-stranded regions (Mougin et al., 1996; Jacquenet et al., 2001), V1 RNase was used to probe helical and stacked regions (Mougin et al., 1996). Two chemical probes for single-stranded regions were also used: DMS acting on A and C residues and CMCT acting on U and G residues (Mougin et al., 1996; Jacquenet et al., 2001). In vitro transcripts containing the epd-pgk sequence from position +701 to +1500 (Fig. 7A and Fig. S1) and the pgk-fbaA sequence from position +2265 to +2626 (Fig. 8A and Fig. S2), respectively, were subjected to chemical modifications and enzymatic digestions in conditions preserving the RNA secondary structure (Mougin et al., 1996; Jacquenet et al., 2001). The positions of modifications and cleavages were identified by primer extension analysis using the O-GB11, O-PGK, O-GB12, O-GB13 and O-FBA oligonucleotides (Fig. 1A and Table S1). The identified cleavage sites were used to settle constraints in a Mfold search for RNA secondary structure. A unique structure in accord with all the experimental data was established for the epd-pgk region extending from position +954 to +1221 (Fig. 7B). In contrast, two alternatives structures were compatible with the

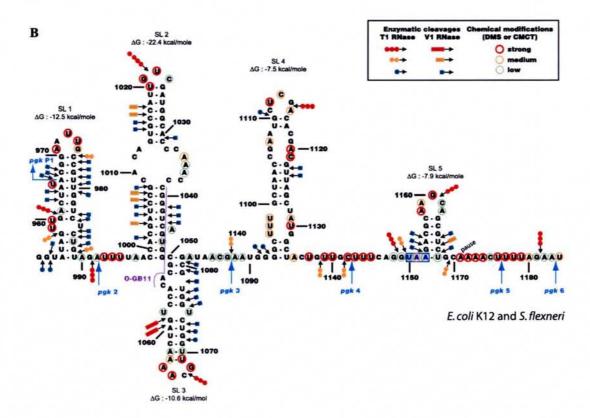
Fig. 7. Secondary structure analysis of the epd-pgk RNA region from position +950 to +1200.

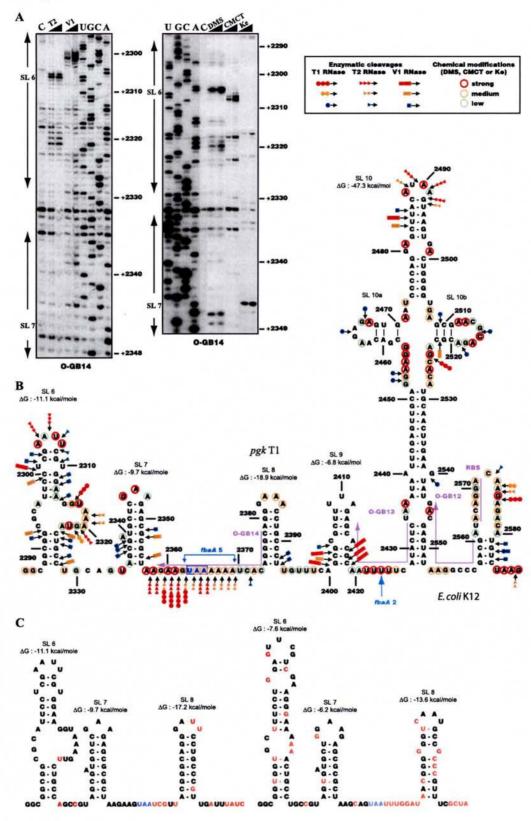
A. The experimental analysis of the secondary structure of the *epd-pgk* region from position +950 to +1140 using chemical and enzymatic probes is shown. The f<sub>epd-pok</sub> RNA produced by *in vitro* transcription was subjected to chemical modifications with DMS (1 μl of a 25% (v/v) DMS-ethanol) and CMCT (1680 μg) and enzymatic cleavages by RNases T1 (2.5 U) and V1 (0.1 U), using conditions described in the *Experimental procedures*. Positions of cleavages and modifications were identified by extension of primers O-PGK and O-GB11 with RT, followed by electrophoresis on sequencing gels in parallel with sequencing ladders obtained with the same oligonucleotides. The identity of the primer used is indicated below each autoradiograms. Positions of nucleotides in the *epd-pgk* sequence and of RNase E cleavage sites are given on the right sides of the autoradiograms. Positions of the identified SL1, SL2, SL3 and SL4 stem-loop structures (B) are indicated on the left sides of the autoradiograms. The identity of the chemical and enzymatic probes used is given above the lanes. Analyses of the region from position +1140 to +1203 are shown in the *Supplementary material*.

B. Schematic representation of the data shown in A and 1B in the *Supplementary material* on the secondary structure model proposed for the *epd-pgk* region extending from position +950 to +1185. The results shown in A and in the *Supplementary material* were used to propose a secondary structure model using the Mfold software. Cleavages of phosphodiester bonds are indicated by arrows surmounted of circles (RNase T1) or squares (RNase V1). As indicated in the inset, the numbers and colours of circles and squares indicate the yields of cleavages. Residues modified by DMS or CMCT are circled and the colour of circles indicate the yields of modifications (see the inset). The free energies at 37°C in 1 M NaCl of SL1 to SL5 stem-loop structures were estimated using the Mfold software. The positions of the *pgk* 2, 3, 4, 5 and 6 RNase E cleavage sites are marked by blue vertical arrows. The *pgk* P1 transcription start site is shown by a bent arrow. The stop codon of the *epd* ORF is boxed in blue.

C. Conservation of the possibility to form the SL1, SL2 and SL3 stem-loop structures in the *epd* ORFs of *Salmonella* and *Yersinia* species. The EMBL database numbers of the sequences used are: NC\_000913 (*E. coli* K12), NC\_004631 (*Salmonella enterica* ssp. *enterica* serovar *Typhi* Ty2), NC\_004337 (*Shigella flexneri* 2a str. 301) NC\_003197 (*Salmonella typhimurium* LT2), NC\_004088 (*Yersinia pestis* KIM) and NC\_006155 (*Yersinia pseudotuberculosis* IP 32953). Base substitutions as compared with *E. coli* are indicated by red nucleotides. The free energies were estimated using the Mfold software.







S. enterica, serovar typhi and S. typhimurium LT2

Y. pestis and Y. pseudotuberculosis

#### epd-pgk-fbaA expression in E. coli 1279

Fig. 8. Secondary structure analysis of the pgk-fbaA mRNA region extending from position +2300 to +2600.

A. The experimental analysis of the secondary structure of the pgk-fbaA region from position +2300 to +2348 using chemical and enzymatic probes is shown. The analysis of the remaining part of the RNA region is given in the Supplementary material. The fpot. to 4265 to +2626) (Fig. 1) produced by in vitro transcription was subjected to enzymatic digestion (T1, T2 and V1 RNases) and chemical modifications [DMS, CMCT, kethoxal (Ke)] using for each probe, the two reagent concentrations given in the Experimental procedures. The positions of cleavages and modifications were identified by extension of primer O-GB14 by RT, followed by electrophoresis on sequencing gels, in parallel with sequencing ladders obtained with the same oligonucleotides. The increasing probe concentrations are represented by triangles. Lanes U, G, C, A correspond to sequencing ladders. Positions of nucleotides in the pgk-fbaA sequence are given on the right sides of the autoradiograms. The identified stemloop structures (B) are indicated on the left sides of the autoradiograms. The pgk stop codon and fbaA start codon are also shown on the right side. B. Schematic representation of the probing data, illustrated in A and shown in the Supplementary material on the secondary structure model proposed for the pgk-fbaA region +2285 to +2600. The same rules as in Fig. 7(B) are utilized for representation of chemical modifications and phosphodiester bonds cleavages (see the insert). Cleavage by the T2 RNase that was not used in the experiment illustrated in Fig. 7(B) are shown by arrows surmounted by triangles. The fbaA 5 series of sites and fbaA 2 site are shown by blue arrows. The pgk stop codon is boxed in blue. C. Conservation of the possibility to form the stem-loop structure SL6, 7 and 8 in Salmonella and Yersinia species. The EMBL database numbers of nucleotide sequences used are the following: NC\_000913 (E. coli K12), NC\_004631 (Salmonella enterica sep. enterica serovar Typhi Ty2), NC\_003197 (Salmonella typhimurium LT2), NC\_004088 (Yersinia pestis KIM) and NC\_006155 (Yersinia pseudotuberculosis IP 32953). Base substitutions as compared with E. coli are indicated by red nucleotides. The free energies were estimated using the Mfold software.

experimental data obtained for the pgk-fbaA region extending from position +2285 to +2400. To choose among them, we performed a phylogenetic analysis in γproteobacteria. Indeed, functionally important secondary structures in RNAs were found to be conserved throughout evolution by the occurrence of compensatory base substitutions (Branlant et al., 1981; Mougin et al., 1996). First, we looked for species that contain an epd ORF upstream of the pgk ORF. To this end, we took into consideration the described signature of E4PDH, namely the presence of E32, M179, Y187, C311 residues and the VDTKL sequence around residue T208 (Boschi-Muller et al., 1997; Fillinger et al., 2000). An epd gene located upstream of the pgk gene was found in Erwinia carotovora, S. typhimurium, Salmonella enterica serovar typhi, Shigella flexneri, Y. pestis and Yersinia pseudotuberculosis. We aligned the corresponding epd-pgk-fbaA sequences with that of E. coli. Then, for each species, we looked for the possibility to form the secondary structures established for E. coli. As shown in Fig. 7B and C and Fig. S1B, the three stem-loop structures found in the E. coli +954 to +1081 region can be formed with the Erwinia, Salmonella, Shigella and Yersinia RNAs. Taking into account the large number of base substitutions, this observation is a strong argument in favour of a biological significance of the proposed structure. In this structure, the RNase E cleavage site pgk 2 is present in a singlestranded sequence that fits perfectly to the consensus sequence of the E. coli RNase E cleavage sites. This sequence is flanked by two conserved stable stem-loop structures (-12.5 and 22.4 kcal mol-1) (Fig. 7B and C and Fig. S1B). The nucleotide sequence and secondary structure of the epd-pgk region containing sites pgk 3, 4, 5 and 6 are not phylogenetically conserved. Compared with site pgk 2, the hairpins flanking sites pgk 3, 4 and 5 are less stable and are more distant from the cleaved bond. Thus, only the upstream pgk 2 site has the hallmarks of the efficiently cleaved RNase E sites. Only one of the two secondary structures possible for the +2288 to +2393 pak-

fbaA segment of *E. coli* can be formed with the corresponding RNA region of *Salmonella* and *Yersinia* species (Fig. 8C). In this structure, the *fbaA* 5 series of sites is located in a single-stranded sequence that fits perfectly to the consensus sequence of the *E. coli* RNase E sites. This segment is flanked by two stable hairpin structures (–9.7 and –18.9 kcal mol<sup>-1</sup>) (Fig. 8B and C). Site *fbaA* 2 is located in a variable region and as will be discussed below, this site does not fit all the criteria needed to be efficiently cleaved. Hence, in the two regions of the *epd-pgk-fbaA* transcripts, which are cleaved by RNase E, the upstream cleavage site is optimized and phylogenetically conserved, the downstream sites are located in variable regions and are not optimized.

#### Discussion

The E. coli pgk and fbaA ORFs are cotranscriptionally expressed

Here, we show that the pgk and fbaA genes, which are among the more highly expressed genes in E. coli (Peng and Shimizu, 2003), are cotranscriptionally expressed by utilization of two transcriptional promoters, epd P0 and pgk P1. The newly identified pgk P1 promoter belongs to the class of extended -10 promoters without characteristic -35 sequence (Keilty and Rosenberg, 1987). Promoters of this class are more easily accommodated in the ORF sequences and are expected to ensure the basal expression of genes independently of the growth conditions. Utilization of promoter of the extended -10 class for expression of genes coding for glycolytic enzymes is rather frequent: in addition to the pgk and fbaA genes, the gapA (GAPDH) (Thouvenot et al., 2004), pfkB (phosphofructokinase) (Mitchell et al., 2003), ppsA (Phosphoenolpyruvate synthase) (Negre et al., 1998) and pckA (phosphoenolpyruvate carboxykinase) (Prost and Cozzone, 1999) genes are also under the control of promoters with an extended -10 box. Our data show that promoter epd P0 is required to get a strong production of both PGK and FBA in the presence of glucose in the medium. Through its regulation by the Cra protein, promoter epd P0 is also expected to be modulated by the metabolic flux during both glycolysis and gluconeogenesis (Ramseier et al., 1995; Saier and Ramseier, 1996). These important roles of promoter epd P0 probably explain its conservation upstream of the epd ORF, which is poorly expressed.

Several parameters contribute to the low epd expression

Our data show that the low epd expression is due to both a low translational efficiency and RNase E cleavages. Translation efficiency in bacteria depends upon codon usage and efficiency of the TIR (Ma et al., 2002). Whereas the pgk and fbaA ORFs only contain codons corresponding to highly abundant tRNAs in E. coli (Karlin et al., 2001), the epd ORF contains nine codons corresponding to minor tRNAs. They are spread all over the ORF. However, an arginine AGG codon, corresponding to a very rare tRNA in E. coli is located immediately upstream of the TAA terminator codon (Fig. 4D). This is a very unfavourable position, recently found to lead to mRNA cleavage by a yet unidentified mechanism (Sunohara et al., 2004). As in addition, our data reveal a very low efficiency of the epd TIR, both codon usage and the poor TIR efficiency should contribute to the low translational efficiency of the epd ORF. The short distance between the RBS and AUG initiation codon is in part responsible for the poor TIR efficiency and this short distance is strongly conserved in epd ORF (Fig. 9). Increasing this distance from 4 to 8 nts in the epd-pgk context increased E4PDH production by a factor of 2 (Table 4). As expected based on the higher stability of the lacZ ORF compared with the epd ORF, the same mutation increased the translation efficiency of the epd-lacZ translational fusion by a factor of 10 in

ARNr 16S anti-RBS 3'AUUCCUCCA 5' 5'UAAGGAGGU 3' RBS (5/5) 5 'ACAAGGAACGCCAUC E. carotovora E. coli K12 5 'UGCAGGAGAUCUAUG (2/5)S. enterica Typhi 5 'UGCAGGAGACCUAUC (4/5) 5 'UGCAGGAGACCUAUG (4/5) S. typhimurium 5 'UGCAGGAGAUCUAUG S. flexneri (2/5)Y. pestis 5 GCAGGGAGUCACAUC (4/5) Y. pseudotuberculosis 5 'GCAGGGAGUCACAU (4/5)

Fig. 9. The short distance between the RBS and AUG initiation codon of the epd TIR is conserved in Erwinia, Escherichia, Salmonella and Yersinia genera. The EMBL database numbers of the sequences used are the same as in Fig. 7. Residues complementary to the anti-RBS sequence of 16S rRNA are boxed in grey. The initiator AUG is shown in dark grey. The distance and aligned distance between the RBS and initiator AUG, as defined by (Ma et al., 2002), are given on the right side of the panel respectively.

monolysogenic cells (Table 3). We noticed a strong difference of  $\beta$ -galactosidase activities between monolysogenic cells containing the transcriptional or translational fusions, compared with transformed cells (factor of 150, against a factor of 10 respectively). This difference may be due to a translational repressor, that is titrated in transformed cells. Further experiments are needed to test this hypothesis. The very low level of *epd* translation is also evidenced by the finding of nearly identical patterns of RNase E cleavage *in vivo* and *in vitro* (Fig. 6C). Due to their limited number, the ribosome translating the *epd* ORF do not protect the transcript against RNase E *in vivo*.

RNase E plays a critical role in the epd-pgk-fbaA and pgk-fbaA RNA segmentation

Among the four E. coli RNAses (E, G, III and P) that are involved in transcript segmentation and mRNA decay (for review: Grunberg-Manago, 1999; Wachi and Nagai, 2001; Li and Altman, 2003), RNase E plays a key role in segmentation of the epd-pgk-fbaA, epd-pgk and pgk-fbaA transcripts. RNase E is a single-stranded specific RNase, that acts within A and/or U rich segments (Ehretsmann et al., 1992; Kaberdin, 2003) and cleaves RNAs with a 5' monophosphate group in preference to RNAs which are triphosphorylated at their 5'-end (Mackie, 1998; Jiang and Belasco, 2004). In spite of this specificity, mRNA with 5'triphosphate can be cleaved at internal (or ectopic) sites (Marchand et al., 2001; Baker and Mackie, 2003). It is known that the RNase E catalytic activity depends both on the sequence and the 2D structure of the target RNA target (Mackie, 1998; Kaberdin, 2003). Nevertheless, the structural differences between the RNase E ectopic sites and conventional sites have not yet been defined. Site pgk 2 in the epd ORF, which may be an ectopic site, has all the characteristic features needed to be efficiently cleaved: it is located in a 9 nt long single-stranded A/U rich segment, flanked by two stable stem-loop structures (12.5 kcal mol-1 and 22.4 kcal mol-1, respectively, Fig. 7B), and it is located 2 nts downstream from a G residue (Fig. 7B). The presence of a G residue 1 or 2 nts upstream of the scissile bond has a strong importance for RNase E activity (Kaberdin, 2003; Redko et al., 2003). The presence of the 5' and/or 3' stem-loop structures are expected to mediate association of RNase E with its substrate through an interaction with the ARRBD C-terminal domain of RNase E (Kaberdin et al., 2000) and to influence the rate of cleavage (for review, Grunberg-Manago, 1999). Interestingly, the characteristic features of the pgk 2 site are conserved in other enterobacteria genera: Erwinia, Shigella, Salmonella and Yersinia (Fig. 7B and C and Fig. S1B). In contrast, the downstream pgk 3, 4, 5 and 6 sites are contained in a less conserved RNA

o distance • aligned distance

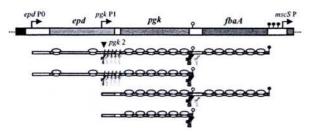


Fig. 10. Transcriptional organization and maturation processes of the *E. coli epd-pgk-fbaA* operon. The *epd-pgk-fbaA* E. *coli* DNA region is shown. The four transcripts produced are drawn below. The RNase E cleavage sites are represented by scissors. The strong *pgk* 2 and *fbaA* 5 RNase E sites are in dark. The numbers of ribosomes drawn on the ORFs reflect the TIR efficiency.

region. They are all located in A/U rich single-stranded regions. However, the flanking stem-loop structures are less stable compared with site pgk 2. In addition, no G residue is located 1 or 2 nts upstream of the pgk 5 scissile bond (Fig. 7B). We found that the pgk P1 transcripts that contain the six RNase E sites is poorly cleaved by RNase E. This may be due to the truncation of the 5' stem-loop structure of site pgk 2 in this transcript (Fig. 7B). This suggests that the structural integrity of site pgk 2 is required for cleavage of the pgk ORF and pgk inter-ORF region by RNase E.

Our current model to explain the present data (Fig. 10) is the following: the *pgk* 2 entry site in the *epd* ORF is efficiently used because of the low *epd* translation efficiency. After cleavage at site *pgk* 2, the released 3' products have a 5' monophosphate, which is favourable for further cleavages by RNase E at sites *pgk* 3–6, in spite of the degeneracy of these sites. We assume that the ribosomes translating the *pgk* ORF constitute a barrier limiting the RNase E action to the *epd* ORF and *epd-pgk* inter-ORF region. Due to the low translational efficiency of the *epd* ORF, the 5' products released by cleavage at site *pgk* 2 are likely rapidly degraded by action of the degradosome.

The fbaA 5 series of RNase E sites may also include an ectopic RNase E site. Indeed, they occur in an A/U rich single-stranded segment containing a G residue, that flanked by two stable stem-loop structures (9.7 kcal mol-1 and 18.9 kcal mol-1, respectively) (Fig. 8B), and these structural features are conserved in other enterobacteria (Salmonella, Yersinia) (Fig. 8C). The fbaA 2 site is also located in an A/U rich single-stranded segment flanked by two stem-loop structures. However, no G residue is present in the single-stranded segment (Fig. 8B), and the sequence and 2D structure of this RNA region is not conserved in other enterobacteria. Taken together, our data suggest that two conserved ectopic RNase E sites are present in the epd-pgk-fbaA transcripts (Fig. 10) and that RNase E ectopic sites carry all

the structural features required for an optimal cleavage by RNase E.

Segmentation of the epd-pgk-fbaA, epd-pgk and pgk-fbaA transcripts is probably involved in the differential expression of the three ORFs

The occurrence of processing events leading to the production of individual pgk mRNAs was also found in bacteria containing a gapA-pgk tandem instead of an epd-pgk tandem. The RNase(s) involved in the cleavage processes were not identified. However, these cleavages were proposed to be a mechanism for the differential expression of the ORFs. For instance, in Z. mobilis, cleavage into individual gapA and pgk mRNAs of the gapA-pgk transcript leads to highly stable gapA mRNAs and less stable pgk mRNAs (Burchhardt et al., 1993). In L. delbrueckii, individual gapA and pgk mRNAs were proposed to be generated by cleavage of the gapA-pgk-tpi transcript and here also, the individual gapA mRNA are more stable than the pgk mRNAs (Branny et al., 1998). More recently, the complex transcriptional organization for the gapA-pgk-tpipgm-eno ORF cluster of B. subtilis was elucidated (Ludwig et al., 2001) and a segmentation process was also detected. It appeared that combination of transcriptional regulation and RNA segmentation allow the adaptation of gene expression to variations of the environmental conditions in this species (Doan and Aymerich, 2003).

In E. coli, taking into account the location of the identified RNase E sites, RNase E is likely involved in both epd and pgk mRNA decay. Indeed, the fbaA 5 sites are located within or very close to the pak UAA stop codon, so that the released pak individual mRNAs are not expected to be stable. We assume that only the individual pak mRNAs produced by transcription termination at the pgk terminator are stable. By rifampicin addition in a culture of pEPF-transformed cells grown in LB, we estimated the half-life of the individual pgk and fbaA mRNAs to 4 and 6 min respectively (not shown). The stronger stability of individual fbaA mRNAs is likely explained by the presence of stable stem-loop structures at their two extremities. Indeed, cleavages at sites fbaA 5 or fbaA 2 both release a 3' product with a stable 5' stem-loop structure (Fig. 8B). In addition, the fbaA ORF is followed by Repeated Extragenic Palindromic sequences (REP) of the BIME 2 class (Bachellier et al., 1999), that forms three successive putative rho-independent terminators (Alefounder and Perham, 1989) and a rho-dependent terminator (Espeli et al., 2001). These structures likely constitute a barrier against exonuclease action (Bachellier et al., 1999). Hence, as already found for other polycistronic transcripts (Selinger et al., 2003), the downstream ORF in the epd-pgk-fbaA triplet is the most stable of the three ORFs.

#### 1282 V. Bardey et al.

In conclusion, we propose that upon conversion of the gapA ORF into an epd ORF, the characteristic features of the ancestral promoter upstream of gapA was conserved due to its high functional importance. Consequently, the codon usage and TIR sequence of the generated epd gene were adapted to the low requirement of E4PDH and an RNase E entry site was selected inside the ORF. In addition, other RNase E sites in the pgk-fbaA inter-ORF region are used for the differential expression of these two ORFs. Our structural analysis also brings important information on the structure of RNase E ectopic sites.

#### **Experimental procedures**

Bacterial strains, media and growth conditions

The genotypes of the various strains used in this study are given in Table 2. The untransformed and transformed E. coli TG1 strain was grown aerobically at 37°C in the minimal M63 medium supplemented with 10  $\mu g \ ml^{-1}$  thiamine, 0.1% (wt/ vol) Casamino Acids and 0.3 mM ZnCl<sub>2</sub> [required for activity of the overproduced class II FBA (Berry and Marshall, 1993)], and either 0.4% (wt/vol) glucose or 0.4% (wt/vol) pyruvate. The isogenic strains IBPC602 (WT) and IBPC600 (rnc105), N3433 (WT) and IBPC935(mg::CmR), and MG1693 (WT) and SK5665 (RNase E15), used for RNase III, RNase G and RNase E assays, respectively, were grown at 30°C in LB, in the presence of 50 µg ml-1 thymine. For RNase E and RNase G assays, when the A600 reached a value of 0.5, half of the culture was incubated at 42°C, while the other half was kept at 30°C. Samples of each culture were collected after a 15 min incubation and total RNA was prepared. The E. coli IBPC5321 and IBPC6311 strains, used for λ phage DNA recombination and production, respectively, were grown in LB medium supplemented with 0.01 M MgSO<sub>4</sub>, and 0.2% maltose. Ampicillin (100 µg ml-1) was added for growth of all the E. coli transformed cells. Growth was monitored by measuring the absorbance at 600 nm (A600).

Construction of plasmids for the study of the epd, pgk and fbaA ORF expression

Standard methods of molecular biology were used for plasmid constructions (Sambrook *et al.*, 1989). The pBluescript derivatives, pEP (previously named pPBK100), pL-EP (previously named pPBK500) and pΔP0-EP (previously named pBK200) (Fig. 1A), that contain the *epd* and *pgk* ORFs, with or without the *epd* P0 promoter were previously described (Charpentier *et al.*, 1998). The previously described M13mp9:PBK100 phage DNA and appropriate oligonucleotides (Table S1) were used to perform site-directed mutageneses by the method of Kramer (Kramer *et al.*, 1984).

To built plasmids pEPF and pΔP0-EPF containing the three epd, pgk and fbaA ORFs, a 1.981 kb long DNA fragment, containing the 3' part of the pgk ORF, the fbaA ORF and the 5' part of the mscS (mechanosensitive channels MscS) ORF, was amplified by PCR using the E. coli TG1 genome as the matrix and the O-GB7 and O-MSCS primers (Fig. 1A, Table S1). After digestion with BamHI, the amplified DNA

fragment was cloned between the EcoRV and BamHI sites of plasmids pEP and pAP0-EP. For production of plasmids pmut-10P1-EP, pmutTGP1-EP and p∆P0-mut-10P1-EP, the 2.7 kb HindIII-BamHI fragment in phage M13mp9:PBK100 DNA was subjected to site-directed mutagenesis using oligonucleotides O-mutTG or O-mut-10 and afterwards, was cloned into plasmid pBluescript (SK)\* (Stratagene). Plasmid pΔP0-mut-10P1-EPF was obtained by exchange of DNA fragments between plasmids p∆P0-EPF and pmut-10P1-EP. To build plasmid pF containing the fbaA ORF under the control of gapA promoter region, a 1.515 kb long DNA fragment was amplified by PCR using plasmid pEPF as the matrix and the O-GB15 and O-MSCS primers (Fig. 1A, Table S1). After digestion with BamHI and Ndel, the amplified DNA fragment was cloned between the Ndel and BamHI sites of plasmid pBS:: EcoGAPDH-mut62 (Charpentier, 1994). The construction of plasmid pE and the production and purification of E4PDH were previously described (Boschi-Muller et al., 1997).

#### Construction of transcriptional and translational fusions

Plasmids pRS415 and pRS414, that contain the lacZ ORF with or without the lacZTIR, were used as recipient plasmids (Simons et al., 1987). To prepare an E. coli DNA fragment containing the epd TIR, a BamHI restriction site was created between positions +183 to +188 in the epd ORF by sitedirected mutagenesis of the phage M13mp9:PBK100 DNA (mutant phage M13mp9:PBK100BamHI). Substitutions of the +58 to +86 epd TIR sequence for the GTCGAC sequence (variant 35) and of the epd TIR sequence from position +130 to +132 for an AAATAAA sequence (variant 408) (Fig. 1B), were made in M13mp9:PBK100BamHI phage DNA, using oligonucleotides O-mut35 and O-mut408 respectively. The HindIII-BamHI DNA fragment from the WT and mutated phage DNAs were inserted into the recipient plasmids together with the Mfel-HindIII fragment from plasmid pL-EP. The resulting pRS415 plasmid derivatives containing the WT, mut35 or mut408 transcriptional fusions were denoted pVBOF2, pVBOFv35 and pVBOFv408 respectively. Similarly, plasmid pRS414 derivatives, containing the WT and mutated translational fusions were denoted pVBGF2, pVBGF2v35 and pVBGF2v408 respectively. To create plasmids pmut35-L-EP and pmut408-L-EP, the 2.7 kb HindIII-BamHI fragments of the mutated phage DNAs, were inserted between the HindIII and BamHI sites of plasmid pL-EP.

The translational and transcriptional fusions obtained with the WT or mutated *epd* genes were transferred into phage λRS45 DNA, by homologous recombination using *E. coli* IBPC5321 cells and a previously described procedure (Simons *et al.*, 1987). Monolysogenic *E. coli* TG1 cells were screened by the PCR method (Powell *et al.*, 1994). A detailed description of all the constructs made in this study can be obtained upon request.

#### Plasmid copy number estimation

The plasmid copy number of the cells was estimated by extraction of the DNA contained in a defined number of cells (about  $1.5 \times 10^9$  cells, corresponding to  $15 A_{600}$ ), followed by

agarose gel electrophoresis and quantification of the ethidium bromid fluorescence using a Gel-Doc1000 Bio-Rad apparatus and the Molecular Analysis software (Bio-Rad). The copy number of plasmid pEPF was taken as the reference.

#### RNA isolation

Total RNA was extracted from 30 ml of a cell culture harvested at an  $A_{600}$  of about 0.5, by two hot phenol extractions (Aiba *et al.*, 1981), followed by two chloroform extractions and one ethanol precipitation. The RNA pellet was washed in 70% ethanol and dissolved in 70  $\mu$ l of sterile MilliQ water. Before freezing at  $-80^{\circ}$ C, the RNA concentration was estimated by  $A_{260}$  measurement.

#### Northern blot analysis

On the basis of the estimated cellular plasmid copy number, 5-9 µg of total RNA extracted from transformed E. coli TG1 cells were fractionated on a 0.8% (wt/vol) agaroseformaldehyde gel containing 0.5 μg ml-1 BrEt (Sambrook et al., 1989). The 16S and 23S rRNAs from the total RNA mixture was used as molecular markers. The fractionated RNAs were transferred on a Biodyne B transfer membrane (0.45 µm, Pall) in 20x SSC [3 M NaCl, 3 M sodium acetate, (pH 7.0)], and were fixed by UV cross-linking for 5 min. Prehybridization and hybridization with the 5'-end-labelled oligonucleotide O-EPD, O-PGK or O-FBA (Fig. 1A and Table S1) were performed at 42°C in 6× SSC solution, containing 0.1% SDS, 5x Denhardt's reagent, 5 mM EDTA and 10 µg ml-1 of sonicated and denatured calf thymus DNA (Sambrook et al., 1989). Hybridizations were performed with 100 ng of 5'-end-labelled oligonucleotide in 10 ml of hybridization buffer. For probe labelling, 100 ng of oligonucleotide were incubated with [\gamma^{32}-P]ATP and T4 polynucleotide kinase, according to Charpentier et al. (1998). After hybridization and before autoradiography, filters were washed in 6x SSC buffer containing 0.1% SDS, once for 2 min at room temperature, then twice for 15 min at 42°C, and once again for 2 min at room temperature.

Northern blot analysis of total RNA from untransformed cells were performed in the same conditions as described above except that 30  $\mu$ g of total RNA were used and instead of 5′-end-labelled oligonucleotide probes, we used uniformly labelled amplified PCR DNA fragments. For amplification, plasmid pEPF was used as the matrix, and the oligonucleotide pairs O-GB16, O-GB17; O-GB18, O-GB19 and O-GB20, O-GB21 (Table S1) were used as primers to get the *epd-*, *pgk-* and *fbaA*-specific probes respectively. Amplification was performed in the presence of dATP, dGTP and dTTP at 0.2 mM concentrations, dCTP at a 8  $\mu$ M concentration and 1  $\mu$ I [ $\alpha$ - $^{32}$ P] dCTP at 3000 Ci mmole $^{-1}$ .

#### Primer extension analysis

Primer extension analyses were performed with 5 ng of 5'end-labelled oligonucleotide and 5 µg of total RNAs extracted from transformed cells or 20 µg of total RNAs extracted from untransformed cells, using conditions previously described (Charpentier et al., 1998). The resulting cDNAs were fractionated on a 6% sequencing gel, in parallel with the products of dideoxy-sequencing reactions obtained with the same primer.

#### S1 mapping experiments

The 5'-end-labelled probe (Fig. 1A) used for analysis of *pgk* mRNA extremities was amplified by PCR, by using a cold O-GB1 primer and a 5'-end-labelled O-PGK primer (Table S1). It was purified on a 0.8% (wt/vol) agarose gel made up in TBE buffer (Sambrook *et al.*, 1989). About 1.5 × 10<sup>5</sup> cpm of the probe were annealed with 5 μg of total RNA extracted from TG1 cells transformed with plasmid pEP or 30 μg of total RNA extracted from untransformed cells. The S1 nuclease treatment was performed according to Sambrook (Sambrook *et al.*, 1989). The resistant DNA fragments were fractionated on a 6% sequencing gel and compared with the products of a sequencing reaction.

#### In vitro transcription

Two DNA fragments containing a T7 promoter were PCR amplified by using the primer pair O-GB9 and O-GB10 or O-GB8 and O-FBA respectively (Fig. 1A and Table S1). They were then, respectively, used as the matrix for *in vitro* transcription of RNA  $f_{opd-pgk}$  and  $f_{pgk-fbaA}$  by T7 RNA polymerase. The transcription conditions were as previously described (Mougin *et al.*, 1996). The RNA transcripts were dissolved in milliQ water and their concentrations were estimated by  $A_{260}$  measurement.

#### In vitro maturation by RNase E

We used a degradosome fraction from the *E. coli* SK5665 (RNase E<sup>ts</sup>) strain generously provided by A. J. Carpousis and prepared according to Carpousis *et al.* (1994). After a 1/32 (v/v) dilution in 500 mM NaCl, 5% glycerol, 0.5% Triton X-100, 1 mM EDTA, 1 mM DTT, 10 mM Tris-HCl (pH 7.5) buffer, degradosome fractions were incubated for 10 min at 30°C or 45°C (for thermal inactivation) and 2 μl of the incubated preparations were used to digest 0.8 pmol of f<sub>epd-pot</sub> transcript in the conditions previously described (Hajnsdorf *et al.*, 1994). The digestion products were analysed by extension of oligonucleotide O-PGK with RT.

#### RNA secondary structure analysis

Chemical modifications and enzymatic digestion of RNA were all performed on about 100 ng of cold f<sub>epd-pgk</sub> or f<sub>pgk-thuk</sub> transcript, in the presence of 2 μg of yeast tRNA mixture (Roche Diagnostics) and 10 mM Mg<sup>2+</sup> (Mougin *et al.*, 1996). For enzymatic digestion with V1, T1 or T2 RNase, the RNA was preincubated for 5 min at 65°C in 50 μl of a 100 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM Tris–HCl (pH 7.5) solution. After slow cooling at 20°C, the RNA was incubated for 10 min at 20°C in the presence of T1 RNase (Roche Diagnostics) 1.25 or 2.5 U, T2 RNase (Gibco BRL) 2.4 or 4.8 U or V1 RNase (prepared from *Naja oxiana* venom) 0.05, 0.10 U. V1 RNase action was stopped by addition of 100 mM EDTA, T1 and T2

RNase digestions were stopped by adding 20 µg of tRNA. The digestion products were phenol-extracted and ethanol precipitated. For chemical modification with dimethylsulphate (DMS), 1-cyclohexyl-3-(2-morpholinoethyl)-carbodiimide-methop-toluene sulphonate (CMCT) and kethoxal, the RNA was preincubated for 5 min at 65°C in 50 µl of a 100 mM KCl, 10 mM MgCl<sub>2</sub>, 500 mM sodium cacodylate (pH 7.5) solution (DMS and kethoxal) or 50 µl of a 100 mM KCl, 10 mM MgCl<sub>2</sub>, 50 mM sodium borate (pH 8) solution (CMCT). After slow cooling at 20°C, incubation was at 20°C for 6 min in the presence of DMS [0.5 or 1 µl of a 25% (v/v) DMS-ethanol] or 10 min in the presence of kethoxal (74 or 185 μg) or CMCT (840 or 1680 μg). DMS, CMCT and kethoxal reactions were stopped as previously described (Marmier-Gourrier et al., 2003). RNA was ethanol precipitated, then dissolved in 4 μl of distilled water. Positions of chemical modifications and enzymatic cleavages were identified by primer extension analysis on 1 µl of RNA solution, performed as previously described (Mougin et al., 1996), using the 5'-end-labelled oligonucleotides O-PGK and O-GB11 for the  $f_{\it epd-pgk}$  RNA and O-GB12, O-FBA, O-GB13 and O-GB14 for the fpgk-tbaA RNA.

#### RNA secondary structure prediction

Predictions of RNA secondary structure were made with the Mfold software (http://www.bioinfo.rpi.edu/applications/mfold/old/rna/form1.cgi) by Zuker (2003). The results of the experimental analysis were introduced as constraints in the computer search. The free energies at 37°C in 1 M NaCl of the proposed stem-loop structures were estimated using the Mfold software.

The GenBank accession numbers of sequences used in the comparative phylogenetic analysis are NC\_000913 (*E. coli* K12), NC\_004547 (*E. carotovora* ssp. *atroseptica* SCRI1043), NC\_004337 (*S. flexneri* 2a str. 301), NC\_004631 (*S. enterica* ssp. *enterica* serovar *Typhi* Ty2), NC\_003197 (*S. typhimurium* LT2), NC\_004088 (*Y. pestis* KIM) and NC\_006155 (*Y. pseudotuberculosis* IP 32953).

#### Enzymatic-specific activity measurements

For measurements of the PGK and FBA activities, 30 ml fractions of cell cultures, that reached an A600 of 0.5 (about 1.5 × 109 cells) were centrifuged. The cell pellet was washed and sonicated in 3 ml of 50 mM Tris-HCl (pH 8.0) buffer, as previously described (Charpentier et al., 1998). The PGK and FBA activities were measured by their coupling with the Bacillus stearothermophilus GAPDH activity. For PGK activity measurement, NADH disappearance was followed as previously described (Charpentier et al., 1998). The specific activity was expressed in µmol NADH/min/A280/A600. For class II FBA-specific activity determination, we used a protocol adapted from the one of Christian et al. (Christian, 1955): 500 μl of a 2 mM fructose-1,6-bisphosphate (Sigma), 2 mM NAD\*, 17 mM Na arsenate, 50 mM Tris-HCI (pH 8.0) solution, containing 0.2 A280 ml-1 of B. stearothermophilus GAPDH, was preincubated 5 min at 30°C and an aliquot of the cell sonication soluble fraction (2-20 µI) was added. The NADH appearance was followed at 340 nm. The FBA activity was expressed in µmol NADH/min/A<sub>280</sub>/A<sub>600</sub>. To compare the expressions of the pgk and fbaA ORFs from plasmid pEPF,  $p\Delta P0$ -EPF and  $p\Delta P0$ -mut-10P1-EPF, the PGK and FBA activities measured for cells transformed with these two latter plasmids were multiplied by the ratio of the plasmid copy number in pEPF-transformed cells versus that in  $p\Delta P0$ -EPF and  $p\Delta P0$ -mut-10P1-EPF-transformed cells respectively.

E4PDH activity was measured in the stationary phase of growth on a volume of cell culture corresponding to  $60 A_{600}$ . The cell pellet was washed and sonicated in 3 ml of 2 mM EDTA, 50 mM Tris-HCl (pH 8.0) buffer and E4PDH-specific activity was measured as described (Boschi-Muller *et al.*, 1997). The PGK-specific activities of the samples were also determined, using the same conditions as above.

For measurement of the  $\beta$ -galactosidase-specific activity, the *E. coli* TG1 cells, transformed with the different *epd-lacZ* plasmid constructs, or lysogenic TG1 cells were harvested at A<sub>600</sub> of both 0.2 and 0.4 Units. The  $\beta$ -galactosidase activity was measured on toluene-lysed cells, as previously described (Miller, 1972).

#### Acknowledgements

M. Graffe is thanked for her help in the construction of *epd-lacZ* fusions. A. J. Carpousis is greatly acknowledged for providing us with the degradosome fraction, containing the RNase E. S. Boshi-Muller and S. Azza are thanked for their gift of *Bacillus stearothermophilus* GAPDH and purified *E. coli* E4PDH. D. Ropers is thanked for her advice in the experimental secondary structure analysis. This work was supported by the French Ministère Délégué à la Recherche et aux Nouvelles Technologies, the Centre National de la Recherche Scientifique, the PRST Bioingénierie of Région Lorraine. Vincent Bardey and Corinne Vallet had a CNRS-BDI fellowship, cofinanced by the Région Lorraine for Vincent Bardey. Benoit Thouvenot was supported by a fellowship from the French Ministère Délégué à la Recherche et aux Nouvelles Technologies.

#### References

Aiba, H., Adhya, S., and de Crombrugghe, B. (1981) Evidence for two functional gal promoters in intact Escherichia coli cells. J Biol Chem 256: 11905–11910.

Alefounder, P.R., and Perham, R.N. (1989) Identification, molecular cloning and sequence analysis of a gene cluster encoding the class II fructose 1,6-bisphosphate aldolase, 3-phosphoglycerate kinase and a putative second glyceraldehyde 3-phosphate dehydrogenase of *Escherichia coli*. *Mol Microbiol* 3: 723–732.

Arraiano, C.M., Yancey, S.D., and Kushner, S.R. (1988) Stabilization of discrete mRNA breakdown products in ams pnp rnb multiple mutants of Escherichia coli K-12. J Bacteriol 170: 4625–4633.

Bachellier, S., Clement, J.M., and Hofnung, M. (1999) Short palindromic repetitive DNA elements in enterobacteria: a survey. Res Microbiol 150: 627–639.

Baker, K.E., and Mackie, G.A. (2003) Ectopic RNase E sites promote bypass of 5'-end-dependent mRNA decay in *Escherichia coli. Mol Microbiol* **47:** 75–88.

Bardey, V. (2000) Organisation et régulation de l'expression

- des gènes de la Glycéraldéhyde-3-phosphate déshydrogénase, la Phosphoglycérate kinase et l'Erythrose-4-phosphate déshydrogénase chez les protéobactéries de la subdivision gamma. Thesis, UHP-Nancy I.
- Berry, A., and Marshall, K.E. (1993) Identification of zincbinding ligands in the class II fructose-1,6-bisphosphate aldolase of Escherichia coli. FEBS Lett 318: 11-16.
- Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., et al. (1997) The complete genome sequence of Escherichia coli K-12. Science 277: 1453– 1474
- Boschi-Muller, S., Azza, S., Pollastro, D., Corbier, C., and Branlant, G. (1997) Comparative enzymatic properties of GapB-encoded erythrose-4-phosphate dehydrogenase of *Escherichia coli* and phosphorylating glyceraldehyde-3-phosphate dehydrogenase. *J Biol Chem* 272: 15106–15112
- Bown, J.A., Barne, K.A., Minchin, S.D., and Busby, S.J.W. (1997) Extended -10 promoters. In *Nucleic Acids and Molecular Biology*. Eckstein, F., and Lilley, D.M.J. (eds). Berlin: Springer, 11: 41–52.
- Branlant, G., and Branlant, C. (1985) Nucleotide sequence of the Escherichia coli gap gene. Different evolutionary behavior of the NAD\*-binding domain and of the catalytic domain of p-glyceraldehyde-3-phosphate dehydrogenase. Eur J Biochem 150: 61–66.
- Branlant, C., Krol, A., Machatt, M.A., Pouyet, J., Ebel, J.P., Edwards, K., and Kossel, H. (1981) Primary and secondary structures of *Escherichia coli* MRE 600, 23S ribosomal RNA. Comparison with models of secondary structure for maize chloroplast 23S rRNA and for large portions of mouse and human 16S mitochondrial rRNAs. *Nucleic Acids Res* 9: 4303–4324.
- Branlant, C., Oster, T., and Branlant, G. (1989) Nucleotide sequence determination of the DNA region coding for Bacillus stearothermophilus glyceraldehyde-3-phosphate dehydrogenase and of the flanking DNA regions required for its expression in Escherichia coli. Gene 75: 145–155.
- Branny, P., de la Torre, F., and Garel, J.R. (1998) An operon encoding three glycolytic enzymes in *Lactobacillus del-brueckii* subsp. *bulgaricus*: glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase and triosephosphate isomerase. *Microbiology* 144: 905–914.
- Braun, F., Hajnsdorf, E., and Regnier, P. (1996) Polynucleotide phosphorylase is required for the rapid degradation of the RNase E-processed *rpsO* mRNA of *Escherichia coli* devoid of its 3' hairpin. *Mol Microbiol* 19: 997–1005.
- Burchhardt, G., Keshav, K.F., Yomano, L., and Ingram, L.O. (1993) Mutational analysis of segmental stabilization of transcripts from the *Zymomonas mobilis gap-pgk* operon. *J Bacteriol* 175: 2327–2333.
- Carpousis, A.J., Van Houwe, G., Ehretsmann, C., and Krisch, H.M. (1994) Copurification of *E. coli* RNAase E and PNPase: evidence for a specific association between two enzymes important in RNA processing and degradation. *Cell* **76**: 889–900.
- Charpentier, B. (1994) Etude de promoteurs transcriptionnels des gènes gapA et gapB codant la Glycéraldéhyde-3-phosphate déshydrogénase d'Escherichia coli. Relation structure/fonction. Thesis, UHP-Nancy I.
- Charpentier, B., Bardey, V., Robas, N., and Branlant, C.

- (1998) The EII<sup>Glc</sup> protein is involved in glucose-mediated activation of *Escherichia coli gapA* and *gapB-pgk* transcription. *J Bacteriol* **180**: 6476–6483.
- Christian, W. (1955) Aldolase from yeast. Methods Enzymol 1: 315–320.
- Condon, C., Squires, C., and Squires, C.L. (1995) Control of rRNA transcription in *Escherichia coli. Microbiol Rev* 59: 623–645
- Conway, T., and Ingram, L.O. (1988) Phosphoglycerate kinase gene from Zymomonas mobilis: cloning, sequencing, and localization within the gap operon. J Bacteriol 170: 1926–1933.
- Della Seta, F., Boschi-Muller, S., Vignais, M.L., and Branlant, G. (1997) Characterization of *Escherichia coli* strains with gapA and gapB genes deleted. J Bacteriol 179: 5218– 5221.
- Doan, T., and Aymerich, S. (2003) Regulation of glycolytic genes in *Bacillus subtilis*: binding of the repressor CggR to its single DNA target sequence is modulated by fructose-1,6-bisphosphate. *Mol Microbiol* 47: 1709–1721.
- Ehretsmann, C.P., Carpousis, A.J., and 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.
- Eikmanns, B.J. (1992) Identification, sequence analysis, and expression of a *Corynebacterium glutamicum* gene cluster encoding the three glycolytic enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate kinase, and triosephosphate isomerase. *J Bacteriol* 174: 6076–6086
- Espeli, O., Moulin, L., and Boccard, F. (2001) Transcription attenuation associated with bacterial repetitive extragenic BIME elements. *J Mol Biol* **314**: 375–386.
- Fillinger, S., Boschi-Muller, S., Azza, S., Dervyn, E., Branlant, G., and Aymerich, S. (2000) Two glyceraldehyde-3-phosphate dehydrogenases with opposite physiological roles in a nonphotosynthetic bacterium. *J Biol Chem* 275: 14031– 14037.
- Fraenkel, D.G. (1996) Glycolysis. In Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology, 2nd edn. Vol. 1. Neidhardt, F.C. (ed.). Washington, DC: American Society for Microbiology Press, pp. 189–198.
- Gebbia, J.A., Backenson, P.B., Coleman, J.L., Anda, P., and Benach, J.L. (1997) Glycolytic enzyme operon of *Borrelia burgdorferi*: characterization and evolutionary implications. *Gene* 188: 221–228.
- Grunberg-Manago, M. (1999) Messenger RNA stability and its role in control of gene expression in bacteria and phages. Annu Rev Genet 33: 193–227.
- Hajnsdorf, E., Steier, O., Coscoy, L., Teysset, L., and Regnier, P. (1994) Roles of RNase E, RNase II and PNPase in the degradation of the *rpsO* transcripts of *Escherichia coli*: stabilizing function of RNase II and evidence for efficient degradation in an *ams pnp rnb* mutant. *EMBO J* 13: 3368–3377.
- Harley, C.B., and Reynolds, R.P. (1987) Analysis of *E. coli* promoter sequences. *Nucleic Acids Res* **15**: 2343–2361.
- Hillman, J.D., and Fraenkel, D.G. (1975) Glyceraldehyde 3phosphate dehydrogenase mutants of *Escherichia coli. J Bacteriol* 122: 1175–1179.
- Jacquenet, S., Ropers, D., Bilodeau, P.S., Damier, L., Mou-

- gin, A., Stoltzfus, C.M., and Branlant, C. (2001) Conserved stem-loop structures in the HIV-1 RNA region containing the A3 3' splice site and its cis-regulatory element: possible involvement in RNA splicing. *Nucleic Acids Res* 29: 464–478.
- Jiang, X., and Belasco, J.G. (2004) Catalytic activation of multimeric RNase E and RNase G by 5'-monophosphorylated RNA. Proc Natl Acad Sci USA 101: 9211–9216.
- Kaberdin, V.R. (2003) Probing the substrate specificity of Escherichia coli RNase E using a novel oligonucleotidebased assay. Nucleic Acids Res 31: 4710–4716.
- Kaberdin, V.R., Walsh, A.P., Jakobsen, T., McDowall, K.J., and von Gabain, A. (2000) Enhanced cleavage of RNA mediated by an interaction between substrates and the arginine-rich domain of *E. coli* ribonuclease E. *J Mol Biol* 301: 257–264.
- Karlin, S., Mrazek, J., Campbell, A., and Kaiser, D. (2001) Characterizations of highly expressed genes of four fastgrowing bacteria. J Bacteriol 183: 5025–5040.
- Keilty, S., and Rosenberg, M. (1987) Constitutive function of a positively regulated promoter reveals new sequences essential for activity. J Biol Chem 262: 6389–6395.
- Kramer, W., Drutsa, V., Jansen, H.W., Kramer, B., Pflug-felder, M., and Fritz, H.J. (1984) The gapped duplex DNA approach to oligonucleotide-directed mutation construction. *Nucleic Acids Res* 12: 9441–9456.
- Lee, K., Bernstein, J.A., and Cohen, S.N. (2002) RNase G complementation of *rne* null mutation identifies functional interrelationships with RNase E in *Escherichia coli. Mol Microbiol* 43: 1445–1456.
- Lesnik, E.A., Sampath, R., Levene, H.B., Henderson, T.J., McNeil, J.A., and Ecker, D.J. (2001) Prediction of rhoindependent transcriptional terminators in *Escherichia coli*. *Nucleic Acids Res* 29: 3583–3594.
- Li, Y., and Altman, S. (2003) A specific endoribonuclease, RNase P, affects gene expression of polycistronic operon mRNAs. Proc Natl Acad Sci USA 100: 13213–13218.
- Lisser, S., and Margalit, H. (1993) Compilation of E. coli mRNA promoter sequences. Nucleic Acids Res 21: 1507– 1516.
- Ludwig, H., Homuth, G., Schmalisch, M., Dyka, F.M., Hecker, M., and Stulke, J. (2001) Transcription of glycolytic genes and operons in *Bacillus subtilis*: evidence for the presence of multiple levels of control of the *gapA* operon. *Mol Micro-biol* 41: 409–422.
- Ma, J., Campbell, A., and Karlin, S. (2002) Correlations between Shine–Dalgarno sequences and gene features such as predicted expression levels and operon structures. J Bacteriol 184: 5733–5745.
- Mackie, G.A. (1998) Ribonuclease E is a 5'-end-dependent endonuclease. Nature 395: 720–723.
- Marchand, I., Nicholson, A.W., and Dreyfus, M. (2001) Bacteriophage T7 protein kinase phosphorylates RNase E and stabilizes mRNAs synthesized by T7 RNA polymerase. *Mol Microbiol* 42: 767–776.
- Marmier-Gourrier, N., Clery, A., Senty-Segault, V., Charpentier, B., Schlotter, F., Leclerc, F., et al. (2003) A structural, phylogenetic, and functional study of 15.5-kD/Snu13 protein binding on U3 small nucleolar RNA. RNA 9: 821–838.
- Martin, W., Brinkmann, H., Savonna, C., and Cerff, R. (1993) Evidence for a chimeric nature of nuclear genomes: eubac-

- terial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. *Proc Natl Acad Sci USA* **90**: 8692–8696.
- Meijer, W.G., van den Bergh, E.R., and Smith, L.M. (1996) Induction of the gap-pgk operon encoding glyceraldehyde-3-phosphate dehydrogenase and 3-phosphoglycerate kinase of Xanthobacter flavus requires the LysR-type transcriptional activator CbbR. J Bacteriol 178: 881–887.
- Miller, J.H. (1972) Experiments in Molecular Genetics. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Mitchell, J.E., Zheng, D., Busby, S.J., and Minchin, S.D. (2003) Identification and analysis of 'extended -10' promoters in *Escherichia coli*. Nucleic Acids Res 31: 4689–4695.
- Mougin, A., Gregoire, A., Banroques, J., Segault, V., Fournier, R., Brule, F., et al. (1996) Secondary structure of the yeast Saccharomyces cerevisiae pre-U3A snoRNA and its implication for splicing efficiency. RNA 2: 1079–1093.
- Negre, D., Oudot, C., Prost, J.F., Murakami, K., Ishihama, A., Cozzone, A.J., and Cortay, J.C. (1998) FruR-mediated transcriptional activation at the *ppsA* promoter of *Escherichia coli*. *J Mol Biol* 276: 355–365.
- Nellemann, L.J., Holm, F., Atlung, T., and Hansen, F.G. (1989) Cloning and characterization of the *Escherichia coli* phosphoglycerate kinase (*pgk*) gene. *Gene* 77: 185–191.
- Peng, L., and Shimizu, K. (2003) Global metabolic regulation analysis for *Escherichia coli* K12 based on protein expression by 2-dimensional electrophoresis and enzyme activity measurement. *Appl Microbiol Biotechnol* 61: 163–178.
- Plumbridge, J., and Soll, D. (1989) Characterization of cisacting mutations which increase expression of a glnS-lacZ fusion in Escherichia coli. Mol Gen Genet 216: 113–119.
- Powell, B.S., Rivas, M.P., Court, D.L., Nakamura, Y., and Turnbough, C.L., Jr (1994) Rapid confirmation of single copy lambda prophage integration by PCR. *Nucleic Acids Res* 22: 5765–5766.
- Prost, J.F., and Cozzone, A.J. (1999) Detection of an extended-10 element in the promoter region of the pckA gene encoding phosphoenolpyruvate carboxykinase in Escherichia coli. Biochimie 81: 197–200.
- Ramseier, T.M., Bledig, S., Michotey, V., Feghali, R., and Saier, M.H., Jr (1995) The global regulatory protein FruR modulates the direction of carbon flow in *Escherichia coli*. *Mol Microbiol* 16: 1157–1169.
- Redko, Y., Tock, M.R., Adams, C.J., Kaberdin, V.R., Grasby, J.A., and McDowall, K.J. (2003) Determination of the catalytic parameters of the N-terminal half of *Escherichia coli* ribonuclease E and the identification of critical functional groups in RNA substrates. *J Biol Chem* 278: 44001– 44008.
- Romby, P., and Springer, M. (2003) Bacterial translational control at atomic resolution. *Trends Genet* 19: 155–161.
- Saier, M.H., Jr, and Ramseier, T.M. (1996) The catabolite repressor/activator (Cra) protein of enteric bacteria. J Bacteriol 178: 3411–3417.
- Salgado, H., Gama-Castro, S., Martinez-Antonio, A., Diaz-Peredo, E., Sanchez-Solano, F., Peralta-Gil, M., et al. (2004) Regulondb (version 4.0): transcriptional regulation, operon organization and growth conditions in *Escherichia coli* K-12. *Nucleic Acids Res* 32 Database issue: D303–306.

- Sambrook, J., Fritsch, F.F., and Maniatis, T. (1989) Molecular Cloning: a Laboratory Manual, 2nd edn. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.
- Schlapfer, B.S., and Zuber, H. (1992) Cloning and sequencing of the genes encoding glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase and triosephosphate isomerase (gap operon) from mesophilic Bacillus megaterium: comparison with corresponding sequences from thermophilic Bacillus stearothermophilus. Gene 122: 53–62.
- Selinger, D.W., Saxena, R.M., Cheung, K.J., Church, G.M., and Rosenow, C. (2003) Global RNA half-life analysis in Escherichia coli reveals positional patterns of transcript degradation. Genome Res 13: 216–223.
- Simons, R.W., Houman, F., and Kleckner, N. (1987) Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene* 53: 85–96.
- Stokes, N.R., Murray, H.D., Subramaniam, C., Gourse, R.L., Louis, P., Bartlett, W., et al. (2003) A role for mechanosensitive channels in survival of stationary phase: regulation of channel expression by RpoS. Proc Natl Acad Sci USA 100: 15959–15964.
- Sunohara, T., Jojima, K., Yamamoto, Y., Inada, T., and Aiba, H. (2004) Nascent-peptide-mediated ribosome stalling at a stop codon induces mRNA cleavage resulting in nonstop mRNA that is recognized by tmRNA. RNA 10: 378–386.

- Thouvenot, B., Charpentier, B., and Branlant, C. (2004) The strong efficiency of the *Escherichia coli gapA* P1 promoter depends upon a complex combination of functional determinants. *Biochem J* 383: 371–382.
- Wachi, M., and Nagai, K. (2001) Escherichia coli ribonuclease G. Methods Enzymol 342: 55-63.
- Zhao, G., Pease, A.J., Bharani, N., and Winkler, M.E. (1995) Biochemical characterization of gapB-encoded erythrose 4-phosphate dehydrogenase of Escherichia coli K-12 and its possible role in pyridoxal 5'-phosphate biosynthesis. J Bacteriol 177: 2804–2812.
- Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31: 3406– 3415

#### Supplementary material

The following supplementary material is available for this article online:

- Fig. S1. Complementary data on the structural analysis and phylogenetic comparison of the *epd-pgk* RNA region from position +950 to +1200.
- Fig. S2. Complementary data on the structural analysis of the *pgk-fbaA* intergenic region from position +2330 to +2597. Table S1. Nucleotide sequence<sup>a</sup> of the oligonucleotide primers used in this study.

