HAL
open science

Etude fonctionnelle du génome de Bacillus subtilis : de nouvelles régulations transcriptionnelles du métabolisme central du carbone.

Thierry Doan

To cite this version:

Thierry Doan. Etude fonctionnelle du génome de Bacillus subtilis : de nouvelles régulations transcriptionnelles du métabolisme central du carbone.. Life Sciences [q-bio]. INAPG (AgroParisTech), 2003. English. NNT : . pastel-00001193

HAL Id: pastel-00001193

https://pastel.hal.science/pastel-00001193v1

Submitted on 11 Apr 2005

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.

INSTITUT NATIONAL AGRONOMIQUE PARIS-GRIGNON

Ecole Doctorale ABIES

(Biologie)

THÈSE

pour obtenir le grade de

Docteur de l'Institut National Agronomique Paris-Grignon

Discipline : Génétique

présentée par

Thierry DOAN

Etude fonctionnelle du génome de *Bacillus subtilis* : de nouvelles régulations transcriptionnelles du métabolisme central du carbone.

Soutenue le : 25 avril 2003, devant le jury composé de

Pr. Claude Gaillardin	Président
Dr. Philippe Glaser	rapporteur
Dr. Jorg Stülke	rapporteur
Dr. Frédéric Barras	Examinateur
Dr. Stéphane Aymerich	Directeur de thèse

© UMLV

WANTED
Fructose
Glucose
Malate
Succinate
Ribose
Bacillus subtilis

Remerciements

Ce travail a été réalisé dans le Laboratoire de Génétique Moléculaire et Cellulaire à Thiverval-Grignon (INRA – CNRS – INAP-G) dirigé par le Professeur Claude Gaillardin. Je souhaite lui témoigner ma gratitude. Je le remercie aussi d'avoir accepté de présider le jury de ma thèse.

Je tiens à te remercier très sincèrement, Stéphane, de m'avoir accueilli et véritablement élevé (scientifiquement parlant !) au sein de l'équipe *Bacillus subtilis* du laboratoire. Cela n'a certainement pas été facile au début, moi qui n'avais même pas coulé une seule boîte de Pétri auparavant !

Je suis très honoré, M. Glaser, M. Stülke et M. Barras, que vous ayez acceptés d'être les rapporteurs et examinateur de cette thèse.

Je remercie Georg Homuth et Michael Hecker pour le chaleureux accueil qu'ils m'ont offert pendant mon séjour à Greifswald.

Je remercie aussi Harald Putzer et Ciaran Condon pour leur gentillesse et l'aide technique qu'ils m'ont apportée.

Je remercie chaleureusement Nathalie et Pascale. Que saurais-je faire d'autre que de la génétique classique (merci Stéphane !) si, à présent, j'ignorais ce que vous m'avez enseigné ? Merci, Pascale, d'avoir partagé ta bonne humeur et ton sens du rangement.

Merci Dominique pour le temps que tu m'as consacré. Merci pour tes bonnes idées et tes critiques pertinentes.

Je remercie l'ensemble des membres du laboratoire, à qui je dois l'ambiance sympathique dans laquelle j'ai travaillé pendant ces trois années. Merci à l'équipe « Bacillacées » pour tous les échanges fructueux : Josef, Sandrine, Annie, Alain, Emmanuel, Grégory et Ivan. Bonne continuation Guillaume ! Merci à mes proches voisins, Brigitte, Marie-Thérèse, Solange, Jean-Marc et Michel, chez qui je venais souvent demander un petit quelque chose. Je te remercie particulièrement Cécile pour l'aide en bioinformatique et la radioactivité (et aussi pour les parties de Age avec Anna, Mathias et Mehdi). Un grand merci à Denise et Guy, Anne, Béatrice, Caroline, Delphine et Kounvary pour l'indispensable « logistique ».

Maman, papa, merci de m'avoir soutenu tout au long de mes études et au cours de cette thèse. Je sais combien cela a été difficile de rebâtir une vie en ayant ses mains nues comme seule fortune. Merci Annick, ma tendre compagne, pour ton soutien et ta patience. Enfin, je n'oublie pas de remercier Adrien, Antoine, Hélène, Lionel, Rachel, et Yannick pour leur amitié et leur aide.

Et il reste de la place pour celles et ceux que je n'ai pas mentionnés !

Résumé

Chez *Bacillus subtilis*, la transcription de l'opéron *gapA*, comprenant les gènes de la partie centrale de la glycolyse, est stimulée en présence de sources de carbone glycolytiques. Nos études *in vivo et in vitro* de CggR, le répresseur qui contrôle cette stimulation, ont démontré d'une part que celui-ci a la capacité de se lier à une séquence d'ADN inhabituellement longue, consistant en une répétition directe de deux motifs ($CGGGACN_6TGTCN_4CGGGACN_6TGTC$) et située entre le promoteur et le codon d'initiation de l'opéron *cggR-gapA*, et d'autre part que son activité est inhibée par le fructose-1,6-biphosphate. Des analyses de séquence et des expériences de transcriptome ont indiqué que CggR, qui est très conservé chez les bactéries à Gram positif et qui définit une sous-famille de la famille de régulateurs transcriptionnels SorC/DeoR, est spécialisé dans le contrôle des gènes de la glycolyse. Ainsi, une collaboration a été engagée avec des structuralistes (CBS, Montpellier) pour aller plus loin dans la connaissance de ce prototype d'une famille encore peu connue de régulateurs.

Deux paires de gènes paralogues ont été détectés dans le génome (*ywkA* et *malS*, *ytsJ* et *mleA*) dont les produits sont homologues à des enzymes maliques. L'analyse transcriptomique globale d'une souche sauvage cultivée en présence de glucose ou de malate comme seule source de carbone montre que l'expression d'*ywkA* est induite en présence de malate. En collaboration avec l'équipe de Y. Fujita, nous avons montré qu'*ywkA* codait bien une enzyme malique NAD-dépendante dont l'expression est spécifiquement induite par le malate extracellulaire et insensible à la répression catabolique. De plus, nous avons montré que le système à deux composants YufL-YufM active directement la transcription d'*ywkA* en présence de malate. Cependant, YwkA n'est pas requis pour la croissance en présence de malate comme seule source de carbone.

La technique d'analyse du transcriptome au moyen de membranes à haute densité est maintenant acquise au laboratoire. Nous avons commencé à mettre à profit cet outil pour une étude globale de l'expression génique en fonction de différentes sources de carbone.

Mots clés : *Bacillus subtilis*, répression/activation de la transcription, métabolisme carboné, glycolyse, cycle de Krebs, voie des pentoses-phosphate, analyse transcriptomique, CggR, Fructose-1,6-Biphosphate, enzyme malique, *ywkA*.

Abstract

In *Bacillus* subtilis, the transcription of the *gapA* operon, including the genes of the central part of glycolysis, is stimulated during growth on glycolytic carbon sources. Our *in vivo* and *in vitro* studies of CggR, the repressor that controls this stimulation, have demonstrated on one hand that it is able to bind a long single DNA sequence made of two directly repeated motifs ($CGGGACN_6TGTCN_4CGGGACN_6TGTC$) and located between the promoter and the translation codon of the *cggr-gapA* operon and on the other hand that its activity is modulated by fructose-1,6-bisphosphate. Sequence analysis and array hybridization indicated that CggR, which is highly conserved in Gram positive bacteria and defines a sub-family of the SorC/DeoR family of transcriptional regulators, is dedicated to the control of the glycolytic genes. A collaboration with structural biologists (CBS, Montpellier) has been engaged to further study this prototype of a not well known sub-family of transcriptional regulators.

Two groups of paralogous genes have been detected on *Bacillus subtilis* chromosome (*ywkA* et *malS*, *ytsJ* and *mleA*) which products are homologous to malic enzymes. Global transcriptome analysis of a wild-type strain in response to growth in the presence of glucose or malate as sole carbon source has revealed that *ywkA* expression is induced in the presence of malate. Our experiments have demonstrated that *ywkA* encodes a NAD-dependent malic enzyme and that its expression is specifically induced by malate and is not subjected to carbon catabolite repression. Moreover, the YufL/YufM two component system has been demonstrated to directly activate *ywkA* transcription in response to the presence of malate. However, YwkA is not required for efficient growth on malate as sole carbon source.

Actually, transcriptome study using macroarray technique is available for our laboratory. Hence, we have started the global comparison of genes expression profile in response to different carbon sources.

Key Words: *Bacillus subtilis*, transcription repression/activation, carbon metabolism, glycolysis, Krebs cycle, pentoses phosphate pathway, transcriptome analysis, CggR, Fructose-1,6-Bisphosphate, malic enzyme, *ywkA*.

Table des matières

Remerciements ____ *3*

Résumé ____ ***5***

Abstract ____ ***6***

Table des matières ____ ***7***

Tables ____ ***10***

Figures ____ ***11***

Abréviations ____ ***14***

Introduction bibliographique ____ ***17***

1. Les voies centrales du métabolisme du carbone chez Bacillus subtilis ____ ***22***

1.1. La glycolyse : une voie universelle ____ **22**

1.1.1. Description générale ____ 22

1.1.2. Les étapes de la glycolyse ____ 23

1.1.3. Comparaison de la glycolyse chez *Bacillus subtilis*, *Lactococcus lactis, Escherichia coli, Helicobacter pylori* et un Archée *Methanobacterium thermoautotrophicum* ____ 34

1.1.4. Le bilan de la glycolyse ____ 35

1.2. La voie des pentoses phosphate ____ **37**

1.2.1. Description générale ____ 37

1.2.2. Les étapes de la voie des pentoses phosphate ____ 37

1.2.3. Comparaison de la voie des pentoses phosphate chez *Bacillus subtilis*, *Lactococcus lactis, Escherichia coli, Helicobacter pylori* et *Methanobacterium thermoautotrophicum* ____ 41

1.3. Le cycle de Krebs ____ **43**

1.3.1. Description générale ____ 43

1.3.2. Les enzymes du cycle de Krebs 44
1.3.3. Organisation comparée du cycle de Krebs chez *Bacillus subtilis*, *Lactococcus lactis*, *Escherichia coli*, *Helicobacter pylori* et *Methanobacterium thermoautotrophicum* 50
1.3.4. Fonctionnement du cycle de Krebs en aérobiose et en anérobiose 51
1.3.5. Rôle du cycle de Krebs dans la sporulation 53

1.4. A la jonction entre la glycolyse/néoglucogenèse et le cycle de Krebs 54
1.4.1. Le complexe pyruvate déshydrogénase (EC 1.2.4.1) *pdhABCD* 54
1.4.2. La pyruvate carboxylase *pycA* (EC 6.4.1.1) 54
1.4.3. La phosphoénolpyruvate carboxykinase *pckA* (EC 4.1.1.49) 55
1.4.4. L'enzyme malique (EC 1.1.1.38-40) 55

2. Les différents niveaux de régulation du métabolisme central du carbone 56

2.1. Un prélude au transport 56

2.2. Le transport : une étape de tri préférentiel des sources de carbone et de transduction du signal 57
2.2.1. Les systèmes bactériens de transport membranaire des sources de carbone 57
2.2.2. Le PTS : un système prédominant de contrôle du transport des glucides doublé d'un système de transduction du signal chez *Bacillus subtilis* 63

2.3. Les régulations de l'expression des gènes impliqués dans le métabolisme du carbone 73
2.3.1. L'organisation génétique des gènes du métabolisme central du carbone (catalytiques et régulateurs) 75
2.3.2. La répression/activation catabolique médiée par CcpA 75
2.3.3. Le rôle direct du PTS dans la régulation des gènes du transport et du catabolisme primaire de sources spécifiques de carbone 88
2.3.4. Les régulations des gènes des voies centrales du métabolisme du carbone indépendantes du PTS (ou indirectement affectées par le PTS) 95

3. Vers l'intégration des données de régulation dans un modèle global à contraintes 97

Présentation du travail 100

Résultats 103

4. CggR : un nouveau régulateur de l'expression des gènes de la partie centrale de la glycolyse ____ *104*

4.1. Présentation synthétique des résultats ____ **104**

4.2. Article n°1 : Regulation of the central glycolytic genes in *Bacillus subtilis*: binding of the repressor CggR to its single DNA target sequence is modulated by fructose-1,6-bisphosphate. ____ **106**

4.2.1. Article ____ 106

4.2.2. "Supplementary material" de l'article ____ 120

4.3. Travaux complémentaires ____ **126**

4.3.1. Résultats ____ 126

4.3.2. Matériel et Méthodes ____ 131

5. Etude transcriptionnelle et physiologique d'ywkA, l'un des quatre gènes paralogues pouvant coder des enzymes maliques chez Bacillus subtilis ____ *134*

5.1. Présentation synthétique des résultats ____ **134**

5.2. Article n°2 : The *Bacillus subtilis ywkA* gene encodes a malic enzyme and its transcription is activated by the YufL-YufM two component system in response to malate. ____ **136**

5.2.1. Article ____ 136

5.2.2. Supplementary Material ____ 150

5.3. Résultats complémentaires ____ **152**

5.4. Matériel et Méthodes ____ **155**

Etude du transcriptome : mise au point de la technique et résultats préliminaires ____ *156*

6. Présentation générale ____ *157*

7. Glycolyse versus Néoglucogenèse : un début d'analyse ____ *160*

7.1. Introduction ____ **160**

7.2. Matériel et Méthodes ____ **160**

7.3. Résultats préliminaires ____ **163**

Discussion générale et perspectives ____ *168*

Bibliographie ____ *193*

Tables

(Ne sont répertoriées ici que les Tables ne faisant pas partie des articles publiés ; le numéro de page indiqué est en fait celui de la page de texte adjacente à la Table).

Table 1 : Systèmes de transport des glucides chez *Bacillus subtilis*. **p. 61**

Table 2 : Séquences *cre* des gènes et opérons chez *Bacillus subtilis* citées dans la revue bibliographique. **p. 76**

Table 3 : Gènes pour lesquels un site *cre* a été identifié ou dont l'expression est sensible à une mutation de *ccpA* ou de *ptsH1* (*crh1*). **p. 76**

Table 4 : Liste des gènes dont l'expression est régulée par le régulateur transcriptionnel Cra. **p. 88**

Table 5 : Régulateurs transcriptionnels contenant des PRD dont l'activité a été étudiée de manière détaillée. **p. 89**

Table 6 : Paramètres cinétiques d'YwkA, de MalS et d'YtsJ. **p. 154**

Table 7 : Exemple de normalisation des données brutes de transcriptome. **p. 163**

Table 8 : Liste incomplète des gènes différentiellement exprimés (signal/bruit>3 et ratio>2) dans les expériences H2 et H3. **p. 165**

Table 9 : Liste des gènes significativement plus exprimés en présence de glucose que de succinate dans l'expérience H3 et croisement de cette liste avec les résultats d'autres comparaisons de ce type. **p. 166**

Table 10 : Liste des gènes significativement plus exprimés en présence de succinate que de glucose dans l'expérience H3 et croisement de cette liste avec les résultats d'autres comparaisons de ce type. **p. 166**

Figures

(Ne sont répertoriées ici que les Figures ne faisant pas partie des articles publiés ; le numéro de page indiqué est en fait celui de la page de texte adjacente à la Figure).

Figure 1 : Les trois voies centrales du métabolisme du carbone chez *Bacillus subtilis*. **p. 22**

Figure 2 : La glycolyse chez *Bacillus subtilis*. **p. 22**

Figure 3 : Comparaison des étapes de la glycolyse chez 5 bactéries dont le génome est disponible. **p. 24**

Figure 4 : La voie des pentoses phosphate chez *Bacillus subtilis*. **p. 37**

Figure 5 : Voie des pentoses phosphate classiquement décrite chez les bactéries lactiques. **p. 42**

Figure 6 : Comparaison des étapes de la voie des pentoses phosphate chez 5 bactéries modèles. **p. 42**

Figure 7 : Le cycle de Krebs chez *Bacillus subtilis*. **p. 43**

Figure 8 : Shunt du glyoxylate chez *Escherichia coli.* **p. 47**

Figure 9 : Comparaison des étapes du cycle de Krebs présentes chez différentes espèces bactériennes. **p. 50**

Figure 10 : Fonctionnement du cycle de Krebs lors de la croissance en anaérobiose chez *Escherichia coli*. **p. 52**

Figure 11 : Les quatre types de système de transport décrits chez les bactéries. **p. 57**

Figure 12 : Les mécanismes de transport secondaires décrits chez les bactéries. **p. 59**

Figure 13 : Organisation générale du PTS chez *Bacillus subtilis.* **p. 63**

Figure 14 : Phosphorylation d'HPr PEP-dépendante par P-EI ou ATP-dépendante par HPrK/P. **p. 65**

Figure 15 : Systèmes de transport PTS chez *Bacillus subtilis.* **p. 67**

Figure 16 : Modèle de la régulation de l'activité de l'Hpr kinase/phosphatase chez *Bacillus subtilis.* **p. 68**

Figure 17 : Modèle d'exclusion de l'inducteur chez les bactéries à Gram négatif. **p. 69**

Figure 18 : Modèle d'exclusion de l'inducteur chez les bactéries à Gram positif. **p. 70**

Figure 19 : Modèle de modulation de l'activité du symporteur lactose par HPr-Ser46-P chez *Lactococcus brevis*. **p. 70**

Figure 20 : Modèle du mécanisme d'expulsion de l'inducteur chez les bactéries à Gram positif. **p. 72**

Figure 21 : phénomène de diauxie observé par Monod dans les années 1940. **p. 73**

Figure 22 : Organisation chromosomique des gènes impliqués dans la glycolyse et la néoglucogenèse chez *Bacillus subtilis*. **p. 75**

Figure 23 : Organisation chromosomique des gènes impliqués dans la voie des pentoses phosphate chez *Bacillus subtilis*. **p. 75**

Figure 24 : Organisation chromosomique des gènes impliqués dans le cycle de Krebs chez *Bacillus subtilis*. **p. 75**

Figure 25 : Organisation chromosomique des gènes codant les enzymes générales du systèmes PTS et le régulateur CcpA chez *Bacillus subtilis*. **p. 75**

Figure 26 : Mécanisme de la répression/activation catabolique médiée par le régulateur transcriptionnel pléïotropique CcpA chez *Bacillus subtilis*. **p. 82**

Figure 27 : Modèle de régulation de l'expression des gènes *citZ* et *citB* par CcpC en présence et en absence de citrate. **p. 83**

Figure 28 : Modèle de régulation des gènes *citZ, citB* et *ccpC* par CcpA et CcpC chez *Bacillus subtilis*. **p. 84**

Figure 29 : Rôles du PTS et de la protéine CAP dans la régulation du métabolisme carboné chez *Escherichia coli*. **p. 86**

Figure 30 : Rôles du PTS et de CcpA dans la régulation du métabolisme carboné chez *Bacillus subtilis*. **p. 87**

Figure 31 : Modèle de répression catabolique impliquant le régulateur transcriptionnel Cra chez *Escherichia coli*. **p. 88**

Figure 32 : Mécanisme général de régulation de l'expression génique médiée par un antiterminateur transcriptionnel. **p. 90**

Figure 33 : Structure secondaire des régulateurs possédant le domaine PRD. **p. 90**

Figure 34 : Modèle de régulation de l'activité de l'antiterminateur LicT par le système PTS chez *Bacillus subtilis*. **p. 91**

Figure 35 : Modèle de régulation de l'expression de l'opéron *ptsGHI* par l'antiterminateur GlcT en réponse à la présence de glucose chez *Bacillus subtilis*. **p. 92**

Figure 36 : Modèle de régulation de l'expression de l'opéron *lev* par l'activateur transcriptionnel LevR chez *Bacillus subtilis*. **p. 93**

Figure 37 : Distribution des flux métaboliques de carbone chez *Bacillus subtilis*. **p. 98**

Figure 38 : Schéma combiné du métabolisme central du carbone et de sa régulation transcriptionnelle chez *Escherichia coli*. **p. 99**

Figure 39 : Détermination du site de démarrage de la transcription en aval de *cggR*. **p. 126**

Figure 40 : Vérification de la structure du mutant *cggR* (GTD030) par Southern Blot. **p. 126**

Figure 41 : Test *in vitro* par titration en retard de migration électrophorétique du rôle de la série conservée de Thymines sur l'interaction CggR/ADN. **p. 127**

Figure 42 : Surproduction et purification de His_6-CggR. **p. 128**

Figure 43 : Détermination des oligomères possibles de CggR par pontage chimique au glutaraldéhyde. **p. 128**

Figure 44 : Test de l'effet de différents métabolites de la glycolyse sur l'interaction CggR/ADN par mesure de variation de résonance plasmonique de surface. **p. 129**

Figure 45 : Effet de la concentration en FBP sur l'interaction CggR/ADN (WT) par mesure de variation de résonance plasmonique de surface. **p. 129**

Figure 46 : Test de l'effet du Fructose-1,6-Biphosphate (FBP) sur l'interaction CggR/ADN en retard de migration électrophorétique. **p. 130**

Figure 47 : L'étape enzyme malique et son contexte métabolique chez *Bacillus subtilis*. **p. 134**

Figure 48 : Organisation génétique des gènes *ywkA*, *malS* et *ytsJ* chez *Bacillus subtilis*. **p. 152**

Figure 49 : Alignement et comparaison des produits des quatre gènes paralogues *ywkA*, *malS*, *ytsJ* et *mleA* chez *Bacillus subtilis* avec SfcA (enzyme malique NAD-dépendante) d'*Escherichia coli* et Maox (enzyme malique NAD-dépendante) de *Bacillus stearothermophilus*. **p. 152**

Figure 50 : Organisation génétique de la région *yufL-yufM* chez *Bacillus subtilis*, *Bacillus halodurans*, *Bacillus anthracis*, *Oceanobacillus hieyensis*, *Streptoccocus agalactiae* et *Streptococcus pyogenes*. **p. 153**

Figure 51 : Profil d'expression d'une fusion rapportrice *ytsJ'::lacZ* en milieu de sporulation (SP) et en milieu minimum supplémenté en glucose ou en malate. **p. 154**

Figure 52 : Schéma expérimental d'une comparaison transcriptomique entre une souche sauvage et un mutant *cggR*. **p. 158**

Figure 53 : Image brute du marquage de la membrane « H3-Glucose » après révélation au PhosphoImager (Molecular Dynamics). **p. 164**

Figure 54 : Résultats globaux des trois comparaisons transcriptomique « Succinate/Glucose » sous forme de graphique de dispersion en échelle logarithmique. **p. 164**

Figure 55 : Etude de l'interaction CggR/ADN par anisotropie de fluorescence. **p. 170**

Figure 56 : Test de l'effet du FBP et d'autres métabolites ou sucres sur l'interaction ADN/CggR par mesure de variation d'anisotropie de fluorescence. **p. 171**

Figure 57 : Modèle hypothétique de l'interaction entre CggR et l'ADN en fonction de la concentration en FBP. **p. 171**

Figure 58 : Alignement du domaine régulateur de CggR de *Bacillus subtilis* avec les domaines homologues chez *Listeria monocytogenes* (LMONO), *Bacillus megaterium* (BACME), *Bacillus halodurans* (BHAL), *Staphylococcus aureus* (SAUR) et *Clostridium acetobutilinum* (CACE), les domaines régulateurs de DeoR ou SorC de *Bacillus subtilis* (BACSU), *Rhizobium meliloti* (RHIME), *Bacillus halodurans*, *Listeria monocytogenes*, *Escherichia coli* (ECOLI), *Lactobacillus casei* (LCASEI) et *Streptococcus pneumoniae* (SPNEUM), et les séquences des glucosamine-6-phosphate déaminases d'*Escherichia coli, Haemophilus influenzae* (HAEIN), *Borrelia burgdorferi* (BORBU), *Caenorhabditis elegans* (CAEEL), *Homo sapiens* (HUMAN), *Candida albicans* (CANAL) *et Bacillus subtilis.* **p. 173**

Figure 59 : Modélisation du site actif de la glucosamine déaminase d'*Escherichia coli* (en bleu clair) montrant le F6P, superposé au site actif potentiel de CggR de *Bacillus subtilis* (en noir) modélisé à partir de 1fqo. **p. 173**

Figure 60 : Modèle hypothétique de la régulation de l'expression de l'opéron *gapA* prenant en compte l'interaction entre CggR et l'ADN en fonction de la concentration en FBP et le rôle de CcpA dans le transport des sucres PTS et la formation de FBP. **p. 175**

Figure 61 : Modèle hypothétique de la régulation de l'expression de l'opéron *gapA* intégrant les différents niveaux de régulation décrits. **p. 176**

Figure 62 : Schéma combiné du métabolisme central du carbone et de sa régulation transcriptionnelle chez *Bacillus subtilis*. **p. 191**

Abréviations

ADN	acide désoxyribonucléique
ADP	adénosine diphosphate
AMPc	adénosine monophosphate cyclique
Arg	arginine
ARN	acide ribonucléique
ARNm	acide ribonucléique messager
ATP	adénosine triphosphate
Ca	Calcium
CAP	Catabolite Activator Protein
CAT	Co Anti Terminator
CcpA	Catabolite control protein A
CoA	Coenzyme A
Cra	Catabolite activator/repressor
cre	catabolite responsive element
Crh	Catabolite repression HPr
CRP	cAMP receptor protein
Da	dalton
DO	densité optique
EI	enzyme I
EII	enzyme II
FAD	flavoadénine dinucléotide
Kb	kilobase
GAPDH	glycéraldéhyde-3-phosphate déshydrogénase
GDP	guanosine diphosphate
GTP	guanosine triphosphate
His	histidine
HPr	heat resistant protein
HPr-His15-P	forme phosphorylée d'HPr sur l'histidine 15
HPr-Ser46-P	forme phosphorylée d'HPr sur la sérine 46
HPrK/P	HPr kinase/phosphatase
IPTG	isopropyl-β-D-thiogalactopyranoside
Lys	Lysine

Mg	Magnésium
mM	millimolaire
Mn	Manganèse
MS	spectrométrie de masse
nM	nanomolaire
NAD	nicotinamide adénine dinucléotide
NADP	nicotinamide adénine dinucléotide phosphate
ORF	open reading frame / phase ouverte de lecture
SDS-PAGE	électrophorèse de protéines en gel de polyacrylamide dénaturant
Ser	sérine
Pase	sucre-phosphate phosphatase
Pb	paire de base
PGM	phosphoglycérate mutase
Pi	phosphate inorganique
PRD	PTS Regulation Domain
PTS	phosphoénolpyruvate sucre phosphotransférase
RAT	Ribonucleic Anti Terminator
RMN	résonance magnétique nucléaire
RT-PCR	Reverse Transcription-Polymerase Chain Reaction
RU	responsive units
TCS	système à deux composants
TMG	méthyl-β-D-thiogalactopyranoside
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
Xt	extérieur

Glucides et Métabolites

P	phosphate/phospho
2PG	2-phosphoglycérate
6PGA	6-phosphogluconolactate
6PG	6-phosphogluconate
AC	acétate
ACTP	acétyl phosphate
AKG	2-oxoglutarate
BPG	2,3-biphosphoglycérate
DHAP	dihydroxyacétone phosphate

E4P	érythrose-4-phosphate
ETH	éthanol
F26P	fructose-2,6-biphosphate
F6P	fructose-6-phosphate
FBP	fructose-1,6-biphosphate
FOR	formate
FUM	fumarate
G3P	glycéraldéhyde-3-phosphate
G6P	glucose-6-phosphate
GL3P	glycérol-3-phosphate
GLC	glucose
GLX	glyoxylate
ICIT	isocitrate
LAC	lactate
LCTS	lactose
MAL	malate
OAA	oxaloacétate
PEP	phosphoénopyruvate
PRPP	α-D-5-phosphoribosylpyrophosphate
PYR	pyruvate
R5P	ribose-5-phosphate
RL5P	ribulose-5-phosphate
S7P	séduheptulose-7-phosphate
SUCC	succinate
X5P	xylulose-5-phosphate

Introduction bibliographique

***Bacillus subtilis* : un modèle d'étude du métabolisme carboné.**

Bacillus subtilis est sans conteste considéré comme la bactérie modèle des bactéries à Gram positif, mais aussi certainement et tout simplement comme l'un des organismes les plus étudiés et les mieux « compris ». Ainsi pourrait-on traduire les premières lignes de *Bacillus subtilis* and Its Closest Relatives : from genes to cells (Sonenshein *et al.*, 2002), le dernier grand ouvrage – presque une bible – essayant de faire la synthèse de toutes les connaissances acquises concernant *Bacillus subtilis*. En effet, seule *Escherichia coli* et *Salmonella typhimurium* ont fait l'objet d'autant d'études (Neidhardt *et al.*, 1987). Historiquement, en 1961, la découverte de la transformabilité naturelle de cette bactérie non pathogène a très tôt rendu possible son étude génétique. Sa capacité à sporuler a été étudiée comme un modèle simple pour comprendre les mécanismes de la différenciation cellulaire. Par ailleurs, *Bacillus subtilis* et d'autres Bacillacées proches ont aussi un intérêt économique considérable : l'industrie exploite ces microorganismes comme "usine biologique" pour différentes bioconversions (exemple : la production de riboflavine) et pour la synthèse d'enzymes massivement utilisées, notamment les protéases intégrées dans les lessives et les amylases utilisées pour extraire des sucres "simples" de l'amidon de maïs ou d'autres plantes de grande culture. Les énormes quantités de bactéries et d'enzymes en jeu représentent une valeur considérable et engendrent des coûts de production énormes que les industriels cherchent à réduire, même d'une fraction faible. Aussi, ces coûts peuvent être minimisés grâce, entre autres, à une meilleure connaissance et une meilleure maîtrise du métabolisme.

Toutes ces raisons ont motivé la mise en oeuvre d'un effort commun, européen et japonais, de séquençage complet du génome de *Bacillus subtilis*. Ainsi, dès 1997, ce fut la première bactérie à Gram positif dont l'intégralité du chromosome était enfin révélée. De l'accès à cette séquence dépendait la mise au point de nouvelles approches expérimentales dites globales. Ce voile enfin levé, on allait pouvoir regarder en même temps l'expression de tous les gènes et atteindre l'organisation physique du chromosome. On allait aussi enfin pouvoir chiffrer, par l'intermédiaire du nombre de gènes connus sur cette carte, l'étendu des connaissances génétiques accumulées sur *Bacillus subtilis*. L'analyse bioinformatique de cette séquence d'environ 4 millions de nucléotides a permis de dénombrer quelques 4200 gènes. Parmi eux, 60 % codent des protéines dont la ou les fonctions étaient reconnues de façon plus ou moins précise. Restaient par conséquent 40 % des gènes pour lesquels aucun indice suffisamment précis n'était disponible quant à leur fonction. En Europe, des

laboratoires se sont associés au sein d'un premier programme nommé BSFA (Bacillus subtilis functional analysis) pour développer l'étude systématique du génome de cette bactérie. Un programme jumeau, intitulé JAFAN (Japan Functional Analysis Network) a aussi été bâti au Japon. Ces deux programmes ont d'abord visé à l'élucidation de la fonction biologique des gènes dits "inconnus". L'idée était de construire un mutant pour chacun d'entre eux et de repérer, à travers toute une panoplie de cribles phénotypiques, la ou les fonctions métaboliques qui pourraient être affectées. Ainsi, on espérait au moins assigner chaque gène à une grande catégorie fonctionnelle. Sur la base de cette première classification, une analyse plus approfondie a permis de déterminer la fonction précise de certains gènes. A titre d'exemple, au laboratoire, il a été observé qu'un mutant *ytvP* était parfaitement viable sur milieu riche mais incapable de croître sur milieux minimum. Des analyses complémentaires ont permis de relier ce défaut de croissance à une auxotrophie pour un acide aminé, l'histidine. Par la suite, il a été démontré que ce gène codait l'avant dernière étape de la voie de biosynthèse de l'histidine et a été renommé *hisJ* (le Coq *et al.*, 1999). Cependant, aucun phénotype n'a pu être associé à la mutation d'un certain nombre de gènes. On peut évoquer deux raisons simples pour l'expliquer : d'une part, les conditions requises pour détecter l'effet de la mutation du gène considéré n'était peut-être pas rassemblées ; d'autre part, du fait de l'existence de plusieurs gènes paralogues et de redondance fonctionnelle, l'inactivation d'un seul gène peut ne pas conférer de phénotype décelable. Suite à cela, s'appuyant sur les banques de connaissances et de mutants acquises durant le premier programme, un deuxième programme européen, Bacell Network, a été lancé en 1999, dont l'ambition était de découvrir de nouvelles régulations de l'expression génétique concernant divers processus généraux comme les métabolismes du carbone, de l'azote et la réponse à divers stress environnementaux.

Bien que les cellules présentent des morphologies et des structures différentes, leur fonctionnement de base est toujours le même. Leur existence même, dans la durée, nécessite une certaine dissipation d'énergie, dont rend compte le deuxième principe de la thermodynamique. Elles extraient cette énergie principalement de la dégradation de molécules organiques : c'est le catabolisme. D'autre part, elles doivent fabriquer des molécules de base, matière première de leur construction, c'est-à-dire fabriquer les lipides, glucides et acides aminés nécessaires pour construire les macromolécules : c'est l'anabolisme. Les deux grands ensembles de réactions qui réalisent ces opérations sont regroupés sous le terme de métabolisme.

Dans leur habitat naturel, les bactéries sont très souvent confrontées à une multitude d'autres organismes, bactéries, champignons et autres. Dans ce contexte , la capacité à coloniser l'écosystème rapidement dépend beaucoup de la capacité de la bactérie à orienter son métabolisme vers l'utilisation de sources de nourritures rapidement métabolisables et fournissant le maximum d'énergie. Cela implique donc l'existence d'un système de sélection des sources de carbone possédant ces critères. En outre, afin de préserver, d'économiser les ressources cellulaires, la bactérie doit être capable de réguler l'expression des gènes de telle sorte que (idéalement) seuls ceux requis pour l'utilisation préférentielle de telle ou telle source de carbone soient induits. En effet, parmi les nutriments, les molécules sources de carbone ont une importance cruciale, double. D'une part, toute structure vivante est basée sur le carbone ; le métabolisme est donc chargé de produire les composés carbonés intermédiaires rentrant dans la constitution cellulaire. D'autre part, les composés carbonés sont la source privilégiée d'énergie ; c'est au niveau du métabolisme que cette énergie sera d'abord convertie puis utilisée pour le fonctionnement des processus cellulaires.

Dans ce cadre, nous nous sommes consacrés à la recherche de nouveaux gènes, catalytiques ou régulateurs, pouvant être impliqués dans le métabolisme du carbone. Nous nous sommes plus particulièrement orientés vers la recherche des mécanismes qui permettent de réguler l'expression des gènes impliqués dans les trois voies majeures du métabolisme carboné chez *Bacillus subtilis*, à savoir la glycolyse, le cycle de Krebs et la voie des pentoses phosphate. En effet, tous les composés carbonés utilisés par la bactérie et à partir desquels seront construites toutes les molécules constituant la cellule traversent aux moins l'une de ces trois voies. L'introduction essaiera dans un premier temps de décrire l'étendue actuelle des connaissances de ces voies métaboliques, chez *Bacillus subtilis* en particulier, mais présentera aussi certains éléments établis chez *Lactococcus lactis*, une autre bactérie à Gram positif et un modèle des bactéries lactiques, et chez d'autres bactéries à Gram négatif comme *Escherichia coli*, cela afin de mieux faire apparaître la spécificité de *Bacillus subtilis*. Il s'agira de désigner les éléments-clés de chacune des voies, les « carrefours métaboliques » les reliant mais aussi de mettre en exergue les éléments manquants ou incertains qui demeurent dans notre vision du métabolisme central du carbone chez *Bacillus subtilis*. Puis, et là se précise notre intérêt, les mécanismes de régulation du métabolisme carboné développés chez *Bacillus subtilis* seront décortiqués, surtout au niveau du transport des sources de carbone à l'intérieur de la cellule et du contrôle de l'expression des gènes impliqués dans ce métabolisme. Il ne s'agira pas d'être exhaustif mais d'essayer

de présenter un éventail significatif des mécanismes prédominants chez *Bacillus subtilis* qui permettent de contrôler l'expression des gènes impliqués dans le métabolisme du carbone en fonction de la disponibilité des sources présentes dans le milieu de croissance.

1. Les voies centrales du métabolisme du carbone chez *Bacillus subtilis*

La glycolyse, le cycle de Krebs et la voie des pentoses phosphate sont les trois voies qui constituent le métabolisme central du carbone (Figure 1). Elles sont conservées dans tout le monde vivant bien que, dans le détail, certaines classes d'organismes possèdent des spécificités. Dans le monde microbien, chaque voie a fait l'objet d'études biochimiques très poussées. Ainsi, dès le début des années 1960, l'ossature du métabolisme central du carbone chez les microorganismes était déjà connue : les étapes enzymatiques avaient été caractérisées et les intermédiaires métaboliques identifiés.

Si l'étude de ces voies métaboliques a passionné tant de scientifiques, c'est probablement parce qu'elles constituent tout simplement les étapes absolument nécessaires à la synthèse des intermédiaires métaboliques qui serviront à la construction des constituants plus élaborés de la cellule. En effet, toutes les voies métaboliques secondaires (ou intermédiaires) se branchent à un point ou un autre de ces trois voies majeures.

1.1. La glycolyse : une voie universelle

1.1.1. Description générale

La glycolyse est sans doute l'une des voies métaboliques les mieux conservées chez tous les organismes (Dandekar *et al.*, 1999). Elle fut l'une des premières connues car son étude a commencé au siècle dernier avec les travaux d'Embden et de Meyerhof, c'est pourquoi elle est aussi nommée voie d'Embden-Meyerhof. Comme le suggère son étymologie (glycos : sucre et lysis : dissolution), elle consiste en une série de réactions biochimiques au cours de laquelle une molécule de glucose, préalablement phosphorylée en glucose-6-phosphate, est transformée en deux molécules de pyruvate (Figure 2), avec production d'énergie, sous forme de molécules d'ATP (adénosine triphosphate), et de pouvoir réducteur.

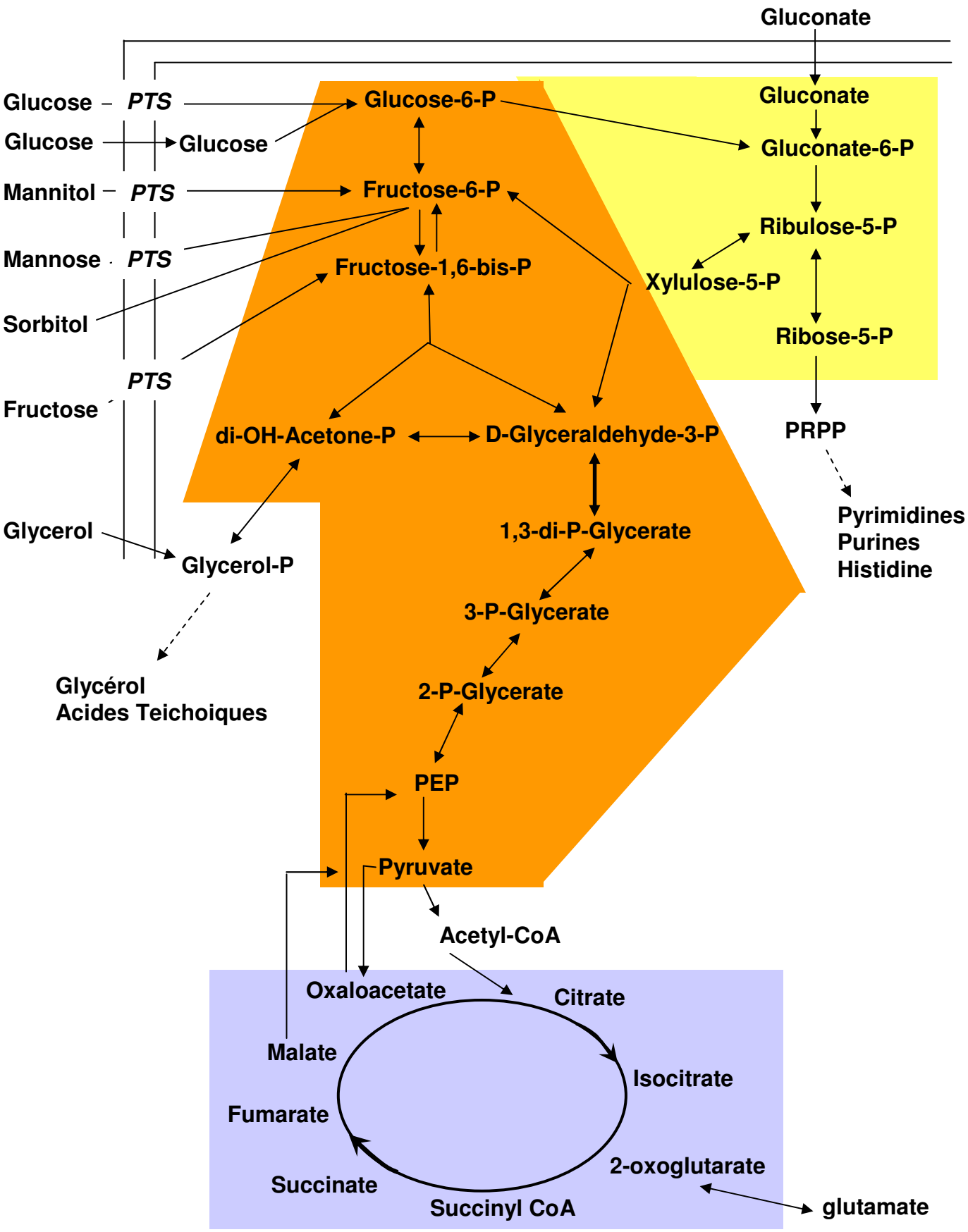


Figure 1. Représentation schématique des trois voies centrales du métabolisme du carbone chez *Bacillus subtilis*. Les sources de carbone sont indiquées à titre d'exemple.

Glycolyse et néoglucogénèse ***Voie des pentoses phosphates***

Cycle de Krebs

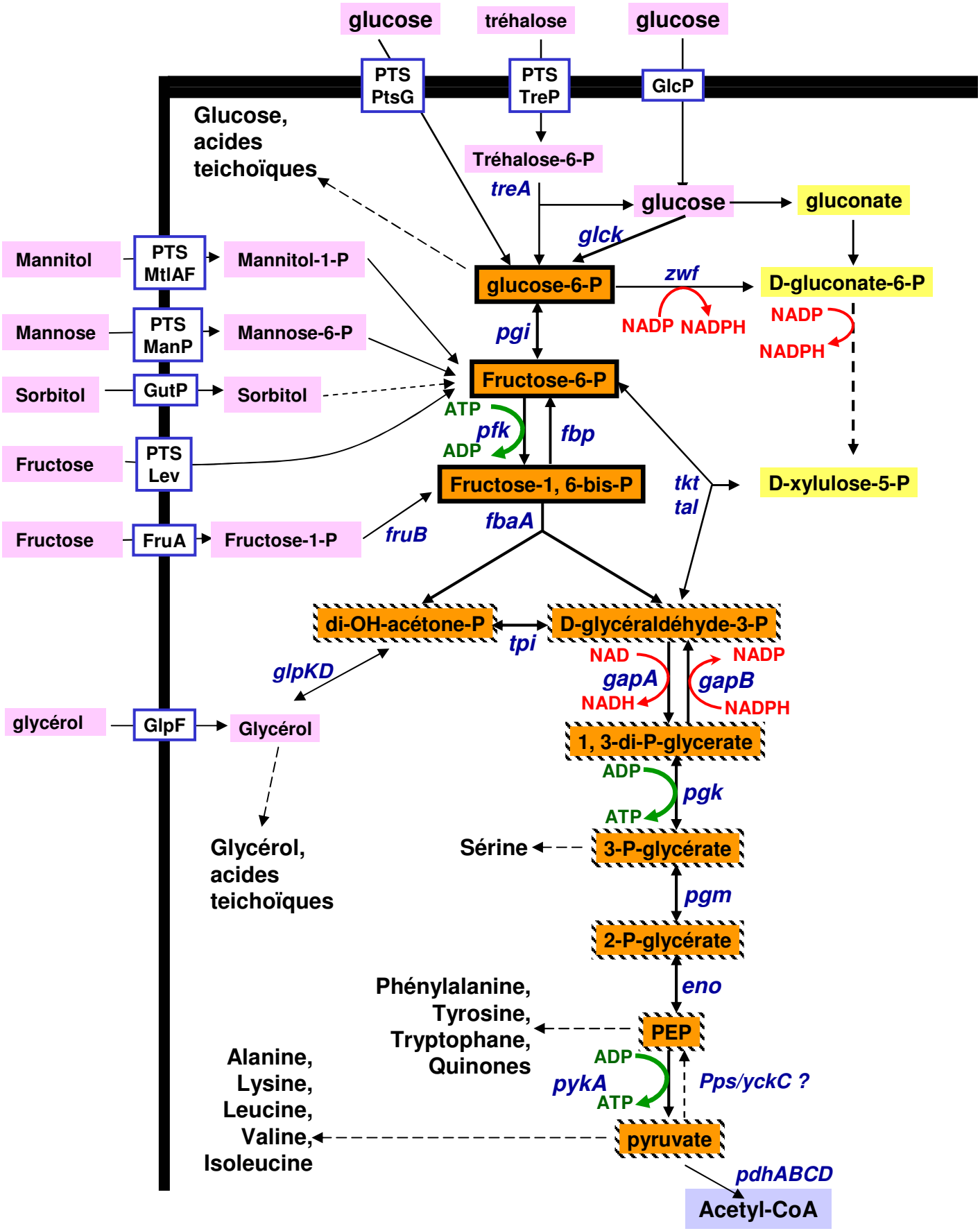


Figure 2. La glycolyse, les principales voies d'utilisation des sources de carbone entrant dans la glycolyse et les voies anaboliques dérivées chez *Bacillus subtilis*. Les intermédiaires glycolytiques sont indiqués en orange. Les hexoses phosphates et les trioses phosphates sont encadrés en trait plein noir et en rainures noires, respectivement. Les sources de carbone indiquées à titre d'exemple sont en mauve. Les intermédaires de la voie des pentoses phosphates et du cycle de Krebs qui se trouve à l'interface avec la glycolyse sont indiqués en jaune et en bleu, respectivement.

C'est la voie principale d'utilisation des sucres chez la très grande majorité des bactéries. Elle est commune à la fermentation et à la respiration oxydative. La glycolyse joue un rôle central dans le puzzle du métabolisme carboné puisqu'elle est la seule route possible pour relier le cycle de Krebs et la voie des pentoses phosphate (Figure 1). Lorsque la bactérie pousse sur des sources de carbone autres que les sucres qui entrent dans le métabolisme en aval, elle doit « remonter » la glycolyse pour synthétiser certains précurseurs. C'est le processus inverse de la glycolyse. On parle alors de néoglucogenèse. Aussi, presque toutes les enzymes impliquées dans la glycolyse ont la capacité de fonctionner de façon réversible, aussi bien dans le sens glycolytique que dans le sens néoglucogénique. Cependant, dans le cas de quelques réactions, il existe des enzymes « spécialisées » qui, dans les conditions physiologiques et thermodynamiques de la cellule, catalysent la réaction uniquement dans un sens. *Bacillus subtilis* possède donc certaines enzymes exclusivement requises pour la glycolyse et d'autres exclusivement requises pour la néoglucogenèse.

1.1.2. Les étapes de la glycolyse

Les composantes de la glycolyse sont très bien décrites sur le plan biochimique chez les bactéries modèles que sont *Escherichia coli* et *Bacillus subtilis* et plus récemment chez les bactéries lactiques (Fraenkel, 1987; Fortnagel, 1993; Thompson & Gentry-Weeks, 1994). Cependant, les données complémentaires concernant les flux métaboliques n'ont pu être obtenues que récemment, notamment grâce à l'avènement de techniques comme la résonance magnétique nucléaire et la spectrométrie de masse (Wiechert, 2001). Aujourd'hui, il s'agit de tisser les liens entre chaque enzyme, ses niveaux de régulation et la variation du flux glycolytique en fonction des conditions de croissance.

Classiquement, la glycolyse est divisée en deux sous-parties (Gunsalus & Shuster, 1961) (Figure 2). Les réactions de la partie supérieure permettent la conversion des hexoses-phosphates, c'est-à-dire la transformation d'une molécule de glucose-6-phosphate (G6P) en une molécule de dihydroxyacétone-phosphate (DHAP) et une molécule de glycéraldéhyde-3-phosphate (G3P). Celles de la seconde partie, ou partie inférieure, de la glycolyse permettent la conversion des trioses-phosphates, c'est-à-dire la transformation du DHAP et du G3P en pyruvate.

Les noms de gènes indiqués sont tels que décrits sur le serveur de base de données SubtiList (Data Release R16.1, April 26, 2001 ; (Moszer *et al.*, 1995)).

1.1.2.1. La glucokinase (ou hexokinase) *glcK* (EC 2.7.1.1)

Le glucose qui entre dans la voie glycolytique est phosphorylé sous forme de G6P cytoplasmique. Chez les bactéries, le glucose extracellulaire peut, soit être concomitamment transloqué et phosphorylé par un transporteur spécifique de type PTS (Phospho Transferase System) (cf partie 2.2), soit entrer sans modification *via* un transporteur secondaire. Chez *Bacillus subtilis*, il a été montré que le premier mode cité de transport et de phosphorylation concomitante du glucose joue un rôle majeur alors que le second mode de transport est minoritaire (Paulsen *et al.*, 1998; Fiegler *et al.*, 1999). Lorsqu'il est perméé par le transporteur secondaire, GlpP chez *Bacillus subtilis*, ou lorsqu'il est issu du catabolisme de disaccharides comme le maltose et le tréhalose (Schonert *et al.*, 1998), le glucose intracellulaire est phosphorylé de façon non réversible par la glucokinase pour entrer dans la glycolyse sous forme de G6P. Cette seconde voie de phosphorylation est essentielle pour les bactéries chez lesquelles le système PTS est manquant ou déficient (Fraenkel, 1987). Le substrat ainsi modifié ne peut plus ressortir de la cellule et est immédiatement intégré dans la voie glycolytique.

La glucokinase existe chez la très grande majorité des Eubactéries (Figure 3), à Gram positif comme à Gram négatif, mais semble être beaucoup moins répandue et peu spécifique chez les Archées (Dandekar *et al.*, 1999; Hansen *et al.*, 2002).

Chez *Bacillus subtilis*, Skarlatos et Dahl ont montré que la délétion du gène *glcK* résultait en l'inactivation complète de l'activité glucokinase dans des extraits bruts (Skarlatos & Dahl, 1998). Il semble donc que GlcK soit la seule glucokinase majeure de *Bacillus subtilis*. Il s'agit d'une enzyme dimérique, d'un poids moléculaire de 33,5 kDa, ATP-dépendante, nécessitant le NAD comme cofacteur (Skarlatos & Dahl, 1998), et qui est exprimée de façon constitutive (Meyer *et al.*, 1997; Spath *et al.*, 1997; Tobisch *et al.*, 1999b). Cependant, une souche mutée pour le gène *glcK,* qui code cette enzyme, croît à la même vitesse en présence de différentes sources de carbone, glucose, fructose, maltose, sucrose ou tréhalose (Skarlatos & Dahl, 1998). Ces sucres sont transportés par un système de type PTS et rentrent donc dans la cellule sous forme phosphorylée. GlcK n'est donc pas nécessaire dans ces conditions. Par ailleurs, un mutant de la voie PTS de transport du glucose pousse très lentement en présence de glucose comme seule source de carbone (Gonzy-Treboul *et al.*,

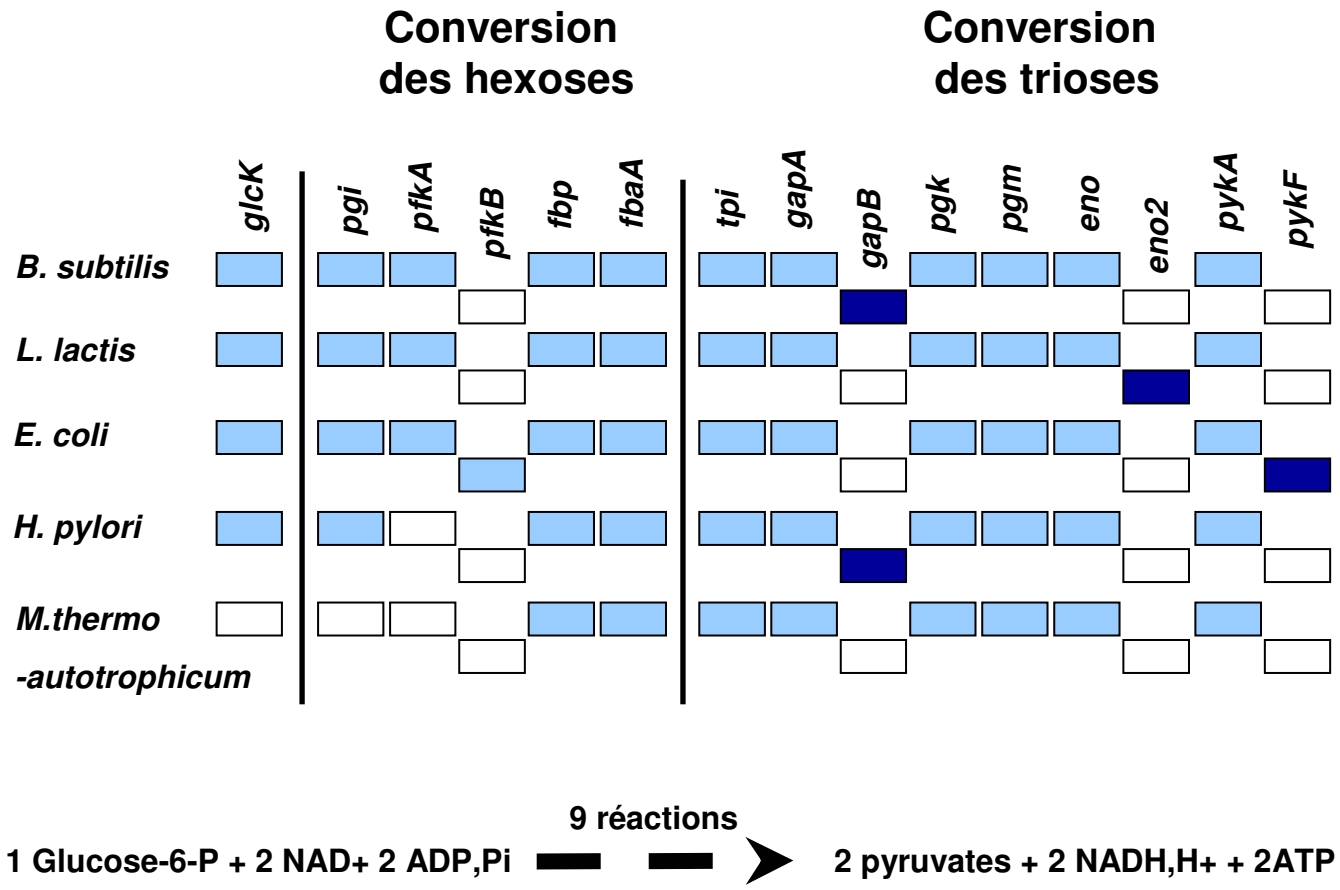


Figure 3. Comparaison des étapes de la glycolyse chez 5 bactéries dont le génome est disponible. Les cases blanches indiquent les étapes enzymatiques pour lesquelles aucun gène n'a pu être identifié. Lorsqu'il existe un gène paralogue, celui-ci est indiqué en bleu foncé. *B.* : *Bacillus* ; *L.* : *Lactococcus* ; *E.* : *Escherichia* ; *H.* : *Helicobacter* ; *M.* : *Methanobacter*.

D'après Dandekar *et al*, 1999.

1991; Bachem *et al.*, 1997). Ces résultats suggèrent que la glucokinase joue un rôle métabolique très mineur et donc que le transport et la phosphorylation du glucose sont essentiellement assurés par le système PTS chez *Bacillus subtilis*.

Chez *Escherichia coli*, dans un contexte où le transporteur PTS du glucose est déficient, la glucokinase Glk joue un rôle plus important dans la mesure ou elle permet une croissance significative en présence de glucose (Fraenkel *et al.*, 1964; Curtis & Epstein, 1975). Cependant, un mutant *glk* croît aussi bien que la souche sauvage en présence de glucose (Curtis & Epstein, 1975).

1.1.2.2. La partie supérieure de la glycolyse : conversion des hexoses

1.1.2.2.1. La phosphoglucoisomérase pgi (EC 5.3.1.9)

L'interconversion du G6P en fructose-6-phosphate (F6P) est catalysée par la phosphoglucoisomérase. Il s'agit de la première étape qui va diriger le flux dans le sens de la glycolyse. Elle existe chez toutes les Eubactéries étudiées mais est absente chez plusieurs Archées (Dandekar *et al.*, 1999) (Figure 3).

Chez *Bacillus subtilis*, les résultats de nos tests de croissance indiquent que la délétion du gène *pgi* entraîne une croissance très sévèrement ralentie en présence de glucose comme seule source de carbone (Doan et Aymerich, non publié). Cela indique que, chez *Bacillus subtilis*, l'utilisation du glucose pour la croissance est beaucoup plus efficace lorsqu'il est catabolisé à travers la glycolyse qu'à travers la voie des pentoses phosphate (cf partie 1.2). Chez *Escherichia coli*, un mutant *pgi* pousse lentement en présence de glucose en utilisant la voie des pentoses phosphate (Fraenkel *et al.*, 1964; Canonaco *et al.*, 2001).

Par ailleurs, la phosphoglucoisomérase a été détectée comme protéine de surface chez *Streptococcus agalactiae* (Hughes *et al.*, 2002).

1.1.2.2.2. La 6-phosphofructokinase pfk (EC 2.7.1.11)

La conversion du F6P en fructose-1,6-bisphosphate (FBP) est catalysée par la 6-phosphofructokinase de façon irréversible et ATP-dépendante.

Cette enzyme est l'une des mieux caractérisées sur le plan biochimique parmi les enzymes glycolytiques du fait de sa position clé. Elle peut utiliser l'ATP, l'ADP (chez les Archées genres *Pyrococcus*, *Thermococcus* et *Methanococcus*) ou le pyrophosphate (chez les plantes) comme donneur de groupement phosphate. On peut distinguer trois isoformes différentes : l'isoforme PfkA (parfois dupliquée comme chez *Treponema pallidum*), forme ATP-dépendante majoritaire des Eucaryotes supérieurs et des Procaryotes, ou pyrophosphate-dépendante des plantes ; l'isoforme PfkB qui existe chez les bactéries ; et l'isoforme PfkC qui ne compte que la phosphofructokinase ADP-dépendante des Archées des genres Pyrococcus, Thermococcus et Methanococcus (Ronimus & Morgan, 2001; Verhees *et al.*, 2001).

Bacillus subtilis et *Lactococcus lactis* possèdent l'isoforme PfkA ATP-dépendante, codée par le gène *pfk*. Un mutant *pfk* est incapable d'utiliser tous les composés (glucose, mannitol, mannose, saccharose et sorbitol) qui sont intégrés dans la glycolyse en amont ou au niveau du F6P comme seule source de carbone (Gay, 1979; Andersen *et al.*, 2001). Ce phénotype est intimement lié à l'effet toxique de l'accumulation progressive du F6P dans le cytoplasme. Cette accumulation a aussi pour effet de retarder la sporulation (Freese *et al.*, 1974; Freese & Heinze, 1984). Par contre, un tel mutant pousse très bien en présence de fructose ou de glycérol, qui entre dans la glycolyse en aval de cette étape, ou en présence des sources de carbone néoglucogéniques comme le malate (Doan et Aymerich, non publié).

Chez *Escherichia coli*, il existe deux 6-phosphofructokinases. L'enzyme majoritaire, Pfk-1, est codée par le gène *pfkA*. C'est une enzyme tétramérique allostérique de type PfkA, activée par les nucléotides biphosphates et inhibée par le PEP (Blangy *et al.*, 1968; Auzat *et al.*, 1997). Une deuxième forme de type PfkB, Pfk-2, est codée par le gène *pfkB*. Son activité est responsable d'environ 10 % de l'activité 6-phosphofructokinase et permet une croissance résiduelle dans un mutant *pfkA* (Daldal, 1983).

On peut remarquer qu'*Helicobacter pylori* fait exception et ne possède pas d'homologue de cette enzyme. Par ailleurs, comme la phosphoglucoisomérase, la phosphofructokinase est aussi absente chez de nombreux Archées (Figure 3).

1.1.2.2.3. La fructose-1,6-biphosphatase fbp (EC 3.1.3.11)

Cette enzyme est l'une des étapes clés de la néoglucogenèse. En effet, la phosphofructokinase n'est pas capable de catalyser la conversion du FBP en F6P chez *Bacillus subtilis*. Cette étape de déphosphorylation nécessite l'activité fructose-1,6-biphosphatase.

Chez *Bacillus subtilis* et les bactéries du même genre, cette enzyme semble se distinguer des fructose-1,6-biphosphatases présentes dans les bases de données recouvrant les entérobactéries, les animaux et les plantes (Fujita *et al.*, 1998). Son activité, constitutive et dépendante de l'ion manganèse Mn^{2+}, n'est pas absolument requise pour la croissance sur des sources de carbone néoglucogéniques comme le malate (Fujita & Freese, 1981; Fujita & Fujita, 1983). En effet, il existe une voie de contournement permettant de produire du G6P à partir du FBP *via* la voie des pentoses phosphate chez *Bacillus subtilis* (partie 1.2).

Chez *Escherichia coli*, deux fructose-1,6-biphosphatases ont été révélées. La première enzyme est codée par le gène *fbp*. L'expression de ce gène est nécessaire à la croissance en présence de substrats néoglucogéniques. (Fraenkel & Horecker, 1965; Yu *et al.*, 1965; Donahue *et al.*, 2000). Il code une enzyme tétramérique qui nécessite la présence d'ions Mg^{2+} pour son activité (Babul & Guixe, 1983). Récemment, il a été montré que le gène *glpX*, qui est le dernier gène de l'opéron *glpFKX* (métabolisme du glycérol), code une fructose-1,6-biphosphatase de classe distincte qui ne possède que 10 % d'identité avec l'enzyme codée par *fbp* (Donahue *et al.*, 2000). Par ailleurs, certaines bactéries ne possèdent que l'enzyme de type GlpX. Chez *Escherichia coli*, GlpX ne semble pas nécessaire pour la croissance mais pourrait intervenir dans la régulation du flux glycolytique en coordination avec Fbp (Donahue *et al.*, 2000).

1.1.2.2.4. La fructose-1,6-biphosphate aldolase fba (EC 4.1.2.13)

Cette enzyme catalyse la conversion réversible du FBP, un hexose, en dihydroxyacétone-phosphate (DHAP) et en glycéraldéhyde-3-phosphate (G3P), deux trioses très proches (Fraenkel, 1987). Elle est présente chez toutes les bactéries étudiées et largement caractérisée. Son activité aldolase est métal-dépendante (Baldwin *et al.*, 1978).

Il existe deux classes communes de FBP aldolases : les Archées possèdent une FBP aldolase de type I (Siebers *et al.*, 2001), alors que l'enzyme qui intervient dans la glycolyse chez *Bacillus subtilis* (Fortnagel, 1993) ou *Escherichia coli* (Fraenkel, 1987) est de classe II.

Chez *Bacillus subtilis*, l'expression de *fbaA* semble constitutive (Tobisch *et al.*, 1999b; Ludwig *et al.*, 2001). Des mutations ponctuelles ont été obtenues qui conduisent à la perte quasi-totale de l'activité aldolase dans des extraits protéiques obtenus à partir des souches portant ces mutations cultivées en milieu riche (Mitchell *et al.*, 1992). Il semble donc que l'activité enzymatique de la FBP aldolase ne soit pas essentielle. Alors, le fait qu'aucun mutant de délétion n'ait été obtenu jusqu'à présent suggère que cette protéine pourrait être impliquée dans une autre fonction qui, elle, serait requise pour la survie (Kobayashi *et al.*, 2003 sous presse).

Chez *Escherichia coli*, un mutant *fda* est incapable de pousser en présence de sucres mais le peut en présence de glycérol ou de succinate, qui intègrent le métabolisme en aval du FBP (Irani & Maitra, 1977; Fraenkel, 1987). Or, dans ces conditions, la formation de FBP est nécessaire à la croissance. Cette activité pourrait être codée par le gène *dhnA* dont le produit est une FBP aldolase de classe I (Thomson *et al.*, 1998). Un lien entre la FBP aldolase et la synthèse des ARNm a aussi été évoqué mais sans explication : une mutation thermosensible du gène *fda* inhibe la synthèse des ARNm à température permissive (Singer *et al.*, 1991). Un phénomène similaire a aussi été observé chez *Bacillus subtilis* (Trach *et al.*, 1988).

1.1.2.3. La partie inférieure de la glycolyse : conversion des trioses

1.1.2.3.1. La triose-phosphate isomérase (EC 5.3.1.1)

L'interconversion du DHAP et du G3P est catalysée par la triose-phosphate isomérase. Thermodynamiquement, elle favorise la formation de DHAP plutôt que celle de G3P. Le flux glycolytique est « tiré » par la réaction suivante qui consomme du G3P et déplace ainsi l'équilibre DHAP/G3P. *In vivo*, cette enzyme peut se trouver sous forme homodimérique, par exemple chez *Lactococcus lactis* (Cancilla *et al.*, 1995), ou bien activée sous forme d'un complexe bifonctionnel avec la phosphoglycérate kinase, PGK, chez la bactérie hyperthermophile *Thermotoga maritima* (Schurig *et al.*, 1995).

Chez *Bacillus subtilis*, le gène *tpi* a été cloné et séquencé (Leyva-Vazquez & Setlow, 1994) après les gènes homologues chez *Bacillus megaterium* et *Bacillus stearothermophilus* (Schlapfer & Zuber, 1992). Aucun mutant d'inactivation n'a pu être obtenu suggérant que la fonction de Tpi est essentielle (Kobayashi *et al.*, 2003 sous presse).

Chez *Escherichia coli*, un mutant *tpi* est capable de croître lorsque le milieu de croissance contient une source de carbone glycolytique et une source de carbone néoglucogénique. Ainsi, malgré l'absence de *tpi*, le métabolisme sera encore capable de produire tous les intermédiaires de la glycolyse (Anderson & Cooper, 1969; Irani & Maitra, 1977). Dans un tel mutant, l'accumulation de DHAP est toxique parce qu'il est converti *via* la méthylglyoxal synthase en méthyglyoxal qui est un agent bactéricide (Hopper & Cooper, 1971; Cooper, 1984).

1.1.2.3.2. La glycéraldéhyde-3-phosphate déshydrogénase gap (EC 1.2.1.12)

Cette enzyme catalyse la phosphorylation oxydative du G3P et du 1,3-biphosphoglycérate (1,3-BPG) en présence de NAD(P) comme cofacteur. Elle existe chez toutes les espèces étudiées et sa séquence est très bien conservée (Figge *et al.*, 1999). Il s'agit sans doute de l'une des enzymes de la glycolyse les mieux caractérisées au niveau biochimique et aussi structural (Corbier *et al.*, 1990a; Corbier *et al.*, 1990b; Clermont *et al.*, 1993; Corbier *et al.*, 1994; Eyschen *et al.*, 1996; Didierjean *et al.*, 1997; Boschi-Muller & Branlant, 1999; Bulatnikov *et al.*, 1999; Levashov *et al.*, 1999; Roitel *et al.*, 1999; Fillinger *et al.*, 2000; Iddar *et al.*, 2002; Didierjean *et al.*, 2003; Roitel *et al.*, 2003). Sa forme active est tétramérique. Chez les plantes, deux GAPDH coexistent : une GAPDH cytoplasmique strictement NAD-dépendante dont l'activité permet la glycolyse et la néoglucogenèse, et une GAPDH chloroplastique utilisant préférentiellement le NADP et uniquement requise pour le processus d'assimilation du CO_2 lors de la photosynthèse. Chez les cyanobactéries, il existe aussi deux GAPDH (voire trois), Gap1 et Gap2, toutes deux cytoplasmiques. Il a été montré chez *Synechocystis* sp. PC6803 que Gap1 est requise lors de la glycolyse et qu'à l'opposé, Gap2 est requise pour l'anabolisme dans le cycle de Calvin et dans la néoglucogenèse en l'absence de photosynthèse (Valverde *et al.*, 1997; Koksharova *et al.*, 1998).

Chez *Bacillus subtilis*, seul le gène *gapA* avait été séquencé (Viaene & Dhaese, 1989) et clairement identifié comme codant une GAPDH canonique NAD-dépendante par homologie avec les GAPDH eucaryotiques. Or, après le séquençage, l'analyse complète du génome a permis de révéler que *Bacillus subtilis* possédait une deuxième copie de ce gène, nommée *gapB*. Des études génétiques et biochimiques ont permis de démontrer que ces deux enzymes avaient des rôles physiologiques opposés, de par la régulation de leur synthèse et de par leurs caractéristiques enzymatiques : GapA intervient lors de la glycolyse et utilise le NAD comme cofacteur ; à l'opposé, GapB est requise lors de la néoglucogenèse et possède le

NADP comme cofacteur (Fillinger *et al.*, 2000). L'étude détaillée du mécanisme qui régule l'expression de *gapA* fait l'objet de la première partie de cette thèse. Ce paragraphe sera donc développé ultérieurement.

D'autres espèces bactériennes possèdent deux gènes paralogues codant des GAPDH. L'analyse de leur séquence montre que l'une possède une signature NAD comme GapA de *Bacillus subtilis* et que l'autre possède une signature NADP, ce qui suggère une spécialisation physiologique comparable à celle qui existe chez *Bacillus subtilis*. C'est très probablement le cas chez *Bacillus stearothermophilus*, *Bacillus anthracis*, *Bacillus halodurans*, trois Bacillacées très proches de *Bacillus subtilis* et peut-être aussi chez *Neisseiria meningitidis*, *Neisseiria gonorrhoeae* et *Helicobacter pylori*, trois bactéries à Gram négatif.

Chez *Escherichia coli*, il n'existe qu'une seule GAPDH codée par le gène *gapA* et dont l'activité est réversible. Il existe une protéine, codée par le gène *gapB*, dont la séquence protéique est très proche de celle de GapA (40 % d'identité) (Zhao *et al.*, 1995). Cependant, il a été montré que cette enzyme codait une érythrose-4-phosphate déshydrogénase (Boschi-Muller *et al.*, 1997).

Chez *Lactococcus lactis*, des expériences ont été menées qui montrent l'importance de l'activité de la GAPDH dans l'orientation du flux métabolique et le métabolisme du pyruvate particulier chez les bactéries lactiques (Cocaign-Bousquet *et al.*, 1996; Even *et al.*, 1999). La GAPDH constitue une étape limitante pour le flux glycolytique. De plus, il a été montré que l'inhibition de l'activité GAPDH conduisait, d'une part, à une baisse du flux à travers la glycolyse qui est due en partie à une modification du ratio NADH/NAD qui perturbe l'activité de certaines enzymes, et d'autre part, à l'augmentation du pool de FBP qui est un effecteur de plusieurs enzymes allostériques et de certains régulateurs.

Par ailleurs, d'autres rôles de la GAPDH ont été mis en évidence. Chez *Streptococcus pyogenes* et *Streptococcus agalactiae*, la GAPDH est aussi présente à la surface cellulaire pour jouer des rôles de protéine de liaison aux plasmines et de protéine de signalisation (D'Costa *et al.*, 2000; Hughes *et al.*, 2002). Chez *Schistosoma japanicum*, la GAPDH a été caractérisée pour sa capacité à lier l'ARN. En effet, le domaine d'interaction avec le NAD est capable d'interagir spécifiquement avec une séquence de l'ARNm riche en adénines et en uraciles (Nagy *et al.*, 2000).

1.1.2.3.3. La 3-phosphoglycérate kinase pgk (EC 2.7.2.3)

La 3-phosphoglycérate kinase catalyse l'interconversion du BPG et du 3-phosphoglycérate (3PG). Cette enzyme fonctionne aussi bien lors de la glycolyse que lors de la néoglucogenèse. Elle est très bien conservée et existe chez toutes les espèces étudiées. Elle se trouve parfois en complexe avec la triose-phosphate isomérase (Schurig *et al.*, 1995).

Chez *Bacillus subtilis*, un mutant *pgk*, gène qui code cette enzyme, est incapable de croître en présence d'une seule source de carbone. En effet, sa croissance requiert un approvisionnement en carbone par l'amont de cette étape, c'est-à-dire une source de carbone capable de fournir du BPG, et un approvisionnement en carbone par l'aval, capable de fournir du 3PG. Récemment, un mutant *pgk* non létal a été obtenu qui permet la croissance en milieu anaérobie en l'absence de la kinase ResE (Nakano *et al.*, 1999) : en l'absence de respiration oxydative, l'expression de nombreux gènes permettant la croissance cellulaire est activée *via* le système à deux composants ResED (senseur-kinase/régulateur transcriptionnel) ; dans un mutant ResE, cette mutation dans le gène *pgk* rend constitutif l'expression des gènes sous contrôle du régulateur ResD.

Par ailleurs, la 3-phosphoglycérate kinase a été détectée comme protéine de surface chez *Streptococcus agalactiae* (Hughes *et al.*, 2002).

1.1.2.3.4. La phosphoglycérate mutase pgm (EC 5.4.2.1)

L'isomérisation du 3PG et du 2-phosphoglycérate (2PG) est catalysée par la phosphoglycérate mutase de façon réversible. Il existe deux isoformes de structures différentes chez les Eucaryotes et les bactéries (Jedrzejas, 2000) , l'une nécessitant le 2,3-biphosphoglycérate comme donneur de groupement phosphate (PGM) et l'autre indépendante de ce cofacteur (iPGM). PGM est la phosphoglycérate mutase majeure voire unique chez les vertébrés et les levures tandis que iPGM est l'enzyme prédominante chez les plantes (Jedrzejas, 2000). Chez les bactéries, on retrouve le gène codant soit PGM soit iPGM soit les gènes codant les deux en même temps. Chez les Archées, il semble que la forme iPGM soit la plus répandue (van der Oost *et al.*, 2002).

Chez *Bacillus subtilis*, l'enzyme majeure, codée par le gène *pgm*, est de type iPGM. C'est une enzyme de 74 kDa dont l'activité nécessite l'ion Mn^{2+} (Watabe & Freese, 1979; Chander *et al.*, 1998; Jedrzejas *et al.*, 2000). Cependant, il existe une deuxième enzyme,

codée par le gène *yhfR*, de type PGM (Pearson *et al.*, 2000) mais dont la fonction physiologique n'est pas déterminée. Un mutant *pgm* existe qui a été construit par remplacement d'un fragment interne du gène par une cassette de résistance à l'érythromycine (Leyva-Vazquez & Setlow, 1994). La présence de l'enzyme n'est pas détectable dans des extraits de cellules cultivées dans un milieu YT (ref). La croissance de ce mutant est extrêmement altérée par rapport à celle de la souche sauvage dans ce milieu. Mais elle est significativement stimulée en présence d'un supplément de malate (Leyva-Vazquez & Setlow, 1994). Un tel mutant n'a pas été décrit chez *Escherichia coli*.

1.1.2.3.5. L'énolase eno (4.2.1.11)

L'énolase catalyse l'interconversion du 2PG et du phosphoénolpyruvate (PEP) de façon réversible. C'est une enzyme multimérique, très souvent octamérique (Brown *et al.*, 1998). Elle est très bien conservée.

Chez *Bacillus subtilis*, l'énolase, codée par le gène *eno*, possède une structure quaternaire de 8 sous-unités de 46,6 kDa (Brown *et al.*, 1998). Le Mg^{2+} semble être un cofacteur nécessaire à son activité (Singh & Setlow, 1978). De même que le mutant *pgm* précédemment décrit, un mutant *eno* croît extrêmement lentement dans un milieu YT et sa croissance est significativement stimulée par l'ajout de malate (Leyva-Vazquez & Setlow, 1994). Chez *Escherichia coli*, il existe un mutant *eno* thermosensible (Klein *et al.*, 1996).

Par ailleurs, l'énolase est impliquée dans d'autres fonctions physiologiques. D'une part, son interaction avec la RNase E a été mise en évidence dans le complexe de dégradation des ARNm ou dégradosome chez *Escherichia coli* ; mais aucune explication de cette observation n'a été fournie (Kuhnel & Luisi, 2001; Carpousis, 2002). D'autre part, chez *Streptococcus pneumoniae* et *Streptococcus agalactiae*, l'énolase est présente en surface où elle joue le rôle de protéine de liaison aux plasmines (Bergmann *et al.*, 2001; Hughes *et al.*, 2002).

1.1.2.3.6. La pyruvate kinase pykA (EC 2.7.1.4)

Dans les conditions physiologiques classiques de croissance, cette enzyme déphosphoryle irréversiblement le PEP en pyruvate avec production d'une molécule d'ATP. C'est une enzyme allostérique dont les activateurs peuvent être soit le FBP, le ribose-5-phosphate, l'AMP ou le PEP (Waygood *et al.*, 1975; Waygood *et al.*, 1976; Boiteux *et al.*, 1983; Lovell *et al.*, 1998; Valentini *et al.*, 2000; Knowles *et al.*, 2001).

Chez *Bacillus subtilis*, un mutant *pykA* poussant en présence de glucose comme source de carbone montre une croissance ralentie, une production très réduite d'acides et une production de CO_2 plus importante (Fry *et al.*, 2000; Phalakornkule *et al.*, 2000). La faible vitesse de croissance de ce mutant pourrait être liée à un déficit de formation du pyruvate et à l'accumulation de PEP et de G6P à l'intérieur de la cellule. En effet, en l'absence de PykA, la seule réaction qui permette encore la conversion du PEP en pyruvate est le transfert du groupement phosphate du PEP vers la chaîne de phosphorylation du système PTS. Cette activité ne serait pas suffisante pour fournir la quantité de pyruvate dont la cellule a besoin pour croître rapidement.

Chez *Escherichia coli*, et contrairement à *Bacillus subtilis* (Munoz et al, 1997), il existe deux pyruvate kinases, PykA et PykF. PykA est activée par le 5'-adénosinemonophosphate alors que PykF est activée par le FBP (Vinopal & Fraenkel, 1974; Waygood *et al.*, 1976; Boiteux *et al.*, 1983). La délétion de *pykF* ou de *pykA* n'entraîne pas de défaut prononcé de croissance (Garrido-Pertierra & Cooper, 1983; Ponce *et al.*, 1995). En outre, un double mutant *pykA pykF* est toujours capable de croître efficacement en milieu minimum supplémenté en glucose et n'accumule pas de PEP (Zhu *et al.*, 2001). Cette absence de phénotype, comparé au défaut partiel de croissance d'un mutant *pykA* de *Bacillus subtilis*, est due à la possibilité de pouvoir directement convertir le PEP en oxaloacétate *via* la phosphoénolpyruvate carboxylase pour faire tourner le cycle de Krebs (cf partie 1.3 et 1.4). Le pyruvate pourra être régénéré à partir de l'oxaloacétate *via* la malate déshydrogénase (oxaloacétate donne malate) et l'enzyme malique (cf partie 1.3 et 1.4). *Bacillus subtilis* ne possède pas de phosphoénolpyruvate carboxylase ; le PEP doit d'abord être converti en pyruvate qui sera ensuite converti en oxaloacétate par la pyruvate carboxylase. Ainsi, chez *Escherichia coli*, cette activité permet de fournir du pyruvate, en sus de celui qui provient de la déphosphorylation du PEP par le système PTS.

Cette enzyme est absente chez certaines bactéries, dont *Helicobacter pylori* et certains Archées (Dandekar *et al.*, 1999). Dans le cas de cette première bactérie, le PEP est uniquement déphosphorylé *via* le système PTS puisqu'elle ne possède pas de phosphoénolpyruvate carboxylase. Par ailleurs, il a été montré que l'ajout de pyruvate de sodium (0,05 %) augmentait significativement la vitesse de croissance d'une souche d'*Helicobacter pylori* en milieu riche (Jiang & Doyle, 2000), suggérant qu'elle est en partie auxotrophe pour le pyruvate.

1.1.2.3.7. La phosphoénolpyruvate synthase et la pyruvate dikinase

Ces deux enzymes catalysent la conversion d'une molécule de pyruvate en une molécule de PEP en utilisant un ATP. Ce sont donc des enzymes qui permettent à la cellule de réaliser la néoglucogenèse à partir de sources de carbone préalablement converties en pyruvate. Cependant, aucune activité de ce type n'a été détectée chez *Bacillus subtilis*, bien que des gènes codant des protéines homologues aux enzymes caractérisées d'*Escherichia coli* (*pps*) (Chao *et al.*, 1993) et de *Rhizobium meliloti* (*yvkC*) (Osteras *et al.*, 1997) existent.

1.1.3. Comparaison de la glycolyse chez *Bacillus subtilis*, *Lactococcus lactis, Escherichia coli, Helicobacter pylori* et un Archée *Methanobacterium thermoautotrophicum*

La Figure 3 récapitule la présence ou l'absence des étapes de la glycolyse chez quatre eubactéries modèles et une Archée dont le génome est également disponible. Cette comparaison montre la conservation remarquable de la voie d'Embden-Meyerhof chez les Procaryotes qui apparaît comme l'une des voies métaboliques les plus anciennes pour l'assimilation des sucres (Dandekar *et al.*, 1999).

La partie supérieure de la glycolyse n'est pas absolument requise pour convertir le glucose en pyruvate puisque la voie des pentoses phosphate est une alternative. Elle semble bien conservée chez toutes les Eubactéries mais pas chez les Archées. La seconde partie de la glycolyse est très bien conservée et opérationnelle chez toutes les bactéries. Seule la pyruvate kinase est parfois absente, chez *Helicobacter pylori* et *Methanobacterium thermoautotrophicum* notamment (Dandekar *et al.*, 1999). Dans ce cas, le PEP peut être converti en pyruvate au cours de son utilisation par le système PTS (Postma *et al.*, 1993; Deutscher *et al.*, 2002) (cf partie 2.2). Il existe aussi chez *Rhizobium meliloti*, *Mycobacterium tuberculosis* et *Treponema pallidum*, une pyruvate dikinase active qui pourrait remplir ce rôle (Osteras *et al.*, 1997; Dandekar *et al.*, 1999).

Chez *Bacillus subtilis*, chaque étape est catalysée par une seule enzyme, capable de fonctionner lors de la glycolyse ou lors de la néoglucogenèse, excepté dans deux cas : l'interconversion du F6P et du FBP et l'interconversion du G3P et du BPG. La

néoglucogenèse n'est fonctionnelle que si les enzymes Fbp et GapB sont actives (et aussi PckA, qui ne fait pas partie de la voie d'Embden-Meyerhof (cf partie 1.4)).

1.1.4. Le bilan de la glycolyse

Au cours des 4 étapes enzymatiques (dans le sens glycolytique) de la partie supérieure de la glycolyse, le G6P est converti en G3P. Cette partie est consommatrice d'énergie, puisqu'il faut une molécule d'ATP pour phosphoryler le F6P en FBP. Si l'on tient compte de l'étape de phosphorylation précédente du glucose, soit par la voie du PTS, soit par la glucokinase, 2 molécules d'ATP sont nécessaires pour transformer une molécule de glucose en deux molécules de G3P.

Sur le plan énergétique, la conversion des trioses conduit, à partir d'une molécule de glucose préalablement transformée en deux molécules de G3P (DHAP), à la production de 4 molécules d'ATP et de deux molécules de NADH,H+. Chez *Bacillus subtilis*, en présence de glucose, on peut considérer qu'une molécule de PEP n'est pas convertie en pyruvate et est accaparée par le système PTS pour le transport et la phosphorylation d'une molécule de glucose suivante.

Une équation de la glycolyse chez *Bacillus subtilis*, considérée ici comme la transformation d'une molécule de G6P en 2 molécules de pyruvate, est donc la suivante :

$$1\ G6P + 2\ NAD^+ + 2\ ADP,Pi \quad (9\ \text{réactions}) \rightarrow \quad 2\ \text{pyruvate} + 2\ NADH,H^+ + 2ATP$$

Cette équation fait apparaître le faible rendement énergétique de la glycolyse. Une grande partie de l'énergie est en fait accumulée dans le NAD réduit et dans le pyruvate. Ce dernier est un intermédiaire métabolique très important. Il est utilisé pour produire de l'acétyl-coenzymeA (partie 1.3) qui permet le fonctionnement du cycle de Krebs et ainsi la production d'énergie et d'autres intermédiaires de l'anabolisme (partie 1.3). Lorsque la production de pyruvate par la glycolyse excède les besoins anaboliques, celui-ci est transformé en dérivés qui seront excrétés dans le milieu extracellulaire (partie 2.3.2). Chez les bactéries lactiques possédant un métabolisme fermentaire mixte, l'acétyl-CoenzymeA formé est converti en acétate et en éthanol fournissant une molécule d'ATP supplémentaire. Par ailleurs, le

NADH(H+) utilisé dans les réactions d'oxydation de la glycolyse est la source majeure de pouvoir réducteur qui sera utilisé lors de la respiration aérobie pour générer un gradient de protons converti en ATP (von Wachenfeldt & Hederstedt, 2002). En l'absence de respiration, le NADH,H+ est reoxydé au cours de la fermentation.

1.2. La voie des pentoses phosphate

1.2.1. Description générale

Utilisée pour la conversion du glucose en F6P et G3P, la voie des pentoses phosphate remplit néanmoins deux autres fonctions métaboliques importantes (Figure 4). Une partie du G6P est engagée dans la production d'une grande variété de molécules dont essentiellement le ribose-5-phosphate (ribose-5P), à la base de la biosynthèse de l'histidine, des nucléotides, du NAD, du tryptophane et de la thiamine, et l'érythrose-4-phosphate (érythrose-4P) qui sert à la synthèse du noyau aromatique dont sont constitués certains acides aminés et qui rentre dans la composition des quinones. En outre, la voie des pentoses phosphate est impliquée dans la génération de pouvoir réducteur, sous la forme de NADPH, qui est requis pour les réactions de biosynthèses.

La voie des pentoses phosphate classiquement décrite peut être divisée en deux parties (Figure 4) : la branche oxydative (*sensu stricto*) constituant la voie de conversion du G6P en un pentose phosphate, le D-ribulose-5-phosphate (ribulose-5P), avec formation de NADPH et dégagement de CO_2, et la branche non-oxydative constituant la voie de conversion de ce ribulose-5P en F6P et en G3P. De plus, la voie des pentoses phosphate est la seule voie d'utilisation des sources de carbone comme le gluconate, des sucres comme le ribose, l'arabinose et le xylose. Le catabolisme du gluconate conduit à la formation de ribulose-5P. Les trois dernières molécules conduisent respectivement à la formation de ribose-5P ou de xylulose-5-phosphate (xylulose-5P) qui sont interconvertibles dans la branche non-oxydative de la voie des pentoses phosphate (Figure 4).

1.2.2. Les étapes de la voie des pentoses phosphate

1.2.2.1. Branche oxydative

1.2.2.1.1. la glucose 1-déshydrogénase gdh (EC 1.1.1.47) et la glucose-6-phosphate 1-déshydrogénase zwf (EC 1.1.1.49)

Chez *Bacillus subtilis*, le gène *gdh* a été annoté comme codant une glucose 1-déshydrogénase NADP-dépendante par homologie avec les gènes de *Staphylococcus xylosus*

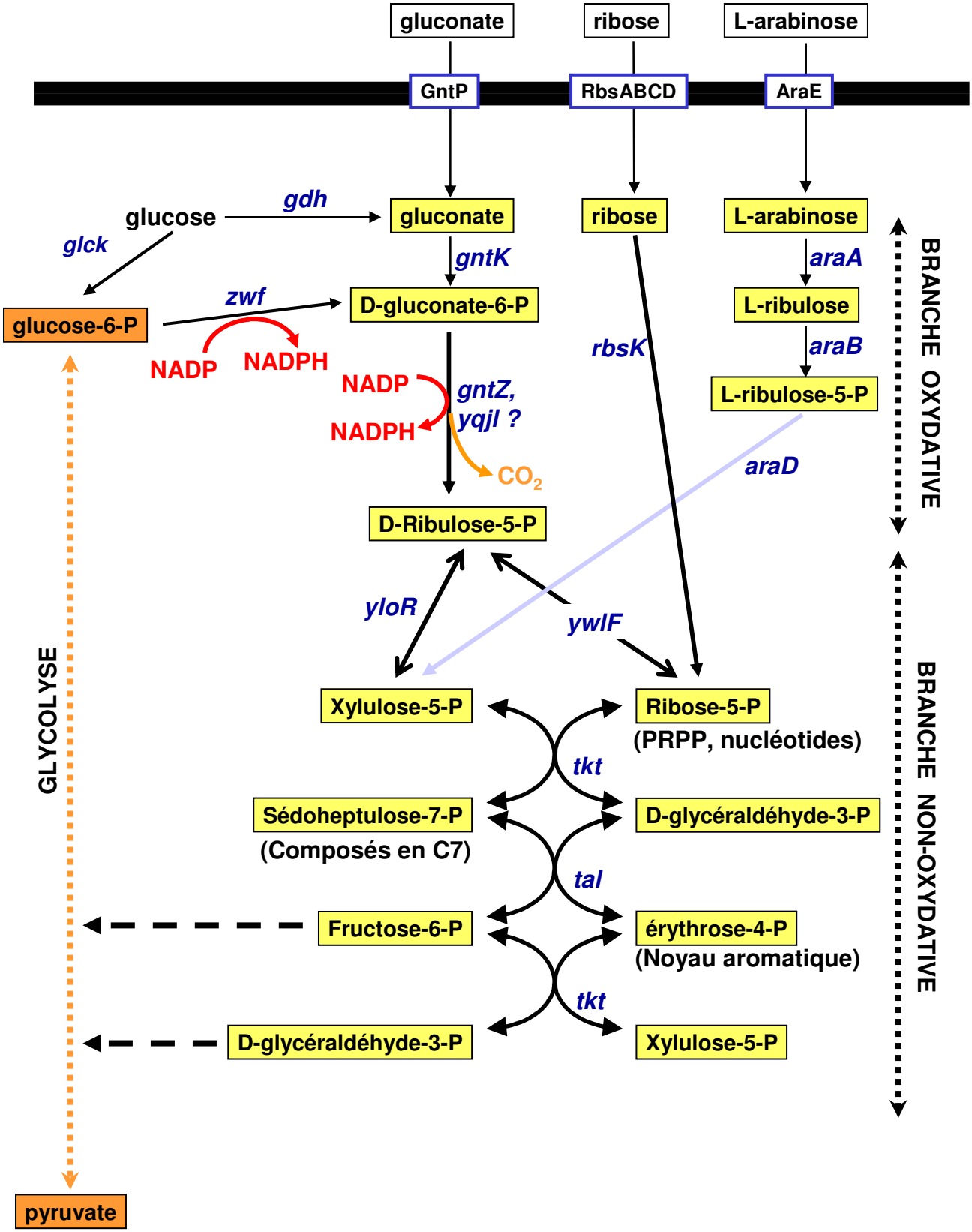


Figure 4. Voie des pentoses phosphates, principales sources de carbone y entrant et voies anaboliques rattachées chez *Bacillus subtilis*.

et *Bacillus megaterium* (Kataoka *et al.*, 1998) convertissant le glucose en gluconate. Ce gluconate cytoplasmique entre alors dans la voie des pentoses phosphate de façon classique, grâce à sa conversion en gluconate-6-phosphate (gluconate-6P).

En présence de glucose comme seule source carbone, une partie du G6P formé est détourné de la glycolyse et sert à la production des dérivés de la voie des pentoses phosphate. Ce G6P est tout d'abord oxydé en gluconate-6-phosphate (gluconate-6P). Chez *Bacillus subtilis*, cette conversion est catalysée de façon irréversible par la glucose-6-phosphate 1-déshydrogénase, codée par le gène *zwf*. Cette enzyme utilise comme cofacteur le NADP qui sera réduit au cours de l'oxydation du G6P. Aussi, cette étape constitue une source importante de NADPH nécessaire à certaines réactions anaboliques lorsque la bactérie croît en présence de glucose.