#### Supplementary Material

#### FIGURE LEGENDS

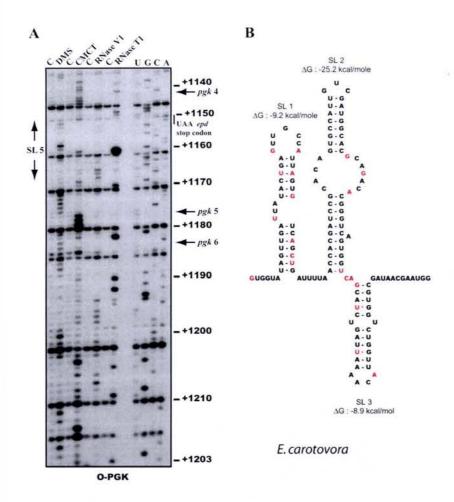
**Figure 1:** Complementary data on the structural analysis and phylogenetic comparison of the *epd-pgk* RNA region from position +950 to +1200.

Panel A: Experimental analysis of the secondary structure of the segment extending from position +1140 to +1203 using chemical (DMS and CMCT) and enzymatic (RNase V1 and T1) probes. The legend is the same as in the Figure 7 panel A. The oligonucleotide O-PGK was used as the primer for reverse transcription. The stem-loop structure 5 shown in Figure 7 Panel B is indicated on the left of the Panel. Nucleotide numbering and positions of the RNase E cleavage sites pgk 4 to pgk 6 are indicated on the right of the figure, as well as, the epd UAA stop codon.

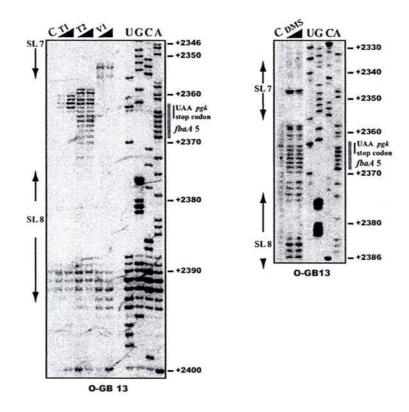
Panel B: Conservation of the possibility to form the SL1, SL2 and SL3 stem-loop structures in the *epd* ORF of *Erwinia carotovora*. The *Erwinia carotovora* subsp. astroseptica SCRI1043 is numbered NC 004337 in the EMBL database. The legend is the same as in Figure 7, Panel B.

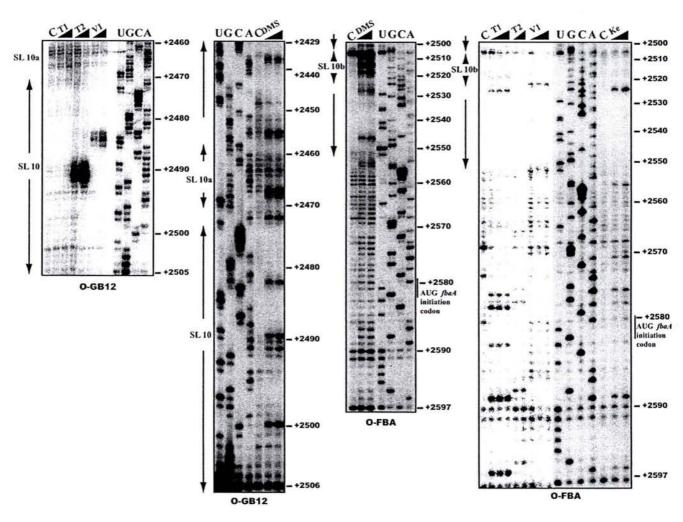
**Figure 2:** Complementary data on the structural analysis of the *pgk-fbaA* intergenic region from position 2330 to 2597.

Experimental analysis of the secondary structure of the segment extending from position +1140 to +1203 using chemical (DMS and CMCT) and enzymatic (RNase V1, T1 and T2) probes. The legend is the same as in the Figure 8 panel A. The oligonucleotide O-GB12, O-GB13 and O-FBA were used as the primers for reverse transcription. The stem-loop structures 7 to 10 shown in Figure 8 Panel B are indicated on the left of the Panels. Nucleotide numbering and positions of the series of RNase E *fbaA* 5 cleavage sites is indicated on the right side of the autoradiograms, as well as, the *pgk* UAA stop codon and *fbaA* AUG initiation codon.



Supplementary Material Figure 1 Bardey *et al.* 





Supplementary Materia Figure 2 Bardey *et al*.

#### **Supplementary Material**

TABLE S1: Nucleotide sequence<sup>a</sup> of the oligonucleotide primers used in this study

| Primers uses        | d for PCR amplification                                    | targeted sequ | ences <sup>b</sup> |
|---------------------|--|---------------|--------------------|
| O-GB1               | TGGAATAAAGCTTCCCACAA                                       |               | (-91)              |
| O-GB7               | ATCGGTGATGCTTCCGCT   | (+2057) -     | (+2074)            |
| O-GB8c              | TCTAATACGACTCACTATAGGGAAAATCTCCTACATCTCCAC                 | (+2264) -     | (+2283)            |
| O-GB9c              | TCTAATACGACTCACTATAGGGTTATTGATGCATACCATC                   | (+701) -      | (+718)             |
| O-GB10              | CAACGCCGTCGAGGTAA  | (+1499) -     | (+1483)            |
| O-GB15              | GACATATGTCTAAGATTTTTGATTTCG                                | (+2580) -     | (+2601)            |
| O-MSCS <sup>d</sup> | CCCGCCGGCGATCCGCAGCAGCGCCTGGTTA                            | (+4089) -     | (+4073)            |
| Primers use         | d for site-directed mutageneses                            |               |                    |
| O-mut-10            | TCCGTATAGTCAGCTGTGCCATGAAATGCA                             | (+971) -      | (+942)             |
| O-mutTG             | AGTCAACTATACGTTGAAATGCACCT                                 | (+964) -      | (+939)             |
| O-mutBamH           | H ATACAAAGCAGGATCCACATTACG                                 | (+198) -      | (+175)             |
| O-mut35             | TTGCCAGACAGAGTAGTCGACCTGTCGCAAACTGCG                       | (+101) -      | (+43)              |
| O-mut408            | ACGCGTACGGTCATTTTATTTTCTCCTGCAAGGT                         | (+146) -      | (+117)             |
| Primers use         | d for primer extension analyses and Northern blot analyses |               |                    |
| O-EPD               | GCACGAACCACATTACG  | (+191) -      | (+175)             |
| O-PGK               | TCCCAGCAAGATCCAGATCG                                       | (+1241) -     | (+1222)            |
| O-GB11              | GTGCGCCACTGACCCG   | (+1039) -     | (+1054)            |
| O-FBA               | TCACCAGTGATTACGCCAG  | (+2626) -     | (+2606)            |
| O-GB12              | CGTCGGGCCTTATACTCGTC                                       | (+2563) -     | (+2544)            |
| O-GB13              | GAGTATGAAAAATTGCGTGC                                       | (+2433) -     | (+2414)            |
| O-GB14              | GCCCTGTGATTTTTTACTTC                                       | (+2382) -     | (+2361)            |
| O-GB16              | CATGGCCGTTTTGCATGG   | (+292) -      | (+309)             |
| O-GB17              | GGGTTTGGGTGCCATCGACAAT                                     | (+1039) -     | (+1018)            |
| O-GB18              | GATTCACCAGGTCTGTAAT  | (+1194) -     | (+1212)            |
| O-GB19              | TTACTTCTTAGCGCGCTCTT                                       | (+2365) -     | (+2346)            |
| O-GB20              | TTACAGAACGTCGATCGCG  | (+3659) -     | (3641)             |
| O-GB21              | ATGTCTAAGATTTTTGATTTCG                                     | (+2580) -     | (+2601)            |

<sup>&</sup>lt;sup>a</sup> The nucleotide sequence are written from 5' to 3'. <sup>b</sup> The positions of the extremities of the targeted DNA or RNA sequences in the *epd-pgk-fbaA* region are given taking the *epd* P0 transcription start-site as position +1 (Fig. 1A). <sup>c</sup> The T7 promoter sequences are shown in italic.

# Travaux complémentaires réalisés sur les régulations de l'expression des opérons epd-pgk-fbaA

Au cours de cette étude, nous avons montré que, chez *E. coli*, les gènes *pgk* et *fbaA* sont transcrits à partir des mêmes promoteurs: le promoteur *epd* P0, fortement activé en présence de glucose, et le promoteur étendu *pgk* P1. Le taux de production des transcrits *epd-pgk-fbaA* et *epd-pgk* à partir du promoteur *epd* P0 dépend des conditions de croissance, alors que le promoteur *pgk* P1 semble assurer un taux basal de transcription des ORF *pgk* et *fbaA*, quelques soient les conditions de croissance. Ceci est à rapprocher de la situation observée pour l'opéron *pts*. En effet, les gènes *ptsH* et *ptsI*, codant les enzymes générales du système PTS, sont exprimés à partir du promoteur *ptsH* P1 assurant une expression de base de l'opéron et d'un promoteur régulé en fonction des conditions de croissance *ptsH* P0 (voir Introduction, Partie E.II.4.a).

Etant donné que l'ORF fbaA d'E. coli est transcrite à partir des mêmes promoteurs que l'ORF pgk et du fait de la présence d'un terminateur rho-indépendant potentiel pgk T1 localisé en aval de l'ORF pgk, on devrait s'attendre à une plus faible transcription de l'ORF fbaA que de l'ORF pgk. Or, les protéines PGK et FBA sont produites en quantité comparables chez E. coli qu'elle soit non transformée (Peng et Shimizu, 2003) ou transformée par les constructions pEPF et pΔP0-EPF (Fig. 3 de l'article 1). En complément du travail décrit dans l'article, nous nous sommes donc intéressés aux mécanismes qui permettraient de compenser la plus faible transcription de l'ORF fbaA par rapport à l'ORF pgk. Une différence marquée d'efficacité de traduction était peu vraisemblable, vue l'utilisation optimale des codons dans chacune des 2 ORF et les caractéristiques canoniques des séquences TIR des 2 ORF (séquence GAGG et AGGA de 4 nts séparé de 7 et 8 nts en distance corrigée de l'AUG initiateur). Une hypothèse raisonnable pour expliquer la production de taux équivalent de PGK et de FBA était une différence de stabilité des ARNm pgk par rapport aux ARNm fbaA. Nous avons testé cette hypothèse.

### I / Etude de la stabilité relative des ARNm monocistroniques pgk et fbaA chez E. coli

Pour étudier la stabilité relative des ARNm pgk et fbaA, nous avons utilisé la souche d'E. coli SK5665 transformée par le plasmide pEPF et nous l'avons cultivée en milieu LB, à 30°C. En effet, V. Bardey, N. Robas et B. Charpentier avaient montré précédemment que le promoteur epd P0 est faiblement utilisé en milieu LB. Dans ces conditions, la composition en ARNm était donc moins

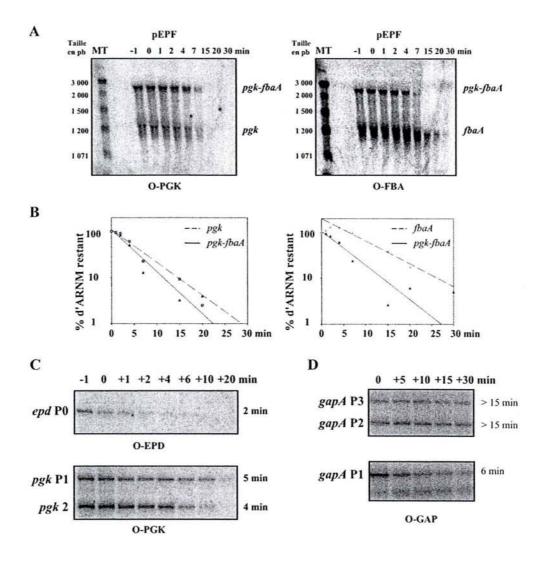


Figure 62 : Détermination de la stabilité relative des transcrits pgk et fbaA chez E. coli.

A : Analyse de la durée de vie des transcrits pgk et fbaA par la méthode Northern-blot. La souche SK5665 transformée par le plasmide pEPF a été cultivée en milieu LB en présence d'ampicilline à 30°C jusqu'en milieu de phase exponentielle de croissance. La rifampicine a été ajoutée à la culture au temps 0, à une concentration finale de 0,5 mg/ml. Les échantillons de 11 ml de culture ont été prélevés avant (-1 min) ou après l'ajout de la rifampicine. Les ARN totaux des différents échantillons ont été extraits, puis 15 μg d'ARN ont été fractionnés en gel d'agarose-formaldéhyde 0,8 %. Un marqueur de taille d'ADN (100 bp DNA Ladder plus mixture de MBI Fermentas) radiomarqué en 5' a été fractionné en parallèle. Les ARN ont ensuite été transférés sur une membrane de nylon, puis hybridés aux oligonucléotides radiomarqués O-PGK et O-FBA.

B : Estimation de la demi-vie des transcrits. La radioactivité contenue dans les bandes des gels du panneau A a été quantifiée à l'aide d'un Phosphorimager, et l'évolution du pourcentage d'ARN résiduel a été représenté en fonction du temps.

C: Analyse par la méthode de transcription inverse de la stabilité des transcrits issus des promoteurs *epd* P0 et *pgk* P1 et du produit de maturation *pgk* 2. Les ARN issus de la culture bactérienne réalisée en A ont été analysés par extension d'amorce avec la transcriptase inverse en utilisant les oligonucléotides radiomarqués O-EPD et O-PGK. Les ADNc ont ensuite été séparés par électrophorèse en gel de polyacrylamide à 5 %.

D : Analyse par extension d'amorce de la stabilité des transcrits du gène *gapA* réalisée par la même approche par B. Thouvenot. Les cellules TG1 ont été cultivées en milieu LB à 30 °C jusqu'en milieu de phase exponentielle, puis traitées à la rifampicine (0,2 mg/ml). Les échantillons ont été prélevés à différents temps. Les ARN extraits des différents échantillons ont ensuite été analysés par élongation d'amorce à la transcriptase inverse avec l'oligonucléotide O-GAP radiomarqué.

faiblement utilisé en milieu LB. Dans ces conditions, la composition en ARNm était donc moins complexe du fait du faible taux de production des transcrits epd-pgk-fbaA et epd-pgk. Nous nous attendions essentiellement à avoir des transcrits pgk-fbaA et pgk et les produits de clivage des transcrits pgk-fbaA. La souche d'E. coli transformée a été cultivée jusqu'en milieu de phase exponentielle de croissance (A<sub>600nm</sub> de 0,5). Nous avons ensuite ajouté de la rifampicine pour bloquer la synthèse de nouveaux ARN. Cet antibiotique bloque spécifiquement l'ARN polymérase bactérienne, en se fixant au niveau de la sous-unité β, à proximité du site catalytique (Campbell et al., 2001). Il inhibe l'élongation des transcrits dès que leur longueur atteint deux à trois nucléotides. Dans ces conditions, la concentration cellulaire des différents transcrits évolue au cours du temps uniquement en fonction de leur vitesse de dégradation par les RNases, ce qui permet de calculer leur demi-vie. Pour établir cette demi-vie, nous avons prélevé des échantillons à des temps définis après l'ajout de rifampicine (temps 0). Les ARN de ces différents échantillons ont ensuite été analysés par Northern-Blot, en utilisant comme sondes les oligonucléotidiques marqués O-PGK et O-FBA (Fig. 62, A et B). La quantité de radioactivité correspondant à chaque transcrit a été déterminée à l'aide d'un Phosphorimager (Amersham Biosciences), en utilisant le logiciel ImageQuant (version 5.2). La demivie apparente déterminée pour le transcrit bicistronique pgk-fbaA (de l'ordre de 4 minutes) était similaire à celle déterminée pour le transcrit pgk, alors que la demi-vie apparente du transcrit fbaA était trois fois plus importante, de l'ordre de 12 min. Les demi-vies des ARNm monocistronique fbaA et pgk sont dites apparentes car les transcrits monocistroniques pgk proviennent à la fois des transcrits initiés au promoteur pgk P1 et se terminant au terminateur pgk T1 et des transcrits pgk-fbaA clivés par la RNase E dans la région intergénique. Les transcrits monocistroniques fbaA résultent eux uniquement du clivage du transcrit bicistronique pgk-fbaA. Nous avons alors déterminé la demi-vie réelle du transcrit fbaA en utilisant le modèle mathématique proposé par Smolke et Keasling (2000). Ainsi, nous avons:

$$d[pgk-fbaA] / dt = -k_{dpgk-fbaA}[pgk-fbaA]$$
 (1)

$$d[fbaA] / dt = -(k_{dpgk-fbaA}[pgk-fbaA] - k_{dfbaA}[fbaA])$$
 (2)

La concentration du transcrit bicistronique pgk-fbaA est représentée par [pgk-fbaA], celle du transcript monocistronique fbaA par [fbaA]. Les constantes de vitesse de dégradation sont k<sub>d pgk-fbaA</sub> pour le transcrit bicistronique pgk-fbaA et k<sub>d fbaA</sub> pour le transcript monocistronique fbaA. La résolution des équations conduit à une évaluation de la demi-vie réelle du transcrit fbaA de l'ordre de 6 minutes.

Cette plus forte stabilité des ARNm fbaA monocistroniques par rapport aux ARNm pgk monocistroniques peut s'expliquer par la présence d'éléments stabilisateurs tels que les structures tigeboucle en amont et en aval de l'ARNm fbaA. En effet, la structure secondaire que nous avons établie en solution pour la région intergénique pgk-fbaA montre que la région située entre le site de clivage fbaA 2 et le codon d'initiation de la traduction de l'ORF fbaA se structure en tige-boucle (SL10, Fig.

dépendant au niveau de la répétition de type BIME-1 (voir Introduction, Partie B.III). Dans tous les cas, des éléments structurés sont présents en 3' de l'ARNm fbaA et doivent aussi contribuer à sa stabilité. Les transcrits pgk issus d'un arrêt au site potentiel de terminaison de la transcription pgk T1 ont eux aussi une structure tige-boucle stable en 3' et c'est eux qui doivent avoir une demi-vie d'environ 4 min. Par contre, les ARNm monocistroniques pgk issus du clivage à la série de sites fbaA 5 sont vraisemblablement peu stables du fait de l'absence du codon de terminaison dans certains ou d'une coupure proche de ce codon de terminaison (Fig. 8A de l'article 1).

En complément de l'analyse par Northern présentée dans la Figure 62, nous avons réalisé sur les mêmes échantillons d'ARN une étude par extension d'amorce en utilisant les oligonucléotides O-EPD et O-PGK. Nous obtenons des résultats concordants pour les ARNm pgk avec des demi-vies de respectivement 5 et 4 minutes pour les ARNm pgk P1 et pgk 2. Dans ces conditions, la durée de demi-vie du fragment 5' terminal de l'ARN initié au promoteur epd P0 est de 2 minutes. B. Thouvenot avait réalisé une expérience similaire sur les ARNm gapA issus d'une culture de la souche TG1 en milieu LB et soumis à un ajout de rifampicine. De manière intéressante, les ARNm gapA P1 se sont avérés avoir une stabilité de 6 minutes alors que les ARNm gapA P2 et P3 respectivement initiés aux promoteurs gapA P2 et P3 étaient eux les plus stables, avec une durée de vie supérieure à 15 minutes (Fig. 62).

Notre observation d'une plus forte stabilité de l'ARNm fbaA par rapport à l'ARNm pgk est en accord avec les données de l'équipe de Cohen (Bernstein et al., 2002). Les demies-vies estimées par ces auteurs par l'approche puces à ADN sont très similaires pour l'ARNm pgk (4,7 minutes) par contre, ces auteurs proposent une stabilité plus grande pour l'ARNm fbaA (8,1 minutes) (Bernstein et al., 2002). Ces auteurs ont aussi réalisé une étude en milieu minimum M9 supplémenté en glucose. Ils obtiennent alors une demi-vie de 2.5 et 7.2 minutes pour les transcrits renfermant respectivement pgk et fbaA, et d'environ 2.8 minutes pour les transcrits contenant epd. Il faut remarquer que ces demies-vies sont apparentes puisque par l'approche puces à ADN, ce qui est mesuré est la globalité des ARNm epdpgk-fbaA, epd-pgk, pgk-fbaA et pgk ainsi que certains produits de dégradation pour pgk et les ARNm epd-pgk-fbaA, pgk-fbaA et fbaA intacts et éventuellement d'autres produits de dégradation pour fbaA. En effet, comme nous l'avons observé pour l'ORF epd, des clivages internes peuvent faire que certaines parties des transcrits sont plus stables que d'autres (forte stabilité de la région 5' d'epd alors que le reste est dégradé). De ce fait, en fonction des oligonucléotides ou fragment PCR utilisés pour construire les puces à ADN, on peut très bien avoir une réponse positive pour une ORF, qui correspond en fait à l'existence d'un petit fragment très résitant. Ceci explique sans doute les résultats que les mêmes auteurs ont obtenus pour les transcrits contenant l'ORF epd. La confrontation de nos données détaillées sur l'opéron epd-pgk-fbaA à celles obtenues globalement par l'approche puces à ADN montre les limites de l'approche puces à ADN et l'importance d'études approfondies, telles que celle que nous avons fait. En tout cas, un point commun à nos données et celles de l'équipe de Cohen que les mêmes auteurs ont obtenus pour les transcrits contenant l'ORF epd. La confrontation de nos données détaillées sur l'opéron epd-pgk-fbaA à celles obtenues globalement par l'approche puces à ADN montre les limites de l'approche puces à ADN et l'importance d'études approfondies, telles que celle que nous avons fait. En tout cas, un point commun à nos données et celles de l'équipe de Cohen est la plus forte stabilité des ARNm fbaA, correspondant à la dernière ORF de l'opéron. Ce phénomène de plus forte stabilité de la dernière ORF en général est lié à la plus faible transcription du fait d'arrêt de transcription assez grand (Selinger et al., 2003).

## II / Poursuite de l'étude phylogénique des opérons *epd-pgk-fbaA* en rapport avec les caractéristiques mises en évidences précédemment

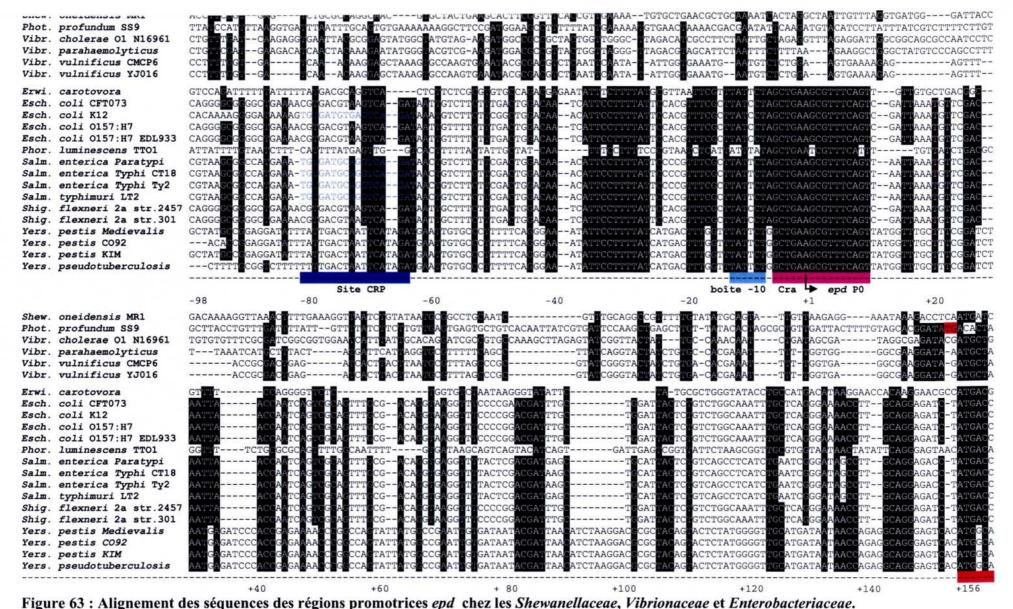
En parallèle de cette étude sur *E. coli*, j'ai poursuivi l'étude phylogénique sur les γprotéobactéries décrite ci-dessus en tenant compte des résultats décrits dans l'article. En effet, vu
l'importance des 2 promoteurs *epd* P0 et *pgk* P1, il était intéressant de rechercher s'ils étaient très
conservés chez les γ-protéobactéries, renfermant une ORF *epd*. Il était aussi intéressant de savoir si la
présence d'un terminateur de la transcription de type rho-indépendant était conservée en aval de l'ORF *pgk*. Enfin, nous avons vu que la faible expression de l'ORF *epd* chez *E. coli* est en partie due à la
distance élevée entre la séquence SD et l'AUG initiateur au sein de l'élément TIR. Là aussi, la
question se posait de savoir si ce mode de régulation est spécifique à *E. coli*, ou si comme le suggérait
les premières comparaisons faites par V. Bardey, cette régulation était retrouvée chez les autres
espèces bactériennes ayant une ORF *epd*.

#### II.1 / Etude de la conservation des régions promotrices epd PO et pgk P1

Parmi les espèces bactériennes où une ORF epd est mise en évidence (Tableau 4), nous n'observons une conservation de la région située en amont de l'ORF epd que chez les Enterobacteriaceae, ceci à l'exception de Photorhabdus luminescens, espèce dans laquelle de grandes séquences conservées sont observées en amont de l'ORF epd (Fig. 63). Il est à noter que chez Yersinia, un promoteur de type étendu peut être proposé (TGGTTTATT). Il est situé légèrement en amont du promoteur présent chez E. coli. Le site de fixation pour la protéine Cra est très fortement conservé chez toutes les Enterobacteriaceae, la bactérie P. luminescens présente néanmoins une variation au niveau des résidus +2 et +10 de cette séquence. Le site de fixation du complexe CRP-AMPc situé en amont du promoteur epd P0 est faiblement conservé. Cette faible conservation est à relier au faible rôle de ce site de fixation pour la régulation du promoteur epd P0 chez E. coli

| Espèces bactériennes  |           |
|---|-----------|
| Erwinia carotovora subsp. atroseptica SCRI1043                        | NC_004547 |
| Escherichia coli K12  | NC_000913 |
| Escherichia coli CFT073   | NC_004431 |
| Escherichia coli O157:H7  | NC_002695 |
| Escherichia coli O157:H7 EDL933                                       | NC_002655 |
| Photobacterium profundum SS9  | NC_006370 |
| Photorhabdus luminescens subsp. laumondii TTO1                        | NC_005126 |
| Salmonella enterica subsp. enterica serovar Paratypi A str. ATCC 9150 | NC_006511 |
| Salmonella enterica subsp. enterica serovar Typhi str. CT18           | NC_003198 |
| Salmonella enterica subsp. enterica serovar Typhi Ty 2                | NC_004631 |
| Salmonella typhimurium LT2  | NC_003197 |
| Shewanella oneidensis MR-1  | NC_004347 |
| Shigella flexneri 2a str. 2457T                                       | NC_004741 |
| Shigella flexneri 2a str. 301   | NC_004337 |
| Vibrio cholerae O1 biovar eltor str. N16961                           | NC_002505 |
| Vibrio parahaemolyticus RIMD 2210633                                  | NC_004603 |
| Vibrio vulnificus CMCP6   | NC_004459 |
| Vibrio vulnificus YJ016   | NC_005139 |
| Yersinia pestis biovar Medievalis str. 91001                          | NC_005810 |
| Yersinia pestis CO92  | NC_003143 |
| Yersinia pestis KIM   | NC_004088 |
| Yersinia pseudotuberculosis IP 32953                                  | NC_006155 |

 $Tableau\ 5: Liste \ des\ génomes\ utilisés\ lors\ de\ l'étude\ de\ la\ conservation\ des\ caractéritiques\ mises\ en\ évidence\ pour\ l'opéron\ epd-pgk-fbaA\ \ d'E.\ coli\ .$ 



Les séquences ont été alignées en utilisant le programme ClustalW. Les tirets indiquent les délétions introduites pour optimiser l'alignement. Les bases indiquées en blanc sur fond noir sont les plus conservées. La numérotation est indiquée par rapport au site +1 d'initiation de la transcription epd P0 chez E.

coli. La boîte -10 du promoteur epd P0 est représentée par un rectangle bleu turquoise. Les positions des sites de fixation du complexe CRP-AMPc et de la protéine Cra sont respectivement indiquées par des rectangles bleu foncés et violets. La séquence correspondant à l'ORF epd est indiquée par un rectangle

rouge. Les génomes utilisés sont répertoriés dans la Figure 65. Les génomes utilisés ont été répertoriés dans le tableau 5.

(Charpentier *et al.*, 1998; Bardey, 2000). La région autour de la position -35 (résidu -24 à -43) est fortement conservée, sans pour autant correspondre à une boîte -35 classique de facteur σ. Cette conservation suggère un rôle important de cette séquence, peut être la fixation d'un facteur d'activation.

L'alignement des régions 3' des ORF epd contenant chez E. coli le promoteur pgk P1, révèle une forte conservation du promoteur étendu pgk P1 chez les Enterobacteriaceae (Fig. 64). Cette forte conservation est supérieure à ce que l'on peut s'attendre à avoir même pour une région codante, et donc, l'utilisation du promoteur pgk P1 doit être conservée chez ces bactéries. Les Shewanellaceae et les Vibrionaceae présentent par contre des variations au niveau de la boîte -10.

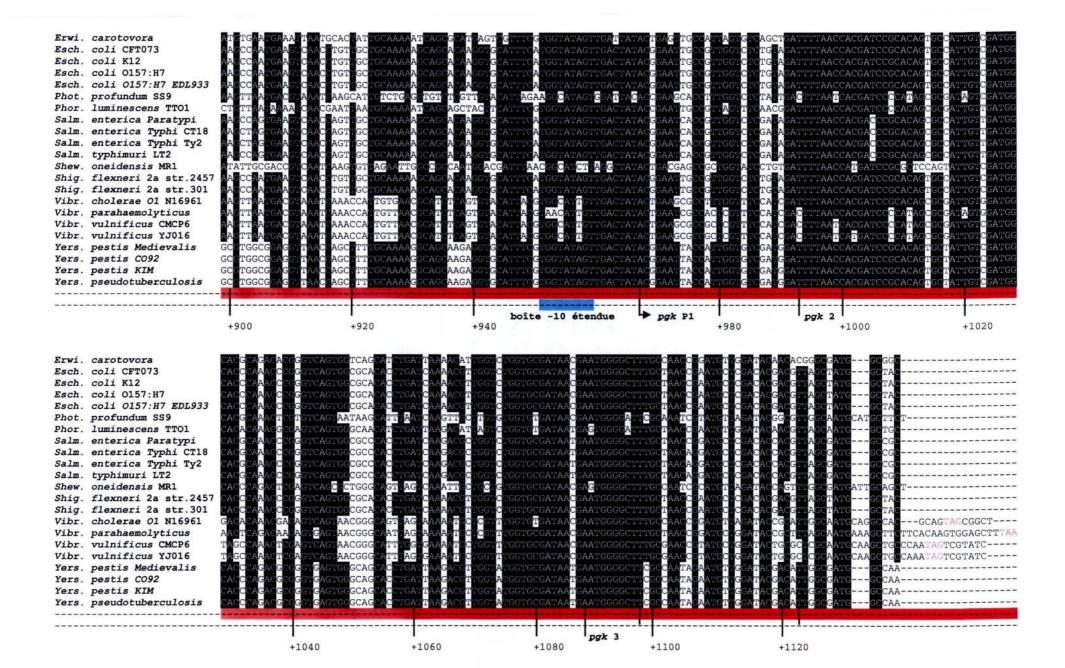
Globalement, nous observons donc une plus forte conservation du promoteur étendu pgk P1 que du promoteur epd P0. L'ORF epd est sans doute transcrite différemment chez les Enterobacteriaceae, les Vibrionaceae et les Shewanellaceae, ce qui est en accord avec le fait que V. Bardey avait mis en évidence des sites d'initiation de la transcription du tandem epd-pgk différents chez E. coli K12 et V. cholerae (Bardey, 2000).

#### II.2 / Conservation du terminateur potentiel rho-indépendant pgk T1

L'alignement des séquences situées en aval du codon de terminaison de l'ORF pgk permet de distinguer trois groupes d'espèces bactériennes (Fig. 65) : celles présentant une séquence homologue à celle du terminateur pgk T1 : il s'agit des bactéries des genres Escherichia, Salmonella et Shigella ; celles ayant un terminateur potentiel ayant une autre séquence et qui appartient aux genres Yersinia, Photobacterium et Vibrio ; enfin, celles n'ayant pas de terminateur potentiel (genres Erwinia, Photorhabdus et Shewanella). Ainsi, la présence d'un terminateur en aval de l'ORF pgk est fortement conservée chez Escherichia, Salmonella et Shigella d'une part, et Yersinia, Photobacterium et Vibrio d'autre part. Dans les bactéries du genre Erwinia, Photorhabdus et Shewanella, aucun arrêt de la transcription n'est attendu à ce niveau.

#### II.3 / Conservation des sites d'entrée interne de la RNase E

L'alignement des séquences des régions des ORF epd correspondant à la région de l'ORF epd d'E. coli contenant les sites de clivages par la RNase E (pgk 2 et 3) montre que ces sites de clivage sont relativement bien conservés chez toutes les espèces bactériennes étudiées (Fig. 64). Par contre, les séquences des sites de clivages pgk 4 à 6, situés en aval sont peu conservées chez les bactéries des genres Shewanella, Photorhabdus et Vibrio. Une part de la conservation des séquences des sites 2 et 3



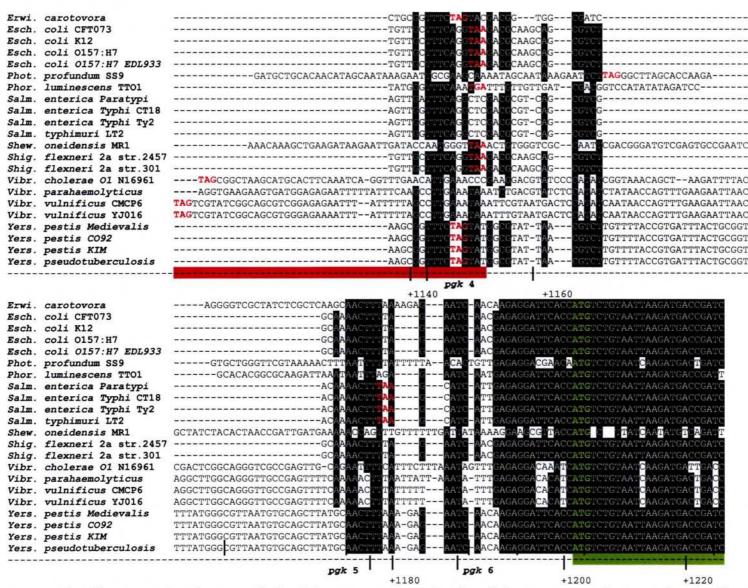


Figure 64 : Alignement des séquences de la région promotrice pgk et des régions correspondant aux sites de clivages pgk 2 à 6 chez les Enterobacteriaceae, Shewanellaceae et Vibrionaceae.

Les séquences ont été alignées en utilisant le programme ClustalW. Les tirets indiquent des délétions introduites pour optimiser l'alignement. Les bases indiquées en blanc sur fond noir sont les plus conservées. Le codon stop de l'ORF epd est indiqué en rouge, le codon initiateur de l'ORF pgk en vert. Les ORF epd et pgk sont respectivement indiquées par des rectangles rouge et vert. La numérotation est établie par rapport au site d'initiation de la transcription du promoteur epd P0 chez E. coli K12. Les génomes utilisés ont été répertoriés dans le tableau 5.

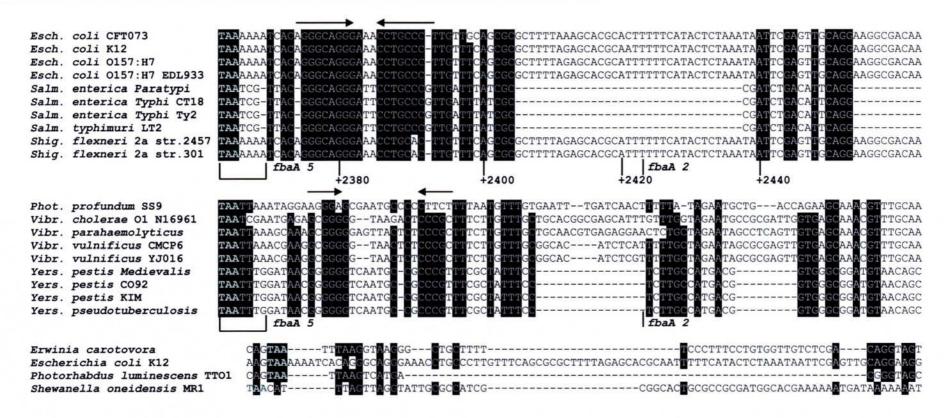


Figure 65 : Conservation de séquences dans la région intergénique pgk-fbaA.

Les régions situées en aval du codon de terminaison de l'ORF pgk, indiqué en vert, ont été alignées par le logiciel clustalW. Les tirets indiquent les délétions introduites pour optimiser l'alignement. Le codon stop de l'ORF pgk est en vert. Les séquences indiquées en blanc sur fond noir sont les plus conservées dans les différentes espèces étudiées. Les flèches en sens opposées représentent le terminateur potentiel pgk T1 rhoindépendant.

Les génomes utilisés ont été répertoriés dans le tableau 5.

peut être due au fait qu'ils sont inclus dans une région codante. Néanmoins, le degré de conservation chez les γ-protéobactéries de la séquence du site pgk 2 est supérieur à ce que l'on peut s'attendre à avoir, même dans une région codante, ce qui est significatif d'une importance fonctionnelle de ce site. Au niveau de la région interORF pgk-fbaA, la séquence renfermant la série de sites fbaA 5 est conservée : chez les bactéries des genres Escherichia, Shigella, Yersinia et Vibrio, à l'exception de Vibrio cholerae et Photobacterium profundum (Fig. 65). Par contre, la séquence contenant le site fbaA 2 est uniquement conservée chez les bactéries des genres Escherichia et Shigella. Donc, seuls les sites situés les plus en amont et considérés comme des sites d'entrée interne de la RNase E sont des séquences conservées dans plusieurs des genres bactériens.

#### II.4 / Sites d'initiation de la traduction

Nous avons montré dans l'article 1 que la taille faible de l'écartement corrigé entre le site de fixation au ribosome (RBS) et le codon initiateur est sans doute un facteur défavorable pour la traduction de l'ORF epd (Tableaux 4 et 5 de l'article). L'alignement des régions amont des ORF epd identifiées précédemment (Tableau 4) montre que comme pour E. coli, un écartement corrigé sousoptimal, de l'ordre de 5 nucléotides voire 3 nucléotides est observé pour les ORF epd des bactéries de la famille des Vibrionaceae (Fig. 66). Ainsi, une limitation de la traduction par un faible écartement entre la séquence SD et le codon initiateur AUG pourrait être une propriété commune des ORF epd.

Chez toutes les bactéries étudiées, l'écartement corrigé dans le cas des ORF pgk et fbaA est optimal, de l'ordre de 8 et 9 nucléotides (Fig. 66).

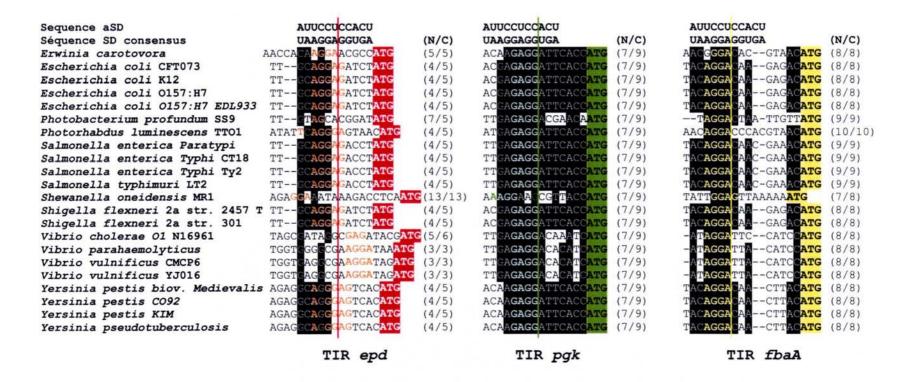


Figure 66 : Etude de l'écartement entre la séquence SD et le codon d'initiation de la traduction des régions TIR des ORF epd, pgk et fbaA chez les  $\gamma$ -protéobactéries.

Les régions TIR des ORF epd, pgk et fbaA ont été alignées par le logiciel ClustalW. Les séquences indiquées en blanc sur fond noir sont les plus conservées chez les différentes espèces étudiées. La séquence SD des ORF epd, pgk et fbaA sont respectivement indiquées en rouge, vert et jaune. Les codons initiateurs des ORF epd, pgk et fbaA sont surlignés respectivement en rouge, vert et jaune. L'écartement non corrigé (N) et l'écartement corrigé (C) définis selon Ma et al. (2002) sont indiqués entre parenthèse, à droite de chaque région TIR. Les génomes utilisés sont répertoriés dans le tableau 5.

## III / Etude visant à tester si la protéine Hfq peut moduler l'expression des gènes epd, pgk et fbaA

Comme nous l'avons vu dans l'introduction, la protéine Hfq peut agir en tant que régulateur post-transcriptionnel (pour revue, Valentin-Hansen et al., 2004). Elle peut interagir directement avec l'ARNm et inhiber sa traduction, comme c'est le cas pour l'ARNm ompA (Vytvytska et al., 2000; Moll et al., 2003a; Moll et al., 2003b). Elle peut aussi favoriser l'interaction d'un ARN régulateur avec son ARNm cible. L'ARN régulateur peut ensuite favoriser ou inhiber la traduction (voir Partie D.I.5 de l'Introduction). Par ailleurs, du fait d'une similarité de séquence des sites reconnus par Hfq et la RNase E, la protéine Hfq peut se fixer au niveau de sites de clivage par cette enzyme, ce qui limite son activité et augmente la stabilité de l'ARN (Moll et al., 2003b). Dans d'autres cas, la protéine Hfq peut recruter la poly (A) polymérase et favoriser l'allongement des queues poly (A). Elle inhibe alors l'action des exonucléases se liant aux séquence poly (A) (Hajnsdorf et Regnier, 2000; Folichon et al., 2003). La protéine Hfq interagit aussi avec l'ARN polymérase et la protéine ribosomique S1, dont elle pourrait moduler l'activité.

Par analyse du protéome de cellules d'*E. coli* cultivées en milieu minimum contenant du glucose, il a été récement montré que l'expression de la protéine Hfq augmente la production de GAPDH (Vecerek *et al.*, 2003). Cet effet n'est pas direct car des études par traduction *in vitro* montrent qu'une forte concentration en protéine Hfq inhibe la traduction de l'ORF *gapA*. Ces observations nous ont conduit à nous intéresser à l'influence possible de la protéine Hfq sur l'expression de l'opéron *epd-pgk-fbaA*. Une action possible de Hfq était d'autant plus vraisemblable que les transcrits de cet opéron comportent des sites de clivages par la RNase E.

Comme nous allons le décrire, nous avons dans un premier temps testé si la protéine Hfq forme in vitro des complexes avec des ARN correspondant à la région intergénique epd-pgk (+896 à +1222) et à la région intergénique pgk-fbaA (+2264 à +2626), ceci par emploi de la technique de retard sur gel. Des complexes étant en effet formés, nous avons étudié leurs stabilités relatives, puis identifié le site de fixation de la protéine Hfq sur le segment epd-pgk. Ayant détecté des interactions in vitro, nous avons ensuite débuté une étude in vivo de l'influence de la protéine Hfq sur l'expression de l'ORF gapA et du triplet d'ORF epd-pgk-fbaA, ceci grâce à une souche hfq qu'E. Hajnsdorf nous a fournie en fin de thèse.

### III.1 / In vitro, la protéine Hfq se fixe au niveau des régions epd-pgk et pgk-fbaA

Comme la protéine Hfq est connue pour interagir avec les sites de clivages par la RNase E,

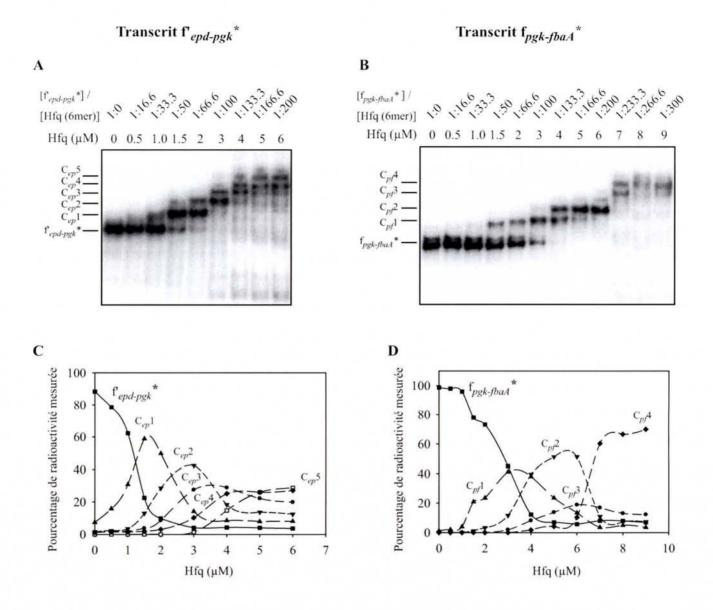


Figure 67: La protéine Hfq se fixe au niveau des régions epd-pgk et pgk-fbaA in vitro.

A et B : Expériences de gel retard réalisées avec les transcrits f'epd-pgk et fpgk-fbaA et la protéine Hfq. Les deux transcrits uniformément marqués au <sup>32</sup>P (6 ng), ont été incubés pendant 10 min à 37°C, en présence de 20 ng d'ARNt, et de concentrations croissantes en protéine Hfq, dans le tampon suivant : MgCl<sub>2</sub> 10 mM ; NaCl 50 mM ; KCl 50 mM et Tris-HCl 10 mM, pH 8.0. Les complexes formés ont été fractionnés par électrophorèse en gel de polyacrylamide non dénaturant à 6 % pendant 3 heures à 6 V/cm. Les différentes concentrations en protéine Hfq, ainsi que l'excès molaire en hexamère Hfq par rapport à l'ARN d'intérêt sont indiqués en haut des pistes. Le transcrit libre radiomarqué et le ou les complexes formés avec la protéine Hfq sont numérotés à gauche de chaque gel.

C et D : Quantification des différents complexes formés aux différentes concentrations en protéine Hfq. La radioactivité présente dans les bandes de gels illustrées dans les panneaux A et B, a été quantifiée à l'aide d'un PhosphorImager.

nous nous sommes intéressés à la fixation de cette protéine au niveau des sites de clivages mis en évidence dans l'article 1, c'est-à-dire dans les régions epd-pgk (+896 à +1222) et pgk-fbaA (+2264 à +2626). J'ai utilisé une protéine Hfq recombinante, produite et purifiée par B. Thouvenot et B. Charpentier au laboratoire (Thouvenot, 2004). Avec D. Hartemann, nous avons tout d'abord synthétisé des ARN uniformément marqué par incorporation d'ATP[α-32P], ceci en utilisant les matrices ADN f'epd-pgk et fpgk-fbaA. La matrice ADN fpgk-fbaA a été décrite précédemment dans l'article 1 (Fig. 1). La matrice ADN f'epd-pgk contient un promoteur reconnu par l'ARN polymérase T7, ainsi que la région epd-pgk s'étendant des positions +896 à +1222, le site +1 étant le site d'initiation de la transcription par le promoteur epd P0. Elle produit un ARN plus court que le transcrit  $f_{epd-pgk}$  décrit dans l'article 1. Les transcrits ainsi obtenus, f'epd-pgk et fpgk-fbaA, ont respectivement 327 et 321 nt. Ils ont été renaturés comme décrit dans l'article 1, puis des échantillons (6 ng) ont été incubés 10 min à 37°C dans le tampon MgCl<sub>2</sub> 10mM; NaCl 50 mM et KCl 50 mM; Tris-HCl 10 mM, pH 8.0 avec différentes concentrations de protéines Hfq, allant de 0.5 à 9.0 µM, et en présence de 20 ng d'ARNt. Les ARNt compétiteurs permettaient de limiter les risques de fixation non spécifique de la protéine Hfq sur les ARN radiomarqués présents à faible concentration. Les complexes ARN-protéine formés ont été fractionnés en gel de polyacrylamide non dénaturant à 6 %. Lorsqu'une protéine se fixe sur un ARN, elle entraîne un retard de migration de cet ARN lors d'une électrophorèse en gel non dénaturant. La protéine Hfq est connue pour former des hexamères (Arluison et al., 2002; Sauter et al., 2003) et ce sont les hexamères qui reconnaissent l'ARN (Mikulecky et al., 2004), l'ajout séquentiel de chaque hexamère entraîne un nouveau retard sur gel. Quand plusieurs complexes sont mis en évidence sur le gel, ces complexes peuvent correspondre à la fixation d'hexamères sur différents sites ou à l'association de plusieurs hexamères au niveau d'un site donné.

Un grand nombre de complexes (C<sub>ep</sub>1, C<sub>ep</sub>2, C<sub>ep</sub>3 et C<sub>ep</sub>4 et C<sub>ep</sub>5) ont été mis en évidence aussi bien pour le transcrit f'<sub>epd-pgk</sub> (Fig. 67A), que pour le transcrit f<sub>pgk-fbaA</sub> (C<sub>pf</sub>1, C<sub>pf</sub>2, C<sub>pf</sub>3 et C<sub>pf</sub>A) (Fig. 67B). Nous avons quantifié la quantité de radiactivité contenue dans chaque bande de gel à l'aide d'un Phosphorimager (Fig. 67, C et D), ce qui nous a permis de représenter de manière graphique et quantitative les résultats obtenus. Vu le chevauchement de la formation des différents complexes, il est seulement possible d'estimer approximativement le K<sub>D</sub> pour le premier complexe formé pour chacun des deux ARN. Il est de l'ordre de 1,2 μM pour l'ARN f'<sub>epd-pgk</sub>, et de 3 μM pour l'ARNm f<sub>pgk-fbaA</sub>. Il ne s'agit donc pas d'affinité très forte, néanmoins elles peuvent être suffisantes pour moduler l'activité RNase E si ces sites de fixation de Hfq correspondent à des sites de clivages par la RNase E.

Nous avons ensuite étudié la stabilité relative des sites de fixation mis en évidence, ceci en réalisant des expériences de compétition avec différents ARN froids. Tout d'abord, nous avons fait varier la concentration en ARNt compétiteur (de 20 ng à 400 ng), soit un rapport molaire ARNt/ARN d'intérêt de 10 à 200. La concentration choisie pour la protéine Hfq était de 3 μM pour l'essai avec l'ARN f'<sub>epd-pgk</sub> et de 6 μM pour celui de l'ARN f<sub>pgk-fbad</sub>. Ces concentrations ont été choisies car elles

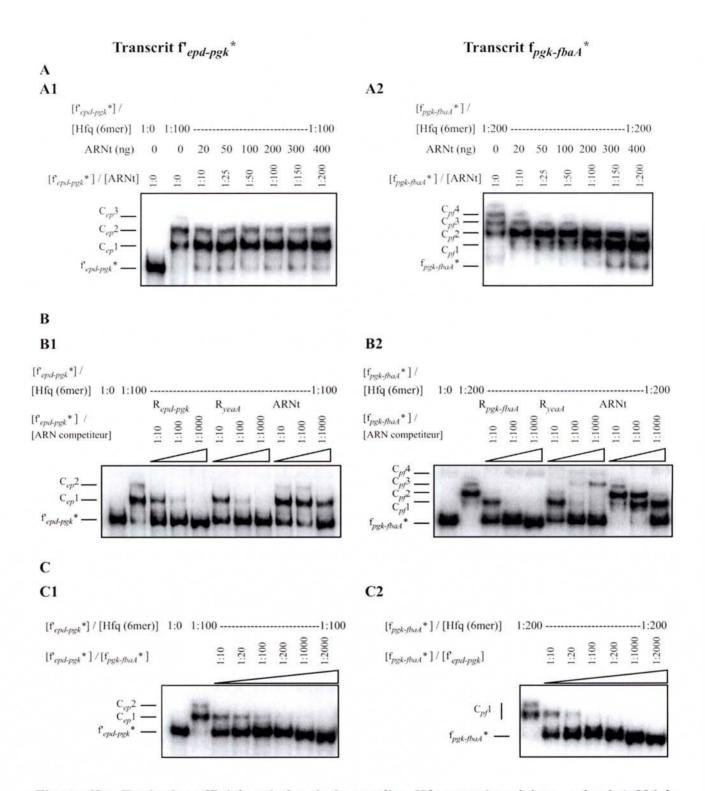


Figure 68 : Etude des affinités relative de la protéine Hfq pour les régions *epd-pgk* (+896 à +1922) et *pgk-fbaA* (+2264 à +2626), et pour la région à de l'ARNm *yeaA* (+1 à +123).

Des expériences de compétition ont été réalisées en utilisant la protéine Hfq, les transcrits uniformément marqués f'epd-pgk\* et fpgk-fbaA\* et soit des ARNt de levure (A1, A2, B1 et B2) soit des transcrits froids f'epd-pgk ou fpgk-fbaA (B1, B2, C1 et C2) ou le transcrit fyeaA (Thouvenot, 2004) (B1 et B2). Les ARN compétiteurs ont été ajoutés aux différentes concentrations indiquées en haut de chaque piste, dans un mélange contenant le transcrit radiomarqué f'epd-pgk\* ou fpgk-fbaA\*. Dans les différents tests, la protéine Hfq a été respectivement ajoutée aux concentrations de 3 μM (A1, B1 et C1) et 6 μM (A2, B2 et C2) dans le tampon suivant : MgCl<sub>2</sub> 10 mM ; NaCl 50 mM ; KCl 50 mM et Tris-HCl 10 mM pH 8.0 pendant 10 min à 37°C. Les complexes formés ont été fractionnés par électrophorèse en gel de polyacrylamide non dénaturant à 6 %, à 6 V/cm.

permettent respectivement la formation des complexes de plus forte affinité  $C_{ep}1$  et  $C_{ep}2$  pour l'ARN  $f_{epd-pgk}$  et  $C_{pf}1$  et  $C_{pf}2$  pour l'ARN  $f_{pgk-fbaA}$  (Fig. 67). L'incubation a été réalisée à 37°C pendant 10 min dans le même tampon que dans les expériences décrites ci-dessus. L'ajout d'un grand excès d'ARNt au complexe  $f'_{epd-pgk}$ -Hfq ne modifie que très légèrement le taux de formation des complexes  $C_{ep}1$  et  $C_{ep}2$  (Fig. 68A). Par contre, les plus nombreux complexes formés avec l'ARN  $f_{pgk-fbaA}$  se dissocient progressivement. Seuls les complexes  $C_{pf}1$  et  $C_{pf}2$  à un moindre degré sont stables.

Nous avons ensuite essayé de comparer l'affinité d'Hfq pour les 2 sites  $C_{ep}1$  et  $C_{pf}1$  et pour le site que B. Thouvenot avait détecté dans l'ARNm yeaA d'E. coli (Thouvenot, 2004). Pour cela, différents transcrits non radioactifs ont été préparés et utilisés comme ARN compétiteur dans les expériences de retard sur gel. Le transcrit f<sub>yeaA</sub> utilisé, en plus des transcrits f'<sub>epd-pgk</sub> et f<sub>pgk-fbaA</sub> dans ces expériences, correspond à la région de l'ARNm yeaA allant de la position +1 (site d'initiation à partir du promoteur yeaA P1) à la position +123. Dans l'expérience représentée dans la Figure 69B1, on voit que l'effet compétiteur du transcrit f<sub>yea</sub> sur la formation du complexe f'<sub>epd-pgk</sub>-Hfq est similaire à celui de l'ARN f'epd-pgk, ce qui suggère une affinité similaire de Hfq pour ces deux ARN. L'expérience contrôle avec les ARNt est une vérification de la validité des résultats obtenus. En ce qui concerne l'ARN f<sub>pgk-fbaA</sub>, un excès de 10 fois de l'ARN f<sub>yea</sub> dissocie tous les complexes de grande taille, mais là aussi un excès de 100 fois est requis pour dissocier le complexe C<sub>pf</sub>1. Nous en concluons qu'Hfq a une affinité du même ordre de grandeur pour les trois ARN. Les affinités relatives d'Hfq pour les régions epd-pgk et pgkfbaA ont enfin été testées par des expériences de compétitions croisées (Fig. 68C). L'ARN f'epd-pgk a sans doute une affinité supérieure à celle de l'ARN fpgk-fbaA, car la concentration en ARN fpgk-fbaA nécessaire pour dissocier le complexe formé avec l'ARN f'epd-pgk est supérieure à celle d'ARN f'epd-pgk requise pour dissocier le complexe formé avec l'ARN f<sub>pgk-fbaA</sub> (facteur 2). En plus, dans ces expériences, les complexes ARN fpgk-fbad-Hfq avaient été formés en présence d'une concentration en Hfq double de celle utilisée pour la formation des complexes ARN f'epd-pgk-Hfq. Ces données suggéraient donc que le site le plus fort de fixation de la protéine Hfq était localisé dans la région epdpgk (+896 à +1222).

## III.2 / Essais de localisation des sites de fixation de la protéine Hfq au niveau des régions epd-pqk et pqk-fbaA

Pour essayer de savoir si les sites forts de fixation de Hfq correspondaient aux sites d'entrée pgk 2 et fbaA 5 de la RNase E dans chacun des deux ARN étudiés, nous avons réalisé des expériences d'empreintes en utilisant la RNase T1. Pour cela, des complexes f'epd-pgk-Hfq et fpgk-fbaA-Hfq ont été formés en présence de différentes concentrations de protéine Hfq, comprises respectivement entre 1,5 et 6,0 μM pour les complexes f'epd-pgk-Hfq et entre 4,0 et 8,0 μM pour les complexes fpgk-fbaA-Hfq. Ces

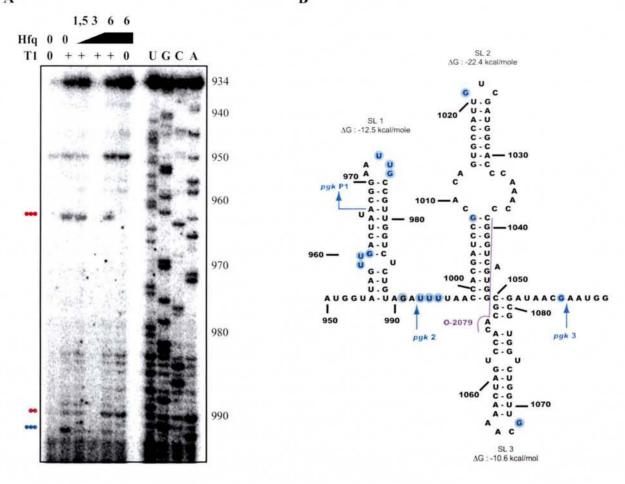


Figure 69 : Localisation de la fixation de la protéine Hfq au niveau du site pgk2 de la région epd-pgk.

A : Empreinte à la RNase T1 de la fixation de la protéine Hfq au niveau du site *pgk2*. Le transcrit froid R<sub>epd-pgk</sub> (100ng) a été incubé pendant 10 min à 37°C, en présence de 1μg d'ARNt, à des concentrations croissante en protéine Hfq, dans le tampon Tris-HCl 10 mM, pH 8.0; MgCl<sub>2</sub> 10 mM; NaCl 50 mM et KCl 50 mM, dans un volume final de 40 μL. Les concentrations de protéine Hfq utilisées sont indiquées en haut de chaque piste. Les différents complexes ont ensuite été soumis à une hydrolyse à la RNase T1 en conditions limitantes. Les produits d'hydrolyse ont été analysés par extension d'amorce à la transcriptase inverse, en utilisant les oligonucléotides O-2979 marqués en 5' par la polynucléotide kinase. Les produits d'extension d'amorce ont énsuite été fractionnés sur un gel de polyacrylamide dénaturant 10 %, avec en parallèle, les produits d'un séquençage réalisé avec le même oligonucléotide. La numérotation à droite est indiquée par rapport au site +1 du site d'initiation de la transcription à partir du promoteur *epd* P0. Les sites de clivages sont indiquées par un rond à gauche du gel, avec en rouge, un site fort de clivage, en bleu, une forte protection par la protéine Hfq.

B: Localisation des sites de modifications engendrés en présence de la protéine Hfq sur le transcrit R<sub>epd-pgk</sub>. La numérotation est indiquée par rapport au site +1 d'initiation de la transcription du promoteur *epd* P0. La position du site d'initiation de la transcription du promoteur *pgk* P1 est indiquée par une flèche brisée, la position des sites de clivages par la RNase E par une flèche.

concentrations en protéine Hfq conduisent d'après les expériences de retard sur gel illustrées dans la figure 67, à la formation dans le cas de l'ARN  $f'_{epd-pgk}$  des complexes  $C_{ep}1$  à 1.5  $\mu$ M,  $C_{ep}1$ , 2 et 3 à 3  $\mu$ M, et  $C_{ep}1$ , 2, 3, 4 et 5 à 6  $\mu$ M, et dans le cas de l'ARN  $f_{pgk-fbaA}$  à la formation des complexes  $C_{pf}1$  et 2 à 4  $\mu$ M et  $C_{pf}1$ , 2, 3 et 4 à 8  $\mu$ M de protéine Hfq. A la concentration de 6  $\mu$ M, une protection générale des deux ARN  $f'_{epd-pgk}$  (Fig. 69) et  $f_{pgk-fbaA}$  (non montré) était observée. Une protection spécifique a seulement été détectée pour le complexe formé entre l'ARN epd-pgk et la protéine Hfq à une concentration de 1,5  $\mu$ M et le seul résidu G protégé dans ces conditions (Fig. 69) correspond au résidu G (+991) de la séquence en simple brin clivée par la RNase E (site pgk 2). Il s'agit donc du site de plus haute affinité au sein des deux ARN. Il doit correspondre à la formation du complexe le plus stable  $C_{ep}1$ . Nous observons donc que le site ayant l'affinité la plus forte pour la protéine Hfq correspond au site d'entrée de la RNase E dans l'ORF epd.

#### III.3 / Test de l'influence de la protéine Hfq sur l'expression des gènes gapA, epd-pgk-fbaA in vivo dans différents milieux de croissance

Une fixation de la protéine Hfq sur le site d'entrée de la RNase E dans les transcrits epd-pgkfbaA, epd-pgk, pgk et pgk-fbaA pouvait moduler leur clivage par cette enzyme, nous avons donc débuté une étude de l'influence de la protéine Hfq sur le taux de transcrits correspondant aux gènes gapA, et epd-pgk-fbaA in vivo, ceci par comparaison du taux des ARN issus de deux souches isogéniques hfq<sup>+</sup> et hfq<sup>-</sup>. Il s'agissait de la souche TX2808 qui porte une insertion d'une cassette kanamycine dans le gène hfq et de la souche isogénique JC7623 qui nous ont été fournies par E. Hajnsdorf (IBPC, Paris). Ces bactéries ont été cultivées en milieu minimum complémenté en glucose ou en milieu riche LB. Les échantillons ont été prélevés en milieu de phase exponentielle de croissance. Les ARN totaux ont été extraits à partir d'échantillons contenant une même quantité de cellules. La quantité d'ARN total a été estimée, puis pour chaque échantillon, 20 µg d'ARN total ont été analysés par extension d'amorce par la transcriptase inverse en utilisant différents oligonucléotides, permettant d'estimer le taux des ARNm gapA, epd, pgk et fbaA (Fig. 70). Lorsque la culture bactérienne a été réalisée en milieu riche, une quantité supérieure d'ARN total (1.8 fois plus) a été obtenue pour la souche  $hfq^+$ , par rapport à la souche  $hfq^-$ , mais l'analyse réalisée sur une même quantité d'ARN total (Fig. 70) ne mettait pas en évidence de fortes variations relative des différents ARNm. Nous notons seulement une légère augmentation des différents ARNm, à l'exception des ARNm gapA P1 à P3, pgk 6 et fbaA 1. Lorsque la culture bactérienne était réalisée en milieu minimum, les quantités d'ARN totaux extraites des deux souches pour une même quantité de cellules étaient identiques, mais nous avons observé par réverse transcription une diminution du taux de l'ensemble des ARNm de la souche hfq. Ces données sont préliminaires : elles devront être

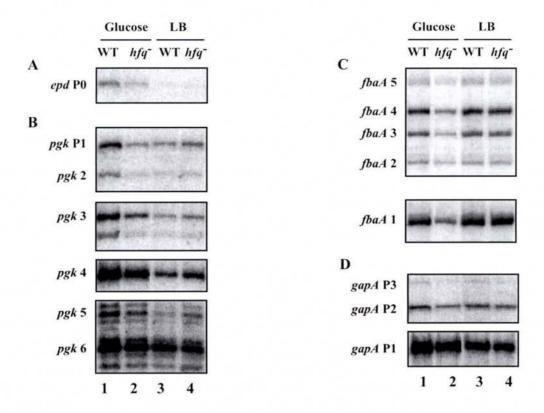


Figure 70 : Etude de l'influence de l'expression de la protéine Hfq sur l'expression des gènes gapA et epd-pgk-fbaA in vivo.

La souche d'*E. coli* JC7623 et la souche isogénique TX2808 *hfq*- ont été cultivées en milieu minimum complémenté avec 2 g de glucose par litre de culture ou en milieu riche (LB), ceci jusqu'en phase exponentielle de croissance (A<sub>600</sub> = 0,5). Des quantités identiques d'ARN totaux (20 μg) d'ARN extraits des différentes cultures ont ensuite été analysés par extension d'amorce par la transcriptase inverse, ceci en utilisant les oligonucléotides O-EPD (A), O-PGK (B), O-FBA (C) et O-GAP (D). Les ADNc ont ensuite été séparés par électrophorèse en gel de polyacrylamide 5 %-urée 8 M.

Les quantités de radioactivitées contenues dans chaque arrêt d'élongation ont été quantifiées à l'aide du Phosphorimager, puis rapportées par rapport à la quantité de radioactivité obtenue dans le cas de la souche sauvage en milieu M63 contenant du glucose.

confirmées en répétant les expériences. Néanmoins, il semble que l'expression ou la non expression de Hfq ait un effet global sur le taux d'ARNm, mais pas un effet spécifique sur certains d'entre eux. En particulier, aucune augmentation marquée des ARNm renfermant l'ORF *epd* n'est observée.

#### IV / Conclusions et perspectives de l'étude de l'opéron epd-pgk-fbaA

L'ensemble des travaux que nous avons réalisés sur l'opéron *epd-pgk-fbaA* chez *E. coli* montre que les ORF de cet opéron sont exprimées à partir de deux promoteurs transcriptionnels, un promoteur fort et régulé *epd* P0 et un promoteur étendu peu régulé *pgk* P1 (Fig. 60). Nous avons montré, par alignement de séquences que la région renfermant le promoteur *epd* P0 est conservée chez les *Enterobacteriaceae*. La séquence conservée inclue la région -35 bien qu'elle ne correspond à aucune des séquences -35 définies pour les différents facteurs  $\sigma$  d'*E. coli*. Il serait donc intéressant de confirmer par mutation que cette région est nécessaire pour l'expression et la régulation de l'ORF *epd*. Il serait alors intéressant de rechercher si un régulateur de la transcription s'y fixe et de l'identifier. La technique utilisée par Jeong *et al*. (2004) pour isoler les facteurs régulateurs de la région promotrice du gène *ptsG* pourrait être utilisée. Le principe est d'amplifier par PCR la région promotrice étudiée avec des oligonucléotides biotinylés de façon à pouvoir immobiliser ce fragment PCR sur une résine de NeutraAvidine, qui est ensuite incubée avec un lysat cellulaire. Les protéines fixées à la région promotrice peuvent ensuite être fractionnées par électrophorèse puis identifiées par spectrométrie de masse.

Nous avons mis en évidence deux sites d'entrée de la RNase E, l'un dans l'ORF epd, l'autre dans la région interORF pgk-fbaA, qui sont eux aussi bien conservés chez les Enterobacteriaceae. Ces sites sont caractérisés par la présence d'une structure tige-boucle riche en A et U avec un résidu G un à deux nucléotides en amont, l'ensemble étant flanqué par deux structures tiges-boucles. Il sera intéressant de tester par mutagenèse dirigée si les caractéristiques structurales de ces sites ont une influence sur l'efficacité de clivage par la RNase E et si leur mutation peut conduire à une stabilisation des transcrits.

Les travaux de V. Bardey montraient une faible efficacité de la traduction de l'ORF epd et une partie de cette faible efficacité semble liée à la faible distance entre la séquence SD et le codon d'initiation de la traduction. Nos études comparatives montrent que cette faible distance est un paramètre conservé des ORF epd. Néanmoins, les résultats des fusions transcriptionnelles et traductionnelles réalisées par V. Bardey suggèrent l'existence possible d'un régulateur traductionnel négatif qui serait titré dans le cas où le gène est porté par un plasmide multicopie. Il serait intéressant d'essayer d'identifier ce régulateur. Il pourrait s'agir d'un petit ARN régulateur.

Enfin, nous avons seulement débuté l'étude du rôle possible de la protéine Hfq *in vivo* sur l'expression de l'opéron *epd-pgk-fbaA*, et ces travaux devront être complétés.

En conclusion, l'ensemble de nos travaux sur l'opéron epd-pgk-fbaA montre comment l'implication des processus de clivages des transcrits permet une expression différentielle des ORF d'un même transcrit. C'est une notion de plus en plus admise dans le domaine de l'expression des gènes bactériens. Par clivage, des ARNm individuels sont produits à partir des transcrits polycistroniques et le taux d'expression de ces ARNm individuels dépend de leur stabilité et de leur efficacité de traduction, paramètres qui peuvent d'ailleurs être liés.

# Partie B : Etude de la régulation de l'expression du gène gapA et de l'opéron epd-pgk-fbaA en fermentation continue dans des conditions de limitation en nutriments

Etant donné que la GAPDH et la PGK agissent en succession dans la glycolyse, en parallèle à l'étude détaillée sur l'expression de l'opéron epd-pgk-fbaA, que nous venons de décrire, l'équipe de C. Branlant mène depuis plusieurs années une étude similaire sur l'expression du gène gapA et ce travail a été plus particulièrement réalisé au laboratoire par B. Charpentier et B. Thouvenot. Un point commun de la régulation de l'expression du gène gapA et de l'opéron epd-pgk-fbaA est leur activation en présence de glucose, lorsque la protéine EIIBC<sup>Gle</sup> du système PTS spécifique du glucose est exprimée (Charpentier et al., 1998). Les deux promoteurs responsables de cette régulation sont les promoteurs epd P0 pour l'opéron epd-pgk-fbaA et gapA P1 pour le gène gapA. D'après les travaux récents de B. Thouvenot (Thouvenot, 2004), cette régulation pourrait à la fois s'exercer au niveau transcriptionnel et au niveau post-transcriptionnel (stabilité des ARNm). Son mécanisme est inconnu.

Au niveau transcriptionnel, on aurait pu penser que, comme pour le promoteur ptsH P0 activé en présence de glucose lorsque le transporteur EIIBCGlc est exprimé (De Reuse et Danchin, 1988), l'activation transcriptionnelle qui s'exerce sur les promoteurs epd P0 et gapA P1 puisse être médiée par la protéine Mlc (Plumbridge, 1999). Cette protéine se fixe sur un élément de régulation du promoteur ptsH P0 en absence de glucose et limite son activité (Plumbridge, 1999). Elle est dissociée en présence de glucose et reste associée à la protéine EIIBCGlc (Nam et al., 2001). Une séquence de forte affinité pour la protéine Mlc a pu être établie par la méthode SELEX (5'-TTATTTN<sub>9</sub>AAATAA-3') (Plumbridge, 2001). Aucune séquence présentant une similarité significative avec ce consensus n'est détectée au niveau du promoteur epd P0, ni au niveau du promoteur gapA P1. B. Thouvenot a vérifié cet état de fait en utilisant une souche Mlc, l'absence d'expression de Mlc n'affecte pas la régulation des promoteurs epd P0 et gapA P1 par le glucose. L'activité d'opérons codant les ARN ribosomiques est aussi stimulée par la présence de glucose dans le milieu de culture (Zhang et Bremer, 1996). Il a été proposé que l'activité des promoteurs de ces opérons dépende de la vitesse de croissance des cellules (Schaechter et al., 1958). Les expériences étaient faites sur le promoteur rrnB P1 de l'opéron rrnB qui est le plus étudié. Une explication qui avait été donnée mais qui est maintenant remise en question est que le taux de transcription des ARNr dépende de la concentration en nucléotide initiant la transcription à ces promoteurs (Gaal et al., 1997). B. Thouvenot a testé si l'effet de la concentration en nucléotide initiateur pouvait expliquer la régulation par le glucose observée pour les promoteurs epd P0 et gapA P1, ceci par des expériences de transcription in vitro. Il a bien retrouvé l'effet de cette concentration pour le promoteur rrnB P1, mais pas pour les promoteurs gapA P1 et epd P0. Les données actuelles sur les

ARNr suggèrent en fait que la régulation en fonction de la concentration en nucléotide initiateur n'intervienne non pas en phase exponentielle de croissance, mais plutôt dans la phase de latence précédent la phase de croissance exponentielle (Murray et al., 2003).

En fait, aussi bien pour les ARNr, que pour les transcrits qui nous intéressent issus des promoteurs *gapA* P1 et *epd* P0, rien n'est clair sur les mécanismes qui font que le taux des transcripts est apparemment lié à la disponibilité en source de carbone rapidement métabolisable et par la même à la vitesse de croissance.

Un des facteurs connu pour réguler négativement la transcription à partir du promoteur epd P0 est la protéine Cra (Ramseier et al., 1995). Sa fixation à l'ADN est inhibée en présence de fructose 1-phosphate et à moindre degré, par le fructose-1,6-bisphosphate (Ramseier et al., 1993). De ce fait, l'activité de la protéine Cra est régulée par le flux de carbone (pour revue, Saier, 1996). Dans une souche n'exprimant pas la protéine EIIBC<sup>Glc</sup> active, on peut imaginer que du fait de la plus faible perméation du glucose, la concentration intracellulaire en fructose-1,6-bisphosphate soit faible et donc que la protéine Cra soit fixée sur son site régulateur au niveau du promoteur epd P0, ce qui limiterait son activité. Néanmoins, ce type d'explication n'est pas valable pour expliquer de façon directe la faible activité du promoteur gapA P1 dans les mêmes conditions, car aucun site de fixation pour la protéine Cra n'est présent au voisinage du promoteur gapA P1.

Aucun des mécanismes connu de régulation par le glucose de la transcription de gènes ne pouvait donc expliquer l'activation par le glucose des promoteurs *epd* P0 et *gapA* P1, dépendantes de l'expression de la protéine EIIBC<sup>Glc</sup>.

Une question importante était de savoir si la protéine EIIBC<sup>Glc</sup> joue un rôle direct ou indirect dans ce mécanisme, ceci par le fait que sa non expression entraîne un ralentissement de la perméation du glucose, qui est alors transporté par le transporteur EII<sup>Man</sup> spécifique du mannose. Le transport du glucose est de ce fait nettement moins efficace, ce qui limite la vitesse de croissance (Curtis et Epstein, 1975; Picon et al., 2005).

Un moyen de comparer les taux de transcrits *gapA* de souches EIIBC<sup>Glc+</sup> et EIIBC<sup>Glc-</sup>, en fonction du taux de croissance en s'affranchissant du rôle de EIIBC<sup>Glc</sup> dans le transport spécifique du glucose, était de réaliser avec ces souches des fermentations en conditions très limitantes en glucose. En effet, dans ces conditions, *E. coli* utilise majoritairement le transporteur ABC Mgl/LamB qui a une plus forte affinité pour le glucose que le système PTS spécifique du glucose (pour revues, Ferenci, 1996; Gosset, 2005). Ceci permet à la bactérie d'avoir une meilleure efficacité d'absorption du glucose, en contre partie, plus de molécules d'ATP sont consommées par molécule de glucose consommée (voir Partie A de l'introduction).

# I / Etude des transcrits gapA dans des bactéries cultivées en fermentation continue en condition de limitation en glucose

#### I.1 / Souches utilisées

Nous avons utilisé la souche *E. coli* TP 2503 et une souche isogénique *E. coli* TP 2504, exprimant une protéine EIIBC<sup>Glc</sup> déficiente pour le transport et la phosphorylation du glucose (Curtis et Epstein, 1975). Le gène muté est appelé *ptsG22*. La mutation portée *ptsG22* correspond à une délétion de 4 nts dans le segment codant le domaine EIIC<sup>Glc</sup> (Plumbridge, 1999). Cette délétion entraîne un décalage du cadre de lecture, conduisant au remplacement des acides aminés du domaine C-ter 220 à 477, par un fragment de 63 acides aminés. Ainsi, la mutation *ptsG22* conduit à la production d'une enzyme EIIC<sup>Glc</sup>, sans domaine IIB<sup>Glc</sup> fonctionnel. Cette enzyme non fonctionnelle pourrait être localisée au niveau de la membrane (Kawamoto *et al.*, 2005). Il semblerait cependant qu'une réinitiation de la traduction ait lieu, en aval du nouveau codon stop, conduisant à la synthèse entière ou partielle du domaine EIIB<sup>Glc</sup> (Plumbridge, 1999).

#### I.2 / Rappel sur les fermentations continues

Nous avons réalisé les fermentations continues en condition de limitation en glucose à différents taux de croissance. Pour cela, nous avons mis au point les conditions de travail pour avoir un volume utile de 2 l dans le fermenteur de type BiofloIII (New Brunswick Scientific). Parallèlement, nous avons calibré les différentes vitesses d'alimentation : en effet, à l'état stationnaire, la vitesse de croissance des bactéries est proportionnelle au débit d'alimentation du milieu. Ceci s'explique par le bilan de biomasse suivant :

Entrée de biomasse + Production = Sortie + Variation  

$$0 + r'''_X \cdot V = Q \cdot X + d(V \cdot X)/dt$$

avec: V: volume utile (volume de milieu contenu dans le fermenteur)

X : concentration en biomasse (g.l<sup>-1</sup>)

Q : débit d'alimentation et de soutirage (l.h<sup>-1</sup>)

t: temps (h)

r "x: vitesse de production de biomasse par litre de milieu de culture et par heure (g.l-1.h-1)

 $\mu$  : taux de croissance ( $\mu$  = r  $\stackrel{""}{_{X}}/$  X en g.g-¹.h-¹)

Lorsque le fermenteur est à l'équilibre, la quantité de biomasse ne varie pas, le volume utile étant fixé à 2 l, le terme d(V . X)/dt est nul. L'équation précédente devient :

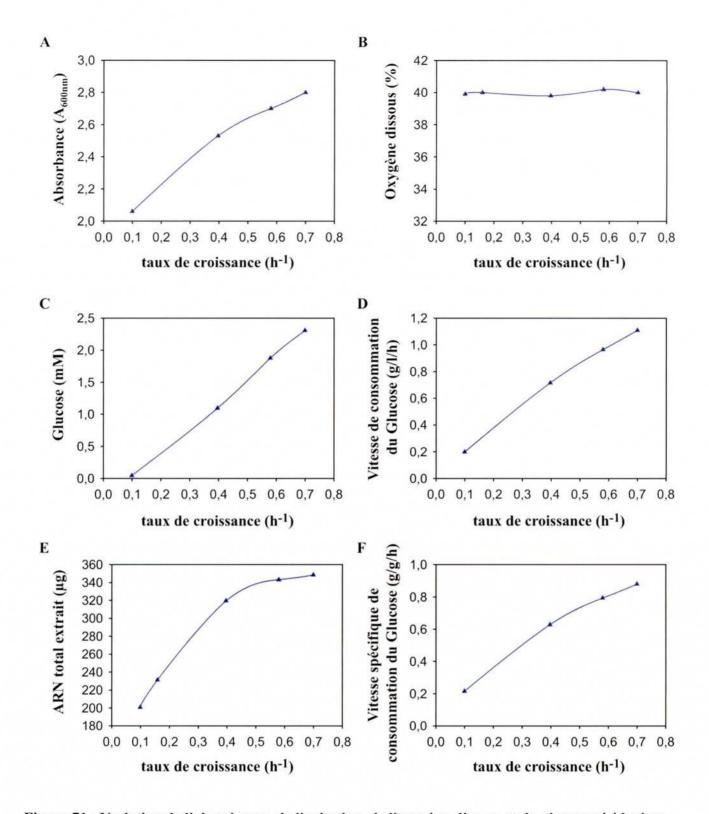


Figure 71 : Variation de l'absorbance, de l'agitation, de l'oxygène dissous et du glucose résiduel en fonction du taux de croissance, pour une fermentation continue limitée en glucose (2g/l).

La souche d'*E. coli* TP 2503 a été cultivée en fermentation continue en conditions de limitation en glucose (2 g/l). Les différents paramètres ont été mesurés après 5 temps de passage suivant l'établissement de chaque taux de croissance.

- A: Absorbance à 600 nm.
- B: Oxygène dissous.
- C: Glucose résiduel.
- D : Quantité d'ARN total extrait par échantillon.
- E: Vitesse de consommation du glucose.
- F: Vistesse spécifique de consommation du glucose.

$$r'''_X \cdot V = Q \cdot X$$

Le taux de croissance est alors directement proportionnel au débit d'alimentation :

$$\mu = r'''_X / X = Q / V$$

Nous avons mis au point les conditions pour avoir un débit d'alimentation tel que pour un volume utile de 2 l, nous puissions obtenir un taux de croissance variant de 0,1 h<sup>-1</sup> à 1 h<sup>-1</sup>. Ceci représente des débits d'alimentation variant de 0,2 l.h<sup>-1</sup> à 2 l.h<sup>-1</sup> pour un volume V de milieu de 2 l.

#### I.3 / Fermentation continue réalisée en condition de limitation en glucose avec 2 g/l de glucose

Nous avons réalisé une première fermentation continue, avec la souche sauvage TP 2503 en milieu minimum M63, supplémenté avec une concentration en glucose de 2 g/l. Le taux de croissance maximale pendant la phase de batch était de 1.01 (+/-0.01) h-1. Les différentes mesures ont été réalisées pour des taux de croissance compris entre 0,1 h<sup>-1</sup> et 0,7 h<sup>-1</sup>, après 5 temps de passage. Un temps de passage correspond au temps nécessaire pour que le volume apporté par l'alimentation soit égal au volume de milieu présent dans le fermenteur, appelé aussi volume utile. En fermentation continue, l'équilibre est considéré être atteint pour chaque taux de croissance, après 3 à 5 temps de passage, ce qui permet une adaptation du métabolisme des cellules aux conditions de croissance données et, statistiquement, un taux de renouvellement de la biomasse et du milieu de culture dans le fermenteur d'environ 94 %. La mesure de l'absorbance montre une augmentation de la biomasse avec le taux de croissance (Fig. 71A). En parallèle, pour un volume d'alimentation en air de 3 vvm constant au cours de la culture, la vitesse d'agitation augmente lorsque le taux de croissance augmente, ceci pour garder un pourcentage d'oxygène dissous de l'ordre de 40 % (Fig. 71B). Ainsi, l'augmentation de la biomasse est accompagnée d'une augmentation de la consommation d'oxygène. Nous avons ensuite mesuré la concentration en glucose résiduel (Fig. 71C). Elle varie de 50 µM à 2,3 mM, lorsque le taux de croissance est augmenté de 0,1 à 0,7 h<sup>-1</sup>. La vitesse spécifique de consommation du glucose augmente aussi avec le taux de croissance, passant de 0,2 à 0,8 g.g-1.h-1 (Fig. 71F). Or, pour avoir une induction du transporteur ABC mgl/LamB, le taux de croissance doit être compris entre 0,2 et 0,7 h<sup>-1</sup>, et la concentration en glucose résiduel entre 0,3 et 2 µM (pour revue, Ferenci, 2001). La concentration en glucose utilisée semblait a priori trop élevée pour que l'expression du système ABC Mgl/LamB soit induite. Nous avons alors réalisé les fermentations continues avec une concentration en glucose de 0,2 g/l.

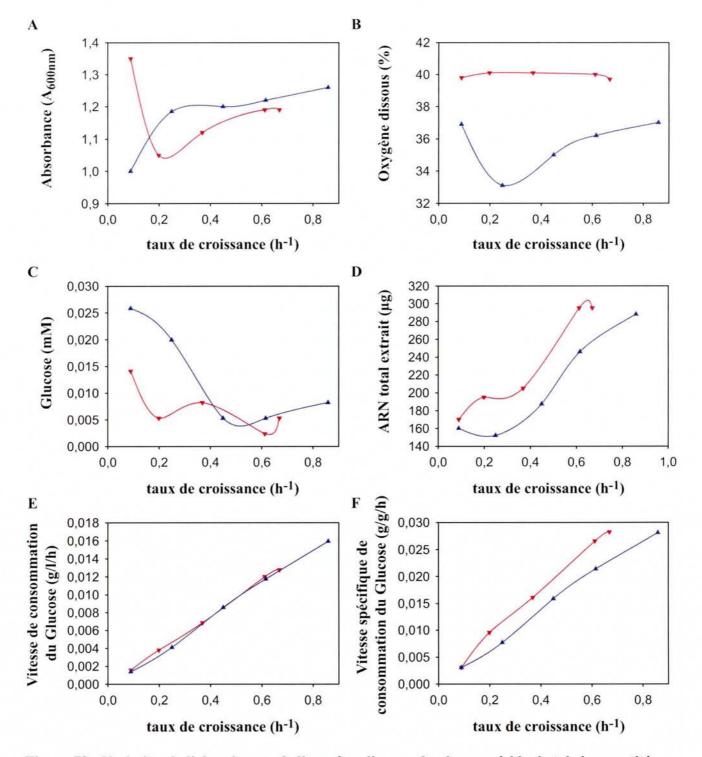


Figure 72 : Variation de l'absorbance, de l'oxygène dissous, du glucose résiduel et de la quantité d'ARNm extrait en fonction du taux de croissance, pour une fermentation continue en conditions de très forte limitation en glucose (0,2 g/l).

Les souches d'*E. coli* TP 2503 (bleu) et TP 2504 (rouge) ont été cultivées en fermentation continue limitée en glucose à 0,2 g/l. Les différents paramètres ont été mesurés après 5 temps de passage suivant l'établissement de chaque taux de croissance.

- A: Absorbance à 600 nm.
- B: Oxygène dissous.
- C : Glucose résiduel.
- D : Quantité d'ARN total extrait par échantillon.
- E : Vitesse de consommation du glucose.
- F: Vistesse spécifique de consommation du glucose.

# I.4 / Fermentation continue réalisée en condition de limitation en glucose (0,2 g/l)

Dans la période de batch du fermenteur, le taux de croissance de la souche sauvage était de 1,07 (+/-0,016) h<sup>-1</sup>. Les différentes mesures ont été réalisées pour des taux de croissance allant de 0,1 à 0,8 h<sup>-1</sup>, après 5 temps de passage. Comme précédemment, la biomasse augmentait avec le taux de croissance (Fig. 72A). L'alimentation en air était fixée au maximum, à 4 vvm, l'agitation était maintenue à son maximum (800 tours/min). De façon surprenante, pour des taux de croissance supérieur ou égal à 0,2 h<sup>-1</sup>, la concentration en oxygène dissous augmentait avec le taux de croissance (Fig. 72B). La concentration en glucose résiduel pour un taux de croissance supérieur ou égal à 0,2 h<sup>-1</sup> était de l'ordre de 5 μM, alors qu'à un taux de croissance de 0,1 h<sup>-1</sup>, elle était de 25 μM (Fig. 72C). Ceci pourrait s'expliquer par la transcription de l'ORF *lamB* pour un taux de croissance supérieur ou égal à 0,2 h<sup>-1</sup>, ce qui permettrait de mieux absorber le glucose présent dans le milieu de culture.

Nous avons ensuite réalisé une fermentation continue avec la souche mutante TP 2504 (ptsG22), avec une concentration en glucose de 0,2 g/l. Le taux de croissance déterminé pendant la phase batch en fermenteur était de 0,72 (+/-0,01) h<sup>-1</sup>. Les différentes mesures ont été réalisées pour des taux de croissance allant de 0,1 à 0,67 h<sup>-1</sup>, après 5 temps de passage. Contrairement aux résultats obtenus avec la souche sauvage, la biomasse était maximale pour un taux de croissance de 0,1 h<sup>-1</sup> (A<sub>600nm</sub> de 1,35) (Fig. 72A). Elle augmentait de 1,05 à 1,2 A<sub>600nm</sub> pour des taux de croissance compris entre 0,2 et 0,6 h<sup>-1</sup>. L'alimentation en oxygène dissous était fixée à 3 vvm, l'agitation était augmentée avec le taux de croissance, de 215 à 380 tours par min (Fig. 72B). La concentration en glucose résiduel, comme pour la souche sauvage, était maximale au taux de croissance de 0,1 h<sup>-1</sup> et atteignait la valeur de 14,1 μM, puis variait entre 2 et 6 μM pour des taux de croissance allant de 0,2 h<sup>-1</sup> à 0,67 h<sup>-1</sup> (Fig. 72C). Ceci est en faveur d'une meilleure utilisation du glucose pour un taux de croissance supérieur ou égal à 0,2 h<sup>-1</sup>, qui pourrait s'expliquer par l'utilisation du transporteur ABC Mgl/LamB.

Pour les deux souches, aucune production d'acide acétique n'était détectée, alors qu'avec une concentration en glucose de 2 g/l, une production d'acide acétique avait été observée. Ceci vérifiait le fait qu'il y avait bien limitation en glucose.

Nous avons ensuite déterminé les vitesses de consommation du glucose. De façon surprenante, les vitesses de consommation du glucose étaient identiques pour les souches sauvages et mutées, alors que la vitesse spécifique de consommation du glucose pour la souche mutée était légèrement plus élevée, que celle de la souche sauvage, ceci pour les taux de croissance supérieur ou égal à 0,2 h<sup>-1</sup> (Fig. 72, E et F).

Plusieurs échantillons de culture, correspondant à 15 Unités de A<sub>600nm</sub> ont été prélevés pour chaque taux de croissance à la fois pour la souche sauvage (TP 2503) et pour la souche mutée (TP

2504; ptsG22). Les ARN totaux ont ensuite été extraits de ces échantillons. La quantité d'ARN total extraite augmentait avec le taux de croissance pour chaque souche (Fig. 72D). Pour chaque échantillon, 20 µg d'ARN totaux ont été analysés par extension de l'amorce O-GAP par la transcriptase inverse (Fig. 73). Les ARNm initiés au promoteur gapA P1 présents de manière prépondérante en phase exponentielle de croissance en présence de glucose sont, aussi bien pour la souche sauvage que mutée, présents en très faible quantité dans les conditions de fermentations utilisées. Les ARNm les plus abondants sont ceux initiés au promoteur gapA P2. Ceci s'explique par le fait que le promoteur gapA P2 est transcrit par l'holoenzyme  $E\sigma^{32}$ , dont le taux augmente lors d'une carence en carbone (Jenkins et al., 1991). Bien que dans des proportions plus faibles, les transcrits initiés au promoteur gapA P3 sont eux aussi plus abondants que ceux initiés au promoteur gapA P1. Or, le promoteur gapA P3 est activé par le complexe CRP-AMPc (Charpentier et Branlant, 1994), dont la concentration diminue lorsque le taux de croissance de la bactérie augmente (Ferenci, 2001 ; Kuo et al., 2003). On peut supposer qu'un ou plusieurs autres facteurs sont aussi impliqués dans la régulation de l'expression du promoteur gapA P3. Par ailleurs, le taux de chacun des 3 transcrits augmente avec la vitesse de croissance, et ceci pour les 2 souches. Lorsque les données obtenues pour les souches sauvages et mutées sont comparées, nous notons seulement une augmentation plus faible des transcrits gapA P1 lors de l'augmentation du taux de croissance de 0,45 à 0,6 par rapport à la variation observée pour la souche mutée lors de l'augmentation du taux de croissance de 0,4 à 0,6. Le comportement des deux souches est donc très voisin dans les conditions de fermentations utilisées, ce qui est en accord avec une très faible expression du transporteur EIIBCGlc dans ces conditions.

L'information importante apportée par ces expériences était que le taux des 3 transcrits du gène gapA augmente de manière parallèle à l'augmentation du taux de croissance et ceci de manière indépendante de l'expression de EIIBC<sup>Glc</sup>. La question était donc de savoir si la modulation du taux de croissance par un paramètre autre que le glucose conduirait aux mêmes variations du taux des ARNm gapA.

# II / Etude des transcrits epd et gapA dans des bactéries cultivées en fermentation continue en condition de limitation en phosphate

#### II.1 / Choix de l'élément limitant la fermentation continue

Pour effectuer ce choix, nous nous sommes placés en excès de glucose. Nous avons alors testé si le potassium ou le phosphate pouvaient être utilisés pour limiter le taux de croissance lors de fermentations continues en milieu minimum complémenté avec des casa amino acides. L'utilisation de

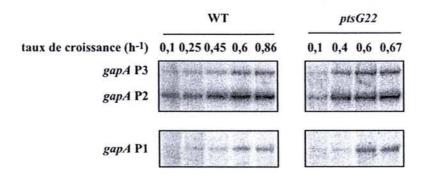


Figure 73 : Analyse par extension d'amorce des transcrits initiés au niveau des promoteurs du gène gapA lors de fermentations continues en conditions de limitation en glucose

Les bactéries WT (TP 2503) et mutées *ptsG22* (TP 2504) ont été cultivées en fermentation continues, à différents taux de croissance, ceci en limitation en glucose. Après établissement d'un nouveau taux de croissance et de 5 temps de passage, un échantillon de cellules a été prélevé. Les ARN totaux ont ensuite été extraits à partir des différents échantillons puis 20 µg d'ARN totaux issus de chaque échantillon ont été analysés par extension de l'oligonucléotide O-GAP radiomarqué par la transcriptase inverse. Les ADNc résultants ont ensuite été fractionnés par électrophorèse en gel de polyacrylamide dénaturant à 5 %.

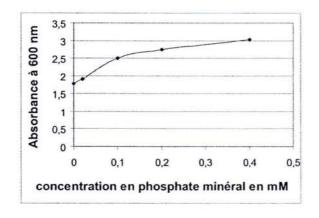


Figure 74 : Evolution de l'absorbance à 600 nm en fonction de la concentration en phosphate minéral

La souche bactérienne TP 2503 a été cultivée en fermentation continue, à un taux de croissance de 0,2 h<sup>-1</sup>, en limitation de phosphate, ceci avec différentes concentrations en phosphate minéral. L'absorbance a été mesurée après 5 temps de passage pour chaque concentration en phosphate.

ces derniers et l'auxotrophie des souches utilisées pour l'arginine ne nous permettaient pas de réaliser une limitation en azote. La fermentation continue que nous avons réalisée en absence de potassium conduisait à une absorbance deux fois plus importante que celle obtenue en limitation en glucose (de l'ordre de 3,2 au lieu de 1,2 unités A<sub>600nm</sub>) ceci pour un taux de croissance de 0,2 h<sup>-1</sup>. Comme nous ne pouvions pas par ce type de limitation obtenir des densités cellulaires comparables à celles obtenues en limitation de glucose, nous avons abandonné l'idée de faire une limitation en potassium. Nous avons alors testé l'effet d'une limitation en phosphate minéral. Il est à noter que les casa-amino acides ajoutés dans le milieu comportent du phosphate. Nous avons mesuré la densité cellulaire obtenue pour différentes concentrations en phosphate minéral ajouté, ceci à un taux de croissance de 0,2 h-1 (Fig. 74). Comme précédemment, les mesures ont été réalisées pour chaque concentration en phosphate minéral utilisé, après 5 temps de passage. Nous avons observé une augmentation de la densité cellulaire en augmentant la concentration en phosphate minéral, apporté sous la forme KH<sub>2</sub>PO<sub>4</sub>, ce qui confirmait l'existence d'une limitation en phosphate dans les conditions utilisées. Lors d'une limitation en phosphate, en présence de glucose en excès, la bactérie excrète de l'acide acétique, ce qui acidifie le milieu de culture. Pour que le phosphate ajouté puisse exercer un effet tampon et pour pouvoir réguler efficacement le pH du milieu de culture à 7,0, j'ai choisi une concentration en phosphate de 0,02 mM pour réaliser les fermentations continues en phosphate limitant.

### II.2 / Fermentation continue réalisée en condition de limitation en phosphate minéral (0,02 mM)

Des fermentations continues en phosphate limitant ont été réalisées pour les 2 souches *E. coli* TP 2503 et *E. coli* TP 2504 (*ptsG22*), avec 4 g/l de glucose. Le taux de croissance pendant la phase de batch du fermenteur était de 1,07 h<sup>-1</sup> pour la souche sauvage (TP 2503) et de 0,60 h<sup>-1</sup> pour la souche mutée *ptsG22* (TP 2504). Nous avons alors prélevés des échantillons de chaque culture pour des taux de croissance allant de 0,1 à 0,8 h<sup>-1</sup> pour la souche sauvage et de 0,1 à 0,6 h<sup>-1</sup> pour la souche mutée. Comme précédemment, les échantillons, qui correspondaient chacun à 15 Unités d'A<sub>600nm</sub>, ont été prélevés après 5 temps de passage de façon à ce que le taux de croissance soit stabilisé.

Pour les deux souches, les densités cellulaires finales diminuaient avec le taux de croissance (Fig. 75A). Avec une alimentation en oxygène de 1 vvm, l'agitation restait comprise entre 301 et 316 tours/min pour la souche sauvage et variait entre 276 et 310 tours/min pour la souche mutée *ptsG22*. Ainsi, la consommation d'oxygène semblait augmenter avec le taux de croissance. La consommation d'oxygène par unité de densité cellulaire est plus importante pour la souche mutée *ptsG22* par rapport à la souche sauvage.

La concentration en glucose résiduel se situait autour de 5 mM (0,9 g/l) pour les deux souches

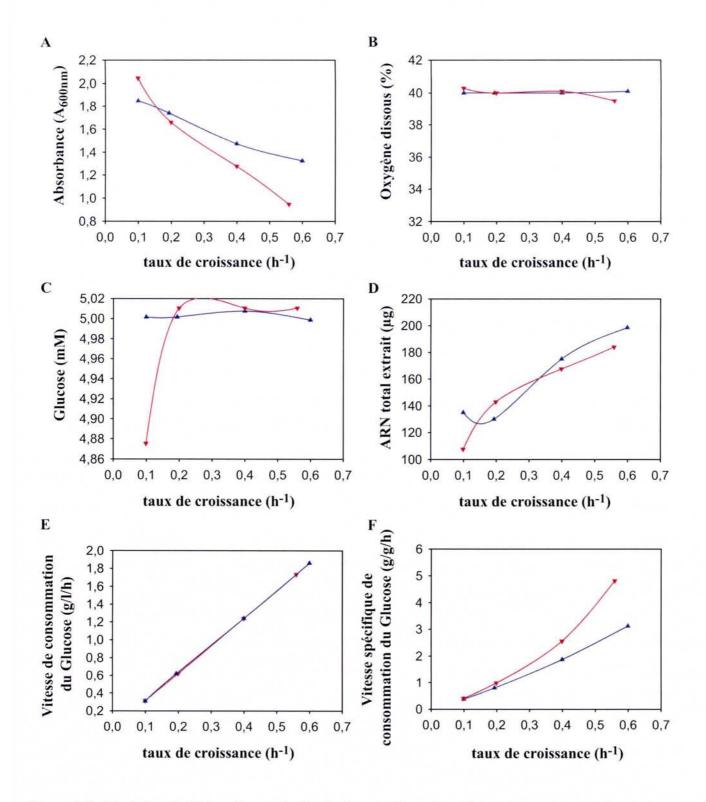


Figure 75 : Variation de l'absorbance, de l'agitation, de l'oxygène dissous et du glucose résiduel en fonction du taux de croissance, pour une fermentation continue limitée en phosphate (0,02 mM).

Les souches d'*E. coli* TP 2503 (WT) et TP 2504 (*ptsG22*) ont été cultivées en fermentation continue limitée en phosphate à 0,02 mM. Les différents paramètres ont été mesurés après 5 temps de passage suivant l'établissement de chaque taux de croissance.

- A: Absorbance à 600 nm.
- B: Oxygène dissous.
- C : Glucose résiduel.
- D : Quantité d'ARN total extrait par échantillon.
- E: Vitesse de consommation du glucose.
- F : Vitesse spécifique de consommation du glucose.

étudiées, ceci quelque soit le taux de croissance (Fig. 75C). Ceci est en accord avec notre observation de vitesse de consommation du glucose identique pour les deux souches aux mêmes taux de croissance (Fig. 75E). Par contre, pour les taux de croissance supérieurs ou égal à 0,2 h<sup>-1</sup>, les vitesses spécifiques de consommation du glucose pour la souche mutée étaient plus importantes que celles obtenues pour la souche sauvage (Fig. 75F), ce qui signifie que prises individuellement, les cellules sauvages consomment moins de glucose que les cellules mutées. Le rendement de biomasse par rapport au glucose était plus faible pour la souche mutée que pour la souche sauvage.

De l'acide acétique était retrouvé dans le milieu de culture à une concentration comprise entre 0,02 et 0,07 mM dans le cas de la souche sauvage, alors qu'aucune production d'acide acétique n'était détecté pour la souche mutée lorsque le taux de croissance est supérieur ou égal à 0,2 h<sup>-1</sup>, c'est-à-dire lorsque les cellules étaient en phase exponentielle de croissance.

Nous avons ensuite extrait les ARN totaux pour chaque échantillon prélevé. Pour les 2 souches étudiées, la quantité totale d'ARN extrait augmentait avec le taux de croissance. Nous avons ensuite, pour chaque échantillon, analysé 20 μg d'ARN totaux par extension d'une amorce radiomarquée par la transcriptase inverse (Fig. 76). Nous avons utilisé les amorces O-EPD, O-PGK, O-FBA, O-GAP et O-RRNB. L'amorce O-RRNB permet de suivre la variation du taux de transcrits initiés au promoteur rrnB P1 de l'opéron rrnB. Comme attendu, nous observons une augmentation du taux des transcrits rrnB P1 avec le taux de croissance. Cette augmentation était cependant plus importante pour la souche sauvage que pour la souche mutée, mais le degré de reproductibilité de cette observation devra être testé. L'analyse des ARN totaux avec l'amorce O-GAP montre que l'augmentation du taux des transcrits initiés aux 3 promoteurs suit l'augmentation du taux de croissance. Dans ces conditions de fermentation, quel que soit le taux de croissance, les transcrits gapA P1 sont ceux présents en plus grande quantité, et c'est tout particulièrement marqué pour la souche sauvage. Pour la souche sauvage, le taux de transcrit initiés au promoteur epd P0 et ses produits de clivages analysés par les amorces O-EPD, O-PGK et O-FBA, augmentent aussi avec le taux de croissance, alors que pour la souche mutée, le taux des transcrits initiés au promoteur epd P0 augmente, le taux des produits de clivages semblent rester stables.

Le taux des transcrits initiés au promoteur pgk P1 est faible au taux de croissance 0,1 h<sup>-1</sup> par rapport aux autres taux de croissance. Le faible taux d'utilisation de ce promoteur pour un très faible taux de croissance pourrait s'expliquer comme suit : pour un taux de croissance inférieur ou égal à 0,1 h<sup>-1</sup>, la bactérie entre en phase stationnaire de croissance et utilise de façon plus importante la forme  $E\sigma^S$  de l'ARN polymérase au détriment de l'holoenzyme  $E\sigma^{70}$ , ce qui entraîne une expression différentielle des gènes (voir Introduction, Partie B). Ainsi, le promoteur pgk P1 pourrait être moins utilisé lorqu'une forte proportion de l'ARN polymérase est sous forme  $E\sigma^S$ .

Nos données suggèrent donc que l'augmentation des taux de transcrits gap en milieu minimum en présence de glucose résulte sans doute de deux effets l'un lié aux variations du taux de croissance et

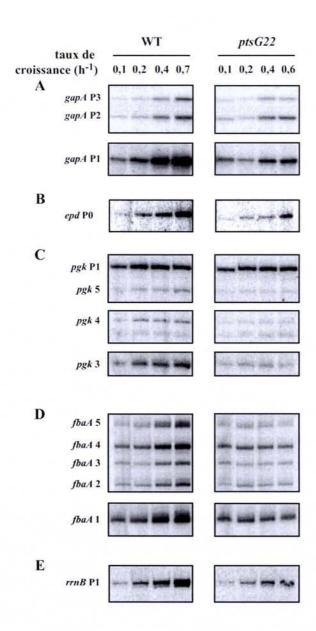


Figure 76 : Evolution du taux des transcrits gapA, epd, pgk, fbaA et rrnB en fonction du taux de croissance, lors de fermentations continues des souches E. coli TP 2503 et TP 2504 réalisées en condition de limitation en phosphate.

Les fermentations continues ont été réalisées en milieu minimum M63 supplémenté avec 4 g/l de glucose, en limitation en phosphate minéral (0,02 mM). Pour chaque taux de croissance, les échantillons de cellules correspondant à 15 unités A<sub>600nm</sub> ont été prélevés après 5 temps de passage. Les ARN totaux ont été extraits des échantillons puis 20 μg d'ARN total ont été analysés par extension des oligonucléotides radiomarqués O-GAP (A), O-EPD (B), O-PGK(C), O-FBA (D) et O-RrnB (E) à la transcriptase inverse. Les ADNc synthétisés ont ensuite été fractionnés par électrophorèse sur gel de polyacrylamide dénaturant 5 %.

l'autre à la présence du transporteur EIIBC<sup>Glc</sup> actif. Le taux de transcrits pgk P1 ne dépend ni de l'expression du transporteur EIIBC<sup>Glc</sup>, ni du taux de croissance à partir d'un taux de croissance supérieur ou égal à  $0.2 \text{ h}^{-1}$ , ce qui est en accord avec les données intégrées dans l'article inséré dans la partie A des résultats. Les travaux présentés ici renforcent donc l'idée que le promoteur epd P0 a été conservé au cours de l'évolution du fait de sa haute importance fonctionnelle. On voit en effet ici que la dépendance du taux des ARNm pgk et fbaA vis-à-vis du taux de croissance repose sur l'existence de ce promoteur.

#### III / Conclusion et Perspectives

L'utilisation des fermentations continues nous a permis de contrôler le taux de croissance cellulaire, et de montrer que le taux des transcrits initiés aux promoteurs gapA P1, gapA P2, gapA P3 et epd P0 augmente lorsque le taux de croissance augmente, quelques soient les conditions de croissance. Les fermentations continues limitées en phosphate nous ont permis de différencier l'effet du taux de croissance de l'inactivation de la protéine EIIBC<sup>Glc</sup> en présence de glucose : l'inactivation du transporteur EIIBC<sup>Glc</sup> semble conduire à une diminution du taux des transcrits initiés aux promoteurs gapA P1 et epd P0, pour un même taux de croissance. Il serait intéressant par la suite d'étudier si cette diminution résulte d'un phénomène transcriptionnel ou post-transcriptionnel. Pour cela, les souches monolysogéniques contenant les fusions transcriptionnelles epd-lacZ ou gapA-lacZ pourront être cultivées en fermentations continues en limitation phosphate. Les échantillons prélevés à différents taux de croissance permettront l'analyse des transcrits endogènes et des transcrits issus des fusions transcriptionnelles, ainsi que l'analyse de l'activité β-galactosidase.

Les résultats obtenus par B. Thouvenot à partir des fusions transcriptionnelles epd-lacZ et gapA-lacZ montrent que l'augmentation du taux des transcrits issus des promoteurs gapA P1 et epd P0 en présence de glucose résulte majoritairement d'un mécanisme post-transcriptionnel. Ceci est à rapprocher des stabilités déterminées en milieu minimum complémenté en glucose ou glycérol, par l'approche des puces à ADN (Bernstein et al., 2002, Bernstein et al., 2004). Ainsi, les transcrits gapA, pgk et fbaA semblent avoir une plus forte stabilité dans un milieu complémenté en glucose par rapport à un milieu complémenté en glycérol. Dans l'article 1, il a été montré que la RNase E initie la dégradation des transcrits issus de l'opéron epd P0, comme pour la majeure partie des transcrits d'E. coli (voir Introduction, Partie D). La stabilisation des différents transcrits peut résulter d'une diminution de l'activité de la RNase E. En effet, l'activité de la RNase E varie avec les conditions de croissance et peut être modulée par l'expression des gènes rraA et rraB (voir Introduction, partie D.I.1.a.ɛ).

La régulation post-transcriptionnelle s'excerçant sur les messagers gapA et epd-pgk-fbaA pourrait être à rapprocher du mécanisme permettant une variation de la stabilité de l'ARNm du gène ptsG en fonction du flux carboné (Kimata et al., 2001; Morita et al., 2003). Lors d'un bloquage au niveau des premières étapes de la glycolyse, réalisé par mutation des gènes pgi, pfkA ou fbaA ou lors de l'utilisation de l'intermédiaire non métabolisable α-méthylglucoside, l'ARNm ptsG est déstabilisé par l'intermédiaire de l'énolase et de l'hélicase (Morita et al., 2004a). Au laboratoire, B. Charpentier a montré une diminution du taux des transcrits initiés au promoteur gapA P1, epd P0 et rrnB et du taux de croissance avec une augmentation de la concentration en α-méthylglucoside. L'énolase, composant du dégradosome et enzyme de la glycolyse, pourrait faire le lien entre la stabilité des transcrits de la glycolyse et le flux de la glycolyse. Lorsque le flux de la glycolyse est important, l'activité de l'énolase serait orientée vers la glycolyse. Lorsque le flux de la glycolyse diminue, l'énolase pourrait favoriser la dégradation des transcrits codant les enzymes de la glycolyse.

Enfin, un dernier paramètre pouvant intervenir dans la stabilité des transcrits est la PAP I. L'activité de cette enzyme est inversement proportionnelle au taux de croissance (Jasiecki et Wegrzyn, 2003). Ainsi, la plus faible stabilité des différents transcrits pourrait provenir d'une plus forte polyadénylation par la PAP. Comme, elle se fixe préférentiellement aux extrémités 3' des transcrits possédant une extrémité 5' monophosphate, elle favoriserait la dégradation des transcrits maturés par la RNase E de l'opéron epd-pgk-fbaA, ce qui expliquerait une dégradation préférentielle des transcrits initiés au promoteur epd P0 et plus rapidement clivés par la RNase E que les transcrits initiés au promoteur pgk P1. Ainsi, une augmentation du taux de croissance pourrait conduire à une diminution du taux de polyadénylation et une stabilisation des transcrits maturés par la RNase E, en particulier les transcrits issus du promoteur epd P0. Cette hypothèse pourra être testée en comparant la stabilité des transcrits gapA et epd-pgk-fbaA de souches pcnB<sup>+</sup> et pcnB<sup>-</sup> cultivées à différents taux de croissance.

# Partie C: Etude comparative des souches d'E. coli TP 2503 et TP 2504

Nous venons de voir que dans la souche TP 2504 portant la mutation *ptsG22*, le glucose est moins rapidement consommé que dans la souche sauvage : d'après les données de la littérature, le glucose en excès entre principalement par le transporteur PTS spécifique du mannose, dans la souche PPA 291, dérivée de la souche MC 4100 portant la mutation Δ*ptsG::Km*, à un taux beaucoup plus faible qu'en présence du transporteur EIIBC<sup>Glc</sup> actif (Picon *et al.*, 2005; pour revue, Gosset, 2005). Au laboratoire, B. Charpentier avait observé une diminution des activités enzymatiques GAPDH et PGK lorsque le transporteur EIIBC<sup>Glc</sup> est inactivé (Charpentier *et al.*, 1998). Ainsi, la plus faible consommation de glucose et la diminution des activités GAPDH et PGK suggèrent que le flux carboné au travers de la glycolyse est faible dans la souche mutée par rapport à la souche sauvage.

Lorsque la bactérie *E. coli* est cultivée sur un milieu riche en glucose, elle excrète 10 à 30 % du carbone apporté sous forme d'acétate (Farmer et Liao, 1997). Cette production d'acide acétique semble résulter d'une inadéquation entre l'entrée du glucose et les besoins en sources de carbone des cellules pour la production de biomasse et d'énergie (Farmer et Liao, 1997; pour revue Gosset, 2005). En particulier, le cycle TCA serait limitant par rapport au flux carboné qui passe au travers de la glycolyse. Dans les fermentations continues en condition de limitation en phosphate que nous avons réalisées, le glucose était en excès. Cependant, pour un taux de croissance au moins égal à 0,2 h<sup>-1</sup> la souche mutée *ptsG22* ne conduisait à aucune production d'acide acétique, ceci contrairement à ce qui était observé pour la souche sauvage (TP 2503). Ces résultats sont des arguments supplémentaires en faveur du passage d'un faible flux carboné au travers de la glycolyse, ce flux pouvant être orienté préférentiellement vers le cycle TCA.

La production d'acide acétique dans le milieu de culture conduit à une acidification de ce dernier, les cellules sont alors en état de stress, ce qui limite la production de protéines recombinantes et de la croissance bactérienne (Jensen et Carlsen, 1990; Luli et Strohl, 1990; Gschaedler et al., 1999). Or, lorsqu'E. coli est cultivé sur différentes sources de carbones, à l'exception du fructose, il produit de l'acide acétique lorsque son taux de croissance est supérieur ou égal à 0,35 h<sup>-1</sup> (El-Mansi, 2004). Différentes approches ont été développées pour diminuer cette production d'acide acétique, telles que la modification génétique des souches d'E. coli, l'utilisation de différentes stratégies d'alimentation, la croissance des souches d'E. coli à un faible taux de croissance (pour revue, Gosset, 2005).

L'observation d'une absence de production d'acide acétique en milieu minimum complémenté en glucose pour un taux de croissance supérieur ou égal à 0,2 h<sup>-1</sup> pour la souche d'*E. coli* TP 2504 (ptsG22) suggérait qu'elle pourrait être intéressante pour la production de protéines recombinantes.

Avant de tester si c'était le cas, nous avons développé une étude approfondie du comportement des souches TP 2503 et TP 2503 (ptsG22) dans différentes conditions de cultures.