Chez *Escherichia coli* et *Lactococcus lactis*, le G6P est d'abord converti en gluconolactone-6-phosphate (gluconolactone-6P) par la glucose-6-phosphate déshydrogénase, qui est codée par le gène *zwf* (Rowley & Wolf, 1991). La fonction de ce gène a été mis en évidence dans un mutant *pgi* : en l'absence de la phosphoglucoisomérase, la totalité du G3P formé à partir du glucose provient de la voie des pentoses phosphate ; ainsi, un mutant *pgi-zwf* est incapable de croître en présence de glucose comme seule source de carbone (Fraenkel, 1968a; Fraenkel, 1968b). Plus récemment, des études de flux ont montré l'importance de *zwf* dans l'initiation de la voie des pentoses phosphate (Edwards & Palsson, 2000). Le gluconolactone-6P est ensuite converti en gluconate-6P par la phosphogluconolactanase, codée par le gène *pgl*. Cette deuxième étape n'est pas absolument requise puisque le gluconolactone-6P peut apparemment être transformé en gluconate-6P de façon non-enzymatique, ce qui permet à un mutant *pgl* de croître en présence de glucose (Kupor & Fraenkel, 1969; Kupor & Fraenkel, 1972).

1.2.2.1.2. La gluconokinase gntK (EC 2.7.1.12)

Chez *Bacillus subtilis* et *Escherichia coli*, le gluconate est perméé par un transporteur secondaire, codé respectivement par les gènes *gntP* et *gnt*, sans modifications. La gluconokinase, codée par le gène *gntK*, catalyse la conversion du gluconate intracellulaire en gluconate-6-phosphate.

Cette enzyme a été biochimiquement et structurellement caractérisée chez *Escherichia coli* (Izu *et al.*, 1996; Kraft *et al.*, 2001; Kraft *et al.*, 2002). Elle est active sous la forme d'un dimère de deux fois 18,4 kDa.

1.2.2.1.3. La 6-phosphogluconate déshydrogénase gntZ (EC 1.1.1.44)

La décarboxylation oxydative du gluconate-6P en D-ribulose-5-phosphate est catalysée par la phosphogluconate déshydrogénase en présence de NADP. Cette réaction est importante parce qu'elle fournit du NADPH, source de pouvoir réducteur pour l'anabolisme.

Chez les Eucaryotes, l'activité de cette enzyme est augmentée en présence de gluconate-6-P et est inhibée par le NADPH, cofacteur réduit au cours de la réaction (Rippa *et al.*, 1998).

S'agissant de *Bacillus subtilis*, la phosphogluconate déshydrogénase est codée par le gène *gntZ*. Cependant, il existe un autre gène, *yqjI*, dont le produit est homologue à une 6-phosphogluconate déshydrogénase (54 % d'identité avec GntZ) mais dont l'activité n'est pas démontrée.

Chez *Escherichia coli*, le gène *gnd* code la seule enzyme de la voie des pentoses phosphate dont l'activité est irréversible (Sprenger, 1995). L'expression de *gnd* est régulée par la vitesse de croissance mais le mécanisme transcriptionnel sous-jacent n'est pas bien établi (Sprenger, 1995). Une mutation de ce gène ne conduit pas à un phénotype observable lors de la croissance en présence de glucose. En effet, la bactérie est encore capable de produire les intermédiaires pentoses-phosphates nécessaires en utilisant le F6P et le G3P provenant de la glycolyse (Figure 1). Cependant, combinée avec une mutation du gène *pgi*, qui empêche la formation de F6P à partir du glucose *via* la glycolyse, cette mutation rend la souche incapable de pousser en présence de glucose comme seule source de carbone (Fraenkel, 1987).

1.2.2.2. Branche non-oxydative

1.2.2.2.1. La ribulose-5-phosphate épimérase (EC 5.1.3.1) et la ribose-5-phosphate isomérase (EC 5.3.1.6)

De façon réversible, 2 molécules de ribulose-5P sont converties en D-xylulose-5-phosphate par la ribulose-5-phosphate épimérase et en ribose-5P par la ribose-5-phosphate isomérase. Ces trois pentoses-phosphates peuvent être interconvertis dans un cycle à partir d'un seul d'entre eux, suivant la source de carbone assimilée (Figure 4) : le ribulose-5P à partir du gluconate-6P, le ribose-5P à partir du ribose et le xylulose-5P à partir du xylose ou de l'arabinose comme source de carbone.

Chez *Escherichia coli*, il existe deux gènes, *rpiA* et *rpiB*, codant des ribose-5-phosphate isomérases présentant très peu d'homologies entre-elles (Essenberg & Cooper, 1975; Hove-Jensen & Maigaard, 1993; Sprenger, 1995). RpiA est exprimée de façon constitutive et semble responsable à 99 % de l'activité ribose-5-phosphate isomérase lors de la croissance sur un milieu riche (Skinner & Cooper, 1971). Ainsi, un mutant *rpiA* est auxotrophe pour le ribose (Skinner & Cooper, 1971; Skinner & Cooper, 1974; Sorensen & Hove-Jensen, 1996). Un mutant *rpiB* ne montre pas de phénotype observable, bien que *rpiB* code une enzyme présentant une activité isomérase significative (Sorensen & Hove-Jensen, 1996). La séquence de RpiB est homologue à celle des sous-unités de la galactose-6-phosphate de *Streptococcus mutans* (40 % d'identité). L'expression de *rpiB* est induite par le ribose qui a pour effet d'inhiber l'activité du répresseur RpiR (Sorensen & Hove-Jensen, 1996). La structure tridimensionnelle de RpiA a été élucidée et montre clairement la structure dimérique de l'enzyme (Rangarajan *et al.*, 2002).

Un mutant *rpe* d'*Escherichia coli* est incapable de pousser sur des sucres en C5 et présente une croissance très altérée sur d'autres sources de carbone glycolytiques (Lyngstadaas *et al.*, 1998). Par ailleurs, *rpe* fait parti d'un large opéron de gènes de la biosynthèse des acides aminés aromatiques (Sprenger, 1995).

Chez *Bacillus subtilis*, ces deux enzymes ne sont pas caractérisées. Mais, l'analyse de leur séquence suggère que les gènes *yloR* et *ywlF* (47 % d'identité avec RpiB d'*Escherichia coli*) coderaient respectivement des enzymes homologues de Rpe et de RpiB d'*Escherichia coli*.

1.2.2.2.2. La transkétolase (2.2.1.1) et la transaldolase (EC 2.2.1.2)

Les étapes de la conversion de deux molécules de D-ribulose-5P en G3P et en F6P sont catalysées successivement par une transkétolase, une transaldolase puis, de nouveau, une transkétolase, selon un schéma complexe (Sprenger, 1995) (Figure 4). Au cours de cette conversion, les intermédiaires formés serviront à la biosynthèse d'autres molécules importantes : le ribose-5-P est le précurseur du PRPP qui entre dans la composition des nucléotides, l'érythrose-4-phosphate (érythrose-4P) est un précurseur des composés aromatiques et le sédoheptulose-7-phosphate (sédoheptulose-7P) est un précurseur de dérivés en C7 (Sprenger, 1995). Ainsi, 6 pentoses phosphate seront finalement convertis en 5 hexoses phosphates.

Chez *Escherichia coli*, il existe deux gènes codant deux isoformes de transkétolase, *tktA* et *tktB*, et deux gènes codant deux transaldolases isoformes, *talA* et *talB* (Sprenger, 1995). Un mutant dépourvu d'activité transkétolase est auxotrophe pour les composés aromatiques et incapable de pousser en présence de pentoses (Josephson & Fraenkel, 1969; Josephson & Fraenkel, 1974).

Chez *Bacillus subtilis*, cette voie a été mise en évidence mais pas étudiée en détail (De Wulf, 1998). Un gène a été identifié qui code une protéine présentant respectivement 47 % et 48 % d'identité avec TktA et TktB d'*Escherichia coli* (Rose & Entian, 1996). Le gène *ywjH*, dont le produit présente 33 % d'identité avec l'enzyme TalA d'*Escherichia coli*, pourrait coder une transaldolase.

1.2.3. Comparaison de la voie des pentoses phosphate chez *Bacillus subtilis*, *Lactococcus lactis, Escherichia coli, Helicobacter pylori* et *Methanobacterium thermoautotrophicum*

La voie des pentoses phosphate est essentielle chez les bactéries lactiques qui ne possèdent pas toutes les enzymes de la partie supérieure de la glycolyse (par exemple *Leuconostoc mesenteroïdes*, *Lactococcus brevis*), d'où l'appellation de voie hétérolactique (par contraste avec la voie homolactique/glycolyse). Chez celles-ci, la voie décrite par Thomson (Thompson & Gentry-Weeks, 1994) diffère de celle classiquement décrite chez

Bacillus subtilis et *Escherichia coli* (Fraenkel, 1987) dans la partie non-oxydative (Figure 5). En effet, le xylulose-5P n'est pas transformé en G3P et en F6P mais en G3P et en acétyl-phosphate (acétyl-P) par une seule enzyme, la xylulose-5-phosphate phosphokétolase, en présence de thiamine-pyrophosphorylase (Veiga-da-Cunha *et al.*, 1993; Posthuma *et al.*, 2002). Le G3P rejoint la voie glycolytique qui aboutira à la production de lactate tandis que l'acétyl-P est réduit en acétate, avec production d'une molécule d'ATP, ou en éthanol. Cependant, dans une récente étude, il a été montré que la souche de *Lactococcus lactis* IO-1 possédait deux voies cataboliques du D-xylose, l'une correspondant à la voie précédemment décrite et l'autre correspondant à l'utilisation d'une voie non oxydative comparable à celle de *Bacillus subtilis* (Tanaka *et al.*, 2002). Cela implique donc l'existence d'une transkétolase et d'une transaldolase. L'analyse du génome de la souche LOO43 révèle l'existence d'un gène dont le produit possède respectivement 55 % et 47 % d'identité avec le produit du gène *tkt* de *Bacillus subtilis* et avec TktA et TktB d'*Escherichia coli*. Cependant, aucun gène ne semble coder pour une transaldolase de classe connue.

A l'instar de *Bacillus subtilis*, *Escherichia coli* semble posséder l'une voie des pentoses phosphate la plus complète parmi les 5 considérées (Figure 6). Remarquablement, à l'exception de la ribulose-5-phosphate épimérase, les enzymes de la branche non-oxydative de la voie sont dupliquées. Pour chaque étape, l'une des deux enzymes impliquée pourrait intervenir dans une fonction secondaire.

Helicobacter pylori possède une voie complète (Figure 6), mais ne semble pas capable d'utiliser le gluconate comme source de carbone puisqu'elle ne possède pas de gène codant une protéine homologue à une gluconate kinase connue.

Methanobacterium thermoautotrophicum constitue un cas à part. En effet, cette bactérie ne possède pas de voie des pentoses phosphate semblable à celle décrite chez *Bacillus subtilis* ou *Escherichia coli* (Figure 6). Par contre, elle possède une désoxyribose-phosphate aldolase (EC 4.1.2.4) qui lui permet de convertir le 2-désoxyribose-5-phosphate en G3P, intermédiaire glycolytique. Cette voie existerait aussi chez toutes les bactéries ici considérées à l'exception d'*Helicobacter pylori* qui ne possède pas de gène homologue à une désoxyribose-phosphate aldolase.

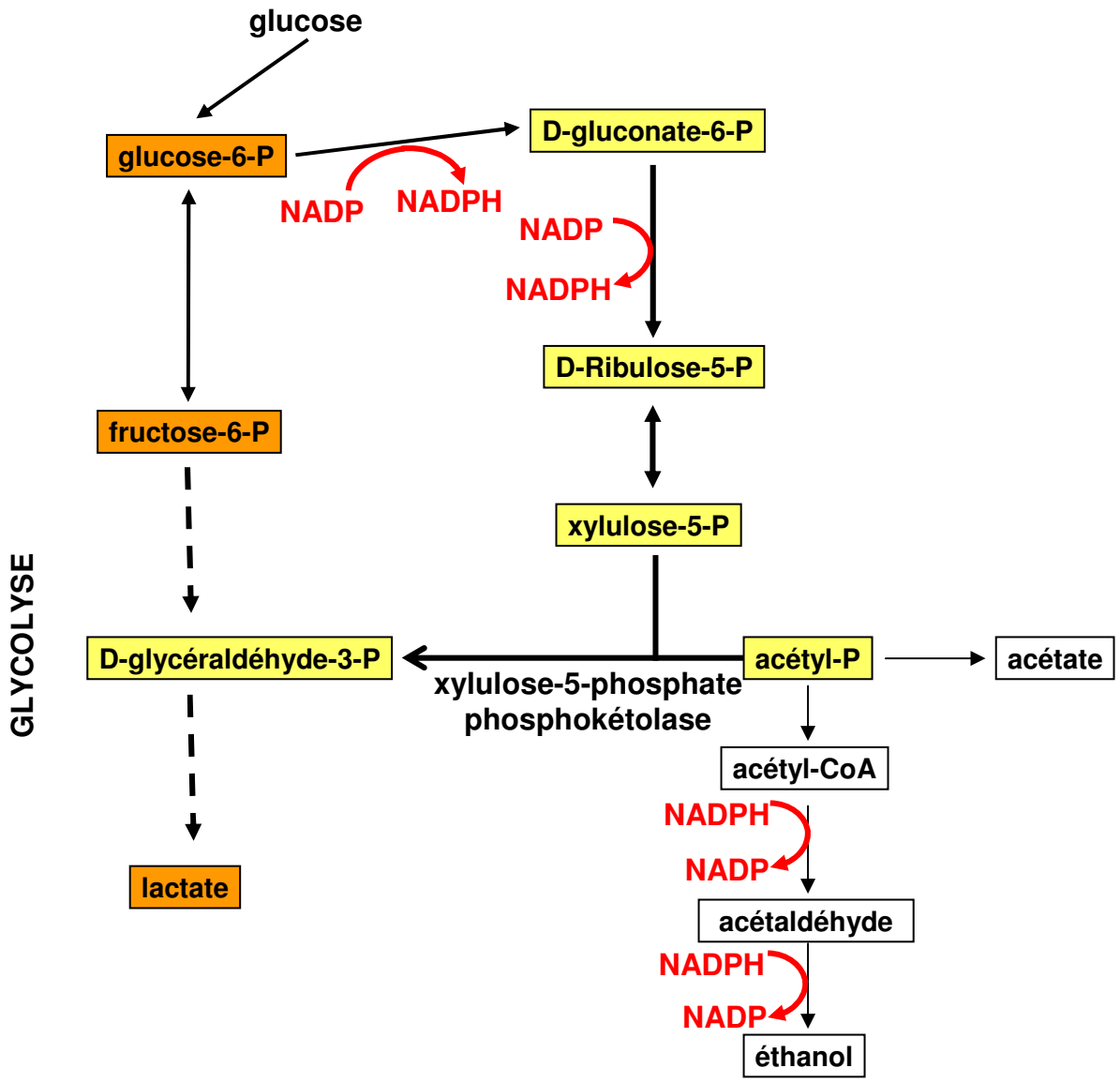


Figure 5. Voie des pentoses phosphates classiquement décrite chez les bactéries lactiques.

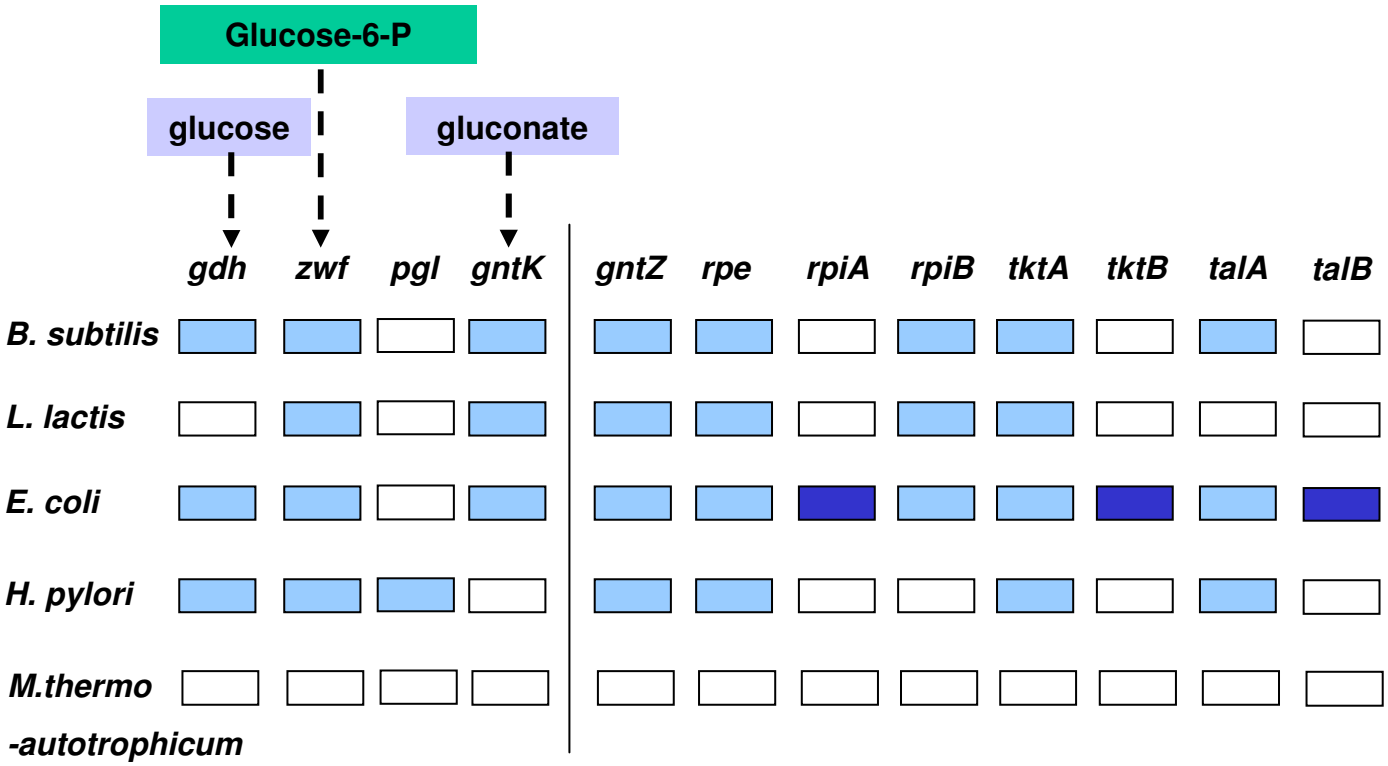


Figure 6. Comparaison des étapes de la voie des pentoses phosphates chez 5 bactéries modèles. Les étapes décrites permettent, d'une part, l'entrée du glucose ou du glucose-6-P dans la voie des pentoses phosphates, et d'autre part, l'assimilation directe du gluconate pour fournir les intermédiaires anaboliques de la voie des pentoses phosphates mais aussi le FBP et le G3P pour la glycolyse. Lorsqu'il existe un gène paralogue, celui-ci est indiqué en bleu foncé. Les cases blanches indiquent les étapes enzymatiques pour lesquelles aucun gène n'a pu être identifié.

1.3. Le cycle de Krebs

1.3.1. Description générale

Décrit pour la première fois en 1937 par Sir Hans Krebs, le cycle de Krebs exerce des fonctions à l'interface du catabolisme et de l'anabolisme. Outre son implication dans le métabolisme du carbone, il fonctionne comme un cycle d'oxydation (en aérobiose) et fournit de l'ATP et du pouvoir réducteur pour la respiration oxydative. Au cours des 8 étapes de ce cycle dit aussi cycle des acides tricarboxyliques, les deux atomes de carbone du groupement acétyl de l'acétyl-coenzymeA seront relachés sous forme de CO_2 et une molécule d'oxaloacétate sera régénérée (Figure 7).

Ce cycle joue plusieurs rôles dans le métabolisme du carbone et plus précisément dans le métabolisme des acides carboxyliques. En premier lieu, il fournit des intermédiaires carbonés à l'anabolisme (Figure 7). L'acétyl-CoA peut entrer dans le cycle de Krebs ou être intégré dans la voie de biosynthèse des acides gras, ou encore être converti en acétate avec production d'ATP. Le 2-oxoglutarate est un précurseur du glutamate, composé très important pour *Bacillus subtilis* puisqu'il est la source d'un très grand nombre de molécules. Il est soit produit à partir des trois premières étapes du cycle, soit fourni par l'assimilation d'une source d'azote comme le glutamate, la glutamine ou l'ammonium (Figure 7). Il se situe donc à la jonction avec le métabolisme de l'azote (Reitzer & Magasanik, 1987; Fisher & Débarbouillé, 2002). L'oxaloacétate est un précurseur de l'aspartate et de l'acétyl-CoA. Il est, soit produit au cours du cycle de Krebs, soit généré par la pyruvate carboxylase (partie 1.4) à partir du pyruvate, étape en fait nécessaire puisqu'une certaine quantité d'oxaloacétate est requise pour que le cycle « démarre » (Figure 1). L'assimilation de l'aspartate est aussi une source d'oxaloacétate (Figure 7). Tous ces métabolites intermédiaires sont nécessaires à l'anabolisme à partir de l'oxydation du pyruvate. Aussi, il suffit théoriquement de fournir ce cycle en un seul intermédiaire pour qu'il puisse démarrer et fonctionner.

Par ailleurs, les réactions permettant de générer l'oxaloacétate, que l'on peut considérer comme le point de départ du cycle lorsque la glycolyse fonctionne, sont appelées réactions anaplérotiques du cycle de Krebs. Parmi ces réactions, trois sont considérées comme majeures, qui sont catalysées par la pyruvate carboxylase à partir du pyruvate, la PEP carboxylase à partir de PEP et l'enzyme malique à partir de malate (Figure 7) (cf partie 1.4).

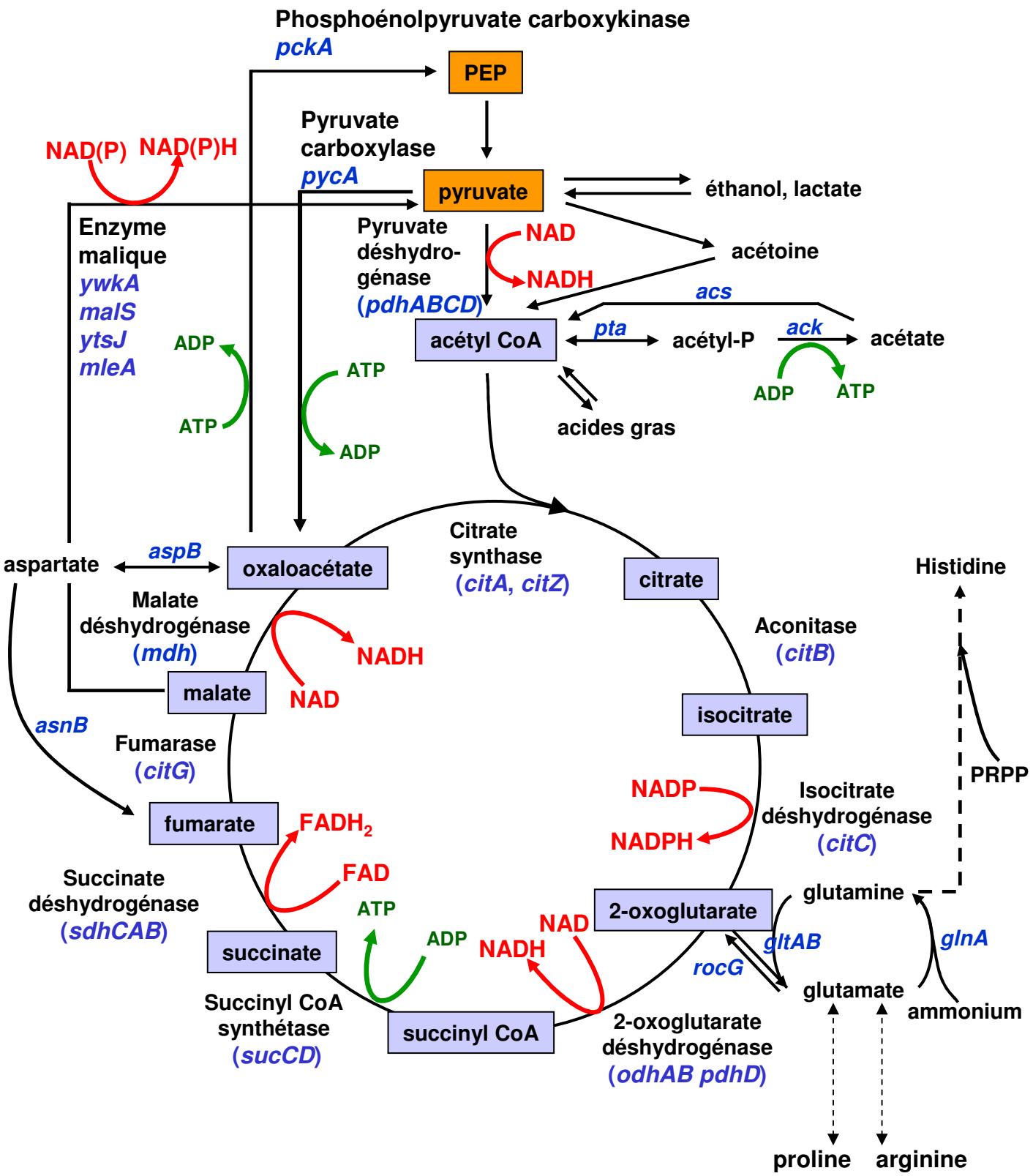


Figure 7. Cycle de Krebs et voies s'y rattachant chez *Bacillus subtilis*. Lors de la respiration aérobie, la voie fonctionne comme un cycle d'oxydation (dans le sens des aiguilles d'une montre).

D'après Sonenshein, 2002.

Cette dernière permet en fait de produire du pyruvate à partir du malate qui sera ensuite converti en oxaloacétate. Quant à la PEP carboxylase, elle n'existe pas chez *Bacillus subtilis*.

Dans un cycle complet, quatre réactions produisent du pouvoir réducteur sous forme de nucléotides réduits lors de la croissance aérobie (NADH, NADPH et $FADH_2$) (Figure 7) : ces réactions impliquent l'isocitrate déshydrogénase, le complexe oxoglutarate déshydrogénase, la succinate déshydrogénase et la malate déshydrogénase. Par ailleurs, la succinate déshydrogénase, également décrite comme une succinate:ménaquinone oxidoréductase, participe directement au transport des électrons dans la chaîne respiratoire (von Wachenfeldt & Hederstedt, 2002), montrant ainsi le lien entre le cycle de Krebs et la respiration oxydative. La succinate thiokinase, ou succinyl-coA synthétase permet également la production directe d'une molécule d'ATP par cycle.

Lorsque *Bacillus subtilis* pousse uniquement sur des sources de carbone qui entrent par le cycle de Krebs, ces substrats sont convertis en oxaloacétate dans le cycle, qui constitue alors la principale source de carbone pour la néoglucogenèse : l'oxaloacétate sera converti en PEP par la phosphoénolpyruvate carboxykinase (partie 1.4).

Chez *Bacillus subtilis*, le cycle de Krebs joue aussi un rôle dans le déclenchement de la sporulation puisque toute délétion d'un gène du cycle entraîne un défaut de sporulation (Sonenshein, 2002).

1.3.2. Les enzymes du cycle de Krebs

Depuis longtemps, les enzymes impliquées dans le cycle de Krebs sont très bien caractérisées sur le plan biochimique chez les Eucaryotes et les bactéries (Weitzman, 1981; Nimmo, 1987; Hederstedt, 1993; Sonenshein, 2002).

1.3.2.1. Les citrate synthases *cytA* et *cytZ* (EC 4.1.3.7)

La citrate synthase catalyse la condensation de l'acétyl-CoA avec l'oxaloacétate en présence d'eau pour former du citrate et libérer le coenzyme A. Dans les conditions physiologiques, cette réaction n'est pas réversible (Wiegand & Remington, 1986) et fait que le cycle complet ne tourne que dans un seul sens.

Chez *Bacillus subtilis*, il existe deux gènes codant deux isoformes de la citrate synthase, *citZ* codant l'enzyme majeure et *citA* (Jin & Sonenshein, 1994; Jin *et al.*, 1996). Le gène *citA* a été identifié par complémentation d'un mutant d'*Escherichia coli* dépourvu d'activité citrate synthase (Jin & Sonenshein, 1996). Mais la délétion de *citA* n'entraîne qu'une diminution faible de l'activité citrate synthase et a peu d'effet sur l'efficacité de sporulation. L'activité citrate synthase restante a donc été assignée au produit du gène *citZ* (Jin & Sonenshein, 1994). Un mutant thermosensible a été décrit qui nécessite l'ajout de citrate ou de glutamate pour la croissance en conditions non permissives (Carls & Hanson, 1971). CitA et CitZ sont des homodimères de deux fois 40,8 et 41,6 kDa, respectivement. CitZ a été étudiée biochimiquement. Ses caractéristiques enzymatiques sont proches de celles de la citrate synthase de *Bacillus megaterium* (Robinson *et al.*, 1983). Son activité n'est notamment pas inhibée par une concentration relativement élevée de NADH comme c'est le cas chez *Escherichia coli* mais par l'ATP et l'acétyl-CoA. Chez cette dernière bactérie et chez la plupart des bactéries à Gram négatif, l'unique citrate synthase, codée par le gène *gltA*, est présente sous la forme un homohexamère régulé allostériquement par le NADH et l'α-oxoglutarate (Flechtner & Hanson, 1970).

1.3.2.2. L'aconitase *cytB* (EC 4.2.1.3)

L'isomérisation du citrate en isocitrate est catalysée par l'aconitase (ou aconitate hydratase) avec comme intermédiaire de réaction l'aconitate :

citrate → aconitate + H2O → isocitrate.

Chez *Bacillus subtilis*, l'aconitase est le produit du gène *citB*. C'est une enzyme vraisemblablement monomérique de 120 kDa qui est probablement active sous une forme complexée fer-soufre (Dingman & Sonenshein, 1987). Un mutant *citB* requiert l'ajout d'isocitrate ou de 2-kétoglutarate pour la croissance et est bloqué à une étape précoce de la sporulation (Dingman *et al.*, 1987; Craig *et al.*, 1997).

Récemment, Alen et Sonenshein ont montré que l'aconitase était capable de se lier à une séquence ARNm-spécifique, appelée iron responsive element (IRE), qui est présente dans la séquence des opérons qui codent la cytochrome oxydase majeure et le système de captage du fer (Alen & Sonenshein, 1999). Cette activité non enzymatique est fonction de la disponibilité en fer, comme c'est le cas des protéines eukaryotiques similaires qui régulent la

traduction et la stabilité des ARNm des gènes du métabolisme du fer en fonction de la disponibilité en fer. De plus, chez *Bacillus subtilis*, il a été montré qu'un mutant dépourvu d'activité aconitase mais possédant toujours la capacité de lier l'ARNm sporule 40 fois mieux qu'un mutant nul de l'aconitase. Cela suggère que l'aconitase est une enzyme bifonctionnelle qui joue un rôle dans le cycle de Krebs en présence de fer et un rôle dans la cascade de sporulation en l'absence de fer (Alen & Sonenshein, 1999). Très récemment, la capacité de l'aconitase à interagir avec l'ARNm a également été étudiée chez *Escherichia coli* qui possède au moins deux gènes paralogues qui codent l'activité aconitase, *acnA* et *acnB*, et dont les produits ont été caractérisés (Jordan *et al.*, 1999; Varghese *et al.*, 2003).

1.3.2.3. L'isocitrate déshydrogénase *cytC* (EC 1.1.1.42)

Cette enzyme catalyse l'oxydation de l'isocitrate en 2-oxoglutarate avec réduction du NADP en NADPH.

L'enzyme de *Bacillus subtilis*, un homodimère de 2 fois 44 kDa (Singh *et al.*, 2001), est probablement allostériquement négativement régulée par l'oxoglutarate et l'oxaloacétate (Shikata *et al.*, 1988). La délétion du gène *cytC* chez *Bacillus subtilis* bloque la sporulation au stade I (Jin *et al.*, 1997), phénotype du à l'accumulation de citrate et d'isocitrate diminuant fortement le pH intracellulaire (Matsuno *et al.*, 1999). Mais la cellule est encore viable. En effet, une partie du cycle peut-être court-circuitée, l'oxaloacétate, notamment, pouvant être directement produit par la pyruvate carboxylase à partir du pyruvate (Fisher & Magasanik, 1984).

Chez *Escherichia coli*, cette enzyme n'est pas allostériquement régulée par un métabolite. Son activité dépend de son état de phosphorylation (Nimmo, 1987; Edlin & Sundaram, 1989). L'enzyme, codée par le gène *icd*, est inactive lorsqu'elle est complètement phosphorylée par l'isocitrate déshydrogénase kinase. Mais son activité est restaurée lorsqu'elle est déphosphorylée par l'isocitrate déshydrogénase phosphatase. En réalité, il s'agit d'une enzyme bifonctionnelle, codée par le gène *aceK*, dont l'activité est soumise à une régulation allostérique par deux classes de métabolites (Miller *et al.*, 1996; Miller *et al.*, 2000). Dans le premier cas, l'effecteur se lie au site actif de la kinase/phosphatase. Le NADPH inhibe les activités kinase et phosphatase et l'isocitrate inhibe l'activité kinase et stimule l'activité phosphatase, résultant en une activité isocitrate déshydrogénase. Dans le second cas, le ligand (pyruvate, AMP, 3PG) se lie à l'enzyme indépendamment du site

ATPase. Un rôle physiologique de ce système de régulation a été proposé (Holms & Bennett, 1971). L'inactivation partielle de l'isocitrate déshydrogénase pendant la croissance en présence d'acétate facilite le flux par la voie cyclique du glyoxylate qui court-circuite complètement les étapes formant le 2-oxoglutarate et le succinyl-Coenzyme A et produit directement du malate (Figure 8). Cette dernière voie est appelée shunt du glyoxylate et est aussi une réaction anaplérotique. Cependant, chez *Bacillus subtilis* et les autres Bacillacées, elle n'est pas fonctionnelle parce que les gènes qui codent l'isocitrate lyase et la malate synthase sont absents.

1.3.2.4. Le complexe oxoglutarate décarboxylase (ODHC) *odhAB*, *pdhD*

Ce complexe multienzymatique catalyse la conversion irréversible du 2-oxoglutarate en succinyl-Coenzyme A avec réduction du NAD en NADH. Une unité enzymatique est composé de trois sous-unités, l'oxoglutarate décarboxylase (E1o, EC 1.2.4.1), la dihydrolipoamide succinyltransférase (E2o, EC 2.3.1.61) et la lipoamide déshydrogénase (E3, 1.8.1.4) commune au complexe pyruvate déshydrogénase (Perham, 1991; Hederstedt, 1993).

Chez *Bacillus subtilis*, de nombreux mutants sont disponibles (Hederstedt, 1993). Ces mutants sont capables d'utiliser le glucose comme seule source de carbone suggérant une faible activité du cycle de Krebs dans des conditions où le flux glycolytique est important. Par contre, ils ne peuvent pas accomplir un cycle complet. Ils sont donc incapables de croître en présence d'acides tricarboxyliques ou de glutamate comme sources de carbone puisque le cycle se trouve bloqué au niveau du 2-oxoglutarate. De plus, la croissance en présence de maltose, une source de carbone utilisée peu efficacement, nécessite l'ajout d'acides dicarboxyliques (succinate, fumarate, malate ou oxaloacétate) ou d'aspartate pour une croissance comparable à celle de la souche sauvage.

Chez *Escherichia coli*, le complexe ODHC est strictement requis même en présence de glucose comme source de carbone (Nimmo, 1987).

1.3.2.5. La succinate thiokinase (ou succinyl-CoA synthétase) *sucCD* (EC 6.2.1.5)

Cette enzyme catalyse l'interconversion du succinyl-CoA et du succinate et produit une molécule d'ATP (ou de GTP) à partir d'ADP (ou de GDP) et par molécule de subtrat transformé. Deux sous-unités, α et β, forment l'enzyme active.

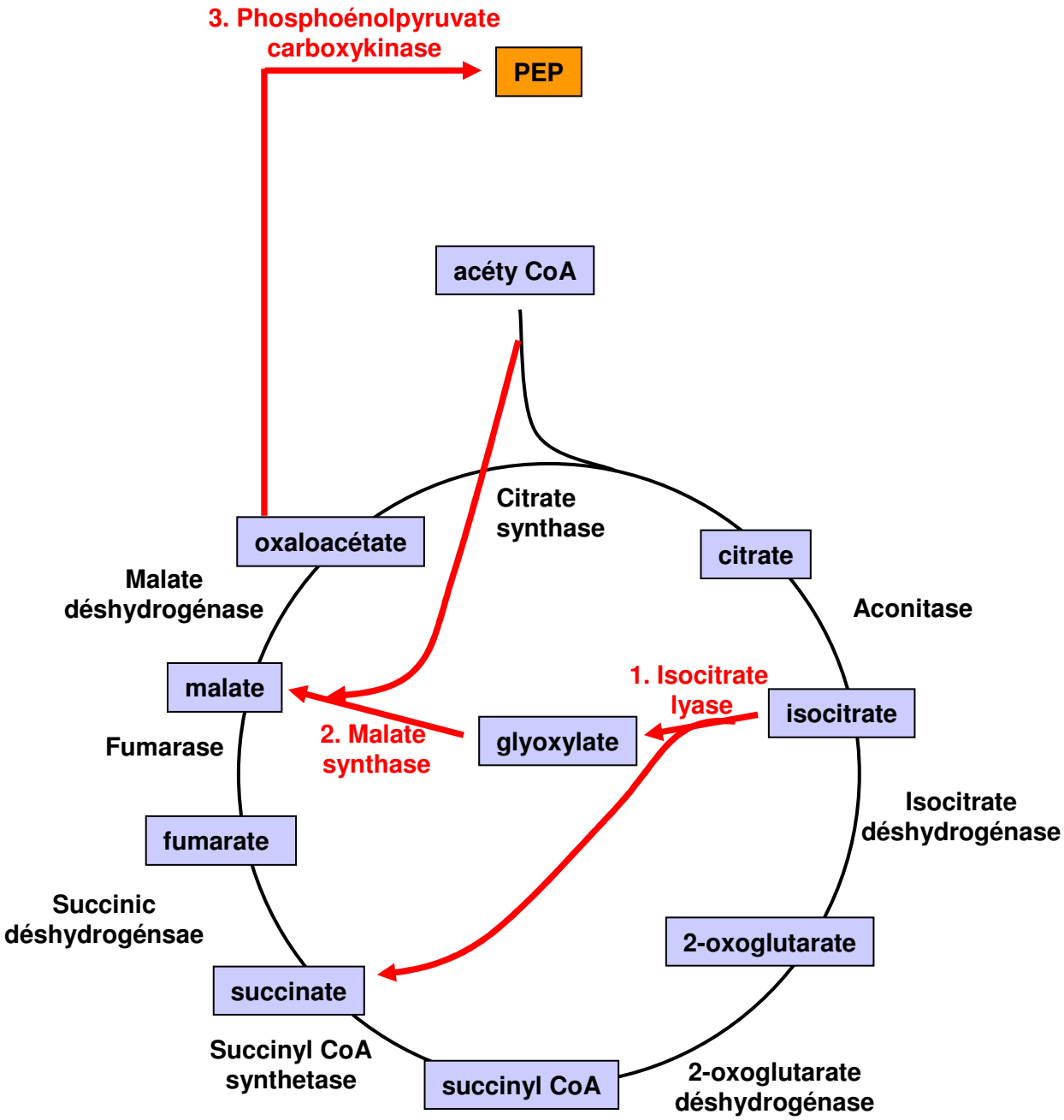


Figure 8. Shunt du glyoxylate chez *Escherichia coli*. Les réactions vont de 1 à 3. D'après Nimmo, 1987.

Chez *Bacillus subtilis* et les bactéries à Gram positif, ces sous-unités sont petites, de l'ordre de 75 kDa (Weitzman, 1981), comparées à celles trouvées chez les bactéries à Gram négatif comme *Escherichia coli* qui font 140 à 180 kDa (Miles & Guest, 1987). L'enzyme de *Bacillus subtilis* semble être ADP-dépendante puisqu'aucune activité en présence de GDP n'a pu être détectée (Weitzman & Jaskowska-Hodges, 1982).

1.3.2.6. La succinate déshydrogénase ou succinate :ménaquinone oxydoréductase *sdhCAB* (EC 1.3.5.1)

Cette enzyme catalyse l'interconversion du succinate et du fumarate avec réduction concomitante d'une molécule de FAD en $FADH_2$. Du fait de son double rôle, dans le cycle de Krebs et dans la respiration (transfert d'électron), cette enzyme membranaire a été largement étudiée (Hederstedt, 1993; von Wachenfeldt & Hederstedt, 2002).

Chez *Bacillus subtilis*, les trois sous-unités de cette enzyme sont codées par les gènes *sdhC* (cytochrome), *sdhA* (flavoprotéine) et *sdhB* (protéine fer-soufre). Chez *Escherichia coli*, le cytochrome est codé par deux gènes, *sucC* et *sucD* (Nimmo, 1987).

1.3.2.7. La fumarase *citG* (EC 4.2.1.2)

La fumarase catalyse l'hydratation réversible du fumarate en malate.

La fumarase de *Bacillus subtilis*, codée par le gène *citG*, est un homotétramère de 4 fois 50,3 kDa. Un mutant *citG* présente une croissance très altérée en milieu riche mais qui peut être restaurée par l'ajout de malate (Ohne, 1975).

Chez *Escherichia coli*, trois fumarases existent et sont distinguables en deux classes (Yumoto & Tokushige, 1988). Les gènes *fumA* et *fumB* codent des enzymes homodimériques qui fonctionnent respectivement dans le cycle de Krebs et dans la production de fumarate comme accepteur d'électron en anaérobiose (Tseng, 1997). Le gène *fumC* code la protéine homologue à celle de *Bacillus subtilis* (60 % d'identité) qui est active dans le cycle de Krebs (Weaver *et al.*, 1993; Weaver *et al.*, 1995).

1.3.2.8. La malate déshydrogénase *mdh* (EC 1.1.1.37)

L'oxydation du malate en oxaloacétate est catalysée par la malate déshydrogénase avec réduction du NAD(P) en NAD(P)H (Musrati *et al.*, 1998). Dauner et ses collaborateurs ont observé que la réaction pouvait aussi avoir lieu dans le sens réducteur lors de la croissance en chémostat et limitation en phosphate (Dauner *et al.*, 2001a). C'est une enzyme dimérique

mdh, le gène codant la malate déshydrogénase de *Bacillus subtilis* a été identifié récemment (Jin *et al.*, 1996) Il code probablement une enzyme homotétramérique de 4 fois 33,5 kDa (Musrati *et al.*, 1998) strictement NAD-dépendante. La disruption de ce gène entraîne une très légère auxotrophie pour l'oxaloacétate et l'aspartate en milieu minimum. En effet, l'oxaloacétate peut être fourni par la combinaison de l'enzyme malique et de la pyruvate carboxylase (Figure 7). Par ailleurs, un tel mutant ne peut pas sporuler sans ajout d'aspartate (Jin *et al.*, 1996).

Chez *Escherichia coli*, deux malate déshydrogénases sont présentes. En effet, cette bactérie possède une malate déshydrogénase cytoplasmique bien étudiée (EC 1.1.1.37) codée par le gène *mdh* et une malate:quinone oxydoréductase (EC 1.1.99.16) codée par le gène *mqo*. Dans les conditions physiologiques, l'enzyme codée par *mdh* permettrait la réduction de l'oxaloacétate alors que celle codée par *mqo* catalyserait l'oxydation du malate (van der Rest *et al.*, 2000). En outre, une délétion de *mdh* conduit à une déficience sévère de croissance en présence de pyruvate ou de lactate comme seule source de carbone et une absence de croissance en présence de malate ou de succinate, alors qu'une délétion de *mqo* ne se traduit pas par un défaut de croissance quelque soit la source de carbone testée, le cycle de Krebs fonctionnant alors *via* le shunt du pyruvate (van der Rest *et al.*, 2000). Toutefois, la délétion combinée des deux gènes se traduit par un défaut de croissance plus important que celui observé avec un simple mutant *mdh*. *mqo* peut donc remplir partiellement le rôle de malate déshydrogénase.

1.3.3. Organisation comparée du cycle de Krebs chez *Bacillus subtilis*, *Lactococcus lactis*, *Escherichia coli*, *Helicobacter pylori* et *Methanobacterium thermoautotrophicum*

La Figure 9 montre l'état des connaissances des étapes et des enzymes du cycle de Krebs chez 5 bactéries modèles. Parmi elles, seules *Bacillus subtilis* et *Escherichia coli* possèdent l'ensemble des gènes codant un cycle de Krebs complet. La partie la mieux conservée du cycle est constituée des étapes de conversion des acides dicarboxyliques, du succinate en oxaloacétate.

Le shunt du glyoxylate est manquant chez *Bacillus subtilis*. En revanche, *Bacillus subtilis* est capable de produire l'oxaloacétate à partir du pyruvate à partir de la pyruvate carboxylase.

Chez *Lactococcus lactis*, il manque certaines enzymes pour réaliser un cycle complet. En effet, la plupart des bactéries lactiques et certaines bactéries à Gram positif à faible GC % ne possèdent pas de cycle de Krebs complet, notamment les enzymes de la branche des acides dicarboxyliques. Ces bactéries sont souvent auxotrophes pour les intermédiaires métaboliques du cycle (Sonenshein, 2002). La partie la plus complète correspond aux deux premières réactions du cycle. Le complexe PDH sert à produire l'acétyl-CoA qui sera par la suite converti en acétate et éthanol avec production d'ATP.

Escherichia coli possède l'un des cycles de Krebs les plus complets parmi les bactéries. De plus, cette voie est fonctionnelle aussi bien lors de la respiration aérobie qu'en l'absence d'oxygène. Elle possède notamment la voie de métabolisation du glyoxylate, voie qui est partiellement utilisée pour court-circuiter une partie du cycle des acides tricarboxyliques en présence d'acétate comme source de carbone. L'activité pyruvate carboxylase est remplacée par l'activité PEP carboxylase qui produit l'oxaloacétate à partir du PEP. Une récente caractérisation de l'activité des enzymes du cycle (Pitson *et al.*, 1999), conjointement aux données génomiques (Tomb *et al.*, 1997), a permis de proposer un modèle du fonctionnement du cycle de Krebs chez *Helicobacter pylori*. Concernant l'activité malate déshydrogénase démontrée (Pitson *et al.*, 1999), aucun gène codant une enzyme homologue à une malate déshydrogénase de classe connue n'a pu être identifié. De même, l'analyse du génome de cette bactérie n'a pas permis d'identifier les gènes codant la malate synthase pour le shunt du glyoxylate et la 2-oxoglutarate oxydase. Cette dernière enzyme semble remplacer les étapes catalysées par la 2-oxoglutarate déshydrogénase et la succinyl-coA synthétase.

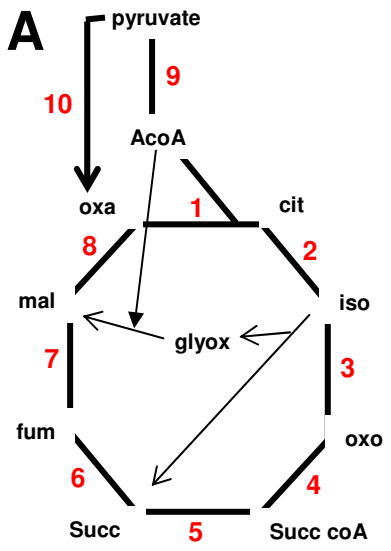
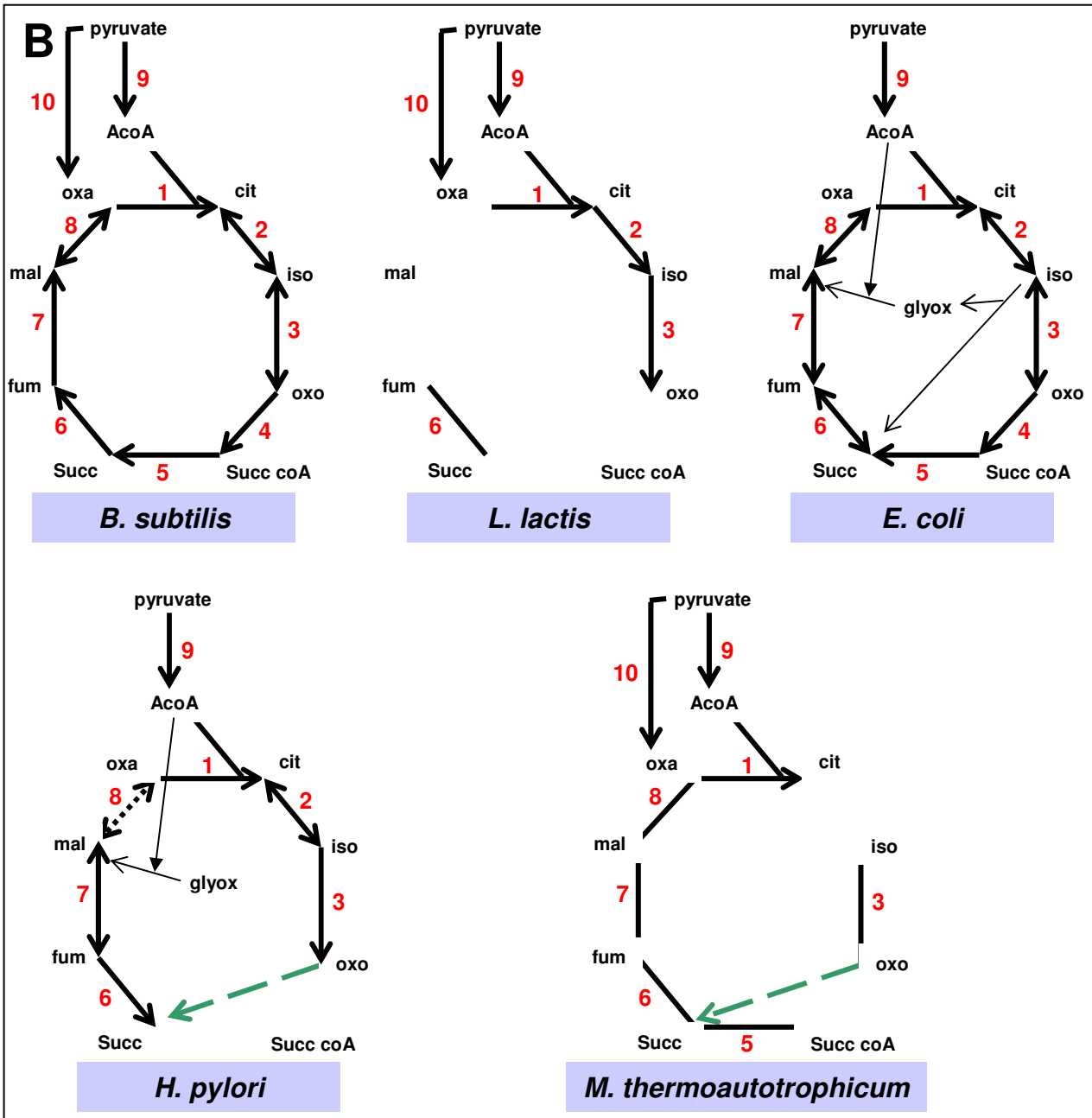


Figure 9. comparaison des étapes du cycle de Krebs présentes chez différentes espèces bactériennes.

A. Schéma théorique d'un cycle de Krebs complet et du shunt du glyoxylate. Les numéros indiqués correspondent aux gènes suivant : 1. Citrate synthase ; 2. Aconitase ; 3. isocitrate déshydrogénase ; 4. 2-oxoglutarate déshydrogénase ; 5. succinyl-coA synthétase ; 6. succinate déshydrogénase ; 7. fumarase ; 8. malate déshydrogénase ; 9. pyruvate déshydrogénase ; 10. pyruvate carboxylase.

B. Cycle de Krebs et shunt du glyoxylate (s'il existe) chez *B. subtilis*, *L. lactis*, *E. coli*, *H. pylori* et *M. thermoautrophicum*. Lorsque les données permettent de le déterminer, les flèches indiquent le ou les sens possibles de la réaction. Chez *H. pylori* et *M. thermoautrophicum*, il semble que le 2-oxoglutarate soit directement converti de façon non réversible en succinate par la 2-oxoglutarate oxydase (EC 1.2.7.3) (flèche verte discontinue). Par ailleurs, une activité malate déshydrogénase a été montrée chez *H. pylori* (traits discontinus), bien qu'aucun gène correspondant n'a pu être reconnu sur la base de la séquence.

D'après Huynen et al, 1999.



Ainsi, *Helicobacter pylori* semble posséder une voie non cyclique qui servirait plutôt à la production d'intermédiaires anaboliques que de source d'énergie et de pouvoir réducteur (Pitson *et al.*, 1999).

Chez *Methanobacterium thermoautotrophicum* et les autres Archées, il n'existe pas de cycle de Krebs complet (Huynen *et al.*, 1999). L'oxoglutarate est produit par réduction du pyruvate et non pas oxydation par la branche des acides dicarboxyliques.

1.3.4. Fonctionnement du cycle de Krebs en aérobiose et en anérobiose

1.3.4.1. Chez *Bacillus subtilis*

Le cycle de Krebs est historiquement lié à l'étude de la respiration oxydative (Sulek, 1969). La découverte relativement récente de la capacité de *Bacillus subtilis*, que l'on croyait jusqu'alors être un organisme aérobie strict, à croître en anaérobiose (Nakano & Zuber, 1998) a ouvert l'étude de l'adaptation de son métabolisme, en particulier le cycle de Krebs, à cette « nouvelle » condition physiologique.

Lors de la croissance en aérobiose en présence d'un excès de glucose comme seule source de carbone et d'une source d'azote (glutamate ou ammonium), la glycolyse suffit à la cellule pour produire de l'énergie directement sous la forme d'ATP par la conversion du glucose en pyruvate et indirectement sous la forme de pouvoir réducteur converti par la voie de phosphorylation oxydative. De plus, la conversion de l'acétyl-CoA en acétate, alors exporté, fournit de l'ATP supplémentaire. Le cycle de Krebs ne fonctionne alors que pour générer suffisamment de précurseurs pour l'anabolisme. Les activités de la citrate synthase et de l'aconitase sont alors très faibles et réprimées par le glutamate (Nakano & Zuber, 1998). Le pool de cofacteurs oxydés, nécessaire à son fonctionnement, est alors reconstitué par la chaîne respiratoire. Si le glucose vient à manquer, le flux à travers le cycle de Krebs peut alors augmenter *via* l'assimilation du glutamate comme source de carbone de rechange ou des produits secondaires précédemment excrétés (Cruz-Ramos *et al.*, 1997; Nakano *et al.*, 1997; Nakano & Hulett, 1997) mais, en contre partie, la croissance sera plus lente.

Dans des conditions de croissance anaérobie, la glycolyse fournit toujours de l'ATP par conversion du glucose en pyruvate et l'accepteur final d'électron au cours de la respiration, qui était l'oxygène, devient le nitrate ou le nitrite. Mais les activités des trois premières enzymes du cycle, la citrate synthase, l'aconitase et l'isocitrate déshydrogénase (essentiellement les deux premières), diminuent drastiquement (Nakano & Zuber, 1998). De plus, cette activité n'est plus stimulée par l'ajout de glucose. Cependant, aucune donnée n'est actuellement disponible quant à l'activité des autres enzymes du cycle, notamment en ce qui concerne la branche de conversion des acides dicarboxyliques, du succinate à l'oxaloacétate qui pourrait alors peut-être fonctionner dans le sens réducteur.

1.3.4.2. Chez *Escherichia coli*

Escherichia coli est capable de croître en milieu anaérobie en présence de glucose. Dans ces conditions, le cycle de Krebs fonctionne de manière non cyclique (Figure 10). L'oxaloacétate est utilisé pour produire du succinyl-CoA et du 2-oxoglutarate (Nimmo, 1987). Les étapes de la branche des acides dicarboxyliques fonctionnent dans le sens inverse de celui utilisé au cours de la respiration aérobie. L'activité de la succinate déshydrogénase est remplacée par celle de la fumarate réductase (Iverson *et al.*, 1999). Le fumarate sert alors d'accepteur final d'électrons pour la chaîne respiratoire anaérobie. De même, la conversion du pyruvate en acétyl-CoA n'est plus uniquement prise en charge par le complexe pyruvate déshydrogénase mais aussi par la pyruvate formate lyase, codée par le gène *pfl* (Rasmussen *et al.*, 1991; de Graef *et al.*, 1999; Alexeeva *et al.*, 2000). Cette dernière présente l'avantage de catalyser cette conversion sans réduire le NAD, qui reste disponible pour les réactions glycolytiques.

Bacillus subtilis ne possède pas de gènes homologues codant ces deux enzymes du métabolisme anaérobie (Sonenshein, 2002). Le complexe PDH est toujours utilisé pour produire l'acétyl-CoA mais sans doute moins efficacement (Nakano *et al.*, 1998; Ye *et al.*, 2000). Ainsi, pour ce qui est de la production d'énergie en milieu anaérobie, le cycle de Krebs ne semble pas jouer un rôle aussi important chez *Bacillus subtilis* que chez *Escherichia coli*, qui croît d'ailleurs bien plus efficacement que *Bacillus subtilis* dans ces conditions.

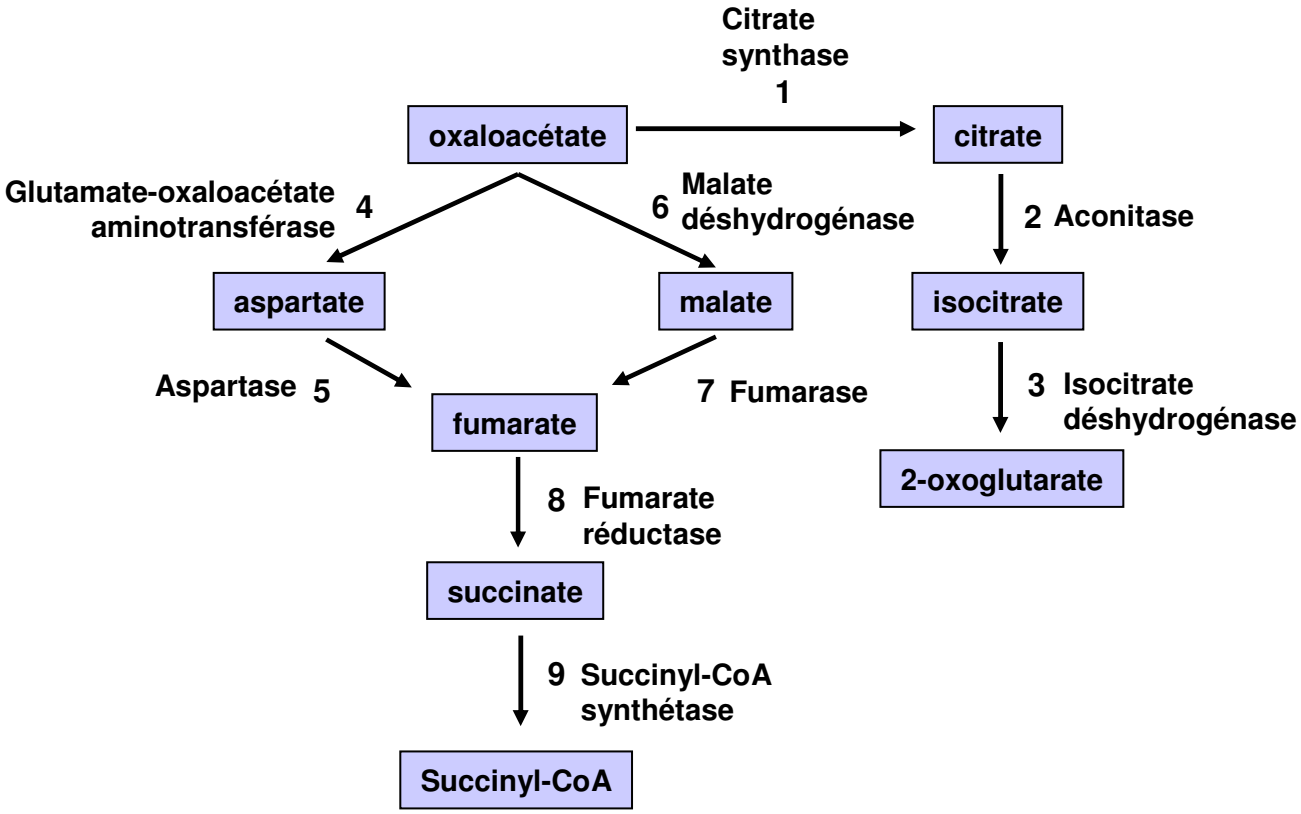


Figure 10. Fonctionnement du cycle de Krebs lors de la croissance en anaérobiose chez *Escherichia coli*. Les réactions fonctionnent selon l'ordre croissant des numéros.

D'après Nimmo, 1987.

1.3.5. Rôle du cycle de Krebs dans la sporulation

Comme décrit précédemment, les mutations des gènes codant les enzymes du cycle de Krebs conduisent à un blocage à différents stades dans la voie de sporulation chez *Bacillus subtilis* (Fortnagel & Freese, 1968a; Fortnagel & Freese, 1968b; Carls & Hanson, 1971; Yousten & Hanson, 1972; Freese & Marks, 1973; Ireton *et al.*, 1995; Jin *et al.*, 1996; Jin *et al.*, 1997). Il paraît évident que le rôle métabolique du cycle est impliqué dans ce processus mais d'autres facteurs interviennent.

La première explication met en jeu la nature des métabolites du cycle de Krebs. En effet, le citrate accumulé joue un rôle de chélateur de cations divalents, Fe^{2+} et Mn^{2+}, et réduit fortement le pH intracellulaire (Craig *et al.*, 1997; Matsuno *et al.*, 1999). L'hypothèse qui en découle explique ce défaut de différenciation par l'inhibition de l'activation du régulateur Spo0A (système de phosphorylation de Spo0A), qui contrôle la cascade d'activation des gènes de la sporulation.

La deuxième explication fait intervenir une fonction non enzymatique de l'aconitase, décrite dans la partie 1.3.2.2. Dans ce cas, il est proposé que cette protéine régule l'expression des gènes de la sporulation *via* son activité de liaison à l'ARNm (Alen & Sonenshein, 1999).

1.4. A la jonction entre la glycolyse/néoglucogenèse et le cycle de Krebs

1.4.1. Le complexe pyruvate déshydrogénase (EC 1.2.4.1) *pdhABCD*

La conversion irréversible du pyruvate en acétyl-CoA avec réduction du NAD est catalysée par un complexe comprenant quatre domaines (Figure 7). De nombreuses études sont disponibles (Fortnagel, 1993; Hederstedt, 1993). Il présente des similitudes avec le complexe 2-oxoglutarate déshydrogénase.

Chez *Bacillus subtilis*, et toutes les autres bactéries étudiées à l'exception d'*Escherichia coli*, le complexe PDH se compose des quatre protéines très bien conservées (Hodgson *et al.*, 1983; Fraenkel, 1987; Hemila *et al.*, 1990) : E1α (42 kDa) et E1β (36 kDa) qui ensemble forment l'activité déshydrogénase ; la protéine E2 (48 kDa) qui possède l'activité acyl transférase ; et la protéine E3 (50 kDa), qui possède l'activité lipoamide déshydrogénase. Un mutant *pdh* ne pousse pas bien sans ajout d'acétate qui fournit l'acétyl-CoA (Freese & Fortnagel, 1969; Gao *et al.*, 2002) (Figure 7). Le complexe pyruvate déshydrogénase est donc nécessaire pour subvenir aux besoins en acétyl-CoA. Par ailleurs, il semble que les sous-unités E1β et E2 de la pyruvate déshydrogénase soit aussi impliquées dans la régulation de la sporulation aux stages II et III. En effet, Gao et ses collaborateurs ont montré qu'une fonction non-enzymatique de ces sous unités était requise dans ce processus de différenciation (Gao *et al.*, 2002).

1.4.2. La pyruvate carboxylase *pycA* (EC 6.4.1.1)

Il s'agit d'une réaction anaplérotique du cycle de Krebs. En effet, la pyruvate carboxylase catalyse la carboxylation irréversible, à partir du CO_2, du pyruvate en oxaloacétate avec l'hydrolyse concomitante d'une molécule d'ATP (Figure 7). La structure et le mécanisme d'action de cette enzyme ont été caractérisés (Attwood, 1995). C'est l'enzyme majeure de fixation du CO_2 chez *Bacillus subtilis* (Fortnagel, 1993). Par ailleurs, elle est essentielle à l'initiation du cycle de Krebs puisqu'elle fournit l'oxaloacétate, substrat de la citrate synthase, à partir du pyruvate (Diesterhaft & Freese, 1973). En combinaison avec l'enzyme malique, elle permet la croissance sur malate comme seule source de carbone en l'absence d'une activité malate déshydrogénase, le malate étant d'abord converti en pyruvate puis en oxaloacétate, ce dernier revenant ensuite dans la néoglucogenèse par la phosphoénolpyruvate carboxykinase (Diesterhaft & Freese, 1973).

1.4.3. La phosphoénolpyruvate carboxykinase *pckA* (EC 4.1.1.49)

Cette enzyme réalise une étape clé permettant en grande partie au carbone d'entrer dans la néoglucogenèse (Diesterhaft & Freese, 1973; Velayudhan & Kelly, 2002). En effet, elle catalyse la conversion irréversible d'une molécule d'oxaloacétate en PEP en consommant une molécule d'ATP. Chez *Bacillus subtilis*, un mutant *pckA* présente une croissance très ralentie en présence de source de carbone néoglucogénique comme les composés du cycle de Krebs mais pousse comme la souche sauvage en présence de glucose (Doan et Aymerich, non publié).

1.4.4. L'enzyme malique (EC 1.1.1.38-40)

L'enzyme malique NAD(P)-dépendante catalyse l'interconversion du malate et du pyruvate. En combinaison avec la pyruvate carboxylase, elle permet de produire de l'oxaloacétate à partir du malate dans un mutant *mdh* pour faire tourner le cycle de Krebs et remonter la glycolyse (Figure 7). De même, elle permet aussi de régénérer le pool d'acétyl-CoA en combinaison avec l'activité pyruvate déshydrogénase. Ainsi, en l'absence de la malate déshydrogénase, elle permet au flux de carbone d'entrer dans la néoglucogenèse.

Chez *Bacillus subtilis*, quatre gènes paralogues peuvent coder des enzymes maliques : *ywkA*, *malS*, *ytsJ* et *mleA*. Aussi, l'étude transcriptionnelle et physiologique d'*ywkA* constitue la troisième partie de cette thèse. Ce paragraphe sera donc développé ultérieurement.

2. Les différents niveaux de régulation du métabolisme central du carbone

2.1. Un prélude au transport

L'habitat naturel le plus courant de *Bacillus subtilis* est le sol. Ce sol colonisé par les racines des plantes est un écosystème complexe et diversifié, d'une part par le nombre d'espèces qui y vivent et d'autre part par le nombre de molécules carbonées disponibles s'y trouvant : les plantes synthétisent notamment de grandes quantités de polysaccharides, polymères plus ou moins complexes de saccharides et de protéines qui sont excrétés sous forme d'exsudats ; la décomposition de la matière principalement dans le sol génère aussi énormément de molécules carbonées réutilisables. Ces polysaccharides et ces polypeptides constituent une source importante de carbone, mais correspondent souvent à des molécules trop grosses pour être « ingérées » directement par la bactérie. Par contre, après dégradation dans le milieu par les enzymes extracellulaires qui sont produites par la multitude de bactéries et champignons du sol, les petits résidus libérés seront transportés et métabolisés. Ainsi, *Bacillus subtilis* se trouve en présence de composés carbonés « végétaux » qu'elle sait dégrader pour certains et de composés déjà « traités », encore décomposables ou directement transportables et métabolisables.

Bacillus subtilis sécrète de nombreuses enzymes, des glucanases, des protéases et des nucléases pour dépolymériser ces complexes carbonés (Priest, 1977; Stulke & Hillen, 2000), ce qui lui vaut en partie d'être une bactérie d'intérêt industriel. S'agissant de la dégradation des polysaccharides, on peut notamment citer comme exemple le rôle de l'α-amylase, codée par le gène *amyE*. Celle-ci catalyse la dégradation des polyglycosyls, comme l'amidon, en molécules de glucose, maltose, maltotriose et maltodextrines (Priest, 1977; Vihinen & Mantsala, 1990). Par ailleurs, de façon remarquable, on observe que les gènes qui codent les enzymes de dégradation sont souvent génétiquement liés aux gènes qui codent les enzymes nécessaires au métabolisme ultérieure des produits correspondants de dégradation.

2.2. Le transport : une étape de tri préférentiel des sources de carbone et de transduction du signal

Toute cellule est constituée d'une enveloppe. Or, si elles peuvent aisément disposer de petites molécules d'eau ou de gaz comme l'oxygène, qui traversent cette barrière complexe par simple diffusion, les bactéries ont dû développer, au cours de l'évolution, des systèmes d'absorption de molécules plus grosses et de nature très diverse leur permettant de coloniser des milieux dont la composition en nutriments est très variable. En effet, les bactéries constituent les espèces vivantes ayant colonisé le plus de niches écologiques. Et en considérant toutes les espèces bactériennes connues, elles sont capables de croître sur des sources extrêmement diverses d'hydrates de carbone. Le rôle des systèmes de transport membranaire est donc de permettre l'internalisation des molécules de façon efficace et régulée. Pour cela, chaque type de molécules est pris en charge par un transporteur/perméase dont l'affinité, plus ou moins forte suivant le subtrat, permet à la bactérie d'utiliser des molécules souvent très diluées. En contrepartie de cette haute affinité, ces perméases sont dites substrat(s)-spécifiques.

2.2.1. Les systèmes bactériens de transport membranaire des sources de carbone

Les transporteurs de glucides sont associés à la membrane cytoplasmique. Ils sont typiquement composés d'une ou de plusieurs protéines qui forment une structure transmembranaire en alpha-hélices. Les mécanismes moléculaires qui régissent le transport sont encore à préciser (Saier *et al.*, 2002). Mais certaines données importantes de structure tridimensionnelle reconstituant l'intégralité de complexe protéique de transport fournissent déjà des éléments de réponse. C'est le cas de la structure de l'ATPase de type P calcium dépendante du réticulum sarcoplasmique du muscle squelettique (Toyoshima *et al.*, 2000).

Quatre types de système de transport ont été décrits chez les bactéries (Figure 11), en majorité chez *Escherichia coli*, et sont classés selon deux critères : d'une part, le mode de translocation, et d'autre part, la source d'énergie nécessaire à ce processus (Dills *et al.*, 1980; Saier *et al.*, 1985; Saier, 2000a; Saier, 2000b; Saier *et al.*, 2002). On distingue donc le transport par diffusion facilitée, le cotransport cation ou soluté-dépendant, le transport ATP-dépendant et enfin le système des phosphotransférases dépendantes du PEP.

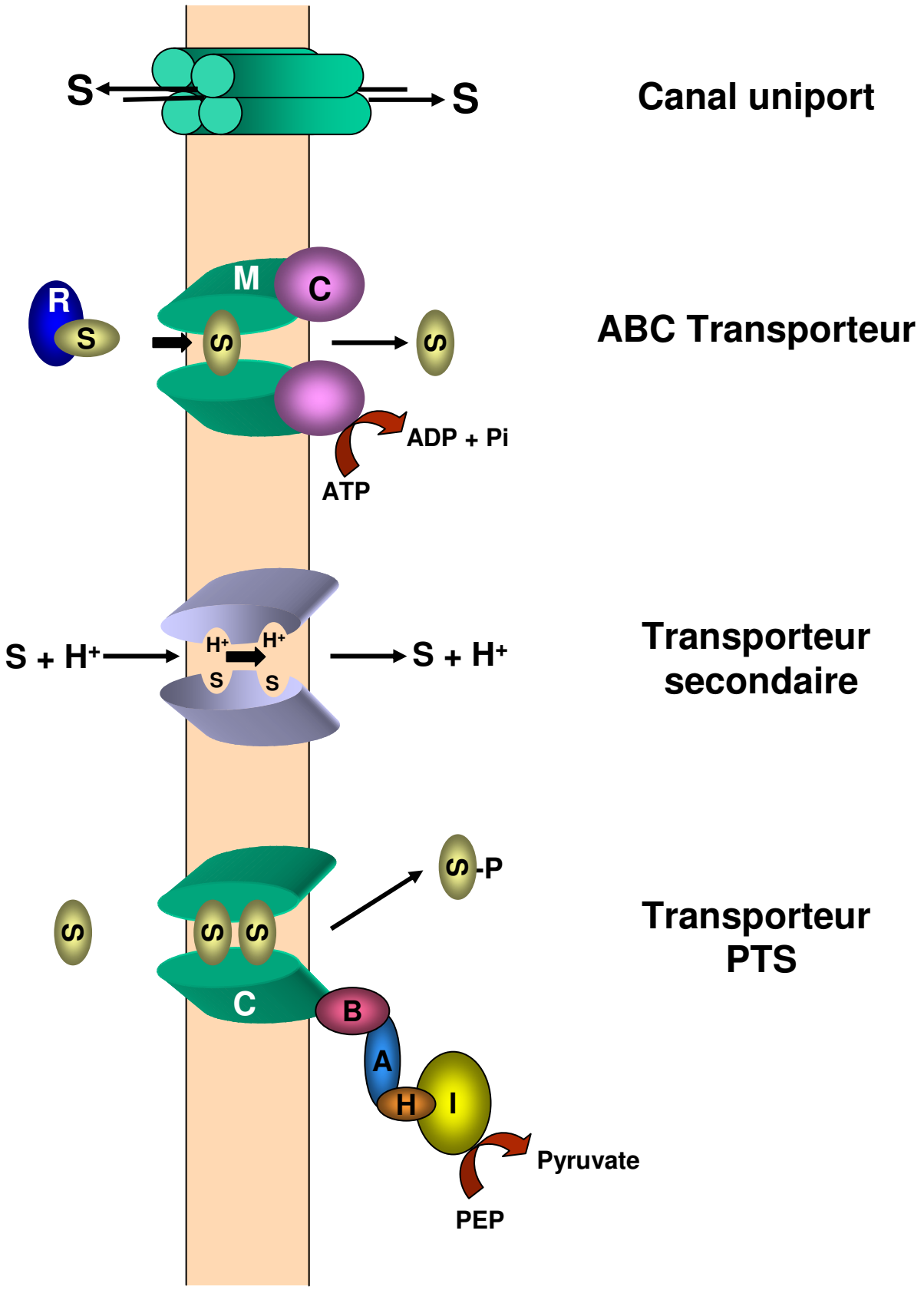


Figure 11. Les quatre types de système de transport décrits chez les bactéries.

Légendes : S (substrat) ; R (protéine réceptrice) ; M (domaine membranaire perméase) ; C (domaine cytoplasmique ATPase) ; H^+ (proton) ; A (domaine cytoplasmique EIIA) ; B (domaine cytoplasmique EIIB) ; C (domaine membranaire EIIC) ; H (HPr) ; I (domaine cytoplasmique EI).

2.2.1.1. Transport par diffusion facilitée

Le premier type de transport connu consiste en la diffusion facilité de solutés à travers un complexe protéique formant un canal, d'où l'expression de transport unipore (Figure 11). Ce système ne diffère du transport par diffusion passive que par la nature enzymatique de son mécanisme. Il est rare chez les Procaryotes (1,1 % des transporteurs chez *Bacillus subtilis* et 2,3 % chez *Escherichia coli*) mais très largement répandu chez les Eucaryotes (Paulsen *et al.*, 2000). L'exemple bactérien le mieux étudié -l'un des seuls- est celui du facilitateur d'entrée du glycérol GlpF, découvert chez *Eschericha coli* (Heller *et al.*, 1980). Le glycérol va pénétrer dans la cellule suivant son gradient de concentration. Tant que la concentration de glycérol dans le milieu extérieur est supérieure à la concentration cytoplasmique, le glycérol se lie au transporteur pour franchir la membrane, sans nécessiter de consommation d'énergie. Récemment, la structure oligomérique de GlpF d'*Escherichia coli* a été résolue (Fu *et al.*, 2000) et les détails moléculaires de la spécificité de cette protéine facilitatrice de la famille des aquaporines pour son substrat carboné ont été élucidés (Jensen *et al.*, 2001; Nollert *et al.*, 2001). Ce système de transport est communément utilisé pour l'entrée des polyols à chaîne linéaire de 3 à 5 carbones comme le glycérol, l'érythritol, le ribitol, le xylitol et l'arabitol.

Les autres systèmes de transport permettent l'entrée d'hydrates de carbone contre leur gradient de concentration mais nécessitent une source d'énergie. En outre, les perméases impliquées ont une très grande affinité pour leur substrat(s). On parle de transport actif de molécules.

2.2.1.2. Transport actif primaire ATP-dépendant

Dans ce deuxième mode, le transport actif de soluté est couplé à une source d'énergie chimique, électrique ou solaire. L'ATP est le principal fournisseur d'énergie utilisé mais il existe aussi d'autres molécules dont la décarboxylation libère une énergie suffisante pour supporter le transport actif de soluté. La superfamille des ABC transporteurs (ATP binding cassette transporteur) a été définie suite à la découverte de l'homologie entre un transporteur dépendant de l'ATP et une pompe assurant l'efflux de multiples drogues (Chen *et al.*, 1986). Classiquement, un ABC transporteur compte plusieurs composants (Figure 11). Le substrat interagit d'abord avec une protéine réceptrice soluble du périplasme qui présente une haute affinité pour ce substrat. Puis, l'interaction de ces deux composants avec le domaine

extracytoplasmique du transporteur va stimuler l'activité ATPase du domaine cytoplasmique fournissant l'énergie nécessaire pour initier le transport à l'intérieur du cytoplasme.

Les Procaryotes possèdent la plus grande diversité de sous-famille d'ABC transporteurs (Paulsen *et al.*, 2000). Ces transporteurs sont aussi impliqués dans l'entrée ou la sortie d'ions, de vitamines, d'acides aminés, de peptides. Chez les bactéries, peu de sucres sont transportés par un tel système (Saier, 2000b; Saier *et al.*, 2002). Il semble que ce système de transport de glucides ne soit relativement développé que chez les bactéries à Gram négatif (Paulsen *et al.*, 2000). L'exemple le plus connu concerne le transport du maltose *via* le complexe MalFGK (MalF, MalG et deux protéines MalK) et la protéine MBP (Maltose Binding Protein) chez *Escherichia coli*, pour lequel les mécanismes de couplage du transport et de l'hydrolyse de l'ATP ont été éclaircis (Davidson & Nikaido, 1990; Dean *et al.*, 1990a; Davidson *et al.*, 1992; Davidson *et al.*, 1996; Chen *et al.*, 2001). Les principaux sucres transportés via ce système sont le maltose, le ribose, le raffinose et le xylose. Certains acides aminés (des oligopeptides, les acides aminés polaires et la proline) sont aussi transportés par des systèmes de type ABC transporteurs (Saier *et al.*, 2002).

2.2.1.3. Transport actif secondaire

Le troisième mode de transport est celui mis en place par les transporteurs actifs secondaires (Figure 11). Ces transporteurs utilisent l'énergie chémostatique fournie par un gradient électrochimique d'ions ou de solutés créé de part et d'autre de la membrane. Le plus souvent, l'énergie provient de la force protonmotrice (gradient de H^+) ou du gradient d'ions sodium (Na^+) allant de l'extérieur vers l'intérieur de la cellule. Cette force protonmotrice peut être générée, soit par la voie oxydative lors du transport d'électrons au cours de la respiration, soit par la voie fermentaire. Cependant, les ions potassium (K^+), lithium (Li^{2+}) ou le phosphate peuvent aussi générer des gradients électrochimiques suffisants.

Trois modes de transport secondaire sont généralement décrits chez les bactéries (Figure 12) qui font intervenir deux voire plus de deux molécules différentes :

- le symport ou cotransport soluté/cation : le gradient ionique permet l'entrée de soluté dans le cytoplasme dans le même sens,
- l'antiport ou contre-transport soluté/cation : le gradient ionique fournit l'énergie nécessaire à l'expulsion de soluté hors du cytoplasme,

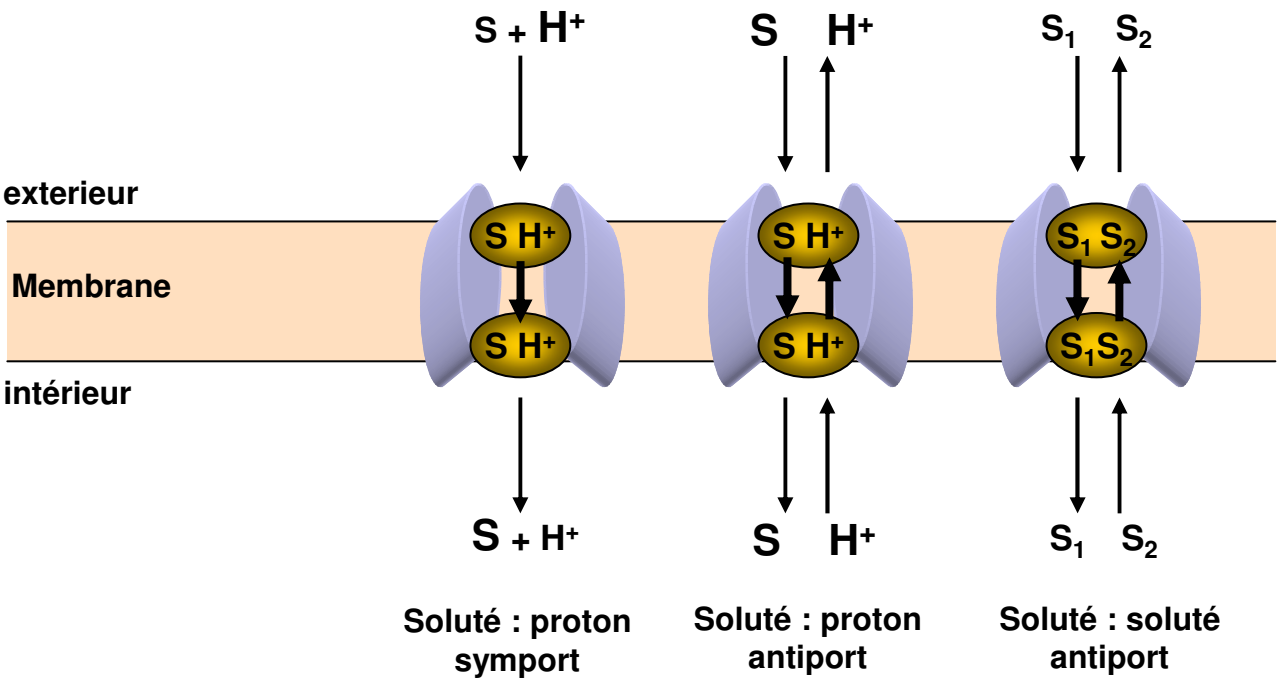


Figure 12. Les mécansimes de transport secondaires décrits chez les bactéries.

- l'antiport ou contre-transport soluté/soluté : l'entrée ou l'expulsion d'un soluté suivant son gradient de concentration permet le transport en sens inverse d'un autre soluté.

Ces systèmes de transport sont très largement répandus chez les bactéries en terme de type de molécules transportées, ions, sucres, acides carboxyliques, acides aminés et drogues en majorité (respectivement 63 % et 64 % des transporteurs membranaires chez *Bacillus subtilis* et *Escherichia coli*) (Paulsen *et al.*, 2000).

2.2.1.4. Translocation PEP-dépendante / système PTS

Ce quatrième système de transport a été découvert chez *Escherichia coli* dans les années 1960 (Kundig *et al.*, 1966). Depuis, il a été mis en évidence chez de très nombreuses bactéries, aussi bien parmi les bactéries à Gram négatif que parmi les bactéries à Gram positif (Saier *et al.*, 1988; Lengeler *et al.*, 1990; Postma *et al.*, 1993; Lengeler *et al.*, 1994; Saier & Reizer, 1994; Lengeler & Jahreis, 1996; Deutscher *et al.*, 2002). Cependant, les Archées, *Lactococcus brevis* et *Lactococcus buchneri* ne possèdent pas de système PTS de transport (Reizer *et al.*, 1988; Paulsen *et al.*, 2000). Des études révèlent que, chez ces bactéries, le PTS est le système prédominant, dédié au transport des sucres, principalement des hexoses et des hexitols (Deutscher *et al.*, 2002; Saier *et al.*, 2002). Le groupe des transporteurs PTS ne représente que 6 % et 7 % de l'ensemble des transporteurs membranaires chez *Bacillus subtilis* et *Escherichia coli*, respectivement (Paulsen *et al.*, 2000). Par ailleurs, il semble que ce soit un système spécifique aux bactéries puisqu'il n'y a aucun système homologue décrit ni chez les Archées ni chez les Eucaryotes (Saier, 2000b).

Le système PTS a pour caractéristique spécifique la simultanéité du transport et de la phosphorylation du substrat carboné selon une chaîne réactionnelle à partir du PEP, donneur initial du groupement phosphate et d'énergie, d'où le terme de phosphoénolpyruvate:sucre phosphotransférase système (PTS) (Figure 11). Ce mécanisme se distingue donc des trois précédents où le substrat n'est pas modifié lors de son entrée dans le cytoplasme. C'est un système régulé dont les caractéristiques seront développées dans la partie suivante.

2.2.1.5. Transport des sources de carbone chez *Bacillus subtilis*

Bacillus subtilis transporte et utilise une multitude de sources de carbone qui sont naturellement présentes dans le milieu extérieur ou qui sont le résultat de la dégradation enzymatique des polysaccharides et autres molécules carbonées (Steinmetz, 1993; Saier *et al.*, 2002). La plupart des composés peuvent être regroupés sous le terme de glucides ou hydrates de carbone. Les monomères d'hydrate de carbone sont définis par la formule générale $(CH_2O)n$. Ils ont la structure chimique d'aldéhydes ou de cétones avec de multiples groupes hydroxyls (aldoses et cétoses). Les monomères et dimères de glucides sont communément appelés sucres. La variété de groupements réactifs et de structures conformationnelles possibles qui caractérisent les glucides fait qu'il existe une multiplicité de molécules différentes. Ainsi, dans la nature, il existe des monomères « purs », des polymères, mais aussi des dérivés carboxyliques et aminés (lorsque les groupements carboxyl et amine sont portés par le même atome de carbone, il s'agit d'un acide aminé), des polyalcools, etc...

Un grand nombre de ces sources de carbone sont transportées et catabolisées : les saccharides (mono, di et oligosaccharides), les sucres aminés et leurs dérivés N-acétylés (par exemple le N-acétyl-glucosamine), les alcools, les opines, les acides carboxyliques et les acides aminés (Deutscher *et al.*, 2002). La Table 1 récapitule les données existantes concernant les systèmes de transport de glucides chez *Bacillus subtilis* (expérimentalement démontrés ou sur la base des homologies de séquence ; l'acide aminé glutamate a été rajouté) (Deutscher *et al.*, 2002; Saier *et al.*, 2002).

Parmi ces sources de carbone, les sucres simples de type hexoses comme le glucose ou le fructose représentent les sources de carbone privilégiées de *Bacillus subtilis* (Steinmetz, 1993). Il est important de remarquer qu'un sucre peut être transporté par plusieurs transporteurs impliquant un ou plusieurs mécanismes différents, mais le plus souvent avec une efficacité très variable. Le glucose, par exemple, est majoritairement transporté et phosphorylé par le système PTS glucose-spécifique codé par le gène *ptsG* (Table 1) et un mutant de ce gène provoque un défaut de croissance très prononcé en présence de glucose comme source de carbone (Gonzy-Treboul *et al.*, 1991; Bachem *et al.*, 1997). Mais il existe deux autres systèmes de transport, codés par les gènes *glcP* et *glcU* qui complémentent très partiellement la mutation de *ptsG*. Le premier est un symporteur glucose:proton également capable de transporter le mannose, le 2-désoxyglucose et le méthyl α-glucoside (Paulsen *et al.*, 1998). Son expression nécessite l'intégrité du système de signalisation du PTS (cf partie 2.2.2). Le second est une protéine orthologue au transporteur GlcU de *Staphylococcus xylosus*

Table 1. Systèmes de transport des glucides chez *Bacillus subtilis*.

Transporteur	Spécificité	Organistaion génétique	Régulation par	Références
Systèmes PTS				
BglP	Aryl-β-glucosides	*bglPH*	LicT, CcpA	Krüger et al, 1995 ; Le Coq et al,1995
FruA	Fructose	*fruRBA*	RfuF, CcpA ?[a]	Martin-Verstraete et Rapoport, non publié
GlvC (MalP)	Maltose	*glvAyfiAglvC*	CcpA ?, ?	Reizer et al, 1999
LevDEFG	Fructose, Mannose	*levDEFGsacC*	LevR, CcpA	Martin-Verstraete et al, 1990 ; Martin-Verstraete et al, 1999
LicABC	Oligo-β-glucosides	*licRBCAH*	LicR, CcpA	Tobisch et al, 1997 ; Tobisch et al, 1999
Mtla, MtlF ?	Mannitol	*mtlAF ?D*	YdaA, CcpA	Deutsher et al, 1999, Kunst et al, 1997 ; Reizer et al, 1992
PtsG	Glucose	*ptsG(HI)*	GlcT	Gonzy-Tréboul, 1991 ; Stülke et al, 1997
SacP	Sucrose	*SacAP ywdA*	SacT, CcpA ?	Arnaud et al, 1992 ; Fouet et al, 1987
SacX	Sucrose	*sacXY*	SacY	Tortosa et Le Coq, 1995 ; Zukowski et al,1990
TreP	Tréhalose	*trePAR*	TreR, CcpA ?	Miwa et al, 2000 ; Schöck et Dahl, 1998
Ybbf	β-glucosides ?	*ybbKJIHFEDCB*	CcpA ?, ?	Kunst et al, 1997 ; Miwa et al, 2000
YbfS (GamP)	Glucosamine	*ybfTS*	YbgA ?	Reizer et al, 1999
YdhNMO	Oligo-β-mannosides ?	*ydhMNOPQRST*	YdhQ ?, CcpA ?	Deutscher et al, non publié ; Kunst et al, 1997 ; Stülke et al, 1998
YflF (NagA)	N-acetylglucosamine	*yflFGH*	?	Reizer et al, 1999
YjdD	Mannose ?	*yjdDEF*	YjdC ?, CcpA ?	Kunst et al, 1997 ; Miwa et al, 2000 ; Stülke et al, 1998
YwbA	?	*ywbAB*	CcpA ?	Deutscher et al, non publié ; Kunst et al, 1997
Facilitateurs				
GlpF	Glycerol	*glpFK (glpD)*	GlpP, CcpA HPr, EI	Rutberg, 1997 ; Darbon et al, 2001
ABC transporteurs				
AmyCD	Melibiose ?	*msmREamyDCmelA*	MsmR, CcpA ?	Kunst et al, 1997
AraPQ	Arabinose	*araABDLMNPQ*	AraR, CcpA	Miwa et al, 2000 ; Mota et al, 1999 ; Sa-Nogueira et al, 1997a
LplBCA	?	*lplABCFyetF*	?	Kunst, 1997
MsmX	?	*yxkFmsmX*	YkxF ?, CcpA	Miwa et al, 2000 ; Yoshida et al, 1996
RbsACD	Ribose	*rbsRKDACBywsB*	RbsR, CcpA ? AbrB	Miwa et al, 2000 ; Woodson et al, 1994 Strauch, 1995
YesPQ	Galactosides ?	*yesJ-S yetA*	YesM/YesN ? YesS ?, CcpA ?	Kunst et al, 1997 ; Miwa et al, 2000
YtePQ	Oligosaccharides ?	*ytePQRSTUV*	YtdP ?	Rong et al, 1991
YtrCD	Acétoine	*ytrABCDEF*	YtrA ?	Yoshida et al, 2000
YufOPQ	Galactose ?	*yufOPQ*	YufL/YufM ?	Oudega et al, 1997
YurMN	Mannopine ?	*yurRQPONML*	YurK ?	Kunst et al, 1997
YurJNM	?	*yurJ*	?	Kunst et al, 1997
YvdHI	Maltodextrine	*yvdEFGHIJKLM*	YvdE ?, CcpA	Miwa et al, 2000 ; Schönert et al, 1998
YvfLM	Maltose, lactose ?	*yvfKLMlacAyvfO*	LacR, CcpA	Daniel et al,1997 ; Deutscher et al, non publié ; Kunst et al, 1997
Transporteurs secondaires				
AraE	L-arabinose	*araE*	AraR	Mota et al, 1999 ; Sa-Nogueira et al, 1997b
CimH (YxkJ)	Malate, Citrate	*yxkJI ?*		Krom et al, 2001 ; Asai et al, 2000
CitH	Citrate (Ca^{2+})	*citH*		Krom et al, 2000
CitM	Citrate (Mg^{2+})	*citMyflNM*	CitST, CcpA	Krom et a, 2000 ; Yamamoto et al, 2000
CsbC (YxcC)	GalP/xylE family	*csbC*	?	Akbar et al, 1999
CsbX	α-oxoglutarate	*yrbEcsbXbofC*	CcpA ?	Gomez et al, 1997 ; Miwa et al, 2000
DctP	Acides dicarboxyliques	*dctP*	DctSR	Asai et al, 2000
GlcU (YcxE)	Glucose	*glcUgdh*	?	Fiegler et al,1999
GlpT	Glycérol-3P	*glpTQ*	GlpP, CcpA ?	Deutscher et al, non publié ; Nilsson et al, 1994 ; Rutberg, 1997

GltP	Glutamate	*gltP*		Saier et al, 2002
GltT	Glutamate	*gltT*		Saier et al, 2002
GntP	Gluconate	*gntRKPZ*	GntR, CcpA	Deutsher et al, 1999 ; Fujita et al, 1986
GlcP (YhjI)	Glucose, mannose	*glcP*	?	Paulsen et al,1998
IolF	Inisitol	*iolABCDEFGHIJ*	IolR, CcpA	Miwa et al, 2000 ; Yoshida et al, 1997
KdgT	Hexuronate	*kdgRKAT*	KdgR, CcpA	Miwa et al, 2000 ; Pujic et al, 1998 ; Rivolta et al, 1998
ExuT (YjmG)	Aldohexuronate	*uxaCyjmBCDuxuA yjmFexuTRuxaBA*	ExuR, CcpA	Mekjian et al,1999 ; Rivolta et al, 1998
LctP	Lactate	*ldhlctP*		Cruz Ramos et al, 2000
MaeN	Malate	*maeN ? mrpABCDEFG*	MaeKR	Doan et al, 2003 ; Ito et al, 2000 ; Wei et al, 2000
YcbE	Glucarate	*ycbABCDEFGHJ*	YcbG ? YcbA/YcbB ?	Ogawa et al, 1995
YceI	?	*ycel*	?	Kunst et al, 1997
YddS	Famille GalP/XylE	*yddQRST*	?	Kunst et al, 1997
YdeG	Famille GalP/XylE	*ydeGHI*	CcpA ?	Kunst et al, 1997
YdjD (GutP)	Sorbitol	*gutB, ydjDE*	GutR, CcpA	Deutsher et al, 199 ; Yoshida et al, 1996
YdjK	Famille GalP/XylE	*ydjK*	CcpA ?, ?	Kunst et al, 1997 ; Miwa et al, 2000
YfiG	Famille GalP/XylE	*yfiFGHI*	YfiF ?, CcpA ?	Kunst et al, 1997 ; Miwa et al, 2000
YflS	Malate	*yflS*		Wei et al, 2000
YhcL	Glutamate ?	*yhcL*		Saier et al, 2002
YjmB	β-glucoronide	*UXACyjmBCDuxuA yjmFexuTRuxaBA*	ExuR, CcpA	Mekjian et al,1999 ; Rivolta et al, 1998 Miwa et al, 2000
YnaJ (XynP)	β-Xylosides	*ynaJxynB*	XylR,CcpA	Galinier et al, 1999 ; Hastrup, 1988
YncC	Famille GalP/XylE ?	*yncC*	CcpA ?, ?	Kunst et al, 1997 ; Miwa et al, 2000
YojA (GntT?)	Gluconate ?	*yojABCDEFG*	CcpA ?	Miwa et al, 2000 ; Orco et al, 1997
YqkI	Malate/lactate	*yqkIJ*	?	Wei et al, 2000
YwbF	Maltose ?	*ywbHGFE*	YwbI ?	Liong et al,1994
YwtG	Famille GalP/XylE ?	*ywtG*	YwtF ?	Kunst et al, 1997
YxjC	β-Hydroxybutyrate ?	*yxjCDEF*	CcpA ?	Kunst et al, 1997 ; Miwa et al, 2000
Yyaj	?	*yyaL*	?	Kunst et al, 1997
YybO	Glucarate	*yybP*	?	Kunst et al, 1997

Table 1. Systèmes de transport des glucides chez *Bacillus subtilis*.
[a] Le point d'interrogation indique qu'aucune preuve expérimentale n'est disponible quant à l'information suggérée. Les couleurs indiquent la voie d'intégration de la source de carbone transportée (orange : glycolyse ; jaune clair : voie des pentoses phosphates ; gris clair : cycle de Krebs).
D'après Deutscher *et al*., 2002 ; Saier *et al*., 2002.