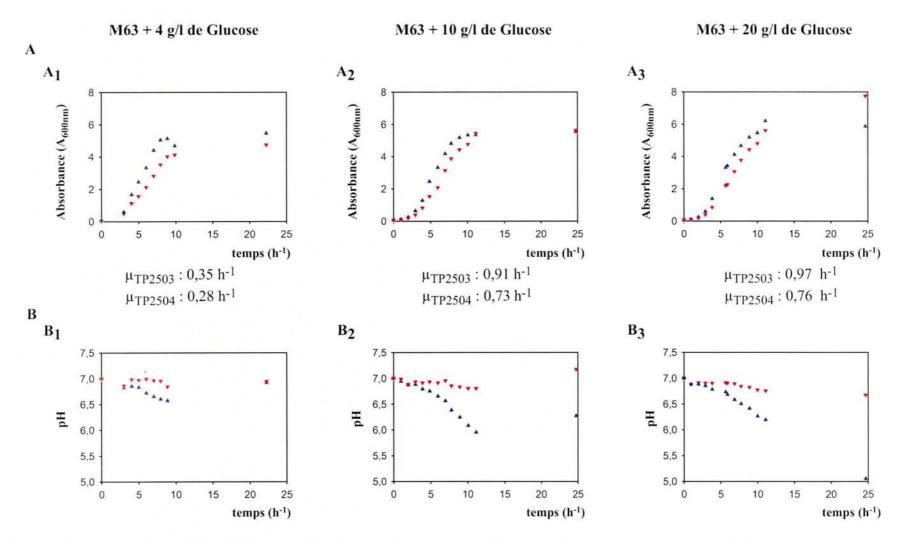


Figure 78 : Cultures des souches TP 2503 et TP 2504 en milieu minimum M63 complémenté avec du glucose à différentes concentrations.

Les souches TP 2503 et TP 2504 ont été cultivées en milieu minimum M63 supplémenté en glucose. Les triangles bleus correspondent à la souche sauvage, les triangles rouges à la souche mutée.

A : Suivi de l'absorbance à 600 nm en fonction du temps.

B: Suivi du pH en fonction du temps.

I / Etude des variations de densité cellulaire des souches *E. coli*TP 2503 et TP 2504 (ptsG22) non transformées dans différentes conditions de croissance

Les suivis de densités cellulaires ont été réalisés en milieu minimum M63 et en milieu riche LB, supplementés par du glucose à différentes concentrations.

## I.1 / Etude de la croissance cellulaire en milieu minimum M63 supplémenté en glucose

L'étude de la croissance des souches sauvages (TP 2503) et mutées ptsG22 (TP 2504) a été réalisée tout d'abord en milieu minimum M63, supplémenté avec du glucose à une concentration de 4, 10 et 20 g/l. Les précultures ont toujours été réalisées en milieu minimum M63 contenant seulement 4 g/l de glucose. Deux expériences différentes ont été réalisées pour chaque condition testée. De manière reproductible, le taux de croissance de la souche mutée était plus faible que celui de la souche sauvage (80 %) (Fig. 78A). Cependant, la densité cellulaire en phase stationnaire de croissance de la souche mutée augmentait lorsque la concentration en glucose dans le milieu de culture augmentait, ce qui n'était pas le cas pour la souche sauvage. Par ailleurs, à forte concentration en glucose, le pH du milieu de culture en phase stationnaire était beaucoup plus bas pour la souche sauvage que pour la souche mutée (Fig. 78B). Les surnageants des cultures ont été analysés par CLHP, ce qui a permis d'associer la diminution du pH du milieu de culture à une augmentation de la concentration en acide acétique (résultats non montrés). Ainsi, aussi bien en conditions limitantes que non limitantes en phosphate, la souche mutée ptsG22 produisait moins d'acide acétique que la souche sauvage. Cependant, l'avantage conféré par la diminution de la production d'acide acétique ne se traduisait pas par un gain marqué de densité cellulaire en phase stationnaire de croissance (seulement de l'ordre de 20 %). Ceci est en accord avec des résultats obtenus par Chou et al. (1994) sur milieu minimum M9 complémenté avec du glucose à 2 ou 5 g/l (souche CBS111 portant la mutation ptsG22 dérivée de la souche MC 4100) et ceux de Picon et al. (2005) pour une souche portant la mutation ΔptsG::Km sur milieu minimum PC avec 2 g/l de glucose (souche PPA 291 dérivée de MC 4100).

# I.2 / Etude de la croissance cellulaire en milieu riche Luria Bertani (LB) supplémenté en glucose et influence de différents paramètres de cultures

Nous avons alors comparé le comportement des 2 souches TP 2503 et TP 2504 en milieu LB

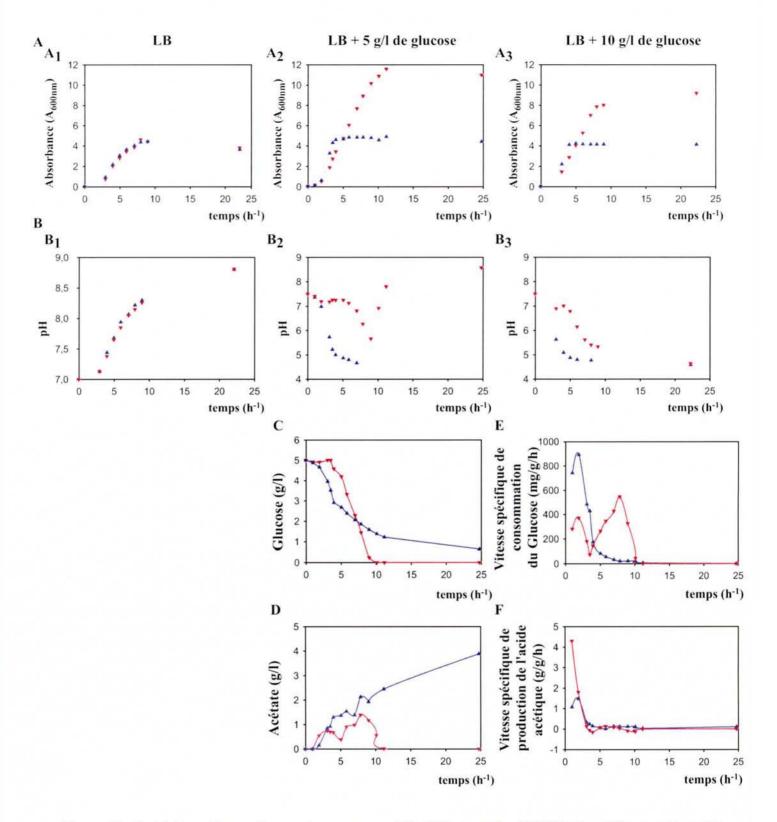


Figure 79 : Suivi des cultures des souches sauvages TP 2503 et mutées TP 2504 (ptsG22) en milieu riche LB complémenté avec différentes concentrations en glucose.

Les souches sauvages et mutées ont été cultivées en milieu riche supplémenté avec 0, 5 ou 10 g/l de glucose. Les triangles bleus correspondent à la souche sauvage, les triangles rouges à la souche mutée *ptsG22*.

- A : Suivi de l'absorbance à 600 nm en fonction du temps.
- B: Suivi du pH en fonction du temps.
- C : Suivi de la concentration en glucose en fonction du temps.
- D : Suivi de la concentration en acide acétique en fonction du temps.
- E : Vitesse spécifique de consommation du glucose en fonction du temps
- F: Vitesse spécifique de production d'acide acétique en fonction du temps.

contenant différentes concentrations en glucose. Les précultures ont été réalisées en milieu LB non complémenté en glucose. Ensuite, nous avons testé l'influence de la concentration en phosphate minéral.

### I.2.a / Etude de la croissance cellulaire en milieu riche (LB) supplémenté avec différentes concentrations en glucose

L'ensemble des résultats obtenus pour deux expériences indépendantes réalisées pour chaque condition de culture est présenté dans le tableau 6. Les cultures ont été réalisées en absence de glucose ou en présence de glucose à 5 et 10 g/l.

|                            | LB Glc 0 g/l      |                   | LB Glc 5 g/l      |                    | LB Glc 10 g/l     |                   |
|----------------------------|-------------------|-------------------|-------------------|--------------------|-------------------|-------------------|
|                            | TP 2503           | TP 2504           | TP 2503           | TP 2504            | TP 2503           | TP 2504           |
| A <sub>600nm</sub>         | 3,42<br>(+/-0,31) | 3,48<br>(+/-0,48) | 4,81<br>(+/-0,77) | 10,37<br>(+/-1,19) | 4,44<br>(+/-0,42) | 9,68<br>(+/-0,68) |
| biomasse X<br>(g/l)        | 1,54<br>(+/-0,14) | 1,32<br>(+/-0,18) | 2,17<br>(+/-0,34) | 3,94<br>(+/-0,45)  | 2,00<br>(+/-0,19) | 3,67<br>(+/-0,26) |
| μ (h <sup>-1</sup> )       | 0,41<br>(+/-0,09) | 0,38<br>(+/-0,07) | 1,43<br>(+/-0,01) | 1,11<br>(+/-0,03)  | 1,38              | 1,07              |
| pH du milieu<br>de culture | 8,85<br>(+/-0,07) | 8,85<br>(+/-0,06) | 4,66<br>(+/-0,08) | 8,64<br>(+/-0,06)  | 4,57<br>(+/-0,02) | 4,57<br>(+/-0,10) |
| R <sub>X/Glc</sub> (g/g)   |                   |                   | 0,43<br>(+/-0,07) | 0,78<br>(+/-0,09)  | 0,19<br>(+/-0,02) | 0,36<br>(+/-0,02) |

Tableau 6 : Rendement en biomasse obtenu pour les cultures des souches d'*E. coli* sauvage TP 2503 et mutée TP 2504 (*ptsG22*) en milieu riche LB supplémenté avec du glucose à différentes concentrations de glucose.

Les cultures des deux souches ont été réalisées au moins deux fois.

En absence de glucose, des courbes de croissance similaires étaient obtenues pour les 2 souches cultivées en milieu LB. Aucune différence marquante n'était observée au niveau du taux de croissance, de la densité cellulaire en phase stationnaire de croissance et de l'évolution du pH du milieu de culture (Fig. 79, A<sub>1</sub> et B<sub>1</sub>, Tableau 6). Ceci pourrait s'expliquer par la faible utilisation de la voie de la glycolyse en milieu LB.

En présence de glucose à 5 ou 10 g/l, le taux de croissance de la souche mutée *ptsG22* (TP 2504) était un peu plus faible que celui de la souche sauvage (TP 2503) (diminution de l'ordre de 20 %) (Fig. 79, A<sub>2</sub> et A<sub>3</sub>, Tableau 6). Par contre, l'ajout de glucose conduisait à une densité cellulaire en phase stationnaire beaucoup plus élevée pour la souche mutée que pour la souche sauvage [facteur 2,2 (+/-0,15)], la plus forte densité cellulaire étant obtenue pour 5 g/l de glucose (10,4 A<sub>600nm</sub>) (Fig. 79A et Tableau 6). Nous avons mesuré la concentration en glucose résiduel et la concentration en acide acétique présents dans les milieux de culture par CLHP (Fig. 79, C et D). Comme attendu, du fait de la

plus forte perméation du glucose pendant la phase exponentielle de croissance, la souche sauvage consommait le glucose plus rapidement que la souche mutée et produisait de l'acide acétique plus rapidement que la souche mutée, ce qui correspondait à une baisse plus rapide du pH du milieu de culture (Fig. 79, B<sub>2</sub>, C et D). Dans le cas de la souche mutée, la diminution de pH du milieu de culture semblait se faire de manière décalée dans le temps par rapport à la souche sauvage (Fig. 79, B<sub>2</sub> et D). La croissance bactérienne des deux souches s'arrêtait dès que le pH du milieu de culture était inférieur à 5 unités de pH, ce qui correspond à une concentration en acide acétique de l'ordre de 1,25 g/l (25 mM). Pour une concentration en glucose de 5 g/l, une remontée du pH du milieu de culture était observée en fin de phase exponentielle de croissance pour la souche mutée. Cette remontée était associée à une diminution de la concentration en acide acétique (Fig. 79, C et D). Ainsi, la souche mutée consommait donc l'acide acétique après avoir consommé le glucose, ceci sans doute parce que le pH du milieu de culture atteint n'inhibait pas la croissance bactérienne. Ceci explique, sans doute, le meilleur rendement en biomasse obtenu par la souche mutée pour cette concentration en glucose (Tableau 6). Les cellules de la souche mutée entraient en phase stationnaire lorsque tout l'acide acétique produit était consommé (Fig. 79, A<sub>2</sub> et D).

Pour la souche mutée, la vitesse spécifique de production d'acide acétique est très faible en phase exponentielle de croissance et élevée lors de la phase de latence qui précédait la phase exponentielle de croissance. Ces résultats sont similaires à ceux que nous avons obtenus en fermentation continue en condition de limitation en phosphate : à savoir, une forte production d'acide acétique lorsque le taux de croissance est faible et une très faible production d'acide acétique lorsque le taux de croissance est supérieur à 0,2 h<sup>-1</sup>.

Le milieu LB supplémenté avec 5 g/l de glucose nous a donc paru intéressant pour l'utilisation de la souche mutée ptsG22 pour la production de protéines recombinantes (TP 2504). En effet, l'acide acétique produit est consommé et on peut atteindre une biomasse élevée dans ces conditions de croissance.

De manière surprenante, en comparant la croissance d'un couple de souches isogéniques portant l'une le gène ptsG sauvage, l'autre le gène ptsG22, (en milieu LB complémenté avec 2 et 4 g/l de glucose), Chou et al. n'avaient observé aucune différence du taux de croissance, ni de la densité cellulaire en phase stationnaire de croissance (Chou et al., 1994). Ceci pourrait s'expliquer peut-être par un contexte génétique différent. Les résultats que nous obtenons sont tout à fait cohérents avec les propriétés respectives de perméation du glucose des deux souches que nous étudions.

### 1.2.b / Influence de la concentration en phosphate minéral sur la croissance en milieu LB supplémenté avec du glucose à différentes concentrations

Les données de la littérature indiquent que pour une concentration en glucose de 4 g/l en milieu

riche LB, l'ajout d'un tampon phosphate minéral (100 mM) permet de maintenir le pH du milieu de culture autour de 7,0 (Hosono et al., 1995). Dans ces conditions, la croissance ne s'arrête que lorsque le glucose et l'acide acétique sont entièrement consommés, ce qui se traduit par une plus forte densité cellulaire en phase stationnaire. Nous avons donc testé l'effet de l'ajout d'un tampon phosphate minéral (100 mM) sur la croissance des souches sauvage et mutées, en milieu LB complémenté avec du glucose à différentes concentrations (Fig. 80 et Tableau 7). Le phosphate était ajouté sous forme H<sub>2</sub>PO<sub>4</sub> à une concentration finale de 0,1 M.

|                            | LBP     |         | LBP G   | lc 5 g/l | LBP Glc 10 g/ |         | LBP Glc 15 g/l |         |
|----------------------------|---------|---------|---------|----------|---------------|---------|----------------|---------|
|                            | TP 2503 | TP 2504 | TP 2503 | TP 2504  | TP 2503       | TP 2504 | TP 2503        | TP 2504 |
| Glucose<br>restant (g/l)   | 0,00    | 0,00    | 0,00    | 0,00     | 0,00          | 2,20    | 4,18           | 6,85    |
| % de glucose<br>consommé   | 0,00    | 0,00    | 100,00  | 100,00   | 100,00        | 78,03   | 72,16          | 54,36   |
| A <sub>600nm</sub>         | 4,96    | 4,48    | 11,90   | 11,58    | 8,80          | 11,68   | 8,52           | 10,84   |
| Biomasse (X)<br>(g/l)      | 2,23    | 1,70    | 5,35    | 4,40     | 3,96          | 4,43    | 3,83           | 4,12    |
| pH du milieu de<br>culture | 8,38    | 8,40    | 6,89    | 7,10     | 5,60          | 6,03    | 4,87           | 5,68    |
| acide acétique<br>(g/l)    | 0,00    | 0,00    | 1,71    | 0,74     | 5,12          | 3,38    | 4,54           | 4,62    |
| R <sub>X/Glc</sub> (g/g)   |         |         | 0,92    | 0,89     | 0,34          | 0,45    | 0,22           | 0,28    |
| R Aac/X (g/g)              | 0,00    | 0,00    | 0,37    | 0,17     | 1,51          | 0,75    | 1,38           | 1,10    |
| R Aac/Glc (g/g)            | 0,00    | 0,00    | 0,34    | 0,15     | 0,51          | 0,43    | 0,42           | 0,57    |

Tableau 7 : Influence de la présence d'un tampon phosphate dans le milieu de culture sur la croissance des souches sauvage TP 2503 et mutée TP 2504 en milieu LB supplémenté par du glucose à différentes concentrations

En milieu LB sans glucose, l'ajout d'un tampon phosphate minéral améliorait faiblement la densité cellulaire en phase stationnaire de croissance (passage de 3,4 à 4,7 A<sub>600nm</sub>), et diminuait légèrement le pH du milieu de culture en phase stationnaire (passage de 8,8 à 8,4) (Fig. 79<sub>1</sub> et 80<sub>1</sub> et Tableau 7).

En milieu LB supplémenté avec 5 g/l de glucose, la densité cellulaire en phase stationnaire de croissance était similaire pour les deux souches de l'ordre de 12 A<sub>600nm</sub> (Tableau 7 et Fig. 80A). Ainsi, pour la souche sauvage, l'ajout d'un tampon phosphate avait des conséquences sur l'augmentation de la densité cellulaire en phase stationnaire de croissance similaire à celle observée sans ajout de phosphate par l'introduction de la mutation ptsG22. Pour la souche sauvage, deux vitesses de croissances successives étaient observées. En début de croissance, entre 0 et 5 h, une forte croissance était associée à une acidification du milieu. Au bout de 5 h de croissance, le pH du milieu de culture atteignait une valeur proche de 6,3 et les cellules semblaient alors modifier leur métabolisme : la

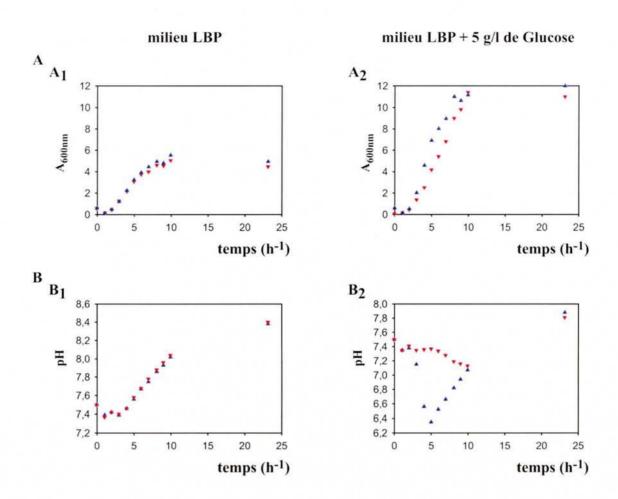


Figure 80 : Influence de la présence d'un tampon phosphate dans le milieu de culture sur la croissance cellulaires des souches TP 2503 et TP 2504 en milieu LB supplémenté ou non avec 5 g/l de glucose.

Les souches sauvages et mutées ont été cultivées en milieu riche LB en présence d'un tampon phosphate minéral (0,1 M), avec ou sans glucose (5 g/l).

A : Suivi de l'absorbance à 600 nm.

B: Suivi du pH.

vitesse de croissance diminuait et devenait inférieure à celle de la souche *ptsG22*, une remontée du pH du milieu de culture était alors observée. Cette remontée du pH du milieu de culture pourrait correspondre à une consommation parallèle du glucose résiduel et de l'acide acétique ou simplement de l'acide acétique. La faible amplitude des variations du pH du milieu de culture en présence de phosphate peut s'expliquer par le pouvoir tampon exercé par l'ajout de phosphate minéral (Fig. 79<sub>2</sub> et 80<sub>2</sub> et Tableau 7).

Lorsque les 2 souches étaient cultivées en milieu LB, en présence d'un tampon phosphate et avec du glucose à plus forte concentration (Tableau 7), le glucose n'était pas entièrement consommé, même pour la souche mutée ; sa consommation maximale était de l'ordre de 9 g/l de glucose. Aucune augmentation significative de la biomasse n'était observée pour la souche mutée, pour des concentrations en glucose supérieures à 5 g/l. Donc, les conditions les plus favorables de croissance aussi bien pour la souche mutée que pour la souche sauvage en présence de 0,1 M de H<sub>2</sub>PO<sub>4</sub> étaient le milieu LB en présence de 5 g/l de glucose.

Comme le phosphate minéral apporté sous forme H<sub>2</sub>PO<sub>4</sub> pouvait agir à la fois par son effet tampon et par le fait qu'il représente une source de phosphate pour les bactéries, nous avons comparé l'effet de l'utilisation de deux tampons, Tris-HCl et HEPES sur la croissance des deux souches à l'effet observé par le phosphate. Comme le montre la Figure 81, l'utilisation du phosphate a un effet positif sur la densité cellulaire supérieur à celui d'un simple tampon.

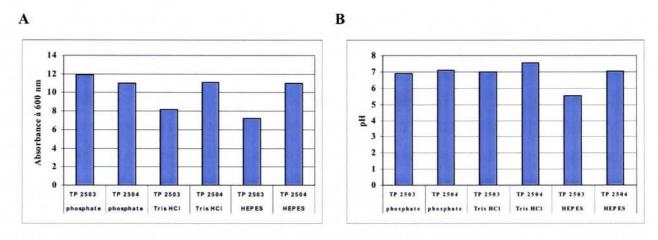


Figure 81 : Influence de la présence de différents tampons dans le milieu de croissance sur la densité cellulaire et le pH du milieu de culture en phase stationnaire de croissance des souches TP 2503 et TP 2504 en milieu LB supplémenté avec du glucose à 5 g/l

Les 2 souches ont été cultivées en milieu LB avec du glucose à une concentration de 5 g/l en présence d'un tampon phosphate minéral, d'HEPES ou de Tris-HCl à une concentration finale de 0,1 M.

A : Absorbance à 600 nm pour les souches souvages et mutées par rapport au tampon utilisé.

B : pH du milieu de culture pour les souches souvages et mutées par rapport au tampon utilisé.

Nous avons alors débuté une étude visant à étudier les mécanismes expliquant la différence de comportement des 2 souches en milieu LB supplémenté en glucose sans ajout de phosphate minéral.

II / Etude des mécanismes expliquant l'obtention d'une forte densité cellulaire en phase stationnaire de croissance pour la souche d'*E. coli* TP 2504

#### II.1 / Influence de l'expression de Mlc sur les propriétés de la souche TP 2504

La mutation ptsG22 consiste en une délétion de 4 bases au niveau de la région IIC de la protéine EIIBCGlc, ce qui entraîne un décalage du cadre de lecture pour la partie C-terminale de la protéine (Plumbridge, 1999). On sait que ce décalage du cadre de lecture invalide les propriétés de transporteur de la protéine EIIBC Glc (Curtis et Epstein, 1975). En plus, cette mutation semble conduire à une absence de fixation de la protéine Mlc sur ce transporteur. En effet, dans une souche sauvage en absence de glucose, l'inhibiteur transcriptionnel Mlc se fixe sur la protéine EIIBCGlc non phosphorylée. Il y a alors levée de l'inhibition transcriptionnelle que Mlc exerce sur de nombreux gènes, tels que ptsHI codant les enzymes générales EI et HPr des systèmes PTS, manXYZ codant les enzymes du système PTS du mannose, malT codant l'activateur transcriptionnel du régulon maltose (pour revue, Plumbridge, 2002). Les propriétés de la souche d'E. coli TP 2504 doivent donc être à la fois liées à l'absence d'activité de EIIBCGle dans le transport du glucose mais aussi à la perte de capacité de cette protéine à piéger Mlc. On peut en effet penser que dans la souche TP 2504, la protéine Mlc qui n'est plus piégée par EIIBC Glc, limite l'expression de différents gènes, en particulier, l'opéron manXYZ codant les enzymes du système PTS du mannose. Or, on suppose que ce système est utilisé par le glucose lorsque le système PTS spécifique du glucose n'est pas actif (Curtis et Epstein, 1975 ; Picon et al., 2005). Afin de vérifier cette hypothèse, nous avons testé l'effet de la non expression de la protéine Mlc sur la croissance d'une souche portant la mutation ptsG22. Afin d'avoir des points de comparaison, les expériences ont été réalisées sur 4 souches : d'une part, la souche JM101 et une souche isogénique IBPC 1012, dont le gène mlc est inactivé, d'autre part, la souche IBPC5321 ne produisant pas la protéine Mlc, et une souche isogénique IBPC1000 portant en plus la mutation ptsG22. Nous avons mesuré la densité cellulaire et le pH du milieu de culture après 14 h de culture en milieu LB supplémenté avec du glucose à une concentration de 5 g/l de glucose (Tableau 8).

|              | JM101 | IBPC 1012 (mlc <sup>-</sup> ) | IBPC 5321(mlc <sup>-</sup> ) | IBPC 1000 (mlc <sup>-</sup> , ptsG22) |
|--------------|-------|-------------------------------|------------------------------|---------------------------------------|
| $A_{600nm}$  | 4,7   | 4,1                           | 5,0                          | 4,7                                   |
| pH du milieu | 4,7   | 4,6                           | 4,75                         | 4,7                                   |
| de culture   |       |                               |                              |                                       |

Tableau 8 : Influence de l'inactivation du gène ptsG22 sur la densité cellulaire et le pH du milieu de culture en phase stationnaire de croissance lorsque les souches testées n'expriment pas la protéine Mlc

Nous constatons que les 4 souches ont des comportements similaires. Donc, lorsque la mutation ptsG22 est accompagnée d'une mutation invalidant le gène mlc, la densité cellulaire et le pH du milieu de culture en phase stationnaire de croissance sont similaires à ceux obtenus pour des souches portant un gène ptsG sauvage. Ces données sont en accord avec l'idée que la protéine Mlc joue un rôle très important dans les propriétés de la souche TP 2504. On peut imaginer que du fait de l'absence de protéine Mlc dans la souche IBPC 1000 (mlc, ptsG22), aucune répression ne s'exerce sur la transcription des protéines spécifiques des systèmes PTS autres que celui spécifique du glucose, en particulier le système PTS spécifique du mannose, d'où une forte perméation du glucose par ce système PTS et l'acidification du milieu liée à son métabolisme. Ceci n'est pas le cas pour la souche TP 2504, où le glucose est métabolisé plus lentement, ce qui ne coduit pas à une rapide acidification du milieu, bloquant la croissance et la consommation de l'acide acétique produit.

Un autre moyen de tester l'implication de la protéine Mlc dans les caractéristiques de la souche TP 2504 était de comparer les effets dans les souches TP 2503 et TP 2504 d'une surexpression de la protéine Mlc. Pour cela, nous avons utilisé le plasmide recombinant pTZ(Mlc) permettant de surexprimer la protéine Mlc. Ce plasmide nous a été fourni par J. Plumbridge (IBPC, Paris), le plasmide vide correspondant (pTZ19R) nous servait de contrôle. Les deux souches TP 2503 et TP 2504 ont chacune été transformées par ces plasmides et les transformants ont été cultivés 14 h en milieu LBA supplémenté avec du glucose à une concentration de 5 g/l.

|              | TP 2503     | TP 2504         | TP 2503   | TP 2504       |
|--------------|-------------|-----------------|-----------|---------------|
|              | WT (pTZ19R) | ptsG22 (pTZ19R) | WT (pMlc) | ptsG22 (pMlc) |
| $A_{600nm}$  | 4,4         | 8,0             | 8,9       | 8,7           |
| pH du milieu | 4,4         | 6,4             | 6,1       | 7,4           |
| de culture   |             |                 |           |               |

Tableau 9 : Influence de la surproduction de la protéine Mlc sur la densité cellulaire et sur le pH du milieu de culture en phase stationnaire de croissance en milieu LBA supplémenté avec du glucose à une concentration de 5 g/l

Les données présentées dans le tableau 9 montrent une nette augmentation du pH du milieu de culture et de la densité cellulaire en phase stationnaire pour la souche TP 2503 transformée par le plasmide surexprimant Mlc par rapport à la même souche transformée par le plasmide vide. La protéine Mlc surexprimée limite l'expression de l'ensemble des systèmes PTS pouvant transporter le glucose, d'où un ralentissement sans doute de la consommation et une faible acidification. On note que la surexpression de Mlc dans la souche TP 2504 conduit à une augmentation du pH du milieu de culture en phase stationnaire. Il semble donc que la quantité de protéine Mlc exprimée à partir du gène endogène ne soit pas suffisante pour qu'un blocage complet de l'expression des gènes des systèmes PTS soit mis en place.

L'ensemble de nos données suggérait très fortement que la protéine Mlc joue un rôle majeur dans les propriétés de la souche ptsG22.

# II.2 / Etude de l'effet de la surproduction d'un transporteur ${\sf EIIBC}^{\sf Glc}$ sauvage ou muté

Une autre possibilité de distinguer les effets liés à l'absence d'activité de transporteur de EIIBC<sup>Glc</sup> à ceux liés à l'absence de piégeage de la protéine Mlc par EIIBC<sup>Glc</sup> était de produire en *trans* dans la souche TP 2504 un transporteur EIIBC<sup>Glc</sup> sauvage ou un transporteur EIIBC<sup>Glc</sup> portant une mutation qui affecte uniquement la fonction de transport. Pour cela, nous avons utilisé les plasmides pTSG4 et pTSG4-421, construits par Nuoffer *et al.*, qui expriment respectivement la protéine EIIBC<sup>Glc</sup> sauvage et la protéine EIIBC<sup>Glc</sup> mutée en position 421 (Cys421Ser). Cette mutation concerne la cystéine phosphorylée lors du transport du glucose. Cette mutation semble abolir spécifiquement la fonction de transporteur (Nuoffer *et al.*, 1988).

Nous avons donc transformé les souches TP 2503 et TP 2504 avec les plasmides pTSG4 et pTSG4-421 qui nous avaient été fournis par B. Erni (Université de Bern, Suisse). Les souches transformées ont été cultivées pendant 14 h en milieu LBA contenant du glucose à 5 g/l. Les résultats obtenus sont présentés dans le tableau 10.

|             | TP 2503     | TP 2504        | TP 2503        | TP 2504            |
|-------------|-------------|----------------|----------------|--------------------|
|             | WT (pTSG4)  | ptsG22 (pTSG4) | WT (pTSG4-421) | ptsG22 (pTSG4-421) |
| $A_{600nm}$ | 3,8         | 3,86           | 4,2            | 5,5                |
| pH du r     | milieu 4,62 | 4,6            | 4,6            | 4,6                |
| de cultu    | ıre         |                |                |                    |

Tableau 10 : Influence de l'expression *en trans* d'une protéine EIIBC<sup>Glc</sup> sauvage ou mutée (Cys421Ser) sur la croissance des souches TP 2503 et TP 2504 et leur propriété d'acidification du milieu

Comme on pouvait s'y attendre, la surexpression du transporteur EIIBC<sup>Gle</sup> actif dans la souche TP 2504 lui confère des propriétés similaires à celles de la souche TP 2503 sauvage, à savoir une faible densité cellulaire en phase stationnaire de croissance et une forte acidification du milieu. La surexpression du transporteur muté dans la souche sauvage a peu d'effet sur la croissance et le pH du milieu de culture. Par contre, l'expression du transporteur non fonctionnel dans la souche TP 2504 confère un phénotype proche de celui de la souche sauvage en ce qui concerne les propriétés d'acidification : à savoir, un fort abaissement du pH du milieu de culture en phase stationnaire. La densité cellulaire en phase stationnaire était plus élevée que celle observée pour les deux autres souches, mais comme l'expérience n'a été réalisée qu'une seule fois, il sera important de la répéter pour savoir si cette augmentation était significative.

Nos données pourraient s'expliquer par la fixation de la protéine Mlc sur le transporteur muté en position 421, ce qui est en accord avec les données de Nuoffer *et al.* (1988). Ces auteurs ont en effet montré que la surproduction de la protéine EIIBC<sup>Glc</sup><sub>Cys421Ser</sub> stimule la croissance sur maltose d'une souche d'*E. coli* dans laquelle le transporteur PTS du glucose n'est pas fonctionnel. L'interaction de la protéine Mlc avec la protéine EIIBC<sup>Glc</sup><sub>Cys421Ser</sub> lèverait la répression transcriptionnelle que Mlc exerce sur le gène *malT* codant l'activateur transcriptionnel du régulon maltose (Decker *et al.*, 1998).

L'ensemble de ces résultats suggère donc très fortement que la faible production d'acide acétique par la souche mutée *ptsG22* résulte sans doute beaucoup plus de l'inhibition générée par la protéine Mlc sur l'expression du transporteur spécifique du mannose, du fait de l'absence d'interaction de cette protéine avec la protéine EIIBC<sup>Glc</sup> que de l'invalidation de la fonction de transport du glucose de EIIBC<sup>Glc</sup>.

## II.3 / Effet d'une invalidation des gènes codant les enzymes EI et HPr et EIIA<sup>Glc</sup> sur la croissance d'*E. coli* en milieu LBGlc (5g/l)

Une autre façon de tester l'hypothèse ci-dessus était d'invalider le gène de la seconde protéine du système spécifique du glucose, EIIA<sup>Glc</sup>. En effet, cette protéine n'est pas impliquée directement dans la rétention de la protéine Mlc en absence de glucose. Au contraire, en son absence, le transporteur EIIBC<sup>Glc</sup> n'est pas phosphorylé. Il piège donc constitutivement la protéine Mlc. Si les propriétés de la souche TP 2504 sont essentiellement dues à l'incapacité de EIIBC<sup>Glc</sup> de piéger Mlc, on peut s'attendre à ce que l'invalidation du gène EIIA<sup>Glc</sup> ne confère pas les caractéristiques observées pour la mutation *ptsG22*. C'est ce que nous avons vérifié en utilisant les souches TP 2503 et TP 2862, dont le gène *crr* codant la protéine EIIA<sup>Glc</sup> est invalidé.

La souche TP 2862 conduit comme la souche TP 2503 à une acidification du milieu en phase

stationnaire de croissance (Tableau 11). La densité cellulaire dans cette phase était légèrement supérieure pour la souche  $\Delta crr$  en comparaison de celle observée pour la souche sauvage, mais là encore, l'expérience n'a été réalisée qu'une seule fois. Elle devra être à nouveau répétée pour vérifier la signification de cette faible différence de densité cellulaire.

Nos résultats sont en accord avec le fait que, en absence de EIIA<sup>Glc</sup>, le transporteur EIIBC<sup>Glc</sup> n'est pas phosphorylé et donc Mlc interagit alors avec le transporteur EIIBC<sup>Glc</sup>. On n'a donc pas d'inhibition du transporteur du mannose et des autres systèmes PTS : le glucose est transporté par le PTS du mannose, ce qui conduit à une acidification du milieu. Ces données sont aussi en accord avec l'idée que la perte des 2 activités de la protéine EIIBC<sup>Glc</sup> dans la souche TP 2504 sont requises pour qu'il n'y ait pas d'acidification du milieu en phase stationnaire de croissance.

Une autre confirmation possible pouvait être obtenue en invalidant les gènes des deux protéines générales à tous les systèmes PTS d'*E. coli*, les protéines EI et HPr. Dans ce cas, on s'attendait à avoir une très faible perméation du glucose, puisque même le système PTS spécifique du mannose emprunté par le glucose lorsque le système PTS spécifique du glucose est inactif n'est plus fonctionnel. Comme on peut le voir dans le tableau 11, l'absence d'expression de la protéine EI dans la souche TP 2819 et l'absence d'expression de la protéine HPr dans la souche TP 2811 s'accompagnent d'une absence d'acidification du milieu en phase stationnaire. Par contre, la densité cellulaire est un peu plus élevée que celle observée pour la souche TP 2862. Ici encore, les expériences devront être répétées une seconde fois.

|              | TP 2503 | TP TP 2862   | TP 2819          | TP 2811   |
|--------------|---------|--------------|------------------|-----------|
|              | WT      | $\Delta crr$ | $\Delta ptsIcrr$ | ΔptsHIcrr |
| $A_{600nm}$  | 4,2     | 5,8          | 5,2              | 5,7       |
| pH du milieu | 4,6     | 4,7          | 8,1              | 8,3       |
| de culture   |         |              |                  |           |

Tableau 11 : Influence des enzymes générales du système PTS, EI et HPr, et de l'enzyme EIIA c spécifique du PTS du glucose, en milieu LB supplémenté avec du glucose à une concentration de 5 g/l sur la densité cellulaire et le pH du milieu de culture obtenus en phase stationnaire de croissance

Les différentes souches ont été cultivées pendant 14 h.

La souche TP 2504 nous paraissait avoir des propriétés intéressantes pour la production de protéines recombinantes, ce que nous avons ensuite testé.

## III / Comparaison des capacités de production de protéines recombinantes par des souches exprimant un gène ptsG sauvage ou muté

Le fait que la souche TP 2504 atteigne des densités cellulaires élevées et n'acidifie pas le milieu de culture après 14 h de culture en milieu LB supplémenté avec du glucose à 5 g/l, paraissait être une propriété intéressante pour la production de protéines recombinantes. En effet, l'acidification du milieu de culture conduit à un état de stress des cellules qui n'est pas propice à une production efficace de protéines recombinantes. D'une manière générale, différents facteurs influencent l'efficacité de production de protéines recombinantes par *E. coli*: le contexte génétique de la souche utilisée, le type de plasmide utilisé, le promoteur permettant d'exprimer le gène d'intérêt, la stabilité de l'ARNm correspondant, l'effet sur la souche de la protéine produite (pour revues, Baneyx, 1999, Hannig et Makrides, 1998, Makrides, 1996, Sorensen et Mortensen, 2005, Weickert *et al.*, 1996).

Nous avons donc comparé la production de protéines dans les deux souches TP 2503 et TP 2504 (ptsG22) transformées par différents plasmides. Pour comparaison, nous avons aussi introduit dans les essais la souche d'E. coli DH5α (supE44, ΔlacU169, hsdR17, recA1, endA1 gyrA96, thi-1, relA1) utilisée régulièrement au laboratoire pour la production de protéines de la famille des Aldéhydes Déshydrogénase. Les tests ont porté sur la production de l'E4PDH, la PGK, la FBA, et la GAPDH d'E. coli, qui ont fait l'objet des études décrites dans le premier chapitre de notre manuscrit. Pour cela, les plasmides pEPF (article 1) et pBS::EcogapA (Charpentier, 1994) ont été utilisés. Dans ce dernier plasmide, l'ORF de la GAPDH d'E. coli est placée sous le contrôle de sa propre région promotrice, comprenant 4 promoteurs (Charpentier, 1994). B. Charpentier avait montré précédemment que cette région promotrice du gène gapA présente un intérêt pour la surproduction de protéines. Cette région avait en effet permis d'obtenir un niveau élevé d'interleukine 1-β chez E. coli en utilisant la souche E. coli TG1 (Charpentier, 1994). L'avantage d'utiliser l'ORF gapA pour les tests de production était de pouvoir mesurer l'activité de la protéine produite et ainsi d'évaluer l'efficacité de surproduction de protéines actives.

Les différents tests ont été réalisés sur des cultures de 50 ml réalisées en erlenmeyer de 250 ml, ensemencée à partir d'un clone obtenu sur boîte. Les cultures étaient réalisées 14 h à 37°C.

## III.1 / Comparaison des capacités de production de protéines recombinantes des 3 souches d'E. coli DH5α, TP 2503 et TP 2504, en milieu LBA, LBAGIc (5 g/l) et LBAPGIc (5 g/l)

En accord avec nos observations sur le plasmide pEPF (article 1), lorsque chacune des trois

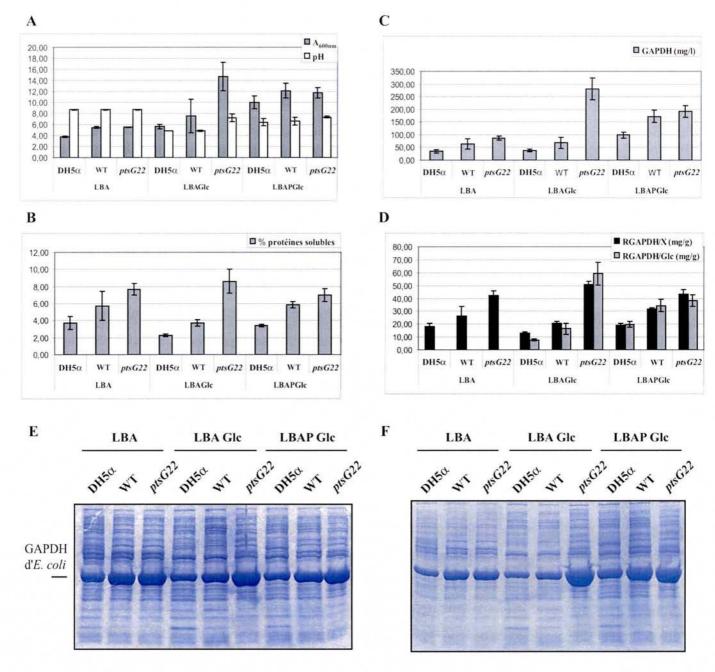


Figure 82 : Etude de la densité cellulaire, du pH du milieu en phase stationnaire de croissance, et du taux de GAPDH rapporté au taux de protéines solubles ou au volume de milieu de culture, lors de cultures des souches DH5α, TP2503 et TP2504 transformées par le plasmide pBS::*EcogapA* en milieu LBA, LBAGlc (5 g/l) ou LBAPGlc (5 g/l)

A : Comparaison des densités cellulaires et pH du milieu de culture obtenus en phase stationnaire de croissance pour les souches DH5α, TP 2503 et TP 2504, transformées par le plasmide pBS::*EcogapA* en milieu LBA, LBAGlc (5 g/l) et LBAPGlc (5 g/l).

B : Pourcentage de protéines solubles correspondant à la GAPDH en g de GAPDH par g de protéines solubles.

C : Quantité de GAPDH produite en mg par litre de milieu de culture.

D : Rendement de production de la GAPDH par rapport au glucose consommé et à la biomasse.

E : Fractionnement des protéines solubles par SDS-PAGE : des aliquots de  $15~A_{600nm}$  de cultures des 3 sauvages, transformées par le plasmide pBS::EcogapA cultivées en milieu LBAGlc (5 g/l) ont été prélevés en phase stationnaire de croissance. Après sonication, un volume de surnageant correspondant à 50 milli $A_{280nm}$  a été déposé sur un gel à 7,5 %.

F : Fractionnement des protéines solubles par SDS-PAGE : un volume de surnageant de sonication correspondant à 10 ml de culture cellulaire a été déposé dans chaque piste.

souches DH5α, TP 2503 et TP 2504 était transformée par le plasmide pEPF et cultivée une nuit en milieu LBA contenant 5 g/l de glucose, nous observions une grande variabilité de la densité cellulaire en phase stationnaire de croissance, d'une expérience à une autre (résultats non montrés). Ceci devait être dû à l'instabilité de ce plasmide que nous avons déjà mentionné dans l'article 1. Cette instabilité est sans doute liée à la très grande taille de l'ADN que nous avons inséré et du fait que c'est un dérivé du plasmide pBlueScript, plasmide à fort nombre de copies.

Nous avons donc concentré notre étude sur les souches DH5α, TP 2503 et TP 2504 transformées par le plasmide pBS::*EcogapA*, pour lesquelles une plus grande reproductibilité des résultats était obtenue.

En milieu LBA, nous notons peu de différences entre les souches, aussi bien au niveau de la densité cellulaire, que du pH du milieu de culture, en phase stationnaire de croissance (Fig. 82A). Nous observions seulement, une plus faible croissance de la souche DH5α par rapport aux deux autres souches, ce qui se traduisait aussi par une quantité plus faible de GAPDH, aussi bien lorsque l'activité GAPDH mesurée était rapportée au volume de culture (Fig. 82, B et E), que lorsqu'elle était rapportée à la quantité de protéines solubles (Fig. 82, C et F). Les deux souches TP 2503 et TP 2504 semblaient donc plus performantes au niveau production de GAPDH recombinante.

Lorsque le milieu LBA était supplémenté avec 5 g/l de glucose, à la fois la densité cellulaire en phase stationnaire et le taux de GAPDH rapporté à la biomasse étaient très supérieurs pour la souche TP 2504 par rapport aux deux autres souches (Fig. 82, A et D). La densité cellulaire en phase stationnaire était doublée en présence de la mutation ptsG22, de même que la production de GAPDH. Un rendement de 281 +/- 44 mg/l de culture était obtenu (Fig. 82C). Le rendement par gramme de biomasse est similaire à celui qui avait été obtenu par A. Gschaedler (Gschaedler, 1994) en utilisant les souches d'E. coli HB101 et C600GalK transformées par le plasmide pBR::EcoGAPDH cultivées en milieu LB modifié contenant 4 g/l de glucose.

L'ajout d'un tampon phosphate à 0,1 M dans le milieu de culture permettait d'obtenir une plus forte densité cellulaire pour les souches sauvages DH5α et TP 2503 [les A<sub>600nm</sub> en phase stationnaire de croissance passaient respectivement de 3,8 (+/-0,08) et 5,4 (+/-0,18) à 10 (+/-1,15) et 12,16 (+/-1,37)] (Fig. 83A). En ce qui concerne la production de GAPDH, les deux souches TP 2503 et TP 2504 avaient des performances beaucoup plus similaires dans ces conditions, ce qui est en accord avec les propriétés de croissance des deux souches décrites ci-dessus. Nous notons que leurs performances étaient nettement supérieures à celles de la souche DH5α. Le rendement de production de GAPDH était de 191 +/-23 mg/l de culture pour la souche TP 2504, contre 98 +/- 11 mg/l pour la souche DH5α.

## III.2 / Etude de l'influence de la concentration en glucose dans le milieu de culture sur la croissance des souches TP 2503 et TP 2504 recombinées et leur production de GAPDH

Nous avons tout d'abord testé si l'ajout de glucose dans le milieu de préculture influençait le comportement des souches TP 2503 et TP 2504, lorsqu'elles étaient cultivées en milieu LB supplémenté avec du glucose. L'avantage observé pour la souche TP 2504 par rapport à la souche TP 2503 au niveau de la densité cellulaire en phase stationnaire de croissance était perdu lorsque les souches étaient pré-cultivées en milieu LBA contenant 5 g/l de glucose (Fig. 83). De ce fait, nous avons réalisé toutes les précultures en absence de glucose.

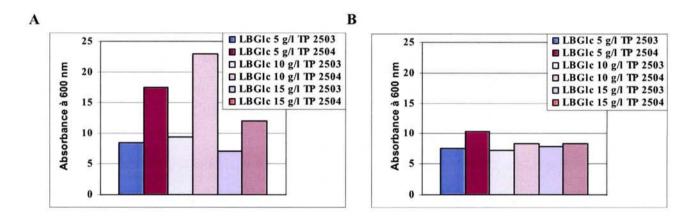


Figure 83 : Influence du milieu de préculture sur la densité cellulaire obtenue en phase stationnaire de croissance pour les souches TP 2503 et TP 2504 tranformées par le plasmide pBS:: EcogapA cultivées en milieu LBA supplémenté avec différentes concentrations en glucose.

A : Absorbance à 600 nm en phase stationnaire des cultures des deux souches TP 2503 et TP 2504 recombinées cultivées en milieu LBA supplémenté avec du glucose à différentes concentrations et avec des précultures réalisées en milieu LBA sans glucose.

B: Absorbance à 600 nm en phase stationnaire des cultures des deux souches TP 2503 et TP 2504 recombinées cultivées en milieu LBA supplémenté avec du glucose à différentes concentrations et avec des précultures réalisées en milieu LBA supplémenté avec du glucose à 5 g/l.

Nous avions vu que sans ajout de tampon phosphate, la concentration de 5 g/l en glucose était plus favorable qu'une concentration de 10 g/l en glucose pour la croissance de la souche TP 2504. Nous avons testé ce qu'il en était pour la production de protéines recombinantes. Comme on peut le constater dans la Figure 84, quelle que soit la concentration en glucose utilisée, une meilleure production de GAPDH était toujours obtenue pour la souche TP 2504 par rapport à la souche TP 2503.

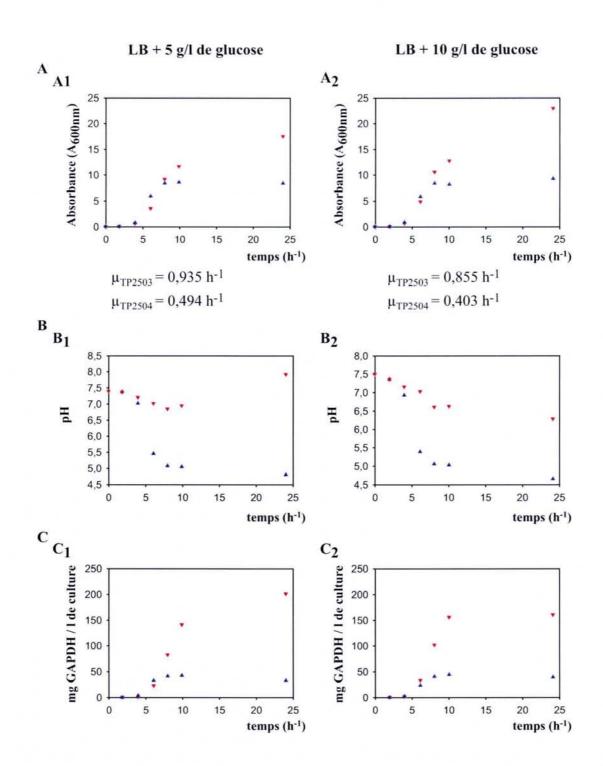


Figure 84 : Suivi des cultures des souches TP 2503 et TP 2504 transformées par le plasmide pBS::*EcogapA* en milieu LB supplémenté avec différentes concentrations en glucose.

Pour chacune des 2 concentrations en glucose, nous avons suivi l'évolution de la densité cellulaire à 600 nm, du pH du milieu de culture et de la quantité de GAPDH produite par litre de culture. Les triangles bleus correspondent à la souche sauvage (TP 2503), alors que les triangles rouges correspondent à la souche mutée (TP 2504).

A : Suivi de l'absorbance à 600 nm en fonction du temps.

B : Suivi du pH du milieu de culture en fonction du temps.

C : Quantité de GAPDH produite en mg par litre de milieu de culture en fonction du temps.

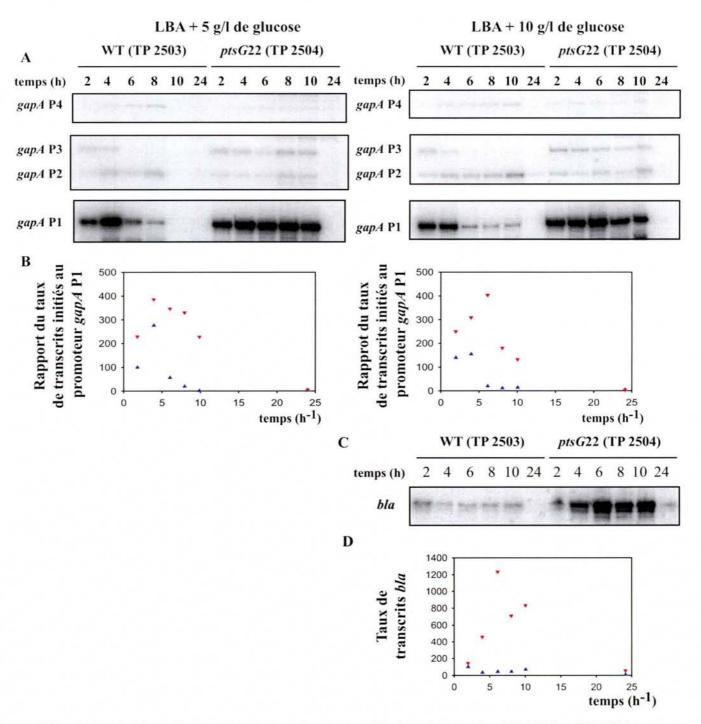


Figure 85 : Analyse du taux des transcrits gapA et bla dans les souches TP 2503 et TP 2504 transformées par le plasmide pBS::EcogapA au cours des cultures réalisées en milieu LBA supplémenté du glucose avec 5 et 10 g/l.

A : Analyse par extension de l'amorce O-GAP des transcrits issus contenus dans les cellules des souches TP 2503 et TP 2504, transformées par le plasmide pBS::*Eco*GAPDH, au cours du temps. Les 2 souches transformées ont été cultivées en milieu LBA contenant 5 ou 10 g/l de glucose. Les échantillons correspondant à 15 Unités A<sub>600nm</sub> ont été extraits au cours du temps, puis les ARN totaux ont été analysés par extension d'amorce O-GAP. Les ADNc ont ensuite été fractionnés en gel de polyacrylamide dénaturant à 5 %.

B : Pourcentage de transcrits initiés au promoteur *gapA* P1. Les taux des différents transcrits ont été quantifiés à l'aide d'un Phosphorimager, puis rapportés par rapport au taux de transcrits initiés au promoteur *gapA* P1 de la souche sauvage à t=2h.

C: Analyse par extension d'amorce du taux des transcrits *bla* codant la  $\beta$ -lactamase.

D : Variation du taux des transcrits *bla* au cours de la croissance. Le taux de transcrits *bla* à t=2h de la souche sauvage a été utilisé comme référence (valeur 100).

Néanmoins, comme attendu, après les résultats précédents, la production finale de GAPDH de la souche TP 2504 était légèrement supérieure pour une concentration en glucose de 5 g/l par rapport à celle obtenue pour une concentration de 10 g/l. D'après l'évolution de l'activité de la GAPDH au cours de la culture, il semble que pour la concentration en glucose de 5 g/l, la production de GAPDH se poursuive au-delà de 10 h de culture.

Afin d'essayer d'expliquer les différences de production de GAPDH dans les deux souches, nous avons étudié l'évolution du taux des ARNm issus du gène gapA dans les échantillons de culture prélevés tout au long des cultures, pour lesquels nous avions déjà mesuré la densité cellulaire et le pH du milieu de culture (Fig. 84). Pour cela, après extraction des ARN totaux, 20 µg d'ARN total ont été analysés par extension d'amorces, ceci en utilisant comme amorce l'oligonucléotide O-GAP complémentaire au segment des ARNm gapA compris entre les positions 69 à 52 de l'ORF gapA. Cet oligonucléotide est complémentaire au 4 ARNm produits à partir des 4 promoteurs du gène gapA (P1 à P4) (Charpentier et Branlant, 1994). En parallèle, afin d'essayer d'avoir une évaluation de la variation possible du nombre de copies de plasmide au cours de la culture, nous avons étudié par la même approche de réverse transcription le taux d'ARNm bla produit à partir du gène bla codant la βlactamase de résistance à l'ampicilline, porté par le plasmide. Les résultats obtenus sont illustrés dans la Figure 85. Nous observions un taux élevé des transcrits gapA P1 et bla dans les échantillons de la souche TP 2504 prélevés après 2, 4, 6, 8 et 10 h de culture en milieu LBAGlc (5 g/l). Aucun transcrit issu du gène gapA n'était détecté après 24 h de culture et seulement une faible quantité d'ARNm bla était observée. En ce qui concerne la souche TP 2503, un faible taux de transcrit bla était détecté tout le long de la culture et le taux de transcrit gapA P1 chutait fortement après 4 h de culture, sans qu'il y ait une augmentation marquée du transcrit gapA P2, correspondant à une transcription par l'enzyme  $E\sigma^{32}$ , comme attendu lorsque les cellules sont soumises à un stress dû à l'acidification du milieu de culture (Gschaedler et al., 1999). Donc, soit les cellules avaient alors un faible taux de transcription général, soit beaucoup de cellules avaient perdu le plasmide durant la phase exponentielle de croissance. Cette expérience devra être répétée en essayant d'apporter des réponses à cette question. Néanmoins, nos données dans l'état montre un fort taux d'ARNm gapA P1 dans les cellules de la souche TP 2504, ce qui est étonnant, puisque le taux de transcription à partir du promoteur gapA P1 et/ou la stabilité des transcrits gapA P1 semblent dépendre de l'expression de la protéine EIIBC Glc d'après les données de B. Charpentier (Charpentier et al., 1998).

## III.3 / Etude de l'effet de l'oxygénation du milieu sur la croissance des souches recombinantes et leur production de GAPDH

On sait que la croissance d'E. coli est fortement dépendante de l'oxygénation du milieu. Il était

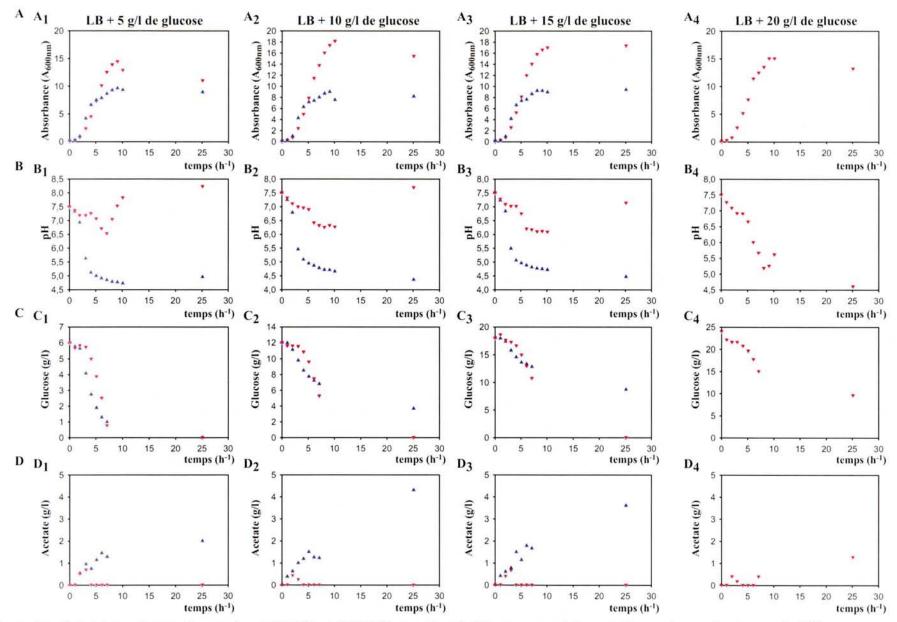


Figure 86 : Suivi des cultures des souches TP 2503 et TP 2504 en milieu LB fortement aéré, complémenté avec du glucose à différentes concentrations Les souches TP 2503 et TP 2504 ont été cultivées en milieu minimum supplémenté en glucose. Les triangles bleus correspondent à la souche TP 2503, alors que les triangles rouges correspondent à la souche TP 2504.

A : Suivi de l'absorbance à 600 nm en fonction du temps. B : Suivi du pH du milieu de culture en fonction du temps. C : Suivi de la consommation du glucose en fonction du temps. D : Suivi de la production d'acide acétique en fonction du temps.

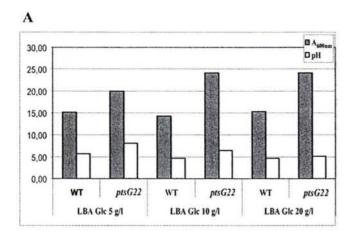
donc intéressant de tester l'effet d'une meilleure oxygénation du milieu de culture sur chacune des 2 souches TP 2503 et TP 2504 transformées par le plasmide pBS::*EcogapA*. Pour augmenter l'oxygénation du milieu, j'ai utilisé des Erlenmeyer à «aillettes».

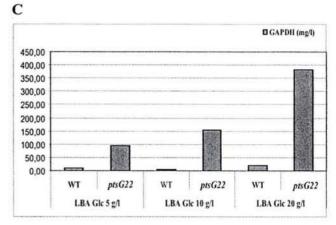
Une première étude réalisée sur les souches non recombinées m'a permis d'observer une augmentation de la biomasse en phase stationnaire de croissance liée à l'augmentation de l'oxygénation et ceci, quelle que soit la souche et les conditions de croissance (Tableau 12). Dans ces conditions d'oxygénation, une croissance optimale des souches était obtenue pour une concentration en glucose de 15 g/l.

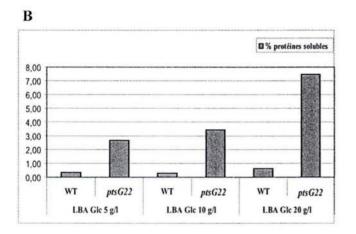
|                                   | LB Glc<br>5 g/l<br>TP 2503 | LB Glc<br>5 g/l<br>TP 2504 | LB Glc<br>10 g/l<br>TP 2503 | LB Glc<br>10 g/l<br>TP 2504 | LB Glc<br>15 g/l<br>TP 2503 | LB Glc<br>15 g/l<br>TP 2504 |
|-----------------------------------|----------------------------|----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| A <sub>600nm</sub>                | 8,9                        | 11                         | 8,15                        | 15,4                        | 9,35                        | 17,3                        |
| biomasse X (g/l)                  | 4,01                       | 4,18                       | 3,67                        | 5,85                        | 4,21                        | 6,57                        |
| μ <sub>1</sub> (h <sup>-1</sup> ) | 1,30                       | 0,72                       | 1,41                        | 0,52                        | 1,39                        | 0,51                        |
| μ <sub>2</sub> (h <sup>-1</sup> ) | 0,07                       |                            | 0,07                        |                             | 0,07                        |                             |
| Glc restant g/l                   | 0                          | 0                          | 3,65                        | 0                           | 8,66                        | 0                           |
| % Glc consommé                    | 100                        | 100                        | 69,56                       | 100                         | 51,86                       | 100                         |
| Acétate (g/l)                     | 2,00                       | 0,00                       | 4,30                        | 0,00                        | 3,60                        | 0,00                        |
| pH du milieu de<br>culture        | 4,95                       | 8,23                       | 4,35                        | 7,69                        | 4,46                        | 7,13                        |
| R <sub>x/s</sub> (g/g)            | 0,47                       | 0,71                       | 0,31                        | 0,49                        | 0,32                        | 0,37                        |
| R <sub>Ac/S</sub> (g/g)           | 0,33                       | 0,00                       | 0,52                        | 0,00                        | 0,39                        | 0,00                        |
| R <sub>Ac/X</sub> (g/g)           | 0,70                       | 0,00                       | 1,65                        | 0,00                        | 1,20                        | 0,00                        |

Tableau 12: Influence d'une augmentation de l'oxygénation sur la densité cellulaire, le taux de croissance, le pH du milieu de culture, la consommation du glucose, le rendement de biomasse par rapport au glucose, les rendements d'acide acétique produit par rapport au glucose et à la biomasse obtenus en phase stationnaire de croissance pour les souches TP 2503 et TP 2504 tranformées par le plasmide pBS::*EcogapA* lors de cultures réalisées en milieu LBA supplémenté avec différentes concentrations en glucose

J'ai donc étudié la croissance des souches recombinées et leur capacité de production de GAPDH recombinante pour 3 concentrations de glucose 5, 10 et 20 g/l de milieu de culture. Comme on peut le constater dans la Figure 85, l'oxygénation supérieure augmentait nettement les performances de production de la souche TP 2504, par rapport à la souche TP 2503. Un taux de GAPDH de 400 mg/l pouvait être atteint en présence de 20 g/l de glucose dans le milieu, ceci malgré le fait que la différence de densité cellulaire en phase exponentielle de croissance entre les deux souches n'était pas très marquée. Les résultats obtenus pourraient encore s'expliquer par une perte de plasmide dans la souche TP 2503, perte qui serait beaucoup plus limitée dans la souche TP 2504. En effet, une forte perte de plasmide dans la souche TP 2503, renforcée par l'oxygénation, est en accord







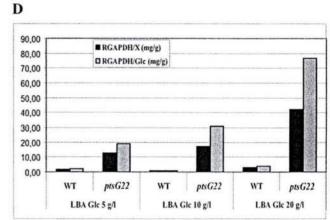


Figure 87: Influence de l'oxygène sur la densité cellulaire, le pH du milieu en phase stationnaire de croissance, et le taux de GAPDH rapporté au taux de protéines solubles ou au volume de milieu des culture, lors de culture des souches TP2503 et TP2504 transformées par le plasmide pBS::*EcogapA* en milieu LBA supplémenté avec du glucose à différentes concentrations

A : Comparaison des densités cellulaires et pH du milieu de culture obtenus en phase stationnaire de croissance pour les souches TP 2503 et TP 2504, transformées par le plasmide pBS::*EcogapA* en milieu LBA supplémenté avec du glucose à différentes concentrations.

B : Pourcentage de protéines solubles correspondant à la GAPDH en g de GAPDH par g de protéines solubles.

C : Quantité de GAPDH produite en mg par litre de milieu de culture.

D : Rendement de production de la GAPDH par rapport au glucose consommé et à la biomasse.

observées. Aucune analyse approfondie de cette observation n'a été faite par ces auteurs.

L'allongement de la taille des cellules, ainsi que la grande stabilité du nombre de copies de plasmides dans la souche TP 2504 cultivée en milieu LBAGlc (5 g/l) étaient inattendues. Il serait intéressant de tester si ces deux phénomènes sont bien liés, ceci en suivant leur évolution au cours d'une culture des souches TP 2503 et TP 2504 recombinées en milieu LBGlc (5 g/l). Ces effets pourraient résulter, soit du défaut du transporteur EIIBC<sup>Gle</sup>, soit de la libération de la protéine Mlc en absence de molécule EIIBC<sup>Gle</sup> active. On peut en effet imaginer que l'une ou l'autre des deux molécules puisse avoir des effets encore non décrits sur les mécanismes de division cellulaire et de partition des plasmides qui sont liés. Dans l'objectif d'obtenir des informations sur ces rôles possibles, il sera intéressant de comparer l'effet de la mutation Cys421Ser de la protéine EIIBC<sup>Gle</sup> à celle de la mutation ptsG22 sur la taille des cellules et la stabilité des plasmides. Il serait aussi intéressant d'inclure dans la comparaison une souche non mutée dans EIIBC<sup>Gle</sup> mais surproduisant la protéine Mlc. Toujours en vue d'avoir des informations dans ce domaine, il sera intéressant de comparer les transcriptomes et les protéomes de souches mlc<sup>+</sup>, mlc<sup>-</sup> et des souches TP 2503 et TP 2504 cultivées en présence ou en absence de glucose.

Les propriétés de la souche TP 2504 pourraient aussi être reliées au flux de glucose dans la voie de la glycolyse, et à la protéine Cra. En effet, comme la souche TP 2504 consomme plus lentement le glucose que la souche sauvage, il est possible que dans cette souche, la concentration des intermédiaires de la glycolyse soit plus faible. En particulier, la concentration de fructose-1,6-bisphosphate et fructose 1-phosphate n'est peut être pas suffisante pour lever l'activation qu'exerce la protéine Cra sur certains gènes du shunt du glyoxylate et la gluconéogenèse. Cette hypothèse pourra être testée en utilisant une souche n'exprimant ni un transporteur EIIBC<sup>Gle</sup> actif, ni la protéine Cra en milieu LB supplémenté avec du glucose.

Plusieurs techniques ont été mises en oeuvre pour inhiber la production d'acide acétique chez E. coli et aussi augmenter la production de protéines recombinantes. Il s'agissait en particulier de la délétion des gènes poxB, ackA, pta, acs codant des protéines qui interviennent dans le métabolisme de l'acide acétique (Fig. 4). La délétion de chacun des ces gènes entraînait une diminution de la densité cellulaire en phase stationnaire de croissance (facteur supérieur ou égal à 2) (Contiero et al., 2000). La souche mutée ptsG22 et les souches surproduisant la protéine Mlc (E. coli SR754 ou E. coli JM 103 surproduisant Mlc à partir d'un plasmide multicopie) ont des propriétés nettement plus intéressantes en terme de production de biomasse et protéines recombinantes que les souches mutées dans les gènes intervenant directement dans la production d'acide acétique, ce qui pourrait en faire des souches intéressantes sur le plan biotechnologique.

SCD UHP NANCY 1 Bibliothèque des Sciences Rue du Jarila Botanique - CS 20148 54601 VILLERS LES NANCY CEDEX

# Conclusion Finale

L'ensemble de mes travaux a porté sur les régulations transcriptionnelles et posttranscriptionnelles de l'opéron *epd-pgk-fbaA* en fonction des conditions de croissance et l'influence de l'inactivation du transporteur EIIBC<sup>Glc</sup> sur la densité cellulaire et la capacité de production de protéines recombinantes chez *E. coli*.

Dans un premier temps, nous avons montré que l'opéron epd-pgk-fbaA d'E. coli est exprimé uniquement à partir de deux promoteurs : epd P0 et pgk P1. Les transcrits issus de ces deux promoteurs sont ensuite maturés par la RNase E au niveau des régions epd-pgk et pgk-fbaA. L'étude de la stabilité des différents transcrits monocistroniques montre une stabilité croissante avec dans l'ordre epd, pgk et fbaA. Nous avons ensuite recherché si les éléments de séquences impliqués dans les régulations transcriptionnelles et post-transcriptionnelles de l'opéron epd-pgk-fbaA étaient conservés chez les γ-protéobactéries. Le triplet d'ORF E4PDH/GAPDH-PGK-FBA n'est en effet trouvé que chez les γ-protéobactéries. La région promotrice epd P0, avec le site de fixation de la protéine Cra et une séquence de 10 pb sont fortement conservés chez les Enterobacteriaceae, alors que le site de fixation du complexe CRP-AMPc est plus faiblement conservé. Le promoteur pgk P1 ainsi que les séquences des sites de clivages localisés dans l'ORF epd et leur structure bidimensionnelle sont aussi bien conservés chez les Enterobacteriaceae, ainsi que le terminateur potentiel rho-indépendant pgk T1 situé en aval de l'ORF pgk.

La protéine Hfq interagit avec les régions *epd-pgk* et *pgk-fbaA*, et nous avons pu mettre en évidence une fixation de la protéine Hfq au niveau du premier site de clivage à la RNase E, localisé dans l'ORF *epd*.

En utilisant la technique des fermentation qui nous a permis de contrôler le taux de croissance cellulaire, nous avons pu montrer que le taux des transcrits initiés aux promoteurs gapA P1, gapA P2, gapA P3 et epd P0 augmente lorsque le taux de croissance augmente.

L'étude des sous-produits du métabolisme a permis de mettre en évidence une faible production d'acide acétique par la souche TP 2504, dont le transporteur EIIBC<sup>Gle</sup> est inactivé. Nous avons alors comparé le comportement des souches portant un transporteur EIIBC<sup>Gle</sup> inactivé par rapport à un transporteur actif, dans différents milieux de croissance. La souche TP 2504, cultivée en milieu LB supplémenté en glucose, permet d'obtenir une plus forte biomasse et une plus faible acidification du pH du milieu de culture. L'augmentation de la biomasse pourrait être reliée à l'augmentation de la taille des bactéries. L'augmentation de la biomasse et la faible acidification du pH du milieu de cultures semblent s'expliquer par l'absence de fixation de la protéine Mlc au transporteur EIIBC<sup>Gle</sup>.

Cette souche TP 2504 possède aussi l'avantage d'avoir une grande stabilité du nombre de copies de plasmides, en milieu LBAGlc (5 g/l). Ceci permet d'obtenir une forte production de protéines recombinantes, d'où un intérêt biotechnologique possible de cette souche.

# Matériels et Méthodes

## Partie A / Matériels

## I / Souches bactériennes utilisées

L'ensemble des souches bactériennes utilisées au cours de ce travail est répertorié dans le Tableau 13.

## II / Milieux de culture utilisés

La composition des différents milieux de cultures utilisés est décrite dans la Figure 88.

## III / Plasmides utilisés

- pBS SK<sup>+</sup> (Clontech), ce plasmide a été utilisé pour les clonages des gènes étudiés.
- pBS::EcogapA contient le gène gapA d'E. coli sous contrôle de sa région promotrice (Charpentier, 1994) (Fig. 89)
- pT-Adv. Ce plasmide a été utilisé pour le clonage de fragments ADN synthétisés par la technique PCR (Clontech).

## IV / Tampons et solutions à usage courant

TE: EDTA 1 mM, Tris-HCl 10 mM pH 8,0.

TBE 1X: acide borique 89 mM, EDTA 2 mM, Tris-Borate 89 mM pH 8,3.

Bleu de dépôt pour gels d'agarose : EDTA 125 mM, SDS 2 %, glycérol 25 %, bleu de bromophénol 0,12 %.

Solution de dépôt pour gel de polyacrylamide dénaturant (bleu de formamide) : EDTA 20 mM, bleu de bromophénol 0,05 %, bleu de xylène cyanol 0,05 % dans de la formamide désionisée.

Bleu de protéine: Tris-HCl 80 mM pH 6,8, SDS 3 %, Bleu de bromophénol 10 mg/ml, DTT 100 mM, glycérol 10 %.

Bleu de coloration de gels de protéines : Bleu de Coomassie G250 0,1 %, acide acétique 10 %, éthanol 40 %.

## Souches bactériennes utilisées

| Souches utilisées   | Génotype  | Utilisation - Caractéristiques   | Source                    | Don de                     |
|---------------------|---|--|---------------------------|----------------------------|
| Souche de référence | pour l'analyse de la régulation de l'expi   | ression des gènes <i>epd</i> , <i>pgk</i> et <i>fbaA</i>                     |                           |                            |
| TGI                 | sup E, hsd D5, thi-, D(lac-pro AB),<br>F'(tra D 36 pro AB +lac I q lacZ DM15)                         |  | Stratagène                |                            |
| Etudes des mécanism | es de maturation  |  | 19                        |                            |
| MG1693              | thyA715   | souche isogénique d'E. coli SK5665, SK5704, IBPC 684                         | Arraiano et al., 1988     | P. Régnier (IBPC, Paris)   |
| SK5665              | thyA715 ams-1   | RNase E thermosensible   | Arraiano et al., 1988     | P. Régnier (IBPC, Paris)   |
| SK5704              | thyA715, ams-1, pnp-7, rnb-500  | RNase E, RNase II et PNPase thermosensibles                                  |                           | P. Régnier (IBPC, Paris)   |
| IBPC684             | thyA715, ams-1, pnp-7, rnb-500,<br>zad::Tn10, pcnB80  | RNase E, RNase II et PNPase thermosensibles ;<br>poly(A) polymérase inactive |                           | P. Régnier (IBPC, Paris)   |
| IBPC600             | thi-1, argE3, \( \Delta lacX74, mtl-1, xyl-5, tsx-29, rpsL, argG6, his-4, nadA51::Tn10, rnc105        | Rnase III inactivée  |                           | M. Springer (IBPC, Paris)  |
| IBPC602             | thi-1, argE3, \(\Delta\lambda\)lacX74, mtl-1, xyl-5, tsx-29,rpsL, argG6, his-4, nadA51::Tn10          | Souche isogénique d'E. coli IBPC 600   |                           | M. Springer (IBPC, Paris)  |
| N3433               | HfrH, lacZ43, λ-, relA 1, spoT 1, thi-1   | WT   | D. apirion                | E. Hajnsdorf               |
| N3431               | N3433, rne 3071(T <sup>s</sup> )  | RNase E thermosensible   | D. apirion                | E. Hajnsdorf               |
| IBPC633             | N3433-rnc 105-nadB ::tetR   | RNase III inactivée  |                           | E. Hajnsdorf               |
| IBPC933             | N3433-rne 3071ts-rng :: CmR   | RNase E thermosensible, RNase G inactivée                                    | communication personnelle | E. Hajnsdorf (IBPC, Paris) |
| IBPC935             | N3433-rng::CmR  | RNase G inactivée  | communication personnelle | E. Hajnsdorf (IBPC, Paris) |
| IBPC5321            | thi-1, argE3, ΔlacX74, mlt-1i, xyl-5, tsx-<br>29, rpsL, argG6, his-4                                  | obtention de phages recombinants   | Plumbridge et Soll, 1989  | M. Springer (IBPC, Paris)  |
| IBPC6311            | thi-1, argE3, \(\Delta\lambda\)lacX74, mlt-1, xyl-5, tsx-29, rpsL, argG6, his-4, recA56, srl-300:Tn10 | obtention de souches monolysogènes   | Plumbridge et Soll, 1989  | M. Springer (IBPC, Paris)  |
| IBPC605             | IBPC5321-rnc+, curee du Tn10-nad B51  | souche WT isogénique d'E. coli IBPC606                                       |                           | E. Hajnsdorf               |
| IBPC606             | IBPC5321-rnc 105curee du Tn10-nad B51   |  |                           | E. Hajnsdorf               |

| Etudes du rôle de | e la protéine Hfq   |   |                            |                                      |
|-------------------|---|---|----------------------------|--------------------------------------|
| JC7623            | recB21 recC22 sbc-15 ara arg his leu pro<br>thr                         | souche WT isogénique d'E. coli TX2808   | Tsui et al., 1994          | E. Hajnsdorf (IBPC, Paris)           |
| TX2808            | JC7623 <i>hfq1</i> ::Ω(Kmr; <i>Bcl</i> I)                               | cassette de clonage insérée en amont du gène hfq :<br>étude de l'effet de la mutation hfq | Tsui et al., 1994          | E. Hajnsdorf (IBPC, Paris)           |
| Souches pour l'ar | nalyse de la mutation ptsG22  |   | T                          |                                      |
| TP 2503           | F xyl argH1 ilvA  | souche isogénique d'E. coli TP 2504   | De Reuse et Danchin, 1988  | H. De Reuse, Institut Pasteur, Paris |
| TP 2504           | TP 2503 zcf-229 ::Tn10 ptsG22   | inactivation du transporteur EIIBC <sup>Glc</sup>   | De Reuse et Danchin, 1988  | H. De Reuse, Institut Pasteur, Paris |
| TP 2006           | F' xyl lac ΔX74 Δcya glp -8306  | délétion amp cyclase  | Roy et Danchin, 1982       | A. Kolb, Paris                       |
| TP 2862           | F', xyl argH1 ilvA aroB Δcrr:: Km                                       | délétion de la production de EIIA Gle   | Levy et al., 1990          | H. De Reuse, Institut Pasteur, Paris |
| TP 2819           | F', xyl argH1 ilvA aroB ΔptsIcrr:: Km                                   | délétion de la production de El et EIIA Gle   | Levy et al., 1990          | H. De Reuse, Institut Pasteur, Paris |
| TP 2811           | F', xyl argH1 ilvA aroB ΔptsHlcrr:: Km                                  | délétion de la production de HPr, EI et EIIA Gle  | Levy et al., 1990          | H. De Reuse, Institut Pasteur, Paris |
| TP2512            | F xyl argH1 ilvA zcf-229 ::Tn10 ptsG22<br>Δcya glp8306                  | délétion cya et mutation ptsG22   | De Reuse et Danchin, 1988  | A. Kolb, Paris                       |
| JM101             | $\Delta(lac-proAB)$ , thi-1, supE, F'(traD36, lacIq, lacZ $\Delta$ M15) | souche WT isogénique d'E. coli IBPC1012   | Yanish-Perron et al., 1985 | J. Plumbridge, Paris                 |
| IBPC1012          | JM101 mlc::Tc   | inactivation de Mlc   | Plumbridge, 1998a          | J. Plumbridge, Paris                 |
| IBPC900           | IBPC5321 mle:: te   | inactivation de Mlc   | Plumbridge, 1998b          | J. Plumbridge, Paris                 |
| IBPC1000          | IBPC5321 ptsG22 zcf229-Tn10<br>mlc::10kan                               | délétion de la production de la protéine Mlc et inactivation du transporteur EIIBC Glc    | Plumbridge et Vimr, 1999   | J. Plumbridge, Paris                 |

Tableau 13 : Souches d'E. coli utilisées au cours de ce travail.

### Milieu riche LB (Luria Bertani broth):

|                        | LB du laboratoire | LB classique | LB + phosphate |
|------------------------|-------------------|--------------|----------------|
| Tryptone               | 20 g/l            | 10 g/l       | 10 g/l         |
| Extrait de levure      | 5 g/l             | 5 g/l        | 5 g/l          |
| NaCl                   | 10 g/l            | 10 g/l       | 10 g/l         |
| KH₂PO₄                 |                   |              | 13,6 g/l       |
| ajustement du pH à 7,5 | NaOH              | NaOH         | кон            |

Le milieu de culture est autoclavé 20 min à 121°C (1 bar).

Le milieu de culture "LB du laboratoire" a été utilisé pour toutes les expériences de biologie moléculaire. Le milieu de culture "LB classique" et "LB + phosphate" ont été employé lors de la comparaison des souches TP 2503 et TP 2504 (Partie C des résultats) afin de pouvoir comparer les résultats avec ceux précédemments publiés.

### Milieu minimum M63:

|   | fermentations continues réalisée<br>en limitation de glucose | Fermentations continues réalisées en<br>limitation de phosphate minéral |
|---|--|---|
| KH₂PO4  | 0,1 M  | 0,02 M  |
| (NH <sub>4</sub> ) <sub>2</sub> 5O <sub>4</sub> | 15 mM  | 15 mM   |
| MgSO <sub>4</sub> , 7H <sub>2</sub> O           | 0,8 mM   | 0,8 mM  |
| FeSO <sub>4</sub> , 7H <sub>2</sub> O           | 2 μΜ   | 2 μΜ  |

Le pH est ajusté à 7,0 par KOH 10 N et le milieu de culture est autoclavé 18 min à 121°C (1 bar).

#### Additifs ajoutés au milieu M63:

| Casaminoacides           | 2,5 % (p/v) |
|--------------------------|-------------|
| Thiamine                 | 0,2 % (p/v) |
| Source de carbone (20 %) | 0,4 % (p/v) |

La source de carbone utilisée est soit le glucose, soit le pyruvate. Les casaminoacides sont autoclavées 20 min à  $121^{\circ}C$  (1 bar), alors que la thiamine et les sources de carbone sont stérilisées par filtration sur filtre Millex, 0.22  $\mu$ m Millipore.

#### Milieux solides :

17 g d'agar sont ajoutés par litre de milieu de culture avant autoclavage.

## Utilisation d'antibiotiques :

Pour assurer une pression de sélection en vue de conserver les plasmides, de l'ampicilline à 100 mg/l ou de la kanamycine à 100 mg/l ont été ajoutés.

Figure 88 : Composition des différents milieux de cultures utilisés.

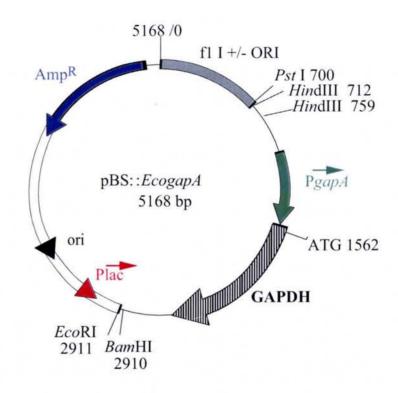


Figure 89: Carte de restriction du plasmide pBS::EcogapA (Charpentier et Branlant, 1994; Charpentier, 1994).

La flèche grisée représente la région codante de la GAPDH et la flèche verte représente la région promoteur du gène gapA (PgapA).

Amp<sup>R</sup> : gène de résistance à l'ampicilline.

ori : origine de la réplication du plasmide ColE1.

Plac: promoteur de l'opéron lactose.

fl I +/- ORI : origine de réplication du phage filamenteux fl.

## Partie B / Méthodes

## I / Cultures des Bactéries

## I.1 / Cultures d'E. coli en mode batch

Une préculture en phase stationnaire de croissance est utilisée pour ensemencer 70 mL de milieu de culture de façon à obtenir une absorbance à 600 nm (A<sub>600nm</sub>) de 0,05. La culture réalisée à 37°C est arrêtée lorsque les bactéries sont en milieu de phase exponentielle de croissance (A<sub>600nm</sub> de 0,5 (comprise entre 0,48 et 0,55). Lorsque l'objectif est à la fois de mesurer les activités enzymatiques intracellulaires et de doser le taux des ARNm, deux échantillons de 30 mL correspondant à environ à 15x10<sup>9</sup> cellules sont prélévés, centrifugés. Les cellules sont ensuite lavées avec du tampon A (EDTA 5 mM, saccharose 25 %, Tris-HCl 50 mM, pH 8,0), puis conservées à -20°C.

Dans le cas où la culture est destinée à la réalisation d'une étude cinétique de l'évolution des transcrits au cours d'une culture, des échantillons correspondant à  $15 \text{ A}_{600}$  soit  $15 \times 10^9$  cellules sont prélevés régulièrement au cours de la culture.

Pour la mesure des capacités de production de la GAPDH des souches E. coli DH5 $\alpha$ , TP 2503 et TP2504, un échantillon correspondant à 15 Unités  $A_{600nm}$ , correspondant à  $15x10^9$  cellules a été prélevé en phase stationnaire de croissance, centrifugé, lavé au tampon A puis conservé à -20°C.

## I.2 / Cultures des bactéries par l'approche des fermentations continues : conduite des fermentations continues

### 1.2.a / Description du fermenteur

Nous avons utilisé le fermenteur BiofloIII de New Brunswick Scientific au cours de ce travail. Il est composé d'une sonde de pH (Ingold), d'une sonde à oxygène dissous (Ingold), d'une sonde de température, chacune de ses sondes a ses systèmes de régulation. Le système comporte aussi une sonde pour réaliser les prélèvements, différentes entrées pour alimenter le système en milieu frais, et d'un système d'aération du milieu (Fig. 90). L'étalonnement de la sonde de pH est réalisé avant l'autoclavage du fermenteur, alors que celui de la sonde à oxygène est réalisé après autoclavage.

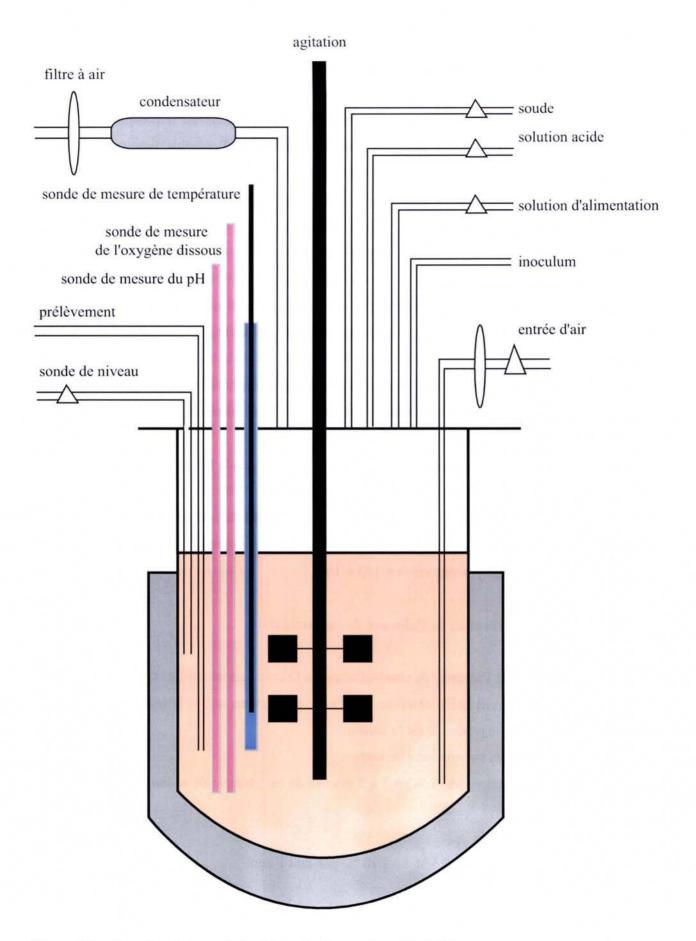


Figure 90 : Représentation schématique du fermenteur BiofloIII.

#### I.2.b / Protocole des fermentations continues

Nous avons travaillé avec un volume utile de 2 l environ dans un fermenteur de 3 l. Les fermentations ont été réalisées à 37°C et à pH 7,0 et à un pourcentage d'oxygène dissous de 36 à 40 %, l'alimentation en air variant de 1 à 4 vvm, et l'agitation de 100 à 800 tours/min. L'ensemencement du fermenteur est réalisé en plusieurs étapes. Une préculture en phase stationnaire de croissance est utilisée pour ensemencer, à une absorbance A<sub>600nm</sub> de 0,1 Unités, un milieu de culture en Erlenmeyer correspondant à un volume de 1/10 du volume utile du fermenteur, soit 200 ml. Lorsque les cellules sont en fin de phase exponentielle de croissance (A<sub>600nm</sub> = 0,8), la culture est utilisée pour ensemencer le fermenteur à une densité cellulaire finale d'environ A<sub>600nm</sub> = 0,08 unités. La phase de croissance batch du fermenteur débute alors. La croissance bactérienne est suivie par mesure de l'absorbance à 600 nm, ce qui permet de mesurer le taux de croissance maximal de la bactérie étudiée dans le milieu de croissance utilisé. Lorsque les bactéries atteignent la fin de phase exponentielle de croissance (A<sub>600nm</sub> de 0,8), les pompes d'alimentation et de soutirage sont mises en route, ce qui permet de maintenir le volume utile du réacteur à 2 l.

Le pH est maintenu à 7,0 en contrôlant l'apport d'une solution de soude 1 N ou de H<sub>2</sub>SO<sub>4</sub> 1 N.

Le chémostat est considéré équilibré lorsqu'un volume égal à 3 à 5 fois le volume utile du fermenteur a traversé le système, soit après 3 à 5 temps de passage. Le temps de passage est défini par le temps nécessaire pour que le volume traversant le chémostat soit égal au volume utile. Lorsque le chémostat est à l'équilibre, le temps de passage est :  $t = 1/D = 1/\mu$ .

## I.2.c / Vérification de l'absence de contamination

Nous avons vérifié l'absence de contamination de la culture à intervalles de temps réguliers au cours de la culture en fermentation continue. Pour cela, nous avons utilisé la technique classique de coloration de Gram, dont le protocole est le suivant :

- fixation des cellules sur une lame de verre
- coloration au cristal violet (1 %) + 3 à 5 gouttes de bicarbonate de sodium (0.5 %) pendant 2
   min
- lavage à l'eau distillée
- ajout de lugol (solution I<sub>2</sub> 0,5 % + KI 10 %) pendant 1 min
- lavage à l'eau distillée
- lavage à l'éther-acétone (1/3-2/3) pendant 10 sec
- lavage à l'eau distillée
- séchage
- observation de la préparation au microscope optique (objectif x100, sous immersion).

Les bactéries à gram négatif sont roses, alors que les bactéries à gram positif sont violettes.

## II / Extraction et analyse de l'ADN

#### II.1 / Extraction d'ADN cellulaire

## II.1.a / Extraction d'ADN chromosomique

Les cellules d'*E. coli* TG1 sont cultivées jusqu'en phase exponentielle de croissance (A<sub>600nm</sub> = 0,5), puis sont broyées par des billes de verre dans le tampon de broyage : Triton X-100 2 % ; SDS 1 % ; NaCl 100 mM ; EDTA 1 mM ; Tris-HCl 10 mM pH 8,0 en présence de phénol-chloroforme. La phase aqueuse, récupérée après centrifugation, est de nouveau extraite au phénol-chloroforme, lavée au chloroforme/alcool isoamylique [24/1 (v/v)] et l'ADN est précipité à l'éthanol.

L'ADN solubilisé dans du tampon TE est traité par la RNase A (20 μg/ml) puis par la protéinase K (250 μg/ml) et enfin précipité à l'éthanol. L'ADN est repris dans du TE, puis quantifié par mesure de l'absorbance à 260 nm, une unité de A<sub>260nm</sub> correspondant à 50 μg/ml.

## II.1.b / Extraction d'ADN plasmidique

La préparation d'ADN plasmidique à petite échelle a été réalisée par la technique de lyse alcaline (Sambrook et al., 1989) sur des cultures bactériennes de 2 mL arrêtées en phase stationnaire de croissance. Pour la production d'ADN plasmidique en grande quantité, le kit "Quick Flow Maxi Kit Protocol" (Hybaid recovery) a été utilisé.

La qualité des préparations est vérifiée par fractionnement des produits dans un gel d'agarose à 0.8 % (p/v). Le gel est préparé dans du tampon TBE  $0.5 \times 10^{10} \, \mathrm{M}$  X additionné de bromure d'éthidium à une concentration finale de  $0.5 \, \mathrm{\mu g/ml}$ . Le tampon TBE  $0.5 \times 10^{10} \, \mathrm{M}$  x est également utilisé comme tampon d'électrophorèse. On ajoute aux échantillons  $1/5^{\mathrm{ème}}$  à  $1/10^{\mathrm{ème}}$  de volume de bleu de dépôt pour gels d'agarose avant de les déposer sur gel en parallèle avec un marqueur de taille. La migration est effectuée à voltage constant (5 V/cm).

## II.2 / Marquage des oligonucléotides

Les oligonucléotides sont achetés purifiés à la société Eurogentec. Ils sont marqués à l'extrémité 5' par incorporation de phosphore 32. La réaction est catalysée par la T4 polynucléotide kinase en présence de  $ATP[\gamma^{-32}P]$ .

### Protocole:

| ATP[γ- <sup>32</sup> P] (10 mCi/ml, 3000 Ci/mmol) (Amersham)                    | 50 μCi    |
|---|-----------|
| Oligonucléotide   | 100 ng    |
| Tampon Enzyme 10X (MgCl <sub>2</sub> 0,1 M; DTT 0,05 M; Tris-HCl 0,7 M; pH 7,6) | 1 μl      |
| T4 polynucléotide kinase (10 U/μl) (Biolabs)                                    | 1 μl      |
| $H_2O$  | qsp 10 μl |

Après une incubation de 1 h à 37°C, le mélange est placé 10 min à 65 °C pour dénaturer l'enzyme.

## II.3 / Amplification d'ADN par la technique Polymerase Chain Reaction (PCR)

Les cycles de synthèse d'ADN à partir d'oligonucléotides amorces et d'ADN polymérase de Thermus brockianus (DyNAzyme, Finnzymes) sont réalisés dans les conditions décrites dans la Figure 91.

## II.4 / Clonage de fragments d'ADN

## II.4.a / Ligation d'un fragment d'ADN issu d'une hydrolyse par des enzymes de restriction

Le vecteur de clonage est hydrolysé par les enzymes de restriction adéquates, puis traité à la phosphatase alcaline (Boehringer) (1 U/μg d'ADN) dans le tampon 10 x suivant : Tris-HCl 500 mM pH 8,5, EDTA 1 mM pendant 1 heure à 37°C. Ce traitement permet d'éviter la religation du vecteur sur lui-même lors de la réaction de ligation. L'enzyme est enfin dénaturée par une incubation de 10 min à 65 °C.

La ligation est catalysée par l'ADN ligase du bactériophage T4. Nous avons travaillé en présence d'un excès de fragment avec un rapport molaire, fragment / vecteur, variant de 4 à 8.

### Protocole:

| Vecteur hydrolysé et déphosphorylé                             | 50 ng     |
|--|-----------|
| Fragment d'ADN à insérer, excès molaire par rapport au vecteur | 4 à 8 x   |
| Tampon de ligation 10X   | 2 μ1      |
| ( MgCl <sub>2</sub> 50 mM; DTT 10 mM; Tris-HCl 660 mM, pH 7,6) |           |
| ATP (10 mM)  | 2 μ1      |
| ADN ligase (1 U/µl) (Roche)                                    | 1 μl      |
| H <sub>2</sub> O   | qsp 20 μl |

## A / Amplification d'ADN par la méthode PCR

## Mélange réactionnel:

| Matrice ADN   | 10   | ng |
|---|------|----|
| Oligonucléotide (100 ng)  | 2    | μl |
| dNTP (20 mM chaque)   | 0,5  | μl |
| Tampon 10x (MgCl <sub>2</sub> 15 mM; KCl 500 mM; Tris-HCl 100 mM, pH 8,3) | 5    | μl |
| DyNAzyme (Finnzymes) (1 U/μl)   | 0,75 | μl |
| H <sub>2</sub> O qsp  | 50   | μl |

L'amplification est réalisée dans un appareil Perkin-Elmer. Les cycles suivant sont réalisés :

| Cycle | Fonction     | Température | Temps       | Nombre de cycles |
|-------|--------------|-------------|-------------|------------------|
| 1     | dénaturation | 94°C        | 2 minutes   | 1                |
| 2-32  | dénaturation | 94°C        | 30 secondes | 30               |
|       | Hybridation  | 48°C        | 30 secondes |                  |
|       | élongation   | 72°C        | 45 secondes |                  |
| 33    | élongation   | 72°C        | 10 minutes  | 1                |
| 34    | stockage     | 4°C         |             |                  |

## B / Amplification d'ADN à partir d'une colonie par la méthode PCR

## Mélange réactionnel :

| Oligonucléotide (20 µM)       | 1,25 μl   |
|-------------------------------|-----------|
| Tampon 10x                    | 2 μ1      |
| dNTP                          | 0,25 μl   |
| DyNAzyme (Finnzymes) (1 U/μl) | 0,25 μl   |
| $H_2O$                        | qsp 25 μl |

La colonie est ensuite ajoutée dans ce mélange réactionnel. Les cycles suivants sont ensuite réalisés :

| Cycle | Fonction     | Température | Temps       | Nombre de cycles |
|-------|--------------|-------------|-------------|------------------|
| 1     | dénaturation | 94°C        | 40 secondes | 1                |
| 2-32  | dénaturation | 94°C        | 30 secondes | 30               |
|       | Hybridation  | 42°C        | 30 secondes |                  |
|       | élongation   | 72°C        | 30 secondes |                  |
| 33    | élongation   | 72°C        | 10 minutes  | 1                |
| 34    | stockage     | 4°C         |             |                  |

Figure 91: Utilisation de la technique PCR pour l'amplification d'ADN

Le mélange, incubé la nuit à 16°C ou 4 h à température ambiante, est utilisé directement pour transformer les cellules d'*E. coli* TG1 préalablement rendues compétentes par un traitement au CaCl<sub>2</sub> (50 mM) (Cohen *et al.*, 1972). Après un choc thermique de 30 sec à 42°C, les cellules renfermant un plasmide sont sélectionnées sur un milieu LB + ampicilline.

## II.4.b / Ligation d'un fragment d'ADN synthétisé par la technique PCR

Les fragments d'ADN db synthétisés par l'ADN polymérase de *T. brockianus* possèdent une adénosine supplémentaire aux extrémités 3'. Le clonage de ces fragments s'effectue dans un plasmide linéarisé appelé plasmide "T", possédant une thymine excédentaire sur chaque extrémité 5'. Le kit de clonage pT-Adv basé sur cette technique a été utilisé.

### II.4.c / Vérification de l'insertion

Pour vérifier l'insertion du fragment, deux techniques peuvent être utilisées : l'analyse par restriction et séquençage du vecteur après extraction des plasmides par la technique de lyse alcaline, ou la technique PCR appliquée directement sur colonie. Cette dernière méthode permet de tester rapidement si le fragment d'ADN a été inséré dans le vecteur. Un transformant est mis en suspension dans le mélange réactionnel, dont la composition est donnée dans la Figure 91. Différents couples d'oligonucléotides peuvent être utilisés pour vérifier si le fragment est inséré. Après fractionnement en gel d'agarose des produits d'amplification, la taille des produits d'amplification permet de sélectionner le vecteur recombinant recherché.

## II.5 / Séquençage d'ADN

Le séquençage de l'ADN double brin, dissocié en simple brin par un traitement alcalin, est réalisé selon la méthode de Tabor et Richardson (1987). Le principe repose sur l'utilisation d'un oligonucléotide de synthèse qui, hybridé en 3' du fragment à séquencer, sert d'amorce à la synthèse d'un ADN complémentaire par l'ADN polymérase du phage T7. La synthèse obtenue par allongement de l'amorce est effectuée en présence de désoxynucléotides (dNTP) et de didésoxynucléotides (ddNTP). L'incorporation de ddNTP arrête la synthèse, l'arrêt est statistique en fonction de la concentration relative des dNTP/ddNTP (Tabor et Richardson, 1987). Les différentes étapes sont indiquées dans la Figure 92.

## A / Dénaturation de l'ADN à séquencer

2-3 μg d'ADN plasmidique double brin 10 μl de NaOH 0,4 N 5 min à température ambiante

8 μl d'AcNH<sub>4</sub> 5 M pH 7,5 110 μl d'EtOH 5 min à -80°C

Centrifugation 10 min à 10 000 rpm Lavage du culot avec 200 µl d'EtOH 70 %, séchage

## B / Hybridation de l'oligodéoxynucléotide à l'ADN simple brin matrice

ADN simple brin 2-3 µg oligodésoxynucléotide 10 ng tampon 5X (MgCl $_2$  50 mM, NaCl 250 mM, Tris-HCl 200 mM, pH 7,5) 2 µl H $_2$ O qsp 10 µl 5 min à 80°C, puis, refroidissement par retour à température ambiante

transfert dans un tube contenant :

1 μl de [α-<sup>33</sup>P]dATP (3000 Ci/mmol, 2 mCi/ml) (Amersham) 1 μl de DTT 0,1 M 2 μl du mélange de dNTP (7,5 μM chaque) 1 μl d'ADN polymérase du phage T7 (2 U/μl) (Pharmacia) 3 min à température ambiante.

## C / Réaction d'élongation

 $3.5~\mu l$  du mélange précédant sont répartis dans 4 tubes (A, C, G, T) contenant respectivement  $2.5~\mu l$  des mélanges de terminaison dATP/ddATP, dCTP/ddCTP, dGTP/ddGTP, dTTP/ddTTP (rapport dNTP/ddNTP = 200~mM/20~mM)

5 min à 37°C

Arrêt de la réaction d'élongation par l'addition de 4  $\mu$ l de bleu de formamide Dénaturation 2 min à 100°C

## D / Fractionnement en gel de polyacrylamide

Dépôt sur un gel de polyacrylamide dénaturant (polyacrylamide 7 %, urée 8 M, TBE 1 X) Migration à 40 V/cm à 50°C.

Fixation du gel dans l'acide acétique 10 % (v/v) pendant 10 min.

Transfert du gel sur papier Whatmann 3MM, recouvert d'un film plastique et séchage sous vide.

Autoradiographie.

## Figure 92 : Protocole utilisé pour le séquençage d'ADN

## III / Extraction et analyse de l'ARN

## III.1 / Extraction des ARN totaux

L'extraction des ARN totaux est réalisée par la méthode basée sur l'emploi du phénol à chaud (65°C) (Aiba et al., 1981). Les ARN extraits sont précipités à l'éthanol et à l'acétate de sodium 0,3 M, puis repris dans de l'eau milliQ stérile. Une quantification est réalisée par mesure de l'absorbance à 260 nm avec un appareil Genequant (Pharmacia).

## III.2 / Technique d'extension d'amorces marquées au <sup>32</sup>P par la transcriptase inverse

Un oligonucléotide marqué en 5' au <sup>32</sup>P est hybridé à l'ARN. L'extrémité 3' de cet oligonucléotide sert d'amorce à la transcriptase inverse qui synthétise, en présence des 4 dNTP, un brin d'ADN complémentaire à l'ARN.

## III.2.a / Hybridation ARN/oligonucléotide

Pour des ARN totaux extraits de cellules non transformées, 20 μg sont incubés en présence de 5 ng d'oligonucléotide marqué (le protocole de marquage est détaillé ci-dessous), alors que pour des ARN totaux extraits de cellules transformées par un plasmide multicopie, 5 μg sont placés en présence de 5 ng d'oligonucléotide marqué, en général. Les quantités d'ARN peuvent être réajustées en fonction de la quantité de plasmides présents dans les cellules, en particulier, lors de l'étude de l'expression des ORF *epd*, *pgk* et *fbaA*. Les mélanges d'hybridation ARN/oligonucléotides sont réalisés dans le tampon de la transcriptase inverse d'AMV (fourni avec l'enzyme) et dans un volume final de 10 μl. Après une incubation de 10 min à 65°C, ils sont placés à température ambiante pendant 10 min.

## III.2.b / Synthèse du brin complémentaire

## Protocole:

| dNTP (dATP, dCTP, dGTP, dUTP 1,25 mM chacun) | 1,5 μ1        |
|--|---------------|
| Tampon transcriptase inverse AMV 10x         | 0,5 μ1        |
| Transcriptase inverse AMV (Life Science)     | 0,1 µl (0,1U) |
| Eau milliQ stérile                           | qsp 5 μl      |

Ce mélange réactionnel est ajouté à chaque mélange d'hybridation ARN-oligonucléotide radiomarqué. L'élongation des amorces est effectuée à 42°C pendant 30 min, puis est arrêtée par l'addition de 100 µl d'acétate de sodium 0,3 M et 500 µl d'éthanol absolu. Après 30 min à -80°C, le mélange est centrifugé, lavé avec 100 µl d'éthanol à 70 %, puis séché.

#### III.2.c / Fractionnement des ADNc obtenus

Les précipités sont repris dans 8 μL de bleu de formamide. Après chauffage 2 min à 95°C, les ADNc obtenus sont déposés sur un gel de polyacrylamide dénaturant (polyacrylamide 5 %, urée 8 M, TBE 1X) ayant subi une pré-électrophorèse d'une heure à 45 V/cm. La migration est réalisée à 45 V/cm, puis le gel est transféré sur papier Whatman 3MM, recouvert d'un film plastique et séché. Les bandes de radioactivité peuvent être quantifiées par utilisation du logiciel Image Quant (version 5.2) de l'appareil PhosphorImager (Molecular dynamics).

## III.3 / Analyse des ARN par Northern-blot

## III.3.a / Séparation des ARN suivant leur taille

Entre 5 et 10 μg d'ARN total extrait de cellules transformées par un plasmide multicopie ou 30 μg d'ARN total extrait de cellules non transformées sont fractionnés en gel d'agarose 0,8 % - formaldéhyde, dans les conditions précédemment décrites (Sambrook *et al.*, 1989).

#### III.3.b / Transfert électrophorétique sur une membrane Biodyne

Après la migration, le gel est incubé dans une solution basique NaOH 0,05 N pendant 20 min pour hydrolyser les ARN et faciliter leur transfert sur la membrane. Le gel est ensuite rincé avec de l'eau milliQ stérile, puis neutralisé dans une solution SSC 20x (NaCl 3M, citrate de sodium 0,3 M) pendant 45 min. Les ARN sont ensuite transférés sous vide en utilisant l'appareil VacuGene XL (Biorad) sur une membrane de nylon (Biodyne B transfer membrane, 0,45 μm, Pall) en tampon SSC 20x durant 3 heures. Les ARN sont ensuite fixés covalemment sur la membrane par pontage photochimique aux UV, durant 5 min.

## III.3.c / Hybridation

La pré-hybridation et l'hybridation avec l'oligonucléotide marqué en 5' sont réalisées à 42°C dans le tampon d'hybridation [SSC 6x, SDS 0,1 %, Réactif de Denhardts 5x (Ficoll 0,1 %, polyvinylpyrrolidone 0,1 %, BSA 0,1 %), EDTA 5mM, ADN de thymus de veau soniqué et dénaturé 10 μg/ml]. Après 4 heures de préhybridation à 42 °C, l'oligonucléotide marqué en 5' est ajouté dans le liquide de préhybridation à une concentration finale de 10 ng/ml. L'hybridation est réalisée durant 16 heures à 42°C sous agitation. La membrane est lavée 2 fois à température ambiante dans le tampon de lavage (SSC 6x; SDS 0,1 %) puis deux fois dans le même tampon durant 20 min à 42°C. La membrane est ensuite scellée dans un film plastique puis placée en fluorographie à -80°C.

## III.3.d / Déshybridation et réhybridation

La sonde est décrochée de la membrane de nylon en plongeant la membrane de nylon 2 fois dans un bain de SDS à 0,1 % porté à ébullition. La déshybridation est vérifiée par exposition en fluorographie à -80°C pendant une nuit. La membrane peut à nouveau être utilisée pour une hybridation avec une autre sonde marquée.

## III.4 / Analyse de la structure secondaire d'ARN par sondes enzymatiques et chimiques

## III.4.a / Production d'ARN froids par transcription *in vitro* avec l'ARN polymérase du phage T7

La matrice ADN pour la transcription est obtenue par amplification avec un couple d'oligonucléotides amorces générant une séquence promoteur reconnue par l'ARN polymérase du phage T7. L'ADN amplifié est extrait par un mélange phénol-chloroforme, précipité à l'éthanol en présence d'acétate d'ammonium (2 M) puis lavé à l'éthanol 70 %. L'ADN est ensuite repris dans de l'eau stérile MilliQ et sa concentration est ajustée à 0,5 μg/μl. La transcription est réalisée dansle mélange réactionnel suivant :

### Protocole:

| ADN $(0.5 \mu g/\mu l)$                                     | 1      | μl      |
|---|--------|---------|
| Tampon de transcription 3x                                  | 10     | μl      |
| (MgCl <sub>2</sub> 60 mM; Spermidine 3 mM; DTT 15 mM;       |        |         |
| Triton 0,03 %; PEG 8000 240 mg/ml; Tris-HCl 120 mM pH 8,0). |        |         |
| rNTP (ATP, CTP, GTP, UPT 20 mM)                             | 6      | μl      |
| RNasin (36,1 U/μl) (Amersham Biosciences)                   | 0,5    | μl      |
| ARN polymérase T7 (purifiée au Laboratoire)                 | 1      | $\mu l$ |
| BSA (2 mg/ml)   | 3      | μl      |
| $H_2O$  | qsp 30 | μl      |

Le mélange réactionnel est incubé à 37°C pendant 2 h, puis la matrice ADN est dégradée par incubation pendant 30 min à 37°C en présence de 10 U de DNase I exempte de RNase. Les transcrits sont purifiés par extraction avec un volume de mélange phénol-chloroforme-Alcool isoamylique (25/24/1) (v/v/v), puis avec un volume de mélange chloroforme-Alcool isoamylique (24/1) (v/v). Ils sont précipités à l'éthanol en présence d'acétate d'ammonium à une concentration finale de 2,5 M, puis lavés à l'éthanol 70 %. Les transcrits sont ensuite repris dans de l'eau milliQ stérile à raison de 100 ng/μl [quantification par mesure de l'absorbance à 260 nm avec un appareil Genequant (Pharmacia)] et conservés à -80°C.

## III.4.b / Analyse de la structure secondaire des ARN par sondes enzymatiques et chimiques

### Principe:

Pour identifier les régions en simple-brin ou double-brin dans l'ARN, nous avons utilisé des sondes enzymatiques (nucléases V1, T1 et T2) et des sondes chimiques [diméthylsulfate (DMS), le 1-cyclohexyl-3-(2-morpholinoéthyl)-carbodiimide-métho-p-toluène sulfonate (CMCT) et le kéthoxal]. Les cibles de ces sondes sont présentées dans la Figure 61 de la partie des Résultats. Les hydrolyses enzymatiques et les modifications chimiques sont réalisées dans des conditions ménagées en présence d'ARNt, afin d'avoir une coupure ou une modification par molécule d'ARN. Les liaisons phosphodiesters clivées par les enzymes et les bases modifiées par les sondes chimiques sont mises en évidence par extension d'une amorce oligonucléotidique par la transcriptase inverse. Cet oligonucléotide est marqué en 5' au <sup>32</sup>P suivant le protocole décrit dans le paragraphe II.2. Une réaction d'élongation est effectuée sur l'ARN intact, ce qui permet de différencier les arrêts d'élongation dus aux modifications chimiques et coupures enzymatiques des pauses de la transcriptase inverse en cours d'élongation. Les produits d'élongation sont séparés par électrophorèse en gel de polyacrylamide dénaturant, en parallèle avec les produits de la réaction de séquençage réalisée avec la

transcriptase inverse et le même oligonucléotide amorce. La position des bases modifiées et des coupures enzymatiques est déduite par comparaison avec le séquençage.

## Protocole expérimental pour les hydrolyses enzymatiques :

Cent ng de transcrit froid et 2 μg d'ARNt de levure sont préincubés 5 min à 65°C, puis 20 min à 20°C, dans 50 μl de tampon KCl 350 mM, MgCl<sub>2</sub> 10 mM, Tris-HCl 10 mM, pH 7,5 avant l'ajout d'une RNase. Les hydrolyses enzymatiques sont réalisées de façon ménagée par incubation, 10 min à 20°C en présence de 0,05 U ou 0,10 U de RNase V1, 1,25 U ou 2,5 U de RNase T1, ou 2,4 U ou 4,8 U de RNase T2. L'action de la RNase V1 est arrêtée par addition de 100 μl d'EDTA à 100 mM, alors que les RNases T1 et T2 sont stoppées par l'ajout de 20 μg d'ARNt. Après extration phénolique, les ARN sont précipités en présence d'acétate de sodium 0,3 M final et d'éthanol, puis lavés à l'éthanol 70 %, et enfin repris dans 4 μl d'eau milliQ stérile. Les élongations d'amorce par la transcriptase inverse sont réalisées avec 1 μl de la préparation ainsi obtenue.

### Protocole expérimental des modifications chimiques :

Le transcript étudié est préincubé 5 min à 65°C dans 50 μl de tampon 100 mM KCl, 10 mM MgCl<sub>2</sub>, 500 mM cacodylate de sodium, pH 7,5 pour les modifications chimiques par le DMS ou kéthoxal, ou dans 50 µl de tampon 100 mM KCl, 10 mM MgCl<sub>2</sub>, 50 mM borate sodium, pH 8,0 pour les modifications chimiques par le CMCT, puis 20 min à 20°C (refroidissement lent). Les modifications chimiques au DMS sont réalisées de facon ménagée pendant 6 min après ajout de 0,5 ou 1 µl d'une solution de DMS 100 % diluée au ¼ dans de l'éthanol absolu (concentration finale de 0,03 M). Les modifications chimiques par le kéthoxal sont réalisées par incubation pendant 10 min après l'ajout de 2 ou 5 μl de kéthoxal à 37 mg/l, dilué au ¼ dans de l'eau à 20°C. Les modifications au CMCT sont réalisées par incubation pendant 10 min à 20°C après ajout de 840 ou 1680 µg de CMCT. Les modifications chimiques au DMS sont arrêtées par l'ajout de 10 µl de solution G-stop (Acétate de sodium 1,5 M pH 5,3; β-mercaptoéthanol 1 M; EDTA 0,1 mM; Tris-acétate 1 M, pH 7,5). Les modifications chimiques au kéthoxal sont stoppées par l'ajout de 10 µl de solution borate de potassium 20 mM, pH 7,0. Les modifications chimiques au CMCT sont arrêtées par l'ajout d'1 µg d'ARNt. Les transcrits sont ensuite précipités à l'éthanol en présence d'acétate de sodium 0,3 M final, puis lavés à l'éthanol, et repris dans 4 µl d'eau milliQ stérile. Un volume de 1 µl de cette préparation est ensuite analysé par élongation d'amorce par la transcriptase inverse.

## Prédictions de structure secondaire par ordinateur :

Nous avons utilisé le programme Mfold disponible à partir du site internet <a href="http://www.bioinfo.rpi.edu/applications/mfold/old/rna/form1.cgi">http://www.bioinfo.rpi.edu/applications/mfold/old/rna/form1.cgi</a> (Zuker, 2003). Ce programme est basé

sur les estimations de la stabilité thermodynamique des appariements de base de l'ARN pour un ARN en solution dans un tampon NaCl 1M à 37°C. Les résultats des analyses expérimentales précédentes sont introduits en tant que contraintes dans le logiciel de recherche.