(Fiegler *et al.*, 1999). Dans les deux cas, le glucose qui entre n'est pas modifié. Il sera phosphorylé par la glucokinase GlcK (cf partie 1.1.2). Dans le même ordre d'idée, le fructose peut être transporté *via* le PTS par deux systèmes, le PTS Lev et le PTS FruA, le premier étant très peu efficace et transportant aussi le mannose et le deuxième étant le transporteur principal spécifique et efficace (Martin-Verstraete *et al.*, 1990; Reizer *et al.*, 1999).

Le transport des acides carboxyliques comme le citrate (CitM, CitH), le succinate et le fumarate (DctP) et le malate (MaeN, YxkJ), qui sont tous des intermédiaires du cycle de Krebs et donc des substrats carbonés de type néoglucogénique, est de type transport secondaire (Saier *et al.*, 2002). Par ailleurs, chez *Bacillus subtilis* mais aussi chez d'autres bactéries, on peut remarquer que l'expression d'une partie des gènes codant ces transporteurs est activée par des systèmes spécifiques senseurs/régulateurs à deux composants en réponse à la présence de leurs substrats dans le milieu (Janausch *et al.*, 2002) (cf article n°2).

La plupart des acides aminés est aussi perméée par des transporteurs secondaires plus ou moins spécifiques et intègrent le métabolisme du carbone *via* le cycle de Krebs ou au niveau du pyruvate, l'intermédiaire clef reliant la glycolyse au cycle de Krebs (Saier *et al.*, 2002). Par ailleurs, les acides aminés constituent une source d'azote privilégiée de *Bacillus subtilis* (Fisher, 1993). Aussi, la régulation de l'expression des gènes codant des transporteurs caractérisés d'acides aminés chez *Bacillus subtilis* est intimement liée à la régulation générale du métabolisme azoté.

2.2.1.6. Considérations thermodynamiques et coût énergétique du transport

Au cours du transport d'un sucre PTS, une molécule de PEP est transformée en pyruvate. L'énergie libre du groupement phosphate du PEP est de –61,5 kJ/mol, c'est-à-dire environ 5 fois plus que l'énergie libre d'un groupement phosphoester (Postma *et al.*, 1993). Cette simple considération laisse penser qu'un transport par PTS est plus consommateur d'énergie qu'un transport de type actif primaire ou secondaire, bien que dans ce dernier cas, il faille prendre la peine de se demander d'où provient la force ionique utilisée. Mais la caractéristique à l'avantage du système PTS est la phosphorylation concomitante au transport à partir d'une molécule de PEP. Par conséquent, on peut considérer que la conversion de l'énergie d'une molécule de PEP en ATP est épargnée, le PEP faisant alors office de source d'énergie. Le coût du transport et de la phosphorylation par le PTS revient donc à l'utilisation de l'énergie d'une seule molécule d'ATP.

S'agissant des sucres transportés par les systèmes non-PTS, plus d'une molécule d'ATP est nécessaire pour le transport et la phosphorylation ATP-dépendante ultérieure d'une molécule de monosaccharide (Postma *et al.*, 1993). En effet, une molécule d'ATP est consommé pour le transport d'une molécule de sucre par un ABC transporteur (Muir *et al.*, 1985) ou pour le transport d'au plus trois molécules de sucre par un symporteur à proton (Kashket, 1983). Il faudra une molécule d'ATP supplémentaire pour phosphoryler puis métaboliser ce sucre. Par conséquent, Le coût du transport de type ABC transporteur et de type actif secondaire est respectivement de 2 et 1,3 (au mieux) molécules d'ATP par sucre transporté et métabolisé, ce qui est supérieur au coût énergétique réclamé par le système PTS.

C'est probablement la raison pour laquelle le système PTS est largement et préférentiellement utilisé par les bactéries à métabolisme anaérobie strict ou facultatif. En effet, ces bactéries synthétisent l'ATP par phosphorylation du substrat de façon peu efficace et doivent donc utiliser l'ATP généré sans gaspillage (Postma *et al.*, 1993).

2.2.2. Le PTS : un système prédominant de contrôle du transport des glucides doublé d'un système de transduction du signal chez *Bacillus subtilis*

2.2.2.1. Description générale

Le PTS est constitué de plusieurs composantes protéiques. On peut faire la distinction entre les composantes dites générales du PTS qui participent au transport de tous les sucres PTS et les composantes spécifiques d'un ou de plusieurs sucres (Figure 13).

Les composantes générales du PTS sont l'enzyme I (EI) et la protéine HPr. Elles sont exprimées en quantité plus ou moins importantes selon la source de carbone présente dans le milieu de croissance mais au moins à un niveau basal *via* un promoteur constitutif (cf partie 2.3.3.1.2). Le PEP fournit l'énergie et le groupement phosphate qui va permettre à l'enzyme I de s'autophosphoryler. L'enzyme I va alors catalyser le transfert de son groupement phosphate sur la position Histidine 15 de la protéine HPr qui à son tour va le donner aux enzymes II membranaires qui peuvent alors jouer le rôle de phosphotranslocation spécifique des sucres. L'expression des enzymes II est très généralement induite par la présence du ou des sucres qu'elles transportent. Chaque complexe enzyme II est formé de trois domaines appelés IIA, IIB, IIC et parfois d'un quatrième domaine IID qui sont codés par un ou plusieurs gènes distincts (Figure 13).

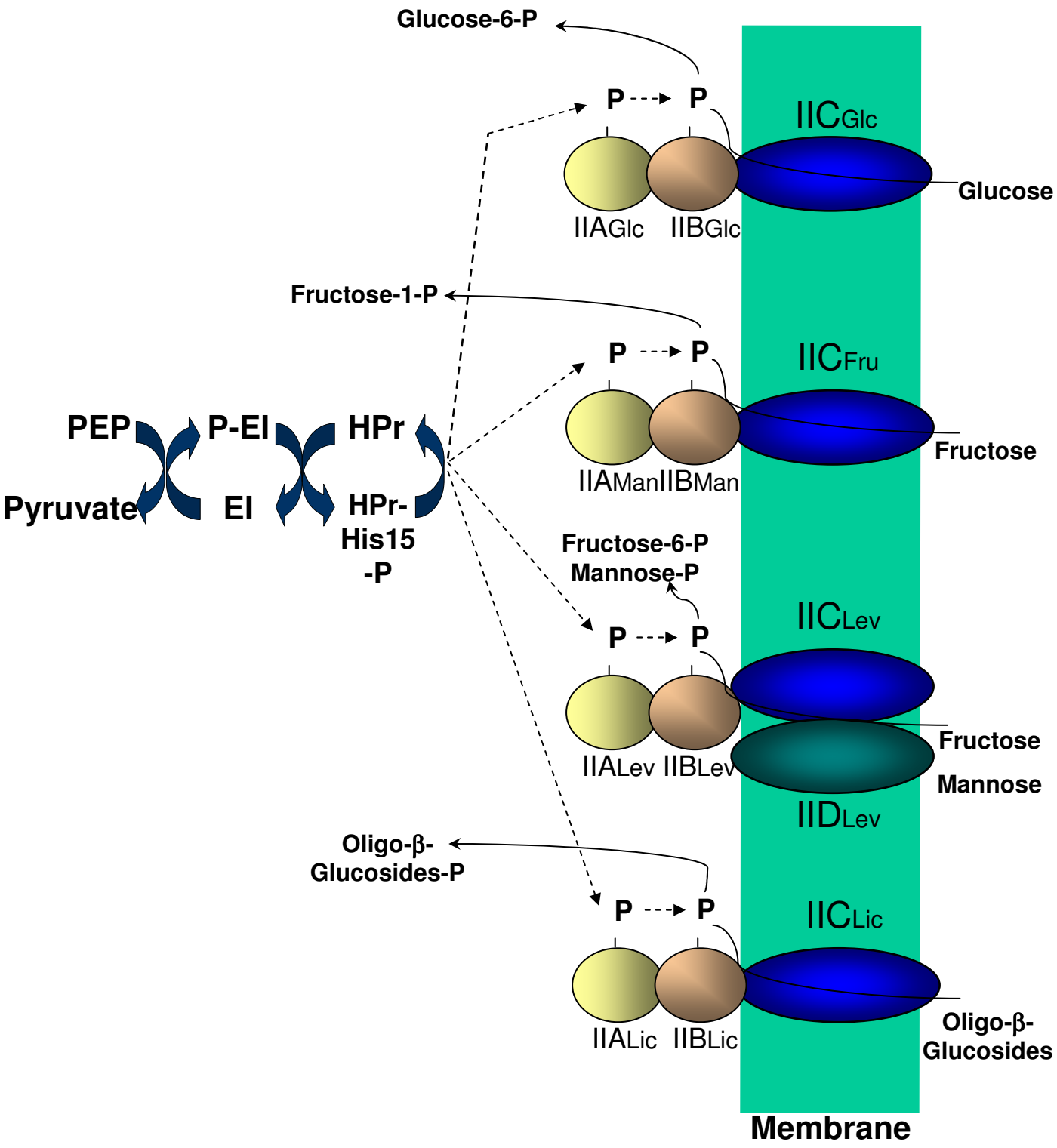


Figure 13. Organisation générale du PTS chez *Bacillus subtilis*.

Les quatre classes de système de transport PTS sucre-spécifiques présentes chez *B. subtilis* sont représentées par quatre exemples. L'enzyme EI catalyse la phosphorylation PEP-dépendante d'HPr sur l'Histidine 15. HPr-His15-P va transférer ce groupement phosphate au domaine EIIA cytoplasmique de l'un des complexes EII spécifiques d'un ou de plusieurs sucres. Le domaine EIIB associé, lui aussi cytoplasmique, va transférer ce groupement de l'enzyme EIIA au sucre correspondant transporté par le domaine EIIC membranaire. Dans le cas des transporteurs de la classe mannose, le complexe de translocation est constitué des domaines EIIC et EIID. Selon les cas, les domaines A,B,C et D peuvent faire partie d'un même polypeptide ou correspondre à plusieurs polipeptides distincts.

D'après Deutscher *et al.*, 2002.

La chaîne réactionnelle aboutissant à la phosphorylation du sucre transporté s'écrit de façon générale :

PEP + EI → pyruvate +P-EI

P-EI + HPr ↔ EI + HPr-His15-P

HPr-His15-P + EIIA ↔ HPr + P-EIIA

P-EIIA + EIIB ↔ EIIA + P-EIIB

P-EIIB + sucre(ext) ↔ EIIB + **P-sucre (int)**

où le PEP est à la fois le donneur initial de groupement phosphate et la source d'énergie pour le transport. Toutes ces réactions de phosphorylation de protéines sont réversibles. C'est donc la concentration en PEP à l'intérieur du cytoplasme, elle même dépendante de l'activité glycolytique et la présence du sucre inducteur ou répresseur dans l'environnement de la bactérie, qui sont les principaux signaux de transduction du signal médié par la phosphorylation. Ainsi, le système PTS peut fonctionner comme un système de transduction du signal médié par l'état de phosphorylation de ces composants (Postma *et al.*, 1993). Suivant leur état, plusieurs des composants du PTS vont exercer une multitude de fonctions régulatrices.

2.2.2.2. Description sommaire des composants du PTS

2.2.2.2.1. Protéines générales du PTS

L'enzyme I (*ptsI*)

L'enzyme I est capable de s'autophosphoryler aux dépens du PEP lorsque celui-ci se trouve en concentration suffisante. Suite à cette première réaction, EI catalyse le transfert du groupement phosphate acquis vers la protéine HPr (Figure 13). Le site d'autophosphorylation se trouve au niveau de l'azote N3 du noyau imidazole d'un résidu histidine (Weigel *et al.*, 1982; Alpert *et al.*, 1985) ; ce processus nécessite la présence de cations bivalents Mg^{2+} et Mn^{2+}.

Chez *Bacillus subtilis*, EI, codée par le gène *ptsI*, est une protéine soluble de 570 acides aminés dont la masse moléculaire est de 66 kDa (Gonzy-Treboul & Steinmetz, 1987;

Gonzy-Treboul *et al.*, 1989; Reizer *et al.*, 1993). Chez *Escherichia coli*, il a été montré que la forme active d'EI pour l'autophosphorylation est dimérique (Misset *et al.*, 1980; Seok *et al.*, 1998). Mais c'est la forme monomérique qui catalyse la phosphorylation d'HPr (Chauvin *et al.*, 1994a; Chauvin *et al.*, 1994b).

HPr (*ptsH*)

En tant que second composant de la chaîne PTS, HPr (pour Heat resistant protein) n'est phosphorylée par P-EI que de façon transitoire. Sa très grande résistance à la chaleur et son faible poids moléculaire, variant de 8 à 10 kDa selon les espèces bactériennes (Postma *et al.*, 1993) a facilité son étude structurale chez de nombreuses bactéries (Klevit & Waygood, 1986; el-Kabbani *et al.*, 1987; Wittekind *et al.*, 1990; Herzberg *et al.*, 1992; Wittekind *et al.*, 1992; Kalbitzer & Hengstenberg, 1993; Jia *et al.*, 1994). Il a été montré que la phosphorylation par EI-P a lieu sur l'histidine en position 15, située en surface de la protéine (Kalbitzer *et al.*, 1982; Weigel *et al.*, 1982) sans changement majeur de conformation structurelle.

Chez les bactéries à Gram positif et faible pourcentage en GC, HPr peut également être phosphorylée sur la sérine en position 46. Cette réaction est catalysée par une autre enzyme bifonctionnelle, l'HPr kinase/phosphatase (Deutscher & Saier, 1983; Deutscher *et al.*, 1986; Galinier *et al.*, 1998; Kravanja *et al.*, 1999) (Figure 14). HPr ainsi phosphorylée joue alors un rôle majeur dans la régulation transcriptionnelle des gènes soumis à la régulation catabolique (cf partie 2.3), dans l'exclusion de l'inducteur (cf partie 2.2.2.3.3) et dans l'expulsion de l'inducteur (cf partie 2.2.2.3.4) (Deutscher *et al.*, 1989; Deutscher *et al.*, 1994; Deutscher *et al.*, 1997).

Chez certaines bactéries, il existe des gènes codant des protéines homologues à HPr. A titre d'exemple, les protéines DTP et MTP, respectivement chez *Escherichia coli* et *Rhodobacter capsulatus*, sont impliquées dans le métabolisme du fructose (Geerse *et al.*, 1989; Wu *et al.*, 1990). S'agissant de *Bacillus subtilis*, le gène *crh* code une protéine présentant 45 % d'identité de séquence avec HPr. Cependant, Crh ne joue pas de rôle dans le système PTS de transport des sucres, l'histidine en position 15 dans HPr étant remplacée par une glutamine dans Crh (il n'y a pas d'histidine dans l'environnement de cette glutamine).

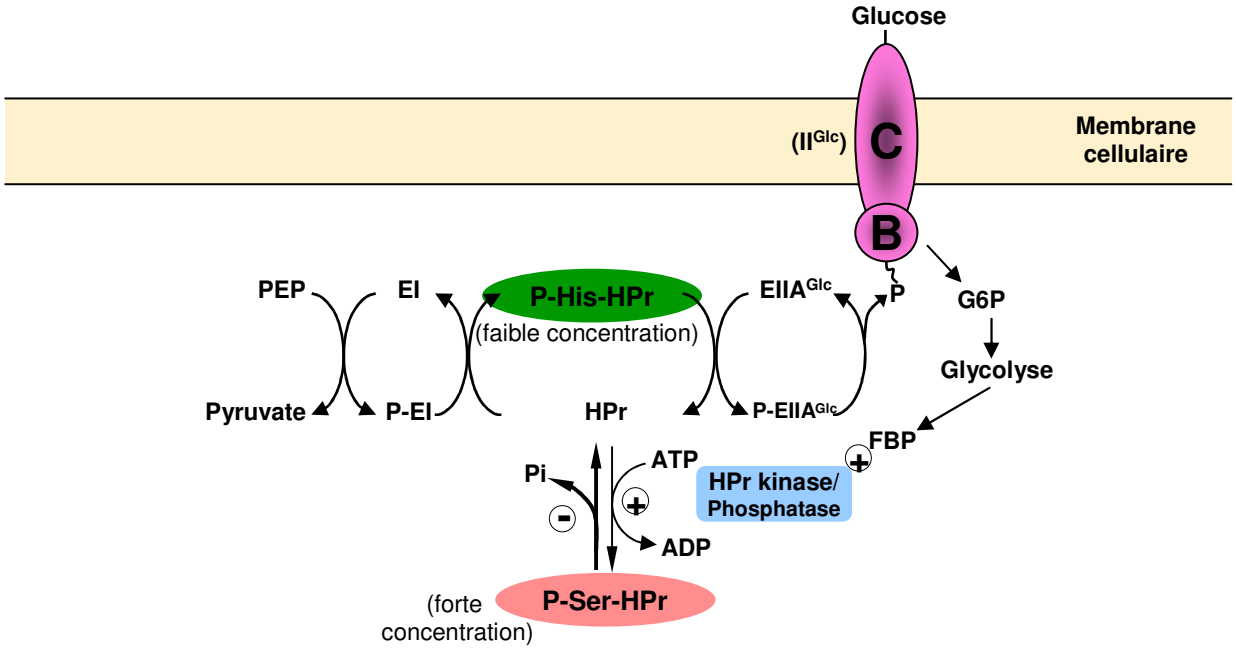


Figure 14. Phosphorylations d'HPr, PEP-dépendante par P-EI ou ATP-dépendante par HPrK/P, en présence de glucose chez *Bacillus subtilis*.

D'après Deutscher *et al.*, 2002.

2.2.2.2.2. Les protéines sucre(s)-spécifiques du PTS

Suivant la chaîne de phosphorylation du PTS, HPr-His15-P va transférer son groupement phosphate à un complexe transporteur, spécifique d'un ou plusieurs sucres, le complexe-perméase Enzyme II (Figure 13).

Le complexe Enzyme II est constitué de protéines multidomaines, de deux ou trois domaines fonctionnels (Figure 13), ou de plusieurs protéines plus petites (Postma *et al.*, 1993; Lengeler *et al.*, 1994; Robillard & Broos, 1999) :

- l'enzyme IIA est l'élément qui reçoit le groupement phosphate de HPr-His15-P. C'est une protéine hydrophile, cytoplasmique, d'environ 100 acides aminés, et toujours phosphorylée sur un résidu histidine.

- l'enzyme IIB est aussi une protéine hydrophile d'une centaine d'acides aminés. Le transfert du groupement phosphate de EIIA vers EIIB a lieu sur une cystéine dans la majorité des cas. EIIB est associée à EIIA et au domaine membranaire EIIC pour catalyser la phosphorylation du sucre ce qui permet le relargage du sucre phosphorylé dans le cytoplasme par le domaine EIIC.

- EIIC est un domaine protéique hydrophobe d'environ 350 acides aminés. Dans la majorité des cas, elle est formée de 6 à 8 hélices transmembranaires et d'une boucle hydrophile interagissant avec le domaine EIIB. EIIC correspond à la perméase spécifique permettant le transport d'un ou de plusieurs sucres et présentant ce sucre au domaine EIIB. Dans certains cas, deux domaines, EIIC et EIID, forment la perméase.

Selon la spécificité par rapport au substrat transporté, l'organisation des domaines et leurs homologies, six classes de complexe EII sont décrites : EII glucose/sucrose (EII^{Glc}), EII fructose/mannitol (EII^{Fru}), EII mannose (EII^{Man}), EII lactose (EII^{Lac}), EII glucitol (EII^{Gut}) et EII galactitol (EII^{Gat}). A l'intérieur d'une même classe, les enzymes possèdent plus de 20 % d'identité entre-elles et peuvent se complémenter au moins partiellement. La figure 13 décrit l'arrangement des différents domaines EIIA, EIIB, EIIC et EIID de ces enzymes II.

2.2.2.2.3. Les transporteurs PTS chez Bacillus subtilis

Chez *Bacillus subtilis*, 16 complexes EII ont été déduits de l'analyse complète du génome qu'il est possible de classer parmi quatre classes existantes (Reizer *et al.*, 1999; Deutscher *et al.*, 2002) : EII glucose/sucrose, EII fructose/mannitol, EII mannose et EII lactose (Figure 13). 14 d'entre eux ont été caractérisés et leur spécificité de substrat étudiée (Figure 15).

Tous les complexes PTS des classes glucose/sucrose, fructose/mannitol et lactose sont composés d'un domaine perméase EIIC. Seule la perméase du complexe PTS-lev (LevDEFG), qui transporte le fructose et le mannose, est composée de deux protéines EIIC et EIID, codées respectivement par les gènes *levF* et *levG*. Il est aussi intéressant de remarquer que seulement 9 systèmes PTS possèdent un ensemble complet de domaines. En effet, dans le cas de 6 systèmes, il manque l'enzyme EIIA et dans un cas, il manque les domaines EIIB et EIIC. Parmi les 10 complexes de la classe glucose/sucrose notamment, seuls trois sont complets et possèdent une enzyme EIIA. Les autres systèmes peuvent ou pourraient utiliser l'enzyme EIIA spécifique du glucose (*ptsG*) ou des aryl-β-glucosides (*bglP*) ou bien encore l'enzyme EIIA codée par le gène *ypqE*, qui est homologue au domaine EIIA glucose (Deutscher *et al.*, 2002). Une enzyme EIIA peut donc fonctionner avec différentes enzymes EIIB. Bien qu'homologue à une enzyme EIIA glucose, le produit du gène *yvzE* ne possède pas la partie N-terminale de ce domaine et aucune donnée expérimentale ne confirme ou n'infirme le rôle de cette protéine dans le système PTS. De même, la fonction du gène *ywbA*, qui code une protéine homologue à une perméase EIIC de la classe lactose, reste à découvrir.

2.2.2.3. Les rôles du PTS dans la régulation du transport des sucres

2.2.2.3.1. La titration de HPr-His15-P par l'HPrK/P

Chez les bactéries à Gram positif à bas GC %, nous avons vu que la protéine HPr pouvait être présente sous trois formes : HPr non phosphorylée, HPr phosphorylée par EI sur l'histidine 15 et HPr phosphorylée par l'HPrK/P sur la sérine 46. Cette dernière enzyme joue un rôle clé dans la régulation du transport par le PTS et dans la répression/activation catabolique. En présence d'un sucre rapidement métabolisable comme le glucose, l'activité glycolytique augmente et conduit à l'augmentation de la concentration en FBP et en ATP (Thompson & Torchia, 1984; Neves *et al.*, 1999; Mijakovic *et al.*, 2002) ; l'HPrK/P, stimulée par ces

Protéines PTS	Domaines structuraux et sites de phosphorylation	Domaines/ sous-unités	Nombre de résidus (aa)	Sucre transporté	Références
Classe glucose/sucrose					
Sous-classe glucose					
PtsG	C46 H62	IICBA	669	Glucose	*Gonzy et al.,1991*
GamP (YbfS)	C41 H56	IICBA	631	Glucosamine	*Gonzy et al.,1991*
NagP (YflF)		IICB	452	N-Acetylgluco-samine ?	*Gonzy et al.,1991*
MalP (GlvC)		IICB	527	Maltose	Kunst *et al.*, 1997
YpqE	H8	IIA	168	Inconnu	Kunst *et al.*, 1998
YyzE	H2	IIA	76	Inconnu	
Sous-classe sucrose					
BglP	C2 H53	IIBCA	609	Aryl-β-gluco-sides ?	Le Coq *et al.*,1995
YbbF	C3	IIBC	466	Inconnu, β-glucosisdes	Kunst *et al.*, 1998
TreP	C2	IIBC	470	Tréhalose	Schöck *et al.*,1996
SacP	C2	IIBC	460	Sucrose	Fouet *et al.*,1987
SacX	C2	IIBC	459	Sucrose	Zukowski *et al.*, 1990
Classe lactose					
LicA, LicB, LicC	H7 C7	IIA, IIB, IIC	110, 102, 452	Oligo-β-glucosides	Tobisch *et al.*, 1997
YdhN, YdhM, YdhO	H8 C8	IIA, IIB, IIC	110, 102, 442	Oligo-β-mannosides	Kunst *et al.*, 1998
YwbA		IIC	444	Inconnu	Kunst *et al.*, 1998
Classe fructose/mannitol					
FruA	H6 C17	IIABC	635	Fructose	Kunst *et al.*, 1998
MtlA, MtlF ?	C? H?	IICB, IIA ?	?	Mannitol	Kunst *et al.*, 1998
YjdD (ManP)	C9 H56	IIBCA	589	Mannose ?	Kunst *et al.*, 1998
Classe mannose					
LevD, LevE, LevF, LevG	H9 H1	IIA, IIB	146, 163	Fructose, mannose	Martin-Vertraete *et al.*, 1990
		IICIID	269, 275		

Figure 15. Systèmes de transport PTS chez *Bacillus subtilis*.

D'après Deutscher *et al.*, 2002.

effecteurs, va phosphoryler les protéines HPr et Crh de façon ATP-dépendante sur la sérine en position 46 (Galinier *et al.*, 1997; Galinier *et al.*, 1998; Reizer *et al.*, 1998; Jault *et al.*, 2000; Ramstrom *et al.*, 2003). Inversement, en présence d'une forte concentration de phosphate inorganique (Pi) ou d'acétyl-P (*in vitro*) (Ramstrom *et al.*, 2003), état physiologique qui correspond à une faible concentration de FBP (Thompson & Torchia, 1984; Mijakovic *et al.*, 2002), l'HPrK/P déphosphoryle HPr-Ser46-P et Crh-Ser46-P (Kravanja *et al.*, 1999; Fieulaine *et al.*, 2002; Lavergne *et al.*, 2002; Ramstrom *et al.*, 2003). La balance entre l'activité kinase et phosphatase est allostériquement régulée. En effet, la conformation d'HPrK/P est affectée par la présence ou l'absence de plusieurs métabolites formés lors de l'activité glycolytique (Ramstrom *et al.*, 2003) (Figure 16). De plus, la présence de cations divalents, d'ATP et d'HPr est absolument requise pour l'activité de cette enzyme (Ramstrom *et al.*, 2003). Ainsi, selon les sources de carbone présentes et utilisées et donc l'état métabolique de la cellule, l'HPrK/P va permettre l'ajustement de l'activité du système PTS.

La phosphorylation d'HPr sur la sérine en position 46 conduit à un changement d'environnement électrostatique (charge négative de la sérine phosphorylée) qui inhibe la phosphorylation d'HPr par EI-P environ 600 fois (Deutscher *et al.*, 1984; Napper *et al.*, 1996). HPr-ser46-P n'est donc plus un bon substrat pour EI-P. En présence de glucose, la forte activité kinase de l'HPrK/P inhibe donc la formation de HPr-His15-P. Par conséquent, la chaîne de phosphorylation des composants du PTS est inhibée et le transport par le PTS est ralenti. Un mutant de *Bacillus subtilis*, qui synthétise l'allèle *hprKV267F* de *Lactobacillus casei* codant une enzyme dépourvue d'activité phosphatase mais dont l'activité kinase est normale, n'est plus capable de transporter des sucres PTS (Monedero *et al.*, 2001). En effet, dans un tel mutant, HPr est sans doute piégée sous la forme HPr-Ser46-P et n'est donc plus disponible pour le transfert du groupement phosphate *via* sa forme HPr-His15-P. Si la formation de HPr-Ser46-P est empêchée par remplacement de la sérine en alanine, la croissance en présence de sucre PTS est restaurée (Eisermann *et al.*, 1988; Monedero *et al.*, 2001). Suivant cette logique, il a été montré qu'en présence de glucose, la formation d'HPr-Ser46-P inhibait le transport du fructose et du mannitol par les systèmes FruA et MltAF. En effet, si l'on perturbe la phosphorylation ATP-dépendante d'HPr en remplaçant la sérine 46 par une alanine (Eisermann *et al.*, 1988), l'effet inhibiteur du glucose sur le transport du mannitol par le PTS est diminué de 3 à 5 fois (Ye & Saier, 1996).

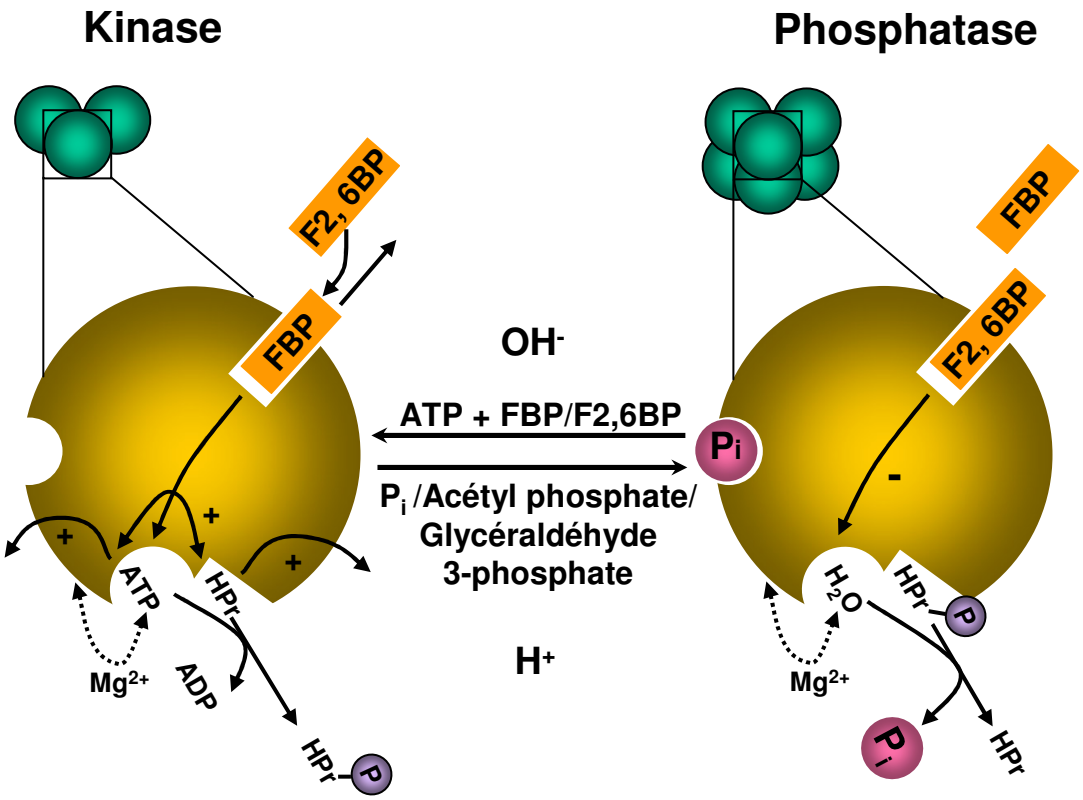


Figure 16. Modèle de la régulation de l'activité de l'Hpr kinase/phosphatase chez *Bacillus subtilis*.

L'activité kinase (trimère) ou phosphatase (hexamère) est probablement corrélée à l'état oligomérique de l'enzyme. Cet état est dépendant des effecteurs présents, qui transmettent l'information de la source de carbone utilisée et reflètent l'état métabolique de la cellule. *In* vitro, le FBP et le F26BP stimulent l'activité kinase et le F26P inhibe l'activité phosphatase.

En présence de Pi, mais aussi d'acétyl phosphate ou de glycéraldéhyde-3-phosphate, l'activité phosphatase est stimulée. Au contraire, le FBP et le F26BP inhibe cette activité.

D'après Ramström *et al.*, 2003.

Ainsi, la dualité fonctionnelle d'HPr rend son rôle central dans la régulation du métabolisme carboné chez les bactéries à Gram positif.

2.2.2.3.2. Contrôle des systèmes sans domaine EIIA de la classe glucose

Nous avons vu que chez *Bacillus subtilis*, parmi les systèmes de transport PTS de la classe glucose, certains ne possédaient pas d'enzyme IIA spécifique et utilisaient l'enzyme fonctionnelle IIA^{Glc} de la protéine PtsG. En présence de glucose, la chaîne de phosphorylation passe par le transfert du groupement phosphate de $EIIA^{Glc}$-P vers $EIIB^{Glc}$ de façon préférentielle puisqu'il s'agit d'un même polypeptide. Par conséquent, le groupement phosphate de $EIIA^{Glc}$-P est peu disponible pour les autre systèmes manquant de domaine EIIA. Ainsi, leurs sucres spécifiques ne sont pas transportés à l'intérieur de la cellule tant que le glucose est présent et assimilé (Deutscher *et al.*, 2002).

2.2.2.3.3. L'exclusion d'inducteur

Il a été observé que la présence de sucre transporté par le système PTS inhibait l'activité de transporteurs de sucres non PTS ainsi que la synthèse de différents inducteurs conduisant à l'absence de stimulation de l'expression des gènes sous contrôle de ces inducteurs (Saier & Roseman, 1976). Ce mécanisme existe mais diffère chez les bactéries à Gram négatif et à Gram positif. De plus, il est étroitement couplé à la répression catabolique, phénomène de régulation de l'expression des gènes impliqués dans le métabolisme des sucres non préférentiels, qui est développé dans la partie 2.3.

Chez *Escherichia coli* et *Salmonella typhimurium*, en présence de glucose, la protéine $EIIA^{Glc}$ se trouve majoritairement sous forme déphosphorylée puisqu'elle transfère son groupement phosphate au sucre entrant et le ratio PEP/pyruvate est faible (Hogema *et al.*, 1998). $EIIA^{Glc}$ déphosphorylée interagit alors avec les perméases spécifiques du transport du lactose (LacY) (Figure 17), du galactose (GalP), du maltose (MalK), du mélibiose (MelB) et du raffinose (RafB) (Kornberg *et al.*, 1980; Osumi & Saier, 1982a; Osumi & Saier, 1982b; Nelson *et al.*, 1984; Dean *et al.*, 1990b; Titgemeyer *et al.*, 1994; Saier & Crasnier, 1996). Cette interaction inhibe l'activité de transport de ces perméases en masquant le site de fixation du substrat (Saier & Crasnier, 1996), aboutissant à une assimilation préférentielle des sucres comme le glucose par le PTS aux dépens des autres systèmes (Figure 17). En présence de l'inducteur, $EIIA^{Glc}$ reconnaît une séquence spécifique du transporteur non PTS

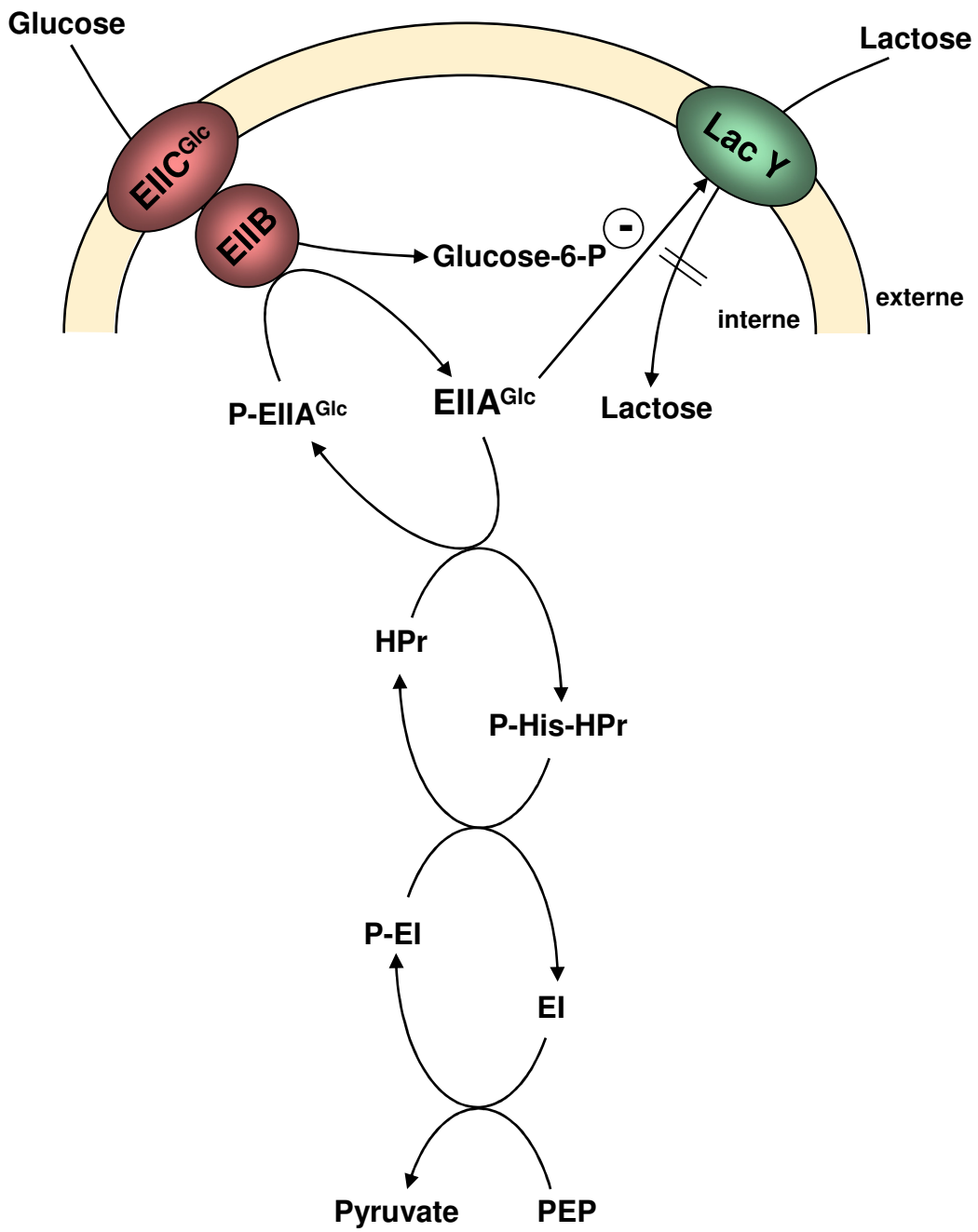


Figure 17. Modèle d'exclusion de l'inducteur chez les bactéries à Gram négatif.

correspondant, VGANXSLXSX (Titgemeyer *et al.*, 1994), d'où le nom d'exclusion de l'inducteur donné à ce mécanisme.

Chez les bactéries à Gram positif, le mécanisme d'exclusion de l'inducteur est médié par HPr-Ser46-P (Figure 18). Toutefois, il n'existe pas d'exemple connu de régulation par exclusion de l'inducteur chez *Bacillus subtilis* bien qu'un tel mécanisme est supposé exister chez cette bactérie (Dahl, 2002; Deutscher *et al.*, 2002).

Le cas le plus étudié est l'exclusion du lactose chez deux bactéries lactiques, *Lactococcus lactis* et *Lactococcus brevis*. De façon surprenante, cette dernière semble dépourvue de système de transport PTS (à l'exception du système PTS spécifique du transport du fructose (Saier *et al.*, 1996)) mais exprime bien HPr et HPrK/P (Reizer *et al.*, 1988). Dans ce cas, on peut remarquer que la régulation de l'expression des gènes *via* le système global dont HPr est un acteur central (cf partie 2.3) est complètement indépendante du système PTS. Cependant, il a été montré que la protéine mutée de *Bacillus subtilis* HPr-Ser46Asp (sérine remplacée par un aspartate mimant la phosphorylation), qui est homologue à HPr-Ser-P de *Lactococcus brevis* (Wittekind *et al.*, 1989), est capable d'interagir avec les symporteurs glucose:proton et lactose:proton et de changer leur conformation (Ye *et al.*, 1994a; Ye & Saier, 1995a; Ye & Saier, 1995b). Cette interaction inhibe l'activité de transport de ces perméases par le découplage du transport des protons fournissant l'énergie électrochimique (Saier & Crasnier, 1996) (Figure 19). Ainsi ces transporteurs ne fonctionnent plus que comme uniporteurs pour permettre l'efflux de lactate provenant du catabolisme du glucose. Chez *Lactococcus lactis*, HPr-Ser46-P interagit avec la perméase PTS spécifique du lactose pour inhiber son activité (Ye *et al.*, 1994b). Enfin, on peut citer l'exemple plus récent de l'inhibition du système de transport du maltose chez *Lactobacillus casei*, restant cependant à préciser (Dossonnet *et al.*, 2000; Viana *et al.*, 2000).

Il est intéressant de remarquer que ce mécanisme n'est pas unique (Deutscher *et al.*, 2002). En effet, chez trois autres bactéries à Gram positif, *Streptococcus thermophilus* (Poolman *et al.*, 1989), *Lactobacillus bulgaricus* (Leong-Morgenthaler *et al.*, 1991) et *Leuconostoc lactis* (Vaughan *et al.*, 1996), l'activité de la lactose perméase LacS est inhibée par la phosphorylation sur un domaine de fusion homologue au domaine $EIIA^{Glc}$ (Gunnewijk & Poolman, 2000a; Gunnewijk & Poolman, 2000b).

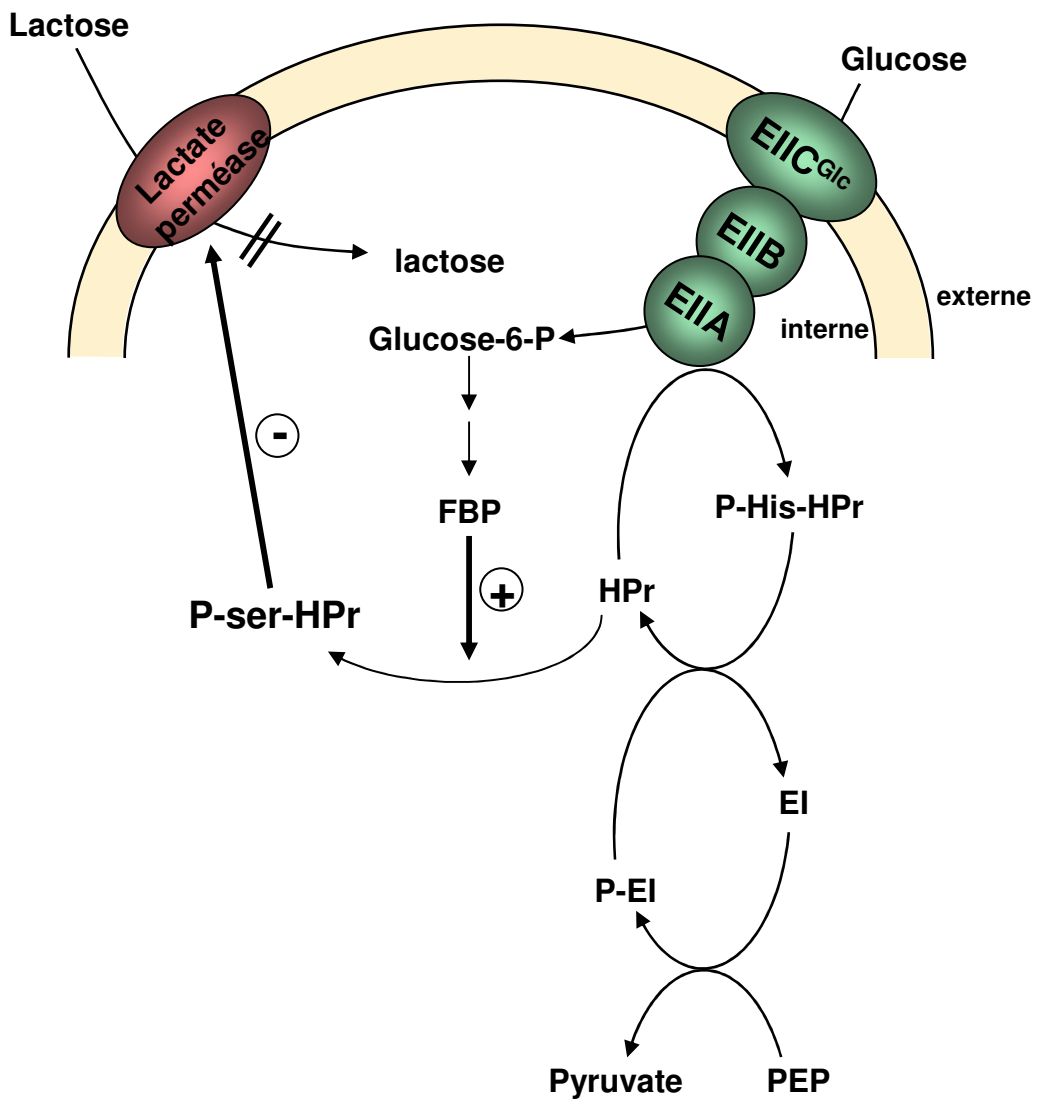


Figure 18. Modèle d'exclusion de l'inducteur chez les bactéries à Gram positif.

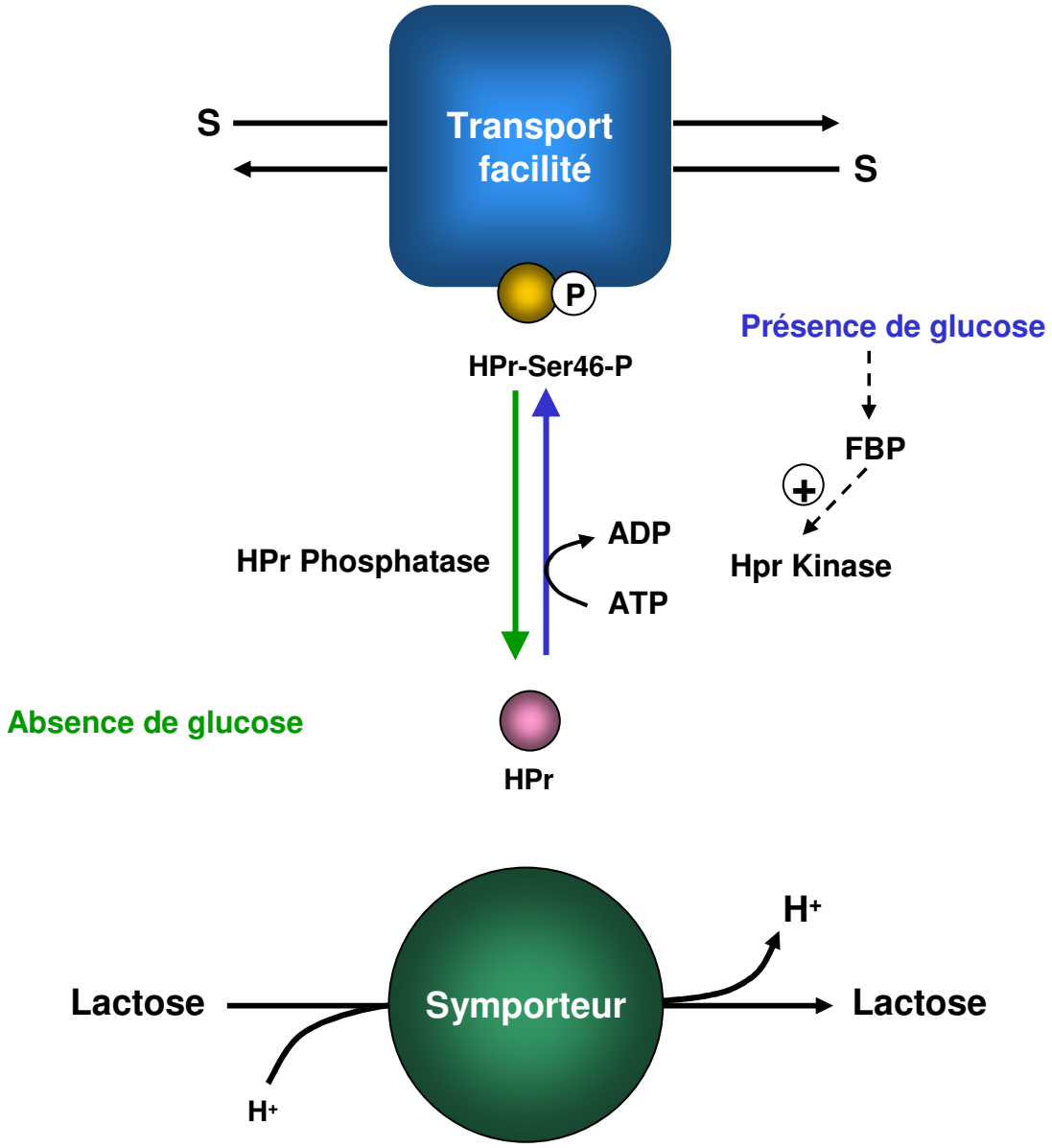


Figure 19. Modèle de modulation de l'activité du symporteur lactose par HPr-Ser46-P chez *Lactococcus brevis*.

2.2.2.3.4. Cas du transport du glycérol

Il a été montré que $EIIA^{Glc}$ déphosphorylée était aussi capable d'interagir avec la glycérol kinase d'*Escherichia coli* et de *Salmonella typhimurium* et d'inhiber son activité en présence de glycérol (Postma *et al.*, 1984; Novotny *et al.*, 1985; Hurley *et al.*, 1993). En l'absence de glycérol kinase active, le glycérol-3-phosphate n'est plus synthétisé. Ce métabolite est l'inducteur de l'expression des gènes du transport et du catabolisme du glycérol. Cette régulation allostérique de la glycérol kinase permet donc de favoriser le métabolisme du glucose en empêchant celui du glycérol. Par ailleurs, Postma et ses collègues ont observé un effet de rétroinhibition de cette régulation. Lorsque le système est préincubé pendant un certain temps dans un milieu contenant du glycérol, la glycérol kinase est capable d'inhiber jusqu'à 30 % du transport PTS (EII^{Glc}) du méthyl α-D-glucopyranoside (Rohwer *et al.*, 1998).

Chez *Bacillus subtilis*, la glycérol kinase joue également un rôle indirect dans la régulation du transport du glycérol. En effet, en phosphorylant le glycérol pénétrant la cellule par diffusion facilitée, elle permet son catabolisme mais aussi le maintien d'un gradient différentiel de concentration en glycérol entre le milieu extérieur et le cytoplasme et donc le maintien du transport. Or, l'absence de glucose ou d'un autre sucre rapidement métabolisable dans le milieu de croissance aboutit à l'accumulation d'HPr-His15-P. Il a été montré qu'HPr-His15-P était capable d'interagir avec la glycérol kinase et ainsi de stimuler son activité kinase et donc le transport du glycérol. Le mécanisme de cette interaction est connu. La phosphorylation d'un résidu Histidine (His230) par HPr-His15-P ou Enzyme I favorise la tétramérisation de la glycérol kinase ; cela augmente l'activité enzymatique et aussi l'affinité de GlpK pour la glycérol perméase GlpF (Reizer *et al.*, 1984; Beijer & Rutberg, 1992; Wehtje *et al.*, 1995; Darbon *et al.*, 2002).

2.2.2.3.5. L'expulsion d'inducteur

Il s'agit du mécanisme par lequel l'ajout d'un sucre rapidement métabolisable dans le milieu de culture conduit à l'expulsion des autres sucres préalablement accumulés dans le cytoplasme. Ce mécanisme n'a pas été détecté chez *Bacillus subtilis*. Cependant, il a été montré chez d'autres bactéries à Gram positif comme *Lactococcus lactis* ou *Streptrococcus pyogenes* lors d'études sur le transport des β-galactosides (Reizer & Panos, 1980; Thompson & Saier, 1981). Il avait alors été observé que des bactéries cultivées en présence de PEP et de méthyl-β-D-thiogalactopyranoside (TMG), un analogue non métabolisable du lactose,

accumulaient le TMG sous forme phosphorylée. Mais l'ajout d'un sucre rapidement métabolisé provoquait une expulsion du TMG sous forme déphosphorylée. Ainsi, en présence de glucose, HPr se trouve majoritairement sous la forme HPr-Ser46-P. Celle-ci stimule l'activité d'une sucre-phosphate phosphatase (Pase II) qui va alors déphosphoryler le sucre-phosphate pour permettre son expulsion hors de la cellule (Reizer *et al.*, 1983). La Figure 20 illustre ce mécanisme.

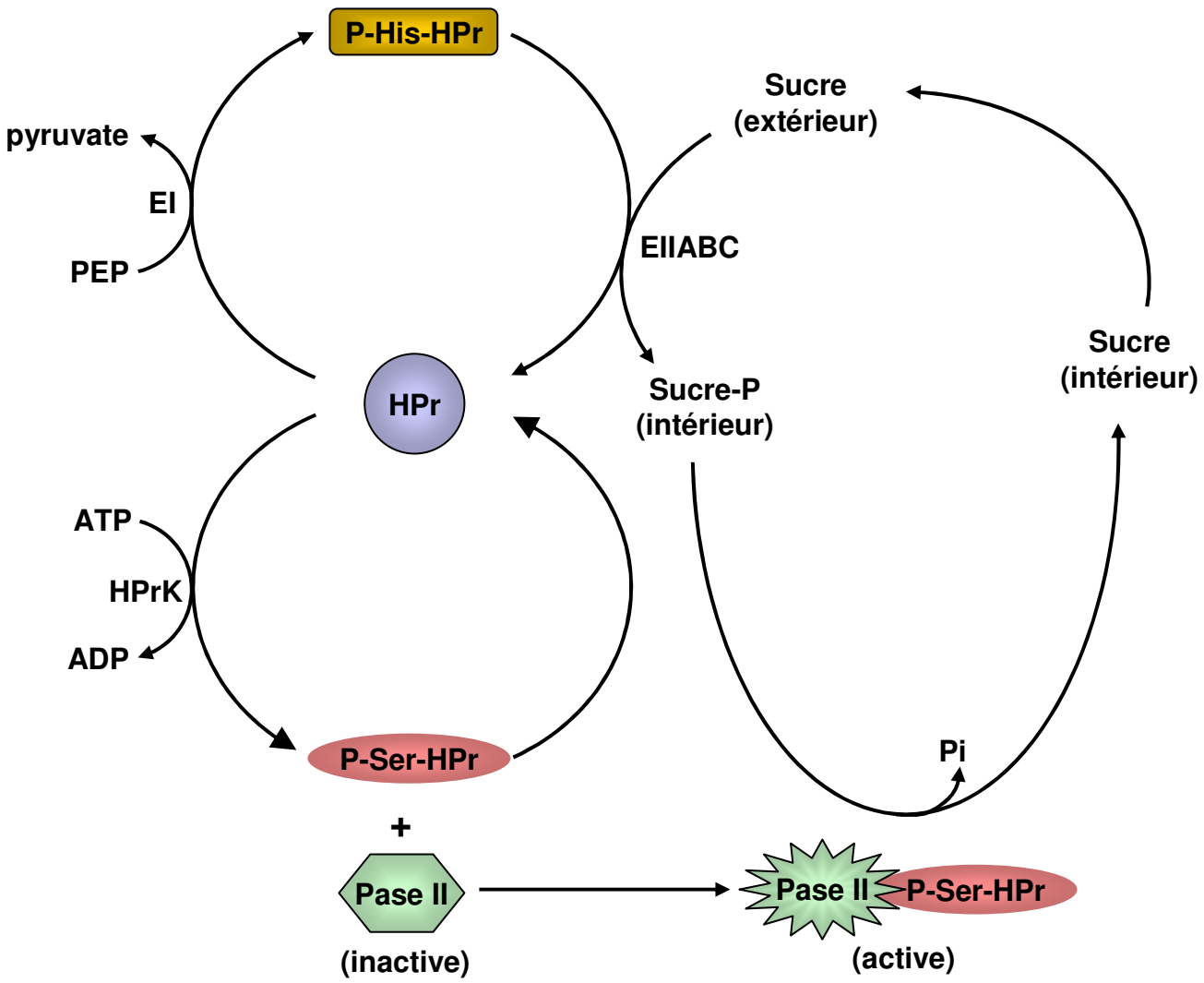


Figure 20. Modèle du mécanisme d'expulsion de l'inducteur chez les bactéries à Gram positif.

D'après Saier *et al.*, 1996.

2.3. Les régulations de l'expression des gènes impliqués dans le métabolisme du carbone

La régulation de l'expression des gènes va de pair avec la régulation du transport des sources de carbone, l'un contrôlant directement l'expression des protéines (enzymes ou régulateurs) et l'autre contrôlant directement leur activité. La synergie de ces deux systèmes, chacun contrôlant des éléments de l'autre, permet de réguler très finement l'ensemble du métabolisme ou d'amplifier tel ou tel phénomène régulatoire. L'observation d'un phénomène de régulation dans l'utilisation des sucres et son analyse ont été faites pour la première fois chez Bacillus subtilis par Jacques Monod (Monod, 1942). Il a observé que lors de la croissance en présence de glucose et de sorbitol dans le milieu de culture, deux phases de croissance étaient distinguables, la première plus rapide que la seconde et séparées par un temps de latence (Figure 21). Ce phénomène a été nommé diauxie. Par la suite, il montra que les sources de carbone était utilisées hiérarchiquement. Le glucose constituait la source de carbone utilisée pendant la première phase de croissance puis le sorbitol était utilisé pendant la seconde phase, lorsque le glucose était épuisé. Ce phénomène d'inhibition de l'utilisation d'une source de carbone « secondaire » par la présence d'un sucre métabolisable de manière plus efficace sera par la suite appelé répression catabolique.

De manière générale, on peut distinguer deux facettes dans la régulation de l'expression des gènes du métabolisme du carbone. D'une part, le mécanisme global de la répression/activation catabolique *sensu stricto* permet de réprimer l'expression des gènes impliqués dans l'utilisation de sources de carbone non préférentielles en présence de sucres rapidement métabolisables comme le glucose et, à l'opposé, de stimuler l'expression des gènes impliqués dans l'utilisation de ces sources préférées de carbone et d'énergie. Ce mécanisme global intimement lié au système PTS existe chez de très nombreuses bactéries dont les autres modèles *Escherichia coli* et *Salmonella typhimurium* (Monod, 1942). D'autre part, dans beaucoup de cas, l'induction de l'expression des enzymes impliquées dans le transport et le catabolisme « préliminaire » d'une source de carbone particulière (ou d'un type de molécules carbonées) pour permettre son entrée dans les voies centrales du métabolisme a lieu en présence de celle-ci. La majorité de ces systèmes spécifiques est directement contrôlée par le système PTS. Ils seront présentés dans ce chapitre. Mais il existe quelques systèmes dont le contrôle ne dépend pas du PTS, du moins directement, mais uniquement de la présence de métabolites inducteurs ou anti-inducteurs. Par exemple, l'expression de l'opéron permettant le transport

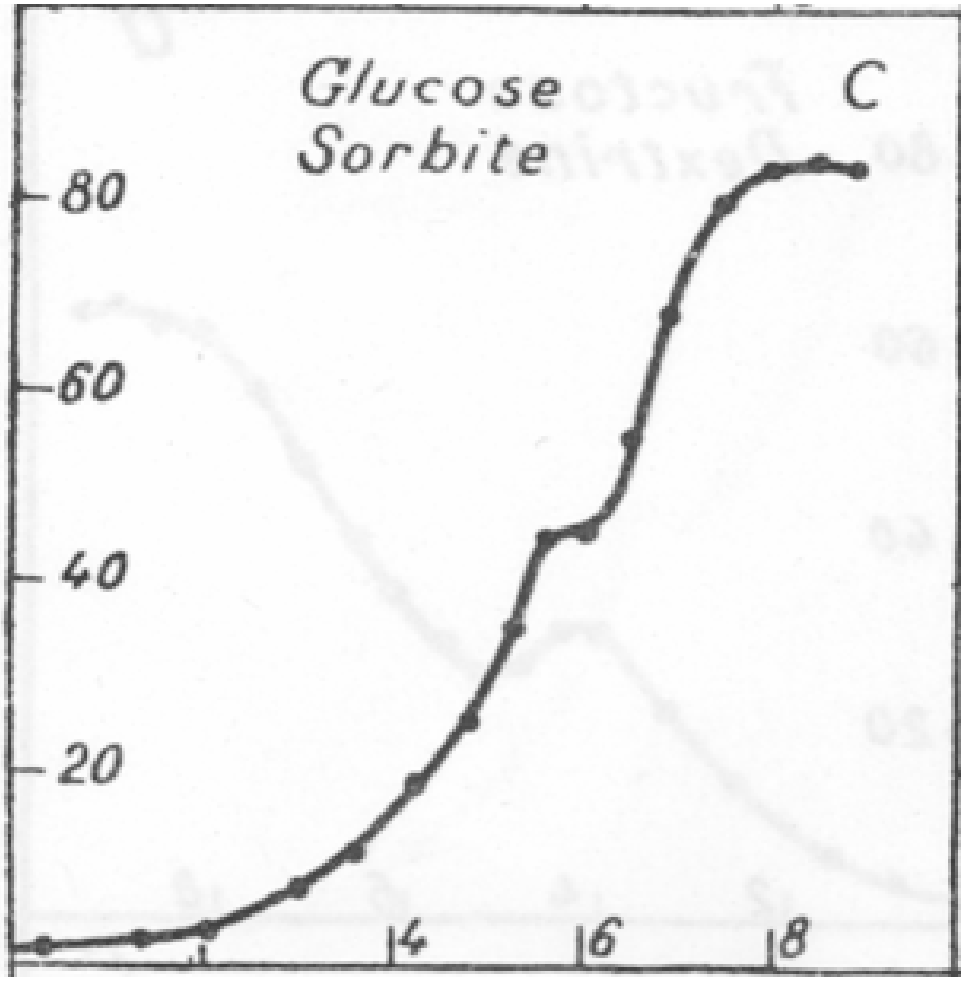


Figure 21. Phénomène de diauxie observé par Monod dans les années 1940.
D'après Monod, 1942.

(bien que PTS) et la conversion du tréhalose en G6P et en glucose est exclusivement régulée par le répresseur transcriptionnel TreR. En présence de tréhalose-6-phosphate (T6P), l'activité de TreR est inhibée, ce qui permet l'expression de l'opéron tréhalose (Burklen *et al.*, 1998). Mais en présence de glucose, le G6P formé empêche l'interaction entre TreR et le T6P (Burklen *et al.*, 1998). Alors, TreR réprime l'expression des gènes de l'utilisation du tréhalose. De même, le répresseur transcriptionnel XylR, en présence de xylose, active l'expression des gènes du transport et du catabolisme de ce sucre (Dahl *et al.*, 1994; Kraus *et al.*, 1994). Mais, la présence de G6P perturbe l'inhibition de XylR par le xylose (Dahl *et al.*, 1995).

Ces deux aspects de la régulation sont intégrés. Par exemple, en présence de glucose, l'exclusion de l'inducteur, médiée par le système PTS, empêche l'expression des gènes sous contrôle de cet inducteur (Saier & Roseman, 1976). Alternativement, un régulateur spécifique peut être sensible à plusieurs signaux dont les effets sont hiérarchisés auquel cas, l'assimilation d'une source de carbone rapidement métabolisable peut aboutir à un signal d'ordre supérieur au signal donné par l'inducteur spécifique (Saier & Crasnier, 1996; Stulke & Hillen, 1999).

S'agissant des enzymes des voies centrales du métabolisme, et surtout celles de la glycolyse/néoglucogenèse, il semble évident que leur expression est nécessaire au fonctionnement de la cellule, mais pas toujours au même niveau, suivant le rôle de chaque voie dans le flux nécessaire à l'élaboration des précurseurs de l'anabolisme et à la production d'énergie.

Chez les bactéries, la régulation de l'expression génique s'exerce le plus souvent au niveau du contrôle de la transcription des gènes. Contrairement au cas des Eucaryotes, les phénomènes de régulation post-transcriptionnels et post-traductionnels semblent peu fréquents chez les Procaryotes. Suivant les cas, les mécanismes de contrôle sont différents et répondent à différents signaux : répression ou activation de la transcription, terminaison/antiterminaison transcriptionnelle, etc.... Par ailleurs, l'organisation en opéron est une caractéristique particulière aux espèces procaryotiques qui permet de réguler concomitamment un ensemble de gènes remplissant souvent la même fonction métabolique.

2.3.1. L'organisation génétique des gènes du métabolisme central du carbone (catalytiques et régulateurs)

La disponibilité de la séquence du génome de *Bacillus subtilis* rend aisée l'analyse de l'organisation des gènes impliqués dans le métabolisme central du carbone. D'une façon générale, on remarque que les gènes codant les enzymes adjacentes dans les voies métaboliques sont souvent organisés en opéron. De plus, cette organisation est relativement bien conservée chez les espèces proches voire beaucoup plus éloignées de *Bacillus subtilis*, de façon cohérente avec le caractère essentiel du métabolisme central du carbone (Deutscher *et al.*, 2002; Sonenshein, 2002). Pour faciliter la lecture et le repérage des gènes et opérons impliqués dans les voies centrales du métabolisme du carbone, les Figures 22, 23, 24 et 25 listent ces gènes et permettent de se rendre compte de leur organisation génétique.

2.3.2. La répression/activation catabolique médiée par CcpA

2.3.2.1. Mécanisme général

Chez les bactéries à Gram positif et à faible pourcentage en GC, la répression/activation catabolique *stricto sensu* implique les protéines HPrKinase/Phosphatase, HPr-Ser46-P et CcpA (Deutscher & Saier, 1983; Deutscher *et al.*, 1984; Deutscher *et al.*, 1989; Deutscher *et al.*, 1994; Deutscher *et al.*, 1995; Deutscher *et al.*, 1997; Stulke & Hillen, 2000; Bruckner & Titgemeyer, 2002; Deutscher *et al.*, 2002).

Historiquement, l'étude des déterminants de la répression catabolique chez les bactéries à Gram positif a été initiée par l'étude transcriptionnelle du gène *amyE* chez *Bacillus subtilis*. Ce gène code une α-amylase dégradant l'amidon en glucose. Ces études permirent de définir les éléments en *cis* et en *trans* du mécanisme de répression. L'intégrité d'une séquence ADN, située juste après le promoteur du gène *amyE* et chevauchant le site de début de la transcription, est indispensable au phénomène de répression catabolique (Nicholson & Chambliss, 1985; Nicholson & Chambliss, 1986; Nicholson *et al.*, 1987; Weickert & Chambliss, 1990; Hueck *et al.*, 1994). Cette séquence, baptisée *cre*, pour catabolite responsive

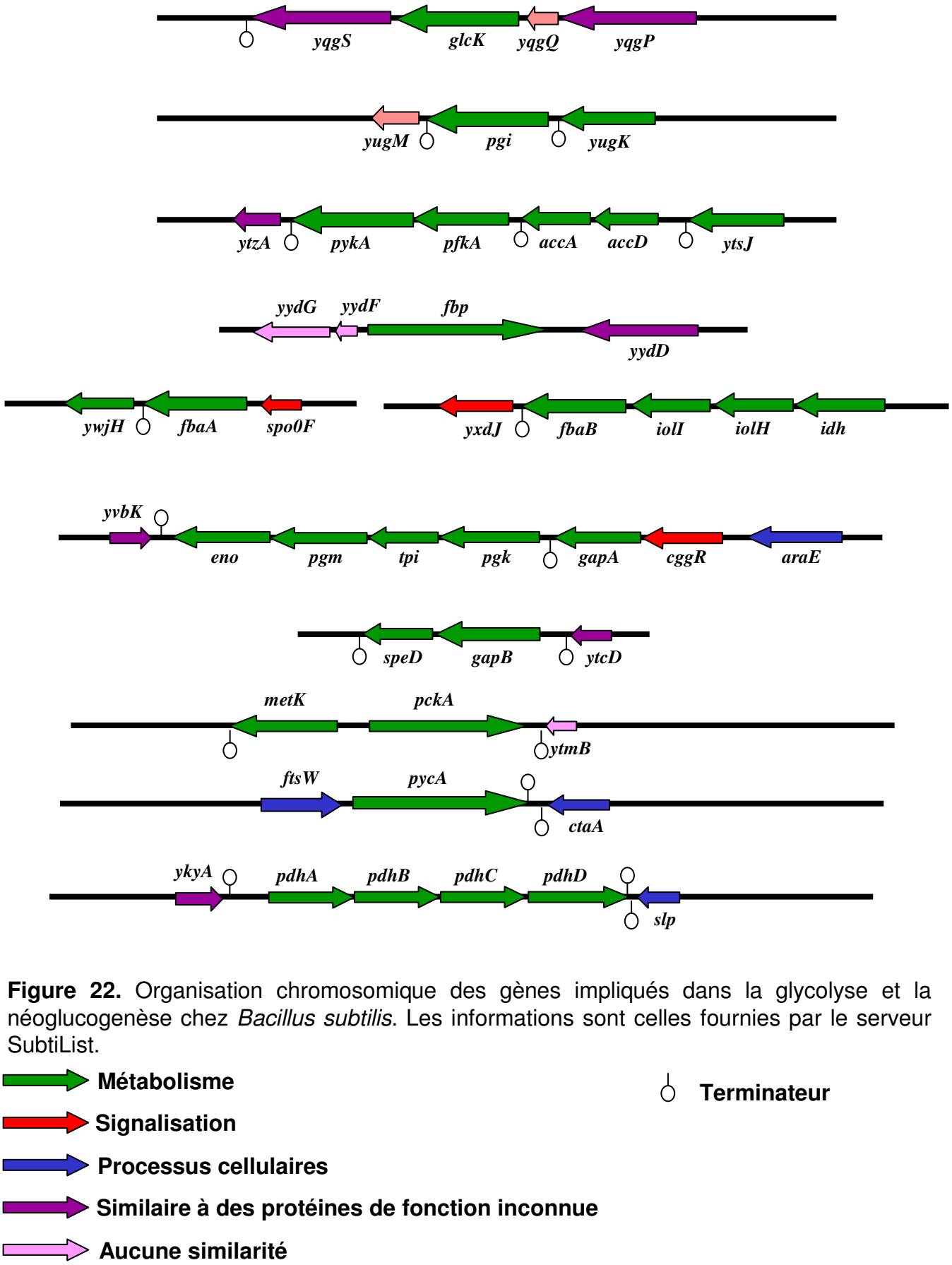


Figure 22. Organisation chromosomique des gènes impliqués dans la glycolyse et la néoglucogenèse chez *Bacillus subtilis*. Les informations sont celles fournies par le serveur SubtiList.

Métabolisme

Signalisation

Processus cellulaires

Similaire à des protéines de fonction inconnue

Aucune similarité

Terminateur

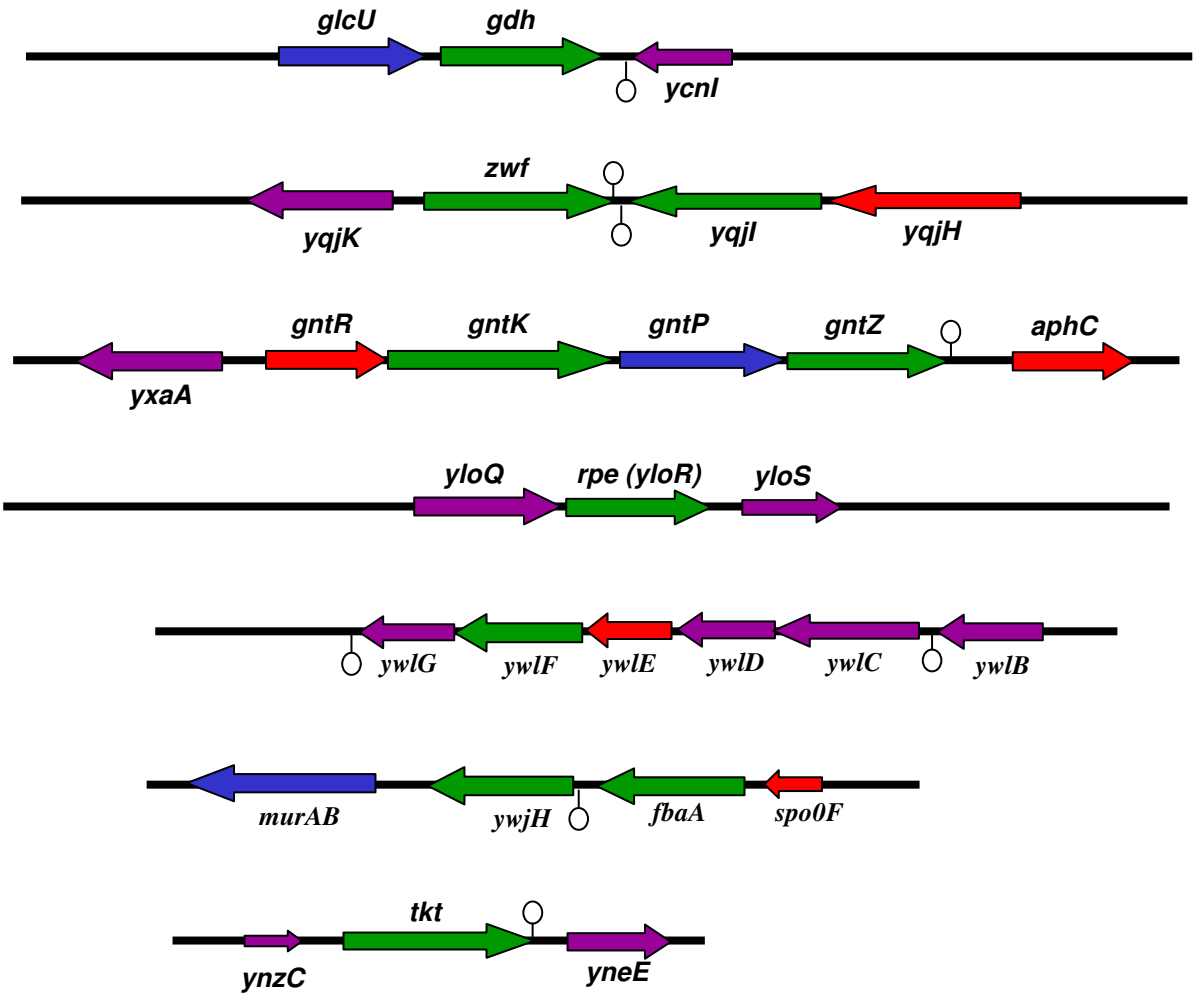


Figure 23. Organisation chromosomique des gènes impliqués dans la voie des pentoses phosphates chez *Bacillus subtilis*. Les informations sont celles fournies par le serveur SubtiList.

Métabolisme

Signalisation

Processus cellulaires

Similaire à des protéines de fonction inconnue

Aucune similarité

Terminateur

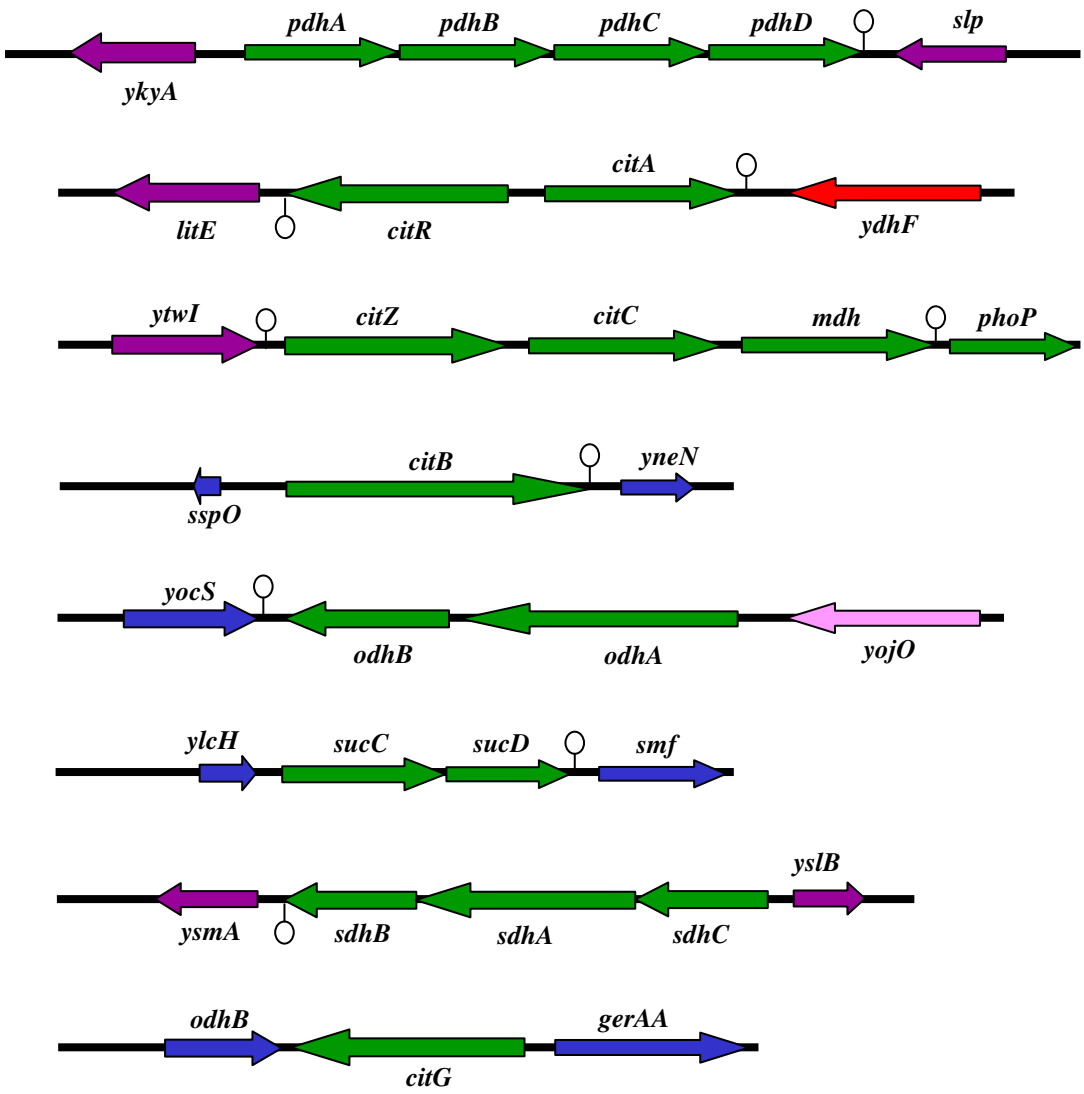


Figure 24. Organisation chromosomique des gènes impliqués dans le cycle de Krebs et les enzymes rattachées chez *Bacillus subtilis*. Les informations sont celles fournies par le serveur SubtiList.

Métabolisme

Signalisation

Processus cellulaires

Similaire à des protéines de fonction inconnue

Aucune similarité

Terminateur

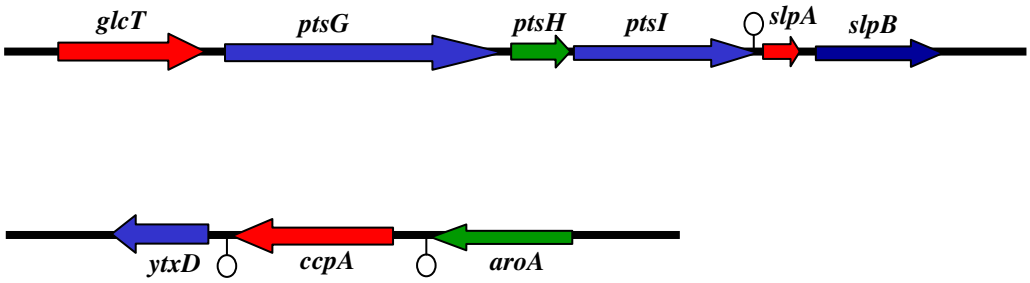


Figure 25. Organisation chromosomique des gènes codant les enzymes générales du systèmes PTS et le régulateur CcpA chez *Bacillus subtilis*. Les informations sont celles fournies par le serveur SubtiList.

Métabolisme

Signalisation

Processus cellulaires

Similaire à des protéines de fonction inconnue

Aucune similarité

Terminateur

element, constitue l'opérateur du régulateur transcriptionnel pléïotropique CcpA, pour catabolite control protein A, aussi requis pour la répression catabolique (Henkin *et al.*, 1991).

L'ensemble des études par mutagenèse établissant la séquence *cre* dans la région régulatrice des opérons *amyE* (α-amylase), *acsA* (acétyl-CoenzymeA synthétase), *hut* (métabolisme des β-xylosides) et *lev* (métabolisme du fructose et de ses polymères) de *Bacillus subtilis* a permis de définir le consensus suivant : TGWAARC│GYTWNCW (le symbole │ représentant l'axe de symétrie du palindrome ; W correspond à A ou T, R à A ou G et Y à C ou T) (Stulke & Hillen, 2000). Ce consensus est un peu plus précis que celui défini par Weickert et Chambliss (1990), TGWNANC│GNTNWCA, qui considéraient alors les résultats de mutagenèse du site *cre* devant le gène *amyE* et les alignements de séquence chez plusieurs Bacillacées. Après eux, d'autres auteurs ont montré de façon expérimentale le rôle de la séquence *cre* dans la régulation de certains gènes soumis à la répression catabolique chez *Bacillus subtilis*, ce qui a permis de confirmer la séquence consensus (Table 2) (Hueck & Hillen, 1995; Miwa *et al.*, 2000). A partir de ce motif légèrement dégénéré, une recherche systématique d'une boîte *cre* dans le génome de *Bacillus subtilis* a permis de dénombrer 126 séquences *cre*, dont certaines sont expérimentalement démontrées et les autres putatives (Miwa *et al.*, 2000). Dans cette dernière étude, l'équipe de Yasutaro Fujita rend flagrante la corrélation entre la présence du motif *cre* pour la fixation de CcpA et l'existence d'une répression/activation catabolique (Table 3). Par ailleurs, des recherches similaires, qui sont basées sur un consensus *cre* un peu différent, ont permis d'identifier d'autres motifs *cre* potentiels dans la région régulatrice de certains opérons connus pour être soumis à la répression catabolique en présence de glucose (Miwa *et al.*, 2000; Miwa & Fujita, 2001).

Récemment, la comparaison transcriptomique de l'expression des gènes dans un mutant *ccpA* et une souche sauvage (souche ST100), en présence ou en absence de glucose (mais dans un milieu riche LB) a permis de compléter les données précédentes (Moreno *et al.*, 2001) (Table 3). En effet, les données obtenues confirment le rôle des séquences *cre* déjà étudiées et suggèrent celui de certains autres motifs *cre* jusqu'alors putatifs. Ils ont aussi montré que CcpA médiait la répression ou l'activation de certains gènes indépendamment de la présence du glucose (bien que celui-ci peut jouer un rôle mineur sur le niveau de régulation de ces gènes) ou du site *cre*. Cependant, ce type de régulation n'intervient pas dans le contrôle du métabolisme du carbone.

Gène ou opéron	Séquence *cre*	position du *cre*		position par rapport au +1	régulation
ackA	TGTAAGCGTTATCA	-155	-142	-64	activation
acsA	TGAAAGCGTTACCA	+7	+20	+37	répression
acuA	TGAAAACGCTTTAT	-74	-61		répression
amyE	TGTAAGCGTTAACA	-124	-111	-3	répression
bglP	TGAAAGCGTTGA	-252	-239	chevauchant	répression
cccA	TGTAAGCGTATACA	-187	-174		répression
gntR	TGAAAGCGGTACCA	+107	+120	+141	répression
hutP	TGAAACCGCTTCCA	+170	+183	+203	répression
levD	TGAAAACGCTT-A-ACA	-79	-65		répression
pta	TGAAAGCGCTATAA	-99	-86	-62	activation
xylA	TGGAAGCGTAAACA	+35	+38	+134	répression
xyn	TGAAAGCGCTTTTA	-97	-85	+229	répression
Consensus	TGWNANCGNTNWCA				

Table 2. Séquences *cre* des gènes et opérons chez *Bacillus subtilis* citées dans la revue bibliographique. La position du site *cre* est indiquée par rapport au premier nucléotide du codon d'initiation du gène ou opéron. La position du site *cre* (premier nucléotide) est aussi donnée par rapport au site de démarrage de la transcription.
D'après Hueck et Hillen, 1995.