## III.5 / Etude d'interaction ARN-protéines par gel retard

## III.5.a / Production d'ARN radiomarqués par transcription *in vitro* avec l'ARN polymérase du phage T7

La synthèse de transcrits radioactifs par l'ARN polymérase du phage T7 a été réalisée en présence de rUTP  $[\alpha^{-32}P]$  dans le mélange réactionnel suivant :

| ADN $(0.5 \mu g/\mu l)$                                      | 1 μ1      |
|--|-----------|
| Tampon de transcription 3x                                   | 10 μ1     |
| (MgCl <sub>2</sub> 60 mM; Spermidine 3 mM; DTT 15 mM;        |           |
| triton 0,03 %; PEG 8 000 240 mg/ml; Tris-HCl 120 mM pH 8,0). |           |
| rNTP (ATP, CTP, GTP 20 mM)                                   | 6 μl      |
| rUTP (5 mM)  | 2,4 μl    |
| [α- <sup>32</sup> P] UTP (800 Ci/mmol; 10 mCi/ml) (ICN)      | 2,4 μl    |
| RNasin (31 U/µl) (Amersham Biosciences)                      | 0,5 μ1    |
| ARN polymérase T7 (purifiée au Laboratoire) (20 U/µl)        | 1 μl      |
| BSA (2 mg/ml)  | 3 μl      |
| $H_2O$   | qsp 30 μl |

La réaction est incubée à 37°C pendant de 2 h, puis la matrice ADN est dégradée par incubation pendant 30 min à 37°C en présence de 10 U de DNase I exempte de RNase.

Les transcrits sont purifiés sur gel de polyacrylamide dénaturant (polyacrylamide 6 %, urée 8 M, TBE 1 X), puis élués une nuit à 4°C dans du tampon d'élution à forte concentration saline (NaCl 300 mM; EDTA 1 mM; SDS 1 %; Tris-HCl 10 mM pH 7,5). Les transcrits sont ensuite précipités à l'éthanol puis lavés à l'éthanol 70 %. Les transcrits sont repris dans 10 µl d'eau stérile. La quantité de transcrits radioactifs obtenue est déterminée par mesure de la radioactivité contenue dans 1 µl d'ARN au compteur à scintillation (Beckman).

# III.5.b / Etude des complexes [ARN-Hfq]

# III.5.b.α / Formation des complexes ARN-protéines

Les complexes [ARN-Hfq] sont formés dans du tampon Hfq 1x (MgCl<sub>2</sub> 10 mM; NaCl 50 mM; KCl 50 mM; Tris-HCl 10 mM, pH 8,0) en incubant 30 min à 37°C, 100 fmol d'ARN marqué uniformément (5000 cpm) en présence de protéine Hfq en concentrations croissantes, de 20 ng d'ARNt de levure (Roche Molecular Biochemicals) et 0,5 μl de RNasin (Roche Molecular Biochemicals, 29,1 U/μl).

# III.5.b. $\beta$ / Fractionnement des complexes [ARN-Hfq] par la technique de retard sur gel

# Principe:

Cette technique de retard sur gel appelée aussi EMSA (Electrophoretic Mobility Shift Assay) est basée sur le changement de mobilité électrophorétique d'un ARN lorsqu'il est lié à une protéine. En absence de protéines, l'ARN supposé porter des sites de fixation pour des protéines migre à une position donnée dans un gel non dénaturant. S'il est fixé par une ou plusieurs protéines, sa migration est retardée.

# Protocole:

Un μL de bleu pour gel retard [TBE 1x ; Glycérol 50 % (p/v) ; bleu de bromophénol 0,1 % (p/v) ; xylène cyanol 0,1 % (p/v)] sont ajoutés au mélange réactionnel ayant servi à former le complexe. L'échantillon est ensuite déposé sur un gel de polyacrylamide natif, acrylamide-bisacrylamide (38:2) 6 %, en présence de TBE 0,5 x. L'électrophorèse se fait à 15 V/cm et à température ambiante. Le gel est ensuite séché puis exposé au PhosphorImager (Molecular Dynamics).

# II.5.b.γ / Expériences d'empreintes

# Principe:

La fixation d'une protéine sur l'ARN protège cet ARN contre l'action des sondes enzymatiques ou chimiques habituellement utilisées dans les analyses de structure secondaire. Ainsi, l'absence d'un clivage enzymatique ou d'une modification chimique au niveau d'un site reflète une protection par la protéine.

#### Protocole:

Au préalable, la formation des complexes ARN-protéines aura été vérifiée dans les différents tampons utilisés pour chaque agent modifiant.

Le transcrit froid est incubé en présence de la protéine à une concentration pour laquelle une interaction avec le transcrit a été détectée, avec en parallèle un témoin de clivage enzymatique ou de modification chimique dans lequel le transcrit est incubé sans la protéine. Les hydrolyses enzymatiques ou les modifications chimiques sont réalisées dans les mêmes conditions que précédemment. Avant la précipitation, une extraction au phénol-chloroforme est réalisée pour éliminer la protéine.

# IV / Analyse des protéines cellulaires

# IV.1 / Sonication des bactéries

Chaque culot de cellules est repris dans 2 ml de tampon de sonication froid, soit classiquement le tampon EDTA 2 mM; Tris-HCl 50 mM pH 7,0, ou, lorsque l'activité FBA dépendante du zinc est mesurée, le tampon de sonication modifié Tris-HCl 20 mM pH 7,0. La sonication est alors réalisée à 0°C sous une puissance de 20 W durant 2 à 3 min par période d'1 sec d'ultrason et 1 sec d'arrêt. Les débris cellulaires sont éliminés par centrifugation et les protéines sont quantifiées par mesure de l'absorbance du surnageant à 280 nm.

# IV.2 / Analyse des protéines par SDS-PAGE

Une quantité de surnageant de sonication correspondant à 30 ou 50 milliA<sub>280nm</sub> de surnageant est additionné de « bleu de protéines », puis déposé sur un gel dénaturant selon la méthode décrite par Laemmli (1970), l'appareillage Mini Protean (Biorad) est utilisé à cet usage. Les protéines sont ensuite colorées en utilisant un tampon contenant du bleu de Commassie R-250 (« bleu de coloration »).

# IV.3 / Mesure des activités enzymatiques

#### IV.3.a / Mesure d'activité GAPDH

**Principe :** La GAPDH catalyse la réaction de conversion du glycéraldéhyde-3-phosphate (G3P) en 1,3-diphosphoglycérate (1,3-DPG) (Fig. 4 et 53 de l'Introduction et Présentation du Sujet). Lors de la

réaction, le NAD<sup>+</sup> est réduit en NADH + H<sup>+</sup>. L'activité enzymatique est donc suivie par apparition du NADH en fonction du temps. Cette apparition est mesurée par la variation de l'absorbance à 340 nm (test de Ferdinand, 1964). Pour chaque détermination du taux de GAPDH, 2 essais enzymatiques sont réalisés avec deux volumes différents d'extrait cellulaire, avec un essai où la pente  $\Delta A_{340}$ /min est comprise entre 0,1 et 0,2. Ceci permet de vérifier la proportionnalité entre la vitesse d'apparition du NADH et le taux de protéines GAPDH.

# Préparation du substrat et du cofacteur :

G<sub>3</sub>P: 200 mg de diéthylacétal D-G3P (DL-glycéraldéhyde-3-phosphate diéthylacétal, sel de monobarium; Sigma Chemical) et 3 g de résine DOWEX 50 (activée) dans 5 ml d'eau sont portés à 100°C au bain marie pendant 3 min. Après décantation de la résine, le surnageant est prélevé puis filtré. La solution est dosée précisément par un test enzymatique avec de la GAPDH purifiée en présence de NAD<sup>+</sup> et d'arsenate (l'arsenate rend la réaction irréversible).

NAD<sup>+</sup>: le NAD<sup>+</sup> (free acid, grade I ; Boehringer) est dissous dans un tampon EDTA 20 mM, Tris-HCl 500 mM, pH 8,0. La concentration initiale est déterminée par mesure de l'absorbance à 259 nm et est ensuite ajustée à 100 mM.

# Test enzymatique:

Le milieu réactionnel suivant est préparé :

G3P 1 mM NAD<sup>+</sup> (100 mM) 5  $\mu$ l

Triéthanolamine 40 mM, pH ajusté à 9,2 par H<sub>3</sub>PO<sub>4</sub><sup>2</sup>-)

La réaction est amorcée par ajout de l'extrait protéique.

Tampon Ferdinand (K<sub>2</sub>HPO<sub>4</sub> 50 mM; EDTA 0,2 mM,

L'activité spécifique (AS) de la GAPDH, mesurée en  $\mu$ moles de NADH / minute /  $A_{280nm}$  est calculée à partir de :

qsp 500 μl

Pente ( $\Delta A_{340}$ /min) x dilution x Volume de la cuve AS = Volume de l'échantillon ( $\mu$ l) x 6,22 x  $A_{280nm}$ 

Le facteur de 6,22 est le coefficient d'extinction molaire du NADH à 340 nm (6220 M<sup>-1</sup>.cm<sup>-1</sup>).

#### IV.3.b / Mesure d'activité PGK

# Principe:

La PGK catalyse la réaction réversible de conversion du 3-phosphoglycérate en 1,3-diphosphoglycérate en présence d'ATP (Fig. 4 et 57 de l'Introduction et Présentation du Sujet). La mesure de l'activité PGK est déterminée par un test impliquant un couplage réactionnel. Le 1,3-diphosphoglycérate produit est aussitôt consommé par la GAPDH, présente en excès dans le mélange réactionnel. Cette seconde réaction conduit à la production de G<sub>3</sub>P et à l'oxydation du NADH + H<sup>+</sup> en NAD<sup>+</sup>. L'activité enzymatique est donc suivie par la disparition du NADH en fonction du temps en mesurant la variation de l'absorbance à 340 nm. Pour chaque détermination du taux de PGK, 2 essais enzymatiques sont réalisés avec deux volumes différents d'extrait cellulaire, avec un essai où la pente ΔA<sub>340nm</sub>/min est comprise entre 0,1 et 0,2. Ceci permet de vérifier la proportionnalité entre la vitesse de consommation du NADH et le taux de protéine PGK.

# Test enzymatique:

Le mélange réactionnel suivant est préparé :

| ATP (500 mM)                             | 5       | μl |
|--|---------|----|
| MgSO <sub>4</sub> (1 M)                  | 2,5     | μl |
| NADH (100 mM)                            | 1       | μl |
| 3-phosphoglycérate (50 mM)               | 20      | μl |
| GAPDH (7 A <sub>280nm</sub> /ml)         | 15      | μl |
| Tampon (EDTA 2mM; Tris-HCl 50 mM pH 8,0) | qsp 500 | μl |

La GAPDH de *B. stearothermophilus* a été purifiée au laboratoire par S. Azza, de l'équipe d'Enzymologie Moléculaire. La réaction est amorcée par ajout de l'extrait protéique.

L'activité spécifique (AS) de la PGK, mesurée en  $\mu$ moles de NADH / minute /  $A_{280nm}$  est calculée à partir de :

Pente (
$$\Delta A_{340nm}$$
/min) x dilution x Volume de la cuve  
AS = - Volume de l'échantillon ( $\mu$ L) x 6,22 x  $A_{280nm}$ 

# IV.3.c / Mesures d'activité FBA

# Principe:

La FBA catalyse le clivage réversible du Fructose-1,6-bisphosphate en dihydroxyacétone phosphate et G<sub>3</sub>P (Fig. 4 et 50 de l'Introduction et Présentation du Sujet). La mesure de l'activité FBA

est réalisée par un test impliquant un couplage réactionnel. Le G<sub>3</sub>P produit est aussitôt consommé par la GAPDH, présente en excès dans le mélange réactionnel. Cette seconde réaction conduit à la production de 1,3-diphosphoglycérate et à la réduction du NAD<sup>+</sup> en NADH + H<sup>+</sup>. L'activité enzymatique est donc suivie par l'apparition du NADH en fonction du temps en mesurant la variation de l'absorbance à 340 nm. Pour chaque détermination du taux de FBA, 2 essais enzymatiques sont réalisés avec deux volumes différents d'extrait cellulaire, avec un essai où la pente ΔA<sub>340nm</sub>/min est comprise entre 0,1 et 0,2. Ceci permet de vérifier la proportionnalité entre la vitesse d'apparition du NADH et le taux de protéine FBA.

# Test enzymatique:

Le mélange réactionnel suivant est préparé :

| Fructose-1,6-bisphosphate (100 mM) | 10 μ1      |
|------------------------------------|------------|
| NAD <sup>+</sup> (100 mM)          | 10 μl      |
| Arsenate de sodium (500 mM)        | 17 μl      |
| GAPDH (7 A <sub>280nm</sub> /ml)   | 15 μΙ      |
| Tampon (Tris-HCl 50 mM; pH 8,0)    | gsp 500 μl |

La réaction est amorcée par ajout de l'extrait protéique.

L'activité spécifique (AS) de la FBA, mesurée en  $\mu$ moles de NADH / minute /  $A_{280nm}$  est calculée à partir de :

# V / Analyse des surnageants de cultures

Nous avons mesuré les concentrations en glucose et acide acétique dans les surnageants de cultures, soit par une méthode enzymatique, soit par chromatographie liquide à haute performance (CLHP). Nous décrivons les méthodes enzymatiques pour chaque composé puis l'analyse par CLHP.

# V.1 / Dosage enzymatique du glucose

La concentration en glucose a été mesurée soit par chromatographie liquide à haute performance (CLHP) ou par une méthode enzymatique utilisant la glucose oxydase couplée à la péroxydase (PGO enzymes, Sigma Diagnostic<sup>®</sup>). Le principe du dosage est le suivant :

- la glucose oxydase catalyse la conversion du glucose en acide gluconique suivant la réaction suivante : Glucose  $+ O_2 + H_2O$   $\longrightarrow$  Acide gluconique  $+ H_2O_2$ 

Le péroxyde d'hydrogène libéré permet, en présence de la peroxydase, d'oxyder l'orthodianisidine ce qui se traduit par une augmentation de l'absorbance à 435 nm.

Pour chaque série de mesures, une gamme étalon est établie, avec des solutions en glucose de 0,1 à 0,5 g/l.

# V.2 / Dosage enzymatique de l'acide acétique

L'acide acétique a été mesuré soit en utilisant le kit de Boehringer Manheim.

# V.3 / Dosage par CLHP du glucose et de l'acide acétique

Le dispositif est composé de :

- un dégazeur en ligne Thermo Separation Products,
- une pompe isocratique SP 8810 LC Pump (Spectra Physics), et un injecteur automatique (boucle d'injection de 20 μl) (Spectra Physics),
- une colonne échangeuse de cations IOA 2000 (150 x 6.5 mm, Alltech) munie d'un chauffecolonne G3080 (Gecko-CIL),
- un détecteur par indice de réfraction SP8430 (Spectra Physics),
- un intégrateur Data Jet integrator (Thermoseparation Products).

L'échantillon de 20  $\mu$ l injecté est dilué au dixième dans de l'eau déionisée et filtrée (Millex, 0,22  $\mu$ m, Millipore). L'élution est réalisée par de l'acide sulfurique ( $H_2SO_4$  0,01 M à un débit de 0,5 ml/min et une pression moyenne de 60 bars. Le solvant a été au préalable filtré sur une membrane nylon Whatmann 0,2  $\mu$ m par un ensemble Millipore filtration/dégazage.

Le profil chromatographique obtenu à 30°C permet d'identifier 2 pics caractéristiques du glucose et de l'acide acétique. Les temps de rétention de chacun de ces deux composés sont de 3,44 pour le glucose et 7,41 pour l'acide acétique. La concentration de chacun des composés est déduite par comparaison des surfaces des pics avec une gamme étalon obtenue par injection de solution de concentration connue pour chacun des composés.

# VI / Exploitations des données expérimentales

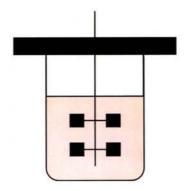
# VI.1 / Détermination du poids sec

Pour déterminer la densité cellulaire d'une culture (en g/l), nous avons établi la corrélation entre l'absorbance à 600 nm et la matière sèche cellulaire. Pour cela, nous avons prélevé un volume de 10 ml d'une suspension cellulaire au cours d'une culture en phase exponentielle de croissance. L'absorbance à 600 nm a été mesurée, et l'échantillon de 10 ml a été centrifugé à 4 000 tours/min pendant 20 min. Le culot a ensuite été repris dans 1,5 mL d'une solution physiologie (9 ‰) puis centrifugé dans des micro-tubes pré-tarés. Cette suspension est centrifugée 2 min à 12 000 tours/min. Après élimination du surnageant, chaque tube est placé dans une étuve à 70°C jusqu'à ce que le contenu ait un poids constant (48 à 72 heures).

Nous avons ainsi pu déterminer les courbes de biomasse en fonction de l'absorbance à 600 nm pour les souches  $E.\ coli\ DH5\alpha$ ,  $E.\ coli\ TP2503$  et  $E.\ coli\ TP2504$ . Nous avons ainsi pu établir qu'une unité  $A_{600}$  correspond à un poids sec bactérien de :

- 0,52 (+/-0,12) g/l pour *E. coli* DH5 $\alpha$
- 0,45 (+/-0,02) g/l pour E. coli TP2503
- 0,38 (+/-0,05) g/l pour E. coli TP 2504.

# VI.2 / Détermination des caractéristiques de croissances des souches d'*E. coli* en cultures batch



V : volume utile (l)

X : concentration en biomasse (g.l<sup>-1</sup>)

S : concentration en substrat (g/l)

P : concentration en produit (g/l)

S0 : Concnetration de substrat dans le fermenteur à t=0 (g/l)

P0 : Concentration de produit dans le fermenteur à t=0 (g/l)

Figure 93 : Schématisation d'une culture batch.

A partir des courbes de croissances [biomasse (X) en fonction du temps], d'assimilation des substrats (S en fonction du temps) et de formation des produits (P), il est possible de déterminer la vitesse

croissance ( $\vec{r}_X = dX / dt$  en  $g.l^{-1}.h^{-1}$ ), la vitesse de consommation du substrat ( $\vec{r}_S = dS / dt$  en  $g.L^{-1}.h^{-1}$ ) et la vitesse de production d'un produit ( $\vec{r}_P = dS$  dt en  $g.l^{-1}.h^{-1}$ ) en déterminant la dérivée de chacune des variables considérées par rapport au temps.

Il est ensuite possible de déterminer les différentes **vitesses spécifiques** par le rapport entre la dérivée de la variable considérée et la valeur de la biomasse au temps considéré. Nous avons alors :

- vitesse de croissance spécifique :  $\mu = r''_X/X$  (g.g<sup>-1</sup>.h<sup>-1</sup>)
- vitesse spécifique de consommation d'un substrat :  $q_S = -r_S^{(1)} / X (g.g^{-1}.h^{-1})$
- vitesse spécifique de production d'un produit :  $q_P = r^{"}_P / X (g.g^{-1}.h^{-1})$

# Les rendements globaux sont définis par les équations suivantes :

rendement apparent de production de biomasse à partir d'un substrat donné (g/g) :

$$R_{X/S} = (X - X0) / (S0 - S)$$

rendement apparent de production de produit à partir d'un substrat donné (g/g) :

$$R_{P/S} = (P - P0) / (S0 - S)$$

Avec S0 la concentration de substrat dans le fermenteur à t = 0 (g/l) et P0, la concentration en produit dans le fermenteur à t = 0 (g/l).

# VI.3 / Détermination des caractéristiques de croissances des souches d'*E. coli* en cultures continues

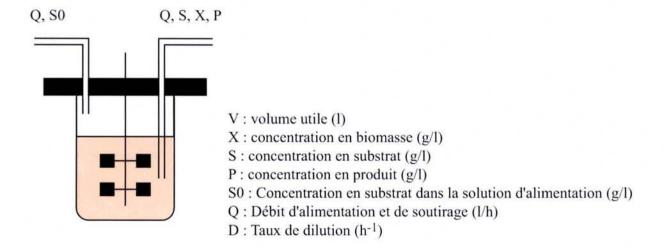


Figure 94 : Schématisation d'une culture continue.

Le débit d'alimentation et de soutirage (l.h-1) est exprimé par  $Q = \Delta V / \Delta t$ .

Le bilan de la biomasse est :  $r_X = Q / V \cdot X + dX / dt$ 

A l'état stable, la biomasse ne varie plus. Nous avons alors :  $r'''_X = Q/V$ . X.

Or Q / V = D et  $\mu = r_X^{\prime\prime\prime}$  X, nous avons :  $\mu$ .X = D.X, soit  $\mu$  = D.

Ainsi, en fermentation continue à l'état stable, le taux de croissance des bactéries est égal au taux de dilution du fermenteur D.

Le bilan du substrat est :  $r_S^{"} = Q / V$ . (S0 - S) - dS / dt

A l'état stable, nous avons :  $r'''_S = Q / V$  .  $(S0 - S) = \mu (S0 - S)$ .

Le bilan du produit est :  $r_P^{"} = Q / V \cdot P + dP / dt$ 

A l'état stable, nous avons :  $r'''_P = Q / V \cdot P = D \cdot P = \mu \cdot P$ 

SCD UHP NANCY 1 Bibliothàque des Sciences Rue du Jardin Botanique - CS 20148 54601 VILLERS LES NANCY CEDEX

# Bibliographie

# A

- Abe, H. et Aiba, H. <u>1996</u>. Differential contributions of two elements of rho-independent terminator to transcription termination and mRNA stabilization. *Biochimie*. **78**:1035-42.
- Ades, S. E., Connolly, L. E., Alba, B. M. et Gross, C. A. 1999. The *Escherichia coli* σ<sup>E</sup>-dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti-σ factor. *Genes Dev.* 13:2449-61.
- Ades, S. E., Grigorova, I. L. et Gross, C. A. 2003. Regulation of the alternative sigma factor σ<sup>E</sup> during initiation, adaptation, and shutoff of the extracytoplasmic heat shock response in Escherichia coli, J Bacteriol. 185:2512-9.
- Ades, S. E. <u>2004.</u> Control of the alternative sigma factor σ<sup>E</sup> in *Escherichia coli*. *Curr Opin Microbiol*. 7:157-62.
- Afonyushkin, T., Vecerek, B., Moll, I., Blasi, U. et Kaberdin, V. R. 2005. Both RNase E and RNase III control the stability of *sodB* mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic Acids Res.* 33:1678-89.
- Aiba, H., Adhya, S. et De Crombrugghe, B. 1981. Evidence for two functional gal promoters in intact Escherichia coli cells. J. Biol. Chem. 256:11905-10.
- Aiba, H. <u>1983.</u> Autoregulation of the *Escherichia coli crp* gene: CRP is a transcriptional repressor for its own gene. *Cell.* 32:141-9.
- Aiba, H. 1985. Transcription of the Escherichia coli adenylate cyclase gene is negatively regulated by cAMP-cAMP receptor protein. J. Biol. Chem. 260:3063-70.
- Alefounder, P. R. et Perham, R. N. 1989. Identification, molecular cloning and sequence analysis of a gene cluster encoding the class II fructose 1,6-bisphosphate aldolase, 3-phosphoglycerate kinase and a putative second glyceraldehyde 3-phosphate dehydrogenase of Escherichia coli. Mol Microbiol. 3:723-32.
- Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S. et Ishihama, A. <u>1999</u>. Growth phase-dependent variation in protein composition of the *Escherichia* coli nucleoid. *J Bacteriol*. **181**:6361-70.
- Alifano, P., Rivellini, F., Piscitelli, C., Arraiano, C. M., Bruni, C. B. et Carlomagno, M. S. 1994. Ribonuclease E provides substrates for ribonuclease P-dependent processing of a polycistronic mRNA. Genes Dev. 8:3021-31.
- Aloy, P., Ciccarelli, F. D., Leutwein, C., Gavin, A. C., Superti-Furga, G., Bork, P., Bottcher, B. et Russell, R. B. <u>2002</u>. A complex prediction: threedimensional model of the yeast exosome. *EMBO Rep.* 3:628-35.
- Altuvia, S., Locker-Giladi, H., Koby, S., Ben-Nun, O. et Oppenheim, A. B. <u>1987</u>. RNase III stimulates the translation of the cIII gene of bacteriophage lambda. *Proc Natl Acad Sci U S A*. **84**:6511-5.
- Anderson, J. S. 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.* 17:1497-506.
- Angerer, A., Enz, S., Ochs, M. et Braun, V. 1995. Transcriptional regulation of ferric citrate transport in *Escherichia coli* K-12. Fecl belongs to a new subfamily of σ<sup>70</sup>-type factors that respond to extracytoplasmic stimuli. *Mol. Microbiol.* 18:163-74.
- Antoun, A., Pavlov, M. Y., Andersson, K., Tenson, T. et Ehrenberg, M. 2003. The roles of initiation factor 2 and guanosine triphosphate in initiation of protein synthesis. *EMBO J.* 22:5593-601.
- Arluison, V., Derreumaux, P., Allemand, F., Folichon, M., Hajnsdorf, E. et Regnier, P. 2002. Structural Modelling of the Sm-like Protein Hfq from Escherichia coli. J Mol Biol. 320:705-12.
- Arnold, T. E., Yu, J. et Belasco, J. G. 1998. mRNA stabilization by the *ompA* 5' untranslated region: two protective elements hinder distinct pathways for mRNA degradation. RNA. 4:319-30.
- Arraiano, C. M., Yancey, S. D. et Kushner, S. R. <u>1988</u>. Stabilization of discrete mRNA breakdown products in *ams pnp rnb* multiple mutants of *Escherichia coli* K-12. *J Bacteriol*. **170**:4625-33.
- Artsimovitch, I., Patlan, V., Sekine, S., Vassylyeva, M. N., Hosaka, T., Ochi, K., Yokoyama, S. et Vassylyev, D. G. <u>2004</u>. Structural basis for transcription regulation by alarmone ppGpp. *Cell*. 117:299-310.
- Avison, M. B., Horton, R. E., Walsh, T. R. et Bennett, P. M. 2001. Escherichia coli CreBC is a global regulator of gene expression that responds to growth in minimal media. J Biol Chem. 276:26955-61.
- Azam, T. A. et Ishihama, A. <u>1999</u>. Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J Biol Chem.* **274**:33105-13.
- Azam, T. A., Hiraga, S. et Ishihama, A. <u>2000</u>. Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes Cells*. 5:613-26.

# В

- Babitzke, P., Granger, L., Olszewski, J. et Kushner, S. R. 1993. Analysis of mRNA decay and rRNA processing in *Escherichia coli* multiple mutants carrying a deletion in RNase III. *J Bacteriol*. 175:229-39.
- Bachellier, S., Clement, J. M. et Hofnung, M. 1999. Short palindromic repetitive DNA elements in enterobacteria: a survey. Res Microbiol. 150:627-39.
- Baker, C. S., Morozov, I., Suzuki, K., Romeo, T. et Babitzke, P. <u>2002</u>. CsrA regulates glycogen biosynthesis by preventing translation of glgC in Escherichia coli. Mol Microbiol. 44:1599-610.
- Baker, K. E. et Mackie, G. A. <u>2003</u>. Ectopic RNase E sites promote bypass of 5'-end-dependent mRNA decay in *Escherichia coli*. *Mol Microbiol*. **47**:75-88.
- Baldwin, N. E. et Dombroski, A. J. <u>2001</u>. Isolation and characterization of mutations in region 1.2 of *Escherichia coli* σ<sup>70</sup>. *Mol Microbiol*. **42**:427-37.

- Baldwin, S. A. et Perham, R. N. 1978. Novel kinetic and structural properties of the class-I D-fructose 1,6bisphosphate aldolase from Escherichia coli (Crookes' strain). Biochem J. 169:643-52.
- Baneyx, F. 1999. Recombinant protein expression in Escherichia coli. Curr Opin Biotechnol. 10:411-21.
- Bardey, V. 2000. Organisation et régulation de l'expression des gènes de la Glycéraldéhyde-3phosphate déshydrogénase, la Phosphoglycérate kinase et l'Erythrose-4-phosphate déshydrogénase chez les Protéobactéries de la subdivision γ, thèse, Université Henri Poincaré.
- Bar-Nahum, G. et Nudler, E. <u>2001</u>. Isolation and characterization of σ<sup>70</sup>-retaining transcription elongation complexes from *Escherichia coli*. *Cell*. **106**:443-51.
- Barne, K. A., Bown, J. A., Busby, S. J. et Minchin, S. D. 1997. Region 2.5 of the *Escherichia coli* RNA polymerase σ<sup>70</sup> subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J.* 16:4034-40.
- Berlyn, M. K. <u>1998.</u> Linkage map of *Escherichia coli* K-12, edition 10: the traditional map. *Microbiol Mol Biol Rev.* **62**:814-984.
- Bermudez-Cruz, R. M., Garcia-Mena, J. et Montanez, C. 2002. Polynucleotide phosphorylase binds to ssRNA with same affinity as to ssDNA. *Biochimie*. 84:321-8.
- Bernstein, B. E., Michels, P. A. et Hol, W. G. <u>1997.</u> Synergistic effects of substrate-induced conformational changes in phosphoglycerate kinase activation. *Nature.* **385**:275-8.
- Bernstein, B. E. et Hol, W. G. <u>1998.</u> Crystal structures of substrates and products bound to the phosphoglycerate kinase active site reveal the catalytic mechanism. *Biochemistry*. **37**:4429-36.
- Bernstein, J. A., Khodursky, A. B., Lin, P. H., Lin-Chao, S. et Cohen, S. N. 2002. Global analysis of mRNA decay and abundance in *Escherichia coli* at singlegene resolution using two-color fluorescent DNA microarrays. *Proc Natl Acad Sci U S A.* 99:9697-702.
- Bernstein, J. A., Lin, P. H., Cohen, S. N. et Lin-Chao, S. <u>2004.</u> Global analysis of *Escherichia coli* RNA degradosome function using DNA microarrays. *Proc Natl Acad Sci U S A.* **101**:2758-63.
- Berry, M. D. <u>2004.</u> Glyceraldehyde-3-phosphate dehydrogenase as a target for small-molecule disease-modifying therapies in human neurodegenerative disorders. *J. Psychiatry. Neurosci.* **29**:337-45.
- Bessarab, D. A., Kaberdin, V. R., Wei, C. L., Liou, G. G. et Lin-Chao, S. 1998. RNA components of Escherichia coli degradosome: evidence for rRNA decay. Proc Natl Acad Sci USA. 95:3157-61.
- Bijlsma, J. J. et Groisman, E. A. <u>2003</u>. Making informed decisions: regulatory interactions between twocomponent systems. *Trends Microbiol*. 11:359-66.
- Binns, N. et Masters, M. <u>2002</u>. Expression of the Escherichia coli pcnB gene is translationally limited using an inefficient start codon: a second chromosomal example of translation initiated at AUU. Mol Microbiol. 44:1287-98.

- Bjornsson, A., Mottagui-Tabar, S. et Isaksson, L. A. 1996. Structure of the C-terminal end of the nascent peptide influences translation termination. *EMBO J.* 15:1696-704.
- Blaszczyk, J., Tropea, J. E., Bubunenko, M., Routzahn, K. M., Waugh, D. S., Court, D. L. et Ji, X. 2001. Crystallographic and modeling studies of RNase III suggest a mechanism for double-stranded RNA cleavage. Structure (Camb). 9:1225-36.
- Blaszczyk, J., Gan, J., Tropea, J. E., Court, D. L., Waugh, D. S. et Ji, X. <u>2004</u>. Noncatalytic assembly of ribonuclease III with double-stranded RNA. *Structure (Camb)*. **12**:457-66.
- Bledig, S. A., Ramseier, T. M. et Saier, M. H., Jr. 1996.
  FruR mediates catabolite activation of pyruvate kinase (pykF) gene expression in Escherichia coli. J Bacteriol. 178:280-3.
- 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. 26:387-98.
- Blum, E., Carpousis, A. J. et Higgins, C. F. 1999. Polyadenylation promotes degradation of 3'structured RNA by the *Escherichia coli* mRNA degradosome in vitro. J Biol Chem. 274:4009-16.
- Boni, I. V., Artamonova, V. S., Tzareva, N. V. et Dreyfus, M. <u>2001</u>. Non-canonical mechanism for translational control in bacteria: synthesis of ribosomal protein S1. *EMBO J.* **20**:4222-32.
- Borukhov, S. et Severinov, K. <u>2002</u>. Role of the RNA polymerase sigma subunit in transcription initiation. *Res Microbiol*. **153**:557-62.
- Boschi-Muller, S., Azza, S., Pollastro, D., Corbier, C. et Branlant, G. 1997. Comparative enzymatic properties of GapB-encoded erythrose-4-phosphate dehydrogenase of *Escherichia coli* and phosphorylating glyceraldehyde-3-phosphate dehydrogenase. *J Biol Chem.* 272:15106-12.
- Botsford, J. L. et Harman, J. G. 1992. Cyclic AMP in prokaryotes. Microbiol. Rev. 56:100-22.
- Bouvet, P. et Belasco, J. G. <u>1992</u>. Control of RNase E-mediated RNA degradation by 5'-terminal base pairing in *E. coli. Nature*. **360**:488-91.
- Boyd, D. A., Cvitkovitch, D. G. et Hamilton, I. R. 1995. Sequence, expression, and function of the gene for the nonphosphorylating, NADP-dependent glyceraldehyde-3-phosphate dehydrogenase of Streptococcus mutans. J Bacteriol. 177:2622-7.
- Branlant, C., Krol, A., Ebel, J. P., Lazar, E., Haendler, B. et Jacob, M. <u>1982</u>. U2 RNA shares a structural domain with U1, U4, and U5 RNAs. *EMBO J*. **1**:1259-65.
- Branlant, G., Flesch, G. et Branlant, C. <u>1983</u>. Molecular cloning of the glyceraldehyde-3-phosphate dehydrogenase genes of *Bacillus stearothermophilus* and *Escherichia coli*, and their expression in *Escherichia coli*. *Gene*. **25**:1-7.
- Branlant, G. et Branlant, C. <u>1985</u>. Nucleotide sequence of the *Escherichia coli gap* gene. Different evolutionary behavior of the NAD<sup>+</sup>-binding domain and of the catalytic domain of D-glyceraldehyde-3-phosphate dehydrogenase. *Eur J Biochem.* **150**:61-6.
- Braun, V., Mahren, S. et Ogierman, M. 2003. Regulation of the FecI-type ECF sigma factor by

- transmembrane signalling. Curr Opin Microbiol. 6:173-80.
- Brescia, C. C., Mikulecky, P. J., Feig, A. L. et Sledjeski, D. D. <u>2003</u>. Identification of the Hfq-binding site on DsrA RNA: Hfq binds without altering DsrA secondary structure. *RNA*. **9**:33-43.
- Brescia, C. C., Kaw, M. K. et Sledjeski, D. D. <u>2004</u>. The DNA binding protein H-NS binds to and alters the stability of RNA in vitro and in vivo. J Mol Biol. <u>339</u>:505-14.
- Briant, D. J., Hankins, J. S., Cook, M. A. et Mackie, G. A. 2003. The quaternary structure of RNase G from Escherichia coli. Mol Microbiol. 50:1381-90.
- Brink, M. F., Verbeet, M. P. et De Boer, H. A. 1995. Specialized ribosomes: highly specific translation in vivo of a single targetted mRNA species. *Gene*. 156:215-22.
- Brown, C. T. et Callan, C. G., Jr. <u>2004</u>. Evolutionary comparisons suggest many novel cAMP response protein binding sites in *Escherichia coli*. *Proc Natl Acad Sci U S A*. **101**:2404-9.
- Brown, N. L., Stoyanov, J. V., Kidd, S. P. et Hobman, J. L. <u>2003</u>. The MerR family of transcriptional regulators. *FEMS Microbiol Rev.* 27:145-63.
- Browning, D. F. et Busby, S. J. <u>2004</u>. The regulation of bacterial transcription initiation. *Nat Rev Microbiol*. 2:57-65.
- Bruckner, R. et Titgemeyer, F. <u>2002</u>. Carbon catabolite repression in bacteria: choice of the carbon source and autoregulatory limitation of sugar utilization. *FEMS Microbiol Lett.* **209**:141-8.
- Buck, M. et Cannon, W. <u>1992</u>. Specific binding of the transcription factor σ<sup>54</sup> to promoter DNA. *Nature*. 358:422-4.
- Buck, M., Gallegos, M. T., Studholme, D. J., Guo, Y. et Gralla, J. D. <u>2000</u>. The bacterial enhancer-dependent σ<sup>54</sup> (σ<sup>N</sup>) transcription factor. *J Bacteriol*. **182**:4129-36.
- Buckle, M., Pemberton, I. K., Jacquet, M. A. et Buc, H. 1999. The kinetics of sigma subunit directed promoter recognition by E. coli RNA polymerase. J Mol Biol. 285:955-64.
- Busby, S. et Ebright, R. H. <u>1994</u>. Promoter structure, promoter recognition, and transcription activation in prokaryotes. *Cell.* **79**:743-6.
- Bycroft, M., Hubbard, T. J., Proctor, M., Freund, S. M. et Murzin, A. G. 1997. The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. Cell. 88:235-42.

#### C

- Cai, M., Williams, D. C., Jr., Wang, G., Lee, B. R., Peterkofsky, A. et Clore, G. M. <u>2003</u>. Solution structure of the phosphoryl transfer complex between the signal-transducing protein IIA<sup>Glucose</sup> and the cytoplasmic domain of the glucose transporter IICB<sup>Glucose</sup> of the *Escherichia coli* glucose phosphotransferase system. *J Biol Chem*. **278**:25191-206.
- Caillet, J., Nogueira, T., Masquida, B., Winter, F., Graffe, M., Dock-Bregeon, A. C., Torres-Larios, A., Sankaranarayanan, R., Westhof, E., Ehresmann, B.,

- Ehresmann, C., Romby, P. et Springer, M. <u>2003</u>. The modular structure of *Escherichia coli* threonyltRNA synthetase as both an enzyme and a regulator of gene expression. *Mol Microbiol.* **47**:961-74.
- Cairrao, F., Chora, A., Zilhao, R., Carpousis, A. J. et Arraiano, C. M. 2001. RNase II levels change according to the growth conditions: characterization of gmr, a new Escherichia coli gene involved in the modulation of RNase II. Mol Microbiol. 39:1550-61
- Calin-Jageman, I. et Nicholson, A. W. <u>2003a.</u> Mutational analysis of an RNA internal loop as a reactivity epitope for *Escherichia coli* ribonuclease III substrates. *Biochemistry*. **42**:5025-34.
- Calin-Jageman, I. et Nicholson, A. W. <u>2003b.</u> RNA structure-dependent uncoupling of substrate recognition and cleavage by *Escherichia coli* ribonuclease III. *Nucleic Acids Res.* **31**:2381-92.
- Callaghan, A. J., Aurikko, J. P., Ilag, L. L., Gunter Grossmann, J., Chandran, V., Kuhnel, K., Poljak, L., Carpousis, A. J., Robinson, C. V., Symmons, M. F. et Luisi, B. F. <u>2004</u>. Studies of the RNA degradosome-organizing domain of the *Escherichia* coli ribonuclease RNase E. J Mol Biol. **340**:965-79.
- Callaghan, A. J., Redko, Y., Murphy, L. M., Grossmann, J. G., Yates, D., Garman, E., Ilag, L. L., Robinson, C. V., Symmons, M. F., Mcdowall, K. J. et Luisi, B. F. 2005. "Zn-Link": A Metal-Sharing Interface that Organizes the Quaternary Structure and Catalytic Site of the Endoribonuclease, RNase E. Biochemistry. 44:4667-75.
- Camarero, J. A., Shekhtman, A., Campbell, E. A., Chlenov, M., Gruber, T. M., Bryant, D. A., Darst, S. A., Cowburn, D. et Muir, T. W. 2002. Autoregulation of a bacterial σ factor explored by using segmental isotopic labeling and NMR. Proc Natl Acad Sci U S A. 99:8536-41.
- Campbell, E. A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A. et Darst, S. A. <u>2001</u>. Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell.* **104**:901-12.
- Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., Trester-Zedlitz, M. L. et Darst, S. A. 2002a. Structure of the bacterial RNA polymerase promoter specificity σ subunit. *Mol Cell*. 9:527-39.
- Campbell, F. E., Jr., Cassano, A. G., Anderson, V. E. et Harris, M. E. <u>2002b</u>. Pre-steady-state and stoppedflow fluorescence analysis of *Escherichia coli* ribonuclease III: insights into mechanism and conformational changes associated with binding and catalysis. *J Mol Biol.* **317**:21-40.
- Cannistraro, V. J. et Kennell, D. <u>1989</u>. Purification and characterization of ribonuclease M and mRNA degradation in *Escherichia coli*. *Eur J Biochem*. <u>181</u>:363-70.
- Cannistraro, V. J. et Kennell, D. <u>1993</u>. The 5' ends of RNA oligonucleotides in *Escherichia coli* and mRNA degradation. *Eur J Biochem.* **213**:285-93.
- Cao, G. J. et Sarkar, N. <u>1992a</u>. Identification of the gene for an *Escherichia coli* poly(A) polymerase. *Proc Natl Acad Sci U S A*. **89**:10380-4.
- Cao, G. J. et Sarkar, N. 1992b. Poly(A) RNA in Escherichia coli: nucleotide sequence at the junction

- of the *lpp* transcript and the polyadenylate moiety. *Proc Natl Acad Sci U S A.* **89**:7546-50.
- Cao, G. J., Pogliano, J. et Sarkar, N. <u>1996</u>. Identification of the coding region for a second poly(A) polymerase in *Escherichia coli*. *Proc Natl Acad Sci* USA. **93**:11580-5.
- Cao, G. J., Kalapos, M. P. et Sarkar, N. 1997. Polyadenylated mRNA in Escherichia coli: modulation of poly(A) RNA levels by polynucleotide phosphorylase and ribonuclease II. Biochimie. 79:211-20.
- Carmell, M. A. et Hannon, G. J. <u>2004</u>. RNase III enzymes and the initiation of gene silencing. *Nat Struct Mol Biol.* 11:214-8.
- Carpousis, A. J., Vanzo, N. F. et Raynal, L. C. 1999. mRNA degradation. A tale of poly(A) and multiprotein machines. Trends Genet. 15:24-8.
- Casaregola, S., Jacq, A., Laoudj, D., Mcgurk, G., Margarson, S., Tempete, M., Norris, V. et Holland, I. B. 1992. Cloning and analysis of the entire Escherichia coli ams gene. ams is identical to hmp1 and encodes a 114 kDa protein that migrates as a 180 kDa protein. J Mol Biol. 228:30-40.
- Casaregola, S., Jacq, A., Laoudj, D., Mcgurk, G., Margarson, S., Tempete, M., Norris, V. et Holland, I. B. 1994. Cloning and analysis of the entire Escherichia coli ams gene. ams is identical to hmp1 and encodes a 114 kDa protein that migrates as a 180 kDa protein. J Mol Biol. 238:867.
- Charpentier, B. et Branlant, C. <u>1994</u>. The *Escherichia coli gapA* gene is transcribed by the vegetative RNA polymerase holoenzyme Eσ<sup>70</sup> and by the heat shock RNA polymerase Eσ<sup>32</sup>. *J. Bacteriol.* **176**:830-9.
- Charpentier, B. 1994. Etude de promoteurs transcriptionnels des gènes gapA et gapB codant la Glycéraldéhyde-3-phosphate déshydrogénase d'Escherichia coli. Relation structure/fonction. Thèse, Université Henri Poincaré.
- Charpentier, B., Bardey, V., Robas, N. et Branlant, C. 1998. The EII<sup>Glc</sup> protein is involved in glucosemediated activation of *Escherichia coli gapA* and gapB-pgk transcription. J Bacteriol. 180:6476-83.
- Chen, C. Y., Gherzi, R., Ong, S. E., Chan, E. L., Raijmakers, R., Pruijn, G. J., Stoecklin, G., Moroni, C., Mann, M. et Karin, M. <u>2001</u>. AU binding proteins recruit the exosome to degrade AREcontaining mRNAs. *Cell.* **107**:451-64.
- Chen, H., Bjerknes, M., Kumar, R. et Jay, E. 1994. Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs. *Nucleic Acids Res.* 22:4953-7.
- Chen, L. H., Emory, S. A., Bricker, A. L., Bouvet, P. et Belasco, J. G. 1991. Structure and function of a bacterial mRNA stabilizer: analysis of the 5' untranslated region of *ompA* mRNA. *J Bacteriol*. 173:4578-86.
- Cheng, Z. F., Zuo, Y., Li, Z., Rudd, K. E. et Deutscher, M. P. 1998. The vacB gene required for virulence in Shigella flexneri and Escherichia coli encodes the exoribonuclease RNase R. J Biol Chem. 273:14077-80
- Cheng, Z. F. et Deutscher, M. P. <u>2002</u>. Purification and characterization of the *Escherichia coli*

- exoribonuclease RNase R. Comparison with RNase II. *J Biol Chem.* **277**:21624-9.
- Cheng, Z. F. et Deutscher, M. P. <u>2005</u>. An important role for RNase R in mRNA decay. *Mol Cell*. **17**:313-8.
- Chin, A. M., Feldheim, D. A. et Saier, M. H., Jr. 1989. Altered transcriptional patterns affecting several metabolic pathways in strains of Salmonella typhimurium which overexpress the fructose regulon. J Bacteriol. 171:2424-34.
- Cho, S., Shin, D., Ji, G. E., Heu, S. et Ryu, S. 2005. High-level recombinant protein production by overexpression of Mlc in *Escherichia coli*. J Biotechnol.
- Chou, C. H., Bennett, G. N. et San, K. Y. <u>1994</u>. Effect of modified glucose uptake using genetic engineering techniques on high-level recombinant protein production in *Escherichia coli* dense cultures. *Biotechnol Bioeng.* **44**:952-60.
- Cisneros, B., Court, D., Sanchez, A. et Montanez, C. 1996. Point mutations in a transcription terminator, lambda tI, that affect both transcription termination and RNA stability. Gene. 181:127-33.
- Clermont, S., Corbier, C., Mely, Y., Gerard, D., Wonacott, A. et Branlant, G. 1993. Determinants of coenzyme specificity in glyceraldehyde-3-phosphate dehydrogenase: role of the acidic residue in the fingerprint region of the nucleotide binding fold. Biochemistry. 32:10178-84.
- 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. 279:1061-74
- Coburn, G. A. et Mackie, G. A. <u>1999.</u> Degradation of mRNA in *Escherichia coli*: an old problem with some new twists. *Prog Nucleic Acid Res Mol Biol.* **62**:55-108.
- Cohen, S. N., Chang, A. C. et Hsu, L. <u>1972</u>. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc. Natl. Acad. Sci. USA.* **69**:2110-4.
- Cohen, S. N. <u>1995.</u> Surprises at the 3' end of prokaryotic RNA. *Cell.* **80**:829-32.
- Cohen, S. N. et Mcdowall, K. J. <u>1997.</u> RNase E: still a wonderfully mysterious enzyme. *Mol Microbiol.* 23:1099-106.
- Cole, J. R., Chai, B., Farris, R. J., Wang, Q., Kulam, S. A., Mcgarrell, D. M., Garrity, G. M. et Tiedje, J. M. 2005. The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. *Nucleic Acids Res.* 33 Database Issue:D294-6.
- Collinet, B., Yuzawa, H., Chen, T., Herrera, C. et Missiakas, D. <u>2000</u>. RseB binding to the periplasmic domain of RseA modulates the RseA:σ<sup>E</sup> interaction in the cytoplasm and the availability of σ<sup>E</sup>.RNA polymerase. *J Biol Chem.* **275**:33898-904.
- Collis, C. M., Molloy, P. L., Both, G. W. et Drew, H. R. 1989. Influence of the sequence-dependent flexure of DNA on transcription in E. coli. Nucleic Acids Res. 17:9447-68.
- Comer, M. M., Dondon, J., Graffe, M., Yarchuk, O. et Springer, M. <u>1996</u>. Growth rate-dependent control, feedback regulation and steady-state mRNA levels

- of the threonyl-tRNA synthetase gene of Escherichia coli. J Mol Biol. 261:108-24.
- Condon, C., Squires, C. et Squires, C. L. <u>1995</u>. Control of rRNA transcription in *Escherichia coli*. *Microbiol Rev.* **59**:623-45.
- Condon, C., Brechemier-Baey, D., Beltchev, B., Grunberg-Manago, M. et Putzer, H. 2001. Identification of the gene encoding the 5S ribosomal RNA maturase in *Bacillus subtilis*: mature 5S rRNA is dispensable for ribosome function. *RNA*. 7:242-53.
- Condon, C. et Putzer, H. <u>2002</u>. The phylogenetic distribution of bacterial ribonucleases. *Nucleic Acids Res.* **30**:5339-46.
- Conrad, C. et Rauhut, R. <u>2002.</u> Ribonuclease III: new sense from nuisance. *Int J Biochem Cell Biol.* **34**:116-29.
- Contiero, J., Beatty, C., Kumari, S., Desanti, C. L., Strohl, W. R. et Wolfe, A. <u>2000</u>. Effects of mutations in acetate metabolism on high-celldensity growth of *Escherichia coli*. *Journal of Industrial Microbiology and Biotechnology*. **24**:421-30.
- Cooper, S. J., Leonard, G. A., Mcsweeney, S. M., Thompson, A. W., Naismith, J. H., Qamar, S., Plater, A., Berry, A. et Hunter, W. N. 1996. The crystal structure of a class II fructose-1,6bisphosphate aldolase shows a novel binuclear metal-binding active site embedded in a familiar fold. Structure. 4:1303-15.
- Corbier, C., Clermont, S., Billard, P., Skarzynski, T., Branlant, C., Wonacott, A. et Branlant, G. 1990. Probing the coenzyme specificity of glyceraldehyde-3-phosphate dehydrogenases by site-directed mutagenesis. *Biochemistry*. 29:7101-6.
- Cormack, R. S. et Mackie, G. A. 1992. Structural requirements for the processing of *Escherichia coli* 5 S ribosomal RNA by RNase E *in vitro*. *J. Mol. Biol.* 228:1078-90.
- Cortay, J. C., Negre, D., Scarabel, M., Ramseier, T. M., Vartak, N. B., Reizer, J., Saier, M. H., Jr. et Cozzone, A. J. 1994. In vitro asymmetric binding of the pleiotropic regulatory protein, FruR, to the ace operator controlling glyoxylate shunt enzyme synthesis. J Biol Chem. 269:14885-91.
- Cowing, D. W., Bardwell, J. C., Craig, E. A., Woolford, C., Hendrix, R. W. et Gross, C. A. <u>1985</u>. Consensus sequence for *Escherichia coli* heat shock gene promoters. *Proc Natl Acad Sci U S A*. **82**:2679-83.
- Crasnier, M., Dumay, V. et Danchin, A. <u>1994.</u> The catalytic domain of *Escherichia coli* K-12 adenylate cyclase as revealed by deletion analysis of the *cya* gene. *Mol Gen Genet.* **243**:409-16.
- Crasnier-Mednansky, M., Park, M. C., Studley, W. K. et Saier, M. H., Jr. 1997. Cra-mediated regulation of *Escherichia coli* adenylate cyclase. *Microbiology*. 143 ( Pt 3):785-92.
- Cunningham, L. et Guest, J. R. 1998. Transcription and transcript processing in the *sdhCDAB-sucABCD* operon of *Escherichia coli*. *Microbiology*. **144 ( Pt 8)**:2113-23.
- Curtis, S. J. et Epstein, W. <u>1975.</u> Phosphorylation of Dglucose in *Escherichia coli* mutants defective in glucosephosphotransferase,

mannosephosphotransferase, and glucokinase. *J Bacteriol.* **122**:1189-99.

# D

- Dam Mikkelsen, N. et Gerdes, K. <u>1997.</u> Sok antisense RNA from plasmid R1 is functionally inactivated by RNase E and polyadenylated by poly(A) polymerase I. *Mol Microbiol.* **26**:311-20.
- Danese, P. N. et Silhavy, T. J. <u>1998</u>. CpxP, a stress-combative member of the Cpx regulon. *J Bacteriol*. **180**:831-9.
- Darst, S. A. <u>2001</u>. Bacterial RNA polymerase. *Curr Opin Struct Biol*. **11**:155-62.
- Darst, S. A., Opalka, N., Chacon, P., Polyakov, A., Richter, C., Zhang, G. et Wriggers, W. 2002. Conformational flexibility of bacterial RNA polymerase. *Proc Natl Acad Sci U S A.* 99:4296-301.
- Dartigalongue, C., Missiakas, D. et Raina, S. <u>2001</u>. Characterization of the *Escherichia coli* σ<sup>E</sup> regulon. *J Biol Chem.* **276**:20866-75.
- D'aubenton Carafa, Y., Brody, E. et Thermes, C. <u>1990</u>. Prediction of rho-independent *Escherichia coli* transcription terminators. A statistical analysis of their RNA stem-loop structures. *J Mol Biol*. **216**:835-58.
- De Reuse, H. et Danchin, A. <u>1988.</u> The *ptsH*, *ptsI*, and *crr* genes of the *Escherichia coli* phosphoenolpyruvate-dependent phosphotransferase system: a complex operon with several modes of transcription. *J Bacteriol.* **170**:3827-37.
- De Reuse, H. et Danchin, A. <u>1991</u>. Positive regulation of the *pts* operon of *Escherichia coli*: genetic evidence for a signal transduction mechanism. *J Bacteriol*. 173:727-33.
- De Reuse, H., Kolb, A. et Danchin, A. <u>1992</u>. Positive regulation of the expression of the *Escherichia coli* pts operon. Identification of the regulatory regions. J Mol Biol. **226**:623-35.
- De Smit, M. H. et Van Duin, J. 1990. Control of prokaryotic translational initiation by mRNA secondary structure. Prog Nucleic Acid Res Mol Biol. 38:1-35.
- De Smit, M. H. et Van Duin, J. <u>1993</u>. Translational initiation at the coat-protein gene of phage MS2: native upstream RNA relieves inhibition by local secondary structure. *Mol Microbiol*. 9:1079-88.
- De Smit, M. H. et Van Duin, J. <u>1994a.</u> Control of translation by mRNA secondary structure in *Escherichia coli*. A quantitative analysis of literature data. *J Mol Biol.* **244**:144-50.
- De Smit, M. H. et Van Duin, J. <u>1994b.</u> Translational initiation on structured messengers. Another role for the Shine-Dalgarno interaction. *J Mol Biol.* **235**:173-84.
- De Smit, M. H. et Van Duin, J. 2003. Translational standby sites: how ribosomes may deal with the rapid folding kinetics of mRNA. *J Mol Biol.* 331:737-43.
- Deana, A. et Belasco, J. G. <u>2004</u>. The function of RNase G in *Escherichia coli* is constrained by its amino and carboxyl termini. *Mol Microbiol*. **51**:1205-17.

- Decker, K., Plumbridge, J. et Boos, W. <u>1998.</u> Negative transcriptional regulation of a positive regulator: the expression of *malT*, encoding the transcriptional activator of the maltose regulon of *Escherichia coli*, is negatively controlled by Mlc. *Mol Microbiol.* **27**:381-90.
- Dehaseth, P. L., Zupancic, M. L. et Record, M. T., Jr. 1998. RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. J Bacteriol. 180:3019-25.
- Della-Seta, F., Boschi-Muller, S., Vignais, M. L. et Branlant, G. 1997. Characterization of *Escherichia coli* strains with *gapA* and *gapB* genes deleted. *J. Bacteriol.* 179:5218-21.
- Deutscher, M. P. et Reuven, N. B. 1991. Enzymatic basis for hydrolytic versus phosphorolytic mRNA degradation in *Escherichia coli* and *Bacillus subtilis*. *Proc Natl Acad Sci U S A.* 88:3277-80.
- Didierjean, C., Corbier, C., Fatih, M., Favier, F., Boschi-Muller, S., Branlant, G. et Aubry, A. 2003. Crystal structure of two ternary complexes of phosphorylating glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* with NAD and D-glyceraldehyde 3-phosphate. *J Biol Chem.* 278:12968-76.
- Dimitrova, M. N., Szczepanowski, R. H., Ruvinov, S. B., Peterkofsky, A. et Ginsburg, A. <u>2002</u>. Interdomain interaction and substrate coupling effects on dimerization and conformational stability of enzyme I of the *Escherichia coli* phosphoenolpyruvate:sugar phosphotransferase system. *Biochemistry*. **41**:906-13.
- Diwa, A., Bricker, A. L., Jain, C. et Belasco, J. G. 2000. An evolutionarily conserved RNA stem-loop functions as a sensor that directs feedback regulation of RNase E gene expression. *Genes Dev.* 14:1249-
- Diwa, A. A. et Belasco, J. G. <u>2002</u>. Critical features of a conserved RNA stem-loop important for feedback regulation of RNase E synthesis. *J Biol Chem*. **277**:20415-22.
- Diwa, A. A., Jiang, X., Schapira, M. et Belasco, J. G. 2002. Two distinct regions on the surface of an RNA-binding domain are crucial for RNase E function. *Mol Microbiol.* 46:959-69.
- Donovan, W. P. et Kushner, S. R. <u>1986</u>. Polynucleotide phosphorylase and ribonuclease II are required for cell viability and mRNA turnover in *Escherichia* coli K-12. *Proc Natl Acad Sci U S A*. 83:120-4.
- Dorman, C. J. <u>2004.</u> H-NS: a universal regulator for a dynamic genome. *Nat Rev Microbiol.* **2**:391-400.
- Dove, S. L., Darst, S. A. et Hochschild, A. <u>2003.</u> Region 4 of σ as a target for transcription regulation. *Mol Microbiol.* **48**:863-74.
- Draper, D.E. <u>1996</u>. Translation Initiation In: Neidhart F., Curtiss R., Ingraham J.L., Lin E.C.C., Low K. B., Magasanik B. eds. *Escherichia coli* and *Salmonella*: cellular and molecular biology 2nd edn. Washington DC: American Society for Microbiology press. 1:902-908.
- Dubey, A. K., Baker, C. S., Suzuki, K., Jones, A. D., Pandit, P., Romeo, T. et Babitzke, P. <u>2003.</u> CsrA regulates translation of the *Escherichia coli* carbon

- starvation gene, cstA, by blocking ribosome access to the cstA transcript. J Bacteriol. 185:4450-60.
- Duee, E., Olivier-Deyris, L., Fanchon, E., Corbier, C., Branlant, G. et Dideberg, O. <u>1996</u>. Comparison of the structures of wild-type and a N313T mutant of *Escherichia coli* glyceraldehyde 3-phosphate dehydrogenases: implication for NAD binding and cooperativity. *J Mol Biol*. 257:814-38.
- Dumay, V., Danchin, A. et Crasnier, M. 1996. Regulation of Escherichia coli adenylate cyclase activity during hexose phosphate transport. Microbiology. 142 ( Pt 3):575-83.