Gène[a]	Fonction	Position du site *cre*[b]	RC/AC[c]	Références
ackA	**Acétate kinase**	**-155 -142**	+	Tobisch *et al.*, 1999 ; Turinsky *et al.*, 1998
acoA	**métabolisme de l'acétoine**	**+434 +447**	-	Miwa *et al.*, 2000
acsA	**Acétyl-CoA synthétase**	**+7 +20**	-	Grundy *et al.*, 1994 ; Zalieckas *et al.*, 1998
acuA	Métabolisme de l'acétoine	-74 -61	-	Grundy *et al.*, 1994
alsS	Synthèse de l'acétoine	-127 -114	+	Deutscher *et al.*, non publié ; Miwa *et al.*, 1994 ; Zahler *et al.*, 1976
amyE	**α-amylase**	**-124 -111**	-	Henkin *et al.*, 1991 ; Nicholson *et al.*,1986
araA	**Métabolisme de l'arabinose**	**+679 +692** -50 -37	-	Deutscher *et al.*, non publié ; Miwa *et al.*, 2000
araBDLN				
bglA	6-P-β-Glucosidase	-83 -70	-	Moreno *et al.*, 2001
bglP	**Métabolisme des β-glucosides**	**-252 -239** +1517 +1530	-	Krüger *et al.*,1995 Deutscher *et al.*, non publié
blt	efflux de drogues multiples	non identifié	-	Moreno *et al.*, 2001
bsaA	Glutathione peroxydase	non identifié	-	Moreno *et al.*, 2001
cccA	**Cytochrome c-550**	**-187 -174**	-	Miwa *et al.*, 2000 ; Monedero *et al.*, 2001 ; Moreno *et al.*, 2001
ykuL-ccpC	**Régulation du cycle de Krebs**	**-108 -92**	-	Kim *et al.*, 2002
citB	Aconitase	-181 -168	-	Deutscher *et al.*, non publié ; Tobisch *et al.*,1999b
citG	Fumarase	non identifié	-	Tobisch *et al.*,1999b
citM	**transporteur citrate**	**-31 -18**	-	Miwa *et al.*, 2000 ; Warner *et al.*, 2000
citST	Régulation TCS	+1299 +1312	-	Deutscher *et al.*, non publié ; Moreno *et al.*, 2001
citZCH	**Cycle de Krebs**	**-113 -100**	-	Kim *et al.*, communication personnelle ; Tobisch *et al.*,1999b
comEC	Transport de l'ADN ?	non identifié	-	Moreno *et al.*, 2001
cstA		+2 +15	-	Miwa *et al.*, 2000 ; Moreno *et al.*, 2001
ctaC	Cytochrome c oxydase II	-206 -193	-	Liu *et al.*, 1998
ctaD	Cytochrome c oxydase I	+5 +18	-	Moreno *et al.*, 2001
ctaE	Cytochrome c oxydase III			
cydA	Cytochrome bd ubiquinol oxydase	non identifié	+	Moreno *et al.*, 2001
cysK	Cystéine synthétase	non identifié	-	Moreno *et al.*, 2001
dhaS	Aldéhyde déshydrogénase	non identifié	-	Moreno *et al.*, 2001
dhbB	Isochorismataste	non identifié	-	Moreno *et al.*, 2001
etfA	Transfert d'électron	+363 376	-	Moreno *et al.*, 2001
exu	Métabolisme des hexuronates	non identifié	-	Mekjian *et al.*, 1999
gabP	perméase à l'γ-aminobutyrate	-141 -128	-	Miwa *et al.*, 2000 ; Moreno *et al.*, 2001
galT	**Métabolisme du galactose-1-P**	**+214 +227**	-	Miwa *et al.*, 2000
gapA	GAPDH (glycolyse)	pas de site cre fonctionnel	+	Tobisch *et al.*,1999b
gapB	GAPDH (néoglucogenèse)	+375 +388	-	Moreno *et al.*, 2001
glpF	**Facilitateur d'entrée du glycérol**	**-181 -168**	-	Darbon *et al.*, 2001 ; Miwa *et al.*, 2000 Moreno *et al.*, 2001
glpQT	Perméase au glycérol 3-P	-158 -145	-	Deutscher *et al.*, non publié ; Moreno *et al.*, 2001
glvA	6-P-α-glucosidase	-25 -12	-	Miwa *et al.*, 2000 ; Moreno *et al.*, 2001
glvC	$EIIBC^{Mal}$	+275 +262		Miwa *et al.*, 2000 ; Moreno *et al.*, 2001
gntK	Gluconate kinase	+434	-	
gntR	**Métabolisme du gluconate**	**+107 +120**	-	Deutscher *et al.*,1994 ; Fujita *et al.*,1995
gutB	Métabolisme du sorbitol	-44 -31	-	Deutscher *et al.*,1994 ; Ye *et al.*, 1994b
hutP	**Métabolisme de l'histidine**	**+170 +183**	-	Wray *et al.*, 1994
ivlB	Acétolactate synthase	non identifié	+	Moreno *et al.*, 2001
iolA	Métabolisme de l'inositol	-105 -92	-	Moreno *et al.*, 2001
iolB	**Métabolisme de l'inositol**	**+668 +681**	-	Deutscher *et al.*,1994 ; Miwa *et al.*, 2000
iolC	Métabolisme de l'inositol			
kdgA	Métabolisme des hexuronates	+355 +368	-	Miwa *et al.*, 2000 ; Pujic *et al.*, 1998
lcfA	**Acyl-CoA synthétase**	**+412 +425**	-	Miwa *et al.*, 2000
lctE	Lactate déshydrogénase	-137 -124	+	Deutscher *et al.*, non publié ; Moreno *et al.*, 2001
levD	**Métabolisme**	**-79 –65**	-	Martin-Verstraete *et al.*, 1999 ; Martin-Verstraete *et al.*, 1995
licA				
licB		**-76 –63**	-	Miwa *et al.*, 2000 ; Tobisch *et al.*,1999a
licC				
licH				

mmgA		**-36 –23**	-	Bryan *et al.*,1996
msmRE	Métabolisme du mélibiose	-85 –72	-	Miwa *et al.*, 2000 ; Moreno *et al.*, 2001
msmX	**ABC transporteur de sucre**	**-34 –21**	-	Miwa *et al.*, 2000
mtlA	Métabolisme du mannitol	-94 -81 +1156 +1169	-	Deutscher *et al.*,1994
narH	Nitrate réductase	non identifié	+	Moreno *et al.*, 2001
ndk	Nucléoside diphosphate kinase	non identifié	-	Moreno *et al.*, 2001
nupCpdp	Métabolisme de la pyrimidine	-32 -29	-	Miwa *et al.*, 2000 ; Moreno *et al.*, 2001
odhAB	2-oxoglutarate déshydrogénase	-99 -86	-	Deutscher *et al.*, non publié ; Tobisch *et al.*,1999b
pag *pbp*	Antigène de protection	non identifié	-	Baillie *et al.*,1998
pgk	seconde partie de la glycolyse	-218 –213	+	Deutscher *et al.*, non publié ; Tobisch *et al.*,1999b
phrC	Régulateur de RapC	non identifié	-	Moreno *et al.*, 2001
phrK	Régulateur de RapK	non identifié	-	Moreno *et al.*, 2001
ppsB	Peptide synthétase	non identifié	-	Moreno *et al.*, 2001
pta	**Métabolisme de l'acétate**	**-99 -86**	+	Presecan-Siedel *et al.*, 1999 ; Shin *et al.*, 1999 ; Tobisch *et al.*,1999b
pyrB	Aspartate carbamoyl-transférase	non identifié	+	Moreno *et al.*, 2001
qcrA	Ménaquinol :cytochrome c oxidoréductase	non identifié	-	Moreno *et al.*, 2001
rbsR	Métabolisme du ribose	-33 -20	-	Miwa *et al.*, 2000 ; Moreno *et al.*, 2001
rbsK	Métabolisme du ribose	+251 +264	-	Miwa *et al.*, 2000 ; Moreno *et al.*, 2001
rbsA	Métabolisme du ribose	+347 +360	-	Miwa *et al.*, 2000 ; Moreno *et al.*, 2001
rocG	Glutamate déshydrogénase	non identifié	-	Belitsky *et al.*, communication personnelle
sacP	Métabolisme du sucrose	-150 -137	-	Deutscher *et al.*, non publié ; Moreno *et al.*, 2001
sdhC	Succinate déshydrogénase	-138 -125 +300 +313	-	Deutscher *et al.*, non publié Miwa *et al.*, 2000 ; Moreno *et al.*, 2001
spoVFB	Dipicolinate synthase	non identifié	-	Moreno *et al.*, 2001
spoVG	Sporulation	+50 +63	-	Deutscher *et al.*, non publié ; Moreno *et al.*, 2001
spoVS	Sporulation	non identifié	-	Moreno *et al.*, 2001
sucCD	Succinyl CoA synthétase	-50 -37	-	Deutscher *et al.*, non publié ; Tobisch *et al.*,1999b
sunT	Lantibiotic transporter	+1839 +1852	-	Miwa *et al.*, 2000 ; Moreno *et al.*, 2001
tdh/kbl	Catabolisme de la thréonine	+1464 +1477	-	Moreno *et al.*, 2001
treP	**Métabolisme du tréhalose**	**+319 +332**	-	Miwa *et al.*, 2000
uxaC/exu	**Métabolisme de l'hexuronate**	-122 -109 -44 -31 -38 +51 **+1177 +1190**	-	Deutscher *et al.*, non publié ; 96 Rivolta *et al.*, 1998 Miwa *et al.*, 2000 Miwa *et al.*, 2000
xyn	**Catabolisme des xylosides**	**-98 -85**	-	Galinier et al, 1999
xylA	**Métabolisme du xylose**	**+35 +48**	-	Jacob *et al.*, 1991 ; Kraus *et al.*,1994 ; Kunst *et al.*, 1997
ydbH	Transport des acides dicarboxyliques	-63 -50	-	Miwa *et al.*, 2000 ; Blencke *et al.*, 2003
yebD	inconnue	-172 -159	+	Deutscher *et al.*, non publié ; Moreno *et al.*, 2001
yfmQ	inconnue	non identifié	+	Moreno *et al.*, 2001
yhbI	inconnue	non identifié	-	Moreno *et al.*, 2001
yobO	Protéine de phage	+1178 +1191	-	Miwa *et al.*, 2000
yoxA	inconnue	+500 +513	-	Moreno *et al.*, 2001
yvfL/lacA	Métabolisme du lactose	+1380 +1393	-	Moreno *et al.*, 2001
yxjC	**Métabolisme du β-hydroxybutyrate**	**-40 -27**	-	Miwa *et al.*, 2000

Table 3. Gènes pour lesquels un site *cre* a été identifié ou dont l'expression est sensible à une mutation de *ccpA* ou de *ptsH1* (*crh1*). [a] Les lignes en gras indiquent que la fonctionnalité du site *cre* a été expérimentalement démontrée. [b] La position du site cre est indiquée en fonction du premier nucléotide du premier codon du gène. [c] + et - indiquent que le gène correspondant est soumis à une activation ou une répression catabolique, respectivement.
D'après Deutscher *et al.*, 2002.

La position du site *cre* est déterminante dans le mécanisme moléculaire de répression ou d'activation transcriptionnelle (Table 2). S'il se trouve au niveau du promoteur, l'interaction avec CcpA va inhiber l'initiation de la transcription (Nicholson *et al.*, 1987; Kruger & Hecker, 1995). S'il est situé en aval du promoteur, voire dans le début de la phase codante du gène, le complexe *cre*-CcpA bloquera l'élongation (*acsA*, *hutP*) (Wray *et al.*, 1994; Zalieckas *et al.*, 1998a). S'agissant de la répression de l'expression des opérons *hut* et *gnt*, Zalieckas et ses collègues ont montré que le facteur Mfd est nécessaire à la répression catabolique (Zalieckas *et al.*, 1998b). Cette observation est en accord avec un mécanisme de blocage de la transcription ou « road-block ». En effet, le facteur Mfd aide au détachement de l'ARN polymérase lors de la réparation de l'ADN. Par contre, le facteur Mfd n'est apparemment pas impliqué dans la répression catabolique lorsque le site *cre* se trouve au niveau du promoteur, comme c'est le cas pour les gènes *amyE* et *bglPH* (Zalieckas *et al.*, 1998b). Par ailleurs, la répression de l'expression de l'opéron lévanase illustre un cas particulier (Martin-Verstraete *et al.*, 1995). CcpA va se fixer sur l'opérateur *cre* situé entre le site de fixation de l'activateur transcriptionnel LevR et la région promotrice de l'opéron, et prévient ainsi l'interaction entre LevR et l'ARN polymérase. Au passage, on peut aussi remarquer que du fait de multiples mécanismes de régulation, l'effet de la mutation du gène *ccpA* ou du site *cre* actif n'entraîne pas toujours un effet de « dérépression catabolique » très prononcé voire aucun effet. Par exemple, dans le contrôle du régulon *glp* (métabolisme du glycérol), l'effet de la mutation de *ccpA* (ou de *ptsH*) est masquée par les systèmes spécifiques d'exclusion de l'inducteur et de contrôle par antiterminaison (Deutscher *et al.*, 1994; Darbon *et al.*, 2002).

Par contre, si le site *cr*e est placé en amont de la région promotrice (et ne gène pas l'activation par un autre régulateur), la fixation de CcpA provoque une activation de la transcription (Table 2). En présence d'un excès de glucose ou d'un autre sucre rapidement métabolisable, la glycolyse produit un excès de pyruvate par rapport aux besoins anaboliques. Dans ces conditions, une partie du pyruvate peut être transformée en acétate et en acétoine. L'expression des gènes qui codent ces réactions, dites « d'overflow », est alors activée, phénomène appelé activation catabolique. Il s'agit des gènes *ackA* (acétate kinase) (Turinsky *et al.*, 1998), *pta* (phosphotransacétylase) (Presecan-Siedel *et al.*, 1999; Shin *et al.*, 1999) et *alsS* (acétolactate synthase) (Renna *et al.*, 1993). Les deux produits secondaires sont alors exportés dans le milieu extérieur. Ceci se traduit par une acidification du milieu. Lorsque, par la suite, la source de carbone préférentielle est épuisée, l'acétate et l'acétoine peuvent être réimportés et utilisés. Mais tant que du glucose se trouve dans le milieu de croissance et est

utilisé, l'expression des gènes du métabolisme de l'acétate et de l'acétoine, *acsA* (acétyl-CoA synthétase), *acuA* (acétoine déshydrogénase) et *acoA* (sous-unité acétoine déshydrogénase), est réprimée par CcpA.

Dans certains cas, il peut y avoir plusieurs sites *cre* adjacents au niveau de la région régulatrice d'un opéron. C'est le cas des opérons *gnt* et *xyl* (Gosseringer *et al.*, 1997; Miwa *et al.*, 1997) (Table 3). Dans le cas de l'opéron *xyl*, les auteurs ont montré que CcpA pouvait interagir avec trois sites *cre* : le site *cre1* situé dans le premier gène de l'opéron *xyl* et impliqué dans la répression catabolique par CcpA et deux autres sites *cre*, l'un chevauchant la séquence –35 et l'autre situé en aval du premier *cre* cité. Les séquences de ces trois *cre* diffèrent légèrement, celui décrit par Hueck et ses collaborateurs étant le plus ressemblant à la séquence consensus chez *Bacillus subtilis* (Hueck *et al.*, 1994). Cela explique l'affinité différentielle de CcpA pour ses séquences. Une mutation dans *cre1* entraîne la perte de la régulation par CcpA *in vivo* et la perte de d'interaction avec les trois sites cre *in vitro*, suggérant une dépendance physique entre ses séquences.

L'étude de la répression catabolique du gène *amyE* a aussi permis d'identifier CcpA (Catabolite control protein A), le régulateur transcriptionnel qui reconnaît la séquence *cre*. En effet, l'interruption du gène *ccpA* rend l'expression du gène *amyE* insensible au glucose (Henkin *et al.*, 1991). Par la suite, de nombreuses études ont confirmé le rôle du gène *ccpA* dans la répression catabolique (*acsA* et *acuABC*) (Grundy *et al.*, 1994) (*gnt*) (Fujita & Miwa, 1994).

Le rôle de CcpA a été démontré chez de nombreuses bactéries à Gram positif et faible pourcentage en GC (Kuster *et al.*, 1996), notamment *Bacillus megaterium* (Hueck *et al.*, 1994; Hueck *et al.*, 1995), *Lactobacillus casei* (Monedero *et al.*, 1997), *Lactococcus lactis* (Luesink *et al.*, 1998) et *Enterococcus faecalis* (Leboeuf *et al.*, 2000a; Leboeuf *et al.*, 2000b). Cependant, chez *Streptococcus mutans*, la protéine RegM, homologue de CcpA, ne semble pas fonctionner comme ce dernier régulateur. En effet, une délétion du gène *regM* ne conduit pas à l'abolition de la répression catabolique, bien au contraire (Simpson & Russell, 1998). Par ailleurs, chez *Streptococcus pneumoniae*, outre son implication dans la répression de l'expression de l'α-glucosidase et de la β-galactosidase, RegM semble impliqué dans la régulation de la biosynthèse des polysaccharides capsulaires, ce qui suggère un lien entre cette protéine et la virulence de cette bactérie (Giammarinaro & Paton, 2002).

CcpA de *Bacillus subtilis* a été surexprimée et purifiée (Miwa *et al.*, 1994; Kim *et al.*, 1995). Il s'agit d'une protéine de 38 kDa, active sous la forme d'un dimère. Les homologies de séquence protéique montrent qu'elle appartient à la famille des régulateurs transcriptionnels LacI/GalR (25 % d'identité avec la séquence de LacI et 31 % avec celle de GalR). En effet, CcpA, LacI et GalR partagent des séquences aminoterminales très proches, correspondant au domaine hélice-tour-hélice de fixation à l'ADN (Henkin *et al.*, 1991). La partie C-terminale de CcpA est constituée de deux sous-domaines permettant l'intégration du signal donné par le métabolisme (Jones *et al.*, 1997; Kraus *et al.*, 1998).

Par ailleurs, les études concernant l'expression génique indiquent que CcpA est synthétisée de façon constitutive, notamment en présence ou en absence de glucose. Ainsi, il semble que ce soit l'activité biochimique de CcpA qui est modulée, permettant ou pas l'interaction avec son opérateur *cre* (Miwa *et al.*, 1994; Hueck & Hillen, 1995). Cependant, il existe chez *Lactococcus lactis* (Luesink *et al.*, 1999) et *Lactococcus pentosus* (Mahr *et al.*, 2000) un site *cre* dans la région régulatrice de l'expression de *ccpA* qui pourrait permettre son autorégulation.

Chez *Bacillus subtilis*, la mutation *ptsH1*, qui permet d'exprimer la protéine HPr-Ser46Ala, provoque une levée plus ou moins forte de la répression catabolique par le glucose de l'expression d'un certain nombre d'enzymes (gluconate kinase, glucitol déshydrogénase, mannitol-1-P déshydrogénase et mannitol PTS perméase, inositol déshydrogénase) (Deutscher *et al.*, 1994). Le fait que ce mutant présente un phénotype similaire à celui d'un mutant *ccpA* suggère que HPr-Ser46-P est l'effecteur qui module l'activité de fixation à l'ADN de CcpA. Aussi, l'année suivante, l'interaction directe entre CcpA et HPr-Ser-46-P a été démontrée *in vitro* par des expériences de retard d'élution sur colonne (Deutscher *et al.*, 1995). Ces résultats ont été confirmés par la suite par des expériences employant les techniques de résonance magnétique nucléaire et de dichroïsme circulaire (Jones *et al.*, 1997). La spécificité de cette interaction a aussi été testée par retard d'élution sur colonne. En effet, alors qu'elle retarde l'élution d'HPr-Ser46-P, la présence de CcpA ne retarde pas celles d'HPr, d'HPr-His15-P et de HPr doublement phosphorylée sur l'histidine 15 et la sérine 46. Cette découverte renforce le lien entre la répression catabolique directement médiée par CcpA et une forte activité glycolytique puisque la forme HPr-Ser46-P est la conséquence de l'activité kinase d'HPrK/P induite par la forte concentration en FBP.

Mais comment cette interaction module-t-elle l'activité de CcpA ? Il avait été observé que CcpA était capable de se lier à l'ADN mais de manière aspécifique (Miwa *et al.*, 1994). Par la suite, les équipes de J. Deutscher et Y. Fujita ont montré que le complexe CcpA/HPr-Ser46-P était capable de reconnaître spécifiquement la séquence ADN *cre* (Deutscher *et al.*, 1995; Fujita *et al.*, 1995). En effet, Fujita et ses collègues ont montré par des expériences d'empreinte à la DNase I que CcpA protégeait la séquence *cre* situé dans la phase codante de *gntR*, le premier gène de l'opéron *gnt*, uniquement en présence de HPr-Ser46-P (Fujita *et al.*, 1995). La spécificité de cette interaction a été confirmée par mutagenèse dirigée. La substitution d'une seule base dans la séquence *cre* suffit à perturber la fixation de CcpA (Fujita *et al.*, 1995). Cette interaction a ensuite été caractérisée de manière plus détaillée (Jones *et al.*, 1997). Le domaine d'interaction de CcpA avec HPr-Ser46-P a été déterminé. Ce sont les deux hélices α, A et B, d'HPr qui interagissent avec le domaine C-terminal de CcpA. En effet, l'hélice A comprend l'histidine 15 et les résidus voisins du site actif et l'hélice B inclut la sérine 46, permettant ainsi la reconnaissance spécifique d'HPr-Ser46-P par CcpA. Reizer et ses collaborateurs avaient observé qu'une mutation de l'histidine 15 d'HPr en une alanine annulait la répression catabolique (Reizer *et al.*, 1996). Cette mutation diminue probablement la spécificité de reconnaissance de l'interaction de CcpA vis-à-vis d'HPr. Il a été montré que l'interaction entre les hélices α d'HPr et la partie C-terminale de CcpA ne modifiait pas la conformation des deux protéines (Jones *et al.*, 1997). Des expériences de dichroïsme circulaire montrent aussi que le complexe est formé de deux protéines HPr-Ser46-P et d'un dimère de CcpA (Jones *et al.*, 1997). L'alignement de séquences de CcpA avec les protéines très homologues et les protéines de la famille LacI/GalR révèle une séquence conservée spécifique de CcpA qui est impliquée dans la reconnaissance d'HPr-Ser46-P (Kraus *et al.*, 1998). Cette région est chargée positivement (Kraus *et al.*, 1998). CcpA et HPr-Ser46-P sont donc probablement liés par des forces électrostatiques puisque la phosphorylation de la sérine 46 d'HPr rend cette région négativement chargée. En conséquence, la formation du complexe CcpA/HPr-Ser46-P a pour résultat l'augmentation de l'affinité du domaine N-terminal de CcpA pour le motif *cre*.

Il existe d'autres cofacteurs de CcpA. En effet, dans certains cas, il a été observé que la présence de FBP stimulait la formation du complexe CcpA/HPr-Ser46-P (Deutscher *et al.*, 1995; Kim *et al.*, 1998; Galinier *et al.*, 1999). Cependant, le rôle du FBP n'a pu être mis en évidence dans le cas des opérons *lev* et *gnt* (Martin-Verstraete *et al.*, 1999). S'agissant du gène *amyE*, Kim et ses collaborateurs ont montré qu'en présence de FBP uniquement,

l'affinité de CcpA pour sa cible était stimulée d'un facteur 2 alors que HPr-Ser46-P seul élève d'un facteur 10 cette affinité. Mais lorsque l'on combine HPr-Ser46-P au FBP, l'affinité de CcpA est stimulée d'un facteur 300 (Kim *et al.*, 1998). Toutefois, ni l'ajout de FBP ni le complexe CcpA/HPr-Ser46-P ne semblent avoir d'effet sur le niveau de répression de l'expression du gène *amyE* dans des expériences de transcription *in vitro* (Voskuil & Chambliss, 1996; Kim *et al.*, 1998). Par ailleurs, Kim et ses collaborateurs ont montré que l'ajout de NADP, réduit ou oxydé, stimulait d'un facteur 1,5 et 120, respectivement en absence et en présence de HPr-Ser46-P, l'affinité de CcpA pour la séquence *cre*. En outre, l'ajout de NADP stimule spécifiquement l'inhibition de la transcription du gène *amyE* d'un facteur 120 (Kim *et al.*, 1998). Ainsi, le NADP a un effet significatif sur la capacité de CcpA à lier la séquence *cre* mais un effet beaucoup plus important sur l'inhibition de la transcription. Ces auteurs suggèrent que la répression de l'expression du gène *amyE* n'est pas directement liée à la fixation de CcpA sur le site *cre* mais à la stimulation de l'interaction entre CcpA et l'ARN polymérase en présence de NADP(H), comme cela a déjà été proposé auparavant pour le régulateur GalR de l'opéron *gal* (métabolisme du galactose) d'*Escherichia coli* (Chatterjee *et al.*, 1997; Choy *et al.*, 1997). Cependant, il convient de se demander quel signification physiologique le rôle du NAPD(H) supporte.

Par ailleurs, il a été montré que CcpA était capable de lier la séquence *cre* selon deux modes. Au niveau du promoteur de l'opéron *gnt* de *Bacillus subtilis*, il existe un deuxième site *cre* auquel CcpA est capable de se fixer uniquement en présence d'une grande concentration de G6P (Miwa *et al.*, 1997). L'effet du G6P a aussi été observé dans le cas de l'opéron *xyl* de *Bacillus megaterium* (Gosseringer *et al.*, 1997). La fixation de CcpA en présence d'HPr-Ser46-P a lieu de façon non-coopérative. En revanche, elle a lieu coopérativement en présence de G6P mais sur d'autres sites *cre*. Cela suggère l'existence de deux formes de répression catabolique dans le cas de ces opérons. Mais ces résultats sont à considérer avec précaution. Les conditions expérimentales utilisées n'ont sans doute pas de véritable signification physiologique ([G6P]>30 mM et pH acide).

Alors qu'HPr-Ser46-P semblait être le cofacteur principal de CcpA, la persistance du phénomène de répression catabolique des opérons *hut* (Zalieckas *et al.*, 1999), *iol* (Miwa *et al.*, 2000), *lev* (Martin-Verstraete *et al.*, 1995) et *xyn* (Galinier *et al.*, 1999) dans un mutant *ptsH1* mais pas dans un mutant *ccpA* restait à expliquer.

Il a alors été montré que Crh, une protéine homologue à HPr (cf partie 2.2.2.2.1), joue aussi un rôle régulateur *via* sa phosphorylation sur la sérine 46 par l'HPrK/P (Galinier *et al.*, 1997; Turinsky *et al.*, 1998; Zalieckas *et al.*, 1998b; Zalieckas *et al.*, 1998a; Galinier *et al.*, 1999; Martin-Verstraete *et al.*, 1999; Presecan-Siedel *et al.*, 1999). Dans ces cas, la délétion du gène *crh* n'affecte pas la régulation médiée par CcpA. Mais la régulation résiduelle qui persiste dans un mutant *ptsH1* est complètement abolie lorsqu'une mutation de *crh* est ajoutée. De façon surprenante, le gène *crh* est uniquement présent chez les Bacillacées proches de *Bacillus subtilis*, à savoir *Bacillus anthracis*, *Bacillus halodurans* et *Bacillus stearothermophilus*. Pourtant, l'opéron *yvc* dans lequel se trouve *crh* existe chez d'autres bactéries que ces Bacilli. La protéine Crh joue-t-elle un rôle spécifique chez *Bacillus subtilis* et ses proches cousines ?

En résumé, en présence d'un sucre rapidement métabolisable comme le glucose, l'augmentation du pool de FBP stimule l'activité kinase de HPrK/P. Cette enzyme va phosphoryler la protéine HPr (et Crh) sur la sérine 46. HPr-Ser46-P (et Crh-Ser46-P) est alors capable d'interagir avec le régulateur transcriptionnel CcpA. Ce complexe se fixe sur une séquence spécifique située dans la région régulatrice des gènes cibles. S'ensuit la répression (ou l'activation) de l'expression de ces gènes (Figure 26). La protéine HPr joue donc un rôle double de façon alternative dans la régulation du transport par le système PTS (HPr-His15-P) et dans la régulation de l'activité de CcpA (HPr-Ser46-P). La balance entre ces deux fonctions est régulée par l'activité de l'HPrK/P qui est donc le senseur central de l'état métabolique et énergétique de la bactérie pour son adaptation physiologique. Selon une estimation statistique et plusieurs analyses transcriptomiques (Moreno *et al.*, 2001; Yoshida *et al.*, 2001), près de 10 % des gènes de *Bacillus subtilis* pourraient être soumis à la répression/activation catabolique *sensu stricto* (Deutscher *et al.*, 2002) (Table 3). Parmi ces gènes, un très grand nombre appartient au métabolisme du carbone (transport et métabolisme intracellulaire) et au système membranaire de production d'énergie étroitement lié (Table 3). Mais CcpA médierait aussi la régulation d'autres fonctions comme la sporulation, la chémotaxie, le métabolisme des acides nucléiques et des coenzymes (Moreno *et al.*, 2001) (Table 3).

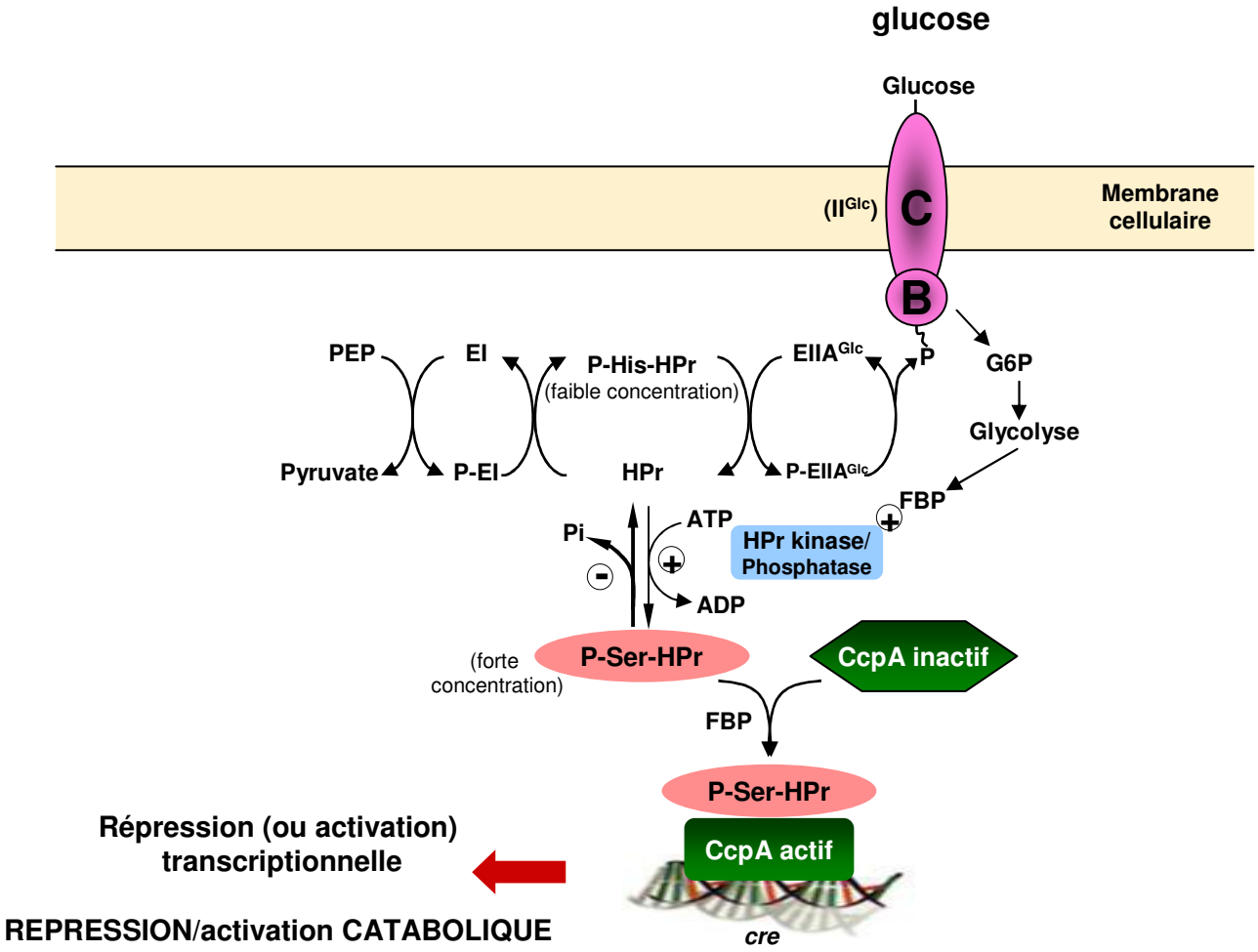


Figure 26. Mécanisme de la répression/activation catabolique médiée par le régulateur transcriptionnel pléïotropique CcpA chez *Bacillus subtilis*.

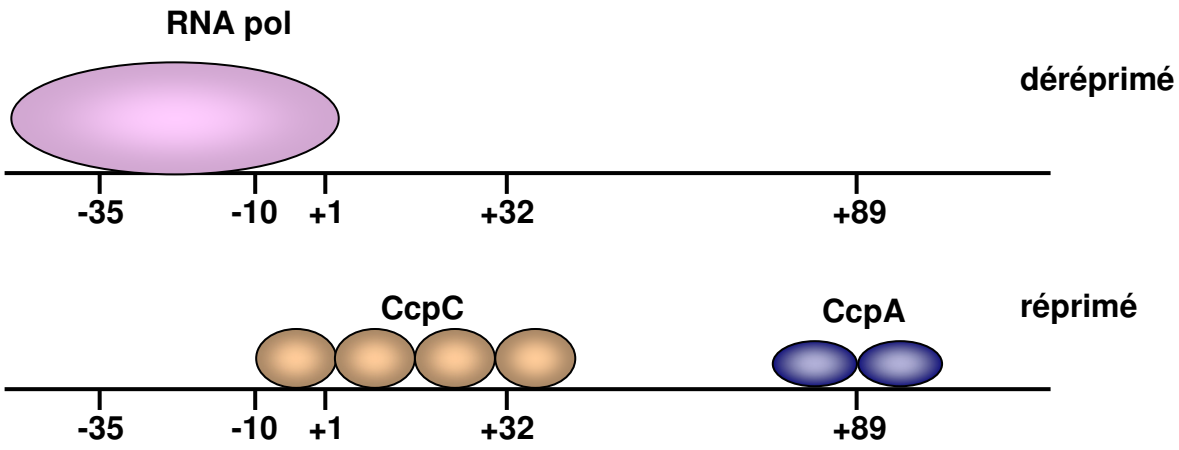
D'après Deutscher *et al.*, 2002.

A côté de ces acteurs principaux, deux autres régulateurs en *trans* ont été décrits qui sont impliqués dans la répression catabolique.

Le premier d'entre eux est codé par le gène *ccpB* (catabolite control protein B). Le rôle partiel de CcpB a été observé dans la régulation des opérons *gnt* et *xyl* de *Bacillus subtilis* lors de la croissance en milieu solide en présence de glucose, de mannitol et de sucrose (Chauvaux *et al.*, 1998). En phase liquide, la levée de la répression catabolique par une interruption de *ccpB* n'est perceptible que dans des conditions particulières de faible agitation conduisant à une faible aération du milieu. CcpB est une protéine synthétisée de façon constitutive qui possède 30 % d'identité avec CcpA. Mais les séquences CcpA-spécifiques ne sont pas conservées ce qui suggère que CcpB n'interagit pas avec HPr ou Crh. L'ensemble de ces éléments suggère que CcpB répond plutôt à un signal de type physico-chimique (oxygénation du milieu, densité cellulaire...).

Le deuxième nouvel acteur en *trans* est codé par le gène *ccpC* (catabolite control protein C). Il a été identifié lors de l'étude de la régulation des gènes codant les enzymes du cycle de Krebs (Jourlin-Castelli *et al.*, 2000; Kim *et al.*, 2002b; Kim *et al.*, 2002a). CcpC appartient à la famille de régulateur transcriptionnel LysR. Il a été montré que CcpC réprimait la transcription des gènes *citZ* (citrate synthase) et *citB* (aconitase) en réponse à la faible disponibilité en citrate du milieu de culture (Jourlin-Castelli *et al.*, 2000). Une mutation de *ccpC* entraîne une levée de la répression de l'expression de *citB* d'un facteur 18 et 2, respectivement, en présence de glucose ou de citrate comme seule source de carbone. CcpC est donc requis pour la répression de l'expression de *citB* de façon dépendante du glucose. Cette mutation entraîne aussi une augmentation d'un facteur 2 de l'expression de *citZ* en présence de glucose mais pas de citrate, ce qui suggère que CcpC est un régulateur négatif de l'expression de *citZ*. En outre, l'expression de *citZ* est soumise à la répression catabolique médiée par CcpA. Au niveau moléculaire, CcpC est capable de lier l'ADN au niveau de deux séquences spécifiques (Figure 27). Mais en présence de citrate, CcpC n'interagit plus qu'avec la première séquence permettant ainsi une modulation de l'expression de *citB* (Figure 27). Les mêmes auteurs ont aussi montré que CcpC était impliqué dans la répression de l'expression des gènes *citB* et *citZ* lors de la croissance anaérobie, condition de croissance où le citrate est vraisemblablement en faible quantité dans la cellule (Nakano *et al.*, 1998; Jourlin-Castelli *et al.*, 2000). Très récemment, la régulation de l'expression de *ccpC* a été étudiée (Kim *et al.*, 2002b). Ces études révèlent une autorégulation de l'expression de *ccpC* par CcpC et une

A. *citZ*



B. *citB*

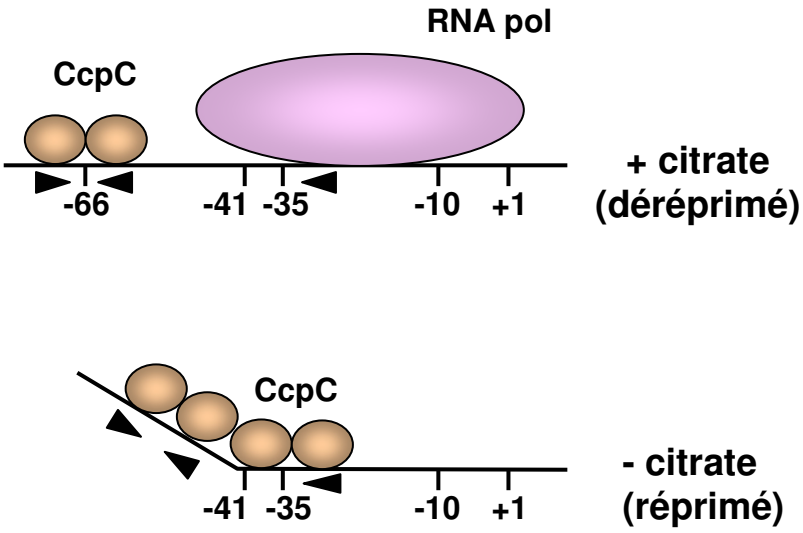


Figure 27. Modèle de régulation de l'expression des gènes *citZ* et *citB* par CcpC en présence et en absence de citrate. Les positions sont indiquées en fonction du site d'initiation de la transcription.

A. L'expression de *citZ* est directement réprimée par CcpA et CcpC.

B. Le promoteur de *citB* est directement réprimé par CcpC et indirectement par CcpA. En présence de citrate, CcpC ne lie que les deux séquences axées en -66 (flèches noires) et l'ARN polymérase peut accéder au promoteur. En l'absence de citrate, CcpC a une plus forte affinité pour l'ADN et interagit avec un motif supplémentaire. Cette interaction aurait pour conséquence une courbure de l'ADN qui empêcherait l'ARN polymérase de reconnaître le promoteur.

D'après Sonenshein, 2002.

régulation par CcpA. L'initiation de la transcription a en fait lieu au niveau de deux promoteurs différents. L'activité du premier promoteur P1, situé juste en amont de la phase codante de *ccpC* est négativement contrôlée par CcpC, indépendamment de la source de carbone disponible dans le milieu. CcpC interagit avec une séquence similaire à celle qu'il reconnaît dans les régions régulatrices de *citB* et de *citZ*. L'activité du promoteur P2, devant le gène amont *ykuL* (gène dont la fonction est encore inconnue), est contrôlée par CcpA en fonction de la présence ou de l'absence de glucose. En effet, il existe une boîte *cre* fonctionnelle juste en amont de la séquence –35 de ce deuxième promoteur. Ainsi, en présence de glucose, CcpA agit comme un répresseur de la transcription de *ccpC*. En conclusion, la Figure 28 illustre la régulation complexe des gènes *citB* et *citZ* par CcpA et CcpC.

2.3.2.2. Régulation CcpA-dépendante des gènes des voies centrales du métabolisme

Nous avons vu à travers la description du métabolisme général de la répression catabolique médiée par CcpA que celui-ci régulait négativement l'expression d'un très grand nombre de gènes impliqués dans le métabolisme de sources de carbone non préférentielles en présence d'un sucre rapidement métabolisable comme le glucose. De cette manière, cette régulation conditionne fortement l'entrée du carbone dans les voies centrales du métabolisme.

En présence de glucose, la machinerie de la répression catabolique va réprimer l'utilisation des autres sources de carbone. Mais elle va aussi diriger le flux de carbone entre la glycolyse, la voie des pentoses phosphate et le cycle de Krebs en contrôlant l'expression des gènes de ces trois voies. En effet, Tobisch et ses collaborateurs ont montré l'importance de CcpA dans la régulation du métabolisme carboné chez *Bacillus subtilis* par des comparaisons protéomiques et des analyses transcriptionnelles (Tobisch *et al.*, 1999b). La mutation de *ccpA* entraîne une altération sévère de la croissance en présence de glucose comme seule source de carbone. Ce phénotype est due à quatre déficiences qui sont liées. Premièrement, Tobisch et ses collaborateurs ont montré que la protéine CcpA était nécessaire à la stimulation de la synthèse des enzymes de la partie centrale de la glycolyse, *gapA*, *pgk*, *tpi*, *pgm* et *eno* (Tobisch *et al.*, 1999b). Cependant, ils n'ont pas observé d'induction par le glucose de la synthèse des gènes de la partie supérieure de la glycolyse, *pgi*, *pfk* et *fbaA*, de

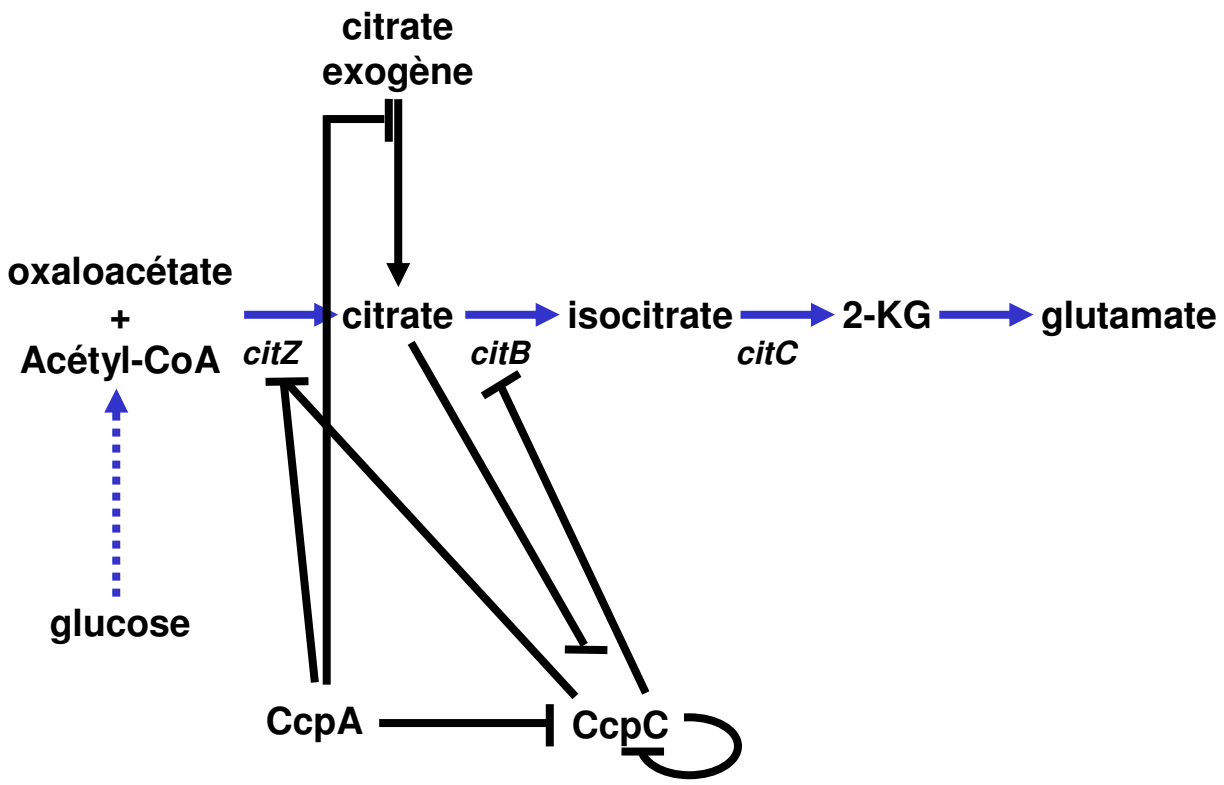


Figure 28. Modèle de régulation des gènes *citZ, citB* et *ccpC* par CcpA et CcpC chez *Bacillus subtilis*. CcpA contrôle l'activité de CcpC en régulant la disponibilité en citrate intracellulaire, un effecteur de CcpC.

D'après Kim *et al.*, 2002.

pykA et de *pdhABCD* contrairement à ce qui avait été suggéré par d'autres études (Moreno *et al.*, 2001; Yoshida *et al.*, 2001). Deuxièmement, Stülke et ses collègues ont montré que dans un mutant *ccpA*, le transport des sucres PTS était fortement perturbé par la titration d'HPr-His15-P (Ludwig *et al.*, 2002b). En effet, en l'absence de CcpA, l'activité kinase d'HPrK/P est telle que HPr est très majoritairement phosphorylée sur la sérine en position 46. Par conséquent, la chaîne de transfert de phosphate du système PTS, du PEP vers l'enzyme IIA, est bloquée et le sucre inducteur de l'expression des gènes codant les enzymes de son catabolisme ne peut pas être perméé (cf partie 2.2). Dans ces conditions, le glucose n'est pas perméé efficacement. Par conséquent, la croissance est très ralentie. Ainsi, en présence de glucose, CcpA est requis pour une activité glycolytique optimale. Troisièmement, il a été montré que CcpA est nécessaire à l'assimilation de l'ammonium comme seule source d'azote (Faires *et al.*, 1999). En effet, CcpA est requis pour l'expression de l'opéron *gltAB* qui code la glutamate synthase. Cette enzyme est indispensable pour l'assimilation de l'ammonium (Figure 7). Enfin, quatrièmement, Ludwig et ses collègues ont montré que la présence de méthionine et des acides aminés branchés permettait une meilleure croissance dans un milieu minimum. En effet, CcpA est nécessaire à l'expression de l'opéron *ilv-leu* impliqué dans la biosynthèse des acides aminés branchés (Ludwig *et al.*, 2002a).

En revanche, en présence de glucose comme source de carbone et de glutamate comme source d'azote, CcpA semble réprimer le fonctionnement du cycle de Krebs. Il a été observé que l'expression de *citH* (malate déshydrogénase), *citC* (isocitrate déshydrogénase), *citB* (aconitase), *odhA* (sous-unité de la 2-oxoglutarate déshydrogénase), *sucCD* (succinyl Coenzyme A synthétase) et *citG* (fumarase) était réprimée par le glucose (Tobisch *et al.*, 1999b; Moreno *et al.*, 2001; Yoshida *et al.*, 2001). Plus récemment, l'équipe de Sonenshein a montré que CcpA, outre l'effet indirect qu'il exerce *via* la régulation de l'expression de *ccpC*, régule aussi directement l'expression des gènes *citZ* (citrate synthase) et *citB* (aconitase) (Kim *et al.*, 2002b; Kim *et al.*, 2002a). Ainsi, l'expression des gènes du cycle de Krebs est très finement régulée par au moins deux régulateurs directement ou indirectement (Figure 28), CcpA et CcpC, tous deux sensibles à la source de carbone disponible dans le milieu de croissance.

Ainsi, en présence de glucose et de glutamate, la régulation catabolique médiée par CcpA stimule la glycolyse et réprime le cycle de Krebs. Physiologiquement, en effet, lorsque *Bacillus subtilis* pousse dans un milieu contenant du glucose comme source de carbone et une

source d'azote, la glycolyse fournit les principaux intermédiaires métaboliques carbonés et une partie suffisante d'énergie (la respiration fournissant de l'ATP par phosphorylation oxydative). Dans ces conditions, le cycle de Krebs ne fonctionne pas pour fournir de l'ATP mais les précurseurs anaboliques dont il a « l'exclusivité ». Le pyruvate et l'acétyl-CoA produits qui ne rentrent pas efficacement dans le cycle de Krebs sont dérivés en acétate, en acétoine et en lactate puis excrétés.

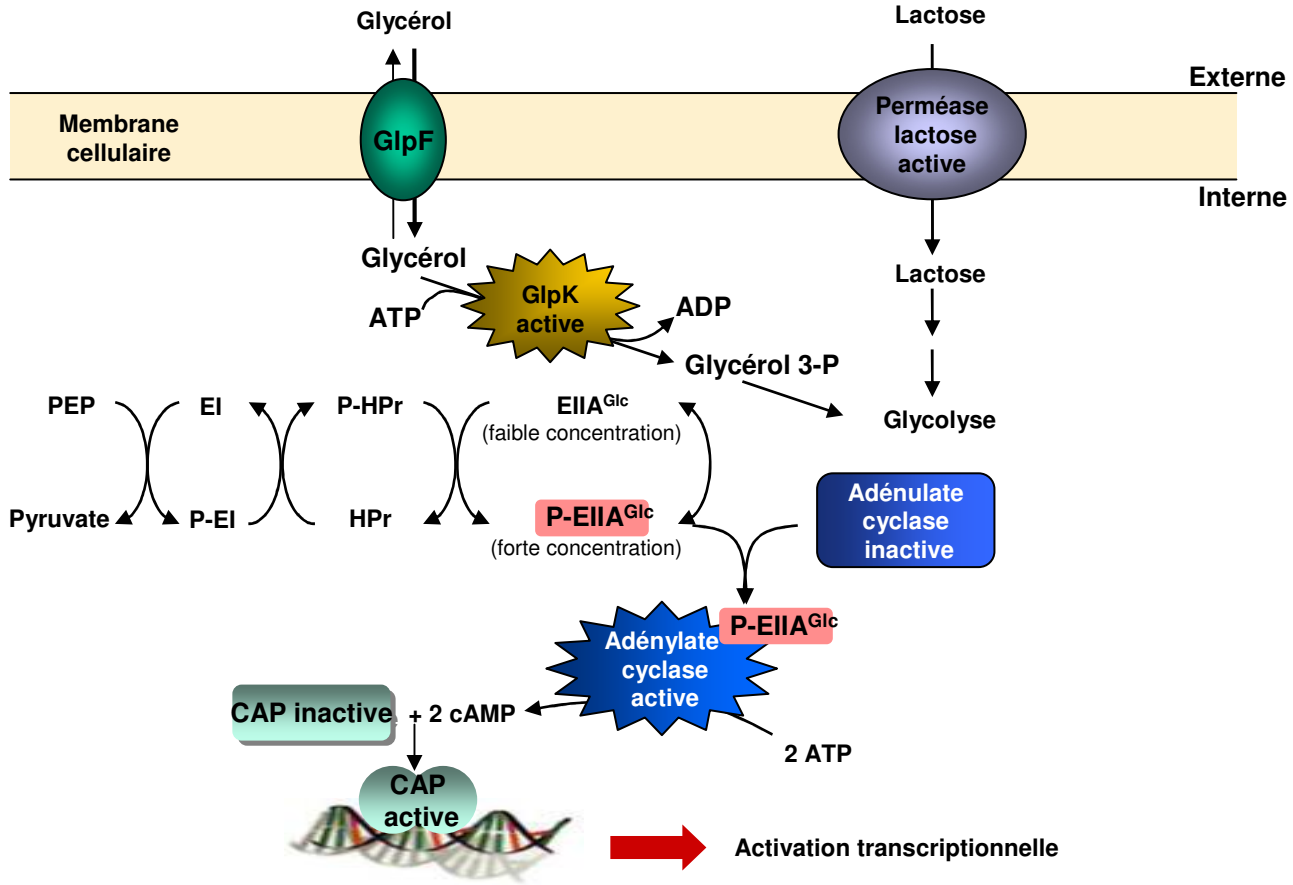
2.3.2.3. Brève comparaison avec la répression catabolique chez les bactéries à Gram négatif

Chez les bactéries à Gram négatif, dont le modèle est *Escherichia coli*, les mécanismes de répression catabolique ont été caractérisés bien avant ceux modulant le même phénomène chez *Bacillus subtilis*. Cependant, contrairement au cas des bactéries à Gram positif et faible pourcentage en GC, où un mécanisme majeur et pléïotropique prédomine, impliquant les protéines CcpA, HPr et HPrK/P, les bactéries à Gram négatif ont développé plusieurs systèmes dont il est plus difficile de quantifier l'importance relative (Stulke & Hillen, 1999).

Chez *Escherichia coli*, le premier mécanisme pléïotropique de la répression/activation catabolique identifié implique le couple Enzyme IIA^{Glc} ($EIIA^{Glc}$) et le régulateur transcriptionnel CAP (Catabolite Activator Protein), aussi appelé CRP (cAMP Receptor Protein) parce que son activité est modulée par l'AMP cyclique (AMPc) (Postma *et al.*, 1993; Bruckner & Titgemeyer, 2002).

Lorsqu'*Escherichia coli* croît en l'absence d'une source de carbone rapidement métabolisable (Figure 29), le ratio PEP-pyruvate est élevé. Par conséquent la protéine $EIIA^{Glc}$ se trouve principalement sous sa forme phosphorylée. P-$EIIA^{Glc}$ a la propriété de stimuler l'activité de l'adénylate cyclase et/ou d'inhiber l'activité de l'AMPc phosphodiestérase (Postma *et al.*, 1993; Takahashi *et al.*, 1998). La forte concentration en AMPc cytoplasmique qui en découle favorise la formation du complexe CAP/AMPc, qui est alors capable d'interagir avec une séquence spécifique au niveau de la région régulatrice des gènes soumis à la répression catabolique et ainsi d'activer leur transcription. Mais, en présence de glucose ou d'un autre sucre rapidement métabolisable (Figure 29), le ratio PEP-pyruvate est faible et provoque la déphosphorylation d'$EIIA^{Glc}$. L'absence de stimulation de l'activité de l'adénylate cyclase résulte en une faible concentration en AMPc et donc l'absence de

A. *Escherichia coli* : absence de glucose



B. *Escherichia coli* : présence de glucose

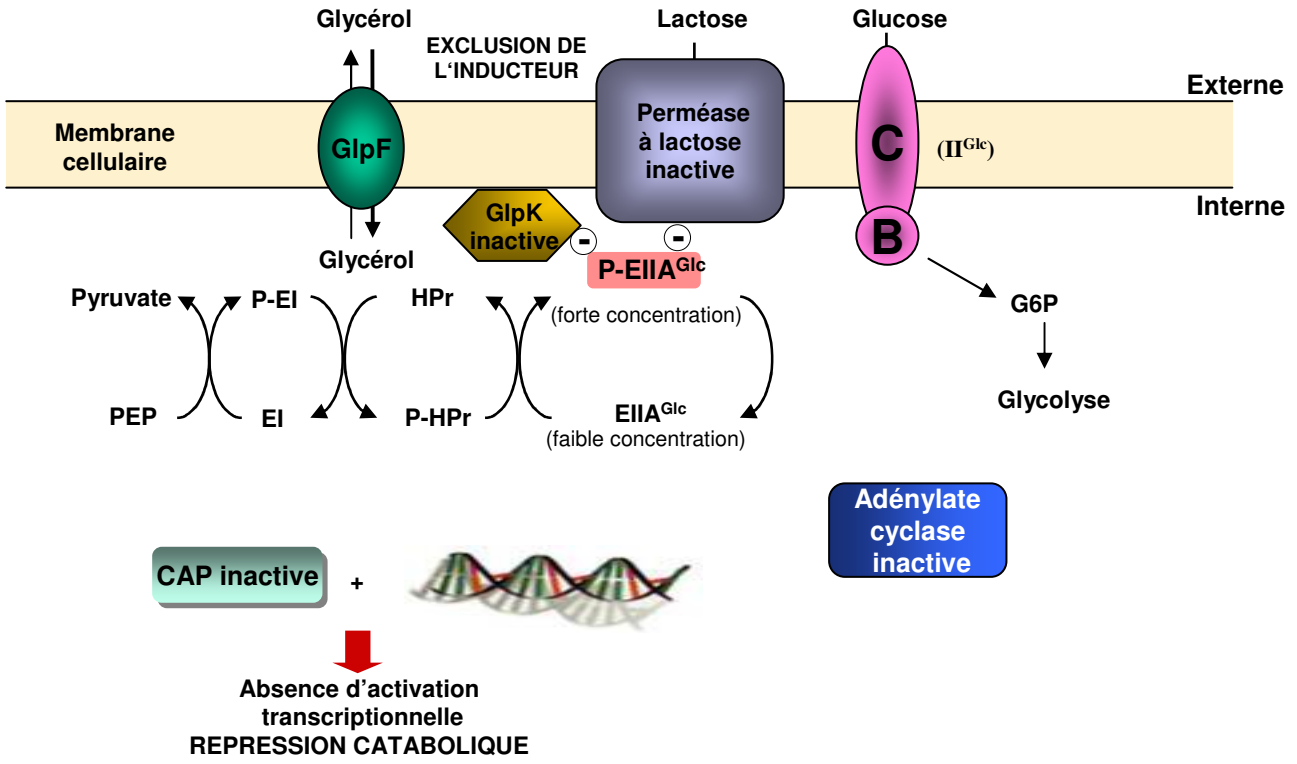


Figure 29. Rôles du PTS et de la protéine CAP dans la régulation du métabolisme carboné chez *Escherichia coli*.

D'après Deutscher *et al.*, 2002.

formation du complexe CAP/AMPc. Dans ces conditions, l'expression des gènes impliqués dans l'utilisation de sources de carbone non-préférentielles n'est pas activée. Par ailleurs, la forme non phosphorylée d'$EIIA^{Glc}$ est aussi impliquée dans les processus d'exclusion et d'expulsion d'inducteurs et dans la régulation de l'activité de la glycérol kinase chez les bactéries à Gram négatif (cf partie 2.2).

Chez *Bacillus subtilis* et les autres bactéries à Gram positif, les fonctions exercées par $EIIA^{Glc}$ chez *Escherichia coli* semblent prises en charge par la protéine HPr (cf parties 2.2) (Figure 30). En outre, ces bactéries à Gram positif ne possèdent pas de gène codant une adénylate cyclase, du moins de classe connue, pour produire l'AMPc (la présence de cette molécule n'a pas pu être détectée expérimentalement chez ces bactéries). L'HPrK/P « sent » la variation de concentration des métabolites de la glycolyse, essentiellement le FBP. Bien que contrôlant les mêmes phénomènes, les déterminants de la répression catabolique sont donc différents entre *Bacillus subtilis* et *Escherichia coli*. Dans le premier cas, il s'agit de contrôler l'inhibition de la transcription par modulation de l'activité d'un répresseur (« accessoirement » activateur) transcriptionnel. Dans le second cas, il s'agit de contrôler l'activation de la transcription par modulation de l'activité d'un activateur transcriptionnel.

Cependant, il existe des homologues d'HPrK/P chez quelques bactéries à Gram négatif : *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Bordetella pertussis*, *Treponema pallidum*, *Chlorobium tepidum* et *Xylella fastidiosa* (Deutscher *et al.*, 2002). Toutefois, chez *Neisseria meningitidis* et *Bordetella pertussis*, il n'y a ni d'enzymes EIIB et EIIC pour le transport PTS-dépendant ni d'homologue de CcpA pour la régulation transcriptionnelle. Par conséquent, chez ces bactéries, HPr-Ser-P n'intervient ni dans la régulation du transport par le système PTS ni dans la régulation catabolique dépendante de CcpA.

Par ailleurs, il existe un second mécanisme pléïotrope de répression catabolique qui est indépendant de l'AMPc chez *Escherichia coli* et d'autres bactéries à Gram négatif. Geerse et ses collègues ont d'abord montré que la protéine FruR réprimait directement l'expression de l'opéron *fru*, impliqué dans le métabolisme du fructose (Geerse *et al.*, 1986). Par la suite, le répresseur FruR a été renommé Cra pour Catabolite repressor/activator. Chin et ses collaborateurs ont montré qu'un mutant nul de *fruR* était incapable de pousser en présence des intermédiaires métaboliques du cycle de Krebs comme seule source de carbone (Chin *et al.*,

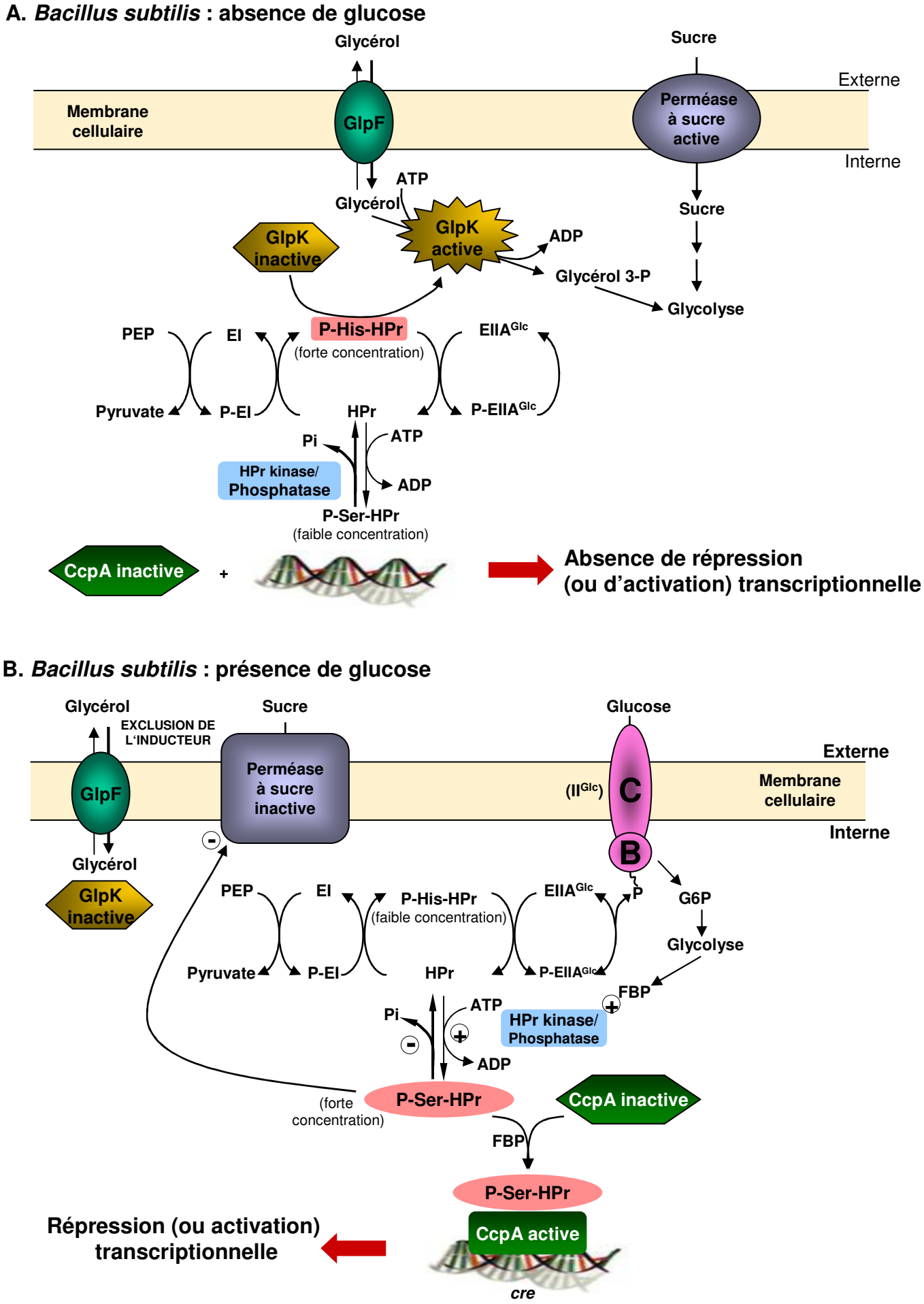


Figure 30. Rôles du PTS et de CcpA dans la régulation du métabolisme carboné chez *Bacillis subtilis*.

D'après Deutscher *et al.*, 2002.

1987). Une autre mutation de *fruR*, qui conduit à l'expression constitutive du régulon fructose par inactivation de Cra, réprime l'expression des enzymes de la néoglucogenèse et active l'expression des enzymes de la glycolyse, de l'enzyme I et des enzymes II du PTS (Chin *et al.*, 1989), reproduisant ainsi l'effet de la présence de glucose dans le milieu. La Table 4 dresse la liste des gènes dont l'expression est régulée, soit positivement, soit négativement, par le régulateur Cra (Saier & Ramseier, 1996).

Sur le plan moléculaire, la manière dont la protéine Cra contrôle la transcription est relativement similaire à celle évoquée précédemment et qui concerne CcpA (Figure 31). Cra reconnaît une séquence pseudo-palindromique asymétrique : RSTGAAWC │ SNTHHW (le symbole │ représentant le centre du palindrome ; R correspond à A ou G, S à C ou G, W à A ou T, N à n'importe quel nucléotide et H à A ou C) (Saier & Ramseier, 1996). La position de cet opérateur dans la région régulatrice des gènes concernés détermine le sens de la régulation. Lorsqu'elle est placée en amont du promoteur, l'interaction entre cette séquence et Cra provoquerait une déformation de l'ADN qui favoriserait l'interaction entre Cra et l'ARN polymérase et ainsi la transcription (Cozzone, 1998). En revanche, lorsqu'elle est placée en aval de la région promotrice, le complexe formé aboutit à l'inhibition de l'expression des gènes concernés. L'activité de liaison à l'ADN de Cra serait elle même inhibée par le FBP et le fructose-1-phosphate, qui lorsque le glucose est rapidement métabolisé, se trouvent en concentration élevée dans la bactérie.

2.3.3. Le rôle direct du PTS dans la régulation des gènes du transport et du catabolisme primaire de sources spécifiques de carbone

Pour éviter de synthétiser constamment les transporteurs et les enzymes cataboliques nécessaires à l'entrée des glucides et en particulier des sucres non préférentiels dans les voies du métabolisme central du carbone (ce que j'appelle catabolisme primaire), la transcription des gènes codant ces protéines est, dans la plupart des cas, induite seulement si leur substrat spécifique est présent dans le milieu de croissance et si le glucose ou un autre sucre préférentiel est absent. Dans un certain nombre de cas, chez *Bacillus subtilis*, le mécanisme de répression catabolique impliquant CcpA (cf partie 2.3.2) empêche la transcription de ces

Gènes régulés positivement par la protéine Cra :

Phosphoénolpyruvate synthétase (*ppsA*)

Phosphoénolpyruvate carboxykinase (*pckA*)

Malate synthétase, isocitrate lyase (*aceBA*)

Isocitrate déshydrogénase (*icd*)

Fructose-1,6-biphosphatase (*fdp*)

+ Cytochrome d oxydase (*cyd*)

Gènes régulés négativement par la protéine Cra :

Enzymes du catabolisme du fructose (*fruBKA*)

HPr, Enzyme I (*ptsHI*)

Enzymes de la voie d'Entner-Doudoroff (*edd-eda*)

Phosphofructokinase (*pfk*)

Enzymes du catabolisme du mannitol (*mltADR*)

Erythrose-4-phosphate déshydrogénase (*gapB*)

Pyruvate kinase (*pykF*)

Table 4. Liste des gènes dont l'expression est régulée par le régulateur transcriptionnel Cra. D'après Saier et Ramseier, 1996.

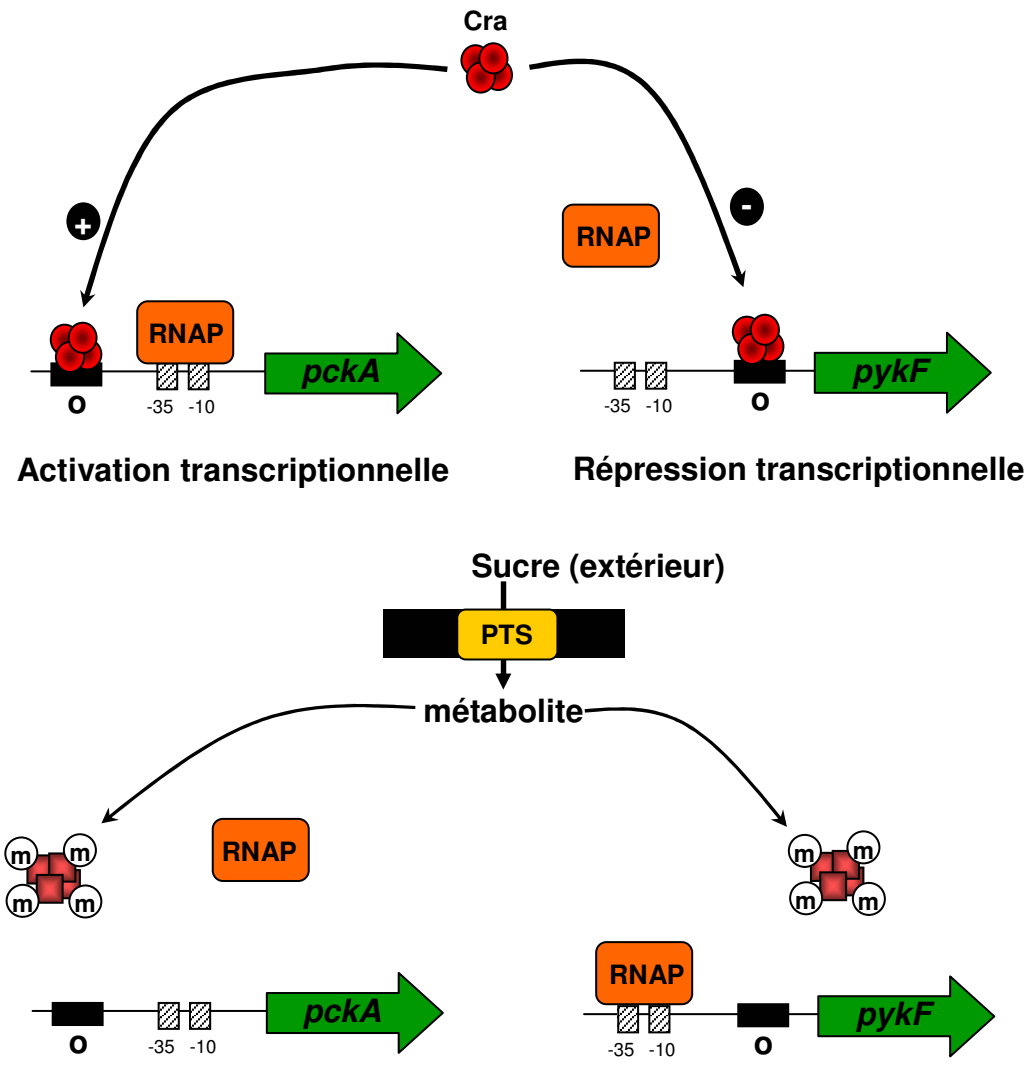


Figure 31. Modèle de répression catabolique impliquant le régulateur transcriptionnel Cra chez *Escherichia coli*.

En absence de sucre PTS, la protéine Cra interagit avec un opérateur spécifique (o) situé en amont du promoteur. Le complexe Cra/opérateur favorise l'activité de l'ARN polymérase. Au contraire, chevauchant ou localisé en aval du promoteur, ce complexe provoque une inhibition de la transcription. Mais en présence d'un sucre rapidement métabolisable, le métabolisme conduit notamment à l'augmentation de la concentration en FBP qui rend Cra inactive et en conséquence perturbe l'interaction entre Cra et son opérateur.

D'après Saier et Ramseier, 1996.

gènes et opérons lorsque le glucose ou un autre sucre préférentiel est présent, et cela même si le substrat spécifique est présent lui aussi (arabinose, β-glucosides, galactose-1-P, glycérol, gluconate, inositol, tréhalose, xylosides, xylose,...). Dans d'autres cas, l'activité des régulateurs spécifiques de la transcription de ces gènes et opérons est directement modulée par le PTS en fonction de la présence ou de l'absence de glucose (ou d'un autre sucre rapidement métabolisable). Dans ces cas-ci, le régulateur spécifique est un régulateur positif, de type antiterminateur ou activateur, possédant un domaine particulier dit PRD (PTS Regulation Domain) qui est la cible du contrôle positif et négatif par les composantes du PTS (Table 5). Je n'évoquerai, dans cette partie, que le rôle du PTS comme voie directe de signalisation vers ces régulateurs de l'état du métabolisme central du carbone et des sources de carbone présentes dans le milieu.

2.3.3.1. Cas des antiterminateurs de l'élongation de la transcription

Parmi les mécanismes bactériens de régulation de l'expression des gènes, la terminaison/antiterminaison transcriptionnelle constitue un système assez commun de contrôle du niveau de transcription, surtout chez les bactéries à Gram positif (Henkin & Yanofsky, 2002; Stülke, 2002). Concernant la régulation du métabolisme carboné chez les bactéries à Gram positif, on peut distinguer deux classes suivant la famille à laquelle appartient la protéine antiterminatrice impliquée : la classe des antiterminateurs possédant un domaine PRD donc sous contrôle du PTS et la classe des régulateurs non contrôlés par le PTS, dont un représentant est GlpP (utilisation du glycérol). Ici, il ne sera question que des antiterminateurs à PRD contrôlés par le PTS. A cet égard, *Bacillus subtilis* possède plusieurs opérons cataboliques qui sont ainsi contrôlés : l'opéron *bglPH* pour l'utilisation des β-glucosides, contrôlé par LicT (Kruger & Hecker, 1995; le Coq *et al.*, 1995; Schnetz *et al.*, 1996) ; le gène *sacB* et l'opéron *sacXY* pour l'expression de la lévane-saccharase, contrôlés par SacY (Steinmetz *et al.*, 1989; Tortosa *et al.*, 1997; Idelson & Amster-Choder, 1998) ; l'opéron *sacPA* qui code un système d'utilisation du saccharose, contrôlé par SacT (Débarbouillé *et al.*, 1990) ; et l'opéron *ptsGHI* codant le transporteur PTS EII^{Glc} et les protéines générales du PTS, contrôlé par GlcT (Bachem *et al.*, 1997; Stülke *et al.*, 1997; Bachem & Stülke, 1998). Les contrôles physiologiques des régulations médiées par LicT et GlcT seront présentés ci-après.

Régulateurs transcriptionnels contenant des PRD dont l'activité a été étudiée de manière détaillée.								
Régulateur possédant des PRD	**Organisme**	**Opérons régulés (sucres métabolisés)**	**Nombre de résidus**	**Régu-lateur négatif**	**Activa-tion HPr dépen-dante**	**Résidus histidine conservés[a]**		**Références**
Antiterminateurs						**PRD1**	**PRD2**	
BglG	*E. coli*	*bglFB* ; *bglG* (β-glucoside)	278	BglF	oui	His101 ; His159	His208(-)	Chen *et al.*,1997 ; Gorke *et al.*,1999
GlcT	*B. subtilis*	*ptsGHI* (glucose)	281	PtsG	non	His104(-) ; His163	His211 ; His272	Bachem *et al.*,1998
LicT	*B. subtilis*	*bglS* ; *bglPH* (β-glucoside)	277	BglP	oui	His100(-)^{P} ; His159(-)^{P}	His2007(+)^{P}; His269(+)^{P}	Tortosa *et al.*, 2001 ; Lindner *et al.*,1999
SacY	*B. subtilis*	*sacB* ; *sacXY* (sucrose)	280	SacX	non	His99(-)^{P} ; Hhis158	His207^{P}; His269^{P}	Tortosa *et al.*, 1997 ; Tortosa *et al.*, 2001
SacT	*B. subtilis*	*sacPA* (sucrose)	276	SacP	oui	His97(-); His156	His205; His267	Arnaud *et al.*,1996
Activateurs								
MtlR	*B. stearo-thermophilus*	*mtlARFD* (mannitol)	697	MtlA	oui	His236(-)^{P}; His295(-)^{P}	His348(+)^{P} ; His405(+)^{P}	Henstra *et al.*, 2000
LicR	*B. subtilis*	*licBCAH* (β-glucoside)	641	LicB	oui	His219(+) ; His278(+)	His333(+)^{P}; His392(+)	Tobisch *et al.*,1999
LevR	*B. subtilis*	*levDEFGsacC* (fructose)	938	LevE	oui	His506^{P} ; His567^{P}	His869(-)^{P}	Martin-Verstraete *et al.*, 1998

Table 5. Régulateurs transcriptionnels contenant des PRD dont l'activité a été étudiée de manière détaillée.
[a] Les résidus histidines dont la phosphorylation a été montrée *in vitro* et qui sont le site d'une régulation négative (-) médiée par EII ou positive (+) médiée par HPr sont indiqués.
D'après van Tilbeurgh et Declerck, 2001.

La Figure 32 illustre le mécanisme général commun aux antiterminateurs possédant un domaine PRD. La transcription des gènes et opérons qui sont contrôlés par ce mécanisme est initiée constitutivement. Mais, en l'absence d'inducteur, l'élongation est très rapidement bloquée par la formation d'une structure terminatrice, localisée entre le promoteur et la phase codante du premier gène. Lorsqu'il est activé par le système PTS en réponse à la présence de l'inducteur spécifique et à l'absence de glucose ou d'un sucre plus efficacement métabolisable, l'antiterminateur se fixe sur l'ARNm naissant au niveau d'une séquence spécifique appelée RAT (Ribonucleic Anti-Terminator) qui précède et chevauche le terminateur (Houman *et al.*, 1990; Aymerich & Steinmetz, 1992; Arnaud *et al.*, 1996; Stülke *et al.*, 1997). La fixation de l'antiterminateur sur sa séquence RAT permet de stabiliser la structure secondaire en tige-boucle du RAT (Manival *et al.*, 1997), empêchant ainsi la formation du terminateur. L'ARN polymérase peut alors poursuivre l'élongation de la transcription.

L'un des mécanismes les plus étudiés impliquant un antiterminateur de ce type a été caractérisé chez *Escherichia coli* : il implique l'antiterminateur BglG qui régule l'expression des gènes du métabolisme des β-glucosides (Mahadevan *et al.*, 1987; Mahadevan & Wright, 1987; Schnetz & Rak, 1988; Amster-Choder *et al.*, 1989; Amster-Choder & Wright, 1990; Houman *et al.*, 1990; Amster-Choder & Wright, 1992; Amster-Choder & Wright, 1993). Il est le seul de ce type contrôlant des gènes du métabolisme carboné chez *Escherichia coli* (Stülke *et al.*, 1998).

Les antiterminateurs possédant un domaine PRD appartiennent à la famille BglG/SacY. Ces antiterminateurs sont constitués de trois domaines : un domaine N-terminal de fixation à l'ARNm d'environ 55 acides aminés nommé CAT (Co Anti-Terminator) et deux domaines PRD d'environ 110 acides aminés chacun, notés PRDI et PRDII (Figure 33) (Manival *et al.*, 1997; Tortosa *et al.*, 1997; Stülke *et al.*, 1998). Chaque PRD contient deux histidines conservées (Table 5). Ces résidus sont phosphorylables de façon réversible par le système PTS. Le domaine CAT reconnaît spécifiquement une séquence RAT donnée. Lorsqu'il est exprimé sous forme de domaine isolé, son activité de liaison à l'ARN RAT est insensible à l'état du PTS ou à la présence/absence de l'inducteur spécifique (Manival *et al.*, 1997; Declerck *et al.*, 1999). La capacité de l'antiterminateur complet à lier l'ARNm est donc modulée par l'état des domaines PRDI et PRDII.

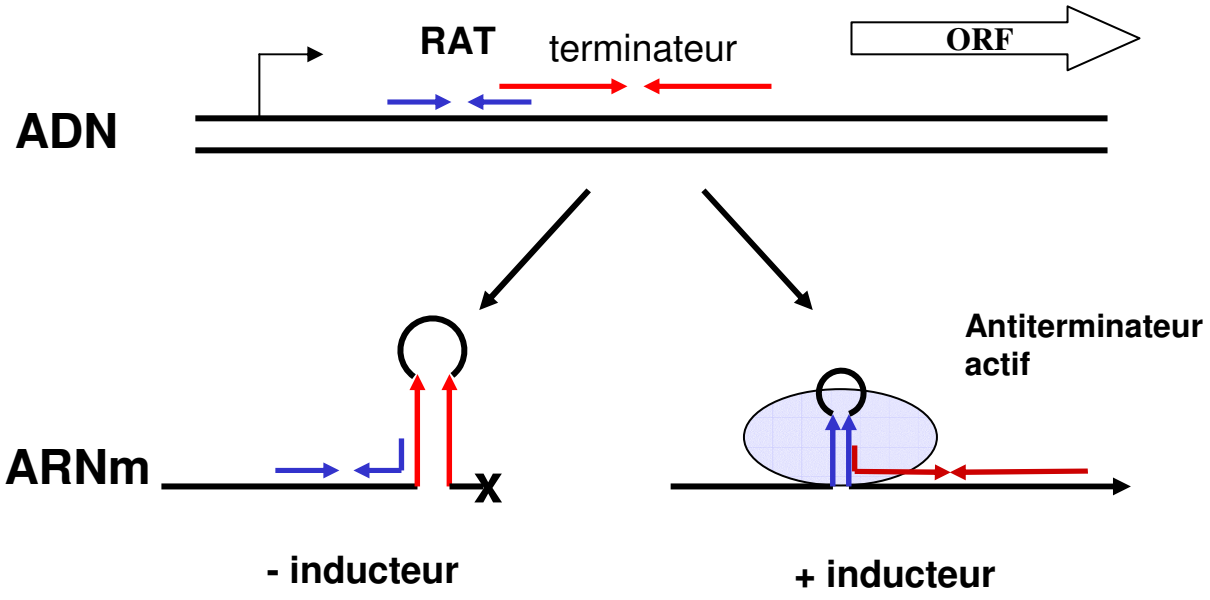


Figure 32. Mécanisme général de régulation de l'expression génique médiée par un antiterminateur transcriptionnel possédant un domaine PRD.

En l'absence de substrat inducteur, la séquence terminatrice forme une structure secondaire qui empêche la transcription de se poursuivre. En présence de l'inducteur et en absence de sucre « préférentiel », la protéine antiterminatrice est activée et se fixe sur la séquence RAT de l'ARNm dès qu'elle est transcrite par la polymérase. Ainsi, la structure en tige boucle formée par la séquence RAT est stabilisée et empêche la formation du terminateur. La transcription se poursuit…

D'après Tortosa *et al.*, 1997.

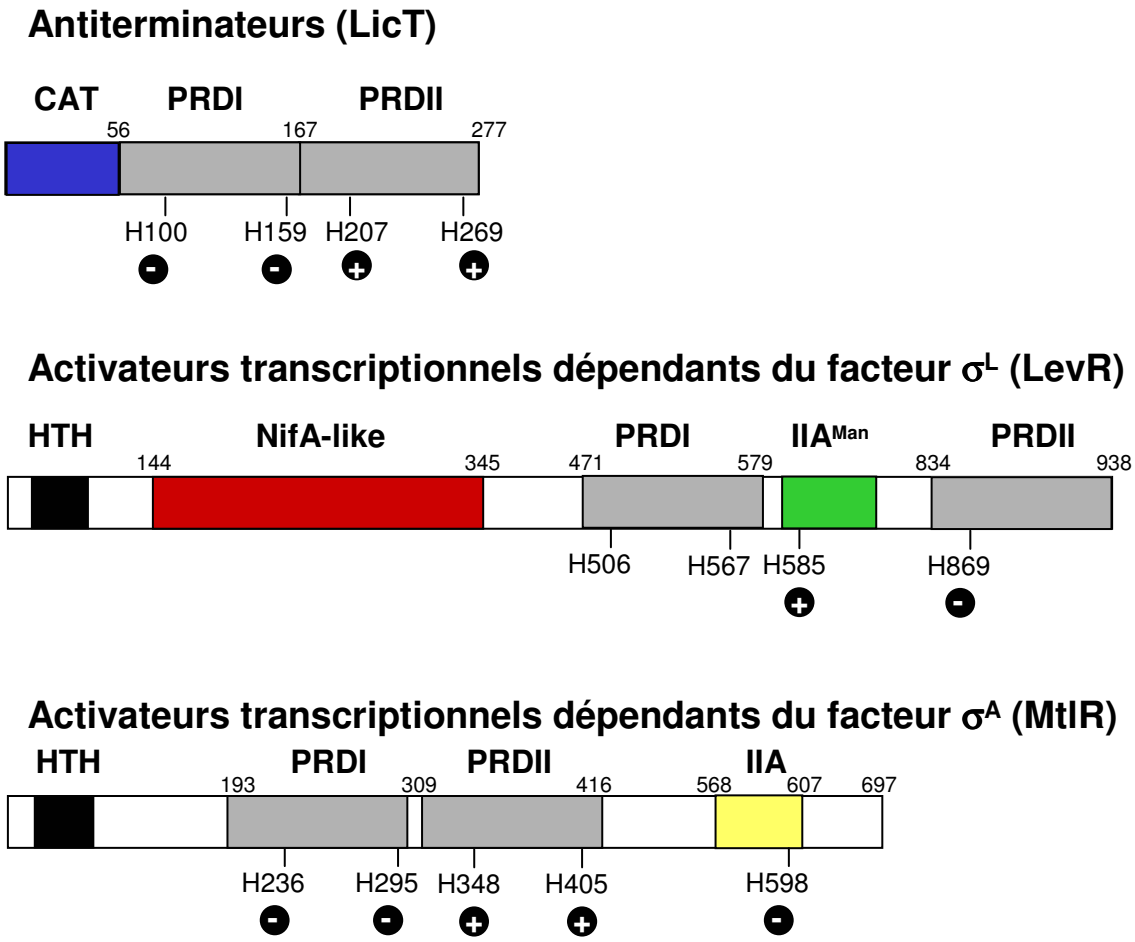


Figure 33. Structure secondaire des régulateurs possédant le domaine PRD.

Le domaine PRD est indiqué en gris. Il est commun aux antiterminateurs, aux activateurs transcriptionnels σL ou σA –dépendants. Les sites de phosphorylation et l'effet positif ou négatif résultant sont indiqués.

En bleu est indiqué le domaine CAT de l'antiterminateur. En noir sont indiqués les domaines hélice-tour-hélice d'interaction avec l'ADN des régulateurs transcriptionnels. En rouge est indiqué le domaine d'interaction avec l'ARN polymérase qui contient le facteur σ54. En vert ou en jaune sont indiqués les domaines homologues à l'enzyme IIA.

D'après van Tilbeurgh et Declerck, 2001.

2.3.3.1.1. Régulation des gènes de l'utilisation des β-glucosides par LicT chez Bacillus subtilis

L'activité de liaison à l'ARNm de LicT est contrôlée par phosphorylation réversible des domaines PRD par le PTS. D'une part, en l'absence de glucose, HPr-His15-P peut phosphoryler le domaine PRDII (C-terminal) sur deux résidus histidine conservés et ainsi stimuler l'activité de l'antiterminateur (Figure 34). Mais en présence de glucose, HPr est majoritairement sous la forme HPr-Ser46-P ; le domaine PRDII ne peut plus être phosphorylé et l'antiterminateur est inactif (Stülke *et al.*, 1998; Lindner *et al.*, 1999; van Tilbeurgh & Declerck, 2001). D'autre part, la présence d'inducteur est aussi nécessaire à la stimulation de l'activité de l'antiterminateur. En présence de salicine dans le milieu de croissance, ce substrat est transporté et phosphorylé par le transporteur PTS EII^{Bgl}. Mais en l'absence de salicine, EII^{Bgl} est constamment phosphorylée. Dans cet état, EII^{Bgl}-P peut phosphoryler le domaine PRDI et inactiver l'antiterminateur (le Coq *et al.*, 1995; Lindner *et al.*, 1999). Finalement, l'antiterminateur n'est actif que s'il est à la fois phosphorylé sur le domaine PRDI et à la fois déphosphorylé sur le domaine PRDII, un état induit par la présence de salicine et l'absence de glucose ou d'un autre sucre rapidement métabolisable (Figure 34). Les différents états de phosphorylation de LicT correspondent à différentes structures quaternaires, à un équilibre déplacé entre une forme dimérique très compacte (état actif) et une forme dimérique plus ouverte et peut-être une forme monomérique (états inactifs) (Declerck *et al.*, 2001; van Tilbeurgh & Declerck, 2001).

Par ailleurs, il faut rappeler que cette régulation de l'expression de l'opéron *bglPH* se double d'une répression catabolique par le glucose ou un autre sucre préférentiel médiée par CcpA agissant au niveau de l'initiation de la transcription (Krüger & Hecker, 1995).

2.3.3.1.2. GlcT : contrôle du transport du glucose et des composants généraux du PTS

Le cas de GlcT est évidemment différent de celui de LicT puisque l'inducteur spécifique est ici aussi un sucre « préférentiel » : le glucose.

Chez *Bacillus subtilis*, le glucose est majoritairement transporté *via* le complexe PTS glucose spécifique EII^{Glc} (Gay, 1979) codé par le gène *ptsG*. *ptsG* est transcrit en opéron avec les gènes *ptsH* et *ptsI* qui codent les composantes générales du PTS, en réponse à la présence

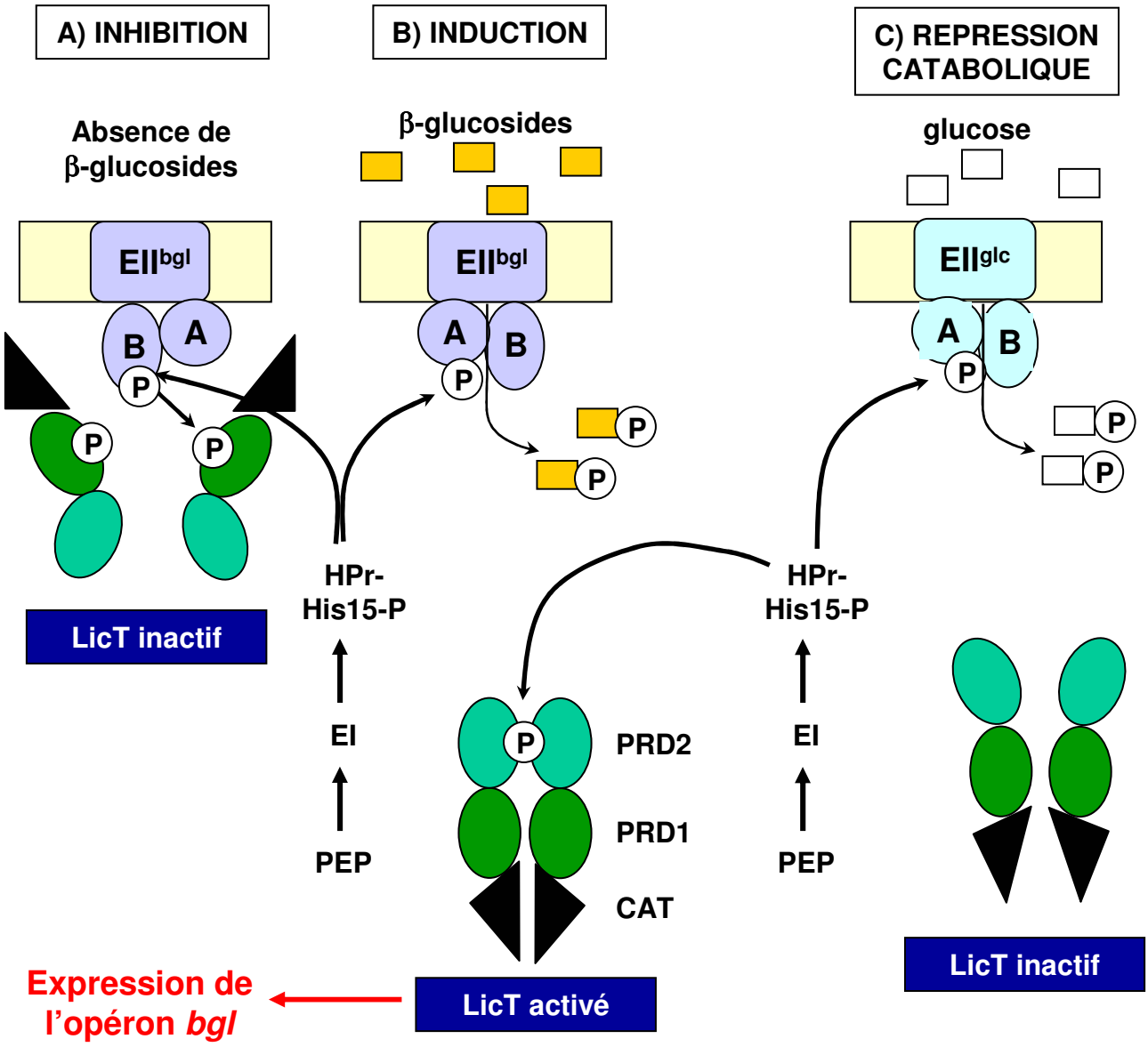


Figure 34. Modèle de régulation de l'activité de l'antiterminateur LicT par le système PTS chez *Bacillus subtilis*.

D'après van Tilbeurgh et Declerck, 2001.

de glucose (Stülke *et al.*, 1997). L'expression de l'opéron *ptsGHI* est contrôlée par un mécanisme d'antiterminaison impliquant GlcT (Figure 35). En présence de glucose, GlcT n'est pas phosphorylé parce que le phosphate est « drainé » par la chaîne PTS. Dans cet état, GlcT forme un dimère actif et empêche la terminaison de l'ARNm *ptsGHI* naissant (Stülke *et al.*, 1997; Knezevic *et al.*, 2000). Mais en l'absence de glucose, le domaine PRDI est phosphorylé sur l'histidine en position 104 et l'antiterminateur est inactif (Bachem & Stülke, 1998). Dans ces conditions, HPr-His15-P est le donneur primaire du groupement phosphate mais il n'est pas établi s'il y a transfert direct du groupement phosphate d'HPr vers le domaine PRDI ou s'il y a d'abord transfert de ce phosphate vers EIIGlc puis transfert d'EIIGlc vers le domaine PRDI. En effet, une analyse génétique montre que l'intégrité d'EII^{Glc} est requise pour réprimer l'expression de l'opéron *ptsGHI* en l'absence de glucose (Bachem & Stülke, 1998). Toutefois, même en l'absence de glucose, l'expression de *ptsH* et de *ptsI* n'est jamais nulle car il existe un promoteur secondaire constitutif entre *ptsG* et *ptsH* (Stülke *et al.*, 1997).

Chez *Bacillus subtilis*, l'activité de l'antiterminateur SacY est essentiellement régulée en fonction de la disponibilité de l'inducteur spécifique (comme GlcT) tandis que celle de l'antiterminateur SacP est l'est en fonction de la disponibilité de son inducteur spécifique et de la présence/absence de glucose (comme LicT) (Stülke *et al.*, 1998).

2.3.3.2. Cas des activateurs de l'initiation de la transcription

Chez *Bacillus subtilis*, quatre activateurs transcriptionnels possédant des domaines PRD sont impliqués dans la régulation de gènes spécifiques de l'utilisation de sources de carbone : MtlR (utilisation du mannitol), LicR (utilisation des hydrolysats de lichénane), ManR (utilisation du mannose) et LevR (utilisation des polymères de fructose et du fructose). Ils peuvent être classés dans deux sous-groupes suivant le facteur σ requis pour la transcription : les activateurs dépendants du facteur σ^A, illustré par LicR et l'activateur dépendant du facteur σ^L, illustré par un unique représentant, LevR (Figure 33). Ces activateurs transcriptionnels contenant un domaine PRD semblent être présents surtout chez les bactéries à Gram positif (Stülke *et al.*, 1998) mais existent aussi chez les bactéries à Gram négatif (Deutscher *et al.*, 2002).

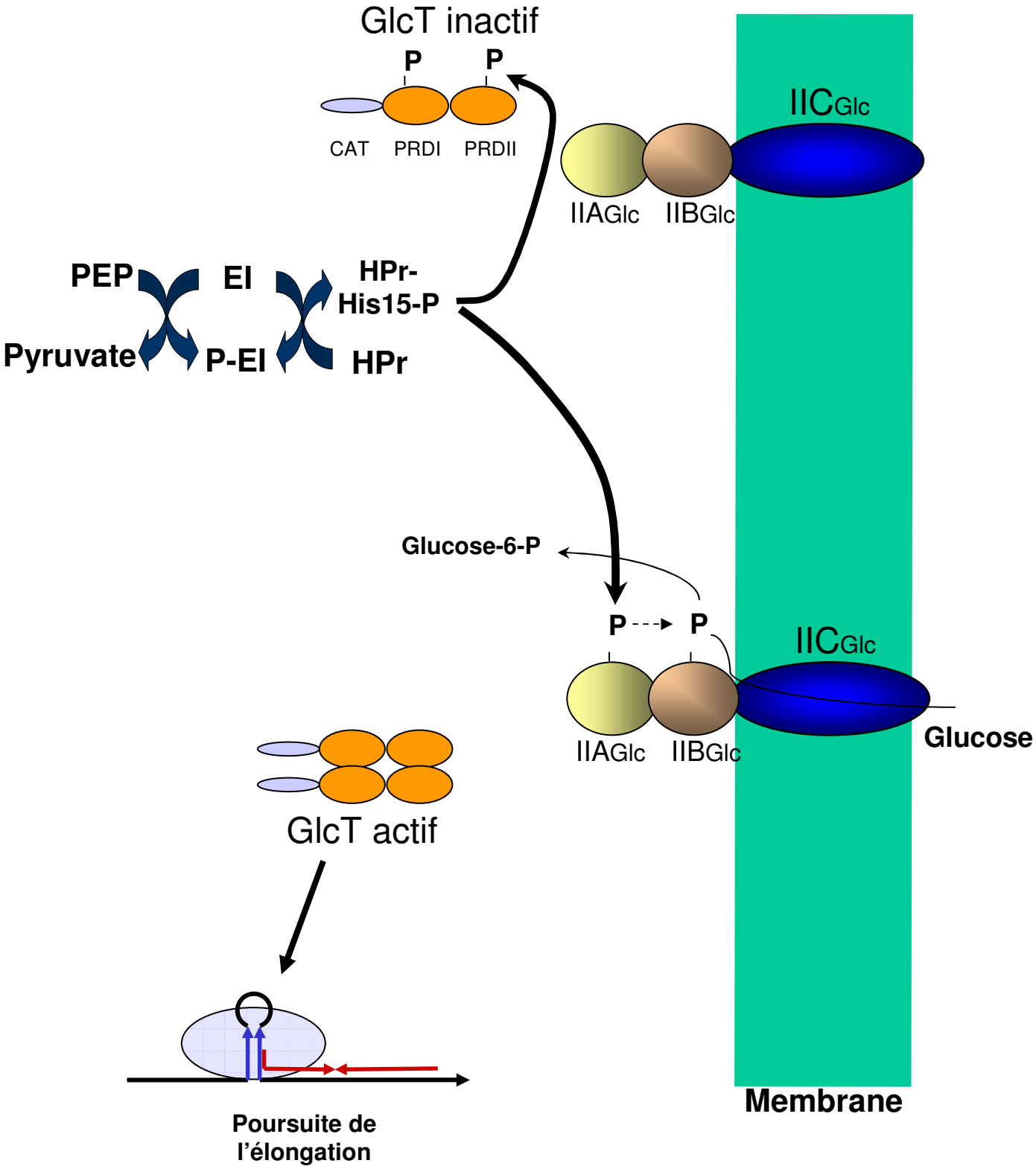


Figure 35. Modèle proposé de régulation de l'expression de l'opéron *ptsGHI* par l'antiterminateur GlcT en réponse à la présence de glucose chez *Bacillus subtilis*.

D'après Bachem et Stülke, 1998.

En réponse à la présence de fructose, la protéine LevR active l'expression de l'opéron *lev*, qui code les enzymes de la dégradation des lévanes (polymères de fructose) et du catabolisme du fructose (Martin-Verstraete *et al.*, 1990) (Figure 36). Comme évoqué précédemment, cette régulation est liée à celle médiée par CcpA (cf partie 2.3.2). LevR est une protéine de 938 acides aminés qui possède un domaine N-terminal hélice-tour-hélice de liaison à l'ADN de la famille NifA/NtrC, un domaine homologue au domaine central de la protéine NifA d'interaction avec l'ARN polymérase, deux domaines PRDI et PRDII en C-terminal et un domaine similaire à un domaine enzyme IIA^{Man} situé entre les deux domaines PRD (Débarbouillé *et al.*, 1991; Stulke *et al.*, 1995; Charrier *et al.*, 1997; Greenberg *et al.*, 2002) (Figure 33). Le domaine de liaison à l'ADN permet de reconnaître une séquence spécifique située en amont du promoteur de l'opéron *lev* (Martin-Verstraete *et al.*, 1992; Martin-Verstraete *et al.*, 1994).

Il a été montré qu'une mutation de *levD* ou de *levE*, qui codent respectivement les protéines $EIIA^{Lev}$ et $EIIB^{Lev}$ d'un système de transport PTS du fructose, résultait en une expression constitutive de l'opéron *lev*. De plus, le test d'une mutation permettant de produire une protéine LevR tronquée du domaine PRDII (C-terminal) a abouti au même résultat (Martin-Verstraete *et al.*, 1990; Débarbouillé *et al.*, 1991). Puis, la phosphorylation d'un résidu histidine conservé du domaine PRDII par $EIIB^{Lev}$ en absence de fructose a été mise en évidence. Cette phosphorylation serait impliquée dans l'inactivation de LevR (Martin-Verstraete *et al.*, 1998). En présence de fructose dans le milieu de croissance, $EIIB^{Lev}$ transfère préférentiellement son groupement phosphate au substrat lié à la perméase $EIIC^{Lev}$ (LevFG) ; le domaine PRDII n'est plus phosphorylé. Mais l'activation de LevR nécessite aussi l'absence de glucose ; dans ces conditions, HPr-His15-P est présent en forte concentration et peut phosphoryler le domaine homologue à l'$EIIA^{Man}$ de LevR. Cette phosphorylation est nécessaire, concomitamment à l'absence de phosphorylation du PRDII, pour une activation de LevR (Martin-Verstraete *et al.*, 1998).

Par ailleurs, l'initiation de la transcription de l'opéron lev est aussi contrôlée par CcpA en fonction de la présence/absence de glucose (Martin-Verstraete *et al.*, 1995; Martin-Verstraete *et al.*, 1999).

Chez *Bacillus stearothermophilus*, il a été montré que MtlR active l'expression de l'opéron *mlt*, impliqué dans l'utilisation du mannitol, en réponse à la présence de mannitol dans le milieu de croissance et à l'absence d'un sucre « préférentiel » (Henstra *et al.*, 1999).

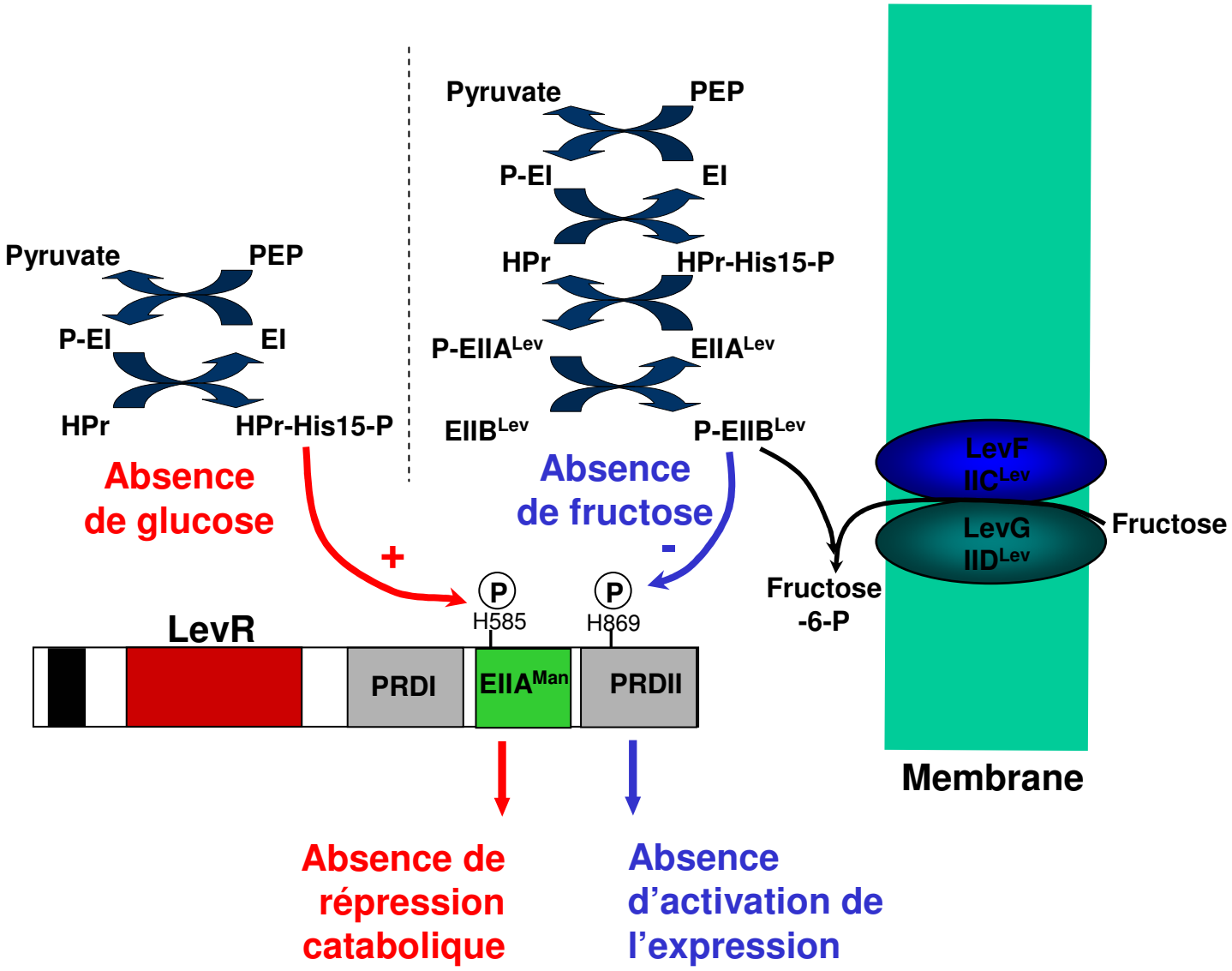


Figure 36. Modèle de régulation de l'expression de l'opéron *lev* par l'activateur transcriptionnel LevR chez *Bacillus subtilis*. Les couleurs des domaines protéiques sont les mêmes que celles utilisées dans la figure 34.

D'après Deutscher *et al.*, 2002.

Chez *Bacillus subtilis*, LicR contrôle l'expression de l'opéron *lic* impliqué dans l'utilisation des hydrolysats de lichénane (Tobisch *et al.*, 1997; Tobisch *et al.*, 1999a) et cette régulation se double de la répression catabolique médiée par CcpA. Ces deux activateurs sont composés d'un domaine N-terminal hélice-tour-hélice, de deux domaines PRD (I et II) et d'un domaine C-terminal homologue à un domaine EIIA (Figure 33).

MtlR peut être phosphorylé sur les domaines PRDI et PRDII et le domaine homologue à un domaine EIIA. La phosphorylation des domaines PRDI par HPr-His15-P et EIIA-like par $EIICB^{Mtl}$-P diminue l'affinité de MtlR pour sa cible tandis que la phosphorylation du domaine PRDII par HPr-His15-P augmente l'affinité de MtlR pour sa cible. Par ailleurs, la stimulation de l'activité de MtlR n'est possible que lorsque les domaines PRDI et homologue à EIIA sont déphosphorylés (Henstra *et al.*, 1999; Henstra *et al.*, 2000). En l'absence de mannitol et d'autres substrats, les composantes du PTS sont constamment phosphorylées ; HPr-His15-P et $EIICB^{Mtl}$-P peuvent alors phosphoryler les domaines PRDI, PRDII et homologue à EIIA. Dans cet état, MtlR est inactif et l'opéron *mtl* n'est pas transcrit. En présence de mannitol et en absence d'un sucre « préférentiel », $EIICB^{Mtl}$ est majoritairement sous la forme déphosphorylée puisqu'elle transfère son groupement phosphate vers le mannitol. Les domaines PRDI et homologue à EIIA sont alors déphosphorylés par $EIICB^{Mtl}$ tandis que le domaine PRDII est phosphorylé par HPr-His15-P. Dans ce second état, MtlR se lie à sa cible avec haute affinité et active la transcription de l'opéron *mtl*. Mais en présence de glucose ou d'un autre sucre « préférentiel », la concentration de HPr-His15-P diminue. En conséquence, le domaine PRDII n'est plus phosphorylé et l'expression de l'opéron *mtl* n'est pas stimulée (Henstra *et al.*, 2000).

Le mécanisme par lequel l'activité de LicR est modulée est semblable (Tobisch *et al.*, 1999a). Il a été proposé qu'en l'absence de β-glucosides, EII^{Lic} phosphorylerait le domaine homologue à EIIA et inactiverait ainsi LicR : une étude génétique montre que lorsque les sites de phosphorylation du domaine homologue à EIIA de LicR sont mutés et ne sont plus phosphorylables, l'activité du promoteur de l'opéron *lic* est constitutive (Tobisch et al, 1999). En présence de β-glucosides, EII^{Lic} transfère préférentiellement son groupement phosphate à son substrat et ne pourrait plus phosphoryler le domaine homologue à EIIA. Concomitamment, les domaines PRDI et PRDII seraient phosphorylés par HPr-His15-P pour activer LicR. Mais en présence de glucose, les domaines PRDI et PRDII ne seraient plus phosphorylés par HPr-His15-P et LicR serait inactif.

2.3.4. Les régulations des gènes des voies centrales du métabolisme du carbone indépendantes du PTS (ou indirectement affectées par le PTS)

Concernant les voies centrales du métabolisme *sensu stricto*, nous avons vu précédemment que l'expression des gènes *citZ* et *citB* du cycle de Krebs est régulée directement par le gène CcpC et que le rôle de CcpA dans la régulation de *citB* est indirect et passe par la régulation de l'expression de *ccpC* (cf partie 2.3.2).

A côté de CcpC, deux autres régulateurs transcriptionnels sont impliqués dans la régulation catabolique de gènes et opérons spécifiques du métabolisme central du carbone : CggR et YqzB.

2.3.4.1. La régulation de l'expression de l'opéron *gapA* : contrôle de l'expression par le répresseur transcriptionnel CggR

Cette régulation constitue l'objet d'une partie de notre travail de thèse et est présenté dans le premier chapitre de la section Résultats.

2.3.4.2. La régulation de la néoglucogenèse chez *Bacillus subtilis*

Dans le premier chapitre, nous avons vu que *Bacillus subtilis* possédait plusieurs gènes nécessaires à la néoglucogenèse. Les diverses analyses génétiques et transcriptomiques montrent que, parmi eux, seuls les gènes *gapB* et *pckA* sont clairement régulés (Fujita *et al.*, 1998; Fillinger *et al.*, 2000; Moreno *et al.*, 2001; Yoshida *et al.*, 2001) (Article n°2, soumis cf section Résultats).

Une grande partie des travaux concernant la régulation de l'expression des gènes *gapB* et *pckA* ont été réalisés au laboratoire par Pascale Servant et sont inclus dans un manuscrit en cours de préparation. Nous n'en ferons donc qu'une brève présentation.

La transcription de *gapB*, le gène codant la glycéraldéhyde-3-phosphate déshydrogénase impliquée dans la néoglucogenèse, est très faible en conditions glycolytiques mais est fortement induite pendant la néoglucogenèse (Fillinger *et al.*, 2000). Une délétion du gène *ccpA* ne supprime que très partiellement cette répression par le glucose ou par les autres sources de carbone glycolytiques. Un crible génétique a donc été réalisé afin d'identifier le régulateur majeur de la transcription de *gapB*. Il a permis de suggérer que la protéine codée par le gène *yqzB* est ce régulateur. Les analyses génétiques ont montré qu'YqzB est un répresseur de l'initiation de la transcription de *gapB* et sa séquence cible a été localisée. Le gène *yqzB* fait partie d'un opéron bicistronique dont le deuxième gène est *yqfL*. Les analyses génétiques ont montré qu'*yqfL* code un régulateur positif de la transcription de *gapB* et qu'il agit *via* YqzB. *pckA*, le gène qui code une phosphoénolpyruvate carboxykinase, est régulé de manière semblable à *gapB*, *via* YqzB et YqfL. Par ailleurs, CcpA n'est nécessaire à la régulation de *gapB* et *pckA* qu'en présence de sources de carbone glycolytiques dont la perméation dépend du PTS, parce qu'il est nécessaire au transport efficace de ces sources de carbone (Ludwig *et al.*, 2002b) (le Coq et Aymerich, non publié).

Enfin, l'analyse de séquence suggère qu'un mécanisme identique est à l'oeuvre chez la plupart des bactéries à Gram positif et faible pourcentage en GC.

2.3.4.3. La régulation de l'expression de l'enzyme malique codée par *ywkA*

L'analyse de séquence suggère que, chez *Bacillus subtilis*, quatre gènes paralogues peuvent coder des enzymes maliques : *ywkA*, *malS*, *ytsJ* et *mleA*. La régulation transcriptionnelle d'*ywkA*, dont nous avons démontré qu'il code bien une enzyme malique, montre qu'il est induit spécifiquement en présence de malate extracellulaire par le système à deux composants YufL-YufM. Cette étude constitue le second objet de cette thèse et est présentée dans le deuxième chapitre de la section Résultats.

3. Vers l'intégration des données de régulation dans un modèle global à contraintes

Nous avons vu que *Bacillus subtilis* possède un réseau de régulations enchevêtrées à plusieurs niveaux permettant semble-t-il de contrôler très finement l'expression des gènes impliqués dans le métabolisme du carbone en réponse à la disponibilité des sources présentes dans le milieu. Aussi, il est apparu de plus en plus nécessaire de proposer une approche intégrée. Pour appréhender la complexité qui caractérise les processus cellulaires à grand nombre d'éléments et de régulations, des suites de bases de données et de logiciels associés de bioinformatique comme KEGG (http://www.genome.ad.jp/kegg/) ont été construites et répertorient de la manière la plus exhaustive possible de très diverses informations d'ordres génétique, génomique et biochimique complémentaires concernant un organisme. Cela permet d'avoir une vision intégrée d'un processus cellulaire et de comparer de façon exhaustive les divers éléments impliqués dans ce processus chez plusieurs bactéries. Cette approche intégrée est bien adaptée à l'étude des voies métaboliques du carbone. Par ailleurs, aussi dans l'optique de mieux comprendre et de prévoir comment tous les éléments intervenant dans un même processus fonctionnent dans un réseau, des modèles simulatoires de fonctionnement du métabolisme prenant en compte les conditions de croissance (les sources de carbone disponibles, aérobiose/anaérobiose, etc...) ont été élaborés. Ces modèles *in silico* sont basés sur les données biochimiques -quelles conditions fondamentales sont nécessaires pour que telle ou telle enzyme fonctionne ? -, les données génomiques qui permettent d'avoir une vision exhaustive des voies métaboliques -pour chaque étape enzymatique, existe-t-il un ou plusieurs gènes codant l'enzyme requise ? -, et les analyses de flux métaboliques qui, en quelque sorte, rendent compte de la distribution des flux cataboliques et anaboliques dans des conditions bien particulières - par exemple, comment sont redistribués les flux de carbone dans tel ou tel mutant et dans telle ou telle condition de croissance ?-. Ces modèles sont dits modèles d'analyse de balance de flux métaboliques.

Les analyses de flux métaboliques ont débuté vers les années 1980 et sont actuellement en plein essor. En effet, elles s'appuient sur des techniques depuis peu bien adaptées à ce sujet, qui sont la résonance magnétique nucléaire (RMN) et la spectrométrie de masse (MS) pour les mesures, et la modélisation mathématique et informatique pour la détermination des flux à partir des mesures (Wiechert, 2001). De façon extrêmement simplifiée, dans le cas de l'analyse

de flux métaboliques basée sur le marquage du carbone, le principe est le suivant : le glucose ou un autre substrat marqué, le [1-C^{13}]-glucose par exemple, est introduit dans le milieu de croissance et sera utilisé par la bactérie. Lors du processus métabolique, les atomes de carbone marqués vont être redistribués entre les différents pools de métabolites intracellulaires à travers les différentes voies empruntées par le carbone. L'enrichissement final en carbone marqué de ces pools est mesuré par RMN et MS. A cela s'ajoute la notion d'isotopomère. Ce terme, issu de la condensation des mots isotope et isomère, signifie l'un des états de marquage isotopique des atomes de carbone dans lequel un métabolite peut être trouvé. Chaque atome pouvant être ou ne pas être marqué (telle est la question !), il est théoriquement possible de trouver un métabolite à n atomes de carbone distribué en 2^n isotopomères différents. La détermination des proportions relatives de ces isotopomères apporte une information supplémentaire quant aux réactions possibles entre deux isotopomères de deux métabolites différents. Puis, à partir de modèles mathématiques très élaborés intégrant sens et vitesse de flux pour chaque étape du métabolisme, des valeurs vont être calculées/simulées qui correspondent le mieux au jeu de données issues des mesures précédemment acquises.

Dans le cas particulier de *Bacillus subtilis*, quelques résultats semblent apporter des informations très intéressantes. Par exemple, Sauer et ses collaborateurs ont déterminé la répartition du flux de carbone entre la glycolyse, la voie des pentoses phosphate et le cycle de Krebs d'une souche produisant de la riboflavine lors de la croissance en chémostat, en présence de glucose en quantité limitée ou en excès (Dauner *et al.*, 2001b). Lorsque la vitesse de croissance est limitée par la richesse du milieu en glucose, il n'y a pas de découplage entre le catabolisme et l'anabolisme dans le sens ou les précurseurs et l'énergie sont limitants. De plus, les trois voies sont actives : 80 % du glucose est métabolisé *via* la glycolyse et 20 % *via* la voie des pentoses phosphate; le flux à travers le cycle de Krebs est relativement élevé ce qui s'explique par le fait que le glucose et le glutamate ne sont pas en quantité suffisante pour observer un effet de répression catabolique médié par CcpA. Cependant, lorsque le glucose est en excès, le flux à travers le cycle de Krebs est faible et la production de produits secondaires sécrétés, acétate et acétoine, augmente fortement, ce qui est en accord avec le phénomène de répression/activation catabolique de l'expression des gènes impliqués dans ces processus métaboliques. Par ailleurs, cette étude a permis d'observer le flux de carbone à l'interface de la glycolyse et du cycle de Krebs qui traverse des voies alternatives aux réactions classiques (Figure 37). La première est un cycle futile qui, à partir de l'oxaloacétate, passe par le PEP, puis le pyruvate, pour reformer de l'oxaloacétate. Par tour, une molécule d'ATP sera donc

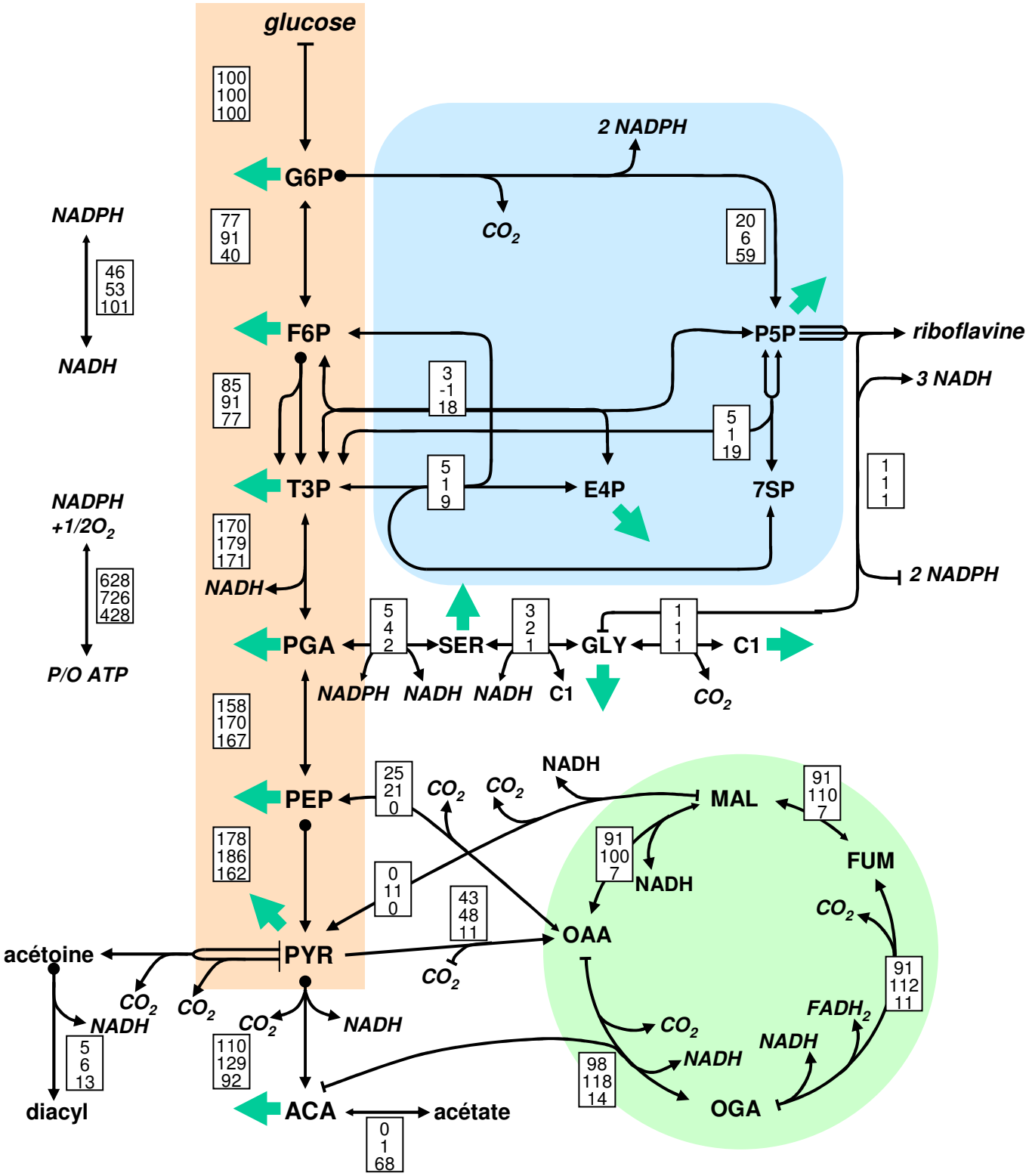


Figure 37. Distribution des flux métaboliques de carbone chez *Bacillus subtilis*. Les cultures ont été réalisées en chémostat avec limitation des sources de carbone (première valeur encadrée), d'azote (deuxième valeur) ou de phosphate (troisième valeur). Les grosses flèches indiquent la direction du flux et les petites flèches que la réaction est considéré comme réversible.

Abbréviations différentes des autres figures : T3P: trioses-phosphates, PGA: 3-phosphoglycérate, PYR: pyruvate, ACA: acétyl-CoA, P5P: pentoses-phosphates, E4P: érythrose-4-phosphate, 7SP: séduheptulose-7-phosphate, SER: sérine, GLY: glycine, C1: groupement méthyl lié au tétrahydrofolate, MAL: malate, FUM: fumarate, OGA: 2-oxoglutarate, OAA: oxaloacétate.

D'après Dauner *et al.*, 2001.

consommé sans bénéfice en retour. La seconde est une voie alternative pour la conversion du malate en oxaloacétate *via* l'enzyme malique et la pyruvate carboxylase qui consomme aussi une molécule d'ATP par molécule de malate.

Jusqu'à récemment, les modèles de balance de flux métaboliques n'intégraient pas les données de régulation de l'expression des gènes. Pourtant, il est légitime de penser que la présence de telle ou telle source de carbone peut entraîner l'expression de telle ou telle enzyme. Par conséquent, la prise en compte des données de régulation permet, en partie, de prédire quelles sont les voies actives et possibles pour les flux de carbone en fonction de la source de carbone présente dans le milieu. Autrement dit, l'ajout de contraintes de régulation permettra de restreindre l'espace de calcul des valeurs de flux dans un modèle donné. L'objectif ultime est donc d'intégrer toutes les données existantes de régulation en fonction des sources de carbone dans un schéma métabolique qui pourra être pris en compte lors de l'analyse de balance de flux.

Récemment, Covert et Palsson ont construit un modèle *in silico* d'analyse de balance de flux du métabolisme central d'*Escherichia coli* qui tient compte de la régulation de l'expression génique (Covert & Palsson, 2002). Ce modèle est basé sur un schéma combiné métabolique et régulatoire qui est presque entièrement décrit dans la Figure 38. Leur modèle inclut 149 gènes dont 73 codent les enzymes du métabolisme et 16 des protéines régulatrices. Ces enzymes catalysent 113 réactions en tout dont 45 sont régulées par un mécanisme transcriptionnel. A partir de ce modèle, il est possible de prédire plusieurs comportements de façon qualitative comme la capacité de souches mutantes à croître dans un milieu ou la source de carbone est définie (glucose, glycérol, succinate, acétate, ribose, ou deux substrats), l'utilisation d'un substrat donné et les produits en dérivant ou encore l'expression d'un gène dans diverses conditions (croissance aérobie ou anaérobie). Recoupées avec les données expérimentales disponibles, les prédictions réalisées se sont révélées très pertinentes. Parmi 116 cas expérimentaux disponibles comparés, ce modèle s'est trompé 10 fois sur l'ensemble des prédictions (106 étaient correctes) alors que le modèle de balance de flux sans contraintes de régulation s'est trompé 9 fois de plus. Le modèle présenté rend donc mieux compte du comportement réel du métabolisme central du carbone d'*Escherichia coli*.

Dans le cas de *Bacillus subtilis*, une telle étude n'est pas disponible. Cependant, un schéma sera présenté dans la partie discussion générale. Il tente de rassembler les principales données de régulation de l'expression des gènes impliqués dans le métabolisme central du carbone chez *Bacillus subtilis*.

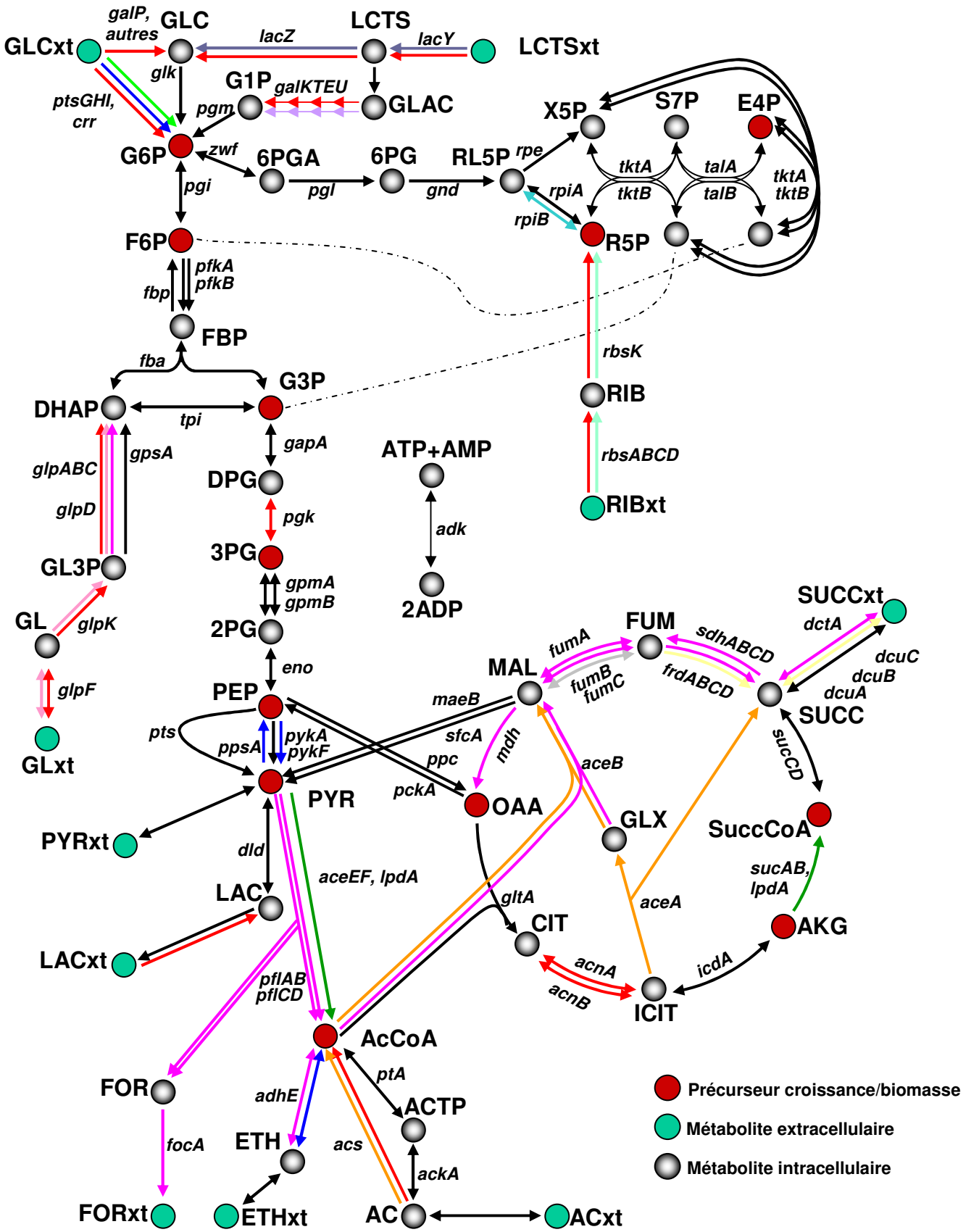


Figure 38. Schéma combiné du métabolisme central du carbone et de sa régulation transcriptionnelle chez *Escherichia coli*.

D'après Covert et Palsson, 2002.

Protéines régulatrices

- ArcA, FNR ou O_2
- Cra
- CRP
- DcuR
- GalRS
- GlpR
- IclR
- LacI
- Mlc
- PdhR
- RbsR
- RpiR
- Radicaux superoxide
- Non régulé

Présentation du travail

Lorsqu'il a été question de déterminer les objectifs de cette thèse, nous avons souhaité mettre en place deux types d'approches, l'une visant à étudier des régulations spécifiques de l'expression génique et l'autre axée sur une étude globale du transcriptome de *Bacillus subtilis*.

L'analyse complète du génome a révélé que *Bacillus subtilis* possède deux gènes, *gapA* et *gapB*, codant tous deux des glycéraldéhyde-3-phosphate déshydrogénases. Les études génétiques et biochimiques récentes du laboratoire ont démontré que ces deux enzymes étaient spécialisées de par la régulation de leur synthèse et de par leurs caractéristiques enzymatiques. GapA intervient lors de la glycolyse tandis que GapB est requise lors de la néoglucogenèse (Fillinger *et al.*, 2000). Pour mieux appréhender cette dualité originale dans le métabolisme central carboné, le laboratoire a entrepris d'élucider les mécanismes de régulation de l'expression de *gapA* et de *gapB*. A cet égard, les résultats acquis au cours de cette thèse révèlent de nouveaux aspects de la famille de régulateurs SorC/DeoR : CggR est un répresseur dédié au contrôle de la stimulation de la transcription des gènes de la partie centrale de la glycolyse chez *Bacillus subtilis* mais aussi probablement chez les bactéries proches et son activité est modulée par le fructose-1,6-bisphosphate (Article n°1 : Doan et Aymerich, 2003).

Deux paires de gènes paralogues (*ywkA* et *malS*, *ytsJ* et *mleA*) ont été identifiés dans le génome de *Bacillus subtilis* dont les produits sont homologues à des enzymes maliques. L'analyse transcriptomique globale d'une souche sauvage en réponse à la croissance en présence de glucose ou de malate comme seule source de carbone montre que l'expression d'*ywkA* est induite en présence de malate. En collaboration avec l'équipe de Yasutaro Fujita (Fukuyama University, Japon), nous avons montré qu'*ywkA* code bien une enzyme malique NAD-dépendante, et dont l'expression est spécifiquement induite par le malate et insensible à la répression catabolique. De plus, nous avons montré que le système à deux composants YufL-YufM régule directement l'expression d'*ywkA* en réponse à la présence de malate extracellulaire. Mais l'expression d'*ywkA* n'est pas requise pour la croissance en présence de malate comme seule source de carbone. En revanche, l'expression d'*ytsJ*, qui code une enzyme malique NADP-dépendante, est nécessaire à la croissance rapide dans ces mêmes conditions.

En parallèle, il était nécessaire de développer une méthode d'analyse globale du comportement de l'ensemble des gènes pour répondre à deux types de questions. *Primo*, comment se comportent les gènes impliqués dans le métabolisme central du carbone (mais

aussi plus généralement l'ensemble des gènes de *Bacillus subtilis*) en réponse à la croissance en présence de différentes catégories de sources de carbone ? *Secundo*, quels sont les effets directs et indirects de la dérégulation de l'expression d'un gène sur l'expression des autres gènes ? Le but principal de ces deux questions est de découvrir de nouveaux gènes, catalytiques ou régulateurs, impliqués dans le métabolisme carboné, ou de déterminer le régulon d'un régulateur transcriptionnel, au sens strict et au sens physiologique. Dans un premier temps, mon travail a donc consisté à apprendre la technique d'analyse du transcriptome au moyen de membranes à haute densité dans l'équipe de Michael Hecker (Greifswald Universität, Allemagne) et par la suite à la mettre en place au laboratoire. Cet objectif est maintenant atteint. Dans un deuxième temps, l'outil a été utilisé pour l'étude du régulon CggR, dont les résultats sont publiés. Par ailleurs, nous avons entamé l'étude comparative de l'expression des gènes du métabolisme central du carbone en réponse à la croissance en présence d'une source de carbone glycolytique, le glucose, et une source de carbone néoglucogénique, le succinate.

Résultats

4. CggR : un nouveau régulateur de l'expression des gènes de la partie centrale de la glycolyse

4.1. Présentation synthétique des résultats

La glycolyse est l'une des voies les mieux conservées du métabolisme. En effet, elle tient une place centrale dans le métabolisme du carbone puisqu'elle permet, à travers le catabolisme d'un très grand nombre de sucres, de produire des précurseurs indispensables pour l'anabolisme et de l'énergie sous forme d'ATP et de pouvoir réducteur. Les gènes *gapA*, *pgk, tpi, pgm et eno* codent les enzymes de la partie centrale de la glycolyse. Chez *Bacillus subtilis*, ils sont physiquement liés et sont cotranscrits. L'expression de *gapA*, basale lors de la néoglucogenèse, est stimulée environ 5 fois lorsque les bactéries sont cultivées en présence d'une source de carbone glycolytique. Les recherches de similarités indiquent que le produit du gène *yvbQ*, situé en amont de *gapA* et faisant partie du même opéron, appartient à la famille SorC/DeoR. C'est une famille de régulateurs transcriptionnels encore peu étudiée. L'analyse génétique a démontré qu'YvbQ est un répresseur de l'opéron *yvbQ-gapA* dont l'activité est inhibée en conditions glycolytiques. Ainsi, il a été renommé CggR pour « **c**entral **g**lycolytic **g**enes **r**egulator ».

Dans un premier temps, des expériences de cartographie par délétion ont permis de délimiter une région de 133 pb en amont du premier codon de *cggR* qui contient tous les éléments nécessaires à l'expression régulée de l'opéron *gapA*. Le site de démarrage de la transcription a également été déterminé 94 pb en amont de l'ATG de *cggR*. Une analyse par délétion a permis de localiser l'opérateur de CggR entre le promoteur et le premier codon de *cggR*. En parallèle, une construction *ad hoc* a été réalisée pour surexprimer et purifier la protéine CggR liée en N-terminal à un motif de 6 histidines (His_6-CggR). Nous avons vérifié que *his_6-cggR* complémentait bien une délétion de *cggR in vivo*. Des expériences de retard de migration électrophorétique ont démontré que His_6-CggR avait la capacité de se lier spécifiquement à la séquence d'ADN de 133 pb précédemment déterminée. Puis, des expériences d'empreinte à la DNase I ont permis de démontrer que CggR protégeait une séquence de 47 pb située entre le promoteur et l'ATG de *cggR* sur le brin codant et une région plus courte, incluse dans la première, sur le brin non codant. Une étude plus approfondie de l'organisation de la région du gène *gapA*, bien conservée chez les bactéries à Gram positif

riches en A-T, montre qu'il existe toujours un homologue de *cggR* en amont de *gapA*. Nous avons donc recherché des motifs conservés entre la région opératrice de CggR de *Bacillus subtilis* et la région régulatrice de l'expression de l'opéron *gapA* chez d'autres bactéries à Gram positif. Cette comparaison a permis de souligner la conservation de deux motifs répétés une fois ($CGGGACN_6TGTCN_4CGGGACN_6TGTC$) dans la région protégée par CggR et aussi d'une région riche en Thymines juste en amont de cette empreinte. Des tests génétiques après mutagenèse dirigée de la région protégée par CggR ont permis de démontrer le rôle de ces motifs dans la répression de l'opéron *gapA* en néoglucogenèse. La région riche en Thymines est aussi nécessaire *in vivo* pour réguler l'expression de *gapA*. Cependant, des tests biochimiques *in vitro* ont montré que cette répétition de Thymines n'était pas nécessaire à l'interaction ADN-CggR suggérant qu'elle pourrait avoir un rôle indirect, par exemple de courbure de l'ADN.

Les comparaisons de séquence n'ont pas révélé d'autres copies de la séquence d'ADN reconnue par CggR. Ainsi, l'opéron *cggR-gapA* serait la seule cible de CggR. Pour confirmer cette hypothèse et aussi pour observer quels pourraient être les effets génétiques indirects de la dérégulation de l'expression des gènes de la partie centrale de la glycolyse, nous avons comparé, au moyen de membranes à haute densité, le transcriptome d'un mutant *cggR* non polaire, avec celui d'une souche sauvage cultivée dans des conditions néoglucogéniques. Trois hybridations indépendantes ont été réalisées. Le seul gène dont la transcription est significativement affectée dans les trois expériences par la délétion de *cggR* est *gapA* (et les autres gènes de l'opéron dans une moindre mesure).

Quel est le signal qui module l'activité de CggR ? Un travail génétique préliminaire nous a permis de montrer que l'activité phosphofructokinase (Pfk) était nécessaire à la formation du signal perçu par CggR mais que l'activité de la glycéraldéhyde-3-phosphate déshydrogénase (GapA) ne l'était pas. Par déduction, le G3P, le DHAP et le FBP sont les trois métabolites candidats dont nous avons testé, *in vitro*, la capacité d'interaction avec CggR. Plusieurs techniques (gel retard, empreinte à la DNase I, résonance plasmonique de surface, anisotropie de fluorescence) ont conduits aux mêmes résultats. Ceux-ci permettent de conclure que le FBP inhibe complètement l'interaction CggR-ADN à une concentration de 10 mM. Il est donc intéressant de remarquer que la variation de la concentration intracellulaire de FBP constitue le signal commun qui régule la répression catabolique médiée par CcpA et l'activation des gènes de la seconde partie de la glycolyse.

L'arbre phylogénétique des régulateurs de la famille SorC/DeoR fait apparaître une sous-famille bien distincte composée de CggR de *Bacillus subtilis* et des protéines très similaires à CggR. Les gènes qui codent ces homologues de CggR sont toujours liés à *gapA* et très souvent aux autres gènes centraux de la glycolyse. On peut donc penser que CggR et les régulateurs de la sous-famille qu'il définit sont spécifiquement dédiés au contrôle des gènes de la glycolyse. Motivés par le fait que la famille SorC/DeoR de régulateurs transcriptionnels est encore très peu étudiée, nous avons entamé une collaboration avec des collègues structuralistes du CBS de Montpellier afin de mieux caractériser CggR biochimiquement et biophysiquement. Les résultats préliminaires que j'ai obtenus seront présentés et discutés.

4.2. Article n°1 : Regulation of the central glycolytic genes in *Bacillus subtilis*: binding of the repressor CggR to its single DNA target sequence is modulated by fructose-1,6-bisphosphate.

4.2.1. Article

Molecular Microbiology (2003) **47**(6), 1709–1721

Regulation of the central glycolytic genes in *Bacillus subtilis*: binding of the repressor CggR to its single DNA target sequence is modulated by fructose-1,6-bisphosphate

Thierry Doan and Stéphane Aymerich*
Génétique Moléculaire et Cellulaire, INRA (UMR216) CNRS (URA1925) and INAP-G, F-78850 Thiverval-Grignon, France.

Summary

Glycolysis is one of the best and widely conserved general metabolic pathways. *Bacillus subtilis* enzymes catalysing the central part of glycolysis, gathering the steps of interconversion of the triose phosphates from dihydroxyacetone-phosphate to phosphoenolpyruvate, are encoded by five genes, *gapA*, *pgk*, *tpi*, *pgm* and *eno*. They are transcribed in a hexacistronic operon together with *cggR*, the first cistron, encoding the repressor of this *gapA* operon. Using deletion analysis, we have localized the CggR operator between the promoter and the first gene of the operon. CggR was purified and used in gel mobility shift assays and DNase I footprinting experiments to delimit its target sequence. Site-directed mutagenesis and *in vivo* tests demonstrated that it consists of two direct-repeats (CGGGACN$_6$TGTCN$_4$CGGGACN$_6$TGTC). Sequence analysis and transcriptome comparison of a wild-type and a *cggR* mutant strain strongly suggested that CggR regulates only the *gapA* operon. The presence of glycolytic carbon sources induces expression of the *gapA* operon. Genetic experiments allowed us to identify the metabolic steps required for the formation of the CggR effector. *In vitro* experiments with the suggested candidates allowed us to demonstrate that fructose-1,6-biphosphate (FBP) acts as an inhibitor of CggR DNA-binding activity (10 mM for full inhibition). FBP is thus the major signal for both CcpA-dependent catabolite repression (or activation) and activation of the central glycolytic genes. Genomic sequence comparisons suggest that these results can apply to numerous low-G+C, Gram-positive bacterial species.

Accepted 29 November, 2002. *For correspondence. E-mail stef@grignon.inra.fr; Tel. (+33) 1 30 81 54 49; Fax (+33) 1 30 81 54 57.

© 2003 Blackwell Publishing Ltd

Introduction

A large number of carbohydrates can be used by *Bacillus subtilis* as single sources of carbon and energy for growth. Substrate-specific enzymes are required for the uptake of these carbohydrates and the generation of phosphorylated metabolic intermediates (Stülke and Hillen, 2000; Deutscher *et al*., 2002). These intermediates then enter into a few general metabolic pathways, glycolysis, pentose-phosphate shunt and Krebs cycle, to provide energy and precursors for anabolism (Fortnagel, 1993; De Wulf, 1998; Sonenshein, 2002). The second part of the glycolysis, i.e. the five steps corresponding to the interconversion of dihydroxyacetone-phosphate and phosphoenolpyruvate, plays a central role in the metabolic network not only because it yields ATP, reducing power and important metabolic precursors during carbohydrate catabolism, but also because it is the only route that links the intermediates of the initial part of the glycolysis and of the pentose-phosphate shunt to those of the Krebs cycle. Thus, these central enzymatic steps are necessary not only for sugar breakdown into the organic acids of the Krebs cycle but also in the reverse direction, when *B. subtilis* is using succinate, malate or any other gluconeogenic compound as sole source of carbon and energy, ultimately to provide the cell with glucose-6-phosphate, ribose-5-phosphate and glycerol-3-phosphate, which are required for biosynthesis of diverse macromolecules.

The analysis of the complete sequence of the *B. subtilis* genome has revealed that the genes encoding the enzymes that catalyse the five steps of the central carbon metabolism, *tpi* (triose phosphate isomerase), *gap* (glyceraldehyde-3-phosphate dehydrogenase), *pgk* (phosphoglycerate kinase), *pgm* (phosphoglycerate mutase) and *eno* (enolase), are physically linked on the chromosome (Leyva-Vazquez and Setlow, 1994; Kunst *et al*., 1997). Moreover, it has also revealed the presence of a second gene, which putatively encodes a glyceraldehyde-3-phosphate dehydrogenase (GAPDH), *gapB*, very similar in sequence but physically unlinked to the other *gap* gene. In a previous study (Fillinger *et al*., 2000), we have demonstrated that both these genes, *gap* (renamed *gapA*) and *gapB*, indeed encode functional GAPDH. Interestingly,

these two enzymes have opposite physiological roles: GapA is necessary for growth under glycolytic conditions, whereas GapB is required for growth on gluconeogenic carbon sources. This specialization is first achieved through different cofactor specificity of GapA and GapB enzymes: GapA is a canonical, strictly NAD-dependent GAPDH enzyme, whereas GapB has a 50-fold higher catalytic efficiency with NADP than with NAD. Secondly, *gapA* and *gapB* are submitted to opposite regulation: *gapA* is transcribed at a basal level under gluconeogenic conditions of growth, and its transcription is enhanced in the presence of glucose or other glycolytic carbon sources; in contrast, *gapB* transcription is repressed in the presence of glycolytic carbon sources and strongly induced under gluconeogenic conditions of growth (Fillinger *et al*., 2000).

gapA is co-transcribed with the downstream *pgk*, *tpi*, *pgm* and *eno* genes and also with *yvbQ*, the upstream gene with a product that shares sequence similarities with the members of the SorC family of transcriptional regulators (Ludwig *et al*., 2001). Our genetic analysis has demonstrated that the *yvbQ* gene product is a regulator of the *gapA* operon, acting as a transcriptional repressor. The *yvbQ* gene has thus been renamed *cggR* for central glycolytic genes regulator (Fillinger *et al*., 2000). The role of *cggR* has been confirmed recently by Ludwig *et al*. (2001).

Stimulation of the synthesis of the enzymes of the central carbon metabolism in the presence of glucose has also been documented through a global analysis of the proteome combined with transcriptional analysis (Tobisch *et al*., 1999). Furthermore, Tobisch and coworkers have demonstrated that CcpA, the major pleiotropic control protein for catabolite regulation of genes in *B. subtilis* (Hueck and Hillen, 1995; Henkin, 1996; Deutscher *et al*., 2002), is necessary for that stimulation. Our analysis has shown that CcpA is necessary for the stimulation of *gapA* transcription by glucose, but dispensable for its stimulation in the presence of glycerol, which is not transported by the phosphotransferase system (PTS) (Deutscher *et al*., 2002), and that a deletion of *cggR* is epistatic over a deletion of *ccpA*. Altogether, these results suggested that CcpA has an indirect effect mediated via CggR on the regulation of the *gapA* operon (Fillinger *et al*., 2000). More recently, the indirect role of CcpA in the induction mechanism of the *gapA* operon expression by a PTS-dependent sugar has been described (Ludwig *et al*., 2002).

Stülke and coworkers have systematically characterized the transcriptional pattern of all glycolytic genes and operons and have shown that only the *gapA* operon is strongly regulated (Ludwig *et al*., 2001). Furthermore, in addition to the transcription initiation controlled by CggR, these authors have demonstrated that an mRNA processing event, between *cggR* and *gapA*, and the presence of an additional internal constitutive promoter between *gapA* and *pgk* lead to a stronger regulation of *gapA* expression compared with that of the downstream genes (Ludwig *et al*., 2001). Indeed, although the expression of these downstream genes, *pgk*, *tpi*, *pgm* and *eno*, is necessary for both glycolysis and gluconeogenesis, the expression of *gapA* is required only for glycolysis; under gluconeogenic conditions, the conversion of diphosphoglycerate into glyceraldehyde-3-phosphate is catalysed by the product of the *gapB* gene (Fillinger *et al*., 2000).

The prototype of the SorC family of regulators is the SorC protein of *Klebsiella pneumoniae*, which has been demonstrated to be a transcriptional repressor of the genes involved in L-sorbose uptake and catabolism (Wohrl *et al*., 1990). The regulatory role of some other members of this family, including CggR, has been demonstrated experimentally (Heuel *et al*., 1998; Bekal-Si Ali *et al*., 1999; Sangari *et al*., 2000; Yebra *et al*., 2000), but the target sequence has been identified for only one of them, DeoR of *B. subtilis*, a repressor of the *dra–nupC–pdp* operon involved in deoxyribonucleoside and deoxyribose utilization (Saxild *et al*., 1996; Zeng and Saxild, 1999; Zeng *et al*., 2000).

Here, we describe the repressor activity of CggR through the determination of its target sequence and the identification of the effector that inhibits its binding activity. Furthermore, based on sequence analysis and transcriptome comparison, we propose that CggR and its close homologues found in other low-G+C, Gram-positive species are specifically dedicated to the regulation of the central glycolytic genes.

Results

Deletion analysis of the cggR–gapA *regulatory region*

In a previous report, we showed that a 295 bp DNA fragment upstream of the initiation codon of *cggR* contains the *cis*-acting elements for the regulated expression of *cggR–gapA* genes (Fillinger *et al*., 2000). To determine the minimal region necessary and sufficient for regulated expression, a deletion analysis was performed. Different fragments of the regulatory region were fused to the *lacZ* reporter gene, and the resulting constructs were inserted at the *amyE* locus into the chromosome (Fig. 1A); the β-galactosidase activity present in the resulting strains grown in minimal medium with a gluconeogenic (proline) or a glycolytic (glycerol) carbon source was measured (Fig. 1B). *lacZ* expression was stimulated under glycolytic conditions of growth in strains GM1511, GTD002 and GTD003, whereas it was weak and not stimulated under glycolytic conditions in strain GTD004. In contrast, *lacZ* expression was very high and constitutive in strains GTD005, GTD006, GTD007 and GTD008. Similar results

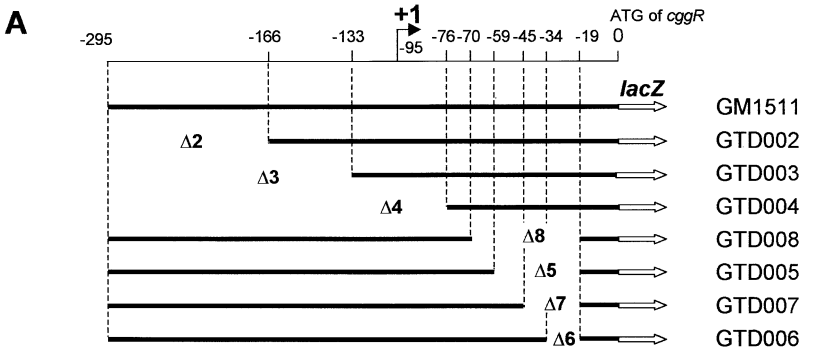


Fig. 1. Deletion analysis of the control region of the *gapA* operon.
A. Limits of the deletions introduced in the P*gapA–lacZ* reporter construct and corresponding strains. Position of the transcription start point determined by Stülke and coworkers (Ludwig *et al.*, 2001) is indicated.
B. Effect of the deletions on the activity of the promoter region of the *gapA* operon as measured by the β-galactosidase activity present in the reporter strains. Strains were cultivated in proline or glycerol CQT media. Activity was assayed at least three times from independent clones; representative results from one series are shown.

B

Strain	Relevant genotype	β-galactosidase activity (U/mg of protein)		Ratio
		proline	glycerol	gly/pro
GM1511	*amyE′::PgapA-lacZ*	57	304	5.3
GTD002	*amyE′::PgapA* $_{\Delta 2}$*-lacZ*	36	299	8.3
GTD003	*amyE′::PgapA* $_{\Delta 3}$*-lacZ*	34	292	8.6
GTD004	*amyE′::PgapA* $_{\Delta 4}$*-lacZ*	38	30	0.8
GTD008	*amyE′::PgapA* $_{\Delta 8}$*-lacZ*	695	795	1.1
GTD005	*amyE′::PgapA* $_{\Delta 5}$*-lacZ*	1170	1185	1
GTD007	*amyE′::PgapA* $_{\Delta 7}$*-lacZ*	991	1162	1.2
GTD006	*amyE′::PgapA* $_{\Delta 6}$*-lacZ*	580	652	1.1

were obtained when succinate plus glutamate or glucose was used instead of proline or glycerol respectively (data not shown). These results indicated that the 133 bp region located upstream of the *cggR* initiation codon contains the *cis*-acting elements necessary for the regulated transcription of the *gapA* operon and that at least part of the CggR operator would be located between positions –34 and –19 upstream of the ATG. They also suggested that promoter sequences are confined between –133 and –76 upstream of the ATG. A primer extension experiment was performed on RNA extracted from the wild-type strain grown on glucose minimal medium, which showed that transcription initiated at an A or a G located at position –95 and –94, respectively (data not shown), in agreement with the result published meanwhile by Ludwig *et al.* (2001).

Specific binding of purified CggR to the gapA *promoter region*

An electrophoretic gel mobility shift assay (EMSA) was used to test CggR capacity to bind specifically to the minimal DNA regulatory sequence defined by deletion analysis. N-terminal His-tagged CggR was overexpressed in *Escherichia coli* and purified to apparent homogeneity (data not shown). The functionality of the His-tagged variant of CggR was verified through the capacity of its encoding gene to complement a *cggR* deletion *in vivo* (Table 1).

Table 1. Complementation of *cggR* disruption by a *his*$_6$*-cggR* variant gene encoding a His-tagged CggR protein.

Strain	Relevant genotype	β-galactosidase activity[a] (U mg^{-1} protein)		Ratio
		glc	succ + glut	glc/succ + glut
GM1511	*amyE′::*P*gapA-lacZ-cat*	220	50	4.4
GM1521	*cggR′::lacZ*, *amyE′::cat*	1127	1558	0.7
GM1520	*cggR′::lacZ*, *amyE′::*P*gapA-cggR*$^+$*-cat*	335	52	6.4
GTD021	*cggR′::lacZ*, *amyE′::*P*gapA-his*$_6$*cggR*$^+$*-cat*	311	70	4.4

a. Activity was assayed at least three times from independent clones; representative results from one series are shown. Strains were cultivated in glucose (glc) or succinate plus glutamate (succ + glut) CQTHC media.

A DNA fragment containing the 133 bp sequence present in the reporter construct of strain GTD003 (Fig. 1A) was labelled and used as a probe. The probe was shifted by CggR in a protein concentration-dependent manner (Fig. 2A). Two shifted bands were visible at weak protein concentration, whereas only the slower migrating one was present at high protein concentration. This suggested that two complexes differing in protein–DNA stoichiometry could form. No shift was detected when using a Ni-NTA-agarose-purified extract from an *E. coli* strain containing an empty expression vector (data not shown). The His_6-CggR protein is thus able specifically to bind the *cggR–gapA* promoter region. The apparent K_d defined as the protein concentration required to shift 50% of the probe is ≈7 nM (of monomer), assuming that all CggR molecules present in the assay were active.

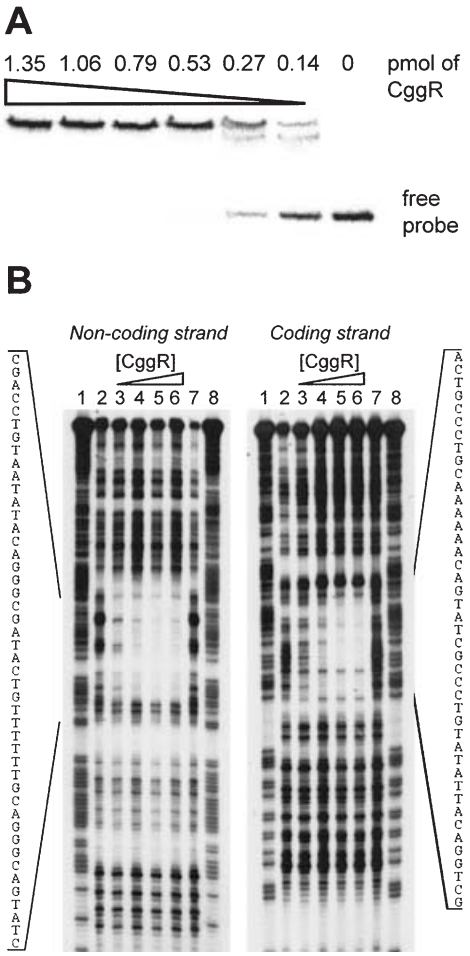


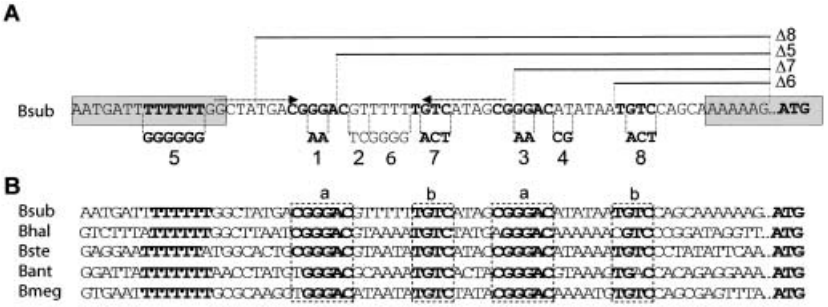
Fig. 2. A. Specific binding of His_6-CggR to the promoter region of the *gapA* operon. A DNA fragment corresponding to the minimal efficient control region (insert of pTD003, ≈0.05 pmol) was labelled and then incubated with increasing amounts of His_6-CggR in the presence of a large excess of unspecific DNA.
B. DNase I footprinting analysis of the interaction between CggR and the promoter region of the *gapA* operon examined on the coding and the non-coding strands. Approximately 0.25 pmol of labelled DNA fragment was incubated with no (lanes 2 and 7) or increasing amounts of His_6-CggR (lanes 3, 0.53 pmol, lanes 4, 1.06 pmol, lanes 5, 2.12 pmol, lanes 6, 4.25 pmol). Lanes 1 and 8, G+A sequencing ladder obtained from the same DNA fragment. Sequence of protected regions is indicated.

Footprint analysis of the interaction between CggR and DNA

To locate the CggR binding site more precisely, DNase I footprinting assays were performed. A DNA fragment encompassing the 133 bp minimal regulatory sequence was labelled on one strand or the other, incubated with an increasing quantity of CggR and then hydrolysed by DNase I. As revealed after electrophoresis (Fig. 2B), CggR protected a common 41 bp sequence on both strands and three further nucleotides located upstream of it on the non-coding strand alone. This protected area is located between the transcriptional start site determined previously (Ludwig *et al.*, 2001; T. Doan and S. Aymerich, data not shown) and the *cggR* coding sequence. Sequence examination of the region revealed an inverted repeat ($GCTATGACN_7GTCATAGC$) not completely included in the protected area and two direct repeats ($CGGGACN_6TGTCN_4CGGGACN_6TGTC$) entirely protected (Fig. 3A). The inverted repeat is not included in the deletions of the reporter constructs present in GTD006 and GTD007 (Fig. 3A) yet leading to a constitutive expression. In contrast, at least one of the direct repeats is included in all the deletions (GTD005 to GTD008) that lead to constitutive expression of the reporter gene. Furthermore, searches in the putative regulatory region of *gapA* in four other *Bacillus* species revealed that the direct repeats are almost perfectly conserved, whereas the inverted repeat is not (Fig. 3B). We also noted that a stretch of T present immediately upstream of the protected area is conserved in the corresponding region of the chromosome of other Bacilli (Fig. 3B).

CggR interacts with two direct repeats

To demonstrate that the direct repeats included in the protection area constitute the CggR binding site, the role of these sequences was tested *in vivo* after directed mutagenesis. Different mutations affecting the direct repeats or the intercalated sequences (Fig. 3A) were introduced into the reporter plasmid pSF111 (Fillinger *et al.*, 2000). The resulting constructs were inserted at the *amyE* locus into the chromosome of the wild-type *B. subtilis* strain. These transformants were grown in minimal medium with proline or glycerol as sole carbon source, and the β-galactosidase activity was measured (Fig. 3C). All the mutations affecting one or other direct repeat led to constitutive expression of the reporter gene (GTD010,



C

Strain	Relevant genotype	β-galactosidase activity (U/mg of protein) proline	glycerol	Ratio gly/pro
GM1511	*amyE'::PgapA-lacZ*	57	304	5.3
GTD010	*amyE'::PgapA* $_{1}$*-lacZ*	1189	889	0.8
GTD011	*amyE'::PgapA* $_{2}$*-lacZ*	68	277	4
GTD012	*amyE'::PgapA* $_{3}$*-lacZ*	870	604	0.7
GTD013	*amyE'::PgapA* $_{4}$*-lacZ*	91	346	4.3
GTD014	*amyE'::PgapA* $_{5}$*-lacZ*	1729	1387	0.8
GTD016	*amyE'::PgapA* $_{6}$*-lacZ*	437	1231	2.8
GTD017	*amyE'::PgapA* $_{7}$*-lacZ*	1156	1059	0.9
GTD018	*amyE'::PgapA* $_{8}$*-lacZ*	1250	813	0.7

Fig. 3. Site-directed mutagenesis analysis of the region protected by CggR to identify the sequence elements necessary for efficient interaction of CggR with DNA *in vivo*.
A. The region protected by CggR in the *gapA* promoter region as determined by DNase I footprinting is shown between the two grey zones. Strongly conserved sequence elements in the corresponding genome region of other Bacilli species are indicated in bold letters. Mutations introduced into the reporter construct are numbered from 1 to 8. An inverted repeat is overlined with two arrows. The limits of the deletions Δ5 to Δ8 present in strains GTD005 to GTD008 (see Fig. 1) are indicated by thin lines above the sequence.
B. Multiple sequence alignment of the region protected by CggR upstream of the *gapA* operon of *B. subtilis* and the corresponding one in the genome of four other *Bacillus* species (Bsub, *B. subtilis*; Bhal, *B. halodurans*; Bste, *B. stearothermophilus*; Bant, *B. anthracis*; Bmeg, *B. megaterium*). Two strongly conserved motifs, *a* and *b*, each one directly repeated once, are in bold letters and boxed. A conserved stretch of T located just upstream of the region protected by CggR is in bold letters.
C. Effect of the mutations introduced into the direct repeats a or b, into the stretch of T or into the intercalated sequences on the activity of the promoter region of the *gapA* operon as measured by β-galactosidase activity synthesized by the corresponding reporter strains. Strains were cultivated in proline or glycerol CQT media. Activity was assayed at least three times from independent clones; representative results from one series are shown.

GTD012, GTD017 and GTD018), whereas *lacZ* expression was still stimulated similarly by glycerol in the mutant strains GTD011, GTD013 and GTD016 affected in the intercalated sequences, as in the control strain GM1511. We concluded that the direct repeats are indeed necessary for the regulation of the *gapA* operon by CggR, whereas the surrounding sequences are not.

The substitution of the stretch of T present upstream of the direct repeats by a stretch of G (GTD014) also led to a constitutive phenotype (Fig. 3C). This suggested that, although clearly lying outside the area protected by CggR, this repetition of thymine nucleotides is necessary for efficient regulation by CggR *in vivo*.

Effect of cggR *inactivation on* B. subtilis *transcriptome*

Searches in the complete *B. subtilis* genome for other occurrences of the target sequence of CggR (CGGGAC N_6TGTCN_4CGGGACN_6TGTC, including 0–4 mismatches) revealed no other putative CggR target genes. However, to detect other genes with transcription possibly controlled, directly or indirectly, by CggR, we performed a comparison of the complete transcriptional pattern between a wild-type strain and a *cggR* mutant strain using DNA macroarrays. A non-polar *cggR* mutant strain (GTD030) was constructed by replacing precisely the *cggR* open reading frame (ORF) with an ORF encoding phleomycin resistance. This mutant strain and the wild-type strain were cultivated in parallel under gluconeogenic conditions until mid-log phase of growth. Total RNA was then extracted, labelled and used as probes for hybridization of a pair of DNA high-density filters on which the complete set of *B. subtilis* ORFs had been spotted. This experiment was repeated three times independently (from three different cultures and using three different batches of filters). Analysis of the results showed that the only

gene differentially transcribed significantly (signal three-fold superior to that of the background, more than three-fold difference between the two strains) in each of the three experiments was *gapA* (complete results in Table S2 in *Supplementary material*). An altered transcription of the four downstream genes of the operon was also detected, but not in every experiment and at a lower level, in accordance with a much stronger regulation of *gapA* compared with the other genes of the operon as a result of the additional levels of regulation (mRNA processing, internal promoter) reported by Ludwig *et al.* (2001).

These global analyses strongly suggest that CggR is specifically dedicated to the regulation of the *gapA* operon and that the artificial (because of *cggR* deletion) overexpression of this operon under gluconeogenic conditions of growth does not significantly affect the transcription of any other gene of *B. subtilis*.

Pfk but not GapA is necessary for the induction of gapA *expression*

We have designed genetic experiments to identify the enzymatic steps required for the formation of the signal that inhibits CggR binding activity during glycolysis. We have compared the kinetics of *gapA* induction by glucose or glycerol in a wild-type strain and *pgi*, *pfk* or *gapA* mutant strains as reported by an ectopic *PgapA–lacZ* fusion. Cells were initially grown in minimal medium with malate as the sole carbon source, and *lacZ* expression was measured after the addition of the glycolytic carbon source. As expected, in the wild-type context, *gapA* expression was induced rapidly after the switch to glycolytic conditions of growth (Fig. 4A). The stimulation was faster and stronger with glycerol than with glucose. In the *pfk* mutant (Fig. 4B), *gapA* expression was still inducible by the addition of glycerol, as in the wild type, but was no longer inducible by the addition of glucose. Similar results were obtained with the *pgi* mutant (data not shown). In contrast, the *gapA* mutant strain behaved exactly as the wild-type strain (Fig. 4C). Thus, Pfk activity is required for the formation of the CggR effector molecule from extracellular glucose, but not from extracellular glycerol, whereas GapA activity is dispensable. This suggests that fructose-1,6-bisphosphate (FBP), dihydroxyacetone phosphate (DHAP) and/or glyceraldehyde-3-phosphate (G3P), the three glycolytic intermediates between the Pfk and GapA steps (Fig. 4D), might be the effector molecule(s) of CggR.

FBP inhibits CggR binding activity

The capacity of FBP, DHAP or G3P to inhibit CggR binding activity was first tested using a gel mobility shift assay.

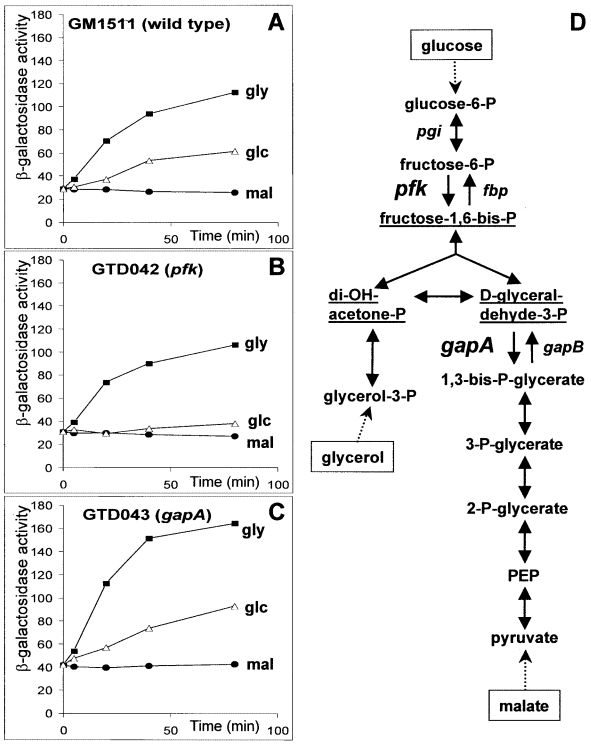


Fig. 4. Induction kinetics of the promoter of the *gapA* operon after the addition of a glycolytic carbon source (glucose or glycerol) in a culture of a *B. subtilis* wild-type strain (A), a *pfk* mutant strain (B) or a *gapA* mutant strain (C) grown under gluconeogenic conditions (CQTHC + malate).
D. Simplified schema of glycolysis and entry points of the gluconeogenic (malate) and glycolytic (glucose or glycerol) carbon sources used in this experiment. The genes encoding the enzymatic steps relevant in this experiment are indicated in italic.

This test revealed an inhibitory effect of FBP (a higher concentration of CggR is required to shift 50% of the labelled probe with FBP than without FBP; similar results were obtained when FBP was also added to the running buffer), but not of G3P or DHAP (data not shown). DNase I footprinting experiments were then used to confirm these results. CggR was preincubated with different concentrations of FBP or other molecules to be tested before adding the DNA labelled probe. These experiments showed that CggR lost its ability to protect its target sequence with increasing FBP concentration (Fig. 5), whereas high concentrations of G3P, fructose-6-phosphate or fructose-2,6-bisphosphate did not interfere with the binding activity of CggR. This demonstrated that FBP is the effector molecule that directly and negatively controls the repressor activity of CggR.

Discussion

Through this work, we identified the DNA target sequence of CggR, the repressor of the *gapA* operon of *B. subtilis*.

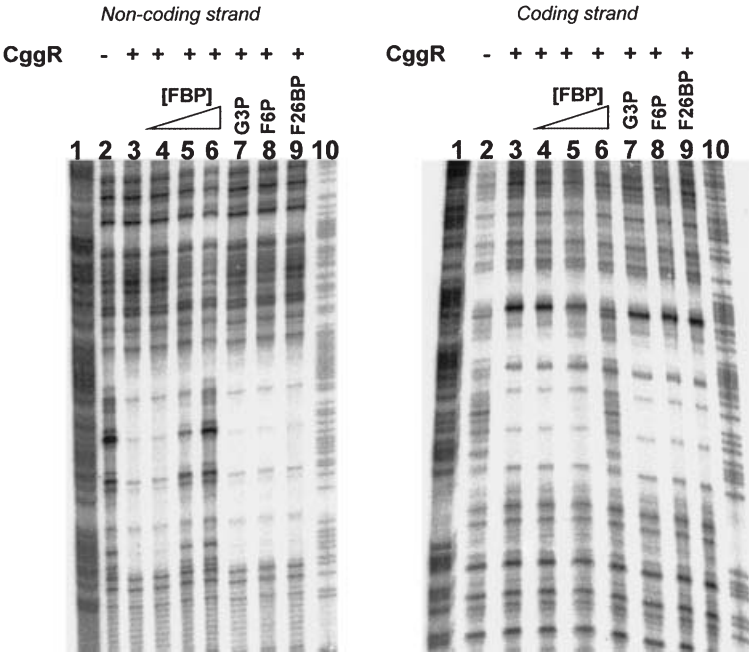


Fig. 5. Effect of fructose-1,6-biphosphate on CggR capacity to bind to the regulatory region of *cggR–gapA*. The experiment was performed as described for DNase I footprinting on the labelled non-coding and coding strands. Varying amounts of FBP were incubated with CggR and the radiolabelled DNA probe before DNase I treatment. Lanes 1 and 10, G+A sequencing; lane 2, 0.25 pmol of DNA; lanes 3–9, 0.25 pmol of DNA + 4 pmol of CggR; lane 4, +1 mM FBP; lane 5, +5 mM FBP; lane 6, +10 mM FBP; lane 7, +10 mM G3P; lane 8, +10 mM F6P; lane 9, +10 mM F26BP.

Deletion analysis, footprinting experiments and site-directed mutagenesis demonstrated that the CggR binding sequence comprises two direct repeats located between the promoter sequence and the first ORF (Fig. 3). The location of this operator sequence, 32 bp downstream of the transcriptional start point (Fig. 1A), strongly suggests that CggR would act as a roadblock to elongation of the leader transcript by the transcriptional machinery, and not at the level of the initiation of the transcription. Sequence analysis of the regulatory region of the *gapA* operon in four other Bacilli showed that these two direct repeats ($CGGGACN_6TGTCN_4CGGGACN_6TGTC$) are perfectly conserved in *Bacillus stearothermophilus* and nearly perfectly conserved in *Bacillus halodurans*, *Bacillus megaterium* and *Bacillus anthracis*, whereas the immediate surrounding sequences are less conserved among these bacterial species (Fig. 3B). The analysis of the interaction between CggR and its target DNA fragment using gel shift assays suggested that CggR would bind the DNA co-operatively to form two different complexes differing in protein–DNA stoichiometry as the protein concentration increases (Fig. 2A). A preliminary biochemical characterization of CggR, through glutaraldehyde cross-linking and gel filtration experiments, suggested that it exists in solution as an equilibrium between a dimeric form and at least one higher oligomeric form (T. Doan, S. Aymerich and N. Declerck, unpublished data). These different oligomeric forms could correspond to the different complexes observed in gel shift assays as the protein concentration increases. The formation of a protein–DNA complex containing more than two monomers of CggR would explain why the operator sequence identified for CggR is longer and more complex than those recognized by most prokaryotic transcription factors. Sequence comparisons revealed that the N-terminal part of CggR contains a DNA-binding motif belonging to the winged helix (WH) family (Brennan, 1993; Gajiwala and Burley, 2000; Huffman and Brennan, 2002). The tridimensional structure of the DNA–protein complex corresponding to some WH-containing transcription factors has been solved and has revealed a clear variability in the detailed mechanisms of DNA recognition as well as in the DNA sequence elements involved in direct contacts with the WH motif (Pohl *et al.*, 1999; Chen *et al.*, 2000; van Aalten *et al.*, 2001; Heldwein and Brennan, 2001). It is therefore impossible to predict from these structural data which particular elements in the CggR operator are directly involved in DNA–protein interaction. Nevertheless, it can be noted that the operator sequence that we identified contains two identical small inverted repeats, the four elements of which are regularly spaced and separated one from one another by one double-helix turn ($GACN_7GTCN_7GACN_7GTC$). A central role of these four triplets would be consistent with our *in vivo* and *in vitro* results. Biochemical and biophysical experiments aimed at the elucidation of the CggR–DNA mechanism of interaction have been undertaken.

The comparison of the sequence of the regulatory region of the *gapA* operon in *B. subtilis* and four other Bacilli species had also revealed a very conserved stretch of T located 8 nucleotides upstream of the operator sequence. Site-directed mutagenesis indicated that these

T residues are necessary for efficient repression by CggR in the absence of glycolytic sugars, but footprinting experiments clearly demonstrated that they lie outside the region protected by CggR. Competition experiments in gel shift assays showed that a DNA fragment containing only the two direct repeats titrates the labelled probe as efficiently as a longer fragment including the direct repeats and the stretch of T (T. Doan and S. Aymerich, unpublished results). Altogether, these results suggest that the stretch of T would be necessary only in the context of the entire chromosome to generate an environment, probably structural, allowing the efficient recognition of the direct repeats by CggR. However, it cannot be ruled out that this stretch of T serves another regulatory purpose.

Sequence analyses of the entire CggR protein showed that it belongs to the SorC family of DNA-binding regulatory proteins. Genes corresponding to 75 members of this family can be recognized unambiguously in the databases in complete and unfinished prokaryote genomes. A phylogenetic tree of these 75 proteins (Fig. S1 in *Supplementary material*) shows that CggR defines a clearly distinct subfamily of the SorC family of proteins including 16 other proteins. To the extent of our knowledge, the biological function of none of these 16 proteins, all encoded by low-G+C, Gram-positive bacterial species, has been demonstrated experimentally. However, careful examination of the sequence environment of their corresponding gene in the chromosome revealed that all (except two for which the information is not available) are the first gene of a putative operon containing at least three central glycolytic genes, *gapA*, *pgk* and *tpi*, and most often also a *pgm* and an *eno* gene. This strongly suggests that the CggR-like protein encoded in these different species, the four other Bacilli already mentioned, *Listeria monocytogenes* and *Listeria innocua*, *Clostridium acetobutylicum*, *Clostridium botulinum*, *Clostridium perfringens* and *Clostridium difficile*, *Enterococcus faecalis*, *Staphylococcus aureus* and *Staphylococcus epidermis*, *Lactobacillus delbrueckii* and *Thermoanaerobacter tengcongensis* are transcriptional repressors of the operon coding for the central steps of glycolysis as in their *B. subtilis* counterpart. The high degree of sequence similarity of these members of the CggR subfamily could also suggest that the activity of each is modulated by the same effector molecule as in *B. subtilis*.

The transcription of the *gapA* operon is derepressed when a glycolytically metabolizable carbohydrate is present in the culture medium (Tobisch *et al*., 1999; Fillinger *et al*., 2000). Ludwig *et al*. (2001) have shown that the level of derepression varies depending on the carbohydrate, but that either carbohydrates catabolized via the upper part of glycolysis, such as glucose, glucitol, fructose or glycerol, or carbohydrates catabolized via the pentose phosphate pathway, such as gluconate or arabinose, did induce the *gapA* operon. In order to identify the signal molecule that inhibits the activity of CggR under glycolytic conditions of growth, we first used a genetic approach (Fig. 4). Then, *in vitro* experiments were performed that revealed that only FBP inhibits CggR DNA-binding activity (Fig. 5); neither DHAP, G3P, F6P nor fructose-2,6-biphosphate had any effect. In these experiments, FBP had no detectable effect at 1 mM and a nearly total effect at 10 mM (Fig. 5); experiments using surface plasmon resonance and fluorescence anisotropy are currently under way to quantify the parameters of the interaction between FBP and CggR. These results are fully consistent with recent results obtained by Mijakovic *et al*. (2002) that showed that the FBP concentration in exponentially growing *B. subtilis* cells may vary from 1.4 mM in the presence of succinate to 14 mM in the presence of glucose as major carbon source, whereas the concentration of the triose-P pool, [DHAP] + [G3P], remains the same in succinate and succinate plus glucose media. Our results also showed that *gapA* induction was faster and stronger when glycerol was added compared with glucose (Fig. 4), suggesting that the FBP concentration increases more rapidly and reaches a higher level after the addition of glycerol compared with glucose. However, Mijakovic *et al*. (2002), determined that the FBP concentration is only 6.2 mM in a *B. subtilis* cell growing on glycerol compared with 14 mM on glucose. We have no definitive explanation for this discrepancy. Nevertheless, it should be mentioned that the culture medium used by these authors differs notably from the one we used (the former contains 1.5% casitone, vitamins and nucleotides, whereas the latter does not). Furthermore, the measures were performed by Mijakovic *et al*. (2002) on steady-state growing cultures, whereas we followed *gapA* induction after a nutrient shift (and no equilibrium was reached, as indicated by the increase in the β-galactosidase activity throughout the experiments). On the other hand, Ludwig *et al*. (2001) found that a higher level of derepression of *gapA* transcription was achieved when *B. subtilis* was growing in a minimal medium containing glucose and amino acids rather than with glucose alone. These authors suggested that CggR could respond synergistically to two different signals, a catabolic one derived from the presence of sugars, and an anabolic one derived from amino acid metabolism (Ludwig *et al*., 2001). Another possibility would be that the decrease in the activity of the TCA cycle in response to the presence of glucose and amino acids (Hederstedt, 1993) would lead to an increase in the glycolytic flux to generate the high quantity of chemical energy necessary for the very fast growth observed with glucose plus amino acids; a higher FBP concentration would correspond to this higher glycolytic flux, leading to stronger CggR inhibition. It will be possible to test this hypothesis by com-

paring the intracellular concentration of FBP when glucose alone or glucose plus amino acids are present as carbon sources.

FBP is also the major signalling molecule for CcpA-dependent catabolite repression and catabolite activation of genes in *B. subtilis* (Stülke and Hillen, 2000; Deutscher *et al*., 2002). An increase in the FBP concentration stimulates the kinase activity of HprK, which phosphorylates the protein HPr on Ser-46; the seryl-phosphorylated HPr (P-Ser-HPr) then interacts with CcpA, which is active for catabolite repression and catabolite activation only in complex with P-Ser-HPr. Furthermore, FBP could also act directly in favouring the formation of the P-Ser-HPr–CcpA complex (Deutscher *et al*., 1995; Aung-Hilbrich *et al*., 2002). Bioinformatic searches for the presence of the DNA sequence specifically recognized by CcpA and global proteomic and transcriptomic analyses have suggested that the CcpA regulon includes a very large number of genes representing about 10% of the genome content (Miwa *et al*., 2000; Moreno *et al*., 2001; Yoshida *et al*., 2001; Deutscher *et al*., 2002). Our findings that FBP is the effector molecule for CggR-dependent activation of the central glycolytic genes clearly further supports the concept that FBP is the key signal molecule for the regulation of carbon metabolism. It also provides an integration point between both CcpA and CggR regulatory networks. A second link exists between these two networks (Tobisch *et al*., 1999; Fillinger *et al*., 2000; Ludwig *et al*., 2001) and has been elucidated recently by Stülke and coworkers: *ccpA* is necessary for the uptake of the PTS sugars and thus for the formation of FBP in the presence of this class of glycolytically metabolizable carbohydrates (Ludwig *et al*., 2002). Indeed, the kinase activity of HprK is increased in a *ccpA* mutant strain to such an extent that nearly all HPr molecules are phosphorylated on Ser-46; no more HPr is thus available for Pi transfer from PEP to the sugars to be taken up via the PTS Enzyme II permeases. Apart from the catabolite repression of genes by CcpA, it appears that *B. subtilis* has two different regulatory pathways using the same direct effector molecule for activation of two distinct subsets of genes in response to the great availability of glycolytic substrates: the CcpA-dependent activation of genes involved in overflow metabolism, *ackA*, *pta*, *alsSD* (Turinsky *et al*., 1998; 2000; Presecan-Siedel *et al*., 1999; Shin *et al*., 1999; Tobisch *et al*., 1999) and putatively *ilvB*, and *lctE* (Moreno *et al*., 2001) and the CggR-dependent derepression of the central glycolytic genes, *gapA*, *pgk*, *tpi*, *pgm* and *eno*.

Genes required for interconversion of hexose phosphates such as FBP and subsequent generation of the triose phosphates G3P and DHAP, *pgi*, *pfk* and *fba*, as well as the *pykA* gene encoding the pyruvate kinase, are constitutively expressed or very weakly induced by glucose (Tobisch *et al*., 1999; Ludwig *et al*., 2001). Therefore, the genes of the *gapA* operon are the only glycolytic genes strongly regulated by a glycolytic carbon source. Moreover, our sequence analysis and our global transcriptomic comparison between a wild-type and a *cggR* mutant strain did not reveal other genes putatively regulated by CggR. This regulatory protein thus appears to be specifically dedicated to the regulation of the central glycolytic genes.

The expression of the genes of the *gapA* operon is not only regulated at the transcriptional level via CggR but also at the post-transcriptional level (Ludwig *et al*., 2001): the transcript originating upstream of *cggR* is processed via an endonucleolytic cleavage that leads to differential stability of the monocistronic *cggR* message and of the mRNA encoding the glycolytic enzymes. Furthermore, an additional weak constitutive promoter is present upstream of *pgk* (Ludwig *et al*., 2001). Altogether, these different regulatory mechanisms result in a stronger regulation of the *gapA* gene than the other genes of the operon and probably allow a fine tuning of the expression of these glycolytic genes that have a central role in carbon metabolism in *B. subtilis*.

Experimental procedures

Bacterial strains, growth conditions and genetic procedures

Escherichia coli TG1 [K12 Δ(*lac-pro*) *supE thi hsd5*/F′ *traD36 proA*$^+$*B*$^+$ *lacI*^{q} *lacZ*Δ*M15*] was used as a general cloning host, and *E. coli* strain M15 pREP4 (Qiagen) was used for overproduction of His-CggR. *E. coli* strains were grown in Luria–Bertani broth (LB) supplemented with antibiotics when necessary (ampicillin 100 mg l^{-1}; kanamycin 25 mg l^{-1}). Conventional calcium shock or electroporation procedures were used for transformation (Sambrook *et al*., 1989).

The *B. subtilis* strains used are listed in Table 2. The standard procedure was used for *B. subtilis* transformation (Anagnostopoulos and Spizizen, 1961). *B. subtilis* strains were grown in LB, except when stated otherwise. Antibiotics for selection were added at 5 mg l^{-1} (chloramphenicol), 0.4 mg l^{-1} (erythromycin), 100 mg l^{-1} (spectinomycin) or 0.25 mg l^{-1} (phleomycin). All cultures were performed at 37°C. Cultures for β-galactosidase assays were performed in CQTHC or CQT minimal medium [C mineral medium (Aymerich *et al*., 1986) supplemented with 0.005% tryptophan, 0.15% glutamine with or without, respectively, 0.05% casein hydrolysate] supplemented with 1% glucose or 1% glycerol (v/v) or 1% proline. Cultures for RNA preparation for transcriptome analysis or for primer extension were performed in TSS minimal medium (Fouet and Sonenshein, 1990) (50 mM Tris-HCl, pH 7.5, 15 mM $(NH4)_2SO_4$, 8 mM $MgSO_4{\cdot}7H_2O$, 27 mM KCl, 7 mM trisodium citrate, 0.6 mM KH_3PO_4, 2 mM $CaCl_2$, trace $FeSO_4{\cdot}H2O$, trace $MnSO_4$, 5 mM glutamate) supplemented with succinate plus glutamate (1%:1%) or 1% glucose.

Table 2. *B. subtilis* strains used in this study.

Strain	Genotype or relevant genotype	Source
168CA	*trpC2*	Laboratory stock
BFA1080	*cggR′*::pMUTIN2	Fillinger *et al.* (2000)
GM 1511	*amyE′*::P*gapA′::lacZ-cat*	pSF111 → 168CA Fillinger *et al.* (2000)
GTD002	*amyE′*::P*gapA*$_{\Delta 2}$*::lacZ-cat*	pTD003 → 168CA
GTD003	*amyE′*::P*gapA*$_{\Delta 3}$*::lacZ-cat*	pTD003 → 168CA
GTD004	*amyE′*::P*gapA*$_{\Delta 4}$*::lacZ-cat*	pTD004 → 168CA
GTD005	*amyE′*::P*gapA*$_{\Delta 5}$*::lacZ-cat*	pTD005 → 168CA
GTD006	*amyE′*::P*gapA*$_{\Delta 6}$*::lacZ-cat*	pTD006 → 168CA
GTD007	*amyE′*::P*gapA*$_{\Delta 7}$*::lacZ-cat*	pTD007 → 168CA
GTD008	*amyE′*::P*gapA*$_{\Delta 8}$*::lacZ-cat*	pTD008 → 168CA
GTD010	*amyE′*::P*gapA*$_1$*::lacZ-cat*	pTD010 → 168CA
GTD011	*amyE′*::P*gapA*$_2$*::lacZ-cat*	pTD011 → 168CA
GTD012	*amyE′*::P*gapA*$_3$*::lacZ-cat*	pTD012 → 168CA
GTD013	*amyE′*::P*gapA*$_4$*::lacZ-cat*	pTD013 → 168CA
GTD014	*amyE′*::P*gapA*$_5$*::lacZ-cat*	pTD014 → 168CA
GTD016	*amyE′*::P*gapA*$_6$*::lacZ-cat*	pTD016 →168CA
GTD017	*amyE′*::P*gapA*$_7$*::lacZ-cat*	pTD017 → 168CA
GTD018	*amyE′*::P*gapA*$_8$*::lacZ-cat*	pTD018 → 168CA
GTD021	*cggR′::lacZ-ery-lacI*-P*spac::′cggR*, *amyE′::PgapA¢::his*$_6$*cggR-cat*	pTD021 → BFA1080
GM1520	*cggR′::lacZ-ery-lacI*-P*spac::′cggR*, *amyE′*::P*gapA′::cggR-cat*	Fillinger *et al.* (2000)
GM1521	*cggR′::lacZ-ery-lacI*-P*spac::′cggR*, *amyE′::cat*	Fillinger *et al.* (2000)
GTD030	*ΔcggR::phleo*	pTD030 → 168CA
QB2066	*pfk met*	Laboratory stock
GTD042	*pfk metC amyE′*::P*gapA′::lacZ-cat*	GM1511 DNA × QB2066
GM1501	*gapA*::pMUTIN2	Fillinger *et al.* (2000)
GTD040	*gapA*::pMUTIN2 *Δ*(*lacZ-ery)::kan*	pEC23 → GM1501
GTD043	*gapA*::pMUTIN2 *Δ*(*lacZ-ery*)*::kan amyE′::PgapA′::lacZ-cat*	GM1511 DNA × GTD043

DNA manipulation

Polymerase chain reaction (PCR) products were purified with the QIAquick PCR purification kit (Qiagen). Plasmid DNA was extracted and purified from *E. coli* with the QIAprep spin miniprep kit (Qiagen). Restriction enzymes and T4 DNA ligase were used as recommended by the manufacturers. Sequencing of DNA was performed by the dideoxy chain termination method (Sambrook *et al.*, 1989) with the Sequenase 2.0 kit (Pharmacia) for primer extension analysis and by the chemical method (Sambrook *et al.*, 1989) for DNase I footprinting. *Pfu* DNA polymerase (Stratagene) was used for preparative PCR as recommended by the manufacturer. DNA fragments were purified from agarose gels using the gel extraction kit (Qiagen).

Plasmid constructions

Reporter fusions of promoter fragments to a promoterless *lacZ* gene were constructed using the vector pDG1661 (Guérout-Fleury *et al.*, 1996). The promoter fragments to be tested were obtained directly by PCR using specific oligonucleotides (see *Supplementary material*) as primers or after site-directed mutagenesis using a PCR procedure; these fragments were inserted between the *Eco*RI and *Hin*dIII sites of pDG1661 to generate plasmids pTD002 to pTD018.