# Ε

- Edwards, J. S., Ibarra, R. U. et Palsson, B. O. <u>2001</u>. *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat Biotechnol.* **19**:125-30.
- Egan, S. E., Fliege, R., Tong, S., Shibata, A., Wolf, R. E., Jr. et Conway, T. 1992. Molecular characterization of the Entner-Doudoroff pathway in *Escherichia coli*: sequence analysis and localization of promoters for the *edd-eda* operon. *J Bacteriol*. 174:4638-46.
- Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J. P. et Ehresmann, B. <u>1987.</u> Probing the structure of RNAs in solution. *Nucleic Acids Res.* **15**:9109-28.
- 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.* 6:149-59.
- Eikmanns, B. J. 1992. Identification, sequence analysis, and expression of a *Corynebacterium glutamicum* gene cluster encoding the three glycolytic enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate kinase, and triosephosphate isomerase. *J. Bacteriol.* 174:6076-86.
- Ellinger, T., Behnke, D., Knaus, R., Bujard, H. et Gralla, J. D. 1994. Context-dependent effects of upstream A-tracts. Stimulation or inhibition of *Escherichia* coli promoter function. J Mol Biol. 239:466-75.
- El-Mansi, M. 2004. Flux to acetate and lactate excretions in industrial fermentations: physiological and biochemical implications. *J Ind Microbiol Biotechnol.* 31:295-300.
- Emory, S. A., Bouvet, P. et Belasco, J. G. <u>1992.</u> A 5'-terminal stem-loop structure can stabilize mRNA in *Escherichia coli. Genes Dev.* **6**:135-48.
- Enz, S., Braun, V. et Crosa, J. H. 1995. Transcription of the region encoding the ferric dicitrate-transport system in *Escherichia coli*: similarity between promoters for *fecA* and for extracytoplasmic function sigma factors. *Gene.* 163:13-8.
- Eppler, T., Postma, P., Schutz, A., Volker, U. et Boos, W. 2002. Glycerol-3-phosphate-induced catabolite repression in *Escherichia coli*. J Bacteriol. 184:3044-52.
- Erickson, J. W. et Gross, C. A. <u>1989</u>. Identification of the sigma E subunit of *Escherichia coli* RNA polymerase: a second alternate sigma factor

- involved in high-temperature gene expression. Genes Dev. 3:1462-71.
- Espeli, O., Moulin, L. et Boccard, F. <u>2001</u>. Transcription attenuation associated with bacterial repetitive extragenic BIME elements. *J Mol Biol*. **314**:375-86.
- Estrem, S. T., Ross, W., Gaal, T., Chen, Z. W., Niu, W., Ebright, R. H. et Gourse, R. L. <u>1999</u>. Bacterial promoter architecture: subsite structure of UP elements and interactions with the carboxy-terminal domain of the RNA polymerase alpha subunit. *Genes Dev.* **13**:2134-47.

# F

- Fargo, D. C., Boynton, J. E. et Gillham, N. W. 1999. Mutations altering the predicted secondary structure of a chloroplast 5' untranslated region affect its physical and biochemical properties as well as its ability to promote translation of reporter mRNAs both in the Chlamydomonas reinhardtii chloroplast and in Escherichia coli. Mol Cell Biol. 19:6980-90.
- Farmer, W. R. et Liao, J. C. 1997. Reduction of aerobic acetate production by Escherichia coli. Appl Environ Microbiol. 63:3205-10.
- Farwell, M. A., Roberts, M. W. et Rabinowitz, J. C. 1992. The effect of ribosomal protein S1 from Escherichia coli and Micrococcus luteus on protein synthesis in vitro by E. coli and Bacillus subtilis. Mol Microbiol. 6:3375-83.
- Feng, Y. et Cohen, S. N. <u>2000</u>. Unpaired terminal nucleotides and 5' monophosphorylation govern 3' polyadenylation by *Escherichia coli* poly(A) polymerase I. *Proc. Natl. Acad. Sci. U S A.* **97**:6415-20.
- Feng, Y., Huang, H., Liao, J. et Cohen, S. N. 2001. Escherichia coli poly(A)-binding proteins that interact with components of degradosomes or impede RNA decay mediated by polynucleotide phosphorylase and RNase E. J Biol Chem. 276:31651-6.
- Feng, Y., Vickers, T. A. et Cohen, S. N. 2002. The catalytic domain of RNase E shows inherent 3' to 5' directionality in cleavage site selection. *Proc Natl* Acad Sci U S A. 99:14746-51.
- Fenton, M. S., Lee, S. J. et Gralla, J. D. 2000. Escherichia coli promoter opening and -10 recognition: mutational analysis of σ<sup>70</sup>. EMBO J. 19:1130-7.
- Ferenci, T. 1996. Adaptation to life at micromolar nutrient levels: the regulation of *Escherichia coli* glucose transport by endoinduction and cAMP. *FEMS Microbiol Rev.* 18:301-17.
- Ferenci, T. <u>2001</u>. Hungry bacteria--definition and properties of a nutritional state. *Environ Microbiol*. 3:605-11.
- Fillinger, S., Boschi-Muller, S., Azza, S., Dervyn, E., Branlant, G. et Aymerich, S. <u>2000</u>. Two glyceraldehyde-3-phosphate dehydrogenases with opposite physiological roles in a nonphotosynthetic bacterium. *J Biol Chem.* **275**:14031-7.
- Folichon, M., Arluison, V., Pellegrini, O., Huntzinger, E., Regnier, P. et Hajnsdorf, E. 2003. The poly(A) binding protein Hfq protects RNA from RNase E

- and exoribonucleolytic degradation. *Nucleic Acids Res.* 31:7302-10.
- Folichon, M., Allemand, F., Regnier, P. et Hajnsdorf, E. 2005. Stimulation of poly(A) synthesis by *Escherichia coli* poly(A)polymerase I is correlated with Hfq binding to poly(A) tails. *Febs J.* 272:454-63.
- Fomenkov, A., Valiakhmetov, A., Brand, L. et Roseman, S. 1998. In vivo and in vitro complementation of the N-terminal domain of enzyme I of the Escherichia coli phosphotransferase system by the cloned Cterminal domain. Proc Natl Acad Sci U S A. 95:8491-5.
- Foussard, M., Cabantous, S., Pedelacq, J., Guillet, V., Tranier, S., Mourey, L., Birck, C. et Samama, J. 2001. The molecular puzzle of two-component signaling cascades. *Microbes Infect*. 3:417-24.
- Fox, D. K., Presper, K. A., Adhya, S., Roseman, S. et Garges, S. 1992. Evidence for two promoters upstream of the pts operon: regulation by the cAMP receptor protein regulatory complex. Proc. Natl. Acad. Sci. USA. 89:7056-9.
- Franze De Fernandez, M. T., Eoyang, L. et August, J. T. 1968. Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature*. 219:588-90.

# G

- Gaal, T., Bartlett, M. S., Ross, W., Turnbough, C. L., Jr. et Gourse, R. L. <u>1997</u>. Transcription regulation by initiating NTP concentration: rRNA synthesis in bacteria. *Science*. **278**:2092-7.
- Gaal, T., Ross, W., Estrem, S. T., Nguyen, L. H., Burgess, R. R. et Gourse, R. L. <u>2001</u>. Promoter recognition and discrimination by Eo<sup>S</sup> RNA polymerase. *Mol Microbiol*. **42**:939-54.
- Galperin, M. Y., Aravind, L. et Koonin, E. V. 2000. Aldolases of the DhnA family: a possible solution to the problem of pentose and hexose biosynthesis in archaea. FEMS Microbiol Lett. 183:259-64.
- Ganoza, M. C., Kiel, M. C. et Aoki, H. <u>2002</u>. Evolutionary conservation of reactions in translation. *Microbiol Mol Biol Rev.* **66**:460-85, table of contents.
- Garcia-Alles, L. F. et Erni, B. <u>2002</u>. Synthesis of phosphoenol pyruvate (PEP) analogues and evaluation as inhibitors of PEP-utilizing enzymes. *Eur J Biochem*. **269**:3226-36.
- Garcia-Alles, L. F., Zahn, A. et Erni, B. <u>2002</u>. Sugar recognition by the glucose and mannose permeases of *Escherichia coli*. Steady-state kinetics and inhibition studies. *Biochemistry*. **41**:10077-86.
- Gase, K., Gase, A., Schirmer, H. et Malke, H. 1996. Cloning, sequencing and functional overexpression of the *Streptococcus equisimilis* H46A *gapC* gene encoding a glyceraldehyde-3-phosphate dehydrogenase that also functions as a plasmin(ogen)-binding protein. Purification and biochemical characterization of the protein. *Eur. J. Biochem.* 239:42-51.
- Geerse, R. H., Van Der Pluijm, J. et Postma, P. W. <u>1989</u>.

  The repressor of the PEP:fructose phosphotransferase system is required for the

- transcription of the *pps* gene of *Escherichia coli*. *Mol Gen Genet.* **218**:348-52.
- Geissmann, T. A. et Touati, D. <u>2004</u>. Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *EMBO* J. <u>23</u>:396-405.
- Georgellis, D., Arvidson, S. et Von Gabain, A. 1992. Decay of *ompA* mRNA and processing of 9S RNA are immediately affected by shifts in growth rate, but in opposite manners. *J Bacteriol.* 174:5382-90.
- Georgellis, D., Barlow, T., Arvidson, S. et Von Gabain, A. 1993. Retarded RNA turnover in *Escherichia* coli: a means of maintaining gene expression during anaerobiosis. Mol Microbiol. 9:375-81.
- Georgellis, D., Kwon, O., De Wulf, P. et Lin, E. C. 1998. Signal decay through a reverse phosphorelay in the Arc two-component signal transduction system. J Biol Chem. 273:32864-9.
- Georgellis, D., Kwon, O. et Lin, E. C. 1999. Amplification of signaling activity of the arc two-component system of *Escherichia coli* by anaerobic metabolites. An *in vitro* study with different protein modules. *J Biol Chem.* 274:35950-4.
- Georgellis, D., Kwon, O. et Lin, E. C. <u>2001</u>. Quinones as the redox signal for the arc two-component system of bacteria. *Science*. **292**:2314-6.
- Ghora, B. K. et Apirion, D. <u>1978.</u> Structural analysis and *in vitro* processing to p5 rRNA of a 9S RNA molecule isolated from an *rne* mutant of *E. coli. Cell.* **15**:1055-66.
- Ghosh, S. et Deutscher, M. P. <u>1999</u>. Oligoribonuclease is an essential component of the mRNA decay pathway. *Proc Natl Acad Sci U S A*. **96**:4372-7.
- Ginsburg, A. et Peterkofsky, A. <u>2002.</u> Enzyme I: the gateway to the bacterial phosphoenolpyruvate:sugar phosphotransferase system. *Arch Biochem Biophys.* 397:273-8.
- Gitelman, D. R. et Apirion, D. <u>1980</u>. The synthesis of some proteins is affected in RNA processing mutants of *Escherichia coli*. *Biochem Biophys Res Commun.* **96**:1063-70.
- Gold, L. <u>1988.</u> Posttranscriptional regulatory mechanisms in *Escherichia coli. Annu Rev Biochem.* 57:199-233.
- Gonzalez-Gil, G., Kahmann, R. et Muskhelishvili, G. 1998. Regulation of *crp* transcription by oscillation between distinct nucleoprotein complexes. *EMBO J.* 17:2877-85.
- Goodrich, A. F. et Steege, D. A. 1999. Roles of polyadenylation and nucleolytic cleavage in the filamentous phage mRNA processing and decay pathways in *Escherichia coli. RNA*. 5:972-85.
- Gosset, G., Zhang, Z., Nayyar, S., Cuevas, W. A. et Saier, M. H., Jr. <u>2004</u>. Transcriptome analysis of Crp-dependent catabolite control of gene expression in *Escherichia coli*. *J Bacteriol*. **186**:3516-24.
- Gosset, G. <u>2005</u>. Improvement of *Escherichia coli* production strains by modification of the phosphoenolpyruvate:sugar phosphotransferase system. *Microb Cell Fact*. **4**:14.
- Gottesman, S. <u>2004.</u> The Small RNA Regulators of Escherichia coli: Roles and Mechanisms. Annu Rev Microbiol. 58:303-28.

- Gourse, R. L., Ross, W. et Gaal, T. <u>2000</u>. UPs and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition. *Mol Microbiol*. **37**:687-95.
- Gralla, J. D. <u>2005</u>. Escherichia coli ribosomal RNA transcription: regulatory roles for ppGpp, NTPs, architectural proteins and a polymerase-binding protein. Mol Microbiol. <u>55</u>:973-7.
- Groeneveld, H., Thimon, K. et Van Duin, J. 1995. Translational control of maturation-protein synthesis in phage MS2: a role for the kinetics of RNA folding? RNA. 1:79-88.
- Grossman, A. D., Erickson, J. W. et Gross, C. A. <u>1984</u>. The *htpR* gene product of *E. coli* is a sigma factor for heat-shock promoters. *Cell*. **38**:383-90.
- Gruber, T. M. et Gross, C. A. <u>2003.</u> Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol.* 57:441-66.
- Grunberg-Manago, M. 1999. Messenger RNA stability and its role in control of gene expression in bacteria and phages. Annu Rev Genet. 33:193-227.
- Gschaedler, A., Robas, N., Boudrant, J. et Branlant, C.

  1999. Effects of pulse addition of carbon sources on continuous cultivation of *Escherichia coli* containing a recombinant *E. coli gapA* gene.

  Biotechnol Bioeng. 63:712-20.
- Gualerzi, C. O. et Pon, C. L. <u>1990</u>. Initiation of mRNA translation in prokaryotes. *Biochemistry*. **29**:5881-9.
- Guenneugues, M., Caserta, E., Brandi, L., Spurio, R., Meunier, S., Pon, C. L., Boelens, R. et Gualerzi, C. O. <u>2000</u>. Mapping the fMet-tRNA<sub>f</sub> binding site of initiation factor IF2. *EMBO J.* **19**:5233-40.
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. et Altman, S. <u>1983</u>. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*. **35**:849-57.
- Guillier, M., Allemand, F., Raibaud, S., Dardel, F., Springer, M. et Chiaruttini, C. 2002. Translational feedback regulation of the gene for L35 in Escherichia coli requires binding of ribosomal protein L20 to two sites in its leader mRNA: a possible case of ribosomal RNA-messenger RNA molecular mimicry. RNA. 8:878-89.

# Н

- Hajnsdorf, E., Braun, F., Haugel-Nielsen, J. et Regnier, P. <u>1995</u>. Polyadenylylation destabilizes the *rpsO* mRNA of *Escherichia coli*. *Proc Natl Acad Sci U S* A. **92**:3973-7.
- Hajnsdorf, E. <u>1999</u>. La dégradation des ARNm chez Escherichia coli. Regard sur la Biochimie. 4:29-34.
- Hajnsdorf, E. et Regnier, P. 1999. E. coli RpsO mRNA decay: RNase E processing at the beginning of the coding sequence stimulates poly(A)-dependent degradation of the mRNA. J Mol Biol. 286:1033-43.
- Hajnsdorf, E. et Regnier, P. <u>2000</u>. Host factor Hfq of Escherichia coli stimulates elongation of poly(A) tails by poly(A) polymerase I. Proc Natl Acad Sci U S A. **97**:1501-5.
- Hanamura, A. et Aiba, H. 1992. A new aspect of transcriptional control of the Escherichia coli crp

- gene: positive autoregulation. *Mol Microbiol*. **6**:2489-97.
- Hannig, G. et Makrides, S. C. 1998. Strategies for optimizing heterologous protein expression in Escherichia coli. Trends Biotechnol. 16:54-60.
- Hansen, A., Pfeiffer, T., Zuleeg, T., Limmer, S., Ciesiolka, J., Feltens, R. et Hartmann, R. K. 2001. Exploring the minimal substrate requirements for trans-cleavage by RNase P holoenzymes from Escherichia coli and Bacillus subtilis. Mol Microbiol. 41:131-43.
- Harley, C. B. et Reynolds, R. P. <u>1987</u>. Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.* **15**:2343-61.
- Harris, R. M., Webb, D. C., Howitt, S. M. et Cox, G. B. 2001. Characterization of PitA and PitB from Escherichia coli. J Bacteriol. 183:5008-14.
- Hayes, C. S., Bose, B. et Sauer, R. T. <u>2002</u>. Stop codons preceded by rare arginine codons are efficient determinants of SsrA tagging in *Escherichia coli*. *Proc Natl Acad Sci U S A*. **99**:3440-5.
- He, W. et Parker, R. <u>2000</u>. Functions of Lsm proteins in mRNA degradation and splicing. *Curr Opin Cell Biol*. **12**:346-50.
- Helmann, J. D. et Chamberlin, M. J. <u>1987.</u> DNA sequence analysis suggests that expression of flagellar and chemotaxis genes in *Escherichia coli* and *Salmonella typhimurium* is controlled by an alternative sigma factor. *Proc. Natl. Acad. Sci. USA*. 84:6422-4.
- Helmann, J. D. et Chamberlin, M. J. <u>1988</u>. Structure and function of bacterial sigma factors. *Annu Rev Biochem.* **57**:839-72.
- Helmann, J. D. <u>2002</u>. The extracytoplasmic function (ECF) sigma factors. Adv Microb Physiol. **46**:47-110.
- Helmreich, E. J. <u>1992.</u> How pyridoxal 5'-phosphate could function in glycogen phosphorylase catalysis. *Biofactors*. 3:159-72.
- Hengge-Aronis, R. 2002a. Signal transduction and regulatory mechanisms involved in control of the σ<sup>S</sup> (RpoS) subunit of RNA polymerase. *Microbiol Mol Biol Rev.* 66:373-95, table of contents.
- Hengge-Aronis, R. <u>2002b.</u> Stationary phase gene regulation: what makes an *Escherichia coli* promoter σ<sup>S</sup>-selective? *Curr Opin Microbiol.* 5:591-
- Henze, K., Morrison, H. G., Sogin, M. L. et Mullerm. 1998. Sequence and phylogenetic position of a class II aldolase gene in the amitochondriate protist, Giardia lamblia. Gene. 222:163-8.
- Hershberg, R., Altuvia, S. et Margalit, H. <u>2003</u>. A survey of small RNA-encoding genes in *Escherichia coli*. *Nucleic Acids Res.* 31:1813-20.
- Higgins, C. F. <u>2001.</u> ABC transporters: physiology, structure and mechanism--an overview. *Res Microbiol.* **152**:205-10.
- Hillman, J. D. et Fraenkel, D. G. <u>1975.</u> Glyceraldehyde 3-phosphate dehydrogenase mutants of *Escherichia* coli. J. Bacteriol. **122**:1175-9.
- Hoch, J. A. et Varughese, K. I. <u>2001</u>. Keeping signals straight in phosphorelay signal transduction. *J Bacteriol*. **183**:4941-9.
- Hogema, B. M., Arents, J. C., Inada, T., Aiba, H., Van Dam, K. et Postma, P. W. 1997. Catabolite

- repression by glucose 6-phosphate, gluconate and lactose in *Escherichia coli*. *Mol Microbiol*. **24**:857-67
- Hogema, B. M., Arents, J. C., Bader, R., Eijkemans, K., Inada, T., Aiba, H. et Postma, P. W. 1998a. Inducer exclusion by glucose 6-phosphate in *Escherichia* coli. Mol Microbiol. 28:755-65.
- Hogema, B. M., Arents, J. C., Bader, R., Eijkemans, K., Yoshida, H., Takahashi, H., Aiba, H. et Postma, P. W. 1998b. Inducer exclusion in *Escherichia coli* by non-PTS substrates: the role of the PEP to pyruvate ratio in determining the phosphorylation state of enzyme IIA<sup>Gk</sup>. *Mol Microbiol.* 30:487-98.
- Hogg, T., Mechold, U., Malke, H., Cashel, M. et Hilgenfeld, R. 2004. Conformational antagonism between opposing active sites in a bifunctional RelA/SpoT homolog modulates (p)ppGpp metabolism during the stringent response [corrected]. Cell. 117:57-68.
- Hoischen, C., Levin, J., Pitaknarongphorn, S., Reizer, J. et Saier, M. H., Jr. 1996. Involvement of the central loop of the lactose permease of *Escherichia coli* in its allosteric regulation by the glucose-specific enzyme IIA of the phosphoenolpyruvate-dependent phosphotransferase system. *J Bacteriol.* 178:6082-6.
- Holtman, C. K., Pawlyk, A. C., Meadow, N. D. et Pettigrew, D. W. <u>2001a</u>. Reverse genetics of Escherichia coli glycerol kinase allosteric regulation and glucose control of glycerol utilization in vivo. J Bacteriol. **183**:3336-44.
- Holtman, C. K., Pawlyk, A. C., Meadow, N., Roseman, S. et Pettigrew, D. W. <u>2001b</u>. IIA<sup>Glc</sup> allosteric control of *Escherichia coli* glycerol kinase: binding site cooperative transitions and cation-promoted association by Zinc(II). *Biochemistry*. **40**:14302-8.
- Hosono, K., Kakuda, H. et Ichihara, S. <u>1995</u>. Decreasing accumulation of acetate in a rich medium by *Escherichia coli* on introduction of genes on a multicopy plasmid. *Biosci Biotechnol Biochem*. **59**:256-61.
- Hu, K. Y. et Saier, M. H., Jr. <u>2002</u>. Phylogeny of phosphoryl transfer proteins of the phosphoenolpyruvate-dependent sugar-transporting phosphotransferase system. *Res Microbiol*. **153**:405-15.
- Huang, H., Liao, J. et Cohen, S. N. <u>1998</u>. Poly(A)- and poly(U)-specific RNA 3' tail shortening by *E. coli* ribonuclease E. *Nature*. **391**:99-102.
- Hunt, T. P. et Magasanik, B. <u>1985</u>. Transcription of *glnA* by purified *Escherichia coli* components: core RNA polymerase and the products of *glnF*, *glnG*, and *glnL*. *Proc Natl Acad Sci U S A*. **82**:8453-7.
- Hurley, J. H., Faber, H. R., Worthylake, D., Meadow, N. D., Roseman, S., Pettigrew, D. W. et Remington, S. J. 1993. Structure of the regulatory complex of *Escherichia coli* III<sup>Glc</sup> with glycerol kinase. *Science*. 259:673-7.
- Huttenhofer, A. et Noller, H. F. <u>1994</u>. Footprinting mRNA-ribosome complexes with chemical probes. *EMBO J.* **13**:3892-901.

# Ι

- Inada, T., Takahashi, H., Mizuno, T. et Aiba, H. 1996. Down regulation of cAMP production by cAMP receptor protein in *Escherichia coli*: an assessment of the contributions of transcriptional and posttranscriptional control of adenylate cyclase. *Mol Gen Genet.* 253:198-204.
- Iost, I., Guillerez, J. et Dreyfus, M. <u>1992</u>. Bacteriophage T7 RNA polymerase travels far ahead of ribosomes in vivo. J Bacteriol. **174**:619-22.
- Iost, I. et Dreyfus, M. <u>1994</u>. mRNAs can be stabilized by DEAD-box proteins. *Nature*. **372**:193-6.
- Iost, I. et Dreyfus, M. 1995. The stability of Escherichia coli lacZ mRNA depends upon the simultaneity of its synthesis and translation. EMBO J. 14:3252-61.
- Ishihama, A. <u>1981.</u> Subunit of assembly of *Escherichia* coli RNA polymerase. *Adv Biophys.* **14**:1-35.
- Ishihama, A. <u>2000.</u> Functional modulation of Escherichia coli RNA polymerase. Annu Rev Microbiol. 54:499-518.
- Ishizuka, H., Hanamura, A., Kunimura, T. et Aiba, H. 1993. A lowered concentration of cAMP receptor protein caused by glucose is an important determinant for catabolite repression in *Escherichia* coli. Mol Microbiol. 10:341-50.
- Iuchi, S. <u>1993.</u> Phosphorylation/dephosphorylation of the receiver module at the conserved aspartate residue controls transphosphorylation activity of histidine kinase in sensor protein ArcB of *Escherichia coli. J Biol Chem.* **268**:23972-80.

# J

- Jacob, W. F., Santer, M. et Dahlberg, A. E. <u>1987</u>. A single base change in the Shine-Dalgarno region of 16S rRNA of *Escherichia coli* affects translation of many proteins. *Proc Natl Acad Sci U S A*. **84**:4757-61.
- Jager, S., Fuhrmann, O., Heck, C., Hebermehl, M., Schiltz, E., Rauhut, R. et Klug, G. <u>2001</u>. An mRNA degrading complex in *Rhodobacter capsulatus*. *Nucleic Acids Res.* **29**:4581-8.
- Jain, C., Deana, A. et Belasco, J. G. <u>2002</u>. Consequences of RNase E scarcity in *Escherichia coli*. *Mol Microbiol*. **43**:1053-64.
- Jarrige, A. C., Mathy, N. et Portier, C. <u>2001</u>. PNPase autocontrols its expression by degrading a doublestranded structure in the *pnp* mRNA leader. *EMBO* J. **20**:6845-55.
- Jasiecki, J. et Wegrzyn, G. 2003. Growth-rate dependent RNA polyadenylation in *Escherichia coli*. EMBO Rep. 4:172-7.
- Jasiecki, J. et Wegrzyn, G. 2005. Localization of Escherichia coli poly(A) polymerase I in cellular membrane. Biochem Biophys Res Commun. 329:598-602.
- Jauregui, R., Abreu-Goodger, C., Moreno-Hagelsieb, G., Collado-Vides, J. et Merino, E. <u>2003</u>. Conservation of DNA curvature signals in regulatory regions of prokaryotic genes. *Nucleic Acids Res.* 31:6770-7.
- Jenkins, D. E., Auger, E. A. et Matin, A. <u>1991.</u> Role of RpoH, a heat shock regulator protein, in *Escherichia*

- coli carbon starvation protein synthesis and survival. J Bacteriol. 173:1992-6.
- Jensen, E. B. et Carlsen, S. 1990. Production of recombinant human growth hormone in Escherichia coli: Expression of different precursors and physiological effects of glucose, ac etate, and salts. Biotechnol. Bioeng. 36:1-11.
- Jeon, Y., Lee, Y. S., Han, J. S., Kim, J. B. et Hwang, D. S. 2001. Multimerization of phosphorylated and non-phosphorylated ArcA is necessary for the response regulator function of the Arc two-component signal transduction system. *J Biol Chem.* 276:40873-9.
- Jeong, J. Y., Kim, Y. J., Cho, N., Shin, D., Nam, T. W., Ryu, S. et Seok, Y. J. <u>2004a</u>. Expression of *ptsG* encoding the major glucose transporter is regulated by ArcA in *Escherichia coli*. *J Biol Chem*. **279**:38513-8.
- Jeong, Y. J., Kim, D. E. et Patel, S. S. <u>2004b</u>. Nucleotide binding induces conformational changes in *Escherichia coli* transcription termination factor Rho. *J Biol Chem.* **279**:18370-6.
- Jiang, X., Diwa, A. et Belasco, J. G. <u>2000</u>. Regions of RNase E important for 5'-end-dependent RNA cleavage and autoregulated synthesis. *J Bacteriol*. **182**:2468-75.
- Jiang, X. et Belasco, J. G. 2004. Catalytic activation of multimeric RNase E and RNase G by 5'monophosphorylated RNA. Proc Natl Acad Sci U S A. 101:9211-6.
- Jin, H., Bjornsson, A. et Isaksson, L. A. <u>2002.</u> Cis control of gene expression in *E.coli* by ribosome queuing at an inefficient translational stop signal. *EMBO J.* **21**:4357-67.
- Jishage, M. et Ishihama, A. <u>1998</u>. A stationary phase protein in *Escherichia coli* with binding activity to the major σ subunit of RNA polymerase. *Proc Natl Acad Sci U S A*. **95**:4953-8.
- Jishage, M. et Ishihama, A. 1999. Transcriptional organization and in vivo role of the Escherichia coli rsd gene, encoding the regulator of RNA polymerase sigma D. J Bacteriol. 181:3768-76.
- Jishage, M., Kvint, K., Shingler, V. et Nystrom, T. 2002. Regulation of σ factor competition by the alarmone ppGpp. Genes Dev. 16:1260-70.
- Joyce, S. A. et Dreyfus, M. 1998. In the absence of translation, RNase E can bypass 5' mRNA stabilizers in *Escherichia coli*. J Mol Biol. 282:241-54.

# K

- Kaberdin, V. R., Chao, Y. H. et Lin-Chao, S. 1996.
  RNase E cleaves at multiple sites in bubble regions of RNA I stem loops yielding products that dissociate differentially from the enzyme. J Biol Chem. 271:13103-9.
- 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 U S A. 95:11637-42.
- Kaberdin, V. R., Walsh, A. P., Jakobsen, T., Mcdowall, K. J. et Von Gabain, A. <u>2000</u>. Enhanced cleavage of RNA mediated by an interaction between substrates and the arginine-rich domain of *E. coli* ribonuclease E. *J Mol Biol*. **301**:257-64.
- Kaberdin, V. R. <u>2003</u>. Probing the substrate specificity of *Escherichia coli* RNase E using a novel oligonucleotide-based assay. *Nucleic Acids Res.* 31:4710-6.
- Kaga, N., Umitsuki, G., Nagai, K. et Wachi, M. 2002. RNase G-dependent degradation of the eno mRNA encoding a glycolysis enzyme enolase in Escherichia coli. Biosci Biotechnol Biochem. 66:2216-20.
- Kajitani, M. et Ishihama, A. <u>1991</u>. Identification and sequence determination of the host factor gene for bacteriophage Q beta. *Nucleic Acids Res.* **19**:1063-6.
- Kajitani, M., Kato, A., Wada, A., Inokuchi, Y. et Ishihama, A. <u>1994.</u> Regulation of the *Escherichia* coli hfq gene encoding the host factor for phage Q beta. J Bacteriol. **176**:531-4.
- Kanaya, S., Yamada, Y., Kudo, Y. et Ikemura, T. 1999. Studies of codon usage and tRNA genes of 18 unicellular organisms and quantification of *Bacillus* subtilis tRNAs: gene expression level and speciesspecific diversity of codon usage based on multivariate analysis. Gene. 238:143-55.
- Kaplan, D. L. et O'donnell, M. <u>2003</u>. Rho factor: transcription termination in four steps. *Curr Biol*. 13:R714-6.
- Karimi, R., Pavlov, M. Y., Buckingham, R. H. et Ehrenberg, M. 1999. Novel roles for classical factors at the interface between translation termination and initiation. *Mol Cell*. 3:601-9.
- Karlin, S., Mrazek, J. et Campbell, A. M. <u>1998.</u> Codon usages in different gene classes of the *Escherichia* coli genome. *Mol Microbiol.* **29**:1341-55.
- Karlin, S., Mrazek, J., Campbell, A. et Kaiser, D. <u>2001</u>. Characterizations of highly expressed genes of four fast-growing bacteria. *J Bacteriol*. **183**:5025-40.
- Kawamoto, H., Morita, T., Shimizu, A., Inada, T. et Aiba, H. 2005. Implication of membrane localization of target mRNA in the action of a small RNA: mechanism of post-transcriptional regulation of glucose transporter in *Escherichia coli*. Genes Dev. 19:328-38.
- Keener et Nomura, 1996. Regulation of Ribosome Synthesis. In: Neidhart F., Curtiss R., Ingraham J.L., Lin E.C.C., Low K. B., Magasanik B. eds. Escherichia coli and Salmonella: cellular and molecular biology 2nd edn. Washington DC: American Society for Microbiology press. 1:1417-1431.
- Kennell, D. <u>2002.</u> Processing endoribonucleases and mRNA degradation in bacteria. *J Bacteriol*. **184**:4645-57; discussion 65.
- Keyhani, N. O., Bacia, K. et Roseman, S. <u>2000.</u> The transport/phosphorylation of *N,N'*-diacetylchitobiose in *Escherichia coli*. Characterization of phospho-IIB<sup>Chb</sup> and of a potential transition state analogue in the phosphotransfer reaction between the proteins IIA<sup>Chb</sup> AND IIB<sup>Chb</sup>. *J Biol Chem.* **275**:33102-9.

- Khemici, V. et Carpousis, A. J. <u>2004</u>. The RNA degradosome and poly(A) polymerase of Escherichia coli are required in vivo for the degradation of small mRNA decay intermediates containing REP-stabilizers. Mol Microbiol. <u>51</u>:777-90
- Khemici, V., Poljak, L., Toesca, I. et Carpousis, A. J. 2005. Evidence in vivo that the DEAD-box RNA helicase RhlB facilitates the degradation of ribosome-free mRNA by RNase E. Proc Natl Acad Sci USA. 102:6913-8.
- Kimata, K., Takahashi, H., Inada, T., Postma, P. et Aiba, H. 1997. cAMP receptor protein-cAMP plays a crucial role in glucose-lactose diauxie by activating the major glucose transporter gene in *Escherichia* coli. Proc Natl Acad Sci USA. 94:12914-9.
- Kimata, K., Inada, T., Tagami, H. et Aiba, H. <u>1998.</u> A global repressor (Mlc) is involved in glucose induction of the ptsG gene encoding major glucose transporter in Escherichia coli. Mol. Microbiol. <u>29</u>:1509-19.
- Kimata, K., Tanaka, Y., Inada, T. et Aiba, H. 2001. Expression of the glucose transporter gene, ptsG, is regulated at the mRNA degradation step in response to glycolytic flux in Escherichia coli. EMBO J. 20:3587-95.
- Koksharova, O., Schubert, M., Shestakov, S. et Cerff, R. 1998. Genetic and biochemical evidence for distinct key functions of two highly divergent GAPDH genes in catabolic and anabolic carbon flow of the cyanobacterium Synechocystis sp. PCC 6803. Plant. Mol. Biol. 36:183-94.
- Kolb, A., Kotlarz, D., Kusano, S. et Ishihama, A. 1995. Selectivity of the *Escherichia coli* RNA polymerase Eσ<sup>38</sup> for overlapping promoters and ability to support CRP activation. *Nucleic Acids Res.* 23:819-26
- Komarova, A. V., Tchufistova, L. S., Supina, E. V. et Boni, I. V. <u>2002</u>. Protein S1 counteracts the inhibitory effect of the extended Shine-Dalgarno sequence on translation. RNA. 8:1137-47.
- Komine, Y., Kitabatake, M., Yokogawa, T., Nishikawa, K. et Inokuchi, H. 1994. A tRNA-like structure is present in 10Sa RNA, a small stable RNA from Escherichia coli. Proc Natl Acad Sci U S A. 91:9223-7.
- Koo, B. M., Yoon, M. J., Lee, C. R., Nam, T. W., Choe, Y. J., Jaffe, H., Peterkofsky, A. et Seok, Y. J. 2004. A novel fermentation/respiration switch protein regulated by enzyme IIA<sup>Glc</sup> in *Escherichia coli*. J Biol Chem. 279:31613-21.
- Kopke, A. K. et Leggatt, P. A. 1991. Initiation of translation at an AUA codon for an archaebacterial protein gene expressed in E. coli. Nucleic Acids Res. 19:5169-72.
- Kozak, M. <u>1999</u>. Initiation of translation in prokaryotes and eukaryotes. *Gene*. **234**:187-208.
- Krin, E., Sismeiro, O., Danchin, A. et Bertin, P. N. 2002. The regulation of Enzyme IIA<sup>Glc</sup> expression controls adenylate cyclase activity in *Escherichia coli*. *Microbiology*. 148:1553-9.
- Kuo, J. T., Chang, Y. J. et Tseng, C. P. <u>2003.</u> Growth rate regulation of *lac* operon expression in

- Escherichia coli is cyclic AMP dependent. FEBS Lett. 553:397-402.
- Kushner, S. R. <u>2002.</u> mRNA decay in *Escherichia coli* comes of age. *J Bacteriol*. **184**:4658-65; discussion 7.
- Kuznedelov, K., Minakhin, L., Niedziela-Majka, A., Dove, S. L., Rogulja, D., Nickels, B. E., Hochschild, A., Heyduk, T. et Severinov, K. <u>2002</u>. A role for interaction of the RNA polymerase flap domain with the σ subunit in promoter recognition. *Science*. **295**:855-7.
- Kwon, O., Georgellis, D. et Lin, E. C. <u>2000</u>. Phosphorelay as the sole physiological route of signal transmission by the arc two-component system of *Escherichia coli*. *J Bacteriol*. **182**:3858-62.

#### L

- La Teana, A., Brandi, A., O'connor, M., Freddi, S. et Pon, C. L. <u>2000</u>. Translation during cold adaptation does not involve mRNA-rRNA base pairing through the downstream box. *RNA*. **6**:1393-402.
- Lamontagne, B. et Elela, S. A. <u>2004</u>. Evaluation of the RNA determinants for bacterial and yeast RNase III binding and cleavage. *J Biol Chem.* **279**:2231-41.
- Lange, R. et Hengge-Aronis, R. <u>1991</u>. Growth phase-regulated expression of *bolA* and morphology of stationary- phase *Escherichia coli* cells are controlled by the novel sigma factor sigma S. *J. Bacteriol.* **173**:4474-81.
- Lay, A. J., Jiang, X. M., Kisker, O., Flynn, E., Underwood, A., Condron, R. et Hogg, P. J. 2000. Phosphoglycerate kinase acts in tumour angiogenesis as a disulphide reductase. *Nature*. 408:869-73.
- Le Derout, J., Regnier, P. et Hajnsdorf, E. <u>2002</u>. Both temperature and medium composition regulate RNase E processing efficiency of the *rpsO* mRNA coding for ribosomal protein S15 of *Escherichia coli*. *J Mol Biol*. **319**:341-9.
- Lease, R. A. et Belfort, M. 2000. A trans-acting RNA as a control switch in *Escherichia coli*: DsrA modulates function by forming alternative structures. *Proc Natl Acad Sci U S A*. 97:9919-24.
- Lease, R. A. et Woodson, S. A. <u>2004.</u> Cycling of the Smlike protein Hfq on the DsrA small regulatory RNA. *J Mol Biol.* **344**:1211-23.
- Lee, K. et Cohen, S. N. <u>2001</u>. Effects of 3' terminus modifications on mRNA functional decay during in vitro protein synthesis. J Biol Chem. **276**:23268-74.
- Lee, K., Bernstein, J. A. et Cohen, S. N. <u>2002</u>. RNase G complementation of *rne* null mutation identifies functional interrelationships with RNase E in *Escherichia coli*. *Mol Microbiol*. **43**:1445-56.
- Lee, K. et Cohen, S. N. <u>2003</u>. A Streptomyces coelicolor functional orthologue of Escherichia coli RNase E shows shuffling of catalytic and PNPase-binding domains. *Mol Microbiol*. **48**:
- 349-60.
- Lee, K., Zhan, X., Gao, J., Qiu, J., Feng, Y., Meganathan, R., Cohen, S. N. et Georgiou, G. 2003. RraA. a protein inhibitor of RNase E activity that

- globally modulates RNA abundance in *E. coli. Cell.* 114:623-34.
- Lee, S. Y. 1996. High cell-density culture of Escherichia coli. Trends Biotechnol. 14:98-105.
- Lendenmann, U. et Egli, T. 1995. Is Escherichia coli growing in glucose-limited chemostat culture able to utilize other sugars without lag? Microbiology. 141 ( Pt 1):71-8.
- Lesage, P., Chiaruttini, C., Graffe, M., Dondon, J., Milet, M. et Springer, M. 1992. Messenger RNA secondary structure and translational coupling in the *Escherichia coli* operon encoding translation initiation factor IF3 and the ribosomal proteins, L35 and L20. *J Mol Biol.* 228:366-86.
- Lesnik, E. A., Sampath, R., Levene, H. B., Henderson, T. J., Mcneil, J. A. et Ecker, D. J. <u>2001</u>. Prediction of rho-independent transcriptional terminators in *Escherichia coli*. Nucleic Acids Res. <u>29</u>:3583-94.
- Levy, S., Zeng, G. Q. et Danchin, A. <u>1990</u>. Cyclic AMP synthesis in *Escherichia coli* strains bearing known deletions in the *pts* phosphotransferase operon. *Gene*. 86:27-33.
- Li, Y. et Altman, S. <u>2003</u>. A specific endoribonuclease, RNase P, affects gene expression of polycistronic operon mRNAs. *Proc Natl Acad Sci U S A*. **100**:13213-8.
- Li, Z., Pandit, S. et Deutscher, M. P. 1999a. RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. EMBO J. 18:2878-85.
- Li, Z., Pandit, S. et Deutscher, M. P. <u>1999b.</u> Maturation of 23S ribosomal RNA requires the exoribonuclease RNase T. *RNA*. 5:139-46.
- Li, Z. et Deutscher, M. P. <u>2002</u>. RNase E plays an essential role in the maturation of *Escherichia coli* tRNA precursors. *RNA*. **8**:97-109.
- Li, Z., Reimers, S., Pandit, S. et Deutscher, M. P. <u>2002</u>. RNA quality control: degradation of defective transfer RNA. *EMBO J.* **21**:1132-8.
- Lin-Chao, S. et Cohen, S. N. <u>1991</u>. The rate of processing and degradation of antisense RNAI regulates the replication of *ColE1*-type plasmids in vivo. Cell. **65**:1233-42.
- Lin-Chao, S., Wong, T. T., Mcdowall, K. J. et Cohen, S. N. 1994. Effects of nucleotide sequence on the specificity of rne-dependent and RNase E-mediated cleavages of RNA I encoded by the pBR322 plasmid. J Biol Chem. 269:10797-803.
- Lin-Chao, S., Wei, C. L. et Lin, Y. T. <u>1999</u>. RNase E is required for the maturation of *ssrA* RNA and normal *ssrA* RNA peptide-tagging activity. *Proc Natl Acad Sci U S A*. **96**:12406-11.
- Liou, G. G., Jane, W. N., Cohen, S. N., Lin, N. S. et Lin-Chao, S. <u>2001</u>. RNA degradosomes exist in vivo in Escherichia coli as multicomponent complexes associated with the cytoplasmic membrane via the N-terminal region of ribonuclease E. Proc Natl Acad Sci USA. 98:63-8.
- Liou, G. G., Chang, H. Y., Lin, C. S. et Lin-Chao, S. 2002. DEAD box RhlB RNA helicase physically associates with exoribonuclease PNPase to degrade double-stranded RNA independent of the degradosome-assembling region of RNase E. J Biol Chem. 277:41157-62.

- Lithwick, G. et Margalit, H. <u>2003</u>. Hierarchy of sequence-dependent features associated with prokaryotic translation. *Genome Res.* 13:2665-73.
- Liu, X. et De Wulf, P. <u>2004</u>. Probing the ArcA-P modulon of *Escherichia coli* by whole genome transcriptional analysis and sequence recognition profiling. *J Biol Chem.* **279**:12588-97.
- Looman, A. C., Bodlaender, J., Comstock, L. J., Eaton, D., Jhurani, P., De Boer, H. A. et Van Knippenberg, P. H. <u>1987</u>. Influence of the codon following the AUG initiation codon on the expression of a modified *lacZ* gene in *Escherichia coli*. *EMBO J*. 6:2489-92.
- Lopez, P. J., Marchand, I., Joyce, S. A. et Dreyfus, M. 1999. The C-terminal half of RNase E, which organizes the *Escherichia coli* degradosome, participates in mRNA degradation but not rRNA processing in vivo. Mol Microbiol. 33:188-99.
- Ludwig, H., Homuth, G., Schmalisch, M., Dyka, F. M., Hecker, M. et Stulke, J. 2001. Transcription of glycolytic genes and operons in *Bacillus subtilis*: evidence for the presence of multiple levels of control of the gapA operon. Mol Microbiol. 41:409-22
- Lukat, G. S., Mccleary, W. R., Stock, A. M. et Stock, J. B. 1992. Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors. *Proc Natl Acad Sci U S A.* 89:718-22.
- Luli, G. W. et Strohl, W. R. 1990. Comparison of growth, acetate production, and acetate inhibition of Escherichia coli strains in batch and fed-batch fermentations. Appl Environ Microbiol. 56:1004-11.
- Lundberg, U., Von Gabain, A. et Melefors, O. 1990. Cleavages in the 5' region of the ompA and bla mRNA control stability: studies with an E. coli mutant altering mRNA stability and a novel endoribonuclease. EMBO J. 9:2731-41.
- Lundberg, U. et Altman, S. <u>1995</u>. Processing of the precursor to the catalytic RNA subunit of RNase P from *Escherichia coli*. RNA. 1:327-34.

#### M

- Ma, J., Campbell, A. et Karlin, S. <u>2002</u>. Correlations between Shine-Dalgarno sequences and gene features such as predicted expression levels and operon structures. *J Bacteriol*. **184**:5733-45.
- Mackie, G. A. <u>1998</u>. Ribonuclease E is a 5'-enddependent endonuclease. *Nature*. **395**:720-3.
- Mager, W. H. et De Kruijff, A. J. <u>1995</u>. Stress-induced transcriptional activation. *Microbiol Rev.* **59**:506-31.
- Majdalani, N., Hernandez, D. et Gottesman, S. 2002. Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. Mol Microbiol. 46:813-26.
- Makarova, O. V., Makarov, E. M., Sousa, R. et Dreyfus, M. 1995. Transcribing of Escherichia coli genes with mutant T7 RNA polymerases: stability of lacZ mRNA inversely correlates with polymerase speed. Proc Natl Acad Sci U S A. 92:12250-4.

- Makrides, S. C. 1996. Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev.* 60:512-38.
- Malhotra, A., Severinova, E. et Darst, S. A. 1996. Crystal structure of a  $\sigma^{70}$  subunit fragment from *E. coli* RNA polymerase. *Cell.* 87:127-36.
- Malpica, R., Franco, B., Rodriguez, C., Kwon, O. et Georgellis, D. <u>2004</u>. Identification of a quinonesensitive redox switch in the ArcB sensor kinase. *Proc Natl Acad Sci U S A*. **101**:13318-23.
- Mans, R. M., Guerrier-Takada, C., Altman, S. et Pleij, C. W. 1990. Interaction of RNase P from Escherichia coli with pseudoknotted structures in viral RNAs. Nucleic Acids Res. 18:3479-87.
- Marchand, I., Nicholson, A. W. et Dreyfus, M. 2001. Bacteriophage T7 protein kinase phosphorylates RNase E and stabilizes mRNAs synthesized by T7 RNA polymerase. *Mol Microbiol.* 42:767-76.
- Marr, M. T. et Roberts, J. W. <u>1997</u>. Promoter recognition as measured by binding of polymerase to nontemplate strand oligonucleotide. *Science*. **276**:1258-60.
- Martin, R. G. et Rosner, J. L. <u>2004</u>. Transcriptional and translational regulation of the *marRAB* multiple antibiotic resistance operon in *Escherichia coli*. *Mol Microbiol*. 53:183-91.
- Martin, W., Brinkmann, H., Savonna, C. et Cerff, R. 1993. Evidence for a chimeric nature of nuclear genomes: eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. Proc Natl Acad Sci U S A. 90:8692-6.
- Martinez-Antonio, A. et Collado-Vides, J. <u>2003</u>. Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr Opin Microbiol*. **6**:482-9.
- Marujo, P. E., Hajnsdorf, E., Le Derout, J., Andrade, R., Arraiano, C. M. et Regnier, P. <u>2000</u>. RNase II removes the oligo(A) tails that destabilize the *rpsO* mRNA of *Escherichia coli*. RNA. **6**:1185-93.
- Marujo, P. E., Braun, F., Haugel-Nielsen, J., Le Derout, J., Arraiano, C. M. et Regnier, P. 2003. Inactivation of the decay pathway initiated at an internal site by RNase E promotes poly(A)-dependent degradation of the rpsO mRNA in Escherichia coli. Mol Microbiol. 50:1283-94.
- Masse, E. et Gottesman, S. <u>2002</u>. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc Natl Acad Sci U S A*. **99**:4620-5.
- Masters, M., Colloms, M. D., Oliver, I. R., He, L., Macnaughton, E. J. et Charters, Y. 1993. The pcnB gene of Escherichia coli, which is required for ColE1 copy number maintenance, is dispensable. J Bacteriol. 175:4405-13.
- Mathy, N., Jarrige, A. C., Robert-Le Meur, M. et Portier, C. 2001. Increased expression of *Escherichia coli* polynucleotide phosphorylase at low temperatures is linked to a decrease in the efficiency of autocontrol. *J Bacteriol.* 183:3848-54.
- Matsunaga, J., Simons, E. L. et Simons, R. W. <u>1996</u>. RNase III autoregulation: structure and function of *rncO*, the posttranscriptional "operator". *RNA*. 2:1228-40.

- Mcdowall, K. J., Hernandez, R. G., Lin-Chao, S. et Cohen, S. N. 1993. The ams-1 and rne-3071 temperature-sensitive mutations in the ams gene are in close proximity to each other and cause substitutions within a domain that resembles a product of the Escherichia coli mre locus. J Bacteriol. 175:4245-9.
- Mcdowall, K. J., Lin-Chao, S. et Cohen, S. N. 1994. A+U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. J Biol Chem. 269:10790-6.
- Mcdowall, K. J. et Cohen, S. N. 1996. The N-terminal domain of the *rne* gene product has RNase E activity and is non-overlapping with the argininerich RNA-binding site. *J Mol Biol.* 255:349-55.
- Mclaren, R. S., Newbury, S. F., Dance, G. S., Causton, H. C. et Higgins, C. F. 1991. mRNA degradation by processive 3'-5' exoribonucleases in vitro and the implications for prokaryotic mRNA decay in vivo. J Mol Biol. 221:81-95.
- Mcleod, S. M., Aiyar, S. E., Gourse, R. L. et Johnson, R. C. 2002. The C-terminal domains of the RNA polymerase α subunits: contact site with Fis and localization during co-activation with CRP at the Escherichia coli proP P2 promoter. J Mol Biol. 316:517-29.
- Mechold, U., Murphy, H., Brown, L. et Cashel, M. 2002. Intramolecular regulation of the opposing (p)ppGpp catalytic activities of Rel(Seq), the Rel/Spo enzyme from Streptococcus equisimilis. J Bacteriol. 184:2878-88.
- Meyer, D., Schneider-Fresenius, C., Horlacher, R., Peist, R. et Boos, W. <u>1997</u>. Molecular characterization of glucokinase from *Escherichia coli* K-12. *J Bacteriol*. **179**:1298-306.
- 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 U S A.* 93:3865-9.
- Mikulecky, P. J., Kaw, M. K., Brescia, C. C., Takach, J. C., Sledjeski, D. D. et Feig, A. L. <u>2004</u>. Escherichia coli Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. Nat Struct Mol Biol. <u>11</u>:1206-14.
- Mikulskis, A., Aristarkhov, A. et Lin, E. C. 1997. Regulation of expression of the ethanol dehydrogenase gene (adhE) in Escherichia coli by catabolite repressor activator protein Cra. J. Bacteriol. 179:7129-34.
- Minakhin, L., Bhagat, S., Brunning, A., Campbell, E. A., Darst, S. A., Ebright, R. H. et Severinov, K. 2001. Bacterial RNA polymerase subunit ω and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. Proc Natl Acad Sci U S A. 98:892-7.
- Misra, T. K. et Apirion, D. <u>1979.</u> RNase E, an RNA processing enzyme from *Escherichia coli. J Biol Chem.* **254**:11154-9.
- Mitchell, J. E., Zheng, D., Busby, S. J. et Minchin, S. D. 2003. Identification and analysis of 'extended -10' promoters in *Escherichia coli*. Nucleic Acids Res. 31:4689-95.

- Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. et Tollervey, D. <u>1997</u>. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. *Cell.* **91**:457-66.
- Mohanty, B. K. et Kushner, S. R. 1999a. Analysis of the function of *Escherichia coli* poly(A) polymerase I in RNA metabolism. *Mol Microbiol*. 34:1094-108.
- Mohanty, B. K. et Kushner, S. R. 1999b. Residual polyadenylation in poly(A) polymerase I (pcnB) mutants of Escherichia coli does not result from the activity encoded by the f310 gene. Mol Microbiol. 34:1109-19.
- Mohanty, B. K. et Kushner, S. R. <u>2000</u>. Polynucleotide phosphorylase functions both as a 3' right-arrow 5' exonuclease and a poly(A) polymerase in *Escherichia coli*. *Proc Natl Acad Sci U S A*. **97**:11966-71.
- Mohanty, B. K. et Kushner, S. R. 2002. Polyadenylation of *Escherichia coli* transcripts plays an integral role in regulating intracellular levels of polynucleotide phosphorylase and RNase E. *Mol Microbiol*. 45:1315-24.
- Mohanty, B. K., Maples, V. F. et Kushner, S. R. 2004. The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in *Escherichia coli. Mol. Microbiol.* 54:905-20.
- Moll, I., Huber, M., Grill, S., Sairafi, P., Mueller, F., Brimacombe, R., Londei, P. et Blasi, U. <u>2001</u>. Evidence against an Interaction between the mRNA downstream box and 16S rRNA in translation initiation. *J Bacteriol*. **183**:3499-505.
- Moll, I., Grill, S., Gualerzi, C. O. et Blasi, U. 2002. Leaderless mRNAs in bacteria: surprises in ribosomal recruitment and translational control. *Mol Microbiol.* 43:239-46.
- Moll, I., Leitsch, D., Steinhauser, T. et Blasi, U. <u>2003a</u>. RNA chaperone activity of the Sm-like Hfq protein. *EMBO Rep.* 4:284-9.
- Moll, I., Afonyushkin, T., Vytvytska, O., Kaberdin, V. R. et Blasi, U. <u>2003b.</u> Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. RNA. **9**:1308-14.
- Moll, I., Hirokawa, G., Kiel, M. C., Kaji, A. et Blasi, U. 2004. Translation initiation with 70S ribosomes: an alternative pathway for leaderless mRNAs. *Nucleic Acids Res.* 32:3354-63.
- Moller, T., Franch, T., Hojrup, P., Keene, D. R., Bachinger, H. P., Brennan, R. G. et Valentin-Hansen, P. 2002. Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol Cell*. 9:23-30.
- Morita, T., El-Kazzaz, W., Tanaka, Y., Inada, T. et Aiba, H. 2003. Accumulation of glucose 6-phosphate or fructose 6-phosphate is responsible for destabilization of glucose transporter mRNA in Escherichia coli. J Biol Chem. 278:15608-14.
- Morita, T., Kawamoto, H., Mizota, T., Inada, T. et Aiba, H. 2004a. Enolase in RNA degradosome plays a crucial role in the rapid decay of glucose transporter mRNA in response to metabolic stress in Escherichia coli.
- Morita, T., Kawamoto, H., Mizota, T., Inada, T. et Aiba, H. 2004b. Enolase in the RNA degradosome plays a crucial role in the rapid decay of glucose transporter

- mRNA in the response to phosphosugar stress in *Escherichia coli. Mol. Microbiol.* **54**:1063-75.
- Mottagui-Tabar, S., Bjornsson, A. et Isaksson, L. A. 1994. The second to last amino acid in the nascent peptide as a codon context determinant. EMBO J. 13:249-57.
- Mottagui-Tabar, S. et Isaksson, L. A. 1997. Only the last amino acids in the nascent peptide influence translation termination in *Escherichia coli* genes. FEBS Lett. 414:165-70.
- Muffler, A., Fischer, D. et Hengge-Aronis, R. 1996. The RNA-binding protein HF-I, known as a host factor for phage Qβ RNA replication, is essential for rpoS translation in Escherichia coli. Genes Dev. 10:1143-51.
- Muffler, A., Traulsen, D. D., Fischer, D., Lange, R. et Hengge-Aronis, R. 1997. The RNA-binding protein HF-I plays a global regulatory role which is largely, but not exclusively, due to its role in expression of the σ<sup>S</sup> subunit of RNA polymerase in *Escherichia* coli. J Bacteriol. 179:297-300.
- Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O. et Darst, S. A. <u>2002</u>. Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science*. **296**:1285-90.
- Murakami, K. S. et Darst, S. A. <u>2003</u>. Bacterial RNA polymerases: the wholo story. *Curr Opin Struct Biol*. **13**:31-9.
- Murray, H. D., Schneider, D. A. et Gourse, R. L. <u>2003</u>. Control of rRNA expression by small molecules is dynamic and nonredundant. *Mol Cell*. **12**:125-34.

#### N

- Nagano, N., Orengo, C. A. et Thornton, J. M. <u>2002.</u> One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J Mol Biol.* 321:741-65.
- Nam, T. W., Cho, S. H., Shin, D., Kim, J. H., Jeong, J. Y., Lee, J. H., Roe, J. H., Peterkofsky, A., Kang, S. O., Ryu, S. et Seok, Y. J. <u>2001</u>. The *Escherichia coli* glucose transporter enzyme IICB<sup>Gle</sup> recruits the global repressor Mlc. *EMBO J.* **20**:491-8.
- Negre, D., Bonod-Bidaud, C., Geourjon, C., Deleage, G., Cozzone, A. J. et Cortay, J. C. <u>1996</u>. Definition of a consensus DNA-binding site for the *Escherichia coli* pleiotropic regulatory protein, FruR. *Mol Microbiol*. 21:257-66.
- Negre, D., Oudot, C., Prost, J. F., Murakami, K., Ishihama, A., Cozzone, A. J. et Cortay, J. C. 1998. FruR-mediated transcriptional activation at the ppsA promoter of Escherichia coli. J Mol Biol. 276:355-65.
- Newbury, S. F., Smith, N. H. et Higgins, C. F. <u>1987a.</u> Differential mRNA stability controls relative gene expression within a polycistronic operon. *Cell.* **51**:1131-43.
- Newbury, S. F., Smith, N. H., Robinson, E. C., Hiles, I. D. et Higgins, C. F. <u>1987b</u>. Stabilization of translationally active mRNA by prokaryotic REP sequences. *Cell*. **48**:297-310.