The plasmid pTDp1 for His-tagged CggR overproduction was constructed by inserting the c*ggR* coding sequence amplified by PCR using primers CQ-9Bam and CQ-10Hd between the *Bam*HI and *Hin*dIII sites of the expression vector pQE-30 (Qiagen). The plasmid pTD021 used to test *in vivo* the functionality of the His-tagged CggR protein was obtained by inserting a fragment corresponding to the *his*$_6$*-cggR* ORF preceded by the wild-type *cggR* promoter and ribosome binding site region into the vector pDG1662 (Guérout-Fleury *et al.*, 1996).

The plasmid pTD030 for construction of a non-polar *cggR* deletion strain was obtained as follows: a *Bam*HI–*Eco*RI DNA fragment containing the 529 bp present just upstream of the *cggR* ORF in the chromosome and an in frame fusion between the initiation codon of the *cggR* ORF and the entire phleomycin resistance-encoding ORF (including its stop codon) from pIC22 (Steinmetz and Richter, 1994) was amplified in two steps by PCR using the primers FARAEB, CGGR-PHLEO and FPHLEOE; another *Eco*RI–*Aat*II DNA fragment corresponding to the 556 bp present immediately downstream of the *cggR* stop codon was generated by PCR using the primers DGAPAE and MGAPAAK; these two fragments were then inserted into the plasmid pJH101 linearized by *Bam*HI and *Aat*II to generate pTD030. All constructions including a PCR step were checked by sequencing.

In vivo replacement by homologous recombination of a kanamycin resistance cassette for the *lacZ* ORF and erythromycin resistance gene of pMUTIN2 was obtained using pEC23, a plasmid specifically designed for this use, constructed by P. Stragier.

β-Galactosidase assay

β-Galactosidase activities were measured using the method of Miller (1992) on cell extracts prepared by lysozyme treatment and centrifugation. Protein concentration was determined by the Bradford method (Bradford, 1976) using the

Bio-Rad protein assay solution. One unit of β-galactosidase activity is defined as the amount of enzyme that produces 1 nmol of *O*-nitrophenol min^{-1} at 20°C.

Expression and purification of His_6-CggR

Escherichia coli M15 pREP4 strain was transformed by the plasmid pTDp1 to generate the strain GTDp1. For production, an overnight preculture of strain GTDp1 was diluted at 1:50 in 500 ml of LB with appropriates antibiotics. Cells were grown at 37°C to OD_{600} 0.8, and 1 mM IPTG was added to induce the production of His-tagged CggR over a period of 3 h. Cells were harvested by centrifugation at 3000 *g* for 15 min at 4°C. The pellet was washed in 50 mM KH_2PO_4, pH 8. Bacteria were then lysed by incubation for 30 min at 37°C with 0.1 mg ml^{-1} lysozyme in the same buffer supplemented with 1 mM phenylmethylsulphonyl fluoride (PMSF). To complete lysis, samples were sonicated for 1 min. The soluble fraction was then recovered by centrifugation at 12 000 *g* for 30 min at 4°C. The supernatant was filtered through a Ni^{2+}-NTA agarose column equilibrated with the lysis buffer. The column was washed with 100 ml of wash buffer A (50 mM KH_2PO_4, pH 8, 10% glycerol), 100 ml of wash buffer B (wash buffer A + 1 M NaCl, 20 mM imidazole) and 20 ml of wash buffer C (50 mM Tris-HCl, pH 8, 10% glycerol, 50 mM NaCl, 20 mM imidazole). His_6-CggR was then eluted using an increasing imidazole gradient (50 mM to 400 mM). Fractions (10 μl) containing proteins were loaded on SDS-PAGE, and those containing apparently pure His_6-CggR were pooled and dialysed twice against 2 l of storage buffer [50 mM Tris-HCl, pH 8, 50 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol (DTT)].

DNA EMSAs

A DNA fragment containing the promoter region of the *cggR–gapA* operon was generated by PCR with primers PG-3R and PG-4aH and purified by Gene Clean (BIO 101). Approximately 20 pmol of this fragment was end-labelled using T4 polynucleotide kinase (Biolabs) and 30 μCi of [γ-^{32}P]-ATP (Isoblue, ICN). The labelled probe was purified by precipitation. Approximately 10 000 c.p.m. of radiolabelled probe (≈ 0.05 pmol) were incubated for 5 min at room temperature in binding buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 5% glycerol) with 1 μg of herring sperm DNA per reaction. Varying amounts of His_6-CggR were added for a total reaction volume of 20 μl. After incubation for 10 min at room temperature, samples were analysed by 6% native acrylamide gel electrophoresis, 1× Tris-glycine, run at room temperature and constant voltage (14 V cm^{-1}). After migration, gels were dried, and the radiolabelled bands were visualized by PhosphoImager system.

DNase I footprinting

Primers Spgapfoot and Apgapfoot (30 pmol) were end-labelled using T4 polynucleotide kinase (Biolabs) and 30 μCi of [γ-^{32}P]-ATP (Isoblue, ICN) and purified by precipitation. They were used to generate one-strand labelled PCR probes. Approximately 60 000 c.p.m. of labelled DNA probes (≈ 0.2 pmol) were incubated for 15 min at room temperature with varying amounts of His_6-CggR (0–160 ng) as described for the mobility shift assays. DNase I (Amersham Pharmacia Biotech) reactions were then performed as described previously (Derré *et al*., 1999). All samples were analysed on a 6% sequencing gel, which was dried and revealed by autoradiography.

The effect of different effector candidates was tested as follows: CggR, at a concentration just sufficient for full protection, was preincubated for 5 min with varying amounts of each metabolite (FBP, G3P, F6P or F26P) in reaction buffer at room temperature; the labelled probe was then added, and the mixture was treated as described above.

Transcriptome analysis

High-quality RNA preparation, radiolabelled cDNA synthesis and hybridization were performed as described previously (Hauser *et al*., 1998; Eymann *et al*., 2002). Cell pellets for RNA extraction were obtained from cultures grown up to exponential growth phase (OD_{600} = 0.5). Panorama *B. subtilis* gene macroarrays and a cDNA labelling kit were provided by Sigma-Genosys. Data analysis was performed using ARRAY VISION software (Imaging Research) for signals and background quantifications and GENESPRING software (Silicon Genetics) for further comparisons. A normalization procedure was performed in two steps: the overall spot normalization function of ARRAY VISION was used to calculate normalized intensity values of individual spots per array to allow for comparison between two filters for each gene, and GENESPRING was used for normalization of each gene to itself to allow comparison of all values obtained from all the experiments for each gene. Therefore, only genes for which the signal/noise ratio was higher than 3 in at least one condition (wild-type or *cggR* mutant context) were taken into consideration.

Databases and bioinformatic tools

Sequence queries and bioinformatic analysis were performed at NCBI (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi), TIGR (http://www.tigr.org) or SubtiList (genolist.pasteur.fr/SubtiList/index.html) websites.

Acknowledgements

We are greatly indebted to Georg Homuth and Michael Hecker (Greifswald University, Germany) for teaching and technical support on transcriptome analysis. We thank Patrick Stragier for the generous gift of plasmid pEC23 (unpublished). We are very grateful to Pascal Servant and Nathalie Declerck for excellent technical advice and fruitful discussion. We also thank Dominique Le Coq, Ivan Mijakovic, Josef Deutscher and Claude Gaillardin for helpful advice and comments on the manuscript. This work was supported by EU Quality of Life and Management of Living Resources, programme grant QL GZ-CT-1999-O1455. Stéphane Aymerich and Thierry Doan are on the staff at INRA.

Supplementary material

The following material is available from http://www.blackwellpublishing.com/products/journals/suppmat/mole/mole3404/mmi3404sm.htm

Fig. S1. Phylogenic tree calculated at EMBL-EBI site (www.ebi.ac.uk/clustalw) from the amino acid sequence of 75 proteins of the SorC family referenced in Table S3.

Table S1. Oligonucleotides used in this study.

Table S2. Complete results of transcriptome comparisons between wild-type and cggR mutant strains.

Table S3. 75 proteins of the SorC family considered in this study.

References

van Aalten, D.M., DiRusso, C.C., and Knudsen, J. (2001) The structural basis of acyl coenzyme A-dependent regulation of the transcription factor FadR. *EMBO J* **20:** 2041–2050.

Anagnostopoulos, C., and Spizizen, J. (1961) Requirements for transformation in *Bacillus subtilis*. *J Bacteriol* **81:** 741–746.

Aung-Hilbrich, L.M., Seidel, G., Wagner, A., and Hillen, W. (2002) Quantification of the influence of HPrSer46P on CcpA–cre interaction. *J Mol Biol* **319:** 77–85.

Aymerich, S., Gonzy-Treboul, G., and Steinmetz, M. (1986) 5′-noncoding region sacR is the target of all identified regulation affecting the levansucrase gene in *Bacillus subtilis*. *J Bacteriol* **166:** 993–998.

Bekal-Si Ali, S., Divies, C., and Provost, H. (1999) Genetic organization of the citCDEF locus and identification of mae and clyR genes from *Leuconostoc mesenteroides*. *J Bacteriol* **181:** 4411–4416.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72:** 248–254.

Brennan, R.G. (1993) The winged-helix DNA-binding motif: another helix–turn–helix takeoff. *Cell* **74:** 773–776.

Chen, C.S., White, A., Love, J., Murphy, J.R., and Ringe, D. (2000) Methyl groups of thymine bases are important for nucleic acid recognition by DtxR. *Biochemistry* **39:** 10397–10407.

Derré, I., Rapoport, G., and Msadek, T. (1999) CtsR, a novel regulator of stress and heat shock response, controls clp and molecular chaperone gene expression in Gram-positive bacteria. *Mol Microbiol* **31:** 117–131.

Deutscher, J., Kuster, E., Bergstedt, U., Charrier, V., and Hillen, W. (1995) Protein kinase-dependent HPr/CcpA interaction links glycolytic activity to carbon catabolite repression in Gram-positive bacteria. *Mol Microbiol* **15:** 1049–1053.

Deutscher, J., Galinier, A., and Martin-Verstraete, I. (2002) Carbohydrate uptake and metabolism. In *Bacillus subtilis and its Closest Relatives: from Genes to Cells.* Sonenshein, A.L., Hoch, J.A., and Losick, R. (eds). Washington, DC: American Society for Microbiology Press, pp. 129–150.

De Wulf, P. (1998) Presence of the ribulose monophosphate pathway in *Bacillus subtilis*. *Microbiology* **144:** 596–597.

Eymann, C., Homuth, G., Scharf, C., and Hecker, M. (2002) *Bacillus subtilis* functional genomics: global characterization of the stringent response by proteome and transcriptome analysis. *J Bacteriol* **184:** 2500–2520.

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.

Fortnagel, P. (1993) Glycolysis. In *Bacillus subtilis and Other Gram-Positive Bacteria.* Sonenshein, A.L., Hoch, J.A., and Losick, R. (eds). Washington, DC: American Society for Microbiology Press, pp. 171–180.

Fouet, A., and Sonenshein, A.L. (1990) A target for carbon source-dependent negative regulation of the citB promoter of *Bacillus subtilis*. *J Bacteriol* **172:** 835–844.

Gajiwala, K.S., and Burley, S.K. (2000) Winged helix proteins. *Curr Opin Struct Biol* **10:** 110–116.

Guérout-Fleury, A.M., Frandsen, N., and Stragier, P. (1996) Plasmids for ectopic integration in *Bacillus subtilis*. *Gene* **180:** 57–61.

Hauser, N.C., Vingron, M., Scheideler, M., Krems, B., Hellmuth, K., Entian, K.D., and Hoheisel, J.D. (1998) Transcriptional profiling on all open reading frames of *Saccharomyces cerevisiae*. *Yeast* **14:** 1209–1221.

Hederstedt, L. (1993) The Krebs citric acid cycle. In *Bacillus subtilis and Other Gram-Positive Bacteria.* Sonenshein, A.L., Hoch, J.A., and Losick, R. (eds). Washington, DC: American Society for Microbiology Press, pp. 181–198.

Heldwein, E.E., and Brennan, R.G. (2001) Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* **409:** 378–382.

Henkin, T.M. (1996) The role of CcpA transcriptional regulator in carbon metabolism in *Bacillus subtilis*. *FEMS Microbiol Lett* **135:** 9–15.

Heuel, H., Shakeri-Garakani, A., Turgut, S., and Lengeler, J.W. (1998) Genes for D-arabinitol and ribitol catabolism from *Klebsiella pneumoniae*. *Microbiology* **144:** 1631–1639.

Hueck, C.J., and Hillen, W. (1995) Catabolite repression in *Bacillus subtilis*: a global regulatory mechanism for the Gram-positive bacteria? *Mol Microbiol* **15:** 395–401.

Huffman, J.L., and Brennan, R.G. (2002) Prokaryotic transcription regulators: more than just the helix–turn–helix motif. *Curr Opin Struct Biol* **12:** 98–106.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., *et al.* (1997) The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390:** 249–256.

Leyva-Vazquez, M.A., and Setlow, P. (1994) Cloning and nucleotide sequences of the genes encoding triose phosphate isomerase, phosphoglycerate mutase, and enolase from *Bacillus subtilis*. *J Bacteriol* **176:** 3903–3910.

Ludwig, H., Homuth, G., Schmalisch, M., Dyka, F.M., Hecker, M., and Stülke, 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:** 409–422.

Ludwig, H., Rebhan, N., Blencke, H.M., Merzbacher, M., and Stülke, J. (2002) Control of the glycolytic gapA operon by the catabolite control protein A in *Bacillus subtilis*: a novel mechanism of CcpA-mediated regulation. *Mol Microbiol* **45:** 543–553.

Mijakovic, I., Poncet, S., Galinier, A., Monedero, V., Fieulaine, S., Janin, J., *et al.* (2002) Pyrophosphate-producing protein dephosphorylation by HPr kinase/phosphorylase: a relic of early life? *Proc Natl Acad Sci USA* **99:** 13442–13447.

Miller, J.H. (1992) *A Short Course in Bacterial Genetics*. New York: Cold Spring Harbor Laboratory press, pp. 72–74.

Miwa, Y., Nakata, A., Ogiwara, A., Yamamoto, M., and Fujita, Y. (2000) Evaluation and characterization of catabolite-responsive elements (cre) of *Bacillus subtilis*. *Nucleic Acids Res* **28:** 1206–1210.

Moreno, M.S., Schneider, B.L., Maile, R.R., Weyler, W., and Saier, M.H., Jr (2001) Catabolite repression mediated by the CcpA protein in *Bacillus subtilis*: novel modes of regulation revealed by whole-genome analyses. *Mol Microbiol* **39:** 1366–1381.

Pohl, E., Holmes, R.K., and Hol, W.G. (1999) Crystal structure of a cobalt-activated diphtheria toxin repressor–DNA complex reveals a metal-binding SH3-like domain. *J Mol Biol* **292:** 653–667.

Presecan-Siedel, E., Galinier, A., Longin, R., Deutscher, J., Danchin, A., Glaser, P., and Martin-Verstraete, I. (1999) Catabolite regulation of the pta gene as part of carbon flow pathways in *Bacillus subtilis*. *J Bacteriol* **181:** 6889–6897.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press.

Sangari, F.J., Aguero, J., and Garcia-Lobo, J.M. (2000) The genes for erythritol catabolism are organized as an inducible operon in *Brucella abortus*. *Microbiology* **146:** 487–495.

Saxild, H.H., Andersen, L.N., and Hammer, K. (1996) Dra-nupC-pdp operon of *Bacillus subtilis*: nucleotide sequence, induction by deoxyribonucleosides, and transcriptional regulation by the deoR-encoded DeoR repressor protein. *J Bacteriol* **178:** 424–434.

Shin, B.S., Choi, S.K., and Park, S.H. (1999) Regulation of the *Bacillus subtilis* phosphotransacetylase gene. *J Biochem (Tokyo)* **126:** 333–339.

Sonenshein, A.L. (2002) The Krebs citric acid cycle. In *Bacillus subtilis and its Closest Relatives: from Genes to Cells.* Sonenshein, A.L., Hoch, J.A., and Losick, R. (eds). Washington, DC: American Society for Microbiology Press, pp. 151–162.

Steinmetz, M., and Richter, R. (1994) Plasmids designed to alter the antibiotic resistance expressed by insertion mutations in *Bacillus subtilis*, through *in vivo* recombination. *Gene* **142:** 79–83.

Stülke, J., and Hillen, W. (2000) Regulation of carbon catabolism in *Bacillus* species. *Annu Rev Microbiol* **54:** 849–880.

Tobisch, S., Zuhlke, D., Bernhardt, J., Stülke, J., and Hecker, M. (1999) Role of CcpA in regulation of the central pathways of carbon catabolism in *Bacillus subtilis*. *J Bacteriol* **181:** 6996–7004.

Turinsky, A.J., Grundy, F.J., Kim, J.H., Chambliss, G.H., and Henkin, T.M. (1998) Transcriptional activation of the *Bacillus subtilis* ackA gene requires sequences upstream of the promoter. *J Bacteriol* **180:** 5961–5967.

Turinsky, A.J., Moir-Blais, T.R., Grundy, F.J., and Henkin, T.M. (2000) *Bacillus subtilis* ccpA gene mutants specifically defective in activation of acetoin biosynthesis. *J Bacteriol* **182:** 5611–5614.

Wohrl, B.M., Wehmeier, U.F., and Lengeler, J.W. (1990) Positive and negative regulation of expression of the L-sorbose (sor) operon by SorC in *Klebsiella pneumoniae*. *Mol Gen Genet* **224:** 193–200.

Yebra, M.J., Veyrat, A., Santos, M.A., and Perez-Martinez, G. (2000) Genetics of L-sorbose transport and metabolism in *Lactobacillus casei*. *J Bacteriol* **182:** 155–163.

Yoshida, K., Kobayashi, K., Miwa, Y., Kang, C.M., Matsunaga, M., Yamaguchi, H., *et al.* (2001) Combined transcriptome and proteome analysis as a powerful approach to study genes under glucose repression in *Bacillus subtilis*. *Nucleic Acids Res* **29:** 683–692.

Zeng, X., and Saxild, H.H. (1999) Identification and characterization of a DeoR-specific operator sequence essential for induction of dra-nupC-pdp operon expression in *Bacillus subtilis*. *J Bacteriol* **181:** 1719–1727.

Zeng, X., Saxild, H.H., and Switzer, R.L. (2000) Purification and characterization of the DeoR repressor of *Bacillus subtilis*. *J Bacteriol* **182:** 1916–1922.

4.2.2. "Supplementary material" de l'article

Table S1. Oligonucleotides used in this study.

Fig. S1. Phylogenic tree calculated at EMBL-EBI site (www.ebi.ac.uk/clustalw) from the amino acid sequence of 75 proteins of the SorC family referenced in Table S3.

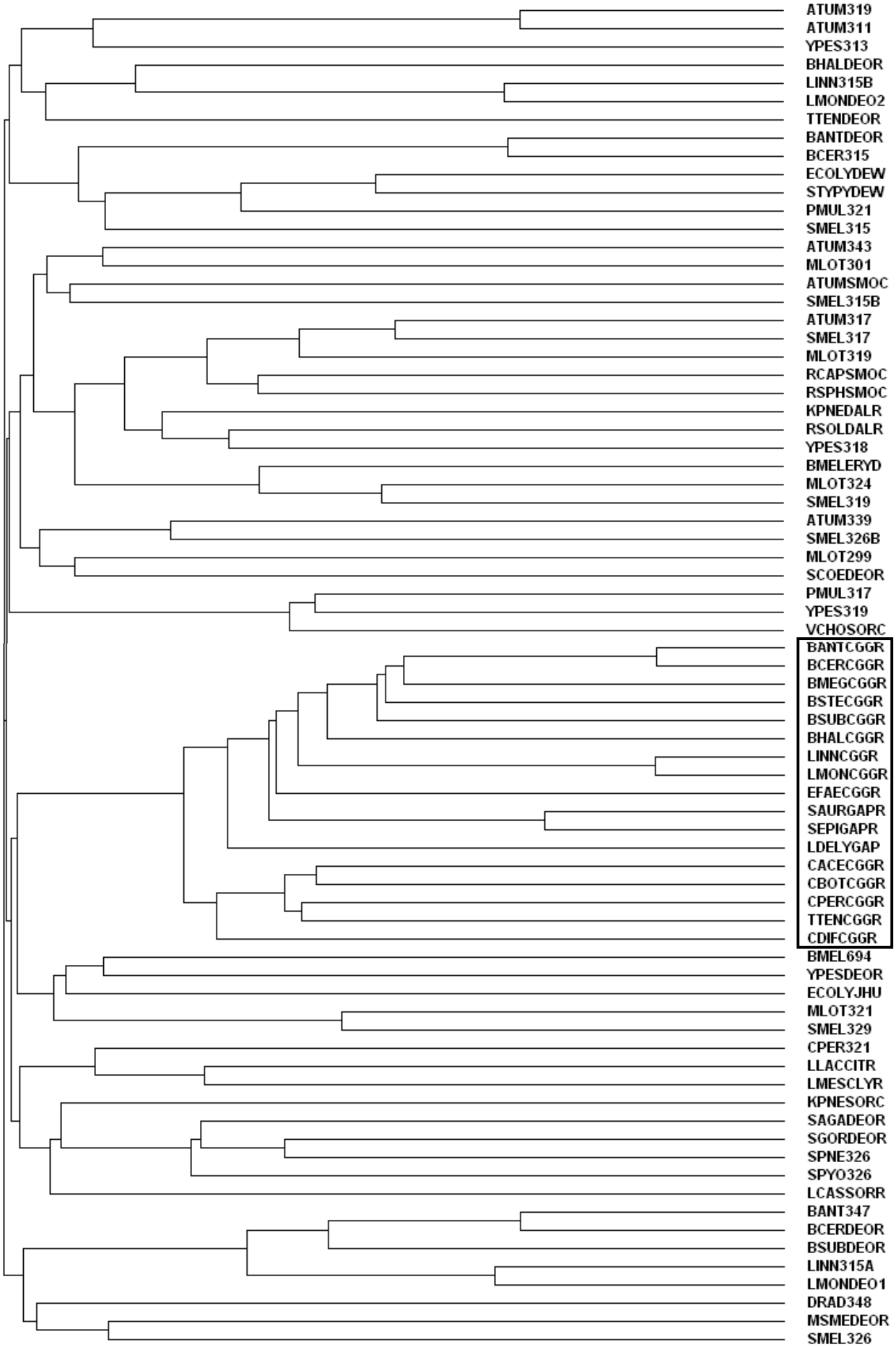
Table S3. 75 proteins of the SorC family considered in this study.

Fig. S2 Graphique de dispersion des trois expériences de transcriptome citées dans l'article n°1 permettant de vérifier la qualité globale de chaque comparaison (cette figure additionnelle n'est pas dans l'article).

TABLE S1. Oligonucleotides used in this study

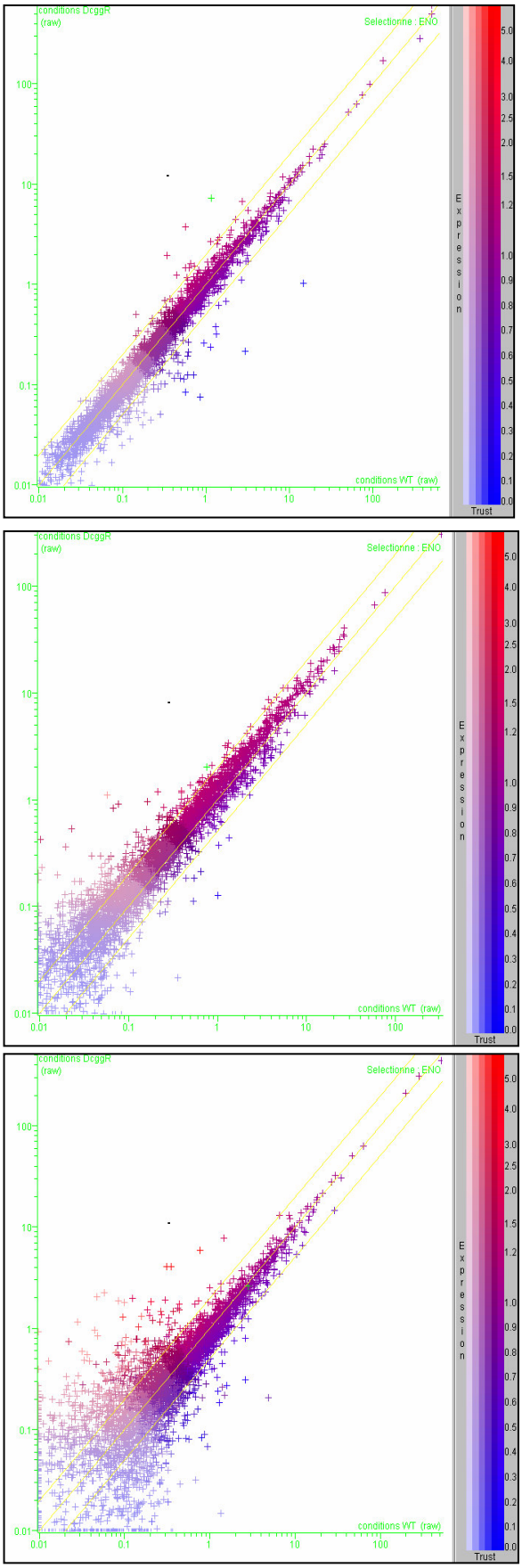
Purpose	sequence (5' ⟶ 3')[a]	Corresponding Plasmid
Fusion		
PG-1R	ggAATTCTCTgCgAATATTgTCgATg	pTD002
PG-2R	ggAATTCAgTTgAATAAACAATTCCACCC	pTD003
PG-3R	ggAATTCTTTTggCTATgACgggACg	pTD004
PG-4aH	gCCAAgCTTCTggTTCATgACTCAAACgTTCC	pTD001 to pTD004
PG-4bH	gCCAAgCTTCTggTTCATgACTCAAACgTTCCTTCCCgTCATAgCCAAAAAAAATC	pTD005
PG-4cH	gCCAAgCTTCTggTTCATgACTCAAACgTTCCTTATgTCCCgCTATgACAAAAAACg	pTD006
PG-4dH	gCCAAgCTTCTggTTCATgACTCAAACgTTCgCTATgACAAAAAACgTCCCg	pTD007
PG-4eH	gCCAAgCTTCTggTTCATgACTCAAACgTTCCTAgCCAAAAAAAATCATTTCTgC	pTD008
Site directed mutagenesis		
PG-d1	CTATgACAAAAAACgT**TT**CgTCATAgCCAAA	pTD010
PG-daval1	CCgCTATgACAAAAA**gT**gTCCCgTCATAgC	pTD011
PG-d2	ggACATTATATgT**TT**CgCTATgACAAAAAAC	pTD012
PG-daval2	gCTggACATTAT**Cg**gTCCCgCTATgAC	pTD013
PG-d7Tamont1	CgTCCCgTCATAgCC**CCCCCCC**ATCATTTCTgCTTTCTTT	pTD014
PG-d5Taval1	TATgTCCCgCTATgAC**CCCCC**ACgTCCCgTCATAgCC	pTD016
PG-damont2	ggACATTATATgTCCCgCTA**CAgT**AAAAAACgTCCCgTCATAg	pTD017
PG-daval4	CgTTCCTTCCTTTTTTgCT**AAgT**ATTATATgTCCCgTATgAC	pTD018
CggR overproduction		
CQ-9Bam	CggggATCCATgAACCAgTTAATACAAgCTC	pTDp1
CQ-10Hd	gCCgAAgCTTATTCATCCCTTAATAACTTCTTTgC	pTDp1
PG-RBS-ATG-Tag	gTgATggTgATgCgATCCTCTCATgACTCAAACgTTCCTTCC	pTD021
Footprint		
Spgapfoot	CTggATAAAACgACgCggCATgAAAAC	
Apgapfoot	ggTTCTgTCAgCCTgATATACTgC	
***cggR* mutant construction**		
FARAEB	CgCggATCCCgAACCATCTgCgTg	pTD030
CGGRPHLEO	CCgggATAgACTgTAACAtgACTCAAACgTTCCT	pTD030
FPHLEOE	CggAATTCTAgATTTTAgCTTTTTATTTgTTgA	pTD030
DGAPAE	CggAATTCTCTCACTTATTTAAAggAgg	pTD030
MGAPAAK	ggggTACCgACgTCATCATACCgCgTTTg	pTD030

[a] Restriction sites are underlined and mutated nucleotides in site-directed mutagenesis are in bold-case.


ATUM319
ATUM311
YPES313
BHALDEOR
LINN315B
LMONDEO2
TTENDEOR
BANTDEOR
BCER315
ECOLYDEW
STYPYDEW
PMUL321
SMEL315
ATUM343
MLOT301
ATUMSMOC
SMEL315B
ATUM317
SMEL317
MLOT319
RCAPSMOC
RSPHSMOC
KPNEDALR
RSOLDALR
YPES318
BMELERYD
MLOT324
SMEL319
ATUM339
SMEL326B
MLOT299
SCOEDEOR
PMUL317
YPES319
VCHOSORC
BANTCGGR
BCERCGGR
BMEGCGGR
BSTECGGR
BSUBCGGR
BHALCGGR
LINNCGGR
LMONCGGR
EFAECGGR
SAURGAPR
SEPIGAPR
LDELYGAP
CACECGGR
CBOTCGGR
CPERCGGR
TTENCGGR
CDIFCGGR
BMEL694
YPESDEOR
ECOLYJHU
MLOT321
SMEL329
CPER321
LLACCITR
LMESCLYR
KPNESORC
SAGADEOR
SGORDEOR
SPNE326
SPYO326
LCASSORR
BANT347
BCERDEOR
BSUBDEOR
LINN315A
LMONDEO1
DRAD348
MSMEDEOR
SMEL326

Name	Bacterial species	Protein Identification		length (aa)
		CggR	SorC/DeoR	
ATUM319	*Agrobacterium tumefaciens C58*		NP 354881	319
ATUM311	*Agrobacterium tumefaciens C58*		NP 532584	311
YPES313	*Yersinia pestis*		NP 406776	313
BHALDEOR	*Bacillus halodurans*		BABO 4499	316
LINN315B	*Listeria innocua*		NP 469894	315
LMONDEO2	*Listeria monocytogenes*		NP 464075	315
TTENDEOR	*Thermoanaerobacter tengcongensis*		NP 621881	313
BANTDEOR	*Bacillus anthracis*		NP 656845	310
BCER315	*Bacillus cereus*		DEOR	315
ECOLYDEW	*Escherichia coli K12*		NP 416029	317
STYPYDEW	*Salmonella typhymurium*		NP 457979	319
PMUL321	*Pasteurella multocida*		NP 246210	321
SMEL315	*Sinorhizobium meliloti*		NP 387076	315
ATUM343	*Agrobacterium tumefaciens C58*		NP 356191	343
MLOT301	*Mesorhizobium loti*		NP 104679	301
ATUMSMOC	*Agrobacterium tumefaciens C58*		AAK88994	317
SMEL315B	*Sinorhizobium meliloti*		NP 437134	315
ATUM317	*Agrobacterium tumefaciens C58*		NP 356209	317
SMEL317	*Sinorhizobium meliloti*		NP 386552	317
MLOT319	*Mesorhizobium loti*		NP 105687	319
RCAPSMOC	*Rhodobacter capsulatum*		T03554	314
RSPHSMOC	*Rhodobacter sphaeroides*		AAC45765	317
KPNEDALR	*Klebsiella pneumoniae*		DALR	314
RSOLDALR	*Ralstonia solanacearum*		NP 520248	315
YPES318	*Yersinia pestis*		NP 405861	318
BMELERYD	*Brucella melitensis*		NP 541405	316
MLOT324	*Mesorhizobium loti*		NP 107618	324
SMEL319	*Sinorhizobium meliloti*		NP 386347	319
ATUM339	*Agrobacterium tumefaciens C58*		NP 356756	339
SMEL326B	*Sinorhizobium meliloti*		NP 437023	326
MLOT299	*Mesorhizobium loti*		NP 104518	299
SCOEDEOR	*Streptomyces coelicolor*		NP 625744	360
PMUL317	*Pasteurella multocida*		NP 246585	317
YPES319	*Yersinia pestis*		NP 406811	319
VCHOSORC	*Vibrio cholerae*		NP 233010	321
BANTCGGR	*Bacillus anthracis*	NP 653588		342
BCERCGGR	*Bacillus cereus*	CggR		342
BMEGCGGR	*Bacillus megaterium*	CggR		342
BSTECGGR	*Bacillus stearothermophilus*	CggR		341
BSUBCGGR	*Bacillus subtilis*	NP 391275		340
BHALCGGR	*Bacillus halodurans*	NP 244428 (BH3561)		339
LINNCGGR	*Listeria innocua*	NP 471884		348
LMONCGGR	*Listeria monocytogenes*	NP 465983		348
EFAECGGR	*Enterococcus faecalis*	CggR		345
SAURGAPR	*Staphylococcus aureus*	NP 371295		337
SEPIGAPR	*Staphylococcus epidermidis*	GapR		337
LDELYGAP	*Lactobacillus delbrueckii*	YgaP		382
CACECGGR	*Clostridium acetobutylicum*	NP 347345		344
CBOTCGGR	*Clostridium botulinum*	CggR		343
CPERCGGR	*Clostridium perfringens*	NP 562221		359
TTENCGGR	*Thermoanaerobacter tengcongensis*	NP 623353		340
CDIFCGGR	*Clostridium difficile*	CggR		340
BMEL694	*Brucella melitensis*		NP 539315	694
YPESDEOR	*Yersinia pestis*		NP 405375	325

ECOLYJHU	*Escherichia coli K12*	NP 418715	266
MLOT321	*Mesorhizobium loti*	NP 107644	321
SMEL329	*Sinorhizobium meliloti*	NP 438071	329
CPER321	*Clostridium perfringens*	NP 562060	321
LLACCITR	*Lactococcus lactis*	NP 267343	309
LMESCLYR	*Leuconostoc mesenteroides*	ClyR	312
KPNESORC	*Klebsiella pneumoniae*	SORC	315
SAGADEOR	*Streptococcus agalactiae*	DeoR	329
SGORDEOR	*Streptococcus gordonii*	DeoR	321
SPNE326	*Streptococcus pneumoniae*	NP 357822	326
SPYO326	*Streptococcus pyogenes*	NP 608072	326
LCASSORR	*Lactobacillus casei*	SorR	318
BANT347	*Bacillus anthracis*	NP 655753	347
BCERDEOR	*Bacillus cereus*	DEOR	315
BSUBDEOR	*Bacillus subtilis*	NP 391822	313
LINN315A	*Listeria innocua*	NP 471438	315
LMONDEO1	*Listeria monocytogenes*	NP 465520	315
DRAD348	*Deinococcus radiodurans*	NP 295658	348
MSMEDEOR	*Mycobacterium smegmatis*	DeoR	314
SMEL326	*Sinorhizobium meliloti*	NP 437684	326



1 : batchD / ARN1

2 : batchH / ARN2

3 : batchD / ARN3

Figure S2. Résultats globaux des trois comparaisons transcriptomiques « souche sauvage/mutant *cggR* » sous forme de graphique de dispersion en échelle logarithmique. En abscisse : expression normalisée dans la souche sauvage. En ordonnée : expression normalisée dans la souche mutante. Pour chaque expérience, le lot de membrane et le lot d'ARN utilisés sont indiqués.

4.3. Travaux complémentaires

4.3.1. Résultats

Détermination du +1 de transcription de l'opéron *gapA*.

L'équipe de Stülke a publié le résultat de leur expérience visant à identifier le site de démarrage de la transcription de l'opéron *gapA* (Ludwig *et al.*, 2001). Elle a identifié deux positions possibles, localisées 96 et 95 pb en amont du premier nucléotide du codon initiateur de la traduction du premier gène de l'opéron *gapA*. Nous avons nous-mêmes réalisé des expériences similaires qui ont abouti à des résultats très légèrement différents. En effet, deux signaux d'égale intensité, correspondant à la synthèse de deux cDNA différant d'un nucléotide en longueur, ont été produits de façon reproductible à partir d'ARN extraits d'une culture en présence de glucose alors qu'aucun signal n'est détectable à partir d'ARN extraits d'une culture en présence de succinate plus glutamate (Figure 39). Ces résultats montrent que la transcription est initiée au niveau d'un A (aussi identifié par Ludwig et ses collègues) ou d'un G, situés à 95 et 94 nucléotides respectivement, en amont du premier codon de *cggR*. Nos résultats et ceux de nos confrères convergent vers l'identification de séquences promotrices –35 et –10 (TTGAATN$_{17}$TAAAAT) correspondant à un promoteur σ^A-dépendant de *Bacillus subtilis*.

Vérification de la structure du mutant *cggR* GTD030.

Le mutant *cggR* dont nous disposions, BFA1080, avait été obtenu au laboratoire par insertion du plasmide pMUTIN (Vagner *et al.*, 1998) dans le cadre du premier programme européen d'analyse fonctionnelle du génome de *Bacillus subtilis* (Fillinger *et al.*, 2000). *cggR* étant le premier gène de l'opéron *cggR-gapA-pgk-tpi-pgm-eno*, l'insertion du plasmide pMUTIN rendait donc la mutation polaire. Pour étudier l'effet d'une inactivation de *cggR* sur l'initiation de la transcription sans empêcher l'expression des gènes suivants de l'opéron, il nous fallait construire un mutant « non polaire ». La stratégie choisie a été de remplacer précisément l'ORF *cggR* par une ORF codant la résistance à la phléomycine. La procédure utilisée est décrite dans l'article n°1. Nous avons vérifié la structure de ce mutant non polaire par Southern Blot (Figure 40).

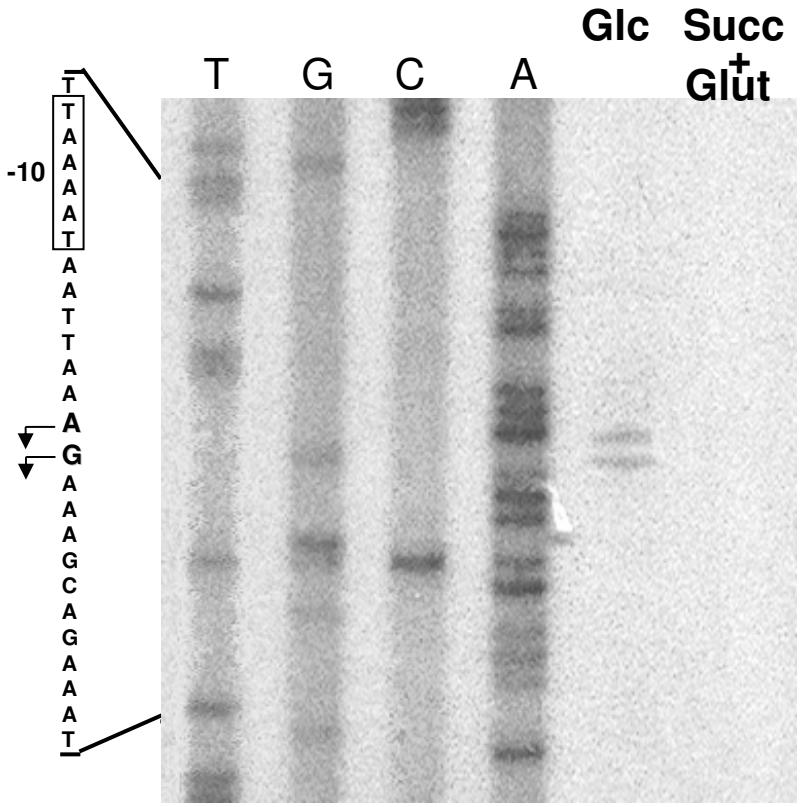
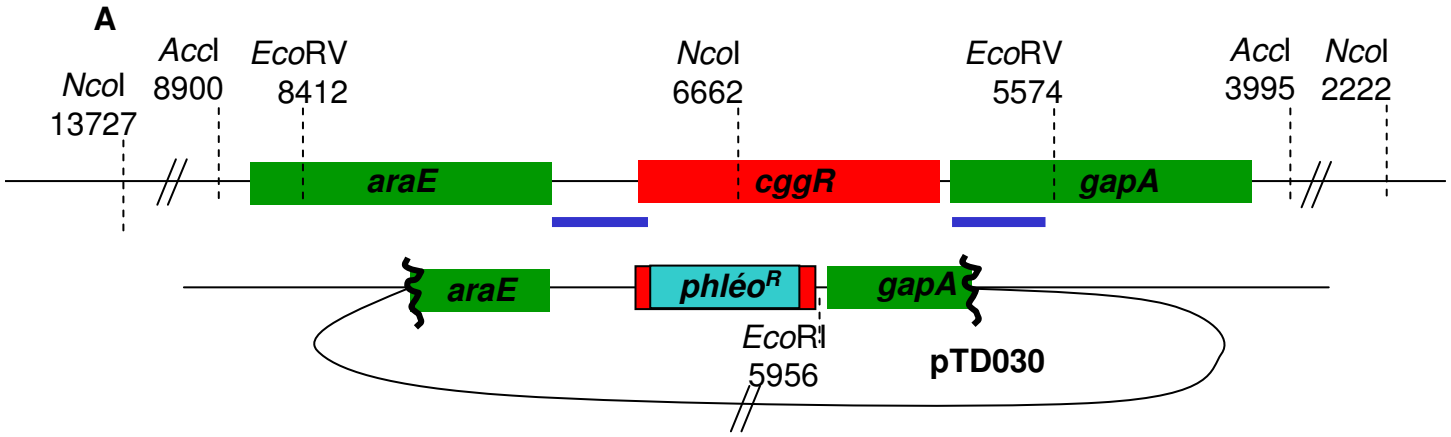


Figure 39. Détermination du site de démarrage de la transcription en amont de *cggR*. Les réactions de reverse-transcription ont été réalisées à partir d'ARN totaux de la souche sauvage 168CA cultivée en présence de glucose (Glc) ou de succinate plus glutamate (Succ+Glut). Les colonnes TGCA correspondent aux produits de séquençage obtenus avec le même oligonucléotide utilisé pour la reverse-transcription. Les sites de démarrage de la transcription sont indiqués par des flèches. La séquence promotrice -10 est encadrée.



B

souche	digestion				
	*EcoR*I+*EcoR*V	*EcoR*V	*Nco*I	*Acc*I	*Acc*I+*Nco*I
WT	2,8	2,8	7,1/4,4	4,9	2,7/2,2
Double événement de recombinaison	1,8/0,4	2,2	10,9	4,3	4,3

C

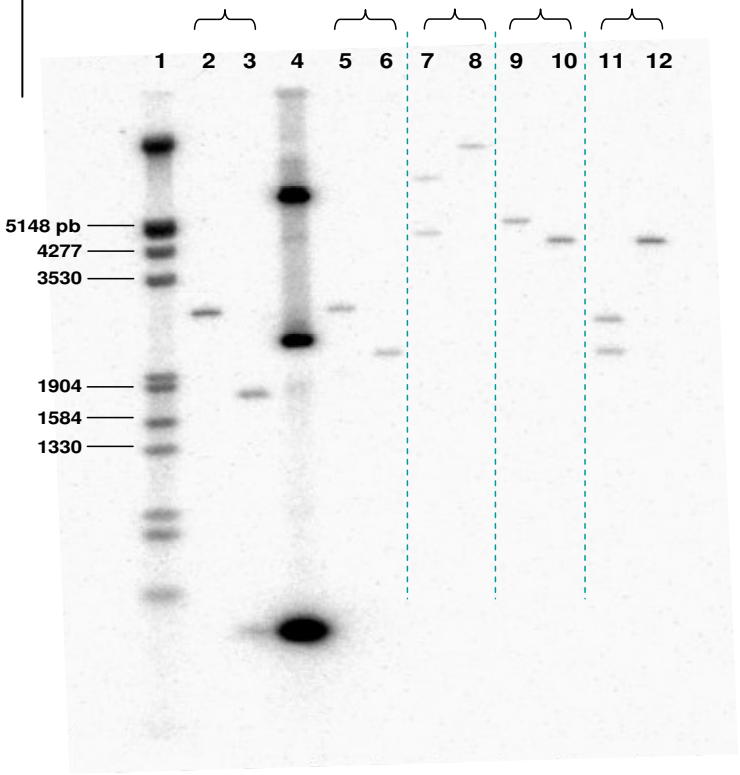


Figure 40. A. Structure chromosomique de la souche sauvage et structure du plasmide pTD030. La position des sondes utilisées est indiquée par les bandes bleues. **B.** Taille en kb des bandes attendues après hybridation et révélation du Southern Blot en fonction des enzymes de restriction utilisées. **C.** Analyse par Southern Blot de la structure chromosomique du mutant GTD030 (*DcggR::phléo*).
Ligne 1 : marqueur de taille I*Eco*RI/*Hind*III. Lignes 2, 5, 7, 9 et 11 : ADN de la souche sauvage. Lignes 3, 6, 8, 10 et 12 : ADN de la souche GTD030. Ligne 4 : ADN plasmidique pTD 030 digéré par *Eco*RI+*Eco*RV (Doan et Aymerich, 2003).
Lignes 2 et 3 : digestion par *Eco*RI+*Eco*RV. Lignes 5 et 6 : digestion par *Eco*RV. Lignes 7 et 8 : digestion par *Nco*I. Lignes 9 et 10 : digestion par *Acc*I. Lignes 11 et 12 : digestion par *Acc*I+*Nco*I.

Le Southern Blot a été réalisé à partir d'ADN génomique extrait de la souche sauvage 168CA et d'un clone obtenu après transformation.

Deux sondes radioactives d'environ 300 pb ont été utilisées simultanément : l'une correspond à la région régulatrice de l'expression de l'opéron *gapA* (insert du plasmide pSF111) (Fillinger *et al.*, 2000) et l'autre correspond aux 300 pb du début du gène *gapA*. Nos résultats démontrent que le transformant testé possède bien la structure chromosomique attendue (Figure 40). Ce clone a été baptisé GTD030.

Etude *in vitro* du rôle de la répétition de Thymines conservées en amont de l'opérateur de CggR.

La comparaison de la séquence de l'opérateur de CggR de *Bacillus subtilis* et de la région régulatrice de l'expression de l'opéron *gapA* chez d'autres bactéries à Gram positif a permis de mettre en évidence quatre motifs conservés dans la région protégée par CggR. Elle a aussi mis en évidence la conservation d'une région riche en Thymines juste en amont de cette empreinte. Nos tests génétiques après mutagenèse dirigée ont montré que cette série de Thymines est nécessaire *in vivo* à la régulation de la transcription de l'opéron *gapA*. Pour savoir si ce motif conservé, bien qu'il se trouve en dehors de la région protégée par CggR dans nos expériences d'empreinte à la DNase I, est nécessaire à l'interaction entre CggR et l'ADN, nous avons réalisé des expériences de titration en retard de migration électrophorétique. Les résultats de ces expériences montrent qu'un fragment qui ne contient pas cette série de Thymines titre aussi bien la sonde sauvage, correspondant a un large fragment de la région régulatrice de l'opéron *gapA*, qu'un fragment d'ADN étendu jusqu'après la série de Thymines (Figure 41). Par ailleurs, un fragment délétée du dernier motif déterminé d'interaction avec CggR titre environ 5 fois moins bien la sonde sauvage qu'un fragment comprenant l'ensemble des motifs conservés (Figure 41). La série de Thymines n'est donc pas nécessaire *in vitro* à l'interaction efficace de CggR avec sa cible. Ces résultats suggèrent que cette répétition de Thymines pourrait *in vivo* être nécessaire à l'établissement d'un contexte favorable, par exemple par courbure de l'ADN, mais pas pour l'interaction directe avec CggR.

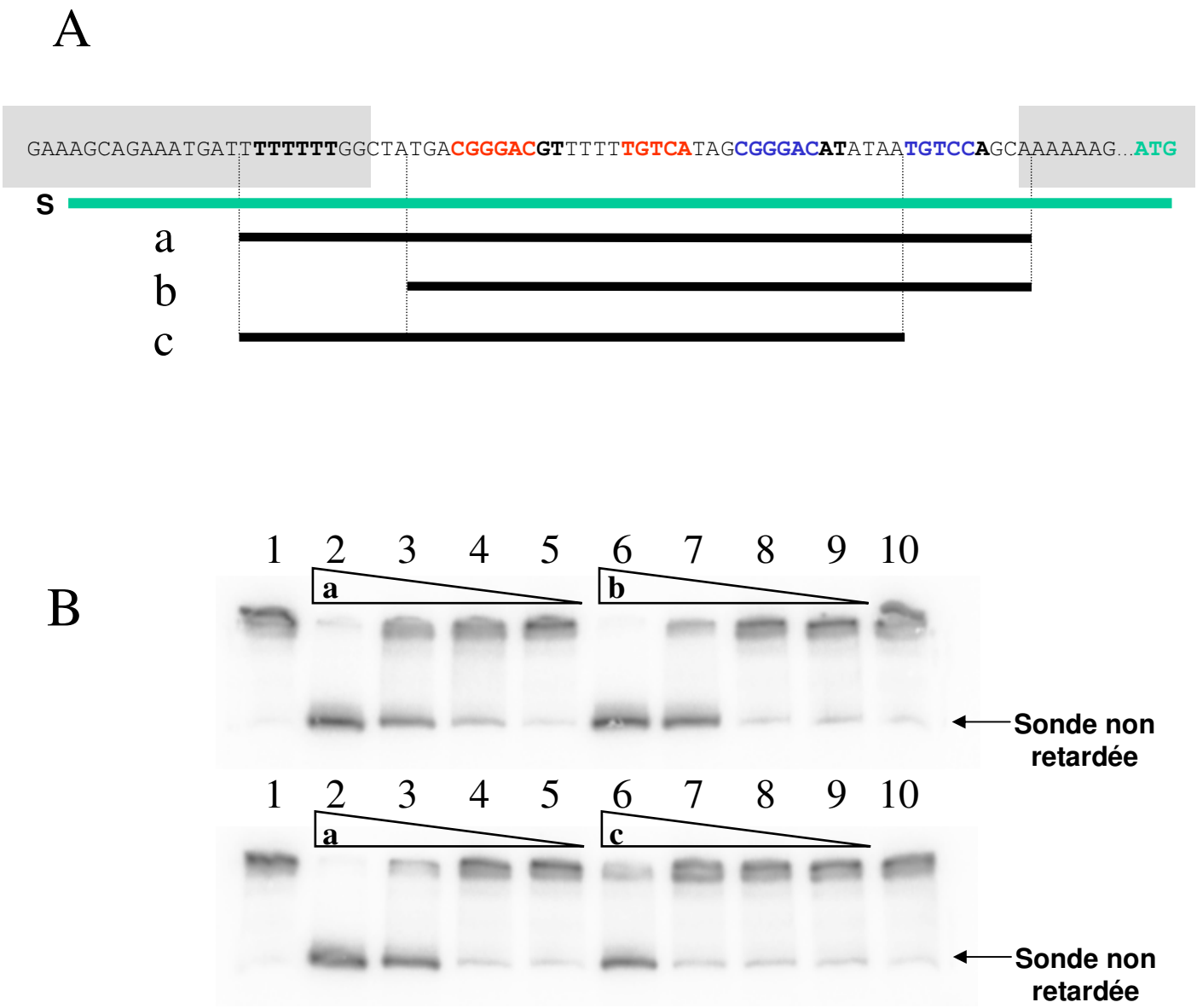


Figure 41. Test *in vitro* par titration en retard de migration électrophorétique du rôle de la série conservée de thymines sur l'interaction CggR/ADN.

A. Séquence de la région protégée par CggR. Les nucléotides non protégés sont indiqués sur fond gris-bleu. Les motifs reconnus par CggR sont en rouge et bleu. La capacité de différents fragments d'ADN non marqués à titrer l'interaction entre CggR (~0,4 pmol) et la sonde radioactive (S ; ~0,05 pmol) a été testée : (a) un fragment comprenant la série de thymines et la région protégée par CggR ; (b) un fragment ne comprenant que la séquence protégée par CggR ; (c) un fragment comprenant la série de thymines mais pas le dernier motif reconnu par CggR dans la région qu'il protège. **B.** Résultats des tests de titration. Puits 1 et 10 : 0,4 pmol CggR + 0,05 pmol sonde marquée. Puits 5 à 2 et 9 à 6 : 0,4 pmol CggR + 0,05 pmol sonde marquée + respectivement (0,01) (0,1) (1) et 10 pmol de fragment ADN a, b ou c, non marqué.

Détermination du degré d'oligomérisation de CggR par pontage chimique au glutaraldéhyde.

La forme monomérique de CggR a un poids moléculaire théorique de 38 kDa. Après purification et analyse d'un aliquot de CggR en gel d'électrophorèse dénaturant (SDS-PAGE), nous avons pu observer une seule bande correspondant à un poids moléculaire d'environ 37 kDa (Figure 42). Cependant, les résultats des expériences de retard de migration électrophorétique réalisées avec cette protéine suggèrent la formation de différents complexes CggR/ADN qui pourraient être la conséquence de l'existence de plusieurs formes oligomériques de CggR en solution (Figure 2 ; Article n°1).

Pour mettre en évidence les différents oligomères de CggR, des expériences de pontage chimique au glutaraldéhyde ont été réalisées. CggR (0,1 ou 0,4 mg/mL) a été incubé dans un tampon Tris contenant du glutaraldéhyde à 0,01 ou 0,1 % suivant une cinétique courte (car le glutaraldéhyde réagit avec le Tris). L'analyse des réactions en SDS-PAGE indiquent l'apparition de plusieurs bandes, bien visibles à 0,4 mg/mL de protéine (Figure 43). Les trois bandes majoritaires d'environ 38, 80 et 110 kDa peuvent correspondre à des formes monomérique, dimérique et trimériques de CggR, respectivement. En outre, au moins deux formes de plus haut poids moléculaire (plusieurs centaines de kDA) sont observables dont l'une forme des agrégats qui ne pénètrent pas dans la partie peu concentrée du gel d'acrylamide (Figure 43). Ces résultats suggèrent donc que CggR peut exister en solution sous une forme dimérique ou trimérique, et au moins une autre forme plus complexe.

Analyse du rôle du FBP par résonance plasmonique de surface.

Alors que nos analyses du rôle du FBP par retard de migration électrophorétique avaient donné des résultats certes encourageants mais pas assez convaincants, nous avons entrepris de tester cette interaction au moyen de deux autres techniques. La première employée est celle d'empreinte à la DNase I dont les résultats sont dans l'article n°1. La seconde alternative est une technique très sensible de mesure de l'interaction entre deux espèces moléculaires : la résonance plasmonique de surface ou « Biacore » (cf matériel et méthodes pour le principe).

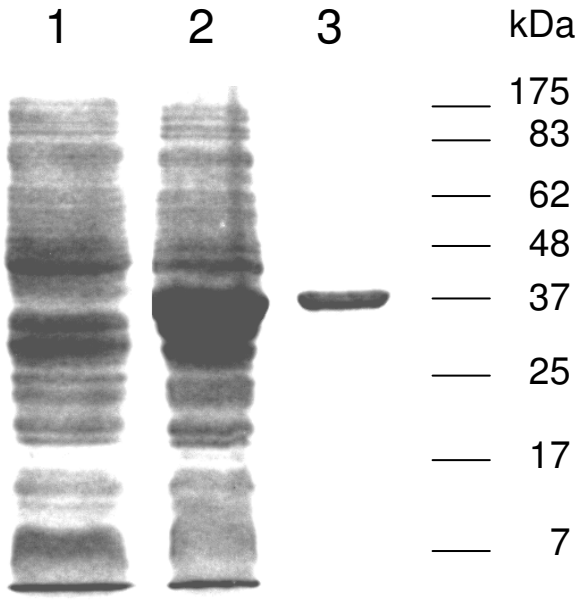


Figure 42. Surproduction et purification de His_6-CggR. Fraction protéique soluble de la souche surproductrice non induite (1) ou induite (2) par addition d'IPTG et protéine His_6-CggR purifiée. Le poids mesuré de CggR est d'environ 37 kDa.

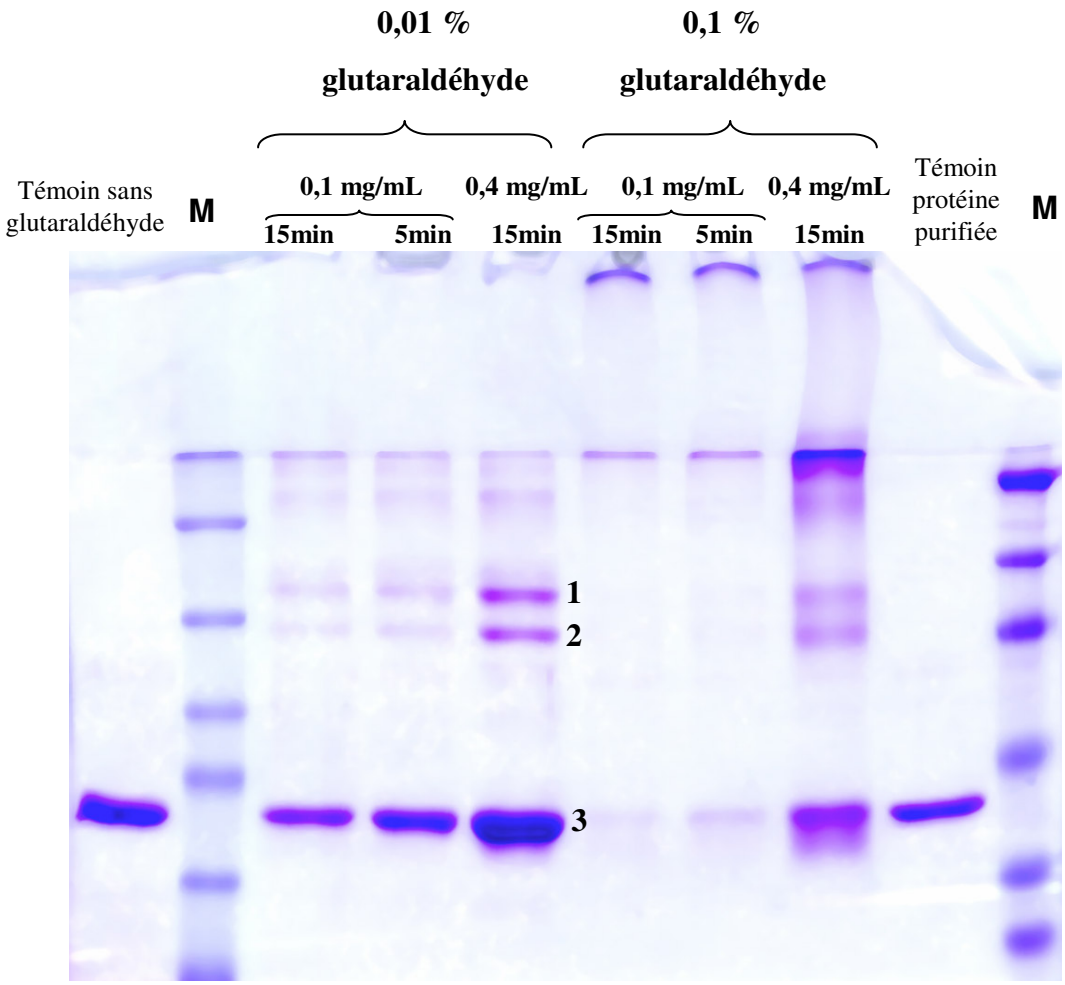


Figure 43. Détermination des oligomères possibles de CggR par pontage chimique au glutaraldéhyde. Les bandes 1 (~38 kDa), 2 (~80 kDa) et 3 (~110 kDa) pourraient correspondre à des formes monomérique, dimérique et trimérique de CggR. **M**: marqueur de poids molculaire.

Dans un premier temps, CggR en solution est accroché sur la surface de mesure recouverte de Ni^{2+}, jusqu'à stabilisation. Lorsque l'ADN portant l'opérateur de CggR est injecté à 10 nM, on observe une augmentation rapide et spécifique de RU (Responsive Unit) qui reste stable tant que l'injection est prolongée. A partir de la fin de l'injection, la RU diminue lentement. L'interaction CggR/ADN est donc clairement détectée. Elle présente une cinétique d'accrochage rapide et une cinétique de décrochage de l'ADN très lente. Ensuite, l'effet des métabolites suivants a été testé : le FBP, le F6P, le G3P et le F26P (fructose-2,6-biphosphate), par injection simultanée d'ADN et de métabolite. Les résultats montrent qu'en présence de 10 mM de FBP, la RU reste « stable » (Figure 44) alors qu'elle augmente de façon comparable au témoin sans métabolite lorsque l'ADN est injecté en présence de la même concentration de n'importe quel autre métabolite testé (Figure 44). Ces résultats indiquent donc que le FBP empêche l'ADN d'interagir avec CggR. Il est intéressant de remarquer que l'ajout de 10 mM de F6P entraîne une faible inhibition de l'interaction CggR/ADN de manière reproductible, ce qui suggère que ce métabolite peut aussi interagir avec CggR bien que de façon beaucoup moins efficace que le FBP.

Dans un second temps, nous avons testé la capacité de CggR à lier l'ADN en fonction de la concentration de FBP. Une gamme de 1 à 10 mM de FBP a donc été injectée avec l'ADN. La RU augmente alors de façon inverse à la quantité de FBP injectée dans la puce (Figure 45), ce qui démontre que le FBP inhibe de manière dépendante de sa concentration l'activité de fixation à l'ADN de CggR.

Etude complémentaire de l'effet du FBP sur l'interaction entre CggR et son ADN cible étudié par retard de migration électrophorétique .

Nos premières expériences qui visaient à tester la capacité du FBP à perturber l'interaction CggR/ADN en retard de migration électrophorétique avaient permis d'obtenir un résultat encourageant mais pas suffisamment convaincant du rôle du FBP. Sur la base de ce résultat, nous avons poursuivi cette étude selon deux alternatives, tout d'abord en utilisant d'autres techniques *in vitro* puis en modifiant légèrement notre protocole expérimental de retard de migration électrophorétique. Nos expériences *in vitro* d'empreinte à la DNase I et de « Biacore » ont permis de montrer que la présence de FBP perturbait la capacité de CggR à se lier à l'ADN. Pour savoir si le FBP a un effet sur la formation de différents complexes

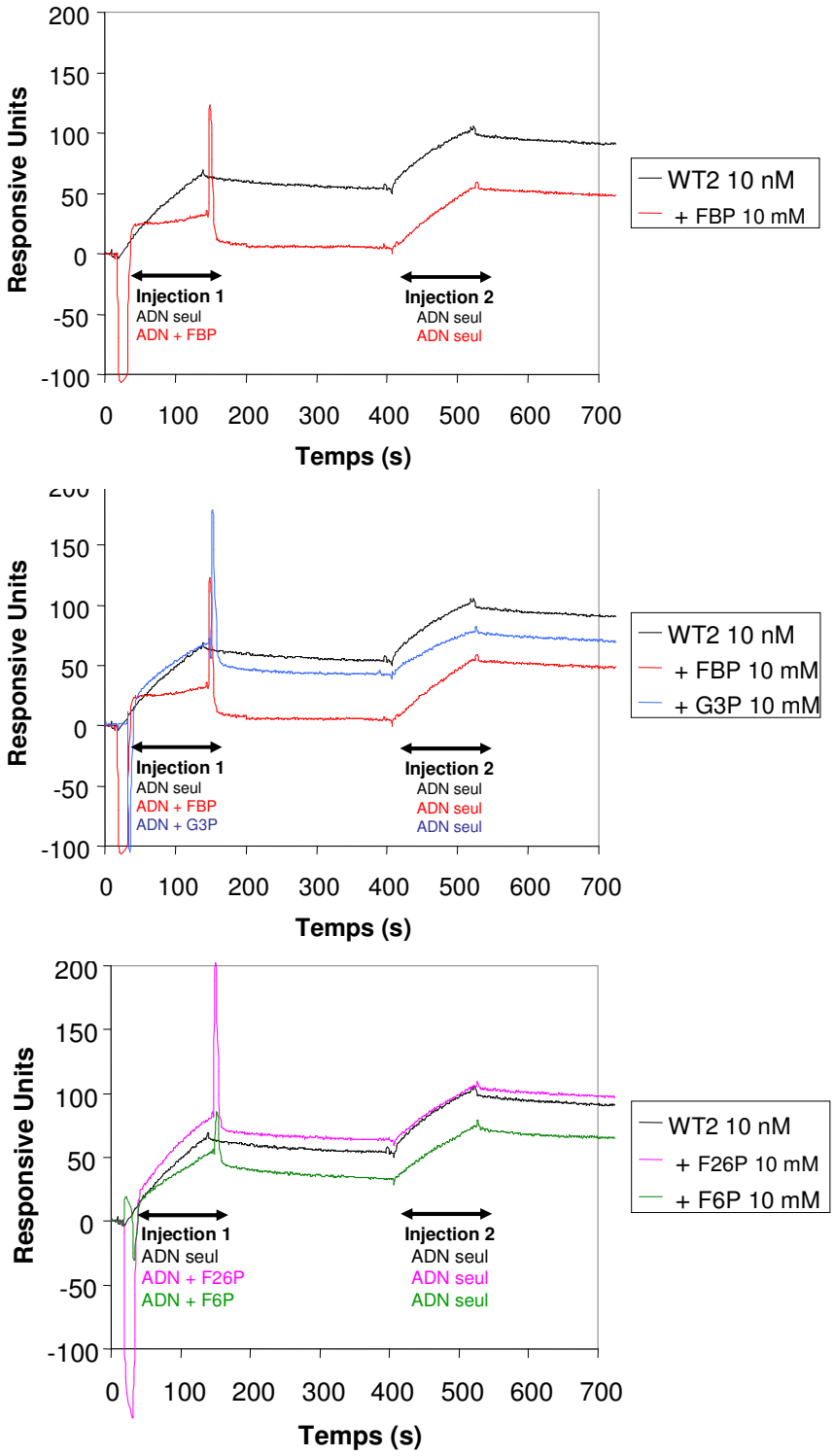


Figure 44. Test de l'effet de différents métabolites de la glycolyse sur l'interaction CggR/ADN par mesure de la variation de résonance plasmonique de surface.

WT : ADN sauvage
FBP : Fructose-1,6-Biphosphate
G3P : Glycéraldéhyde-3-phosphate
F26P : Fructose-2,6-Biphosphate
F6P : Fructose-6-phosphate

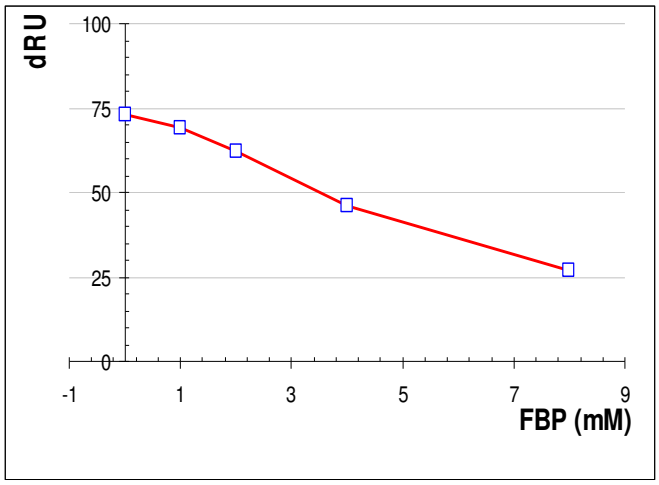


Figure 45. Effet de la concentration en FBP sur l'interaction CggR/ADN (WT) par mesure de la variation de résonance plasmonique de surface.

CggR/ADN, nous avons mis en oeuvre une nouvelle série d'expériences de retard de migration électrophorétique le plus possible identique aux précédentes, à l'exception cette fois-ci, de l'ajout de FBP directement dans le tampon de migration. En effet, si l'affinité de CggR pour le FBP est faible, il est possible que ce dernier soit très rapidement dilué lors du dépôt des réactions dans les puits du gel d'électrophorèse. Ce problème ne se pose pas lors des expériences d'empreinte à la DNase I ou de Biacore.

Dans cette nouvelle série de réactions, 10 mM de FBP ont été ajoutés dans le tampon de migration ainsi que dans l'un des deux gels réalisés. Une gamme croissante de concentration de CggR a été préincubée avec 10 mM de FBP avant l'incubation avec la sonde ADN et le dépôt dans le gel d'électrophorèse. Cependant, l'expérience témoins la plus appropriée, c'est-à-dire la même expérience sans FBP dans le tampon et réalisée en parallèle, n'a pas été effectuée parce qu'il n'était pas possible de séparer le tampon de migration des deux gels dans l'appareil d'électrophorèse. L'expérience témoins sans FBP dans la Figure 46 a été réalisé juste après les expériences avec le FBP, mais donc pas en parallèle. Les résultats de cette nouvelle série d'expérience montrent que le FBP n'a pas d'effet sur la quantité de CggR nécessaire pour retarder 50 % de la sonde marquée (Figure 46). En revanche, ils indiquent une modification du nombre de types de complexes formés en présence de FBP de façon dépendante de la concentration en protéine. En présence de FBP, une bande migrant rapidement (A) est visible dès les petites quantités jusqu'à des quantités relativement grandes de CggR (Figure 46) alors qu'elle n'apparaît que pour des quantités très faibles de protéines en l'absence de FBP (Figure 2 ; Article n°2). Puis, au moins deux bandes sont visibles pouvant correspondre à des complexes CggR/ADN de stœchiométries différentes et dont la plus « lente » (B) (à migrer) semble correspondre à la bande majeure visible dans l'expérience sans FBP (Figure 46). Néanmoins, pour évaluer précisément la nature de ces complexes, il faudrait procéder à un marquage de la protéine de façon à pouvoir, après prélèvement de chacune des bandes « retardées », mesurer les quantités relatives de protéine et d'ADN présentes. Ces résultats sont à interpréter avec beaucoup de précautions, d'une part parce que la bande « rapide » n'est pas toujours visible dans nos expériences faites en l'absence de FBP suggérant une certaine variabilité inhérente à la manipulation, et d'autre part parce la quantité de CggR nécessaire pour retarder 50 % de la sonde peut légèrement varier d'une série d'expériences à une autre. La seule différence clairement observée dans cette nouvelle série de retard de migration électrophorétique se trouve donc au niveau des différents types de complexes formés entre CggR et l'ADN. Nous proposons dans la discussion générale, sur la

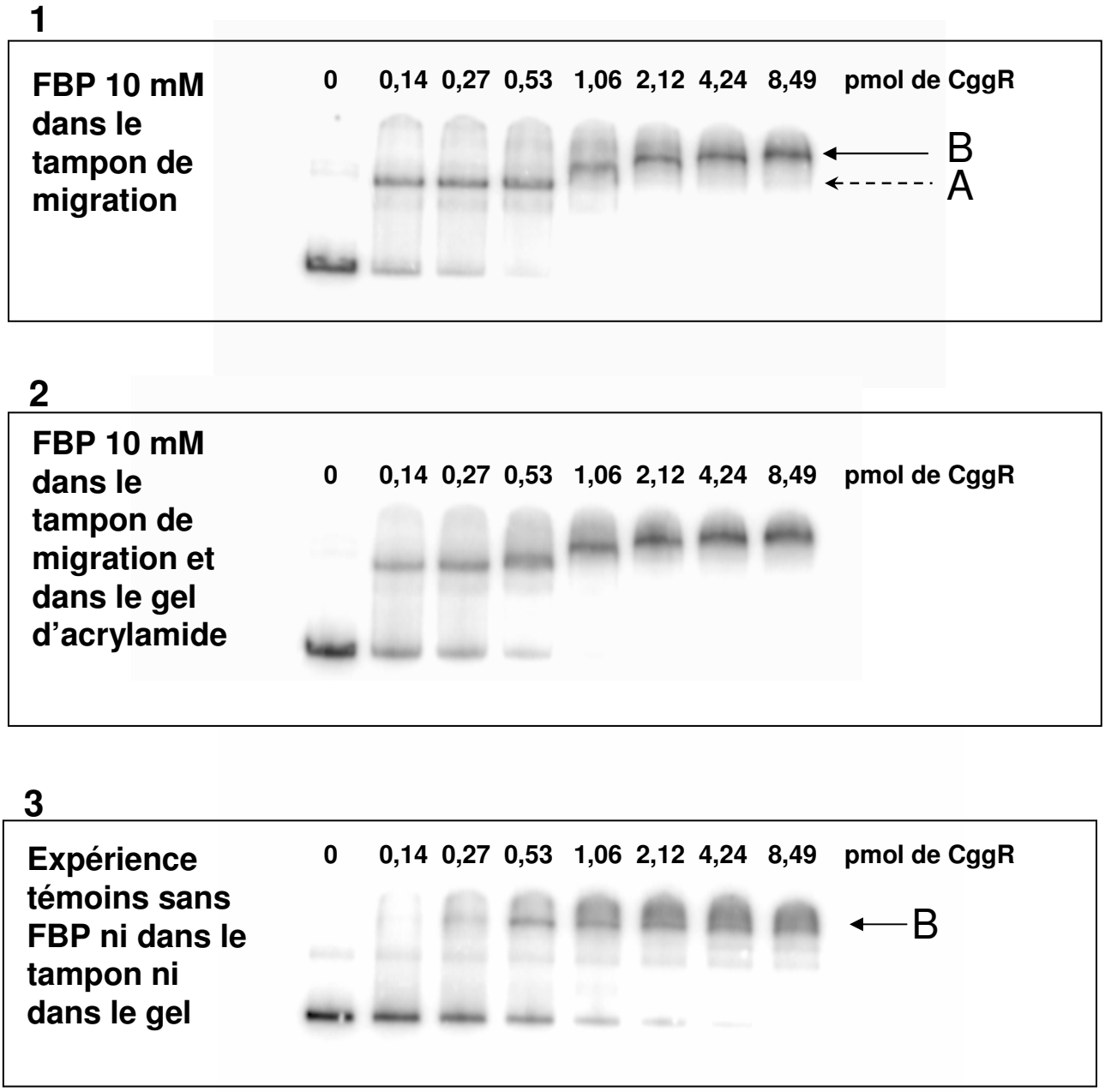


Figure 46. Test de l'effet du Fructose-1,6-Biphosphate (FBP) sur l'interaction CggR/ADN en retard de migration électrophorétique. Deux bandes distinctes sont visibles en présence de FBP, correspondant à un petit complexe CggR/ADN en présence de faible concentration de protéine (A) et un complexe de plus haut poids moléculaire en présence de forte concentration de FBP (B).

base de l'ensemble de nos résultats, un modèle possible rendant compte du mode d'action du FBP sur l'activité du répresseur CggR. Des expériences complémentaires seront néanmoins nécessaires pour confirmer ce qui n'est encore qu'un modèle « de travail ».

4.3.2. Matériel et Méthodes

Détermination du site d'initiation de la transcription.

Les ARN totaux utilisés ont été extraits (Article n°1) à partir de bactéries cultivées dans un milieu minimum CQT (Aymerich *et al.*, 1986) supplémenté avec 1 % de glucose ou 1 % de succinate plus glutamate jusqu'à environ DO 0,5. 20 pmol de l'oligonucléotide +1gapA-1 ont été marqués par incubation en présence de T4 polynucléotide kinase (Biolabs) et de 30 µCi de [γ-^{32}P]-ATP (ICN). Par réaction, 1 pmol d'oligonucléotide a été hybridé avec 20 µg d'ARN totaux. Puis la réverse-transcriptase (Pharmacia) a été utilisée pour la synthèse des ADNc. L'oligonucléotide +1gapA-1 a aussi été utilisé pour réaliser les réactions de séquence avec le kit Sequenase 2.0 (Amersham). Les échantillons d'ADNc et les réactions de séquence ont été analysées (gel d'acrylamide à 6 %). Puis, la radioactivité a été révélée par autoradiographie.

Southern Blot.

L'ADN génomique de la souche sauvage 168CA et du transformant à analyser a été extrait à partir de bactéries cultivées en milieu LB (Luria-Bertani Broth) jusqu'à DO_{600nm} 0,5 et quantifié par spectrophotométrie. Pour chaque vérification, 1 µg d'ADN a été digéré avec les enzymes de restriction appropriées, comme recommandé par les différents fournisseurs, et les produits de la réaction ont été séparés dans un gel d'agarose à 1 %. L'ADN a ensuite été transféré sur membrane Hybond-N (Amersham) et fixé par l'action des UV. La membrane a été incubée dans une solution de préhybridation (formamide 50 %, SSC 5X, Denhardt 5X, SDS 1 %, ADN de sperme de hareng 1 mg/mL) à 37°C pendant 2 h. Les sondes radioactives ont été préparées à l'aide du kit Megaprime (Biolabs). Pour chaque sonde, 100 ng de fragment PCR purifiés et 4 µL de [α-^{32}P]-dCTP (ICN) ont été utilisés. 2 ng d'ADN du marqueur de

taille λ*Eco*RI/*Hind*III ont été directement rajoutés à la réaction. Après dénaturation, les deux sondes ont été mélangées à la solution de préhybridation dans laquelle baignait la membrane et l'hybridation a été prolongée pendant 12 h à 37°C. La membrane a ensuite été lavée (SSC 1X, SDS 0,1 %, 10 min à 37°C; SSC 1X, SDS 0,1 %, 10 min à 45°C ; SSC 0,1X, SDS 0,1 %, 30 min à 55°C). Puis, elle a été scellée et exposée en cassette de PhosphoImager. L'exposition a été révélée au bout de 24 h.

Pontage chimique au glutaraldéhyde.

CggR (à 0,1 ou 0,4 mg/mL final) a été incubé avec du glutaraldéhyde (à 0,01 ou 0,1 % final) dans un volume final de 100 µL dans un tampon Tris 50 mM pH 7,4 pendant 5 ou 15 min. Les réactions de pontage ont été stoppées avec 400 µL de $NaBH_4$ 5%. Les protéines ont été précipitées au TCA 10 % final, incubées 1 h dans la glace, puis culotées par centrifugation. Le culot a été rincé avec 300 µL d'acétone, séché, puis resuspendu dans 20 µL de tampon de dépôt 2X. La solution a revêtu une teinte jaunâtre due à un pH acide. Le pH a été relevé avec une goutte de Tris 50 mM pH 8. Les réactions ont ensuite été dénaturées par la chaleur et analysées en SDS-PAGE (gel à 10 %).

« Biacore ».

Le principe de cette technique est basée sur un phénomène de résonance plasmonique d'une surface qui transforme l'angle d'incidence spécifique de la lumière réfléchie sur une surface métallique en réponse à la présence de molécules attachées à cette surface. Dans la pratique, il s'agit d'une surface intégrée dans une puce qui permet la circulation d'un flux de soluté. Les mesures données par un instrument Biacore sont arbitrairement exprimées en Responsive Units (RU), ou unités de réponse, qui sont proportionnelles à la quantité de molécules liées à la surface de la puce. Ainsi, le suivi de la résonance plasmonique de surface permet d'observer l'interaction entre deux espèces moléculaires, l'une fixée à la surface et l'autre passant dans le flux de soluté du circuit alimentant la surface.

Pour notre étude, une Chip NTA (Biacore) et un appareil Biacore X ont été utilisés. Cette puce possède deux canaux séparés et les mesures ont été réalisées de façon différentielle

entre les résonances des canaux 1 et 2 (ΔRU=RU canal2 –RU canal1), le premier canal constituant une référence (en fait, le flux de solutés passe d'abord dans le canal 1 puis dans le canal 2 ou dans un seul des deux). Pour les deux canaux, la surface de mesure a d'abord été saturée en ions Ni^{2+} par injection d'une solution de $NiCl_2$ 0,5 mM. Puis, une solution de Tris-HCl 50 mM pH 8 contenant His_6-CggR à 100 nM a été injectée uniquement dans le canal 2 jusqu'à atteindre environ 3000 RU. Par simple affinité, les molécules de CggR ont été attachées sur la surface aux ions Ni^{2+} . L'ADN cible double brin utilisé a été obtenu par hybridation de deux oligonucléotides de 56 nucléotides de séquences complémentaires ((TTTTTTTggCTATgA**CgggAC**gTTTTT**TgTC**ATAg**CgggAC**ATATAA**TgTC**CAgCA séquence de l'ADN cible). Juste avant injection, cet ADN est chauffé pour détruire les structures conformationnelles éventuelles. Après stabilisation de la ΔRU, 10 µL d'une solution de Tris-HCl 50 mM pH 8 contenant 10 nM d'ADN a été injectée dans les deux canaux. Cela a provoqué une augmentation de la ΔRU correspondant à une interaction entre les molécules de CggR à la surface et l'ADN que ces premières ont complexé. Les complexes protéines/ADN peuvent être dissociés par injection d'une solution de $NaCl_2$ 0,5 M. Pour tester l'effet des métabolites, 10 nM d'ADN cible ont été préincubés avec 10 mM du métabolite à tester pendant 5 min dans la glace et 10 µL de cette solution ont été injectés dans les deux canaux.

5. Etude transcriptionnelle et physiologique d'*ywkA*, l'un des quatre gènes paralogues pouvant coder des enzymes maliques chez *Bacillus subtilis*

5.1. Présentation synthétique des résultats

L'enzyme malique catalyse l'interconversion du malate et du pyruvate. En combinaison avec la pyruvate carboxylase, elle permet de produire de l'oxaloacétate à partir du malate. Lorsque la bactérie utilise le malate comme seule source de carbone (ou une autre source préalablement transformée en malate), ces deux réactions constituent la seule alternative permettant de court-circuiter la réaction catalysée par la malate déshydrogénase (Figure 47). En combinaison avec l'activité pyruvate déshydrogénase, l'enzyme malique permet aussi de régénérer le pool d'acétyl-CoA (Figure 47).

Au début de mon travail, aucun gène n'avait été clairement identifié pour cette fonction chez *Bacillus subtilis*. Or, l'analyse de la séquence complète du chromosome de *Bacillus subtilis* révèle l'existence de quatre gènes, *ywkA*, *malS*, *ytsJ* et *mleA*, dont les produits ressemblent à des enzymes maliques. En outre, ces gènes peuvent être répartis en deux groupes sur la base de leurs séquences : YwkA et MalS présentent 60 % d'identité, YtsJ et MleA 40 % d'identité (et au mieux 28 % d'identité entre les protéines des deux groupes). Par ailleurs, le gène *mleA* est organisé en opéron avec le gène *mleN*, qui code un antiporteur malate-lactate (Wei *et al.*, 2000). Il est donc très probable que *mleA* code en fait une enzyme malolactique associée à ce transporteur. Nous avons donc laissé l'étude de ce gène de côté, du moins dans un premier temps.

Nos collègues japonais ont réalisé une comparaison transcriptomique de la souche sauvage 168CA de *Bacillus subtilis* cultivée en conditions glycolytique ou néoglucogénique qui a donné une piste quant à l'étude du gène *ywkA*. En effet, les résultats de cette étude révèlent que la transcription d'*ywkA* est 7 fois plus forte lors de la croissance en présence de malate que lors de la croissance en présence de glucose comme seule source de carbone. Toujours selon cette étude, les gènes *malS*, *ytsJ* (et *mleA*), ne seraient pas différentiellement exprimés en fonction de ces conditions. Par la suite, une fusion rapportrice de l'activité du promoteur en amont d'*ywkA* a été construite par Pascale Servant. Nous l'avons utilisée pour montrer que l'expression d'*ywkA* est spécifiquement induite par le malate extracellulaire. De

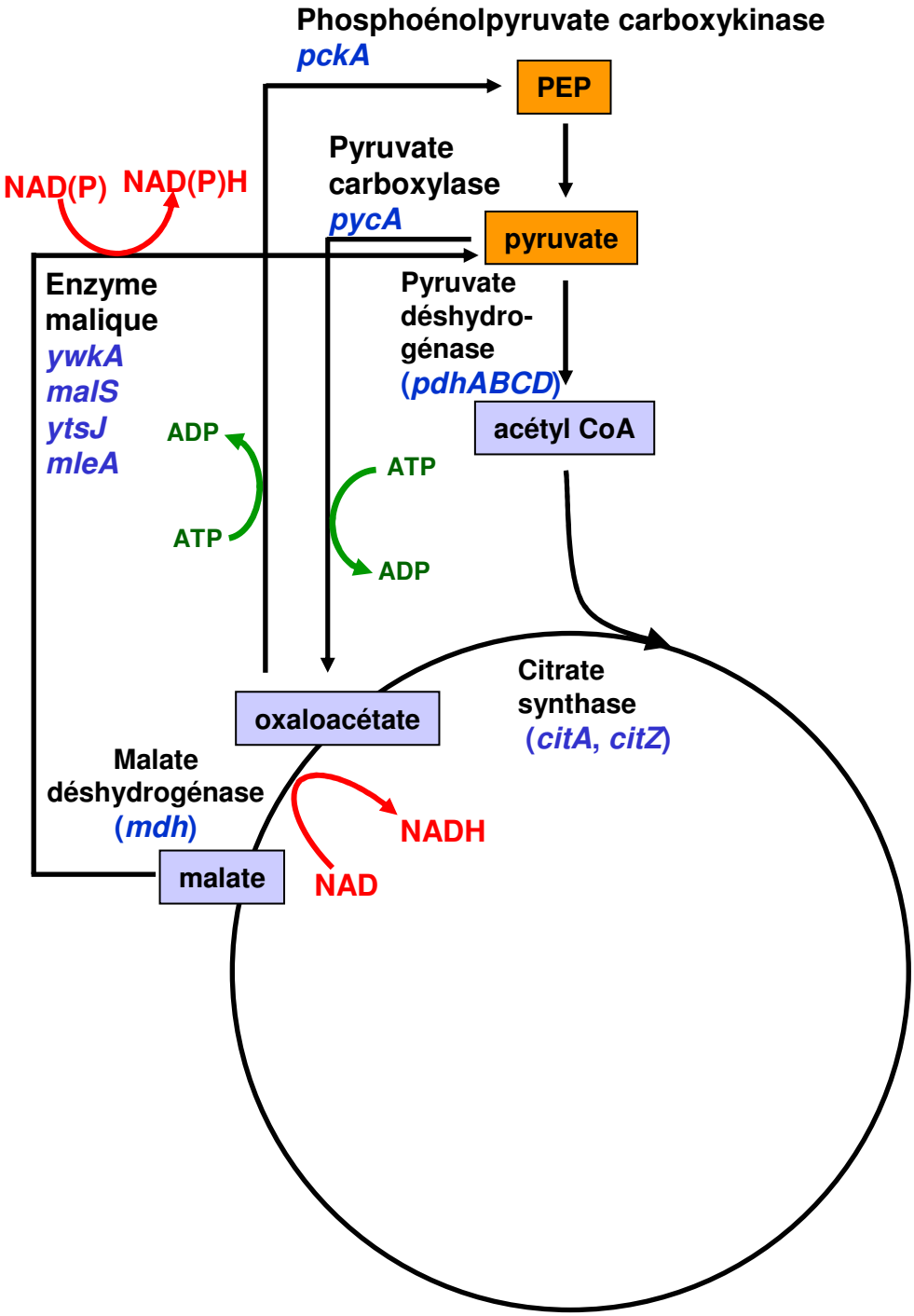


Figure 47. L'étape enzyme malique et son contexte métabolique chez *Bacillus subtilis*.

plus, nous avons montré que cette régulation est indépendante de la répression catabolique en présence de glucose. L'induction de l'expression d'*ywkA* par le malate a été confirmée par des Northern Blot réalisés en parallèle avec la comparaison transcriptomique. Cette expérience permet aussi de démontrer qu'*ywkA* est cotranscrit avec le gène situé en aval, *ywkB*. Le site de démarrage de la transcription a aussi été déterminé et a permis de trouver les séquences –10 et –35 correspondant à un promoteur σ^A-dépendant.

Comme *ywkA* est spécifiquement exprimé en présence de malate, nous avons voulu savoir s'il était impliqué dans l'utilisation de cette source de carbone pour la croissance. Un mutant *ywkA* a donc été construit et sa capacité à croître en présence de différentes sources de carbone a été testée. Toutefois, ce mutant pousse aussi bien que la souche sauvage quelque soit la source de carbone présente dans le milieu. J'ai alors construit des souches mutées pour l'un des paralogues (*malS* ou *ytsJ*), deux paralogues (*ywkA-malS*, *ywkA-ytsJ*, *malS-ytsJ*), ou les trois à la fois. Puis, de même, j'ai testé le phénotype de ces souches pour la croissance en présence de glucose ou de malate. Les résultats de ces tests indiquent que ni *ywkA* ni *malS* ne sont utiles pour la croissance en présence de malate comme seule source de carbone mais que, par contraste, *ytsJ* joue un rôle majeur dans ces conditions. Par ailleurs, aucun de ces gènes ne semble impliqué dans l'utilisation du glucose comme source de carbone puisqu'aucun mutant n'est affecté lors de la croissance en présence de glucose.

Mais *ywkA* code-t-il bien une enzyme malique ? Pour répondre à cette question, j'ai surexprimé et purifié YwkA par fusion avec une étiquette de 6 histidines en N-terminal. Puis, j'ai déterminé les caractéristiques enzymatiques d'YwkA. Les résultats de ces expériences démontrent qu'*ywkA* code bien une enzyme malique dont l'activité est 10 fois plus forte en présence de NAD qu'en présence de NADP (k_{cat}/K_M = 102 and 10 $s^{-1}mM^{-1}$, respectivement). En parallèle, Guillaume Lerondel et moi avons entrepris de réaliser le même travail avec les produits des gènes *malS* et *ytsJ*. MalS et YtsJ ont donc été purifiées et biochimiquement analysées selon les mêmes techniques. Cette seconde série de résultats, qui n'est pas présentée dans l'article n°2, mène aux conclusions suivantes : MalS est une enzyme malique dont les caractéristiques sont très similaires à celles d'YwkA, utilisant donc préférentiellement le NAD comme cofacteur ; au contraire, YtsJ est une enzyme malique clairement NADP dépendante. Par ailleurs, Guillaume a aussi étudié le profil d'expression de *malS* et d'*ytsJ*. Une partie de ces résultats, qui permettent d'avoir une vision plus globale du sujet, sera présentée à la suite de l'article n°2.

Pour tenter de trouver un rôle à *ywkA*, malgré l'absence de phénotype de nos mutants, nous avons étudié plus avant la régulation de l'expression d'*ywkA*. Guillaume, Pascale et Stéphane ont réalisé un crible génétique après mutagenèse par transposition pour identifier un ou plusieurs régulateurs impliqués dans cette régulation. Cette expérience leur a permis de suggérer que le système à deux composants *yufL*/*yufM*, non encore caractérisé, intervenait dans le mécanisme qui contrôle l'expression d'*ywkA*. Par la suite, j'ai effectué une analyse génétique qui montre que *yufL*/*yufM* est nécessaire à l'induction de l'expression d'*ywkA* par le malate. J'ai également réalisé des expériences *in vitro* qui montrent que YufM est capable d'interagir avec une séquence située juste en amont du promoteur d'*ywkA* et d'activer sa transcription en réponse au signal donné par YufL qui percevrait la présence de malate extracellulaire.

5.2. **Article n°2 :** The *Bacillus subtilis ywkA* gene encodes a malic enzyme and its transcription is activated by the YufL-YufM two component system in response to malate.

5.2.1. Article

Microbiology (2003), **149**, 2331–2343 DOI 10.1099/mic.0.26256-0

The *Bacillus subtilis ywkA* gene encodes a malic enzyme and its transcription is activated by the YufL/YufM two-component system in response to malate

Thierry Doan,[1] Pascale Servant,[1]† Shigeo Tojo,[2] Hirotake Yamaguchi,[2] Guillaume Lerondel,[1] Ken-Ichi Yoshida,[2] Yasutaro Fujita[2] and Stéphane Aymerich[1]

[1]Génétique Moléculaire et Cellulaire, INRA (UMR216) CNRS (URA1925) and INAP-G, F-78850 Thiverval-Grignon, France

[2]Department of Biotechnology, Fukuyama University, 985 Sanzo, Higashimura, Fukuyama, Japan

Correspondence
Stéphane Aymerich
stef@grignon.inra.fr

Received 24 January 2003
Revised 13 May 2003
Accepted 1 July 2003

A transcriptome comparison of a wild-type *Bacillus subtilis* strain growing under glycolytic or gluconeogenic conditions was performed. In particular, it revealed that the *ywkA* gene, one of the four paralogues putatively encoding a malic enzyme, was more transcribed during gluconeogenesis. Using a *lacZ* reporter fusion to the *ywkA* promoter, it was shown that *ywkA* was specifically induced by external malate and not subject to glucose catabolite repression. Northern analysis confirmed this expression pattern and demonstrated that *ywkA* is cotranscribed with the downstream *ywkB* gene. The *ywkA* gene product was purified and biochemical studies demonstrated its malic enzyme activity, which was 10-fold higher with NAD than with NADP (k_{cat}/K_m 102 and 10 s^{-1} mM^{-1}, respectively). However, physiological tests with single and multiple mutant strains affected in *ywkA* and/or in *ywkA* paralogues showed that *ywkA* does not contribute to efficient utilization of malate for growth. Transposon mutagenesis allowed the identification of the uncharacterized YufL/YufM two-component system as being responsible for the control of *ywkA* expression. Genetic analysis and *in vitro* studies with purified YufM protein showed that YufM binds just upstream of *ywkA* promoter and activates *ywkA* transcription in response to the presence of malate in the extracellular medium, transmitted by YufL. *ywkA* and *yufL*/*yufM* could thus be renamed *maeA* for malic enzyme and *malK*/*malR* for malate kinase sensor/malate response regulator, respectively.

INTRODUCTION

Malic enzymes catalyse the reversible oxidative decarboxylation of malate into pyruvate with reduction of NAD or NADP. In eukaryotes, several isoforms of this enzyme have been characterized. Three classes of malic enzymes have been defined on the basis of their coenzyme specificity and their capability to also decarboxylate oxaloacetate (Kobayashi *et al.*, 1989). A first class of NADP-dependent malic enzymes (EC 1.1.1.40), decarboxylating malate and also oxaloacetate, are found in cytosol, chloroplasts and mitochondria; a second group of enzymes, preferentially utilizing NAD (EC 1.1.1.38), are also capable of decarboxylating oxaloacetate and are found in bacteria and insects; finally, a third class of NAD-dependent malic enzymes (EC 1.1.1.39), unable to use oxaloacetate, have been found only in mitochondrial matrix. In prokaryotes, malic enzymes are also widely distributed (Murai *et al.*, 1971; Diesterhaft & Freese, 1973; Lamed & Zeikus, 1981; Knichel & Radler, 1982; Bartolucci *et al.*, 1987; Kobayashi *et al.*, 1989; Kawai *et al.*, 1996; Voegele *et al.*, 1999). However, few of them have been characterized so far, and in particular, few reports allowed distinction of whether the observed malic enzyme activity resulted from one or several isozymes. In *Escherichia coli*, both NAD- and NADP-dependent malic enzymes, named SfcA and B2463 respectively, have been characterized (Murai *et al.*, 1971; Hansen & Juni, 1974, 1975; Stols & Donnelly, 1997) but their respective physiological functions have not been precisely documented. It has been suggested that SfcA is involved in gluconeogenesis by providing pyruvate, then converted into

†Present address: Institut de Génétique et Microbiologie, Université Paris-Sud, Bât. 409, F-91405 Orsay, France.

Abbreviations: TCA, tricarboxylic acid; TCS, two-component system.

The online version of this paper (at http://mic.sgmjournals.org) contains a table listing the oligonucleotides used in this study and a figure showing the overproduction and purification of His_6-YufM.

phosphoenolpyruvate by a phosphoenolpyruvate synthase (Hansen & Juni, 1975), and that B2463 is involved in the supply of NADPH and acetyl-CoA, necessary for anabolic reactions during growth on C_4-dicarboxylates as sole carbon sources (Murai *et al.*, 1971; Hansen & Juni, 1974, 1975). The synthesis of both enzymes is repressed in the presence of glucose while the expression of *sfcA*, encoding the NAD-dependent malic enzyme, is inducible by malate (Murai *et al.*, 1971). The combined deletion of the two genes leads to a severe growth defect only on C_4-dicarboxylates, thus indicating that malic enzyme activity is necessary under such growth conditions in *E. coli* (van der Rest *et al.*, 2000). *Rhizobium meliloti*, a nitrogen-fixing Gram-negative symbiont of alfalfa, also synthesizes two malic enzymes named DME and TME, which are NAD- and NADP-dependent, respectively (Driscoll & Finan, 1997; Mitsch *et al.*, 1998). The former has been shown to be involved in symbiotic nitrogen fixation as a result of its role for growth in the nodule, where C_4-dicarboxylates are the major carbon and energy source provided by the host (Driscoll & Finan, 1993; Dunn, 1998), while the latter has been proposed to function as a generator of NADPH necessary for biosynthesis reactions (Driscoll & Finan, 1996).

Interestingly, the analysis of the *Bacillus subtilis* complete genome sequence revealed the presence of four paralogous genes encoding putative malic enzymes: *ywkA*, *malS*, *ytsJ* and *mleA*. None of these genes has been previously characterized. Indeed, little information has been reported about malic enzyme activity in this bacterial species. Diesterhaft & Freese (1973) demonstrated that the presence of malate stimulates the synthesis of an NADP/NAD-dependent malic enzyme activity.

Amino acid sequence comparisons indicate that the *ywkA* and *malS* products, on the one hand, and the *ytsJ* and *mleA* products, on the other hand, belong to two distinct subgroups (60 % identity between YwkA and MalS; 40 % identity between YtsJ and MleA; but 28 % identity between YtsJ and MalS). Furthermore, YwkA and MalS share 40 % identity with SfcA, an NAD-dependent malic enzyme which was found to be more similar to malolactic enzymes than to other malic enzymes (Groisillier & Lonvaud-Funel, 1999). However, malic and malolactic enzymes are difficult to distinguish on the basis of the amino acid sequence because they are highly similar and harbour common conserved sequence motifs (Denayrolles *et al.*, 1994; Kawai *et al.*, 1996; Groisillier & Lonvaud-Funel, 1999) and because only very few malolactic enzymes have been enzymically characterized (Ansanay *et al.*, 1993; Arthurs & Lloyd, 1999). Recently, a function for *mleA* in *B. subtilis* has been proposed by Wei *et al.* (2000). *mleA* is the second gene of a bicistronic operon including the *mleN* gene. *mleN* has been demonstrated to encode an antiporter that achieves simultaneous Na^+/H^+ and malate/lactate antiport. Therefore, MleA is strongly predicted to be a malolactic enzyme.

The examination of the sequence of the *B. subtilis* chromosome indicates that *ywkA* is probably cotranscribed with *ywkB*, which encodes a protein of unknown function. *malS* is linked to *ytnP*, an unknown gene, and *ytsJ* is linked to *dnaE*, which encodes the second essential DNA polymerase (Dervyn *et al.*, 2001). Thus, the genetic organization of these three genes did not give any precise clue about their possible physiological roles.

Bacteria have to adapt to frequent changes in nutritional availability and, more generally, in environmental conditions. This is particularly true for free-living soil bacteria like *B. subtilis*. To accommodate gene expression to environmental conditions, two-component sensor–regulator systems are frequently used by bacteria. The two-component systems (TCSs) are composed of sensory protein kinases and of response regulator proteins. The protein kinases are typically located in the cytoplasmic membrane and comprise an N-terminal sensory and a conserved C-terminal transmitter domain. Binding of the signal ligand to the sensory domain induces ATP hydrolysis and autophosphorylation of the transmitter domain. The phosphoryl group is subsequently transferred to the cognate response regulator, leading to the activation of its regulatory property, most commonly transcriptional activation. More than 30 TCSs can be predicted from the complete *B. subtilis* genome sequence. However, a precise adaptative role has been identified for only a few of them (Perego & Hoch, 2002). Among them, two, YdbFG (DctSR) (Asai *et al.*, 2000) and CitST (Yamamoto *et al.*, 2000), appeared to be able to sense the presence of tricarboxylic acid (TCA) cycle intermediates in the medium and to activate the expression of genes encoding transporters for these compounds. Furthermore, a comprehensive DNA microarray analysis of the *B. subtilis* TCSs has suggested that *ywkA* could be a target of YufL/YufM (Kobayashi *et al.*, 2001).

The starting point of the study reported here is a global transcriptome comparison of a wild-type *B. subtilis* strain growing in minimal medium containing glucose or malate as sole carbon source. Among the genes which appeared to be more transcribed on malate than on glucose was *ywkA*, but not *malS* or *ytsJ*. The regulation of the expression of the *ywkA* gene has therefore been further investigated, and the uncharacterized YufL/YufM TCS has been demonstrated to control directly *ywkA* transcription in response to the presence of malate. The enzymic activity encoded by *ywkA* has been characterized. Finally, single and multiple mutant strains for *ywkA* and the other putative malic enzyme genes have been constructed and studied to investigate the physiological role of YwkA.

METHODS

Bacterial strains and growth conditions. *E. coli* TG1 [K12 Δ(*lac–pro*) *supE thi hsd-5*/F′ *traD36 proA*$^+$*B*$^+$ *lacI*^{q} *lacZ*ΔM15] was used as a general cloning host and *E. coli* strain M15 pREP4 (Qiagen) was used for overproduction of His_6-YufM and His_6-YwkA. *E. coli* strains were grown in Luria–Bertani broth (LB) supplemented with antibiotics when necessary (ampicillin, 100 mg l^{-1};

kanamycin, 25 mg l^{-1}). Conventional calcium-shock or electroporation procedures were used for transformation (Sambrook *et al.*, 1989).

The *B. subtilis* strains used in this work are listed in Table 1. Standard procedures were used to transform *B. subtilis* (Anagnostopoulos & Spizizen, 1961). *B. subtilis* strains were grown in LB, except when stated otherwise. Antibiotics for selection were added at 5 mg l^{-1} (chloramphenicol), 0·4 mg l^{-1} (erythromycin), 100 mg l^{-1} (spectinomycin), or 5 mg l^{-1} (kanamycin). All cultures were performed at 37 °C. Growth tests were performed in MM minimal medium [K_2HPO_4 85 mM, KH_2PO_4 40 mM, $(NH_4)_2SO_4$ 15 mM, trisodium citrate 6 mM, $MgSO_4$ 0·8 mM, pH 7] supplemented with tryptophan (0·005 %) and carbon sources (0·5 %). Cultures for β-galactosidase assays were performed in CQTHC minimal medium [C minimal medium (Aymerich *et al.*, 1986) supplemented with 0·005 % tryptophan, 0·05 % glutamine, 0·05 % casein hydrolysate] supplemented with 1 % glucose, 1 % glycerol, 1 % fumarate, 1 % malate, or 1 % succinate plus 1 % glutamate.

DNA manipulation. PCR products were purified with the QIAquick PCR purification kit (Qiagen). Plasmid DNA was extracted and purified from *E. coli* with the QIAprep spin miniprep kit (Qiagen). Restriction enzymes and T4 DNA ligase were used as recommended by the manufacturers. Pyrobest DNA polymerase (Takara) was used for preparative polymerase chain reactions as recommended by the manufacturer. DNA fragments were purified from agarose gels using the Qiagen Gel Extraction Kit.

Oligonucleotides. The oligonucleotides used in this study are listed in a table available as supplementary data with the online version of this paper (http://mic.sgm.journals.org).

Construction of strains. A fusion of a *ywkA* promoter fragment to a promoter-less *lacZ* gene was constructed using the vector pDG1661 (Guérout-Fleury *et al.*, 1996). The 277 bp promoter fragment to be tested was obtained by PCR using primers PS69 and PS70 (see supplementary table); this fragment was inserted between the *Eco*RI and *Bam*HI sites of pDG1661 to generate plasmid pPS61, then used to insert the reporter construct into the *B. subtilis amyE* locus to generate strain PS1716 (Table 1).

In vivo replacement by homologous recombination of a kanamycin-resistance cassette for the *lacZ* ORF and erythromycin resistance gene of pMUTIN was obtained using pEC23, a plasmid specifically designed for this use, constructed by P. Stragier, Institut de Biologie Physico-Chimique, 75005 Paris, France.

Strain GTD110 was constructed by replacing the *malS* ORF by the spectinomycin-resistance cassette contained between the *Bam*HI and *Xba*I sites of plasmid pIC333 (Steinmetz & Richter, 1994): two PCR fragments, one corresponding to the 500 bp upstream of *malS* ORF and one to the 500 bp downstream of the *malS* ORF, were generated with primer pairs EagI-ytnP/XbaI-malS and BamHI-ytzB/EcoRI-ytzB (see supplementary table), respectively. The two PCR fragments were then ligated together with the spectinomycin cassette flanked by the *Bam*HI and *Xba*I restriction sites into vector pJH101 (Ferrari *et al.*, 1982) to generate plasmid pTD101, which was used to transform *B. subtilis*.

RNA isolation and Northern blot analysis. For RNA preparation, cell pellets were obtained from a mid-exponential-phase (OD_{600} 0·8) culture in S6 medium (Fujita & Freese, 1981) containing 25 mM glucose or 37·5 mM malate. Total RNAs were isolated from cells essentially as previously described (Yoshida *et al.*, 2001). For Northern blot analysis, RNA samples (2 μg each) were electrophoresed in

Table 1. *B. subtilis* strains used in this study

Strain	Relevant genotype	Source
168CA	*trpC2*	Lab. stock
PS1716	*amyE′* : : P*ywkA′* : : *lacZ–cat*	pPS61→168CA
YWKAd	*ywkA′* : : pMUTINT3(*ery*)	JFAN*
BFA91	*ytsJ′* : : pMUTIN2	BSFA† collection
GTD102	*ywkA′* : : *kan*	pEC23→YWKAd
GTD110	Δ*malS* : : *spc*	This study
GTD111	*ywkA′* : : *kan* Δ*malS* : : *spc*	GTD102 DNA × GTD110
GTD121	*ytsJ′* : : pMUTIN2 *ywkA′* : : *kan*	BFA91 DNA × GTD102
GTD122	*ytsJ′* : : pMUTIN2 Δ*malS* : : *spc*	BFA91 DNA × GTD110
GTD123	*ytsJ′* : : pMUTIN2 *ywkA′* : : *kan* Δ*malS* : : *spc*	BFA91 DNA × GTD111
1A250	*alsR1 ivl*Δ*1*	*Bacillus* Genetic Stock Center
1A147	*ccpA1 alsR1 ivl*Δ*1*	*Bacillus* Genetic Stock Center
C1	*yufL′* : : Tn*10 amyE′* : : P*ywkA′* : : *lacZ–cat*	This study
BFA1431	*yufL′* : : pMUTIN2	BSFA collection
GTD133	*yufL′* : : [pMUTIN2 Δ(*lacZ–ery*) : : pEC23(*kan*)]	pEC23→BFA1431
GTD134	*yufL′* : : [pMUTIN2 Δ(*lacZ–ery*) : : pEC23(*kan*)] *amyE′* : : P*ywkA′* : : *lacZ–cat*	GTD133 DNA × PS1716
BFA1432	*yufM′* : : pMUTIN2	This study
GTD131	*yufM′* : : [pMUTIN2 Δ(*lacZ–ery*) : : pEC23(*kan*)]	pEC23→BFA1432
GTD132	*yufM′* : : [pMUTIN2 Δ(*lacZ–ery*) : : pEC23(*kan*)] *amyE′* : : P*ywkA′* : : *lacZ–cat*	GTD131 DNA × PS1716

*Japan Function Analysis Network for *B. subtilis*. Plasmid pMUTINT3 used to construct strain YWKAd is a derivative of the pMUTIN plasmid (Moriya *et al.*, 1998).

†Bacillus Subtilis Functional Analysis programme.

glyoxal gels, transferred to a Hybond-N membrane (Amersham) and hybridized with a labelled probe as described previously (Yoshida *et al.*, 1997). To prepare the probes for *ywkA*, *ywkB* and *gapB*, parts of their coding regions were amplified by PCR using chromosomal DNA of *B. subtilis* 168 as a template and primer pairs ywkA-N1 and ywkA-N2, ywkB-N1 and ywkB-N2, and gapB-N1 and gapB-N2, respectively (see supplementary table). The PCR products were labelled radioactively by using the *Bca* BEST labelling kit (Takara Shuzo) and [α-^{32}P]dCTP (ICN Biomedicals).

Preparation of fluorescently labelled cDNA, hybridization and microarray analysis. The fluorescently labelled cDNA probes used for hybridization to DNA microarrays were prepared by a two-step procedure, as described previously (Ogura *et al.*, 2001).

The microarrays that we used in this study contained 4055 protein genes, 45 not being spotted due to a problem with DNA amplification by PCR, as well as 39 calf thymus DNA spots as negative controls. The hybridization and microarray analysis were performed as described previously (Ogura *et al.*, 2001; Yoshida *et al.*, 2001). The mean Cy3 and Cy5 fluorescence intensities for each spot were calculated, the background being taken as the mean of the intensities of the 39 calf thymus DNA spots. After subtracting the background from all the intensities of the *B. subtilis* gene spots, and their normalization using the total Cy3 and Cy5 intensities, we calculated the expression ratios. To get reliable ratios, we ignored the spots of the intensities used as numerators for this calculation which were less than the background, and replaced the intensities used as denominators with the standard deviation of the mean intensity of the negative controls if they were lower than them.

Primer extension. To map the 5′-end of the *ywkA* transcript by primer extension, 50 μg of each RNA was annealed to primer rev (see supplementary table) that had been labelled at its 5′-end by a MEGALABEL kit (Takara Shuzo) and [α-^{32}P]ATP (Amersham). Primer extension reactions were carried out as described previously (Yoshida *et al.*, 1997).

Construction of the mini-Tn*10* transposition library and genetic screening. The mini-Tn*10* delivery vector pIC333 (Steinmetz & Richter, 1994) was used for transposon mutagenesis. Outside the transposon, this plasmid carries a thermosensitive Gram-positive origin of replication, an erythromycin-resistance gene, and a gene encoding a modified transposase with relaxed target specificity. The mini-transposon itself confers spectinomycin resistance and contains the pUC origin of replication in *E. coli*. pIC333 was introduced into PS1716 at 25 °C using erythromycin selection. Single transformant colonies were used to inoculate 12 independent cultures in 2 ml LB containing spectinomycin. At the beginning of the exponential growth phase, the temperature was shifted from 25 to 37 °C and cells were grown for 4 h more at 37 °C. Dilutions of the cultures were spread on plates of MM plus casein hydrolysate (0·05 %) or LB medium, plus X-Gal supplemented with glucose (0·5 %) or malate (0·5 %) to screen for clones with deregulated *ywkA′* : : *lacZ* expression.

***β*-Galactosidase assays.** β-Galactosidase activities were measured using the method of Miller (1972) on cell extracts prepared by lysozyme treatment and centrifugation. Protein concentration was determined by the method of Bradford (1976) using the Bio-Rad protein assay solution. One unit of β-galactosidase activity is defined as the amount of enzyme that produces 1 nmol *o*-nitrophenol min^{-1} at 20 °C.

Expression and purification of His$_6$-YufM. His$_6$-YufM was overexpressed and purified using the QIAexpress kit (Qiagen). *yufM* coding sequence (705 bp) was integrated into *E. coli* expression vector pQE-30 using primers BamHI-YufM and PstI-YufM (see supplementary table) in order to produce a fusion protein with a hexahistidine tag at the N-terminus. The resulting plasmid, pTDp3, was transformed into *E. coli* M15 pREP4 *groESL* to generate the strain used for production, GTDp3. All the subsequent procedures were performed as described in the QIAexpress manual. His$_6$-YufM was stored in a buffer containing 50 mM Tris pH 8, 150 mM NaCl, 1 mM DTT and 0·5 mM EDTA.

Expression and purification of His$_6$-YwkA. His$_6$-YwkA was overexpressed and purified as described for His$_6$-YufM. Thus, plasmid pTDp3 was built using primers BamHI-YwkA and SphI-YwkA (see supplementary table) and introduced into *E. coli* M15 pREP4 to generate strain GTDp4, used for production of His$_6$-YwkA. Storage buffer was the same as used for His$_6$-YufM.

Determination of kinetic parameters of YwkA. YwkA malic enzyme activity was tested at 37 °C by monitoring spectrophotometrically at 340 nm the NAD(P)H formation during the reductive decarboxylation of malate [malate + NAD(P) → pyruvate + NAD(P)H]. The reaction mixture was composed of 25 nmol YwkA in a Tris buffer containing divalent cations (50 mM Tris/HCl pH 8, 10 mM $MgCl_2$, 10 mM $MnCl_2$, 50 mM KCl, 10 mM β-mercaptoethanol). For determination of the K_m for malate, 5 mM of both cofactors were added to the reaction. For determination of K_m and V_{max} for each cofactor, the malate concentration was raised to 40 mM. For both cofactors, a molar absorption coefficient of $6{\cdot}22 \times 10^6$ cm^2 mol^{-1} was used for calculations. Demonstration of malic enzyme activity was assessed in a coupled reaction: at the end of the first reaction (end of NAD reduction), 40 units of strictly NAD-dependent lactate dehydrogenase from *Bacillus stearothermophilus* (Sigma) and 10 mM $CaCl_2$ were added to the reaction mixture. Then, oxidation of NADH, if produced by the first conversion of malate to pyruvate, was followed.

DNA electrophoretic mobility shift assays. The 277 bp DNA fragment of plasmid pPS61 containing the promoter region of *ywkA* was amplified by PCR with primers PS69 and PS70 (see supplementary table) and purified. Approximately 20 pmol of this fragment was end-labelled using T4 polynucleotide kinase (New England Biolabs) and 30 μCi (1·1 MBq) [γ-^{32}P]ATP (Isoblue, ICN). After precipitation, 10 000 c.p.m. of radiolabelled probe (≈0·05 pmol) was incubated for 5 min at room temperature in binding buffer (10 mM Tris/HCl pH 8·0, 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 5 %, w/v, glycerol) supplemented with 1 μg herring sperm DNA. Varying amounts of His$_6$-YufM were added for a total reaction volume of 20 μl. After incubating for 10 min at room temperature, samples were analysed by native 6 % acrylamide gel electrophoresis, 1 × Tris/glycine, run at 4 °C and constant voltage (14 V cm^{-1}). After migration, gels were dried and the radiolabelled bands revealed by the PhosphoImager system.

DNase I footprinting. Primers PS69 and PS70 (see supplementary table), 30 pmol each, were both end-labelled using T4 polynucleotide kinase (Biolabs) and 30 μCi (1·1 MBq) [γ-^{32}P]ATP (Isoblue, ICN) and purified by precipitation. They were used to generate one-strand labelled PCR probes. The labelled DNA probes (60 000 c.p.m.; ≈0·25 pmol) were incubated for 15 min at room temperature with varying amounts of His$_6$-YufM (the amount of protein was increased fivefold compared to gel mobility shift experiments due to partial loss of YufM binding activity after freezing). DNase I (Amersham Pharmacia Biotech) reactions were then performed as previously described (Doan & Aymerich, 2003). All samples were analysed on a 6 % acrylamide sequencing gel, which was dried and revealed by autoradiography.

RESULTS

Transcriptome analysis reveals *ywkA* induction in the presence of malate

A transcriptome analysis originally designed to identify the *B. subtilis* genes whose expression varies depending on whether the central carbon flux is glycolytic or gluconeogenic was performed. Cells of strain 168CA grew at almost the same rate on either glucose or malate as the sole carbon source; their doubling times in S6 medium containing glucose and malate were 1·37 and 1·25 h, respectively. Thus, these growth conditions are suitable to detect by DNA microarray analysis specific changes in global gene expression during glycolytic growth versus gluconeonenic growth; changes of expression of numerous genes such as ribosome protein genes that are affected by growth rate would not interfere.

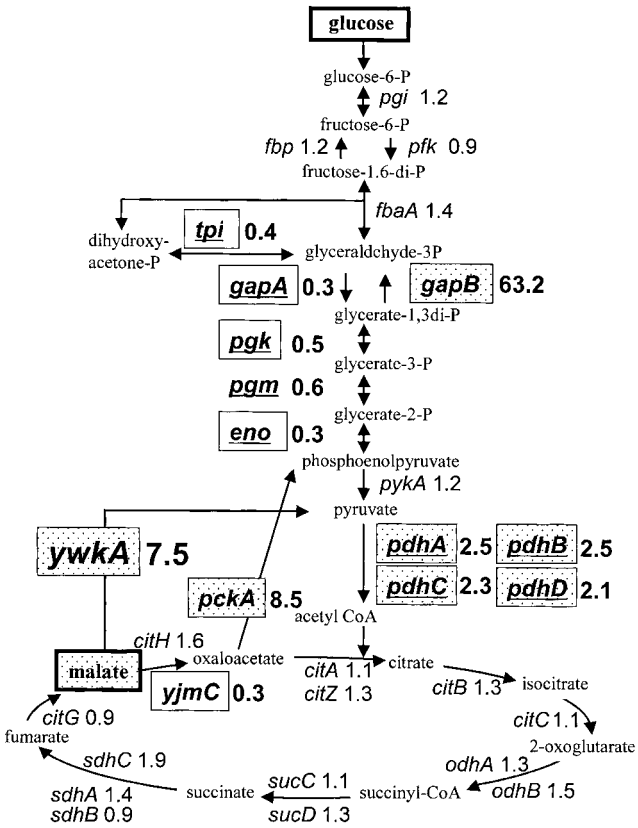


Fig. 1. Expression change of the *B. subtilis* genes involved in glycolysis and gluconeogenesis. Genes involved in the glycolytic pathway and the TCA cycle are followed by their expression ratios (malate/glucose) obtained by the DNA microarray analysis. The expression ratio for each gene was taken as the mean of two ratios obtained from the different hybridization profiles. Genes whose expression was changed at least twofold between glycolytic and gluconeogenic growths, are boxed (glycolytic genes) or boxed and stippled (gluconeogenic genes). The underlined *gapA*, *pgk*, *tpi*, *pgm* and *eno* genes, on the one hand, and *pdhABCD* genes, on the other, are organized in two operons. The complete microarray data are available on the KEGG website (http://www.genome.ad.jp/kegg/expression).

Fig. 1 shows the expression ratio of genes involved in the glycolytic pathway and the TCA cycle between the gluconeogenic and the glycolytic growth conditions, which were inferred from our DNA microarray analysis (the complete microarray data are available on the KEGG web site: http://www.genome.ad.jp/kegg/expression). Expression of genes boxed in the figure was changed at least twofold between glycolysis and gluconeogenesis. *ywkA* transcription appeared to be induced more than sevenfold in the presence of malate as compared to glucose. The transcription signal of *ywkB*, the gene located just downstream and most probably constituting a bicistronic operon with *ywkA*, was very weak. Therefore, the upregulation of *ywkB* transcription found under gluconeogenic growth conditions (ratio 1·96) cannot be considered as significant. The *gapB*, *pckA* and *pdhABCD* genes were also upregulated under gluconeogenic growth conditions. By contrast, the *gapA*, *pgk*, *tpi*, *eno* and *yjmC* genes were upregulated during glycolytic growth. The upregulation of *gapB* (Fillinger *et al.*, 2000) and *pckA* (Yoshida *et al.*, 2001) during gluconeogenesis was reported previously. It has also been reported that transcription of the *gapA*, *pgk*, *tpi* and *eno* genes, as well as that of the *pgm* gene, both belonging to the same *gapA-pgk-tpi-pgm-eno* operon, is stimulated during glycolysis (Fillinger *et al.*, 2000; Ludwig *et al.*, 2001, 2002; Doan & Aymerich, 2003).

ywkA expression is induced only when malate is present in the medium

To confirm that *ywkA* transcription is induced in the presence of malate and to define more precisely the physiological conditions in which it is turned on, a reporter strain of the activity of the *ywkA* promoter region, PS1716, was constructed (Table 1). This strain harbours a *ywkA′* : : *lacZ* fusion inserted into the chromosome at the ectopic *amyE* locus. The β-galactosidase activity synthezised in strain PS716 grown in mineral medium supplemented with different carbon sources was measured (Table 2). The results

Table 2. Expression of a *ywkA′* : : *lacZ* reporter fusion in minimal medium supplemented with various carbon sources

Carbon source	β-Galactosidase activity* [U (mg protein)$^{-1}$]
Glucose	7
Glycerol	9
Fumarate	11
Succinate + glutamate	17
Malate	774
Malate + glucose	764

*Activity was assayed at least three times from independent clones; representative results from one series are shown. Strains were cultivated in CQTHC medium supplemented with the carbon sources indicated.

indicated that *ywkA* promoter activity was very weak in the presence of succinate plus glutamate, or fumarate, and of glucose or glycerol, as sole gluconeogenic and glycolytic carbon source, respectively. Similarly, very weak basal β-galactosidase activity was detected in the presence of other gluconeogenic carbon sources, pyruvate or aspartate (data not shown). By contrast, the expression of the reporter construct was very strong in the presence of malate (and rather constant throughout the exponential growth phase). Moreover, addition of glucose to malate minimal medium did not affect *ywkA* expression. These results showed that *ywkA* transcription is specifically induced in the presence of malate in the culture medium and that it is not subject to carbon catabolite repression. Furthermore, because no expression was detected in the presence of fumarate, a C_4-dicarboxylic acid that enters the TCA cycle and is then converted in one step into intracellular malate, it could be assumed that only extracellular malate induces *ywkA* expression.

ywkA does not contribute to the utilization of malate for growth

As *ywkA* is specifically expressed when malate is present in the culture medium, we wanted to find out whether this gene is involved in the utilization of malate as a carbon source for growth. A *ywkA* mutant strain, GTD102, was constructed and was tested for growth in minimal medium containing either malate, or succinate plus glutamate (entering the TCA cycle and subsequently providing malate), or a glycolytic carbon source, glucose. This growth test showed that strain GTD102 was able to use either malate (Fig. 2), or succinate plus glutamate or glucose (data not shown) as efficiently as the wild-type strain.

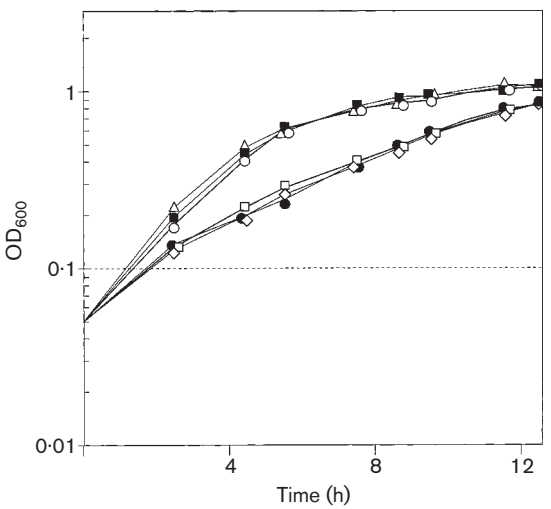


Fig. 2. Comparative growth of strains 168CA (○), GTD102 (*ywkA*, ■), BFA91 (*ytsJ*, □), GTD110 (*malS*, △), GTD121 (*ytsJ ywkA*, ●) and GTD123 (*ytsJ ywkA malS*, ◇) in malate minimal medium (MM malate).

A *ywkA* inactivation might be functionally complemented by one or several other putative malic enzyme encoding genes. Therefore, we also constructed different strains harbouring different combinations of mutations of the *ywkA*, *malS* and *ytsJ* genes: single mutant strains BFA91 (*ytsJ*) and GTD110 (*malS*), double mutant strains GTD111 (*ywkA*, *malS*), GTD121 (*ywkA*, *ytsJ*) and GTD122 (*ytsJ*, *malS*), and a triple mutant strain GTD123 (*ywkA*, *malS*, *ytsJ*). All these strains were able to grow in glucose minimal medium as efficiently as the wild-type strain (data not shown). But when growing in the presence of malate as sole carbon source (Fig. 2), all the strains mutated for *ytsJ*, BFA91, GTD121, GTD122 (data not shown) and GTD123, exhibited the same lower growth rate than the wild-type strain. By contrast, all the other mutants, GTD102, GTD110 and GTD111 (data not shown) grew as well as the wild-type strain. The *ywkA* mutation is supposed to have a polar effect on *ywkB*, and the absence of phenotype could be due to the absence of expression of both *ywkA* and *ywkB*. Because the *ywkA* disruption was constructed using a pMUTIN derivative (Moriya *et al.*, 1998; Vagner *et al.*, 1998), the possible polar effect on *ywkB* can be suppressed by addition of IPTG. However, the growth pattern of the strains harbouring a *ywkA* mutation was identical in the presence and in the absence of IPTG (data not shown). These experiments confirmed that *ywkA* and *malS* are not necessary for growth under either of the conditions tested. By contrast, *ytsJ* appeared to play a major role in the utilization of malate for growth.

YwkA exhibits a malic enzyme activity and preferentially utilizes NAD

Does *ywkA* encode a functional malic enzyme? To answer this question, a characterization of the enzymic properties of the YwkA protein was performed. First, a vector designed to overexpress a YwkA tagged with His_6 was constructed (pTDp3) and introduced into *E. coli* M15 pREP4. Overexpression was induced by cultivating the resulting strain in the presence of IPTG, and the His_6-YwkA protein was purified to apparent homogeneity. Then, the capacity of this protein to reduce NAD or NADP in the presence of malate was assayed. The results showed that YwkA has a strong activity of reduction of either NAD^+ or $NADP^+$ that depends on the presence of malate. To demonstrate that this reducing activity is associated with the conversion of malate into pyruvate, i.e. that YwkA exhibits a malic enzyme activity and not a malate dehydrogenase or a malolactic enzyme activity, the assay of the activity of an enzyme, lactate dehydrogenase, using specifically pyruvate as substrate, was coupled to the assay of YwkA. Indeed, addition of lactate dehydrogenase after the end of the first reaction (no further increase of A_{430}) led to rapid reoxidation of NADH (Fig. 3). This could result only from oxidative conversion of pyruvate into lactate and thus demonstrated that pyruvate is the product of the first reaction catalysed by YwkA. In a second series of experiments, we determined the kinetic parameters of YwkA activity. The K_m value for malate was

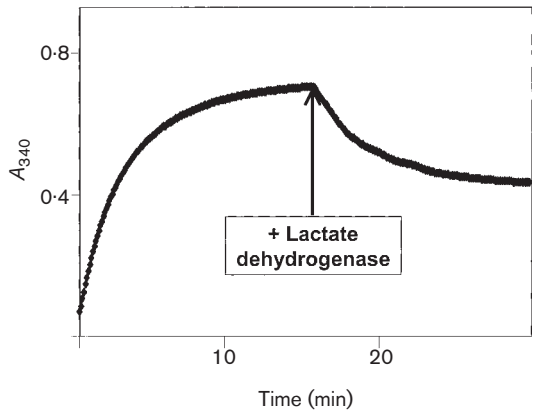


Fig. 3. Characterization of the enzymic activity of YwkA. Addition of lactate dehydrogenase after the end of the first reaction (no further increase of A_{430}) results in reoxidation of NADH, thus showing conversion of pyruvate into lactate.

determined to be 6·6 mM. K_m, k_{cat} and k_{cat}/K_m values were 1·1 mM, 112 s^{-1} and 102 s^{-1} mM^{-1}, respectively, for NAD, and 9·0 mM, 88 s^{-1}, and 10 s^{-1} mM^{-1}, respectively, for NADP. Thus, we concluded that YwkA is a malic enzyme with a dual specificity for NAD and NADP as a cofactor but with a strong preference for NAD [10-fold higher catalytic efficiency (k_{cat}/K_m) with NAD than with NADP].

Transcriptional characterization of *ywkA*

RNA samples prepared from cells of strains 168CA, 1A250 and 1A147 (*ccpA1*) grown in malate or glucose minimal medium were subjected to Northern analysis using specific probes for *ywkA*, *ywkB*, and *gapB*, a key gluconeogenic gene. As shown in Fig. 4, a specific transcript (2·8 kb) was detected with the *ywkA* probe only when cells were grown on malate (lanes 2, 4 and 6). This transcript is able to cover the *ywkA* gene as well as the downstream gene, *ywkB*. When the *ywkB* probe was used, this 2·8 kb transcript was only slightly visible due to the presence of the faint broad bands from rRNAs (3·0 and 1·5 kb), only when cells of the three strains were grown on malate (lanes 2, 4 and 6). These results strongly suggested that the *ywkAB* genes are cotranscribed as a single transcription unit. As a control, a transcript (1·7 kb) of *gapB* was observed, as expected, only when cells were grown on malate (Fig. 4, *gapB* lanes 2, 4 and 6) (Fillinger *et al.*, 2000). Furthermore, this analysis also confirmed that repression of *ywkA-ywkB* transcription as well as *gapB* transcription (Fillinger *et al.*, 2000; Yoshida *et al.*, 2001) is independent of CcpA (compare lanes 4 and 6).

Primer extension analysis mapped a 5′-end of the *ywkAB* transcript only with RNA prepared from cells of strain 168CA grown on malate, not from those grown on glucose (Fig. 5, lane 1), and allowed us to find a corresponding promoter sequence consisting of −10 (TACAAT) and −35 (TTGAAT) regions separated by a 17 bp spacer, which is probably recognized by σ^A RNA polymerase. No other clear 5′-end of the *ywkAB* transcript was detected in the region (nucleotides about +120 to −140) covering the translation initiation nucleotide of *ywkA* (+80) by our primer extension experiments, including ones using another more downstream primer (data not shown).

In conclusion, the *ywkAB* genes are cotranscribed to produce a 2·8 kb transcript from the transcription initiation site (Fig. 5); this transcription is induced during growth on malate and this regulation is independent of CcpA.

Genetic screening for regulator(s) involved in the control of *ywkA* transcription

The *ywkAB* genes are not physically linked to any gene encoding a putative transcriptional regulator. So, we decided to search for a regulator of *ywkAB* transcription by a genetic screening. A mini-Tn*10* transposant library was obtained

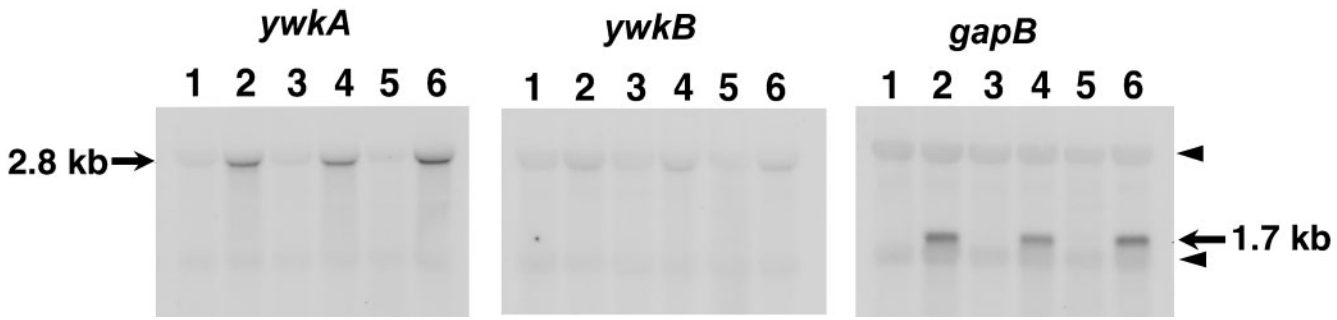


Fig. 4. Northern analysis of the *ywkA*, *ywkB* and *gapB* transcripts. The RNA samples from strains 168CA (lanes 1 and 2), 1A250 (lanes 3 and 4) and 1A147 (lanes 5 and 6) grown on glucose (lanes 1, 3 and 5) or malate (lanes 2, 4 and 6) were subjected to Northern analyses using *ywkA*, *ywkB* and *gapB* probes. The specific bands for *ywkA* and *ywkB* were observed at 2·8 kb, while those for *gapB* were observed at 1·7 kb. Two faint bands indicated by arrowheads at about 1·5 and 3 kb are due to nonspecific hybridization to rRNAs.

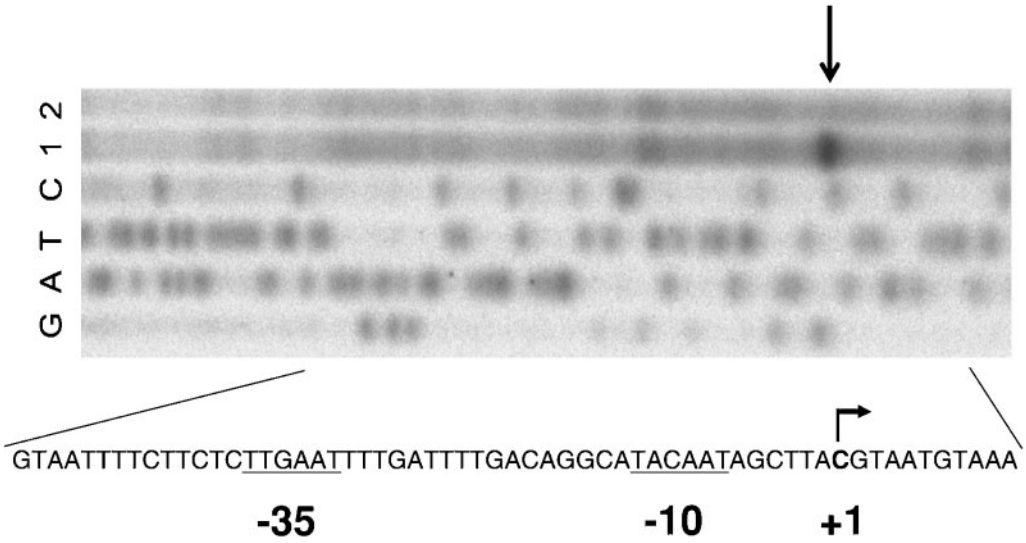


Fig. 5. Mapping of the 5′ end of the *ywkAB* transcript. Reverse transcriptase reaction was performed using total RNA from strain 168CA grown in malate (lane 1) or glucose (lane 2). Lanes G, A, T and C correspond to the dideoxy sequencing ladder obtained with the same primer as in the reverse transcriptase reaction. The transcriptional start site is indicated by an arrow on the right side of the panel. On the left side, the 5′ end position is marked by an arrow (+1) in the nucleotide sequence of the non-coding strand of the *ywkA* promoter region, where probable −10 and −35 regions are underlined.

from *B. subtilis* strain PS1716, which contains an ectopic *ywkA′* : : *lacZ* reporter fusion. This library was then screened on plates for deregulation of *ywkA* expression (see Methods). Several clones from independent pools of transposants were isolated in which *ywkA* was no longer induced in the presence of malate. The gene targeted by the transposon in 18 of these clones was identified. YufL was the gene most frequently found to be interrupted (eight occurrences). Back-crossing experiments from these eight transposants indicated that the loss of *ywkA* inducibility was associated with the insertion of the transposon. As predicted on the basis of its nucleotide sequence, *yufL* would encode the membrane-bound sensor-kinase of a TCS, associated with the transcriptional response-regulator encoded by the putatively co-transcribed gene *yufM*. However, no clone harbouring a transposon inserted into the *yufM* ORF was picked up in this genetic screening.

The TCS YufL/YufM is required for *ywkA* induction by malate

To confirm the role of the YufL/YufM TCS in the regulation of *ywkA* transcription, we introduced a *yufM* disruption into strain PS1716 and analysed the expression of the *ywkA′* : : *lacZ* reporter fusion in the resulting strain (GTD132). Strain GTD134, a PS1716-derived strain in which the *yufL* gene is disrupted by a plasmidic cassette that places the downstream gene *yufM* under the control of a P*spac* promoter, and the parental strain PS1716 were also tested in parallel. The *β*-galactosidase activity synthesized by these strains grown in the presence of glucose or malate as sole carbon source was measured (Table 3). The results showed that *ywkA* expression was no longer induced by malate in a *yufL* or *yufM* mutant strain. In the *yufL* mutant strain (GTD134), expression of *ywkA* was restored after the artificial overexpression of *yufM* by addition of IPTG,

Table 3. Effect of *yufL* or *yufM* mutations on *ywkA′* : : *lacZ* reporter fusion expression

Strain	*β*-Galactosidase activity* [U (mg protein)$^{-1}$]	
	Glucose	Malate
PS1716 [*amyE′* : : P*ywkA* : : *lacZ*]	8	655
GTD134 [*amyE′* : : P*ywkA* : : *lacZ*, *yufL′* : : pEC23(*kan*)]	3	14
GTD134 [*amyE′* : : P*ywkA* : : *lacZ*, *yufL′* : : pEC23(*kan*)]	250 (+IPTG)	368 (+IPTG)
GTD132 [*amyE′* : : P*ywkA* : : *lacZ*, *yufM′* : : pEC23(*kan*)]	5	7

*Activity was assayed at least three times from independent clones; representative results from one series are shown. Strains were cultivated in glucose or malate CQTHC media (+IPTG when indicated).

independently of the carbon source, glucose or malate, present in the medium. These findings thus demonstrated that the YufL/YufM TCS is required for *ywkA* induction in response to the presence of malate.

The growth rate and the biomass yield of the *yufM* and of the *yufL* mutant strains in the malate medium was significantly lower than that of the parental strain (doubling time 1·6 h versus 1·2 h; final OD_{600} 0·7 versus 1·5, respectively). This observation suggests that the YufL/YufM TCS positively controls the expression of other genes that, in contrast to *ywkA*, are involved in the utilization of malate as carbon source for growth.

YufM is able to bind specifically to the regulatory region of *ywkA*

A His-tagged YufM protein was overexpressed in *E. coli* and purified to apparent homogeneity (a figure showing the purification is available as supplementary data with the online version of this paper at http://mic.sgmjournals.org). A gel mobility shift assay was then performed to test the capacity of YufM to bind to the *ywkA* regulatory region. The assay, which was carried out in the presence of a large excess of unspecific DNA competitor, showed that YufM can bind directly and specifically to the regulatory region of *ywkA* in a concentration-dependent manner (Fig. 6). The shift of the probe was complete only when the protein/DNA concentration ratio exceeded 150; this suggests that, as demonstrated for other TCS response regulators (Perego & Hoch, 2002), YufM requires to be phosphorylated to be fully active in DNA binding.

YufM protects a 50 bp DNA sequence upstream of the *ywkA* promoter sequence

To define more precisely the YufM binding region, DNase I footprinting experiments were performed. The 277 bp fragment present upstream of the *ywkA* coding sequence was labelled on one strand or the other, incubated with an increasing amount of YufM, and then hydrolysed by DNase I. As revealed after electrophoresis (Fig. 7a), YufM protected a 50 bp sequence shifted three bases from one strand to the opposite one (Fig. 7b). This protected area is located just upstream of the −35 sequence of the *ywkA* promoter, in agreement with the transcriptional activator activity of YufM. Sequence examination of that protected region revealed two distinct perfect direct repeats (Fig. 7b): $ATTAAAAAATTN_{11}ATTAAAAAATT$ and $TAAGTAN_{11}TAAGTA$. Searches for these motifs (with one mismatch allowed in each sequence) in the *B. subtilis* complete genome sequence revealed no other occurrences. In particular, no sequence similar to these direct repeats can be detected in the region protected by YufM upstream of two other target genes of YufM (Tanaka *et al.*, 2003). This suggests that none of these DNA sequences is strictly required for efficient binding of YufM (see also Discussion).

DISCUSSION

In this study, we have investigated the physiological function and the transcriptional regulation of *ywkA*, one of the four paralogous genes of *B. subtilis* that could be predicted to encode malic enzymes. On the basis of the results of a global transcriptomic analysis using DNA microarrays, we have further studied the expression pattern of the *ywkA* gene through *lacZ* reporter fusion, Northern and primer extension analyses. We have been able to demonstrate that *ywkA* expression is specifically induced by malate. A genetic study then identified the uncharacterized YufL/YufM TCS as the regulatory system necessary for the activation of *ywkA* transcription in response to the presence of malate in the culture medium. *In vitro* experiments demonstrated that YufM is able to bind directly to a region located just upstream of the *ywkA* promoter. Thus we propose to rename this TCS MalK/MalR for malate sensor kinase/malate regulator component. On the other hand, if our enzymic study clearly demonstrated that YwkA is a malic enzyme, our phenotypic analysis showed that *ywkA* is not necessary for utilization of malate as sole carbon source for growth.

The enzymic characterization reported here demonstrated that YwkA exhibits a malic enzyme activity and excludes the possibility that it is a malolactic enzyme or a malate dehydrogenase. Furthermore, it showed that YwkA has a 10-fold higher catalytic efficiency with NAD as a cofactor than with NADP. We therefore propose to class YwkA in the group of NAD-dependent malic enzymes, EC 1.1.1.38, and to rename its structural gene *maeA* for malic enzyme A.

Interestingly, our transcriptome analysis indicated that either *malS* or *ytsJ*, two paralogues of *ywkA*, were significantly and similarly transcribed in the presence of glucose or malate, suggesting that their expression is not regulated by malate nor by the glycolytic or gluconeogenic carbon flux

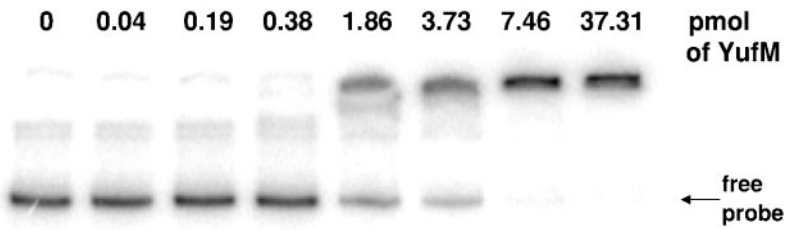


Fig. 6. Specific binding of His_6-YufM to the promoter region of the *ywkA* operon. The 277 bp DNA fragment inserted into plasmid pPS61 (approx. 0·05 pmol) was labelled and then incubated with no or increasing amounts of His_6-YufM in the presence of a large excess of unspecific DNA.

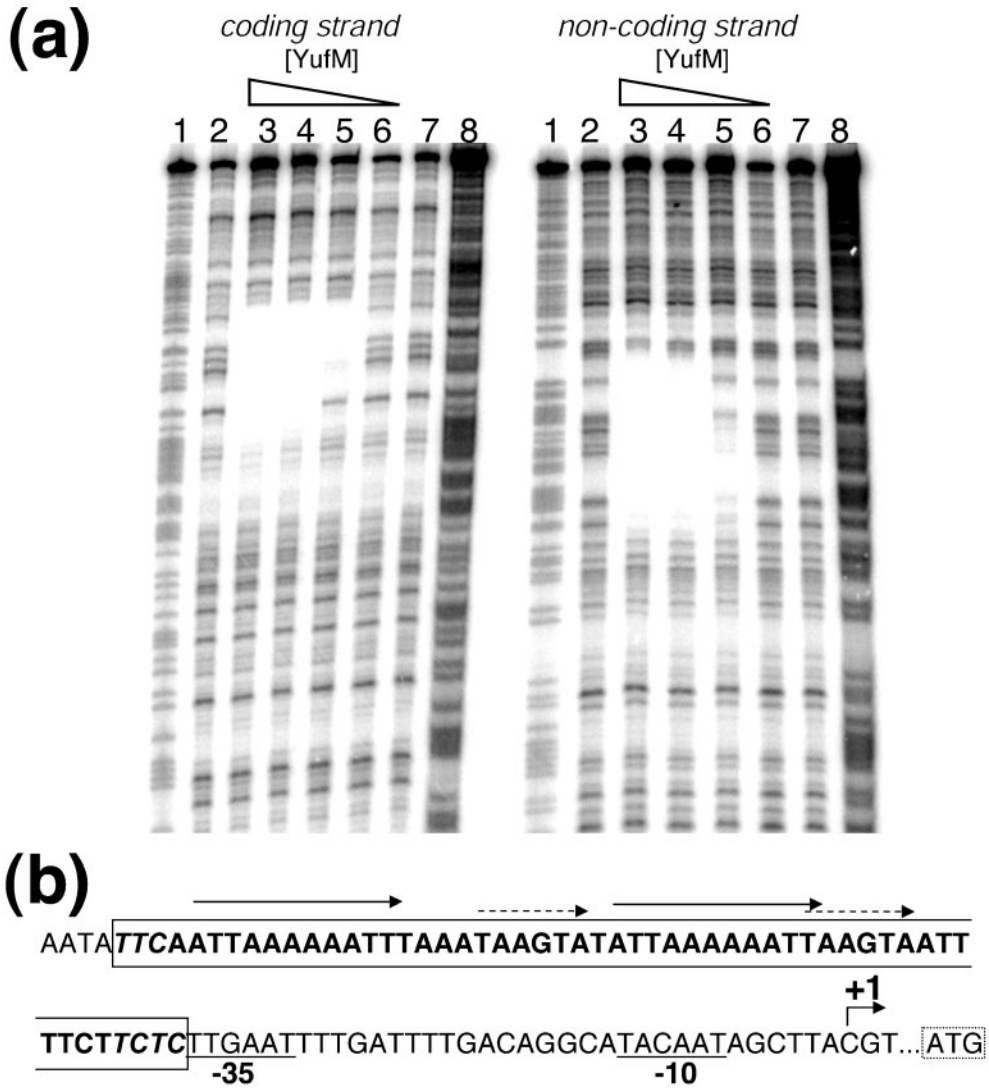


Fig. 7. (a) DNase I footprinting analysis of the interaction between YufM and the promoter region of *ywkA* examined on the coding and the non-coding strands. Approximately 0·25 pmol of labelled DNA fragment was incubated with no (lanes 2 and 7) or increasing amounts of His_6-YufM (lanes 3, 74·58 pmol; lanes 4, 37·29 pmol; lanes 5, 9·32 pmol; lanes 6, 1·86 pmol). Lanes 1 and 8, G+A sequencing ladder obtained from the same DNA fragment. (b) Nucleotide sequence of the *ywkA* promoter region. The protected region is boxed: the protected nucleotides on both strands are in normal bold letters; italic letters indicate nucleotides that are protected only on the coding (5′ side) or the non-coding (3′ side) strand. Two direct repeats are indicated by continued or dashed arrows.

orientation. By contrast, *ywkA* transcription was induced only in the presence of malate in the culture medium. Further addition of glucose or another glycolytic carbon source to the medium did not affect *ywkA* expression. Thus, *ywkA* is not subject to carbon catabolite repression. YwkA was not expressed during growth on fumarate or succinate plus glutamate, although these compounds enter the TCA cycle upstream of malate and thus very probably lead to the elevation of the intracellular malate concentration. It is thus very likely that extracellular malate is the signal for *ywkA* induction. Of course this does not exclude the possibility that the permeation of malate would be necessary for efficient transduction of the signal and thus full induction of *ywkA* transcription. The effects of the inactivation of the different putative malate transporters on the inducibility of *ywkA* need to be studied to test this possibility.

To try to get some clues concerning the biological role of *ywkA*, the regulatory mechanism of its expression has been investigated. Our analysis demonstrated that the YufL/YufM TCS is the positive regulatory system that mediates the induction of *ywkA* expression in response to the presence of malate in the culture medium, YufM being a direct transcriptional activator of *ywkA* transcription. This conclusion is in agreement with the results of a global transcriptome analysis performed by Kobayashi *et al.* (2001)

indicating that *ywkA* expression is ninefold upregulated in a *yufL* mutant strain overexpressing YufM, compared to a wild-type strain. The YufL/YufM TCS belongs to the CitA/CitB family of TCSs. The members of this family that have already been characterized are CitA/CitB from *E. coli* (Kaspar & Bott, 2002) and *Klebsiella pneumoniae* (Bott *et al.*, 1995; Kaspar *et al.*, 1999), CitS/CitT from *B. subtilis* (Yamamoto *et al.*, 2000) and *Streptomyces coelicolor*, YdbF/YdbG from *B. subtilis* (Asai *et al.*, 2000) and DcuS/DcuR from *E. coli* (Zientz *et al.*, 1998; Golby *et al.*, 1999). All these TCSs are involved in the control of genes required for TCA cycle intermediates import and utilization. Sequence analyses strongly suggest that YufL is a transmembrane protein containing an extracellular sensor domain and a cytoplasmic kinase domain, as shown for the highly similar *E. coli* DcuS protein (Golby *et al.*, 1999). Thus, YufL would be able to modulate YufM DNA-binding activity by phosphorylation/dephosphorylation depending on the presence of extracellular malate. The transcriptome analysis performed by Kobayashi *et al.* (2001) has revealed 97 other differentiallly expressed genes in the *yufL* mutant strain overexpressing *yufM* compared to the wild-type strain. Because this list included several competence genes, and especially *comK* and most of the members of the ComK regulon (Berka *et al.*, 2002; Ogura *et al.*, 2002), the authors suggested that the YufL/YufM TCS would be involved in competence development. However, we found no drastic transformability defect, if any, of *yufL*, *yufM* or *ywkA* mutant strains (T. Doan & S. Aymerich, unpublished data). The study of the genome-wide expression changes dependent on ComK, recently reported by Berka *et al.* (2002), led these authors to conclude that the ComK regulon defines a growth-arrested state, distinct from sporulation, of which competence for genetic transformation is but one notable feature. They suggest that this is a unique adaptation to stress and that it be termed the 'K-state'. Thus the YufL/YufM TCS could participate in the signalling of the 'K-state'.

Interestingly, simultaneously with our work, the expression of *maeN* (*yufR*), a gene located 5·3 kb downstream of *yufL-yufM* on the chromosome, has been found to be activated by the YufL/YufM TCS in the presence of malate (Tanaka *et al.*, 2003). Wei *et al.* (2000) had previously reported that MaeN exhibits a malate–Na^+ symporter activity. Ogasawara and colleagues indeed found that MaeN, and thus YufL/YufM, are essential for utilization of malate as sole carbon source for growth (Tanaka *et al.*, 2003). This is in agreement with the growth defect on malate associated with the *yufL* and the *yufM* mutations that we observed during the test of their effects on *ywkA* expression; however, the defect was only partial in our experiments because the medium used contained not only malate but also glutamine and casein hydrolysate as carbon sources (Table 3). The coregulation of *ywkA* and *maeN* is of course physiologically relevant since MaeN could provide the YwkA malic enzyme with cytoplasmic malate. Furthermore, Krulwich and coworkers (Ito *et al.*, 2000) have shown that, in some particular conditions, *maeN* expression was coupled to the expression of the downstream *mrp* gene cluster involved in alkali resistance and pH homeostasis. These authors suggested a possible requirement of coordinated Na^+–malate symport and Na^+ re-extrusion in exchange for H^+. The protonmotive force could be generated by secondary NADH dehydrogenases putatively encoded by some *mrp* genes. Therefore, YwkA could be involved in the regeneration of NADH during this process.

Ogasawara and colleagues have also found that the expression of *yflS*, putatively encoding a 2-oxoglutarate/malate translocator, was under the control of the YufL/YufM TCS (Tanaka *et al.*, 2003). Indeed, *maeN* and *yflS* have been detected as being up-regulated (ratios Mal/Glc 4·2 and 2·7, respectively) during growth on malate in our transcriptomic analysis. By contrast, they were not clearly identified (ratios 1·7 and 1·5) in the transcriptomic analysis of Kobayashi *et al.* (2001), which compared a *yufL* mutant strain overexpressing YufM with the wild-type strain. We have compared the sequence of the region protected by YufM upstream of the *ywkA*, the *maeN* and the *yflS* promoters and found no significant similarities. The most conspicuous feature of these three regions is their very high A + T content (88, 83 and 82 mol%, upstream of *ywkA*, *maeN* and *yflS*, respectively). On the other hand, no other occurrence of one or the other direct repeat included in the region protected by YufM upstream of *ywkA* (Fig. 6) can be detected in the complete *B. subtilis* genome. Thus one should conclude that the YufM DNA-binding activator does not recognize a unique primary sequence, a property similar to that of several DNA-binding components of other well-characterized TCSs (Perego & Hoch, 2002).

YwkA does have malic enzyme activity but it does not contribute to the very efficient utilization of malate as sole carbon source for growth by *B. subtilis*. This is in agreement with the fact that *ywkA* is not subject to carbon catabolite repression. Because neither a *malS* nor a double *ywkA malS* mutant strain showed any growth defect, it can also be concluded that, like *ywkA*, *malS* is not necessary for growth on malate. By contrast, the *ytsJ* gene is required for efficient growth on malate, as well as on other TCA cycle intermediates. *ytsJ* is expressed constitutively during the exponential phase of growth and encodes an NADP-dependent malic enzyme (T. Doan, G. Lerondel & S. Aymerich, unpublished data). YtsJ could thus be required for its capacity to catalyse an anaplerotic reaction that allows the TCA cycle to function as a real cycle when the only carbon source is one intermediate of this cycle, or for its capacity to regenerate NADPH necessary for numerous anabolic reactions, or for both of these activities. We would propose that *ywkA* is involved in a more specific biological process necessary only or principally when malate is present in the extracellular environment. The study of the *ywkB* gene, of which no clear orthologue can be detected in the genome of other bacilli, and the identification of the activity of more genes belonging to the *yufL/YufM* TCS regulon, should help to develop this hypothesis.

ACKNOWLEDGEMENTS

We thank Patrick Stragier for the generous gift of plasmid pEC23 (unpublished). We are very grateful to Naotake Ogasawara and Uwe Sauer for communicating results before publication and for very helpful discussions. We thank Claude Gaillardin and Josef Deutscher for valuable discussions and for their continuous encouragement, and Dominique Le Coq and Ivan Mijakovic for their comments on the manuscript.

This work was supported by EU Quality of Life and Management of Living Resources, programme grant QL GZ-CT-1999-O1455, and by a Grand-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports and Culture of Japan. Stéphane Aymerich, Guillaume Lerondel and Thierry Doan are on the staff at INRA (France).

REFERENCES

Anagnostopoulos, C. & Spizizen, J. (1961). Requirements for transformation in *Bacillus subtilis*. *J Bacteriol* **81**, 741–746.

Ansanay, V., Dequin, S., Blondin, B. & Barre, P. (1993). Cloning, sequence and expression of the gene encoding the malolactic enzyme from *Lactococcus lactis*. *FEBS Lett* **332**, 74–80.

Arthurs, C. E. & Lloyd, D. (1999). Kinetics, stereospecificity, and expression of the malolactic enzyme. *Appl Environ Microbiol* **65**, 3360–3363.

Asai, K., Baik, S. H., Kasahara, Y., Moriya, S. & Ogasawara, N. (2000). Regulation of the transport system for C_4-dicarboxylic acids in *Bacillus subtilis*. *Microbiology* **146**, 263–271.

Aymerich, S., Gonzy-Treboul, G. & Steinmetz, M. (1986). 5′-noncoding region *sacR* is the target of all identified regulation affecting the levansucrase gene in *Bacillus subtilis*. *J Bacteriol* **166**, 993–998.

Bartolucci, S., Rella, R., Guagliardi, A., Raia, C. A., Gambacorta, A., De Rosa, M. & Rossi, M. (1987). Malic enzyme from archaebacterium *Sulfolobus solfataricus*. Purification, structure, and kinetic properties. *J Biol Chem* **262**, 7725–7731.

Berka, R. M., Hahn, J., Albano, M., Draskovic, I., Persuh, M., Cui, X., Sloma, A., Widner, W. & Dubnau, D. (2002). Microarray analysis of the *Bacillus subtilis* K-state: genome-wide expression changes dependent on ComK. *Mol Microbiol* **43**, 1331–1345.

Bott, M., Meyer, M. & Dimroth, P. (1995). Regulation of anaerobic citrate metabolism in *Klebsiella pneumoniae*. *Mol Microbiol* **18**, 533–546.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254.

Denayrolles, M., Aigle, M. & Lonvaud-Funel, A. (1994). Cloning and sequence analysis of the gene encoding *Lactococcus lactis* malolactic enzyme: relationships with malic enzymes. *FEMS Microbiol Lett* **116**, 79–86.

Dervyn, E., Suski, C., Daniel, R., Bruand, C., Chapuis, J., Errington, J., Janniere, L. & Ehrlich, S. D. (2001). Two essential DNA polymerases at the bacterial replication fork. *Science* **294**, 1716–1719.

Diesterhaft, M. D. & Freese, E. (1973). Role of pyruvate carboxylase, phosphoenolpyruvate carboxykinase, and malic enzyme during growth and sporulation of *Bacillus subtilis*. *J Biol Chem* **248**, 6062–6070.

Doan, T. & Aymerich, S. (2003). Regulation of the central 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.

Driscoll, B. T. & Finan, T. M. (1993). NAD^+-dependent malic enzyme of *Rhizobium meliloti* is required for symbiotic nitrogen fixation. *Mol Microbiol* **7**, 865–873.

Driscoll, B. T. & Finan, T. M. (1996). $NADP^+$-dependent malic enzyme of *Rhizobium meliloti*. *J Bacteriol* **178**, 2224–2231.

Driscoll, B. T. & Finan, T. M. (1997). Properties of NAD^+- and $NADP^+$-dependent malic enzymes of *Rhizobium* (*Sinorhizobium*) *meliloti* and differential expression of their genes in nitrogen-fixing bacteroids. *Microbiology* **143**, 489–498.

Dunn, M. F. (1998). Tricarboxylic acid cycle and anaplerotic enzymes in rhizobia. *FEMS Microbiol Rev* **22**, 105–123.

Ferrari, F. A., Ferrari, E. & Hoch, J. A. (1982). Chromosomal location of a *Bacillus subtilis* DNA fragment uniquely transcribed by sigma-28-containing RNA polymerase. *J Bacteriol* **152**, 780–785.

Fillinger, S., Boschi-Muller, S., Azza, S., Dervyn, E., Branlant, G. & Aymerich, S. (2000). Two glyceraldehyde-3-phosphate dehydrogenases with opposite physiological roles in a nonphotosynthetic bacterium. *J Biol Chem* **275**, 14031–14037.

Fujita, Y. & Freese, E. (1981). Isolation and properties of a *Bacillus subtilis* mutant unable to produce fructose-bisphosphatase. *J Bacteriol* **145**, 760–767.

Golby, P., Davies, S., Kelly, D. J., Guest, J. R. & Andrews, S. C. (1999). Identification and characterization of a two-component sensor-kinase and response-regulator system (DcuS-DcuR) controlling gene expression in response to C4-dicarboxylates in *Escherichia coli*. *J Bacteriol* **181**, 1238–1248.

Groisillier, A. & Lonvaud-Funel, A. (1999). Comparison of partial malolactic enzyme gene sequences for phylogenetic analysis of some lactic acid bacteria species and relationships with the malic enzyme. *Int J Syst Bacteriol* **49**, 1417–1428.

Guérout-Fleury, A. M., Frandsen, N. & Stragier, P. (1996). Plasmids for ectopic integration in *Bacillus subtilis*. *Gene* **180**, 57–61.

Hansen, E. J. & Juni, E. (1974). Two routes for synthesis of phosphoenolpyruvate from C4-dicarboxylic acids in *Escherichia coli*. *Biochem Biophys Res Commun* **59**, 1204–1210.

Hansen, E. J. & Juni, E. (1975). Isolation of mutants of *Escherichia coli* lacking NAD- and NADP-linked malic enzyme. *Biochem Biophys Res Commun* **65**, 559–566.

Ito, M., Guffanti, A. A., Wang, W. & Krulwich, T. A. (2000). Effects of nonpolar mutations in each of the seven *Bacillus subtilis mrp* genes suggest complex interactions among the gene products in support of Na^+ and alkali but not cholate resistance. *J Bacteriol* **182**, 5663–5670.

Kaspar, S. & Bott, M. (2002). The sensor kinase CitA (DpiB) of *Escherichia coli* functions as a high-affinity citrate receptor. *Arch Microbiol* **177**, 313–321.

Kaspar, S., Perozzo, R., Reinelt, S., Meyer, M., Pfister, K., Scapozza, L. & Bott, M. (1999). The periplasmic domain of the histidine autokinase CitA functions as a highly specific citrate receptor. *Mol Microbiol* **33**, 858–872.

Kawai, S., Suzuki, H., Yamamoto, K., Inui, M., Yukawa, H. & Kumagai, H. (1996). Purification and characterization of a malic enzyme from the ruminal bacterium *Streptococcus bovis* ATCC 15352 and cloning and sequencing of its gene. *Appl Environ Microbiol* **62**, 2692–2700.

Knichel, W. & Radler, F. (1982). D-Malic enzyme of *Pseudomonas fluorescens*. *Eur J Biochem* **123**, 547–552.

Kobayashi, K., Doi, S., Negoro, S., Urabe, I. & Okada, H. (1989). Structure and properties of malic enzyme from *Bacillus stearothermophilus*. *J Biol Chem* **264**, 3200–3205.

Kobayashi, K., Ogura, M., Yamaguchi, H., Yoshida, K., Ogasawara, N., Tanaka, T. & Fujita, Y. (2001). Comprehensive DNA microarray analysis of *Bacillus subtilis* two-component regulatory systems. *J Bacteriol* **183**, 7365–7370.

Lamed, R. & Zeikus, J. G. (1981). Thermostable, ammonium-activated malic enzyme of *Clostridium thermocellum*. *Biochim Biophys Acta* **660**, 251–255.

Ludwig, H., Homuth, G., Schmalisch, M., Dyka, F. M., Hecker, M. & Stülke, 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**, 409–422.

Ludwig, H., Rebhan, N., Blencke, H. M., Merzbacher, M. & Stülke, J. (2002). Control of the glycolytic *gapA* operon by the catabolite control protein A in *Bacillus subtilis*: a novel mechanism of CcpA-mediated regulation. *Mol Microbiol* **45**, 543–553.

Miller, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Mitsch, M. J., Voegele, R. T., Cowie, A., Osteras, M. & Finan, T. M. (1998). Chimeric structure of the NAD(P)$^+$- and NADP$^+$-dependent malic enzymes of *Rhizobium* (*Sinorhizobium*) *meliloti*. *J Biol Chem* **273**, 9330–9336.

Moriya, S., Tsujikawa, E., Hassan, A. K., Asai, K., Kodama, T. & Ogasawara, N. (1998). A *Bacillus subtilis* gene-encoding protein homologous to eukaryotic SMC motor protein is necessary for chromosome partition. *Mol Microbiol* **29**, 179–187.

Murai, T., Tokushige, M., Nagai, J. & Katsuki, H. (1971). Physiological functions of NAD- and NADP-linked malic enzymes in *Escherichia coli*. *Biochem Biophys Res Commun* **43**, 875–881.

Ogura, M., Yamaguchi, H., Yoshida, K., Fujita, Y. & Tanaka, T. (2001). DNA microarray analysis of *Bacillus subtilis* DegU, ComA and PhoP regulons: an approach to comprehensive analysis of *B. subtilis* two-component regulatory systems. *Nucleic Acids Res* **29**, 3804–3813.

Ogura, M., Yamaguchi, H., Kobayashi, K., Ogasawara, N., Fujita, Y. & Tanaka, T. (2002). Whole-genome analysis of genes regulated by the *Bacillus subtilis* competence transcription factor ComK. *J Bacteriol* **184**, 2344–2351.

Perego, M. & Hoch, J. A. (2002). Two-component systems, phosphorelays, and regulation of their activities by phosphatases. In *Bacillus subtilis and its Closest Relatives: from Genes to Cells*, pp. 423–481. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Steinmetz, M. & Richter, R. (1994). Easy cloning of mini-Tn10 insertions from the *Bacillus subtilis* chromosome. *J Bacteriol* **176**, 1761–1763.

Stols, L. & Donnelly, M. I. (1997). Production of succinic acid through overexpression of NAD$^+$-dependent malic enzyme in an *Escherichia coli* mutant. *Appl Environ Microbiol* **63**, 2695–2701.

Tanaka, K., Kobayashi, K. & Ogasawara, N. (2003). The *Bacillus subtilis* YufLM two-component system regulates the expression of the malate transporters MaeN (YufR) and YflS, and is essential for utilization of malate in minimal medium. *Microbiology* **149**, 2317–2329.

Vagner, V., Dervyn, E. & Ehrlich, S. D. (1998). A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology* **144**, 3097–3104.

van der Rest, M. E., Frank, C. & Molenaar, D. (2000). Functions of the membrane-associated and cytoplasmic malate dehydrogenases in the citric acid cycle of *Escherichia coli*. *J Bacteriol* **182**, 6892–6899.

Voegele, R. T., Mitsch, M. J. & Finan, T. M. (1999). Characterization of two members of a novel malic enzyme class. *Biochim Biophys Acta* **1432**, 275–285.

Wei, Y., Guffanti, A. A., Ito, M. & Krulwich, T. A. (2000). *Bacillus subtilis* YqkI is a novel malic/Na$^+$-lactate antiporter that enhances growth on malate at low protonmotive force. *J Biol Chem* **275**, 30287–30292.

Yamamoto, H., Murata, M. & Sekiguchi, J. (2000). The CitST two-component system regulates the expression of the Mg-citrate transporter in *Bacillus subtilis*. *Mol Microbiol* **37**, 898–912.

Yoshida, K., Aoyama, D., Ishio, I., Shibayama, T. & Fujita, Y. (1997). Organization and transcription of the myo-inositol operon, *iol*, of *Bacillus subtilis*. *J Bacteriol* **179**, 4591–4598.

Yoshida, K., Kobayashi, K., Miwa, Y. & 9 other authors (2001). Combined transcriptome and proteome analysis as a powerful approach to study genes under glucose repression in *Bacillus subtilis*. *Nucleic Acids Res* **29**, 683–692.

Zientz, E., Bongaerts, J. & Unden, G. (1998). Fumarate regulation of gene expression in *Escherichia coli* by the DcuSR (*dcuSR* genes) two-component regulatory system. *J Bacteriol* **180**, 5421–5425.

5.2.2. Supplementary Material

Table S1. Oligonucleotides used in this study.

Table S1. Oligonucleotides used in this study.

	sequence (5' → 3')[a]	Corresponding plasmid or probe
Fusion and Footprinting		pPS61
PS69	TTAGAATTCATCACGAAGTCCCAGGTAAAT	pPS61
PS70	ATTGGATCCCATCAATATGACCGA	
Construction of strain GTD110		
EagI-ytnP	CCCGGCCGGGGACGAAATGCGCCATC	pTD101
MalS-XbaI-SPEC	CTTTATTATACAGATCGATCCTCTAGATTATATCGCTTGTTTCACCCTGCCTCT	pTD101
BamHI-ytzB	CGGGATCCTAAAAAAGAGCCGCTGCA	pTD101
EcoRI-ytzB	GGAATTCGCGCCGGGAAGCTGCG	pTD101
Northern Blot		
ywkA-N1	TACTGGCGCGATTACACTTG	*ywkA*
ywkA-N2	GCACTTTTCCGTCAGTCCAT	*ywkA*
ywkB-N1	TTTGTTATCGTGCTGGGTTG	*ywkB*
ywkB-N2	CTCACAAAGACCATGCAGGA	*ywkB*
gapB-N1	TGTTAAACAGCCGTGATCCA	*gapB*
gapB-N2	AATGAGACGTTCGGAACAGG	*gapB*
+1 Mapping		
rev	CCGATCATGATCTTTATGCGACTT	
protein overexpression		
BamHI-yufM	CGCGGATCCATGATTAATGTACTAATAGTTG	pTDp3
PstI_yufM	AACTGCAGTTATAGATATTGCTTTATTCCG	pTDp3
BamHI-ywkA	CGCGGATCCAAGAAAGAAGGAGATTTGCTTG	pTDp4
SphI-ywkA	ACATGCATGCTTATTTCGCGATGACTTTTTTG	pTDp4

[a] Restriction sites are underlined

5.3.Résultats complémentaires

Organisation génétique des gènes paralogues chez *Bacillus subtilis*.

La Figure 48 montre l'organisation génétique d'*ywkA*, de *malS* et d'*ytsJ* de *Bacillus subtilis. malS* pourrait être cotranscrit avec le gène en amont, *ytnP*, qui code une protéine de fonction inconnue. *ytsJ* est apparemment un gène monocistronique. Mais on peut remarquer que le gène situé en amont est *dnaE*, qui code la deuxième ADN polymérase essentielle de *Bacillus subtilis* (Dervyn *et al.*, 2001).

Comparaison des séquences des produits des quatre gènes paralogues de *Bacillus subtilis* et des enzymes maliques SfcA d'*Escherichia coli* et Maox de *Bacillus stearothermophilus*.

L'alignement des séquences protéiques d'YwkA, de MalS, d'YtsJ et de MleA de *Bacillus subtilis*, de SfcA d'*Escherichia coli* et de Maox de *Bacillus stearothermophilus* montre qu'YwkA présente des similarités significatives avec d'autres enzymes maliques bactériennes. Il permet de mettre en évidence plusieurs motifs très bien conservés (Figure 49). Certains de ces motifs ont été étudiés pour leur implication dans l'activité d'enzymes maliques eucaryotes. La boîte I correspond au site d'interaction avec le malate et contient une Tyrosine (Y) très bien conservée qui est impliquée dans la liaison avec le malate (Hsu, 1982; Viljoen *et al.*, 1994). Cette boîte renferme aussi un site potentiel d'interaction avec le NADP comprenant une cystéine très bien conservée (Hsu, 1982; Groisillier & Lonvaud-Funel, 1999). Dans le cas d'YwkA et de MalS, cette cystéine (C) est remplacée par une isoleucine (I), suggérant que ces deux enzymes n'utilisent pas le NADP comme cofacteur. La boîte II est très bien conservée et est impliquée dans la reconnaissance du NAD(P). La boîte III est la mieux conservée parmi les enzymes maliques connues et est considérée comme une « signature » de l'enzyme malique. Les boîtes IV, V et VI sont conservées parmi les enzymes maliques mais leur rôle spécifique n'est pas défini. Globalement, à l'intérieur de ces motifs conservés, YwkA partage plus de similitudes avec MalS et SfcA qu'avec YtsJ, MleA et Maox. Or, SfcA est une enzyme NAD-dépendante (Stols et Donnelly, 1997). L'alignement suggère donc qu'YwkA et MalS sont elles aussi des enzymes NAD-dépendantes. Cependant, il convient d'être prudent quant à cette

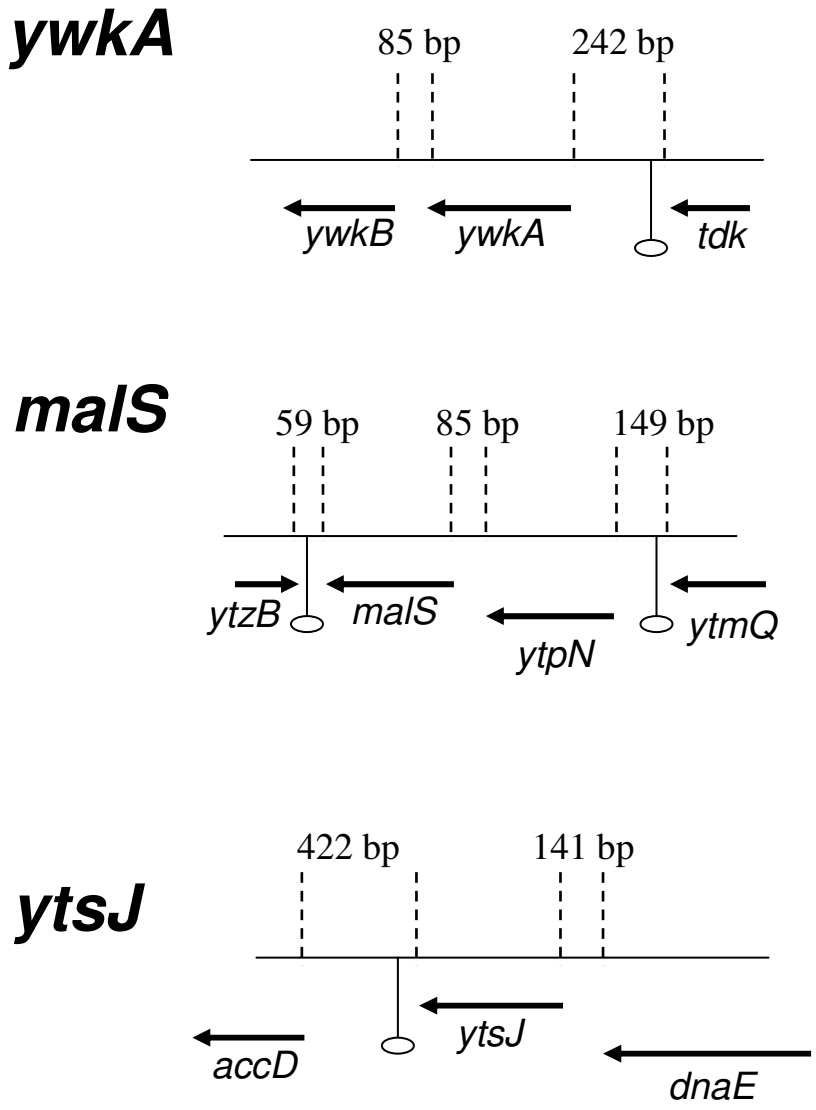
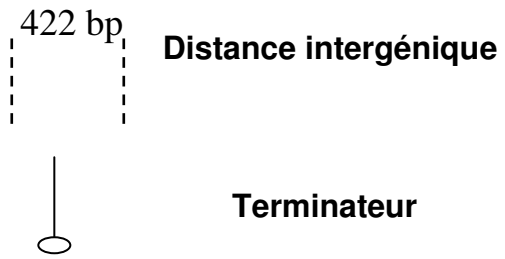


Figure 48. Organisation génétique des gènes *ywkA*, *malS* et *ytsJ* chez *Bacillus subtilis*. Les terminateurs de transcription prédits par l'analyse de séquence sont figurés.



```
            1                                                                                                   100
YWKABACSU   MGYYLTWLTISRKKEICLNNIKKTKEGHLETTLRGKEVLSIPTLNKGVAFSLEERQELGLEGLLPPTVLSLDQQAQRAYEQFQAQPDRLRQNVYLSDLAN
MALSBACSU                   MKQFRVTNEGDIQTTLRGLEVLSVPFLNKGVAFTEEERKELGLKGFLPPKVLTIDDQAKRAYEQYSAQPDDLSKNVYLTALHD
SFCAECOLI           MDIQKRVSDMEPKTKKQRSLYIPYAGPVLLEFPLLNKGSAFSMEERRNFNLLGLLPEVVETIEEQAERAWIQYQGFKTEIDKHIYLRNIQD
YTSJBACSU       MSLREEA-------------------------------------------------------------------------------LHLHKVNQG
MAOXBACST       MALPGGAAMNITIRLQFEKDIVSFSDIAAAIGKAGGDIVGIDVISSSKVHTVRDITVSALDTKQCDLIIEALKKIRGVKIVN-VSDRTFLMHIGG
MLEABACSU          MIAKHMIRTLMIETPSVPGNLGRVATAIGLLGGDIGEVETVKVGPNYTMRNITVQVENEEQLQEVIAAVQALGEGIRLHTVSDEVLSAHEGG
                                  Boîte I                                          Boîte II
            101                                                                                                 200
YWKABACSU   RNEVLFYKLLKNHLREMLPVVYTPTVGEAIQEYSHEYRRPQGIYLSIDNIDGIEKAFENLHATAGDIDLIVATDSESILGIGDWG-VGGINIAIGKLAVY
MALSBACSU   RNETLFYRLLNDHLGEMLPIVYTPTVGTAIQRYSHEYRKPRGLYLSIDDPDGMKEAFKQYKDQSDTIDLIVATDAEGILGIGDWG-VGGIAISIGKLAVY
SFCAECOLI   TNETLFYRLVNNHLDEMMPVIYTPTVGAACERFSEIYRRSRGVFISYQNRHNMDDILQN--VPNHNIKVIVVTDGERILGLGDQG-IGGMGIPIGKLSLY
YTSJBACSU   KLESKSKVEVRNAKD--LSLAYSPGVAEPCKDIHEDINKVYDY-------------------TMKGNMVAVVTDGTAVLGLGNIGPEAALPVMEGKAVLF
MAOXBACST   KIETNSKIPVKTRDD--LSRVYTPGVARVCTAIAEDPRKAYSL-------------------TIKRNTVAVVSDGTAVLGLGDIGPYAAMPVMEGKAMLF
MLEABACSU   KIQMKSKMPIRSLAE--LGRVYTPGVADVCRLIEKEPEKASIY-------------------TTISNSVAIVTDGTAILGLGNIGSVAGMPVMEGKAALF
                                                                                                  Boîte III
            201                                                                                                 300
YWKABACSU   TAAAGIDPSRVIPVVLDVGTNNEKLLNDPLYIGNKHERVQGERYEAFIDAYVKAALKFFPKALLHWEDLGNKNARNIMKKYNHEILTFNDDIQGTGAITL
MALSBACSU   TAAAGIDPSRVLAVVLDAGTNQESLLNDPLYVGNQHSRVRGERYDQFIDDYVALARETFPNALLHWEDFGAKNARSILKRYKDKVCTFNDDIQGTGAVSL
SFCAECOLI   TACGGISPAYTLPVVLDVGTNNQQLLNDPLYMGWRNPRITDDEYYEFVDEFIQAVKQRWPDVLLQFEDFAQKNAMPLLNRYRNEICSFNDDIQGTAAVTV
YTSJBACSU   KSFAGVDAFPIALNTNDVDKIVETVKLLEPTFGGVNLEDIAAPNCFIIEERLKKET---------------------------NIPVFHDDQHGTAIVTV
MAOXBACST   KEFAGVDAFPICLDTKDTEEIIQIVKAIAPAFGGINLEDISAPRCFEIEKRLKEEL---------------------------DIPVFHDDQHGTAVVLL
MLEABACSU   DQLAGISGIPILLDTSDPEEIIKTVKHISPGFSGILLEDIGSPHCFEIEDRLKEEL---------------------------NIPVMHDDQHGTAVVTL
                                  Boîte IV
            301                                                                                                 400
YWKABACSU   AGVLAAMKKTGASIKDQRVVIFGAGSAGIGIADQIRDTMVLAGLSEEEANKRFYTLDYRGLLTEDIEGILDFQKPYLRNADEVKDWKRDEKGQ-IPFDEV
MALSBACSU   AAVLSCAKASKVPLRDHRVVIFGAGTAGIGIAEQLREALVREGLSEEESYKRFWCIDRNGLLTDDMDQLLDFQKPYARSADEVKDYQRNGDGGGIDLLEV
SFCAECOLI   GTLIAASRAAGGQLSEKKIVFLGAGSAGCGIAEMIISQTQREGLSEEAARQKVFMVDRFGLLTDKMPNLLPFQTKLVQKRENLSDWDTDSDV--LSLLDV
YTSJBACSU   AGLVNALKLSGKSMSSIKVVANGAGAAGIAIIKLLHHYGVRDIVM-----------------CD-SKGAIYEGRPNGMNDVKNEVAKFTNQDRKDGSLKD
MAOXBACST   AGLLNALKIVDKKLEDIKVVLTGIGAAGIACTKILLAAGVRNIIG-----------------VD-RHGAIHRDETYENPYWQ-EYAQLTNPDNLKGSLSD
MLEABACSU   AAAISACRSAGVDLKEAKVGQIGLGAAGVAICRMFMAYGVNAVYG-----------------TDKSESAMNRLEQYGGQAV--------------SSIEE
                                                                                                  Boîte V
            401                                                                                                 500
YWKABACSU   VRQAKPTILIGTSGVSGAFTEEIVKEMASHVDRPVIMPMSNPTHLAEAVPEDLFKWTDGKVLIATGSPFDNVEYNGVSYEIGQSNNAFAFPGLGLGSIVA
MALSBACSU   VRQAKPTILIGTSTVSGAFTEEIVKEMASHVKRPAILPMSNPTTLSEAKPEDLIEWTEGRALITTGSPFPPVEYNGVTYHIGQANNALVFPGLGLGTIVT
SFCAECOLI   VRNVKPDILIGVSGQTGLFTEEIIREMHKHCPRPIVMPLSNPTSRVEATPQDIIAWTEGNALVATGSPFNPVVWKDKIYPIAQCNNAFIFPGIGLGVIAS
YTSJBACSU   VI-VDADVFIGVSVAGALTKEMVQSM----AKDPIIFAMANPNP--------------EIMPEDAREAGASVVGTGRSDFPNQVNNVLAFPGIFRGALDV
MAOXBACST   VI-AGADVFIGVSAPGILKVEDVKKM----ARDPIVFAMANPIP--------------EIDPELA-EPYVRVMATGRSDYPNQINNVLCFPGIFRGALDC
MLEABACSU   LM-ETCDIVIATTGVPGLIKPAFVR------SGQVILALSNPKP--------------EIEPEAALQAGAAYAADGRS-----VNNVLGFPGIFRGALNA
                Boîte VI
            501                                                                                 588
YWKABACSU   EARIITPAMFAATADAIAEMVDLETPGAGL-LPSIDKLQEVSIQVAIAVAEAAIKDGVA-NRQPEDVKQAVLDAMWTPEYKKVIAK
MALSBACSU   KSKLITDGMFEACARAIAGMVNVGVPGAPM-LPKVEDLRTVSATVAVEVAKTAMKEGVA-TEEPEDIIQAVQDAMWYPVYKPIRAI
SFCAECOLI   GASRITDEMLMSASETLAQYSPLVLNGEGMVLPELKDIQKVSRAIAFAVGKMAQQQGVAVKTSAEALQQAIDDNFWQAEYRDYRRTSI
YTSJBACSU   RATHINEEMKIAAVEAIASLVSEDELSADYVIPAPFDKR-VAPAVAKAVAKAAMETGVA--------RITVDPEEVAEKTRKLTIIGE
MAOXBACST   RAREINEEMKLAAAKAIASVVTEDELNETYIIPSVFNSK-VVERVRQAVVEAAYRTGVA--------RKDNIPVGGYTGQ
MLEABACSU   KSTEINHDMLVAAAEAIAACTKQGD-----VVPQPLDSK-VHHAVAAAVEHAALTAVK
```

Figure 49. Alignement et comparaison des produits des quatre gènes paralogues *ywkA*, *malS*, *ytsJ* et *mleA* chez *Bacillus subtilis* avec SfcA (enzyme malique NAD-dépendante) d'*Escherichia coli* et Maox (enzyme malique NAD-dépendante) de *Bacillus stearothermophilus*. Abréviations : BACSU: *Bacillus subtilis* ; ECOLI: *Escherichia coli* ; BACST: *Bacillus stearothermophilus*.

interprétation. Par ailleurs, l'alignement montre que les enzymes maliques (YwkA, SfcA) et malolactiques (MleA) partagent les mêmes motifs conservés.

Région chromosomique des gènes *yufL-yufM* chez *Bacillus subtilis*, *Bacillus halodurans*, *Bacillus anthracis*, *Oceanobacillus hieyensis*, *Streptococcus agalactiae* et *Streptococcus pyogenes*.

Afin de savoir si le système à deux composants YufL-YufM est aussi impliqué dans la régulation de l'expression de protéines homologues à YwkA et à MaeN chez d'autres bactéries proches de *Bacillus subtilis* et aussi pour savoir si la fonction de *maeN* a un lien avec celle des gènes *mrp*, nous avons réalisée une comparaison préliminaire de la région chromosomique d'intérêt chez *Bacillus subtilis*, trois de ses proches cousines et deux espèces de Streptocoques.

Les résultats de cette comparaison indique que, chez les bactéries considérés à l'exception de *Bacillus subtilis*, les gènes *yufL-yufM* sont physiquement liés à deux gènes codant une enzyme malique homologue à YtsJ et un transporteur homologue à MaeN (plus de 50 % d'identité de séquence) (Figure 50). Cela suggère fortement que le système YufL-YufM sert à réguler l'expression de ces deux gènes de l'utilisation du malate chez ses bactéries. Bien que l'enzyme malique associée chez les autres bactéries présente toujours plus d'identité avec la séquence d'YtsJ que celle d'YwkA, il n'est pas possible de prédire s'il s'agit d'une enzyme NAD ou NADP-dépendante. Par ailleurs, le lien physique entre le gène codant le transporteur et les gènes *mr*p n'est pas conservé ni chez les proches cousines de *Bacillus subtilis*, ni chez les deux Streptocoques. Il semble que l'organisation des gènes de la région *yufL-yufM*, avec l'insertion des gènes *yufNOPQ* qui codent des ABC transporteurs entre *yufL* et *maeN* et le lien entre *maeN* et les gènes *mrp*, soit spécifique à *Bacillus subtilis*.

Détermination des paramètres cinétiques de MalS et d'YtsJ.

Notre étude montre qu'*ywkA* n'est pas nécessaire pour la croissance en présence de malate comme seule source de carbone. Au contraire, l'expression d'*ytsJ* est requise pour une croissance efficace dans ces conditions. Nous nous sommes donc demandés quelle fonction physiologique ce gène remplit. C'est pourquoi nous avons déterminé les caractéristiques enzymatiques d'YtsJ, mais aussi de MalS.