- Nicholson, A. W. 1999. Function, mechanism and regulation of bacterial ribonucleases. FEMS Microbiol Rev. 23:371-90.
- Nierlich, D. P. et Murakawa, G. J. <u>1996</u>. The decay of bacterial messenger RNA. *Prog Nucleic Acid Res Mol Biol.* **52**:153-216.
- Nilsson, G., Belasco, J. G., Cohen, S. N. et Von Gabain, A. <u>1984</u>. Growth-rate dependent regulation of mRNA stability in *Escherichia coli*. *Nature*. **312**:75-7.
- Nilsson, G., Lundberg, U. et Von Gabain, A. <u>1988</u>. In vivo and in vitro identity of site specific cleavages in the 5' non-coding region of ompA and bla mRNA in Escherichia coli. EMBO J. 7:2269-75.
- Nivinskas, R., Malys, N., Klausa, V., Vaiskunaite, R. et Gineikiene, E. 1999. Post-transcriptional control of bacteriophage T4 gene 25 expression: mRNA secondary structure that enhances translational initiation. J Mol Biol. 288:291-304.
- Nogueira, T., De Smit, M., Graffe, M. et Springer, M. 2001. The relationship between translational control and mRNA degradation for the *Escherichia coli* threonyl-tRNA synthetase gene. *J Mol Biol.* 310:709-22.
- Notley-Mcrobb, L., Death, A. et Ferenci, T. 1997. The relationship between external glucose concentration and cAMP levels inside *Escherichia coli*: implications for models of phosphotransferase-mediated regulation of adenylate cyclase. *Microbiology.* 143 ( Pt 6):1909-18.
- Nudler, E. <u>1999.</u> Transcription elongation: structural basis and mechanisms. *J Mol Biol.* **288**:1-12.
- Nudler, E. et Gottesman, M. E. <u>2002</u>. Transcription termination and anti-termination in *E. coli*. *Genes Cells*. 7:755-68.
- Nudler, E. et Mironov, A. S. <u>2004</u>. The riboswitch control of bacterial metabolism. *Trends Biochem Sci.* 29:11-7.
- Nuoffer, C., Zanolari, B. et Erni, B. <u>1988</u>. Glucose permease of *Escherichia coli*. The effect of cysteine to serine mutations on the function, stability, and regulation of transport and phosphorylation. *J Biol Chem.* **263**:6647-55.

#### 0

- O'connor, M., Asai, T., Squires, C. L. et Dahlberg, A. E.

  1999. Enhancement of translation by the
  downstream box does not involve base pairing of
  mRNA with the penultimate stem sequence of 16S
  rRNA. Proc Natl Acad Sci U S A. 96:8973-8.
- O'donnell, S. M. et Janssen, G. R. <u>2002</u>. Leaderless mRNAs bind 70S ribosomes more strongly than 30S ribosomal subunits in *Escherichia coli*. *J Bacteriol*. **184**:6730-3.
- Oh, M. K. et Liao, J. C. <u>2000</u>. DNA microarray detection of metabolic responses to protein overproduction in *Escherichia coli*. *Metab Eng.* 2:201-9.
- Oh, M. K., Rohlin, L., Kao, K. C. et Liao, J. C. 2002. Global expression profiling of acetate-grown Escherichia coli. J Biol Chem. 277:13175-83.
- O'hara, E. B., Chekanova, J. A., Ingle, C. A., Kushner, Z. R., Peters, E. et Kushner, S. R. 1995.

- Polyadenylylation helps regulate mRNA decay in Escherichia coli. Proc Natl Acad Sci U S A. 92:1807-11.
- Okada, Y., Wachi, M., Hirata, A., Suzuki, K., Nagai, K. et Matsuhashi, M. 1994. Cytoplasmic axial filaments in *Escherichia coli* cells: possible function in the mechanism of chromosome segregation and cell division. *J Bacteriol.* 176:917-22.
- Okamoto, K. et Freundlich, M. <u>1986</u>. Mechanism for the autogenous control of the *crp* operon: transcriptional inhibition by a divergent RNA transcript. *Proc Natl* Acad Sci U S A. 83:5000-4.
- Okamoto, K., Hara, S., Bhasin, R. et Freundlich, M. 1988. Evidence in vivo for autogenous control of the cyclic AMP receptor protein gene (crp) in Escherichia coli by divergent RNA. J Bacteriol. 170:5076-9.
- Ono, M. et Kuwano, M. <u>1979</u>. A conditional lethal mutation in an *Escherichia coli* strain with a longer chemical lifetime of messenger RNA. *J Mol Biol*. 129:343-57.
- Opdyke, J. A., Kang, J. G. et Storz, G. <u>2004</u>. GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J Bacteriol*. **186**:6698-705.
- Osumi, T. et Saier, M. H., Jr. 1982. Regulation of lactose permease activity by the phosphoenolpyruvate:sugar phosphotransferase system: evidence for direct binding of the glucose-specific enzyme III to the lactose permease. *Proc Natl Acad Sci U S A*. 79:1457-61.
- Ow, M. C., Liu, Q. et Kushner, S. R. <u>2000</u>. Analysis of mRNA decay and rRNA processing in *Escherichia* coli in the absence of RNase E-based degradosome assembly. *Mol Microbiol*. **38**:854-66.
- Ow, M. C. et Kushner, S. R. <u>2002</u>. Initiation of tRNA maturation by RNase E is essential for cell viability in *E. coli*. *Genes Dev*. **16**:1102-15.
- Ow, M. C., Liu, Q., Mohanty, B. K., Andrew, M. E., Maples, V. F. et Kushner, S. R. <u>2002</u>. RNase E levels in *Escherichia coli* are controlled by a complex regulatory system that involves transcription of the *rne* gene from three promoters. *Mol Microbiol*. **43**:159-71.
- Ow, M. C., Perwez, T. et Kushner, S. R. <u>2003</u>. RNase G of *Escherichia coli* exhibits only limited functional overlap with its essential homologue, RNase E. *Mol Microbiol*. **49**:607-22.

#### P

- Paget, M. S. et Helmann, J. D. <u>2003</u>. The σ<sup>70</sup> family of sigma factors. *Genome Biol*. **4**:203.
- Panaghie, G., Aiyar, S. E., Bobb, K. L., Hayward, R. S. et De Haseth, P. L. <u>2000</u>. Aromatic amino acids in region 2.3 of *Escherichia coli* σ<sup>70</sup> participate collectively in the formation of an RNA polymerase-promoter open complex. *J Mol Biol*. **299**:1217-30.
- Pannone, B. K. et Wolin, S. L. <u>2000.</u> Sm-like proteins wRING the neck of mRNA. *Curr Biol.* **10**:R478-81.
- Pannucci, J. A., Haas, E. S., Hall, T. A., Harris, J. K. et Brown, J. W. 1999. RNase P RNAs from some

- Archaea are catalytically active. *Proc Natl Acad Sci U S A.* **96**:7803-8.
- Park, S. J., Chao, G. et Gunsalus, R. P. <u>1997</u>. Aerobic regulation of the *sucABCD* genes of *Escherichia coli*, which encode α-ketoglutarate dehydrogenase and succinyl coenzyme A synthetase: roles of ArcA, Fnr, and the upstream *sdhCDAB* promoter. *J Bacteriol*. **179**:4138-42.
- Paulsen, I. T., Nguyen, L., Sliwinski, M. K., Rabus, R. et Saier, M. H., Jr. <u>2000.</u> Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes. *J Mol Biol.* **301**:75-100.
- Paulus, M., Haslbeck, M. et Watzele, M. <u>2004</u>. RNA stem-loop enhanced expression of previously nonexpressible genes. *Nucleic Acids Res.* 32:e78.
- Peck-Miller, K. A. et Altman, S. 1991. Kinetics of the processing of the precursor to 4.5 S RNA, a naturally occurring substrate for RNase P from Escherichia coli. J Mol Biol. 221:1-5.
- Peng, L. et Shimizu, K. <u>2003.</u> Global metabolic regulation analysis for *Escherichia coli* K12 based on protein expression by 2-dimensional electrophoresis and enzyme activity measurement. *Appl Microbiol Biotechnol.* **61**:163-78.
- Petersen, C. <u>1993</u>. Translation and mRNA stability in bacteria: a complex relationship.
- Petersen, C. et Moller, L. B. <u>2000</u>. Invariance of the nucleoside triphosphate pools of *Escherichia coli* with growth rate. *J Biol Chem.* **275**:3931-5.
- Philippe, C., Portier, C., Mougel, M., Grunberg-Manago, M., Ebel, J. P., Ehresmann, B. et Ehresmann, C. 1990. Target site of *Escherichia coli* ribosomal protein S15 on its messenger RNA. Conformation and interaction with the protein. *J. Mol. Biol.* 211:415-26.
- Picon, A., Teixeira De Mattos, M. J. et Postma, P. W. 2005. Reducing the glucose uptake rate in Escherichia coli affects growth rate but not protein production. Biotechnol Bioeng. 90:191-200.
- Plaumann, M., Pelzer-Reith, B., Martin, W. F. et Schnarrenberger, C. 1997. Multiple recruitment of class-I aldolase to chloroplasts and eubacterial origin of eukaryotic class-II aldolases revealed by cDNAs from Euglena gracilis. Curr Genet. 31:430-8
- Plumbridge, J. et Soll, D. <u>1989.</u> Characterization of cisacting mutations which increase expression of a *glnS-lacZ* fusion in *Escherichia coli. Mol Gen Genet.* **216**:113-9.
- Plumbridge, J. <u>1996.</u> How to achieve constitutive expression of a gene within an inducible operon: the example of the *nagC* gene of *Escherichia coli. J Bacteriol.* **178**:2629-36.
- Plumbridge, J. 1998a. Control of the expression of the manXYZ operon in Escherichia coli: Mlc is a negative regulator of the mannose PTS. Mol. Microbiol. 27:369-80.
- Plumbridge, J. 1998b. Expression of ptsG, the gene for the major glucose PTS transporter in Escherichia coli, is repressed by Mlc and induced by growth on glucose. Mol. Microbiol. 29:1053-63.

- Plumbridge, J. et Kolb, A. <u>1998</u>. DNA bending and expression of the divergent nagE-B operons. Nucleic Acids Res. 26:1254-60.
- Plumbridge, J. 1999. Expression of the phosphotransferase system both mediates and is mediated by mlc regulation in Escherichia coli. Mol. Microbiol. 33:260-73.
- Plumbridge, J. et Vimr, E. <u>1999</u>. Convergent pathways for utilization of the amino sugars Nacetylglucosamine, N-acetylmannosamine, and Nacetylneuraminic acid by *Escherichia coli*. J Bacteriol. **181**:47-54.
- Plumbridge, J. <u>2001</u>. DNA binding sites for the Mlc and NagC proteins: regulation of *nagE*, encoding the Nacetylglucosamine-specific transporter in *Escherichia coli*. *Nucleic Acids Res.* **29**:506-14.
- Plumbridge, J. <u>2002.</u> Regulation of gene expression in the PTS in *Escherichia coli*: the role and interactions of Mlc. *Curr Opin Microbiol.* 5:187-93.
- Plumbridge, J. A. <u>1991</u>. Repression and induction of the nag regulon of *Escherichia coli* K-12: the roles of *nagC* and *nagA* in maintenance of the uninduced state. *Mol Microbiol*. 5:2053-62.
- Pollack, J. D., Li, Q. et Pearl, D. K. 2005. Taxonomic utility of a phylogenetic analysis of phosphoglycerate kinase proteins of Archaea, Bacteria, and Eukaryota: Insights by Bayesian analyses. Mol Phylogen et Evol. 35:420-30.
- Poole, E. S., Brown, C. M. et Tate, W. P. 1995. The identity of the base following the stop codon determines the efficiency of in vivo translational termination in Escherichia coli. EMBO J. 14:151-8.
- Postma, P. W., Lengeler, J. W. et Jacobson, G. R. 1993. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol. Rev.* 57:543-94.
- Postma, P.W., Lengeler, J.W., Jacobson G.R. 1996. Phosphoenolpyruvate:carbohydrate phosphotransferase systems. *Escherichia coli* and *Salmonella*: cellular and molecular biology 2nd edn. Washington D.C. American Society for microbiology press. 1:1149-1174.
- Prost, J. F., Negre, D., Oudot, C., Murakami, K., Ishihama, A., Cozzone, A. J. et Cortay, J. C. <u>1999</u>. Cra-dependent transcriptional activation of the *icd* gene of *Escherichia coli*. *J Bacteriol*. **181**:893-8.
- Prud'homme-Généreux, A., Beran, R. K., Iost, I., Shane Ramey, C., Mackie, G. A. et Simons, R. W. 2004. Physical and functional interactions among RNase E polynycleotide phosphorylase and the cold-shock protein, CsdA: evidence for a 'cold shock degradosome'. Mol Microbiol.
- Py, B., Higgins, C. F., Krisch, H. M. et Carpousis, A. J. 1996. A DEAD-box RNA helicase in the Escherichia coli RNA degradosome. Nature. 381:169-72.

# a

Qamar, S., Marsh, K. et Berry, A. <u>1996</u>. Identification of arginine 331 as an important active site residue in the class II fructose-1,6-bisphosphate aldolase of *Escherichia coli*. *Protein Sci.* 5:154-61.

# R

- Raina, S., Missiakas, D. et Georgopoulos, C. <u>1995</u>. The rpoE gene encoding the sigma E (sigma 24) heat shock sigma factor of *Escherichia coli*. *EMBO J*. 14:1043-55.
- Ramseier, T. M., Negre, D., Cortay, J. C., Scarabel, M., Cozzone, A. J. et Saier, M. H., Jr. 1993. In vitro binding of the pleiotropic transcriptional regulatory protein, FruR, to the fru, pps, ace, pts and icd operons of Escherichia coli and Salmonella typhimurium. J Mol Biol. 234:28-44.
- Ramseier, T. M., Bledig, S., Michotey, V., Feghali, R. et Saier, M. H., Jr. 1995. The global regulatory protein FruR modulates the direction of carbon flow in Escherichia coli. Mol Microbiol. 16:1157-69.
- Ramseier, T. M. <u>1996.</u> Cra and the control of carbon flux via metabolic pathways. *Res Microbiol.* **147**:489-93.
- Ramseier, T. M., Chien, S. Y. et Saier, M. H., Jr. 1996. Cooperative interaction between Cra and Fnr in the regulation of the cydAB operon of Escherichia coli. Curr Microbiol. 33:270-4.
- Rauhut, R. et Klug, G. 1999. mRNA degradation in bacteria. FEMS Microbiol Rev. 23:353-70.
- Ray, B. K. et Apirion, D. <u>1981</u>. Transfer RNA precursors are accumulated in *Escherichia coli* in the absence of RNase E. *Eur J Biochem*. **114**:517-24.
- Raynal, L. C. et Carpousis, A. J. <u>1999.</u> Poly(A) polymerase I of *Escherichia coli*: characterization of the catalytic domain, an RNA binding site and regions for the interaction with proteins involved in mRNA degradation. *Mol Microbiol.* 32:765-75.
- Reddy, P. et Kamireddi, M. <u>1998</u>. Modulation of Escherichia coli adenylyl cyclase activity by catalytic-site mutants of protein IIA<sup>Glc</sup> of the phosphoenolpyruvate:sugar phosphotransferase system. J Bacteriol. **180**:732-6.
- Redko, Y., Tock, M. R., Adams, C. J., Kaberdin, V. R., Grasby, J. A. et Mcdowall, K. J. <u>2003</u>. Determination of the catalytic parameters of the Nterminal half of *Escherichia coli* ribonuclease E and the identification of critical functional groups in RNA substrates. *J Biol Chem.* **278**:44001-8.
- Regnier, P. et Portier, C. <u>1986</u>. Initiation, attenuation and RNase III processing of transcripts from the *Escherichia coli* operon encoding ribosomal protein S15 and polynucleotide phosphorylase. *J Mol Biol*. **187**:23-32.
- Regnier, P., Grunberg-Manago, M. et Portier, C. <u>1987</u>. Nucleotide sequence of the *pnp* gene of *Escherichia coli* encoding polynucleotide phosphorylase. Homology of the primary structure of the protein with the RNA-binding domain of ribosomal protein S1. *J Biol Chem.* **262**:63-8.
- Regnier, P. et Arraiano, C. M. <u>2000</u>. Degradation of mRNA in bacteria: emergence of ubiquitous features. *Bioessays*. **22**:235-44.
- Regonesi, M. E., Briani, F., Ghetta, A., Zangrossi, S., Ghisotti, D., Tortora, P. et Deho, G. 2004. A mutation in polynucleotide phosphorylase from Escherichia coli impairing RNA binding and degradosome stability. Nucleic Acids Res. 32:1006-17.

- Reitzer, L. et Schneider, B. L. <u>2001</u>. Metabolic context and possible physiological themes of σ<sup>54</sup>-dependent genes in *Escherichia coli*. *Microbiol Mol Biol Rev*. **65**:422-44, table of contents.
- Reizer, J., Reizer, A., Kornberg, H. L. et Saier, M. H., Jr. 1994. Sequence of the fruB gene of Escherichia coli encoding the diphosphoryl transfer protein (DTP) of the phosphoenolpyruvate: sugar phosphotransferase system. FEMS Microbiol Lett. 118:159-62.
- Repoila, F., Majdalani, N. et Gottesman, S. <u>2003</u>. Small non-coding RNAs, co-ordinators of adaptation processes in *Escherichia coli*: the RpoS paradigm. *Mol Microbiol*. **48**:855-61.
- Rhodius, V. A. et Busby, S. J. <u>1998</u>. Positive activation of gene expression. *Curr Opin Microbiol*. 1:152-9.
- Richardson, J. P. <u>2002.</u> Rho-dependent termination and ATPases in transcript termination. *Biochim Biophys Acta*. **1577**:251-60.
- Richardson, J. P. <u>2003</u>. Loading Rho to terminate transcription. *Cell*. **114**:157-9.
- Richet, E., Vidal-Ingigliardi, D. et Raibaud, O. <u>1991</u>. A new mechanism for coactivation of transcription initiation: repositioning of an activator triggered by the binding of a second activator. *Cell.* **66**:1185-95.
- Ringquist, S., Shinedling, S., Barrick, D., Green, L., Binkley, J., Stormo, G. D. et Gold, L. <u>1992</u>. Translation initiation in *Escherichia coli*: sequences within the ribosome-binding site. *Mol Microbiol*. 6:1219-29.
- Ringquist, S., Jones, T., Snyder, E. E., Gibson, T., Boni, I. et Gold, L. 1995. High-affinity RNA ligands to Escherichia coli ribosomes and ribosomal protein S1: comparison of natural and unnatural binding sites. Biochemistry. 34:3640-8.
- Robillard, G. T. et Broos, J. <u>1999</u>. Structure/function studies on the bacterial carbohydrate transporters, enzymes II, of the phosphoenolpyruvate-dependent phosphotransferase system. *Biochim Biophys Acta*. **1422**:73-104.
- Rocak, S. et Linder, P. <u>2004.</u> DEAD-box proteins: the driving forces behind RNA metabolism. *Nat Rev Mol Cell Biol.* 5:232-41.
- Rodriguez, C., Kwon, O. et Georgellis, D. <u>2004</u>. Effect of D-lactate on the physiological activity of the ArcB sensor kinase in *Escherichia coli*. *J Bacteriol*. **186**:2085-90.
- Rojo, F. <u>2001.</u> Mechanisms of transcriptional repression. Curr Opin Microbiol. **4**:145-51.
- Romby, P., Caillet, J., Ebel, C., Sacerdot, C., Graffe, M., Eyermann, F., Brunel, C., Moine, H., Ehresmann, C., Ehresmann, B. et Springer, M. 1996. The expression of *E.coli* threonyl-tRNA synthetase is regulated at the translational level by symmetrical operator-repressor interactions. *EMBO J.* 15:5976-87.
- Romby, P. et Springer, M. <u>2003</u>. Bacterial translational control at atomic resolution. *Trends Genet.* **19**:155-61.
- Romeo, T. <u>1998.</u> Global regulation by the small RNAbinding protein CsrA and the non-coding RNA molecule CsrB. *Mol Microbiol.* **29**:1321-30.
- Rossmann, M. G., Moras, D. et Olsen, K. W. <u>1974</u>. Chemical and biological evolution of nucleotidebinding protein. *Nature*. **250**:194-9.

- Rouviere, P. E., De Las Penas, A., Mecsas, J., Lu, C. Z., Rudd, K. E. et Gross, C. A. <u>1995.</u> *rpoE*, the gene encoding the second heat-shock sigma factor, σ<sup>E</sup>, in *Escherichia coli. EMBO J.* **14**:1032-42.
- Roy, A. et Danchin, A. <u>1982</u>. The *cya* locus of *Escherichia coli* K12: organization and gene products. *Mol Gen Genet*. **188**:465-71.
- Ryder, A. M. et Roberts, J. W. <u>2003</u>. Role of the nontemplate strand of the elongation bubble in intrinsic transcription termination. *J Mol Biol*. **334**:205-13.
- Ryu, S. et Garges, S. 1994a. Promoter switch in the Escherichia coli pts operon. J. Biol. Chem. 269:4767-72.
- Ryu, S., Ramseier, T. M., Michotey, V., Saier, M. H., Jr. et Garges, S. <u>1995</u>. Effect of the FruR regulator on transcription of the pts operon in Escherichia coli. J Biol Chem. **270**:2489-96.

# S

- Sacerdot, C., Chiaruttini, C., Engst, K., Graffe, M., Milet, M., Mathy, N., Dondon, J. et Springer, M. 1996. The role of the AUU initiation codon in the negative feedback regulation of the gene for translation initiation factor IF3 in Escherichia coli. Mol Microbiol. 21:331-46.
- Sacerdot, C., Caillet, J., Graffe, M., Eyermann, F., Ehresmann, B., Ehresmann, C., Springer, M. et Romby, P. 1998. The Escherichia coli threonyltRNA synthetase gene contains a split ribosomal binding site interrupted by a hairpin structure that is essential for autoregulation. Mol Microbiol. 29:1077-90.
- Saier, M. H., Jr., Simoni, R. D. et Roseman, S. <u>1970</u>. The physiological behavior of enzyme I and heat-stable protein mutants of a bacterial phosphotransferase system. *J Biol Chem.* **245**:5870-3.
- Saier, M. H., Jr., Feucht, B. U. et Mccaman, M. T. 1975. Regulation of intracellular adenosine cyclic 3':5'-monophosphate levels in *Escherichia coli* and *Salmonella typhimurium*. Evidence for energy-dependent excretion of the cyclic nucleotide. *J. Biol. Chem.* 250:7593-601.
- Saier, M. H., Jr., Feucht, B. U. et Hofstadter, L. J. 1976. Regulation of carbohydrate uptake and adenylate cyclase activity mediated by the enzymes II of the phosphoenolpyruvate: sugar phosphotransferase system in *Escherichia coli*. J. Biol. Chem. 251:883-02
- Saier, M. H., Jr., Straud, H., Massman, L. S., Judice, J. J., Newman, M. J. et Feucht, B. U. 1978. Permease-specific mutations in Salmonella typhimurium and Escherichia coli that release the glycerol, maltose, melibiose, and lactose transport systems from regulation by the phosphoenolpyruvate:sugar phosphotransferase system. J. Bacteriol. 133:1358-67
- Saier, M. H., Jr., Novotny, M. J., Comeau-Fuhrman, D., Osumi, T. et Desai, J. D. <u>1983</u>. Cooperative binding of the sugar substrates and allosteric regulatory protein (enzyme III<sup>Glc</sup> of the phosphotransferase system) to the lactose and melibiose permeases in

- Escherichia coli and Salmonella typhimurium. J Bacteriol. 155:1351-7.
- Saier, M. H., Jr. 1996. Cyclic AMP-independent catabolite repression in bacteria. FEMS Microbiol Lett. 138:97-103.
- Saier, M. H., Jr. et Ramseier, T. M. 1996. The catabolite repressor/activator (Cra) protein of enteric bacteria. *J Bacteriol.* 178:3411-7.
- Salgado, H., Gama-Castro, S., Martinez-Antonio, A., Diaz-Peredo, E., Sanchez-Solano, F., Peralta-Gil, M., Garcia-Alonso, D., Jimenez-Jacinto, V., Santos-Zavaleta, A., Bonavides-Martinez, C. et Collado-Vides, J. <u>2004</u>. RegulonDB (version 4.0): transcriptional regulation, operon organization and growth conditions in *Escherichia coli* K-12. *Nucleic Acids Res.* **32 Database issue**:D303-6.
- Salmon, K. A., Hung, S. P., Steffen, N. R., Krupp, R., Baldi, P., Hatfield, G. W. et Gunsalus, R. P. 2005. Global gene expression profiling in *Escherichia coli* K12: effects of oxygen availability and ArcA. *J Biol Chem.* 280:15084-96.
- Sanchez, L., Horner, D., Moore, D., Henze, K., Embley, T. et Muller, M. 2002. Fructose-1,6-bisphosphate aldolases in amitochondriate protists constitute a single protein subfamily with eubacterial relationships. Gene. 295:51-9.
- Sasse-Dwight, S. et Gralla, J. D. <u>1988.</u> Probing the Escherichia coli glnALG upstream activation mechanism in vivo. Proc Natl Acad Sci U S A. 85:8934-8.
- Sauter, C., Basquin, J. et Suck, D. <u>2003.</u> Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from *Escherichia coli*. *Nucleic Acids Res.* **31**:4091-8
- Scamuffa, M. D. et Caprioli, R. M. <u>1980.</u> Comparison of the mechanisms of two distinct aldolases from *Escherichia coli* grown on gluconeogenic substrates. *Biochim. Biophys. Acta.* **614**:583-90.
- Schaechter, M., Maaloe, O. et Kjeldgaard, N. O. 1958. Dependency on medium and temperature of cell size and chemical composition during balanced grown of Salmonella typhimurium. J Gen Microbiol. 19:592-606.
- Schläpfer, B. S. et Zuber, H. 1992. Cloning and sequencing of the genes encoding glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase and triosephosphate isomerase (gap operon) from mesophilic Bacillus megaterium: comparison with corresponding sequences from thermophilic Bacillus stearothermophilus. Gene. 122:53-62.
- Schlax, P. J. et Worhunsky, D. J. <u>2003.</u> Translational repression mechanisms in prokaryotes. *Mol Microbiol.* 48:1157-69.
- Schneider, D. A., Gaal, T. et Gourse, R. L. <u>2002</u>. NTP-sensing by rRNA promoters in *Escherichia coli* is direct. *Proc Natl Acad Sci U S A*. **99**:8602-7.
- Schneider, D. A. et Gourse, R. L. <u>2003</u>. Changes in the concentrations of guanosine 5'-diphosphate 3'-diphosphate and the initiating nucleoside triphosphate account for inhibition of rRNA transcription in fructose-1,6-diphosphate aldolase (*fda*) mutants. *J Bacteriol*. **185**:6192-4.
- Schneider, D. A. et Gourse, R. L. 2004. Relationship between growth rate and ATP concentration in

- Escherichia coli: a bioassay for available cellular ATP. J Biol Chem. 279:8262-8.
- Schroder, O. et Wagner, R. <u>2002</u>. The bacterial regulatory protein H-NS--a versatile modulator of nucleic acid structures. *Biol Chem.* 383:945-60.
- Schubert, M., Edge, R. E., Lario, P., Cook, M. A., Strynadka, N. C., Mackie, G. A. et Mcintosh, L. P. 2004. Structural characterization of the RNase E S1 domain and identification of its oligonucleotidebinding and dimerization interfaces. *J Mol Biol.* 341:37-54.
- Schulz, V. P. et Reznikoff, W. S. 1990. In vitro secondary structure analysis of mRNA from lacZ translation initiation mutants. J Mol Biol. 211:427-45.
- Schumacher, M. A., Pearson, R. F., Moller, T., Valentin-Hansen, P. et Brennan, R. G. <u>2002</u>. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *EMBO* J. **21**:3546-56.
- Schurr, T., Nadir, E. et Margalit, H. <u>1993</u>. Identification and characterization of *E. coli* ribosomal binding sites by free energy computation. *Nucleic Acids Res*. 21:4019-23.
- Segal, H. L. et Boyer, P. D. <u>1953</u>. The role of sulfhydryl groups in the activity of D-glyceraldehyde 3phosphate dehydrogenase. *J Biol Chem.* **204**:265-81.
- Seitz, S., Lee, S. J., Pennetier, C., Boos, W. et Plumbridge, J. 2003. Analysis of the interaction between the global regulator Mlc and EIIB<sup>Glc</sup> of the glucose-specific phosphotransferase system in Escherichia coli. J Biol Chem. 278:10744-51.
- Selinger, D. W., Saxena, R. M., Cheung, K. J., Church, G. M. et Rosenow, C. <u>2003</u>. Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome Res*. <u>13</u>:216-23.
- Senear, A. W. et Steitz, J. A. <u>1976</u>. Site-specific interaction of Qβ host factor and ribosomal protein S1 with Qβ and R17 bacteriophage RNAs. *J Biol Chem.* **251**:1902-12.
- Seok, Y. J., Sun, J., Kaback, H. R. et Peterkofsky, A. 1997a. Topology of allosteric regulation of lactose permease. Proc Natl Acad Sci U S A. 94:13515-9.
- Seok, Y. J., Sondej, M., Badawi, P., Lewis, M. S., Briggs, M. C., Jaffe, H. et Peterkofsky, A. 1997b. High affinity binding and allosteric regulation of Escherichia coli glycogen phosphorylase by the histidine phosphocarrier protein, HPr. J Biol Chem. 272:26511-21.
- Serganov, A., Ennifar, E., Portier, C., Ehresmann, B. et Ehresmann, C. 2002. Do mRNA and rRNA binding sites of *E.coli* ribosomal protein S15 share common structural determinants? *J Mol Biol.* 320:963-78.
- Shean, C. S. et Gottesman, M. E. <u>1992</u>. Translation of the prophage lambda cl transcript. *Cell*. **70**:513-22.
- Shin, D., Cho, N., Heu, S. et Ryu, S. <u>2003</u>. Selective regulation of *ptsG* expression by Fis. Formation of either activating or repressing nucleoprotein complex in response to glucose. *J Biol Chem*. **278**:14776-81.
- Shingler, V. <u>1996</u>. Signal sensing by σ<sup>54</sup>-dependent regulators: derepression as a control mechanism. *Mol Microbiol*. **19**:409-16.

- Siebers, B., Brinkmann, H., Dorr, C., Tjaden, B., Lilie, H., Van Der Oost, J. et Verhees, C. H. 2001. Archaeal fructose-1,6-bisphosphate aldolases constitute a new family of archaeal type class I aldolase. J Biol Chem. 276:28710-8.
- Siebold, C., Flukiger, K., Beutler, R. et Erni, B. 2001. Carbohydrate transporters of the bacterial phosphoenolpyruvate: sugar phosphotransferase system (PTS). FEBS Lett. 504:104-11.
- Sirover, M. A. 1999. New insights into an old protein: the functional diversity of mammalian glyceraldehyde-3-phosphate dehydrogenase. *Biochim Biophys Acta*. 1432:159-84.
- Skarzynski, T., Moody, P. C. et Wonacott, A. J. <u>1987</u>. Structure of holo-glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* at 1.8 A resolution. *J. Mol. Biol.* **193**:171-87.
- Skordalakes, E. et Berger, J. M. <u>2003</u>. Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell*. **114**:135-46.
- Soderbom, F. et Wagner, E. G. <u>1998</u>. Degradation pathway of CopA, the antisense RNA that controls replication of plasmid R1. *Microbiology*. **144** ( Pt 7):1907-17.
- Sondej, M., Sun, J., Seok, Y. J., Kaback, H. R. et Peterkofsky, A. 1999. Deduction of consensus binding sequences on proteins that bind IIA<sup>Glc</sup> of the phosphoenolpyruvate:sugar phosphotransferase system by cysteine scanning mutagenesis of Escherichia coli lactose permease. Proc Natl Acad Sci U S A. 96:3525-30.
- Sondej, M., Vazquez-Ibar, J. L., Farshidi, A., Peterkofsky, A. et Kaback, H. R. <u>2003</u>. Characterization of a lactose permease mutant that binds IIA<sup>Glc</sup> in the absence of ligand. *Biochemistry*. **42**:9153-9.
- Sorensen, H. P. et Mortensen, K. K. <u>2005.</u> Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J Biotechnol*. 115:113-28.
- Sorensen, M. A., Kurland, C. G. et Pedersen, S. <u>1989.</u> Codon usage determines translation rate in *Escherichia coli. J Mol Biol.* **207**:365-77.
- Spickler, C. et Mackie, G. A. <u>2000</u>. Action of RNase II and polynucleotide phosphorylase against RNAs containing stem-loops of defined structure. *J Bacteriol*. **182**:2422-7.
- Spickler, C., Stronge, V. et Mackie, G. A. <u>2001</u>. Preferential cleavage of degradative intermediates of rpsT mRNA by the *Escherichia coli* RNA degradosome. *J Bacteriol*. **183**:1106-9.
- Spira, B. et Yagil, E. <u>1998</u>. The relation between ppGpp and the PHO regulon in *Escherichia coli*. *Mol Gen Genet*. **257**:469-77.
- Sprengart, M. L., Fatscher, H. P. et Fuchs, E. <u>1990</u>. The initiation of translation in *E. coli*: apparent base pairing between the 16srRNA and downstream sequences of the mRNA. *Nucleic Acids Res.* **18**:1719-23.
- Sprenger, G. A. 1995. Genetics of pentose-phosphate pathway enzymes of *Escherichia coli* K-12. Arch Microbiol. 164:324-30.
- Steege, D. A. <u>2000</u>. Emerging features of mRNA decay in bacteria. RNA. 6:1079-90.

- Stenstrom, C. M., Jin, H., Major, L. L., Tate, W. P. et Isaksson, L. A. 2001. Codon bias at the 3'-side of the initiation codon is correlated with translation initiation efficiency in *Escherichia coli*. Gene. 263:273-84.
- Stokes, N. R., Murray, H. D., Subramaniam, C., Gourse, R. L., Louis, P., Bartlett, W., Miller, S. et Booth, I. R. 2003. A role for mechanosensitive channels in survival of stationary phase: regulation of channel expression by RpoS. Proc Natl Acad Sci U S A. 100:15959-64.
- Storz, G., Opdyke, J. A. et Zhang, A. <u>2004</u>. Controlling mRNA stability and translation with small, noncoding RNAs. *Curr Opin Microbiol*. 7:140-4.
- Stulke, J. et Hillen, W. 1998. Coupling physiology and gene regulation in bacteria: the phosphotransferase sugar uptake system delivers the signals. *Naturwissenschaften.* 85:583-92.
- Subbarayan, P. R. et Deutscher, M. P. <u>2001</u>. Escherichia coli RNase M is a multiply altered form of RNase I. RNA. 7:1702-7.
- Sun, W. et Nicholson, A. W. 2001. Mechanism of action of *Escherichia coli* ribonuclease III. Stringent chemical requirement for the glutamic acid 117 side chain and Mn<sup>2+</sup> rescue of the Glu117Asp mutant. *Biochemistry.* 40:5102-10.
- Sun, W., Jun, E. et Nicholson, A. W. <u>2001</u>. Intrinsic double-stranded-RNA processing activity of *Escherichia coli* ribonuclease III lacking the dsRNA-binding domain. *Biochemistry*. **40**:14976-84.
- Sun, W., Li, G. et Nicholson, A. W. <u>2004</u>. Mutational Analysis of the Nuclease Domain of *Escherichia* coli Ribonuclease III. Identification of Conserved Acidic Residues that Are Important for Catalytic Function in Vitro. Biochemistry. **43**:13054-62.
- Sunohara, T., Abo, T., Inada, T. et Aiba, H. <u>2002</u>. The C-terminal amino acid sequence of nascent peptide is a major determinant of SsrA tagging at all three stop codons. *RNA*. 8:1416-27.
- Symmons, M. F., Williams, M. G., Luisi, B. F., Jones, G. H. et Carpousis, A. J. <u>2002</u>. Running rings around RNA: a superfamily of phosphate-dependent RNases. *Trends Biochem Sci.* **27**:11-8.

# т

- Tabor, S. et Richardson, C. C. <u>1987.</u> DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA.* **84**:4767-71
- Takada, A., Wachi, M., Kaidow, A., Takamura, M. et Nagai, K. 1997. DNA binding properties of the hfq gene product of Escherichia coli. Biochem Biophys Res Commun. 236:576-9.
- Takahashi, H., Inada, T., Postma, P. et Aiba, H. 1998.
  CRP down-regulates adenylate cyclase activity by reducing the level of phosphorylated IIA<sup>Glc</sup>, the glucose-specific phosphotransferase protein, in Escherichia coli. Mol Gen Genet. 259:317-26.
- Takeda, S., Matsushika, A. et Mizuno, T. <u>1999.</u> Repression of the gene encoding succinate dehydrogenase in response to glucose is mediated

- by the EIICB<sup>Glc</sup> protein in *Escherichia coli*. *J Biochem (Tokyo)*. **126**:354-60.
- Tanaka, K., Kusano, S., Fujita, N., Ishihama, A. et Takahashi, H. 1995. Promoter determinants for Escherichia coli RNA polymerase holoenzyme containing σ<sup>38</sup> (the rpoS gene product). Nucleic Acids Res. 23:827-34.
- Tanaka, Y., Itoh, F., Kimata, K. et Aiba, H. 2004. Membrane localization itself but not binding to IICB is directly responsible for the inactivation of the global repressor Mlc in *Escherichia coli*. Mol Microbiol. 53:941-51.
- Taraseviciene, L., Bjork, G. R. et Uhlin, B. E. 1995. Evidence for an RNA binding region in the Escherichia coli processing endoribonuclease RNase E. J Biol Chem. 270:26391-8.
- Tchufistova, L. S., Komarova, A. V. et Boni, I. V. <u>2003</u>. A key role for the mRNA leader structure in translational control of ribosomal protein S1 synthesis in γ-proteobacteria. *Nucleic Acids Res.* **31**:6996-7002.
- Tebbutt, J., Rhodius, V. A., Webster, C. L. et Busby, S. J. 2002. Architectural requirements for optimal activation by tandem CRP molecules at a class I CRP-dependent promoter. FEMS Microbiol Lett. 210:55-60.
- Tedin, K., Moll, I., Grill, S., Resch, A., Graschopf, A., Gualerzi, C. O. et Blasi, U. <u>1999</u>. Translation initiation factor 3 antagonizes authentic start codon selection on leaderless mRNAs. *Mol Microbiol*. 31:67-77.
- Thomson, G. J., Howlett, G. J., Ashcroft, A. E. et Berry, A. 1998. The *dhnA* gene of *Escherichia coli* encodes a class I fructose bisphosphate aldolase. *Biochem J.* 331 ( Pt 2):437-45.
- Thouvenot, B., Charpentier, B. et Branlant, C. <u>2004</u>. The strong efficiency of the *Escherichia coli gapA* P1 promoter depends on a complex combination of functional determinants. *Biochem. J.* **383**:371-82.
- Tjaden, B., Haynor, D. R., Stolyar, S., Rosenow, C. et Kolker, E. <u>2002a</u>. Identifying operons and untranslated regions of transcripts using *Escherichia* coli RNA expression analysis. *Bioinformatics*. **18** Suppl 1:S337-44.
- Tjaden, B., Saxena, R. M., Stolyar, S., Haynor, D. R., Kolker, E. et Rosenow, C. <u>2002b</u>. Transcriptome analysis of *Escherichia coli* using high-density oligonucleotide probe arrays. *Nucleic Acids Res*. 30:3732-8.
- Tock, M. R., Walsh, A. P., Carroll, G. et Mcdowall, K. J. 2000. The *CafA* protein required for the 5'-maturation of 16 S rRNA is a 5'-end-dependent ribonuclease that has context-dependent broad sequence specificity. *J Biol Chem.* 275:8726-32.
- Torres-Larios, A., Dock-Bregeon, A. C., Romby, P., Rees, B., Sankaranarayanan, R., Caillet, J., Springer, M., Ehresmann, C., Ehresmann, B. et Moras, D. 2002. Structural basis of translational control by Escherichia coli threonyl tRNA synthetase. Nat Struct Biol. 9:343-7.
- Tsai, H. Y., Masquida, B., Biswas, R., Westhof, E. et Gopalan, V. 2003. Molecular modeling of the threedimensional structure of the bacterial RNase P holoenzyme. J Mol Biol. 325:661-75.

- Tsui, H. C., Leung, H. C. et Winkler, M. E. 1994. Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in Escherichia coli K-12. Mol Microbiol. 13:35-49.
- Tsui, H. C., Feng, G. et Winkler, M. E. 1997. Negative regulation of *mutS* and *mutH* repair gene expression by the Hfq and RpoS global regulators of *Escherichia coli* K-12. *J Bacteriol.* 179:7476-87.
- Typas, A. et Hengge, R. <u>2005</u>. Differential ability of σ<sup>S</sup> and σ<sup>70</sup> of *Escherichia coli* to utilize promoters containing half or full UP-element sites. *Mol Microbiol*. **55**:250-60.

# U

- Ueguchi, C., Misonou, N. et Mizuno, T. <u>2001</u>. Negative control of *rpoS* expression by phosphoenolpyruvate: carbohydrate phosphotransferase system in *Escherichia coli*. *J Bacteriol*. **183**:520-7.
- Umitsuki, G., Wachi, M., Takada, A., Hikichi, T. et Nagai, K. 2001. Involvement of RNase G in *in vivo* mRNA metabolism in *Escherichia coli*. *Genes Cells*. 6:403-10.

# ٧

- Valentin-Hansen, P., Eriksen, M. et Udesen, C. 2004. The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Mol Microbiol.* 51:1525-33.
- Van Den Bergh, E. R., Baker, S. C., Raggers, R. J., Terpstra, P., Woudstra, E. C., Dijkhuizen, L. et Meijer, W. G. 1996. Primary structure and phylogeny of the Calvin cycle enzymes transketolase and fructosebisphosphate aldolase of Xanthobacter flavus. J. Bacteriol. 178:888-93.
- Van Meerten, D., Girard, G. et Van Duin, J. <u>2001</u>. Translational control by delayed RNA folding: identification of the kinetic trap. RNA. 7:483-94.
- Vanderpool, C. K. et Gottesman, S. <u>2004</u>. Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol Microbiol*. **54**:1076-89.
- Vanzo, N. F., Li, Y. S., Py, B., Blum, E., Higgins, C. F., Raynal, L. C., Krisch, H. M. et Carpousis, A. J. 1998. Ribonuclease E organizes the protein interactions in the *Escherichia coli* RNA degradosome. *Genes Dev.* 12:2770-81.
- Vasil'eva Iu, M. et Garber, M. B. <u>2002.</u> [The regulatory role of the Hfq protein in bacterial cells]. *Mol Biol* (*Mosk*). **36**:970-7.
- Vassilieva, I. M., Rouzanov, M. V., Zelinskaya, N. V., Moll, I., Blasi, U. et Garber, M. B. <u>2002</u>. Cloning, purification, and crystallization of a bacterial gene expression regulator--Hfq protein from *Escherichia* coli. Biochemistry (Mosc). **67**:1293-7.
- Vassylyev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S. et Yokoyama, S. 2002. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature*. 417:712-9.
- Vecerek, B., Moll, I., Afonyushkin, T., Kaberdin, V. et Blasi, U. <u>2003</u>. Interaction of the RNA chaperone

- Hfq with mRNAs: direct and indirect roles of Hfq in iron metabolism of *Escherichia coli*. *Mol Microbiol*. **50**:897-909.
- Voges, D., Watzele, M., Nemetz, C., Wizemann, S. et Buchberger, B. 2004. Analyzing and enhancing mRNA translational efficiency in an Escherichia coli in vitro expression system. Biochem Biophys Res Commun. 318:601-14.
- Von Hippel, P. H. et Pasman, Z. <u>2002</u>. Reaction pathways in transcript elongation. *Biophys Chem*. 101-102:401-23.
- Vytvytska, O., Jakobsen, J. S., Balcunaite, G., Andersen, J. S., Baccarini, M. et Von Gabain, A. 1998. Host factor I, Hfq, binds to Escherichia coli ompA mRNA in a growth rate-dependent fashion and regulates its stability. Proc Natl Acad Sci U S A. 95:14118-23.
- Vytvytska, O., Moll, I., Kaberdin, V. R., Von Gabain, A. et Blasi, U. <u>2000</u>. Hfq (HF1) stimulates *ompA* mRNA decay by interfering with ribosome binding. *Genes Dev.* **14**:1109-18.

# W

- Wachi, M., Umitsuki, G. et Nagai, K. <u>1997</u>. Functional relationship between *Escherichia coli* RNase E and the CafA protein. *Mol Gen Genet*. **253**:515-9.
- Wachi, M., Umitsuki, G., Shimizu, M., Takada, A. et Nagai, K. 1999. Escherichia coli cafA gene encodes a novel RNase, designated as RNase G, involved in processing of the 5' end of 16S rRNA. Biochem Biophys Res Commun. 259:483-8.
- Walsh, A. P., Tock, M. R., Mallen, M. H., Kaberdin, V. R., Von Gabain, A. et Mcdowall, K. J. 2001. Cleavage of poly(A) tails on the 3'-end of RNA by ribonuclease E of *Escherichia coli. Nucleic Acids Res.* 29:1864-71.
- Wang, G., Peterkofsky, A. et Clore, G. M. <u>2000a.</u> A novel membrane anchor function for the N-terminal amphipathic sequence of the signal-transducing protein IIA Glucose of the Escherichia coli phosphotransferase system. J Biol Chem. <u>275</u>:39811-4.
- Wang, G., Sondej, M., Garrett, D. S., Peterkofsky, A. et Clore, G. M. <u>2000b.</u> A common interface on histidine-containing phosphocarrier protein for interaction with its partner proteins. *J Biol Chem.* 275:16401-3.
- Wang, G., Louis, J. M., Sondej, M., Seok, Y. J., Peterkofsky, A. et Clore, G. M. 2000c. Solution structure of the phosphoryl transfer complex between the signal transducing proteins HPr and IIA glucose of the Escherichia coli phosphoenolpyruvate:sugar phosphotransferase system. EMBO J. 19:5635-49.
- Wang, J. Y. et Syvanen, M. <u>1992.</u> DNA twist as a transcriptional sensor for environmental changes. *Mol Microbiol.* **6**:1861-6.
- Wang, M. et Cohen, S. N. <u>1994</u>. ard-1: a human gene that reverses the effects of temperature-sensitive and deletion mutations in the *Escherichia coli rne* gene and encodes an activity producing RNase E-like cleavages. *Proc Natl Acad Sci U S A*. **91**:10591-5.

- Weickert, M. J., Doherty, D. H., Best, E. A. et Olins, P. O. 1996. Optimization of heterologous protein production in *Escherichia coli. Curr. Opin. Biotechnol.* 7:494-9.
- West, A. H. et Stock, A. M. <u>2001</u>. Histidine kinases and response regulator proteins in two-component signaling systems. *Trends Biochem Sci.* **26**:369-76.
- Weyens, G., Charlier, D., Roovers, M., Pierard, A. et Glansdorff, N. 1988. On the role of the Shine-Dalgarno sequence in determining the efficiency of translation initiation at a weak start codon in the car operon of Escherichia coli K12. J Mol Biol. 204:1045-8.
- Will, C. L. et Luhrmann, R. <u>2001.</u> Spliceosomal UsnRNP biogenesis, structure and function. *Curr Opin Cell Biol.* 13:290-301.
- Winkler, W., Nahvi, A. et Breaker, R. R. 2002. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature*. 419:952-6.
- Winkler, W. C. et Breaker, R. R. <u>2003</u>. Genetic control by metabolite-binding riboswitches. *Chembiochem*. **4**:1024-32.
- Worhunsky, D. J., Godek, K., Litsch, S. et Schlax, P. J. 2003. Interactions of the non-coding RNA DsrA and RpoS mRNA with the 30 S ribosomal subunit. *J Biol Chem.* 278:15815-24.
- Wosten, M. M. 1998. Eubacterial sigma-factors. FEMS Microbiol Rev. 22:127-50.
- Wrobel, B., Herman-Antosiewicz, A., Szalewska-Palasz, S. et Wegrzyn, G. <u>1998</u>. Polyadenylation of *oop* RNA in the regulation of bacteriophage lambda development. *Gene*. 212:57-65.
- Wu, X. Q., Iyengar, P. et Rajbhandary, U. L. 1996. Ribosome-initiator tRNA complex as an intermediate in translation initiation in *Escherichia* coli revealed by use of mutant initiator tRNAs and specialized ribosomes. *EMBO J.* 15:4734-9.
- Wulczyn, F. G. et Kahmann, R. 1991. Translational stimulation: RNA sequence and structure requirements for binding of Com protein. Cell. 65:259-69.

# X

- Xu, F., Lin-Chao, S. et Cohen, S. N. 1993. The Escherichia coli pcnB gene promotes adenylylation of antisense RNAI of ColE1-type plasmids in vivo and degradation of RNAI decay intermediates. Proc Natl Acad Sci U S A. 90:6756-60.
- Xu, F. et Cohen, S. N. 1995. RNA degradation in Escherichia coli regulated by 3' adenylation and 5' phosphorylation. Nature. 374:180-3.
- Xu, F. F., Gaggero, C. et Cohen, S. N. 2002. Polyadenylation can regulate ColE1 type plasmid copy number independently of any effect on RNAI decay by decreasing the interaction of antisense RNAI with its RNAII target. Plasmid. 48:49-58.

# Y

Yamanaka, K. et Inouye, M. <u>2001</u>. Selective mRNA degradation by polynucleotide phosphorylase in

- cold shock adaptation in Escherichia coli. J Bacteriol. 183:2808-16.
- Yang, Y., Zhao, G., Man, T. K. et Winkler, M. E. 1998. Involvement of the gapA- and epd (gapB)-encoded dehydrogenases in pyridoxal 5'-phosphate coenzyme biosynthesis in Escherichia coli K-12. J Bacteriol. 180:4294-9.
- Yanisch-Perron, C., Vieira, J. et Messing, J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*. 33:103-19.
- Yarchuk, O., Jacques, N., Guillerez, J. et Dreyfus, M. 1992. Interdependence of translation, transcription and mRNA degradation in the *lacZ* gene. *J Mol Biol.* 226:581-96.
- Yoon, S. H., Han, M. J., Lee, S. Y., Jeong, K. J. et Yoo, J. S. <u>2003</u>. Combined transcriptome and proteome analysis of *Escherichia coli* during high cell density culture. *Biotechnol Bioeng*. **81**:753-67.
- Young, B. A., Gruber, T. M. et Gross, C. A. <u>2002</u>. Views of transcription initiation. *Cell*. **109**:417-20.
- Young, B. A., Gruber, T. M. et Gross, C. A. <u>2004</u>. Minimal machinery of RNA polymerase holoenzyme sufficient for promoter melting. *Science*. 303:1382-4.
- Yu, D. et Deutscher, M. P. <u>1995.</u> Oligoribonuclease is distinct from the other known exoribonucleases of *Escherichia coli. J Bacteriol.* 177:4137-9.
- Yu, J. S., Madison-Antenucci, S. et Steege, D. A. 2001. Translation at higher than an optimal level interferes with coupling at an intercistronic junction. *Mol Microbiol.* 42:821-34.
- Yun, M., Park, C. G., Kim, J. Y. et Park, H. W. 2000. Structural analysis of glyceraldehyde 3-phosphate dehydrogenase from *Escherichia coli*: direct evidence of substrate binding and cofactor-induced conformational changes. *Biochemistry*. 39:10702-10
- Yusupova, G. Z., Yusupov, M. M., Cate, J. H. et Noller, H. F. <u>2001</u>. The path of messenger RNA through the ribosome. *Cell*. **106**:233-41.

# Z

- Zamore, P. D. 2001. Thirty-three years later, a glimpse at the ribonuclease III active site. Mol Cell. 8:1158-60.
- Zgiby, S., Plater, A. R., Bates, M. A., Thomson, G. J. et Berry, A. 2002. A functional role for a flexible loop containing Glu182 in the class II fructose-1,6bisphosphate aldolase from Escherichia coli. J Mol Biol. 315:131-40.
- Zgiby, S. M., Thomson, G. J., Qamar, S. et Berry, A. 2000. Exploring substrate binding and discrimination in fructose1, 6-bisphosphate and tagatose 1,6-bisphosphate aldolases. Eur J Biochem. 267:1858-68.
- Zhang, A., Altuvia, S., Tiwari, A., Argaman, L., Hengge-Aronis, R. et Storz, G. 1998a. The OxyS regulatory RNA represses rpoS translation and binds the Hfq (HF-I) protein. EMBO J. 17:6061-8.
- Zhang, A., Wassarman, K. M., Ortega, J., Steven, A. C. et Storz, G. 2002. The Sm-like Hfq protein increases

- OxyS RNA interaction with target mRNAs. Mol Cell. 9:11-22.
- Zhang, K. et Nicholson, A. W. 1997. Regulation of ribonuclease III processing by double-helical sequence antideterminants. Proc Natl Acad Sci U S A. 94:13437-41.
- Zhang, X. et Bremer, H. <u>1996</u>. Effects of Fis on ribosome synthesis and activity and on rRNA promoter activities in *Escherichia coli*. *J Mol Biol*. 259:27-40.
- Zhang, X., Zhu, L. et Deutscher, M. P. 1998b. Oligoribonuclease is encoded by a highly conserved gene in the 3'-5' exonuclease superfamily. J Bacteriol. 180:2779-81.
- Zhao, G., Pease, A. J., Bharani, N. et Winkler, M. E. 1995. Biochemical characterization of gapBencoded erythrose 4-phosphate dehydrogenase of Escherichia coli K-12 and its possible role in pyridoxal 5'-phosphate biosynthesis. J Bacteriol. 177:2804-12.
- Zheng, D., Constantinidou, C., Hobman, J. L. et Minchin, S. D. <u>2004</u>. Identification of the CRP regulon using *in vitro* and *in vivo* transcriptional profiling. *Nucleic Acids Res.* 32:5874-93.
- Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31:3406-15.
- Zuo, Y. et Deutscher, M. P. <u>2001</u>. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Res.* **29**:1017-26.

#### Titre:

Etude de l'organisation transcriptionnelle et de la régulation post-transcriptionnelle du groupe d'ORF codant l'Erythrose 4-phosphate déshydrogénase, la Phosphoglycérate kinase et la Fructose-1,6-bisphosphate aldolase de classe II chez *Escherichia coli*, applications à la production de protéines recombinantes.

# Résumé:

Le triplet d'ORF Erythrose-4-phosphate déshydrogénase (E4PDH)-phosphoglycérate kinase (PGK)-fructose-1,6-bisphosphate aldolase (FBA), correspondant à trois enzymes intervenant dans le métabolisme du carbone, est retrouvé uniquement dans une partie des γ-protéobactéries. Nous avons montré que, chez *Escherichia coli*, ce triplet d'ORF forme un opéron comportant un promoteur fortement régulé *epd* P0 et un promoteur étendu *pgk* P1. Les différents transcrits obtenus sont ensuite maturés par la RNase E au niveau des régions *epd-pgk* et *pgk-fbaA*, ce qui conduit à une stabilité différentielle des différents transcrits, avec par ordre croissant de demie-vie *epd*, *pgk* et *fbaA*. Par des expériences en fermentation continue, nous avons montré que le taux d'expression des ORF *epd*, *pgk*, *fbaA* et *gapA* augmente lorsque le taux de croissance augmente. La faible production d'acide acétique par les bactéries possédant un transporteur EIIBC<sup>Glc</sup> inactif nous a conduit à tester si ces-bactéries pouvaient être intéressantes pour la production de protéines recombinantes. Nous avons mis au point des conditions permettant d'obtenir une forte densité cellulaire et une forte production de protéines recombinantes.

#### Mots clés :

Escherichia coli, phosphoglycérate kinase (PGK), classe II fructose-1,6-bisphosphate aldolase (FBA), structure secondaire des ARN, RNase E, fermentation continue, surproduction de protéines recombinantes.

#### Title:

Study on the transcriptional organization and post-transcriptional regulation of the Erythrose 4-phoshate dehydrogenase, Phosphoglycerate kinase and class II Fructose 1,6-bisphosphate aldolase (FBA) ORF cluster in *Escherichia coli*, applications to recombinant protein production

#### Summary:

The Erythrose-4-phosphate déshydrogénase (E4PDH)-phosphoglycerate kinase (PGK)-fructose-1,6-bisphosphate aldolase (FBA) ORF cluster, corresponding to three enzymes from carbon metabolism, is found in a part of the γ-proteobacteria. We showed that, in *Escherichia coli*, this ORF cluster is an operon with one strong regulated promoter *epd* P0 and one extended promoter *pgk* P1. The different transcripts are maturated by RNase E, in *epd-pgk* and *pgk-fbaA* region. This result in the differential stability of the transcripts, with by increasing half-life, *epd*, *pgk* and *fbaA*. By continuous fermentation, we showed that *epd*, *pgk*, *fbaA* and *gapA* expression increases with the growth rate. The low production of acetic acid form by strains with an inactivated transport protein EIIBC<sup>Glc</sup> led us to test whether they can be used for recombinant protein production. We developed conditions to get high cellular density and high level of recombinant protein production.

#### Key words:

Escherichia coli, phosphoglycerate kinase (PGK), class II fructose-1,6-bisphosphate aldolase (FBA), secondary structure of RNA, RNase E, continuous fermentation, recombinant protein production.