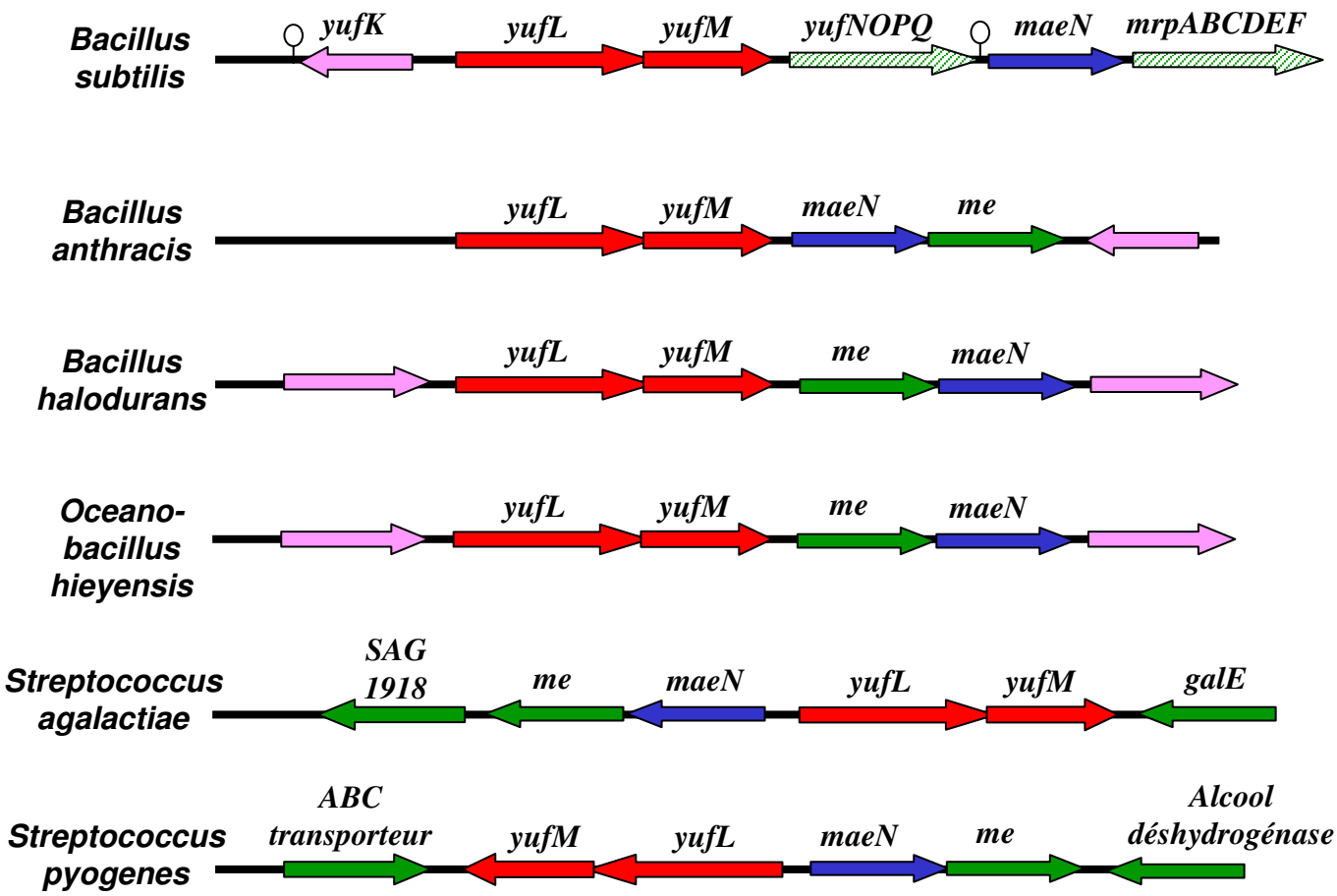


Figure 50. Organisation génétique de la région *yufL-yufM* chez *Bacillus subtilis, Bacillus halodurans, Bacillus anthracis, Oceanobacillus hieyensis, Streptoccocus agalactiae et Streptococcus pyogenes. me* indique un gène dont le produit est clairement homologue à une enzyme malique, plus semblable à YtsJ qu'à YwkA, mais dont on ne peut pas dire si elle est NAD ou NADP dépendante.

Métabolisme

Signalisation

Processus cellulaires

Similaire à des protéines de fonction inconnue

Aucune similarité

opéron

Terminateur

Nous avons surexprimé et purifié MalS et YtsJ en procédant de la même manière que pour YwkA (Article n°2). Les mêmes procédures ont été employées pour déterminer si ces deux protéines possèdent une activité de type enzyme malique et quels sont leurs paramètres cinétiques (Article n°2). Les résultats de nos expériences d'enzymologie démontrent que *malS* et *ytsJ*, comme *ywkA*, codent aussi des enzymes maliques. Cependant, leurs caractéristiques diffèrent (Table 6). MalS est une enzyme malique dont les paramètres cinétiques sont très proches de ceux d'YwkA, ce qui confirme les prédictions de l'analyse de séquence. C'est donc aussi une enzyme malique utilisant préférentiellement le NAD comme cofacteur. En ce qui concerne YtsJ, nos résultats montrent en revanche que son efficacité catalytique est environ 100 fois supérieure en présence de NADP comparé au NAD. Il s'agit donc d'une enzyme malique NADP-dépendante.

Profil d'expression de *malS* et d'*ytsJ*.

Pour déterminer les conditions dans lesquelles ces deux gènes pourraient remplir une fonction, Guillaume a entrepris l'étude de leur expression en réponse à la présence de différentes sources de carbone. Il a donc construit des souches portant les fusions rapportrices de l'activité des régions régulatrices potentielles de la transcription de *malS* et d'*ytsJ* (GM1600) au locus *amyE*. En ce qui concerne *malS*, deux fusions rapportrices ont été construites, la première en considérant que l'activité promotrice est issue de la région juste en amont de la phase codante de *malS* (GM1601) et la seconde en considérant que *malS* pourrait être cotranscrit avec le gène amont *ytnP* (GM1602) (Figure 46). L'activité de la β-galactosidase synthétisée par chaque souche cultivée en milieu riche ou en milieu minimum supplémenté en glucose ou en malate a été mesurée le long de la croissance. S'agissant du gène *malS*, aucune expression ou une expression extrêmement faible a été détectée dans l'ensemble des tests réalisés. Ces résultats sont quelque peu déroutants puisque contradictoires avec l'expression significative obtenue dans diverses études transcriptomiques dont celle figurant dans l'article n°2. Concernant l'expression d'*ytsJ* (Figure 51), nos résultats indiquent qu'en milieu minimum, elle est forte et peu dépendante de la source de carbone présente. En outre, en milieu riche, cette expression chute très rapidement à l'entrée en phase stationnaire.

	Km (mM)	
	NAD	NADP
YwkA	1,10	9,00
MalS	1,30	10,00
YtsJ	1,35	0,55

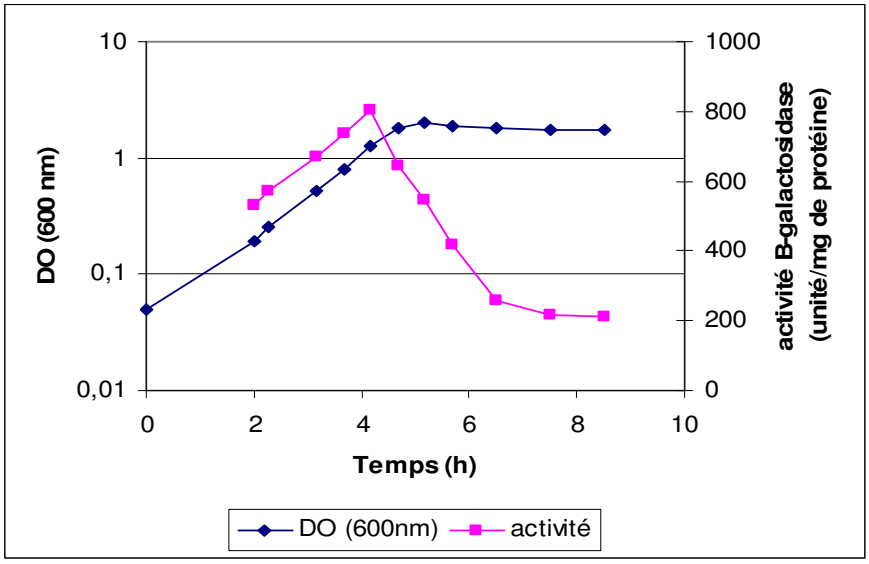
	kcat (s^{-1})	
	NAD	NADP
YwkA	112	88
MalS	117	88
YtsJ	1	40

	Kcat/Km ($s^{-1}.mM^{-1}$)	
	NAD	NADP
YwkA	102	10
MalS	90	8,8
YtsJ	1	73

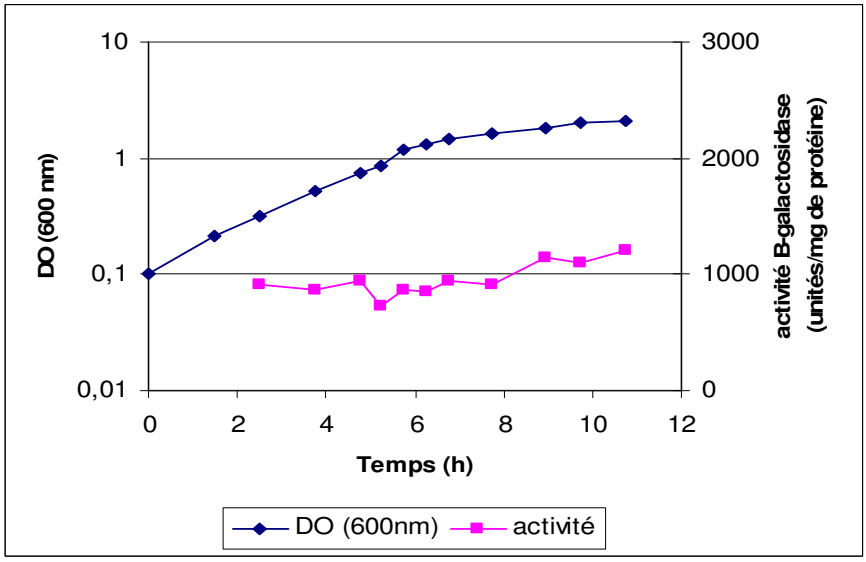
Table 6. Paramètres cinétiques d'YwkA, de MalS et d'YtsJ.

Les tests enzymatiques ont été réalisés dans les mêmes conditions pour les trois protéines (Article n°2).

Milieu SP



MM glucose (1%)



MM malate (1%)

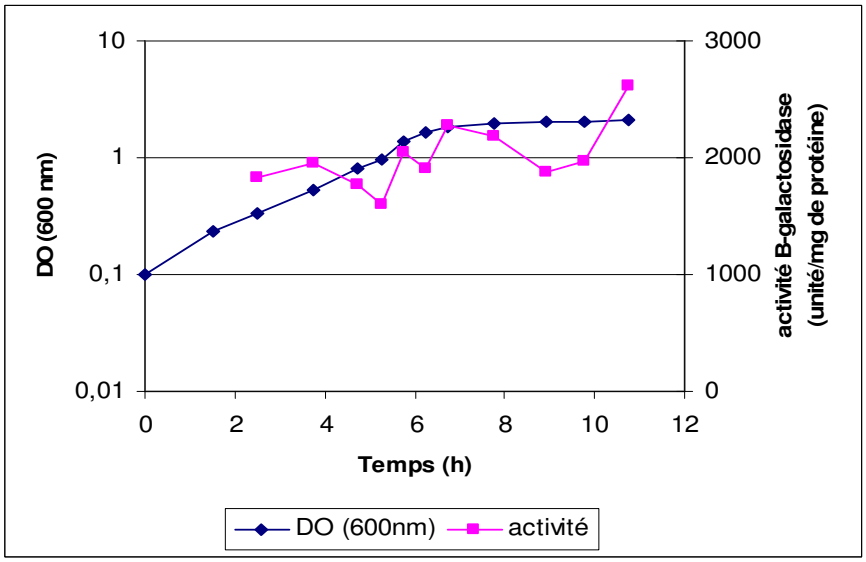


Figure 51. Profil d'expression d'une fusion rapportrice *ytsJ'::lacZ* en milieu de sporulation (SP) et en milieu minimum supplémenté en glucose ou en malate.

5.4. Matériel et Méthodes

Expression, purification de His_6-MalS et de His_6-YtsJ.

Les séquences codantes de *malS* et d'*ytsJ* ont été amplifiées avec les couples d'amorces BamHI-malS/PstI-malS (CgCggATCCgTgAAACAATTCAgAgTAACAAAC / TTCTgCAgTTATATCgCgCgAATCggTTTg) et BamHI-ytsJ/PstI-malS (CgCggATCCATgTCATTAAgAgAAgAAgCATTAC/TTCTgCAgTTATCACCAATAATCg TAAgTTTTC), respectivement, et intégrées dans le plasmide pQE-30 (QIAGEN) pour générer une protéine fusionnée avec une étiquette de 6 histidines en N-terminal. Les plasmides résultants respectifs, pTDp5 et pTDp6, ont été introduits dans la souche M15 pREP4 d'*Escherichia coli* pour produire les souches GTDp5 et GTDp6. Les étapes de surexpression et de purification ont été effectuées suivant les recommandation du manuel QIAEXPRESS (QIAGEN). His_6-MalS et His_6-YtsJ ont été stockées dans le même tampon qu'His_6-YwkA, (Tris 50 mM pH 8, NaCl 150 mM, DTT 1 mM, EDTA 0.5 mM).

Construction des souches GM1601 (*malS'::lacZ*), GM1602 (*ytnP'::lacZ*) et GM1600 (*ytsJ'::lacZ*).

Le vecteur pDG1661 (Guérout-Fleury *et al.*, 1996) a été utilisé pour construire les souches rapportrices de la transcription de *malS* et d'*ytsJ*. Les couples d'amorces lgO6/lgO7 (gCgAATTCTgATgCCgACACATgCCC/CgCggATCCCTgCCTCTTCCTTTCTg), lgO5/lgO7 (gCgAATTCgggCCgTTTgCTACCATg/CgCggATCCCTgCCTCTTCCTTTCTg) et lgO3/lgO4 (gCgAATTCAgCACAgCCAggAgATCC/ CgCggATCCAgTTTCCCCTggTTgAC) ont été utilisés pour amplifier les fragments promoteurs à tester des souches GM1601, GM1602 et GM1600, respectivement. Ces fragments ont été insérés entre les sites *EcoR*I et *BamH*I du vecteur pDG1661 pour générer les plasmides pGL1, pGL2 et pGL3, respectivement. Les constructions résultantes ont été vérifiées par séquençage. Puis, ces plasmides ont été utilisés pour insérer les constructions rapportrices au locus *amyE* de *Bacillus subtilis*.

Etude du transcriptome : mise au point de la technique et résultats préliminaires

6. Présentation générale

La disponibilité de la séquence de génomes entiers a grandement participé à l'avènement de méthodes globales d'analyse. Dès lors, les notions de génomique fonctionnelle, d'analyse transcriptomique et protéomique ont acquis une place très importante dans notre façon d'approcher les questions biologiques. Parmi ces nouvelles approches, l'étude du transcriptome permet de comparer le profil d'expression du génome entier en fonction, par exemple, de valeurs différentes d'un paramètre et ainsi de déterminer les (certains) gènes dont l'expression est affectée par ce paramètre, ou bien encore, de dresser la liste des gènes appartenant à un même régulon.

Dans le cadre de l'étude du métabolisme carboné chez *Bacillus subtilis*, et plus particulièrement des régulations de ce métabolisme, deux questions de type global se présentaient. D'une part, afin de connaître toutes les fonctions physiologiques dans lesquelles il est impliqué, il est important de déterminer le régulon d'un répresseur ou d'un activateur transcriptionnel. D'autre part, beaucoup de gènes ont une fonction encore inconnue mais sont peut-être impliqués dans le métabolisme du carbone. La comparaison du comportement de l'ensemble des gènes de *Bacillus subtilis* en réponse à la croissance en présence de différentes catégories de sources de carbone pourra permettre d'identifier certains de ces gènes. C'est pourquoi l'équipe a décidé de se doter de l'outil transcriptomique.

Lorsque je suis arrivé au laboratoire à l'automne 1999, le deuxième programme européen d'analyse fonctionnelle intitulé Bacell Network s'apprêtait à démarrer. A l'intérieur de ce consortium, le groupe de Michael Hecker à Greifswald (Allemagne) avait déjà engagé les efforts nécessaires pour maîtriser cette technique. C'est donc très près de la Mer Baltique en plein hiver 2000-2001 que je suis joyeusement allé découvrir l'analyse transcriptomique sous la tutelle de Georg Homuth. Il était tellement plus simple d'importer cette technique un peu lourde plutôt que de tout mettre au point tout seul. Je remercie très sincèrement Georg pour le temps qu'il a passé à me former à cette technique. Je fais aussi mes compliments à Georg, bien sûr, mais aussi à l'ensemble des personnes du laboratoire de Michael Hecker que j'ai rencontré pour leur sens de l'accueil.

Le support choisi pour permettre la quantification de l'expression des gènes est une membrane de nylon (Panorama *Bacillus subtilis* Gene Arrays, Sigma-Genosys). On l'appelle quelquefois filtre à haute densité ou macroarray par comparaison avec une microarray, support de type lame de verre et nettement plus miniaturisée. Il s'agit donc d'une membrane de nylon de 12 cm x 24 cm positivement chargée (l'ADN est électronégativement chargé) qui contient 4107 fragments d'amplification PCR correspondant à 4107 ORF de la souche 168CA de *Bacillus subtilis*. Pour la très grande majorité des ORF, l'intégralité de la séquence a été amplifiée. Dans la pratique, pour chacun des ORF, deux spots de 5 ng d'ADN sont déposés côte à côte sur la membrane. Une fois les quelques 8200 dépôts réalisés, l'ADN est fixé aux rayons ultraviolets. Les dépôts sont en fait divisés en trois zones rectangulaires. A chaque coin de chaque zone, est déposé une quantité déterminée d'ADN génomique de *Bacillus subtilis* qui pourra servir de témoins positif d'hybridation et, si nécessaire, de contrôle pour la normalisation (pour comparer deux membranes différentes).

La Figure 52 illustre le principe d'une comparaison transcriptomique d'une souche mutante et de la souche sauvage parentale. Il faut donc une paire de membrane pour réaliser la comparaison de deux conditions expérimentales. Il convient de remarquer dès à présent que deux membranes, bien que fabriquées avec le plus grand soin, ne sont pas strictement identiques. Par conséquent, l'emploi de deux supports distincts introduit un biais lors de la comparaison de l'expression d'un gène entre deux hybridations différentes. Par ailleurs, un certain nombre de facteurs comme la longueur, la séquence ou la conformation d'un fragment PCR particulier peut affecter l'hybridation de telle sorte que celle-ci ne se déroule pas, au niveau moléculaire, de façon similaire à toutes les autres. Il n'est donc pas possible de comparer l'expression de deux gènes différents.

Pour l'heure, au laboratoire, nous avons réalisé une dizaine d'expériences validées (hybridations quantifiables de deux membranes). Les expériences ont porté sur :

- la comparaison transcriptomique de la souche sauvage et d'un mutant *cggR* en conditions néoglucogéniques pour déterminer le régulon de CggR et les éventuels effets indirects de la dérégulation de l'expression des enzymes de la seconde partie de la glycolyse (Article n°1). **(3 comparaisons)**

- la comparaison transcriptomique de la souche sauvage et d'un mutant *yqzB* en conditions glycolytiques pour déterminer le régulon d'YqzB et les

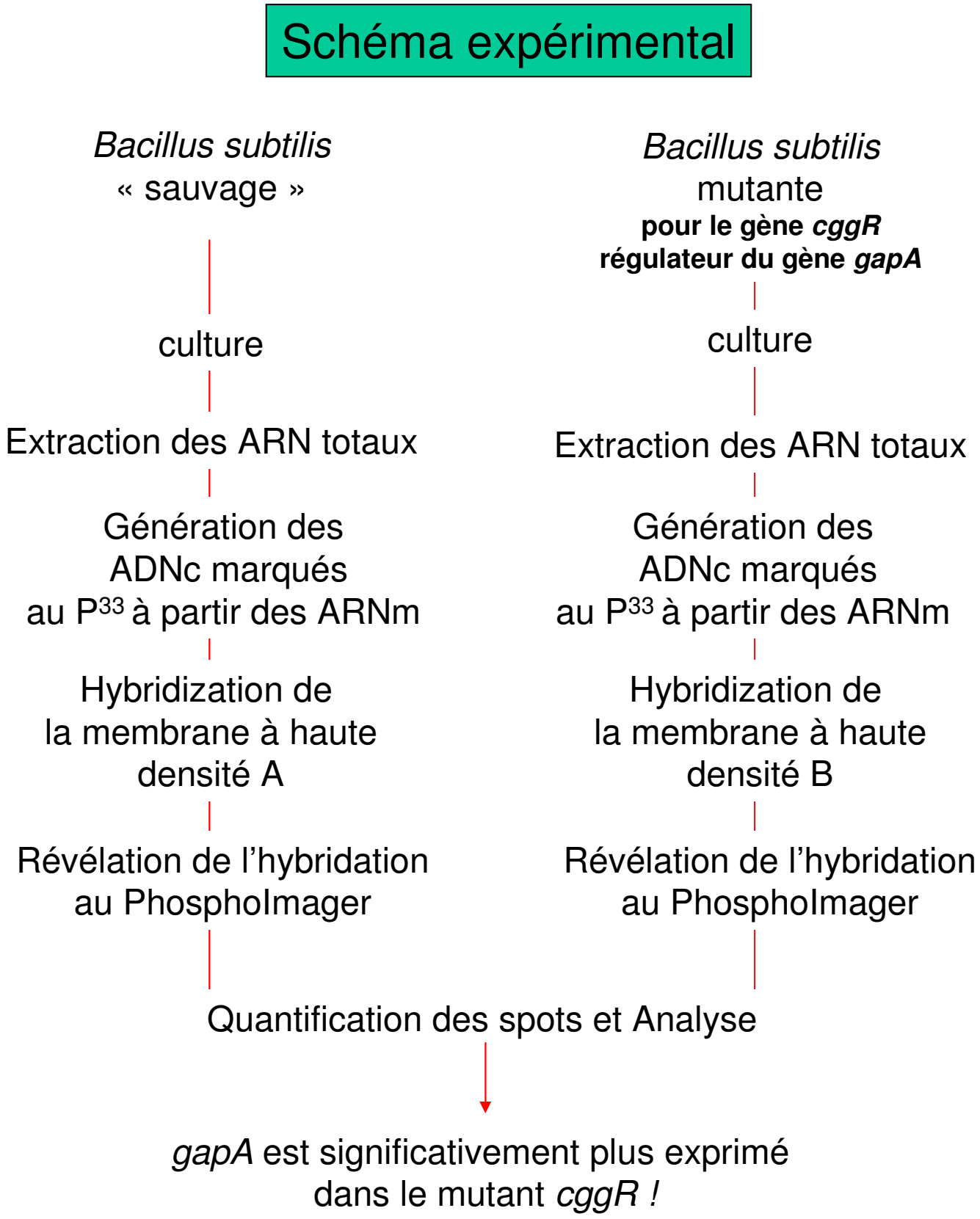


Figure 52. Schéma expérimental d'une comparaison transcriptomique entre une souche sauvage et un mutant *cggR*.

éventuels effets indirects de la dérégulation de l'expression des enzymes de la néoglucogenèse (Servant *et al.*, non publié). **(4 comparaisons + 2 comparaisons 168CA/*yqzByqfL)***

- la comparaison du pattern d'expression de la souche sauvage lors de la croissance en présence d'une source de carbone glycolytique, le glucose, ou néoglucogénique, le succinate. Cette dernière étude fera l'objet de la prochaine partie qui permettra de présenter notre façon de procéder pour chaque étape de l'expérience. **(3 comparaisons).**

7. Glycolyse *versus* Néoglucogenèse : un début d'analyse

7.1. Introduction

Dans le cadre du programme européen, notre but est de découvrir de nouveaux gènes, catalytiques ou régulateurs, impliqués dans le métabolisme du carbone. Nous nous basons sur l'hypothèse qu'au moins une partie des gènes catalytiques ou régulateurs, qui sont impliqués dans le métabolisme carboné, ont une expression différentielle selon le type de source de carbone disponible dans le milieu. Et donc, afin de repérer de tels gènes différentiellement exprimés, nous avons décidé de réaliser une analyse globale du transcriptome de la souche 168CA en fonction de six différentes sources de carbone au moyen des filtres à haute densité. Nous avons choisi trois sources de type glycolytique, le glucose (glycolyse, PTS), le glycérol (glycolyse, non PTS) et le gluconate (voie des pentoses phosphate) et trois sources de type néoglucogénique, le succinate, le malate et le pyruvate. Par ailleurs, l'ampleur d'une telle étude, en terme de quantité de données, nécessitera l'emploi d'un véritable outil statistique pour l'analyse des résultats.

Dans un premier temps, nous avons choisi de comparer une condition glycolytique et une condition néoglucogénique. En effet, les résultats de plusieurs comparaisons transcriptomiques réalisées dans des conditions similaires sont disponibles. Nous pourrons comparer nos données avec ces résultats et vérifier la qualité globale de notre comparaison avant de nous lancer dans une étude globale plus lourde. Les sources de carbone utilisées ont été le glucose pour la condition glycolytique et le succinate pour la condition néoglucogénique.

7.2. Matériel et Méthodes

Le protocole de l'expérience, tel qu'utilisé actuellement, est indiqué de façon non formelle avant de donner les résultats préliminaires pour permettre de mieux cerner les principales difficultés de chaque étape. Les étapes expérimentales sont toujours effectuées en parallèle pour la condition « Succinate » et la condition « Glucose ».

La principale évolution de ce protocole par rapport à celui mis au point à Greifswald concerne l'extraction des ARN totaux. En effet, il ne s'agit pas de la procédure indiquée dans l'article « CggR ». Pascale Servant et moi-même avons modifié le protocole utilisé lors de la première série d'expériences pour rendre les manipulations plus rapides et pour éviter au maximum les contaminations d'ADN.

Cultures bactériennes.

Les cultures ont été réalisées en milieu minimum TSS (Fouet et Sonenshein, 1990) modifié [TrisHCl 50 mM pH 7,5; $(NH_4)_2SO_4$ 15 mM; $MgSO_47H_2O$ 8 mM; KCl 27 mM; trisodium citrate 7 mM; KH_3PO_4 0,6 mM; $CaCl_2$ 2 mM; $FeSO_4$H2O traces; $MnSO_4$ traces; glutamate 5 mM] supplémenté avec du glucose 1 % ou du succinate 1 %. Une préculture de 5 mL est ensemencée à partir d'une colonie fraîche d'un étalement sur boîte LB et deux dilutions successives au 1/10ème et au 1/100ème sont effectuées. Les précultures sont incubées pendant la nuit à 37°C sous agitation. Le lendemain matin, une culture de 30 mL est ensemencée à DO_{600nm} 0,05 à partir d'une préculture qui se trouve en phase exponentielle ($0,4<DO_{600nm}<1$). Lorsque la culture a atteint une DO_{600nm} de 0.6, 12 mL de cellules bactériennes sont prélevés et centrifugés 1 min à 12 000 rpm et le surnageant éliminé. Le culot bactérien est immédiatement congelé à -80°C pour limiter au maximum l'évolution de la population d'ARNm.

Extraction des ARN.

Les ARN sont extraits des culots bactériens au macaloïde selon une méthode développé par l'équipe de Rapoport (ref). L'étape critique est celle de l'extraction au phénol-chloroforme des ARN. Il faut en effet faire très attention de ne pas prélever des protéines et surtout de l'ADN génomique.

Les ARN totaux sont quantifiés par spectrophotométrie. Le calcul du rapport $DO_{260nm}/DO_{280\ nm}$ permet de donner une estimation de leur pureté. Une contamination par les protéines ou le phénol diminue ce rapport et rend difficile la quantification. La qualité des ARN est ensuite vérifiée sur gel d'agarose 1 %. 1 µg de chaque échantillon est déposé.

Finalement, afin de s'assurer de l'excellente qualité des ARN, mais aussi que les cultures ont été réalisées dans les conditions souhaitées, des RT-PCR sont réalisées. En effet, pour cette expérience, nous disposons de gènes témoins dont nous savons qu'ils sont différentiellement exprimés dans ces conditions de croissance : *gapA* et *gapB* ! Il ne s'agit pas de RT-PCR quantitative. Cependant, ces gènes sont suffisamment différentiellement exprimés (*gapA* : x5 ; *gapB* : x50 en ce qui concerne les fusions rapportrices *lacZ*) pour que la différence soit observable. Les RT-PCR sont réalisées avec 0,2 µg d'ARN avec le kit ReadyTo Go (Amersham Pharmacia Biotech). En parallèle, des réactions PCR permettent de contrôler que les produits obtenus sont bien issus de la réaction de reverse transcription.

Synthèse des ADNc marqués au ^{33}P par reverse transcription.

Les ARN sont utilisés pour synthétiser les ADNc qui seront hybridés sur la membrane. Cette synthèse est effectuée avec le kit Panorama *Bacillus subtilis* cDNA Labelling kit (Sigma-Genosys). Les oligonucléotides (labeling primer mix) spécifiques de chaque ORF transcrit sont d'abord hybridés aux ARNm présents dans 2 µg d'ARN totaux. Puis, la réaction de réverse transcription est effectuée en présence de [α-^{33}P]-dCTP après quoi les ARN sont dégradés par hydrolyse alcaline. Enfin, les ADNc sont précipités à l'éthanol et repris dans de l'eau.

Puisqu'il faudra comparer deux hybridations différentes, il est très important de toujours faire en sorte que les conditions expérimentales soient toujours les plus proches possibles. C'est pourquoi l'efficacité de marquage est déterminée pour les deux réactions à l'aide d'un compteur à scintillation. Cette efficacité doit être supérieure à 50 % de radioactivité incorporée et l'écart ne pas dépasser 20 % entre les deux échantillons d'ADNc.

Hybridation des membranes et révélation du marquage.

Les membranes sont d'abord saturées par préhybridation dans la solution d'hybridation (Sigma-Genosys) en présence d'ADN de sperme de saumon à 65°C pendant 2 h. Puis, les ADNc sont dénaturés et ajoutés. L'hybridation dure 20 h. Après lavage, les filtres sont scellés et mis en cassette en contact avec un écran dont les cristaux sont sensibles à

l'émission radioactive (Kodak). L'exposition dure de 1 à 5 jours suivant le lot de membranes et l'activité de la radioactivité. Le marquage est finalement révélé au PhosphoImager (Molecular Dynamics) à une résolution de 50 µm. Encore une fois, vérification est faite que le marquage total est proche pour les deux hybridations et que l'intensité du marquage est restée dans une gamme linéaire (qu'il n'y a pas trop de spots dont les pixels sont hors de la gamme dynamique mesurable par l'appareil).

Quantification et analyse des membranes révélées.

La quantification des signaux d'expression et du bruit de fond est effectuée avec le logiciel ArrayVision (Imaging Research) et l'analyse qui suit est réalisée avec le logiciel GeneSpring (Silicon Genetics) suivant les recommandations indiquées dans les manuels livrés.

Deux procédures de normalisation sont réalisées pour permettre la comparaison directe des expériences. Dans ArrayVision, pour supprimer la différence d'exposition ou d'efficacité d'hybridation globale entre deux membranes, les valeurs d'intensité des spots sont normalisées par rapport à la médiane, ramenée à 1, du signal de tous les spots. Cette fonction présuppose que, quelles que soient les conditions de croissance, l'expression des gènes est globalement toujours la même. Elle n'est donc pas toujours utilisable. Puis, les données sont « injectées » dans GeneSpring. Il faut alors normaliser les signaux d'expression pour supprimer la différence d'échelle d'intensité entre plusieurs expériences. Pour cela, l'expression d'un gène sur une membrane particulière est normalisée par rapport à la médiane de toutes les mesures de ce gène sur toutes les autres membranes. A travers un exemple, la Table 7 permet de mieux comprendre ces procédures de normalisation.

7.3. Résultats préliminaires

Au total, trois comparaisons « Succinate/Glucose » ont été effectuées : H1, H2 et H3. Ces trois expériences sont indépendantes dans le sens où trois lots d'ARN totaux provenant de cultures réalisées des jours différents et trois lots de membranes ont été utilisés. La qualité de la manipulation est très satisfaisante, le marquage total à l'intérieur d'un couple de filtre étant

Signaux d'expression bruts			
Gène	**Expérience 1**	**Expérience 2**	**Expérience 3**
A	1000	2000	1500
B	1000	2000	500
C	100	200	50
D	1000	2000	500
E	10 000	20 000	5000

Normalisation 1: (signal du gène X dans l'expérience X)/(Médiane de toutes les mesures de l'expérience X)

Signaux d'expression après normamisation 1			
Gène	**Expérience 1**	**Expérience 2**	**Expérience 3**
A	1	1	3
B	1	1	1
C	0,1	0,1	0,1
D	1	1	1
E	10	10	10

Normalisation 2 : (signal du gène X dans l'expérience X)/(Médiane de toutes les mesures du gène X dans toutes les expériences)

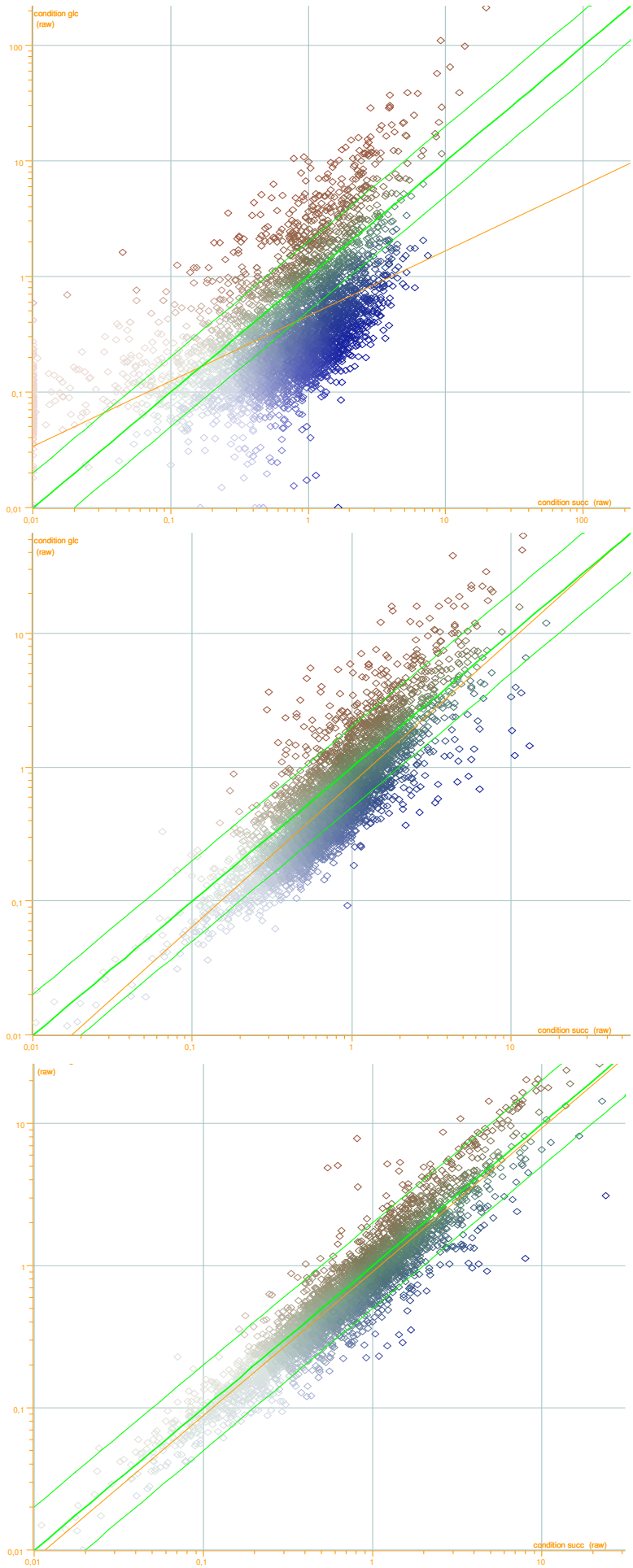
Signaux d'expression après normalisation 2			
Gène	**Expérience 1**	**Expérience 2**	**Expérience 3**
A	1	1	3
B	1	1	1
C	1	1	1
D	1	1	1
E	1	1	1

Table 7. Exemple de normalisation des données brutes de transcriptome.

comparable et le bruit de fond étant très faible, comme en témoigne la Figure 53 qui montre l'image brute du marquage de la membrane correspondant à la culture en présence de glucose de l'expérience H3.

La Figure 54 permet de visualiser les résultats globaux de chacune des expériences sous forme de graphique de dispersion (scatter plot) des signaux d'expression normalisés sous ArrayVision. Pour chaque expérience, la valeur du coefficient de corrélation de Pearson permet d'estimer la qualité de la comparaison. De manière simplifiée, on s'attend à ce que l'allure générale de la dispersion des gènes soit symétrique de part et d'autre de la diagonale représentant les gènes dont l'expression est identique dans les deux conditions et de plus en plus étroite vers les gènes les plus fortement exprimés. Les comparaisons H1 et H2 ont des coefficients de corrélation très faible. L'expérience H1, surtout, pose problème. Notre première hypothèse pour expliquer ce résultat fut de penser que les conditions de croissance devaient être trop différentes et donc entraîner des pattern d'expression très différents. Il est en revanche plus difficile de savoir si cette opposition est due à la différence de source de carbone ou à des aléas de manipulation. En effet, les vitesses de croissance dans le milieu utilisé en présence de glucose ou de succinate comme seule source de carbone sont très différentes (du simple au double). Il aurait été plus judicieux de choisir comme source de carbone néoglucogénique le malate puisqu'il permet une croissance presqu'aussi rapide que le glucose (Article n°2). Mais la comparaison H3, réalisée dans les mêmes conditions de culture, montre une dispersion acceptable des gènes ce qui suggère que la différence de vitesse de croissance n'est pas la cause majeure à ce problème. Il s'est avéré en fait que la différence majeure entre l'expérience H1 et les expériences H2 et H3 se situait au niveau de l'étape de récupération du culot bactérien pour l'extraction d'ARN. Nous supposions que le fait de prélever les cellules et de les centrifuger à 4°C provoquait un choc thermique qui pourrait entraîner certains changements dans le profil d'expression du génome. Il nous a donc semblé plus judicieux de culoter les cellules directement à 37°C puis de rapidement les congeler à –80°C. Pour l'expérience H1, les ARN avaient été extraits de cellules prélevées à 4°C alors qu'ils avaient été extraits de cellules prélevées à 37°C pour les expériences H2 et H3. Quoi qu'il en soit, cette première expérience sera laissée de côté. Concernant l'expérience H2, l'axe de dispersion est aussi un peu écarté de la diagonale, peut-être un peu pour la même raison. Les résultats ont quand même été analysés. Finalement, seule l'expérience H3 semble vraiment correcte et analysable (comme suit).

Figure 53. Image brute du marquage de la membrane « H3-Glucose » après révélation au PhosphoImager (Molecular Dynamics).



H1
r=0,599

H2
r=0,662

H3
r=0,842

Figure 54. Résultats globaux des trois comparaisons transcriptomiques « Succinate/Glucose » sous forme de graphique de dispersion en échelle logarithmique. En abscisse : expression normalisée en condition Succinate. En ordonnée : expression normalisée en condition Glucose. r: coefficient de corrélation de Pearson. La droite orange indique la ligne de meilleure corrélation.

Comme évoqué précédemment, en partie du fait du faible bruit de fond, peu de gènes apparaissent très faiblement ou pas du tout exprimés. Dans les expériences H2 et H3, le rapport signal/bruit de 3 correspond à une valeur d'expression de 0,4 sur l'échelle des scatter plot. Ainsi, dans cette série de comparaison, approximativement 80 % des gènes sont considérés comme significativement exprimés dans au moins l'une des deux conditions et sont pris en compte dans la comparaison. Parmi ces gènes, nous avons ensuite considéré les gènes dont le ratio d'expression entre les deux conditions est supérieur à 2 (A/B>2 et B/A>2). Respectivement 779 (121 avec un ratio de 3) et 145 gènes sont plus exprimés lors de la croissance en présence de succinate dans les expériences H2 et H3. Les résultats indiquent que 184 et 53 gènes sont plus exprimés lors de la croissance en présence de glucose dans les expériences H2 et H3, respectivement. La Table 8 liste l'intersection des gènes significativement différentiellement exprimés dans les deux comparaison : 57 gènes, dont 24 « inconnus », ou 38 gènes, dont 4 « inconnus », sont retrouvés plus transcrits de façon reproductible en présence de succinate ou de glucose, respectivement. On retrouve bien l'opéron *gapA* plus exprimé lors de la glycolyse et inversement les gènes *gapB* et *pckA* plus exprimés lors de la néoglucogenèse.

Pour vérifier la qualité de nos manipulations, les résultats de l'expérience H3 (la plus satisfaisante) ont été comparés à ceux obtenus dans quatre autres expériences transcriptomiques similaires. La première expérience est celle décrite dans l'article « *ywkA* » qui compare l'expression des gènes en milieu minimum supplémenté en glucose ou en malate comme seule source de carbone et qui vise à déterminer les gènes impliqués dans la glycolyse et la néoglucogenèse (Article n°2). La deuxième est une expérience très comparable qui a été réalisée dans un milieu minimum contenant du succinate et du glutamate comme sources de carbone et d'azote, respectivement, supplémenté ou non en glucose (Blencke *et al.*, 2003 sous presse). Les deux autres expériences avaient pur but de lister les gènes soumis à la répression catabolique par le glucose indépendamment ou dépendamment de CcpA (Moreno *et al.*, 2001; Yoshida *et al.*, 2001). On peut remarquer que ces deux dernières analyses ont été effectuées à partir de cultures en milieu complexe, en LB et en milieu Difco de sporulation (DSM) avec ou sans ajout de glucose. Par ailleurs, la souche utilisée par Moreno et ses collaborateurs n'est pas la souche 168CA (*trpC2*) mais ST1OO [*trpC2 pheA1 Δ(spo0E)*]. En outre, il n'est pas mentionné combien de fois ces expériences ont été reproduites.

Gènes plus exprimés en présence de succinate : 57 gènes dont 24 codent une protéine inconnue.

Gene name	Description				
PHOD	phosphodiesterase/alkaline phosphatase D				
YDBH	similar to C4-dicarboxylate transport protein				
YKVU	similar to spore cortex membrane protein				
YURP	similar to opine catabolism				
NARH	nitrate reductase (beta subunit)				
SPSJ	spore coat polysaccharide synthesis				
MSMX	multiple sugar-binding transport ATP-binding protein				
BGLP	phosphotransferase system (PTS) beta-glucoside-specific enzyme IIABC component				
LCTP	L-lactate permease				
TREA	trehalose-6-phosphate hydrolase				
LIPB	lipase				
YGAL	similar to nitrate ABC transporter (binding protein)				
PRKA	serine protein kinase				
GLPD	glycerol-3-phosphate dehydrogenase				
MCPC	methyl-accepting chemotaxis protein				
BGLC	endo-1,4-beta-glucanase				
GGT	gamma-glutamyltranspeptidase	glutathione metabolism			
CGEA	involved in maturation of the outermost layer of the spore				
SIGF	RNA polymerase sporulation-specific sigma factor (sigma-F)				
LCFA	long chain acyl-CoA synthetase	fatty acid metabolism			
ARAB	L-ribulokinase	L-arabinose utilization			
ARAA	L-arabinose isomerase	L-arabinose utilization			
GAPB	glyceraldehyde-3-phosphate dehydrog	glycolysis			
PCKA	phosphoenolpyruvate carboxykinase				
GLGA	starch (bacterial glycogen) synthase				
GERAA	germination response to L-alanine and related amino acids (earliest stage)				
LACA	beta-galactosidase				
RBSR	negative regulation of the ribose operon (rbsRKDACB)				
COTG	morphogenetic protein	required for the incorporation of CotB into the coat			
UNG	uracil-DNA glycosylase				
IOLI	myo-inositol catabolism				
IOLF	myo-inositol catabolism	inositol transport protein			

Gènes plus exprimés en présence de glucose : 38 gènes dont 4 codent une protéine inconnue.

Gene name	Description				
RPLA	ribosomal protein L1 (BL1)				
RPSL	ribosomal protein S12 (BS12)				
RPLC	ribosomal protein L3 (BL3)				
RPLD	ribosomal protein L4				
RPLW	ribosomal protein L23				
RPLB	ribosomal protein L2 (BL2)				
RPSS	ribosomal protein S19 (BS19)				
RPLV	ribosomal protein L22 (BL17)				
RPSC	ribosomal protein S3 (BS3)				
RPLP	ribosomal protein L16				
RPLN	ribosomal protein L14				
RPSN	ribosomal protein S14				
RPLF	ribosomal protein L6 (BL8)				
RPLO	ribosomal protein L15				
INFA	initiation factor IF-I				
SRFAA	surfactin synthetase				
SRFAB	surfactin synthetase				
SRFAD	surfactin synthetase				
PURN	phosphoribosylglycinamide formyltransferase	purine biosynthesis			
PEL	pectate lyase				
PTSH	histidine-containing phosphocarrier protein of the phosphotransferase system (PTS) (HPr protein)				
RPSB	ribosomal protein S2				
YLXS	alternate gene name: ymxA; similar to hypothetical proteins				
RPSO	ribosomal protein S15 (BS18)				
RPSU	ribosomal protein S21				
DNAK	class I heat-shock protein (chaperonin)				
AROA	3-deoxy-D-arabino-heptulosonate 7-phosphate synthase	shikimate pathway			
ENO	enolase	glycolysis			
TPI	triose phosphate isomerase	glycolysis			
PGK	phosphoglycerate kinase	glycolysis			
GAP	glyceraldehyde-3-phosphate dehydrogenase	glycolysis			
SSB	single-strand DNA-binding protein				
COTN	spore coat-associated protein				
COMGF	probably part of the DNA transport machinery	competence			

Table 8. Liste incomplète des gènes différentiellement exprimés (sigan/bruit>3 et ratio>2) dans les expériences H2 et H3. *gapB* et *pckA* (bleu foncé) sont plus exprimés en conditions néoglucogénique et l'opéron *gapA* (orange foncé) est plus exprimé en conditions glycolytiques, comme démontré expérimentalement.

Les résultats de cette comparaison sont listés dans les Tables 9 et 10. Globalement, ils montrent que nous retrouvons la régulation observée expérimentalement et par analyse transcriptomique dans des conditions similaires pour un grand nombre de gènes impliqués dans le métabolisme du carbone.

Les gènes du transport et du métabolisme dont l'induction par le glucose est expérimentalement démontrée sont retrouvés : l'opéron *ptsGH* et l'opéron *gapA* (*gapA*, *pgk*, *tpi*, et *eno*) alors que ceux constitutivement exprimés ou très faiblement régulés dans ces conditions ne sont pas retrouvés (Table 9). Concernant l'induction de l'opéron *pdhABCD* (complexe pyruvate déshydrogénase) (Blencke *et al.*, 2003 sous presse), seul le gène *pdhA* est significativement différentiellement exprimé dans nos conditions. De plus, l'expression de gènes impliqués dans le transport et la biosynthèse des acides aminés a été trouvée induite par le glucose. C'est notamment le cas de l'opéron *gltAB* (glutamate synthase), dont l'expression est dépendante de la présence de CcpA (Faires *et al.*, 1999) (cf partie 2.3). De même, le gène *lysC*, qui code une aspartokinase est aussi induit par le glucose dans notre expérience et celle de Blencke et de ses collaborateurs (Table 9). Enfin, certains gènes de biosynthèse des nucléotides sont induits par le glucose dans notre expérience : *purC*, *purF* et *purN* (Table 9). Ces gènes appartiennent au même opéron (*purEKBCSQLFMNHD*). Les autres gènes de l'opéron sont aussi induits par le glucose mais dans une moindre mesure. L'induction de l'opéron *pur* dans des conditions très similaires a également été observée par Blencke et ses collègues (2003 sous presse).

Dans notre expérience H3, 145 gènes sont significativement plus exprimés (ratio > 2) en présence de succinate que de glucose comme source de carbone (Table 10). Parmi eux, on retrouve, beaucoup de gènes impliqués dans le transport de sources de carbone spécifiques et leur catabolisme dont l'expression est soumise à la répression catabolique en présence de glucose (Table 10). Beaucoup de ces gènes sont directement ou indirectement régulés *via* CcpA (Table 10). Concernant les voies métaboliques centrales, l'expression des gènes *gapB* (GAPDH anabolique) et *pckA* (phosphoénolpyruvate carboxykinase), nécessaires à la néoglucogenèse, est bien induite en conditions néoglucogéniques (Table 10). L'expression de *citZ* (citrate synthase majeure) est bien réprimée par le glucose. De plus, les opérons *sucCD* (succinylCoA synthétase) et *odhAB* sont aussi réprimés par le glucose dans notre expérience et dans au moins deux autres comparaisons transcriptomiques (Table 10).

Gène	Ratio	Fonction	Doan *et al.*	Blencke *et al.*
adaB	2,14	adaptative response to DNA alkylation		
cotN	**8,19**	**spore coat-associated protein**	**D**	
eno	**3,69**	**enolase**	**D**	**B**
gap	**9,68**	**glyceraldehyde-3-phosphate dehydrogenase**	**D**	
gltA	**2,31**	**glutamate synthase (large subunit), gltAB**		**B**
gltB	4,26	glutamate synthase (small subunit)		
hisS	2,17	histidyl-tRNA synthetase		
hisZ	2,58	histidyl-tRNA synthetase		
infA	2,06	initiation factor IF-I		
lctE	**8,93**	**L-lactate dehydrogenase**	**D**	
lysC	**2,51**	**diaminopimelate/lysine biosynthesis**		**B**
pdhA	**2,11**	**pyruvate dehydrogenase (E1 alpha subunit)**		**B**
pel	**2,47**	**pectate lyase**	**D**	
pgk	**3,34**	**phosphoglycerate kinase**		**B**
ptsG	**4,38**	**phosphotransferase system (PTS) glucose-specific enzyme IIABC component**	**D**	**B**
ptsH	3,28	HPr		
purC	2,18	purine biosynthesis		
purF	2,19	purine biosynthesis		
purN	2,59	purine biosynthesis		
pyrE	2,69	pyrimidine biosynthesis		
rplC	2,10	ribosomal protein L3 (BL3)		
rplF	2,19	ribosomal protein L6 (BL8)		
rplL	**2,02**	**ribosomal protein L12 (BL9)**	**D**	
rplQ	2,11	ribosomal protein L17 (BL15)		
rplR	2,09	ribosomal protein L18		
rplV	2,14	ribosomal protein L22 (BL17)		
rpmB	2,12	ribosomal protein L28		
rpmC	2,03	ribosomal protein L29		
rpmF	2,51	ribosomal protein L32		
rpmJ	**2,00**	**ribosomal protein L36 (ribosomal protein B)**	**D**	
rpsC	2,15	ribosomal protein S3 (BS3)		
rpsL	2,15	ribosomal protein S12 (BS12)		
rpsO	2,01	ribosomal protein S15 (BS18)		
ssb	2,12	single-strand DNA-binding protein		
tpi	**3,25**	**triose phosphate isomerase**	**D**	**B**
ycdA	**2,31**		**D**	
ydjL	2,17	similar to L-iditol 2-dehydrogenase		
yjhA	2,51			
yjlC	2,01			
ylnB	**2,45**	**similar to sulfate adenylyltransferase**	**D**	
ylqC	2,38	similar to hypothetical proteins		
ymfE	2,02	similar to phage-related protein		
ypmB	2,02			
ypuC	**2,66**		**D**	
yraM	2,00	similar to hypothetical proteins		
yueB	2,15			
yukC	2,28			
yukE	2,94			
yvzA	2,52			
ywqD	2,01	similar to capsular polysaccharide biosynthesis		
ywqH	2,28			
ywqJ	2,26	similar to hypothetical proteins from B. subtilis		
yxeM	2,45	similar to amino acid ABC transporter (binding protein)		

Table 9. Liste des gènes significativement plus exprimés en présence de glucose que de succinate dans l'expérience H3 et croisement de cette liste avec les résultats d'autres comparaisons de ce type. Pour chaque gène, les lettres D, B, M et Y indiquent s'il est retrouvé régulé de la même manière dans les travaux de Article n°2, Blencke *et al.*, 2003, Moreno *et al.*, 2001 et Yoshida *et al.*, 2001, respectivement.

Gène	Ratio	Fonction	Doan *et al.*	Blencke *et al.*	Moreno *et al.*	Yoshida *et al.*
ald	3,16	L-alanine dehydrogenase required for sporulation				
bglP	**2,38**	**phosphotransferase system (PTS) beta-glucoside-specific enzyme IIABC component**			**M**	
bglS	**2,14**	**endo-beta-1,3-1,4 glucanase**		**B**		
bmrU	2,05	multidrug resistance protein				
cccA	**2,55**	**cytochrome c550**	**D**	**B**	**M**	
citM	2,64	secondary transporter of the Mg2+/citrate complex				
citZ	**2,17**	**citrate synthase II**		**B**		
cotE	**2,20**	**outer coat assembly**			**M**	
cotG	4,37	required for the incorporation of CotB into the coat				
cotJA	2,96	polypeptide composition of the spore coat; required for the assembly of CotJC				
cotJB	2,63	polypeptide composition of the spore coat				
cotJC	2,91	polypeptide composition of the spore coat				
cotW	2,40	spore coat protein (insoluble fraction)				
cotX	**2,89**	**spore coat protein (insoluble fraction)**			**M**	
cotY	3,13	spore coat protein (insoluble fraction)				
cotZ	3,04	spore coat protein (insoluble fraction)				
ctaC	**4,44**	**cytochrome caa3 oxidase (subunit II)**		**B**		**Y**
dhaS	**2,46**	**aldehyde dehydrogenase**			**M**	
dhbF	2,41	involved in siderophore 2,3-dihydroxybenzoate (DHB) synthesis				
dppA	**2,54**	**dipeptide ABC transporter, dppABCDE**		**B**		
dppC	2,10	dipeptide ABC transporter (permease)				
dppD	5,27	dipeptide ABC transporter (ATP-binding protein)				
dppE	**3,32**	**dipeptide ABC transporter (dipeptide-binding protein)**			**M**	
flgC	2,27	flagellar basal-body rod protein				
flgE	2,09	flagellar hook protein				
flhO	2,47	flagellar basal-body rod protein				
fliD	2,38	flagellar hook-associated protein 2 (HAP2)				
fliE	2,30	flagellar hook-basal body protein				
fliL	2,20	flagellar protein				
fliS	2,04	flagellar protein				
fliT	2,65	flagellar protein				
gapB	**2,00**	**glyceraldehyde-3-phosphate dehydrogenase**	**D**	**B**	**M**	**Y**
gerAA	2,13	germination response to L-alanine and related amino acids (earliest stage)				
gntK	**2,24**	**gluconate kinase**	**D**		**M**	
gntP	2,31	gluconate permease				
gsiB	**2,81**	**general stress protein**			**M**	
iolF	**2,28**	**inositol transport protein**	**D**		**M**	**Y**
iolH	**3,09**	**myo-inositol catabolism**	**D**		**M**	**Y**
iolI	**2,28**	**alternate gene name: yxdH; myo-inositol catabolism**	**D**		**M**	**Y**
licC	**2,50**	**phosphotransferase system (PTS) lichenan-specific enzyme IIC component**			**M**	
mmgA	2,09	alternate gene name: yqiL, operon subject to catabolite repression				
moaD	2,29	molybdopterin converting factor (subunit 1)				
nasB	2,12	assimilatory nitrate reductase (electron transfer subunit)				
odhB	**2,30**	**2-oxoglutarate dehydrogenase complex (dihydrolipoamide transsuccinylase, E2 subunit)**		**B**	**M**	
opuBD	2,01	choline ABC transporter (membrane protein)				
pckA	**2,18**	**phosphoenolpyruvate carboxykinase**	**D**	**B**		**Y**
phrA	3,00	inhibitor of the activity of phosphatase RapA				
phrE	**2,06**	**regulator of the activity of phosphatase RapE**		**B**		
pmi	2,01	mannose-6-phosphate isomerase				
qcrA	**2,09**	**menaquinol:cytochrome c oxidoreductase (iron-sulfur subunit)**		**B**	**M**	**Y**
qcrC	**2,38**	**menaquinol:cytochrome c oxidoreductase (cytochrome b/c subunit)**			**M**	
rapA	**4,09**	**response regulator aspartate phosphatase**		**B**		

Gène	Ratio	Fonction	Doan *et al.*	Blencke *et al.*	Moreno *et al.*	Yoshida *et al.*
rapH	2,03	response regulator aspartate phosphatase				
rbsB	**2,12**	**ribose ABC transporter (ribose-binding protein)**	**D**		**M**	**Y**
ruvB	2,12	Holliday junction DNA helicase				
sacT	4,80	transcriptional antiterminator ; positive regulation of sacA and sacP				
sspE	2,10	small acid-soluble spore protein (gamma-type SASP)				
sucC	**2,38**	**succinyl-CoA synthetase (beta subunit), operon *sucCD* ; *sucD* (succ/glc=1.53)**		**B**	**M**	**Y**
zwf	2,02	similar to glucose-6-phosphate 1-dehydrogenase (pentose phosphate)				
ybcO	**7,62**			**B**		
ybcP	2,63					
ybcQ	2,43	similar to hypothetical proteins				
ybdE	2,16	similar to hypothetical proteins				
ybyB	3,58					
ycbA	2,06	similar to two-component sensor histidine kinase [YcbB]				
ycbF	**3,55**	**similar to glucarate dehydratase**	**D**			
ycbH	**2,19**	**similar to hypothetical proteins**	**D**			
yckG	2,34	similar to D-arabino 3-hexulose 6-phosphate formaldehyde lyase				
ydaP	2,96	similar to pyruvate oxidase				
ydbA	2,23	similar to ABC transporter (binding protein)				
ydbD	2,54	similar to manganese-containing catalase				
ydbE	2,08	dctB, C4-dicarboxylate binding protein				
ydiA	2,32	alternate gene name: ydxA; similar to thiamin-monophosphate kinase				
yerA	**2,01**	**similar to adenine deaminase**	**D**			**Y**
yesP	2,28	similar to lactose permease				
yesQ	2,67	similar to lactose permease				
yesR	2,41	similar to hypothetical proteins				
yesS	2,37	similar to transcriptional regulator (AraC/XylS family)				
yetO	2,10	alternate gene name: yfnJ; similar to cytochrome P450 / NADPH-cytochrome P450 reductase				
yfhD	2,71					
yfiS	2,32	similar to multidrug resistance protein				
yfjU	3,11					
yflJ	2,51					
yflN	2,48	cotranscribed with citM, regulated by CcpA (glucose)				
yflT	2,57					
yfmH	5,18					
yfmJ	3,39	similar to quinone oxidoreductase				
yfmK	2,09					
yfmO	2,11	similar to multidrug-efflux transporter				
ygcA	2,07	alternate gene name: yzeC; similar to hypothetical proteins from B. subtilis				
yisG	2,44					
yjcC	2,29					
yjcL	2,44					
yjcQ	2,00					
yjdB	2,34					
yjfB	2,49					
ykfB	2,63	similar to chloromuconate cycloisomerase				
ykzA	2,56	similar to general stress protein				
ykzI	2,91					
ylbA	2,77					
ylqB	**3,01**		**D**			
ymaH	**2,30**	**similar to host factor-1 protein**		**B**		
ymfJ	2,48	similar to hypothetical proteins from B. subtilis				
ymzC	2,12					
yobH	2,57	similar to hypothetical proteins				

Gène	Ratio	Fonction	Doan *et al.*	Blencke *et al.*	Moreno *et al.*	Yoshida *et al.*
yobL	2,40	similar to hypothetical proteins from B. subtilis				
yocH	**2,38**	**similar to cell wall-binding protein**				**Y**
yocN	2,18	similar to permease				
yocK	2,60	similar to macrolide glycosyltransferase				
yojL	**2,49**	**similar to cell wall-binding protein**		**B**		
yqcD	2,00	similar to phage-related protein				
yqfT	2,74					
yqgW	2,15					
yqgZ	2,59	similar to hypothetical proteins				
yqiK	2,01	similar to glycerophosphodiester phosphodiesterase				
yqzH	2,04					
yqzI	2,07					
yrbE	2,07	similar to opine catabolism				
ytxG	2,34	alternate gene name: csb40; similar to general stress protein				
ytxH	2,56	alternate gene name: csb40; similar to general stress protein				
ytzE	2,11	similar to transcriptional regulator (DeoR family)				
yuiB	4,03	similar to hypothetical proteins from B. subtilis				
yurM	2,93	similar to sugar permease				
yurO	2,50	similar to multiple sugar-binding protein				
yurP	7,11	similar to opine catabolism				
yvAA	2,56	similar to hypothetical proteins				
yvaZ	4,04					
yvbA	4,71	similar to transcriptional regulator (ArsR family, negative regulator)				
yvoD	2,19					
yvzB	2,05	similar to flagellin				
ywcI	2,52	alternate gene name: ipa-46d				
ywdK	2,62	alternate gene name: ipa-61d; similar to hypothetical proteins				
ywpH	2,39	similar to single-strand DNA-binding protein				
ywrF	2,25					
ywsB	**2,01**	**similar to hypothetical proteins from B. subtilis**	**D**			**Y**
ywzA	2,50	similar to hypothetical proteins from B. subtilis				
yxaA	**3,23**	**similar to hypothetical proteins**	**D**			
yxaM	2,09	similar to antibiotic resistance protein				
yxbB	**2,40**	**alternate gene name: yxaP; similar to hypothetical proteins**		**B**		
yxbC	2,22	alternate gene name: yxaQ				
yxiQ	2,74	similar to Mg2+/citrate complex transporter				
yxlJ	2,04	similar to DNA-3-methyladenine glycosidase				
yxnB	2,35					
yxzF	2,03					
yycA	2,01	similar to hypothetical proteins from B. subtilis				

Table 10. Liste des gènes significativement plus exprimés en présence de succinate que de glucose dans l'expérience H3 et croisement de cette liste avec les résultats d'autres comparaisons de ce type.
Pour chaque gène, les lettres D, B, M et Y indiquent s'il est retrouvé régulé de la même manière dans les travaux de Article n°2, Blencke *et al.*, 2003, Moreno *et al.*, 2001 et Yoshida *et al.*, 2001, respectivement.

Ces résultats sont cohérents avec la faible activité du cycle de Krebs lors de la croissance en présence de glucose.

De façon intéressante, à l'exception du gène *zwf* codant la glucose-6-phosphate 1-déshydrogénase qui est plus exprimé en présence de succinate que de glucose dans notre comparaison, les gènes impliqués dans la voie des pentoses phosphate ne sont pas différentiellement exprimés dans ces conditions ce qui suggère qu'ils sont constitutivement exprimés (Table 10).

Par ailleurs, les gènes possédant un site *cre* ou sensible à une mutation de *ccpA* ou de *ptsH* (cf Table 3) sont indiqués dans les Tables 9 et 10, ce qui permet de remarquer l'importance de la répression/activation catabolique médiée directement ou indirectement par CcpA en présence de glucose chez *Bacillus subtilis* (Moreno *et al.*, 2001; Yoshida *et al.*, 2001; Blencke *et al.*, 2003 sous presse).

Evidemment, il nous faudra répéter notre expérience de manière satisfaisante et tenir compte de la reproductibilité des résultats. Un certain nombre de gènes différentiellement exprimés de manière « artéfactuelle » dans l'expérience H3, qui ne sont pas retrouvés dans les listes de nos confrères, pourraient être écartés. Ainsi, nous pourrons conclure sur les gènes (inconnus surtout) à retenir comme candidats et dont la régulation pourra être vérifiée expérimentalement par Northern Blot et/ou par mesure de l'activité d'une fusion rapportrice.

Une approche statistique de l'analyse de nos données de transcriptome est aussi envisagée lorsqu'il s'agira de considérer un plus grand nombre d'expériences et de conditions. Malheureusement, notre échantillon est pour l'heure trop petit pour réaliser des tests plus fiables que la simple prise en considération de la reproductibilité. Cependant, à l'aide d'ArrayStat (Imaging Research), nous avons statistiquement analysé les données portant sur la comparaison de la souche sauvage et du mutant *cggR*, décrite dans l'article n°1. Cette analyse permet d'aboutir aux mêmes résultats que ceux obtenus en ne considérant que la reproductibilité, à savoir que les seuls gènes significativement différentiellement exprimés sont *gapA* et *pgk*, qui codent les deux premières enzymes de l'opéron dont l'expression est régulée par CggR.

Discussion générale et perspectives

Dans ce travail, nous avons étudié certains aspects du métabolisme central du carbone par deux approches. Nous nous sommes attachés à analyser le mécanisme de régulation de l'expression de l'opéron *gapA*, dont CggR est l'acteur principal, et à étudier le rôle physiologique et la régulation transcriptionnelle du gène *ywkA*, l'un des gènes codant une enzyme malique chez *Bacillus subtilis*. En parallèle, nous avons commencé l'étude systématique de l'expression des gènes en réponse à diverses sources de carbone par une approche globale d'analyse transcriptomique.

CggR : le prototype d'une sous-famille des régulateurs de type SorC/DeoR.

La première partie de cette thèse a consisté à comprendre comment CggR contrôle la transcription de l'opéron *gapA* en fonction du type de métabolisme central de la cellule, glycolytique ou néoglucogénique. A travers cette étude, nous avons mis en évidence de nouveaux aspects de la famille de régulateurs SorC/DeoR. Dans un premier temps, nous avons identifié la séquence opératrice de CggR, située entre le promoteur et l'ATG de *cggR*, par des approches génétiques et biochimiques. Les résultats montrent qu'il s'agit d'une cible non canonique et inhabituellement longue pour un régulateur transcriptionnel bactérien puisqu'elle est composée de deux séquences répétées directes s'étendant sur près de 40 nucléotides (**CGGGAC**N_6**TGTC**N_4**CGGGAC**N_6**TGTC**). Mais l'agencement de ces séquences et l'étude de la structure secondaire de CggR ne nous permettent pas de prédire plus précisément comment ces motifs interagissent avec les éléments protéiques de reconnaissance de l'ADN. Beaucoup de questions restent donc ouvertes quant aux déterminants de l'interaction CggR/ADN.

Un premier modèle, encore hypothétique, du mécanisme d'interaction entre CggR, l'ADN et le FBP.

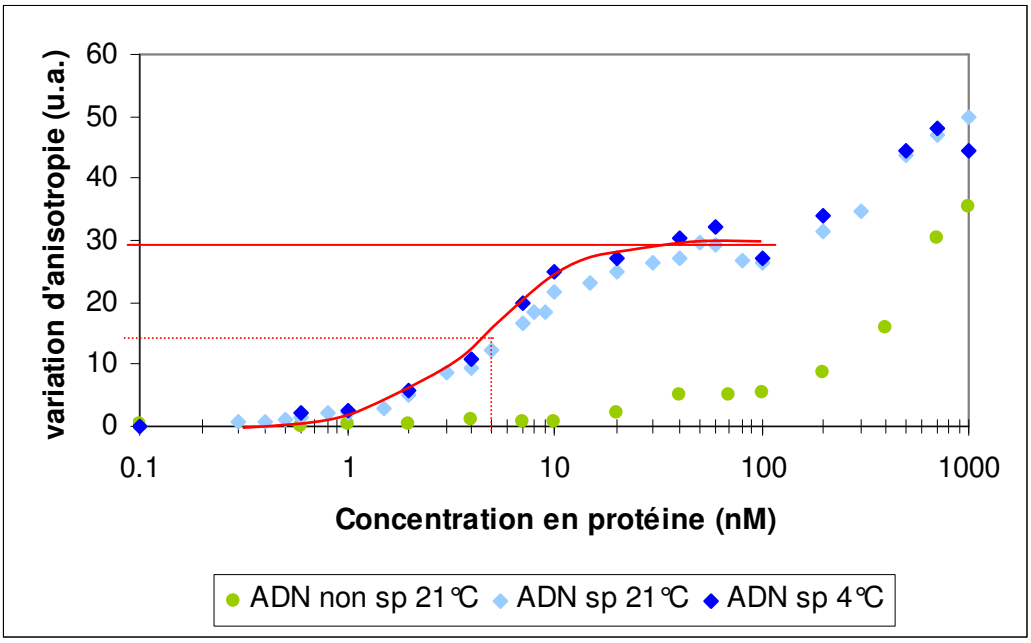
Les premiers résultats de nos expériences de retard de migration électrophorétique suggèrent que plusieurs types de complexes ADN/protéine de poids moléculaires différents peuvent se former en fonction de la concentration de CggR (Figure 2A, Article n°1). Lorsqu'il

se trouve en très faible concentration, CggR se lie à l'ADN pour former un petit complexe. Puis, lorsque sa concentration augmente, CggR se fixerait d'une manière coopérative à l'ADN pour former un gros complexe. Récemment, certains résultats obtenus par Nathalie Declerck permettent de soutenir ce modèle coopératif de fixation. En effet, des expériences d'anisotropie de fluorescence ont été réalisées avec CggR et un ADN cible de 44 pb portant un fluorophore en 5' de la séquence. Les résultats montrent que la courbe de fixation de CggR à son ADN cible suit une forme complexe caractéristique d'un modèle coopératif de fixation (Declerck, communication personnelle) (Figure 55). En outre, le K_d d'environ 5 nM, estimé à partir de la première partie de la courbe (Figure 55), est en accord avec celui de 7 nM estimé dans nos expériences de retard de migration électrophorétique (Figure 2, Article n°1). La cinétique de décrochage de l'ADN a aussi été mesurée par cette technique. Là encore, les résultats sont conformes à ceux observés par résonance plasmonique de surface (Figure 44), à savoir une cinétique de dissociation très lente du complexe déjà formé (Declerck, communication personnelle).

Quelles sont les formes oligomériques de CggR qui peuvent interagir avec l'ADN et quelle est la nature du complexe CggR/ADN qui constitue le « road-block » de transcription ? Au laboratoire, nos expériences suggèrent que CggR peut exister en solution sous les formes monomérique, dimérique, trimérique et au moins une autre d'un poids moléculaire encore plus haut (Figure 43). Des expériences encore préliminaires de chromatographie d'exclusion analytique, effectuées à Montpellier, permettent d'observer des formes trimérique et tétramérique de CggR et un multimère d'encore plus haut poids moléculaire. Toutefois, il n'est pas exclu que CggR puisse avoir une migration aberrante sur la colonne utilisée lors de ces expériences (Declerck, communication personnelle). L'ensemble de ces résultats préliminaires suggèrent donc que CggR pourrait exister en tant que monomère, dimère, trimère, tétramère et un autre oligomère de plus haut poids moléculaire. Bien qu'ils ne permettent pas de prédire quelle est la nature de l'oligomère de CggR qui est impliqué dans le complexe qui bloque efficacement la transcription *in vivo*, ces résultats suggèrent un lien entre les différents types de complexes CggR/ADN visibles dans les expériences de gel retard (Figure 2, Article n°1) et les différentes formes oligomériques existantes de CggR.

Dans un second temps, nous avons montré *in vitro*, par des expériences d'empreinte à la DNase I (Figure 5, Article n°1) et par Biacore (Figure 44 et 45), que le FBP est un effecteur de CggR. Les résultats d'une première série d'expériences de retard de migration

```
5' Fluo-TGACGGGACGTTTTTTGTCATAGCGGGACATATAATGTCCAGCAA 3 ’
      3'ACTGCCCTGCAAAAAACAGTATCGCCCTGTATATTACAGGTCGTT 5'
```



Modèle coopératif, Kd 3-8 nM

Figure 55. Etude de l’interaction CggR/ADN par mesure de variation d’anisotropie de fluorescence. L’interaction est mesurée par la variation d’anisotropie de l’ADN flurorescent en fonction de la concentration de CggR. CggR interagit avec l’ADN selon un modèle coopératif (courbes en bleu, sp : spécifique). Entre 0 et 10 nM, la courbe monte rapidement, s’ensuit un plateau entre 10 et 100 nM, puis la variation d’anisotropie réaugmente au delà de 100 nM. Le Kd estimé à partir de la première partie de la courbe est d’environ 5 nM (courbe en rouge).

électrophorétique avaient aussi suggéré que le FBP augmentait légèrement la quantité de CggR nécessaire pour retarder 50 % de la sonde radioactive. Mais une seule bande était clairement visible sur les gels. Une série d'analyses complémentaires par retard de migration électrophorétique, réalisée dans des conditions différentes, a conduit à des résultats différents. En effet, ceux-ci montrent que la présence de FBP modifie clairement le profil de formation des différents types de complexes CggR/ADN jusqu'alors visibles dans les expériences sans FBP (Figure 46). Le FBP ne semble pas modifier la quantité de CggR requise pour retarder 50 % de l'ADN cible mais stabilise la formation d'une bande « rapide » jusqu'à une concentration de CggR au moins deux fois plus forte qu'en l'absence de FBP. De plus, il semble qu'il existe au moins trois types de complexes différents (Figure 46). Ces différentes formes de complexes pourraient être mises en relation avec les différents degrés d'oligomérisation possibles de CggR. En parallèle, les expériences d'anisotropie de fluorescence effectuées par Nathalie ont conduit à des résultats complémentaires. D'une part, ils confirment les résultats des expériences d'empreinte à la DNase I et de résonance plasmonique de surface : parmi tous les métabolites testés, seul le FBP perturbe nettement la formation d'un complexe CggR/ADN (Figure 56). D'autre part, ils montrent que la présence de FBP fait disparaître le phénomène de coopérativité de fixation de CggR sur son opérateur (Figure 56). L'ensemble de ces résultats suggèrent donc que la présence de FBP n'empêche pas l'interaction entre CggR et l'ADN. En revanche, le FBP favoriserait vraisemblablement la formation d'un type de complexe caractérisé par un ratio CggR/ADN cible faible au détriment de la formation d'un complexe à haut ratio CggR/ADN cible qui serait le seul capable de bloquer la transcription.

Ainsi, l'ensemble de ces résultats permet d'établir un modèle hypothétique du mécanisme d'interaction entre CggR, son opérateur ADN et le FBP (Figure 57). En présence de glucose dans le milieu de croissance, l'activité glycolytique conduit à l'augmentation de la concentration de FBP à l'intérieur de la cellule. CggR n'est alors capable de former qu'un complexe de faible poids moléculaire avec l'ADN. Ce complexe pourrait n'interagir qu'avec une partie de l'opérateur requis pour une régulation efficace de l'expression de l'opéron *gapA*. Le complexe de transcription pourrait déplacer cet édifice peu stable et permettre l'expression de l'opéron *gapA*. A l'opposé, en absence d'une source de carbone glycolytique, la concentration de FBP est faible. Dans ces conditions, CggR n'est pas inhibé par le FBP et « s'oligomériserait » sur sa cible de manière coopérative pour former un complexe de plus haut poids moléculaire, capable de bloquer la machinerie de transcription.

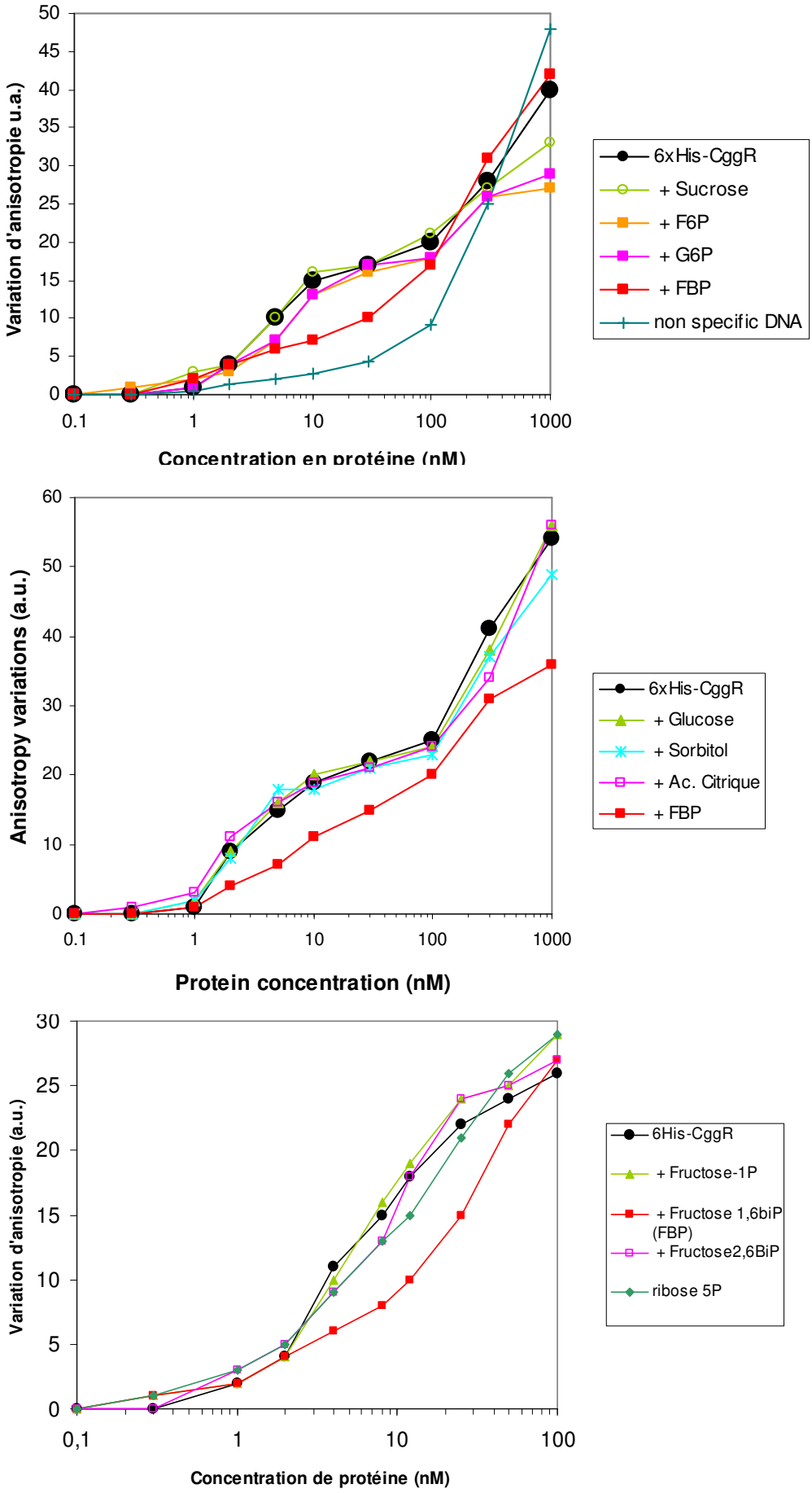


Figure 56. Test de l'effet du FBP et d'autres métabolites ou sucres sur l'interaction ADN/CggR par mesure de variation d'anisotropie de fluorescence.

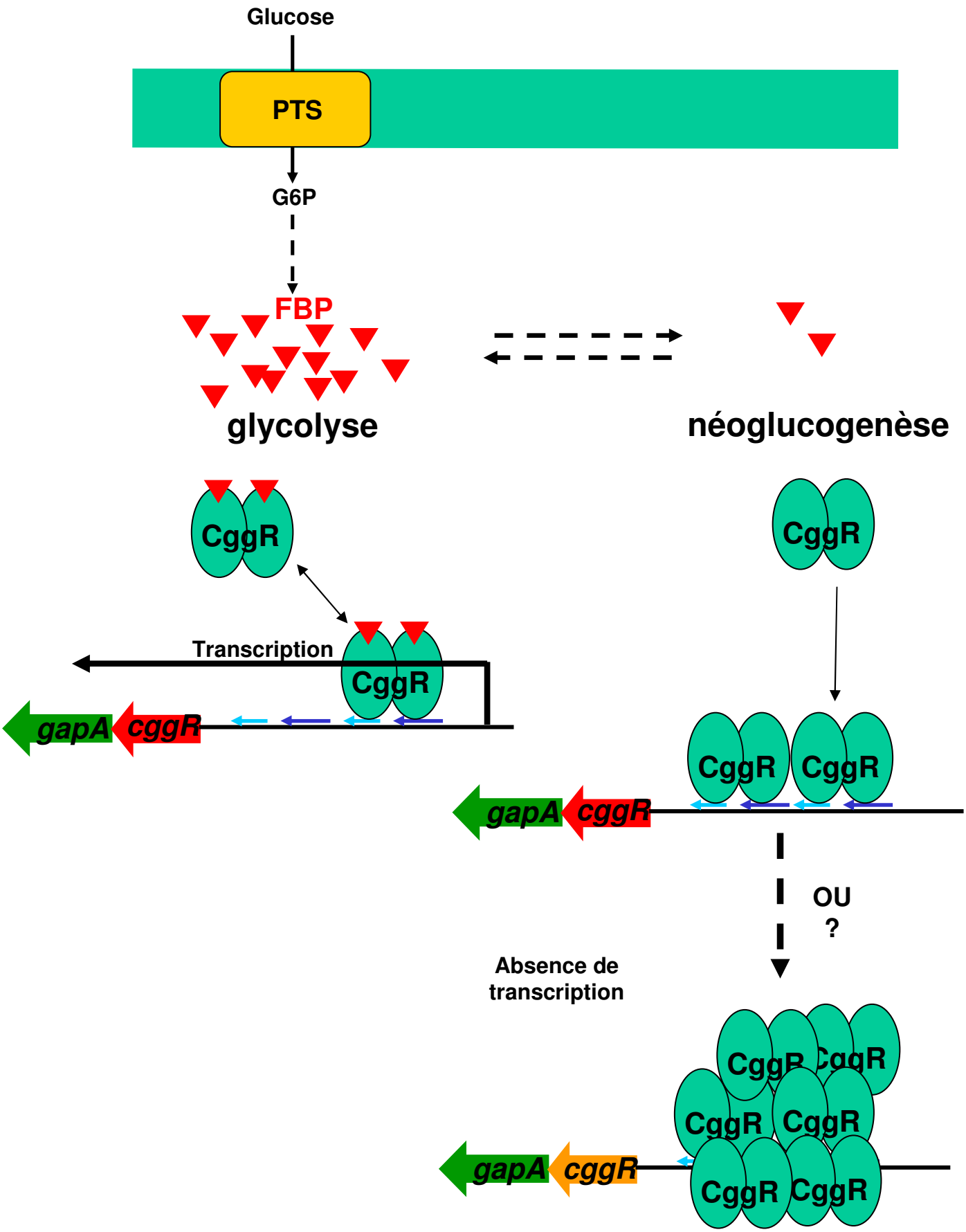


Figure 57. Modèle hypothétique de l'interaction entre CggR et l'ADN en fonction de la concentration en FBP. Les petites flèches en bleu foncé et bleu clair représentent la séquence cible de CggR.

Vers une étude biochimique et structurale de l'interaction entre CggR, son opérateur et le FBP.

Pour l'heure, aucun membre de la famille de régulateurs transcriptionnels SorC/DeoR n'a encore fait l'objet d'une caractérisation poussée sur le plan biochimique et structural. Motivés par ce fait, nous allons mettre en oeuvre une série d'expériences génétiques, biochimiques et biophysiques visant *in fine* à résoudre la structure de CggR et à déterminer les éléments submoléculaires de l'interaction entre CggR, son opérateur nucléique et le FBP.

Nous allons tenter de comprendre comment CggR reconnaît sa séquence cible et comment les différents complexes CggR/ADN se forment en regard des différents éléments de cette cible. Dans un premier temps, des expériences seront menées pour confirmer nos résultats préliminaires quant aux formes oligomériques possibles de CggR. Puis, notre contribution pourra être le test, par retard de migration électrophorétique ou par Biacore, des différentes hypothèses concernant les types d'interaction protéine/ADN envisagés. En particulier, nous allons tester cette hypothèse : l'opérateur que nous avons identifié (**CGGGAC**N_6**TGTC**N_4**CGGGAC**N_6**TGTC**) contient deux petits palindromes identiques, dont les motifs sont régulièrement espacés d'un tour d'hélice (**GAC**N_7**GTC**N_7**GAC**N_7**GTC**) ; ces quatre triplets de nucléotides pourraient jouer un rôle central dans la spécificité de reconnaissance et l'oligomérisation de CggR d'un côté de la molécule d'ADN. L'effet de la mutation de chacun des nucléotides inclus dans ces motifs sera testé. Par ailleurs, dans l'hypothèse ou un premier oligomère de CggR se fixerait d'abord sur une moitié de l'opérateur (**CGGGAC**N_6**TGTC**), on peut se demander quels sont les éléments importants qui permettent de discriminer les deux parties de l'opérateur. Là encore, notre contribution sera de tester le rôle des séquences entourant ces motifs. L'effet de différentes mutations sur l'interaction CggR/ADN sera testé, par exemple en substituant totalement une moitié de l'opérateur par la séquence complète de l'autre partie.

Nous allons aussi focaliser nos efforts sur l'étude de l'interaction entre CggR et le FBP : quels sont les déterminants de l'interaction entre le FBP et CggR et quels changements de conformation provoqués par cette interaction permettent de moduler la capacité de CggR à former un complexe capable de bloquer la transcription ? Pour répondre à ces questions, deux axes sont envisagés. D'une part, les modifications structurales éventuellement entraînées par l'interaction CggR/FBP seront analysées par cinétique de protéolyse ménagée par la trypsine et par spectre de dichroïsme circulaire, en présence de CggR seul, en présence de CggR et

d'ADN, en présence de CggR et de FBP ou en présence de CggR, d'ADN et de FBP. D'autre part, les éléments submoléculaires de l'interaction CggR/FBP seront analysés par des moyens bioinformatiques et génétiques.

Le domaine régulateur C-terminal de CggR présente des homologies de séquences avec le domaine C-terminal des régulateurs de la famille SorC (Figure 58). Cependant, la recherche dans les bases de données PFAM et SMART n'a pas permis de détecter de motif structural connu. Une analyse par le Metaserveur du CBS permet de proposer plusieurs analogues structuraux de CggR, dont la glucosamine-6-phosphate déaminase d'*Escherichia coli* (NagB, entrée PDB : 1fqo) co-cristallisée en présence du substrat fructose-6-phosphate (F6P). CggR et NagB partagent peu d'homologies de séquence (12 % d'identité sur 250 acides aminés) mais les résidus strictement conservés dans la famille des glucosamines déaminases et dans la famille des régulateurs SorC sont tous situés dans le site de fixation du substrat. La représentation du site actif de NagB entourant le F6P, superposé avec le site potentiel d'interaction entre le FBP et CggR (modélisé à partir de 1fqo) (Figure 59) permet de remarquer que la Lysine en position 310 (de CggR), qui est strictement conservée, pourrait interagir avec le groupement phosphate en position 6 du FBP (et du F6P en ce qui concerne NagB). L'Arginine en position 175 (de CggR), qui n'est conservée que dans les CggR (Figure 58), serait le site d'interaction avec le groupement phosphate en position 1 du FBP. Par ailleurs, dans cette modélisation, le F6P est linéaire. CggR pourrait fixer une forme linéaire du FBP, qui est très minoritaire dans la cellule. Au laboratoire, nous envisageons de tester toutes ces hypothèses par des approches génétiques et biochimiques en collaboration avec Nathalie Declerck. Dans un premier temps, nous essaierons de savoir quels sont les effets de la mutation des résidus potentiels d'interaction avec le FBP. La Lys310 et/ou l'Arg175 seraient substituées par mutagenèse dirigée par d'autres acides aminés n'interagissant pas avec le groupement phosphate. Puis, pour déterminer *in vivo* la sensibilité au FBP de ces protéines, leur capacité à complémenter un mutant *cggR* sera testée. Il sera aussi possible d'observer la capacité d'interaction et la spécificité de ces protéines mutées vis-à-vis des métabolites glycolytiques, en particulier le FBP et le F6P, *in vitro* après purification et test d'interaction par mesure de variation de résonance plasmonique de surface ou de variation d'anisotropie de fluorescence.

La détermination des résidus de CggR impliqués dans l'interaction avec CggR sera aussi approché de façon génétique : nous rechercherons des mutants « gain de fonction »,

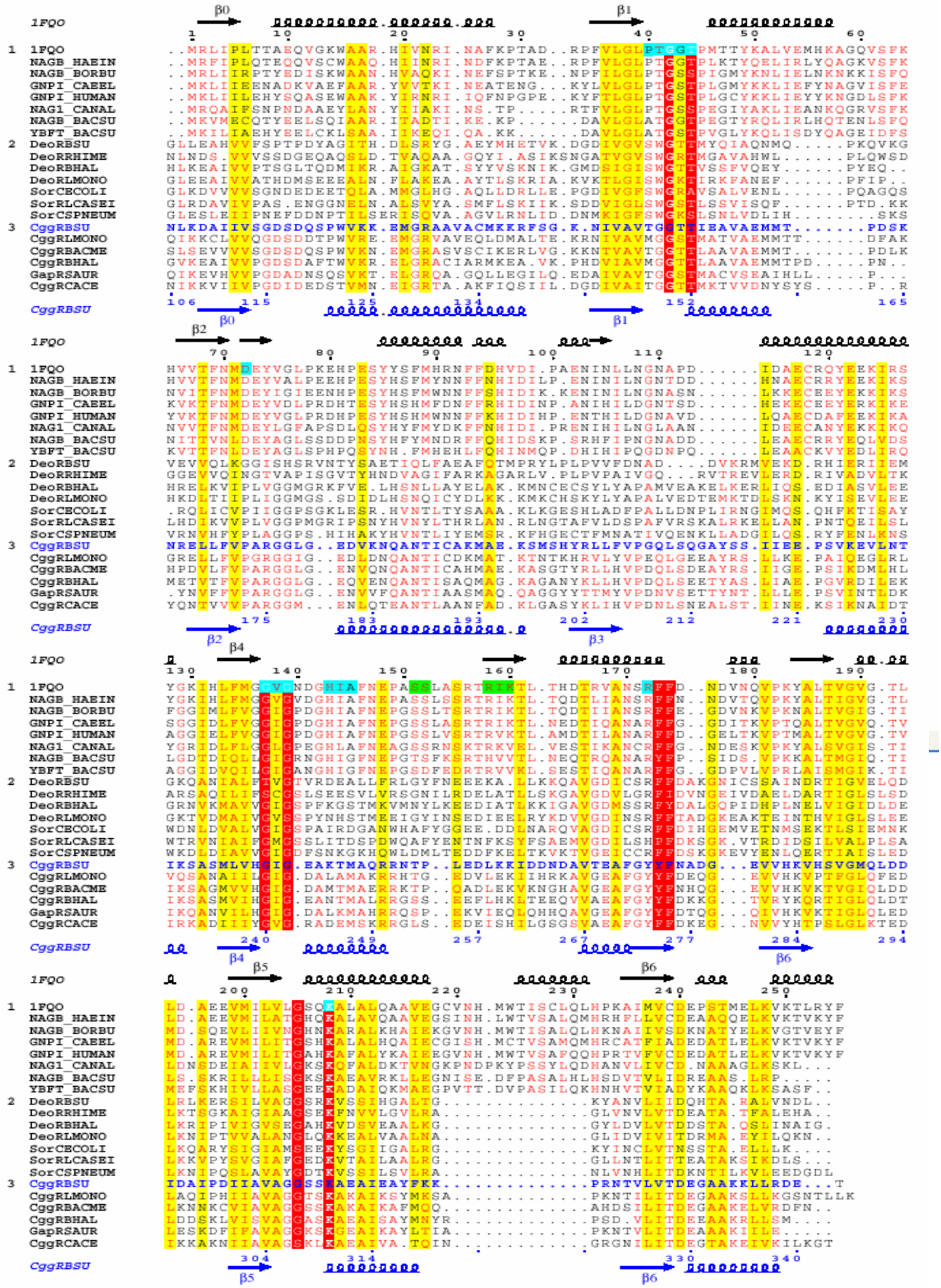


Figure 58. Alignement du domaine régulateur de CggR de *Bacillus subtilis* avec les domaines homologues chez *Listeria monocytogenes* (LMONO), *Bacillus megaterium* (BACME), *Bacillus halodurans* (BHAL), *Staphylococcus aureus* (SAUR) et *Clostridium acetobutilinum* (CACE), les domaines régulateurs de DeoR ou SorC de *Bacillus subtilis* (BACSU), *Rhizobium meliloti* (RHIME), *Bacillus halodurans*, *Listeria monocytogenes*, *Escherichia coli* (ECOLI), *Lactobacillus casei* (LCASEI) et *Streptococcus pneumoniae* (SPNEUM), et les séquences des glucosamine-6-phosphate déaminases d'*Escherichia coli, Haemophilus influenzae* (HAEIN), *Borrelia burgdorferi* (BORBU),*Caenorhabditis elegans* (CAEEL), *Homo sapiens* (HUMAN), *Candida albicans* (CANAL) *et Bacillus subtilis*. Les résidus strictement conservés sont en blan sur fond rouge. Ils sont tous situés au niveau du site actif de NagB.

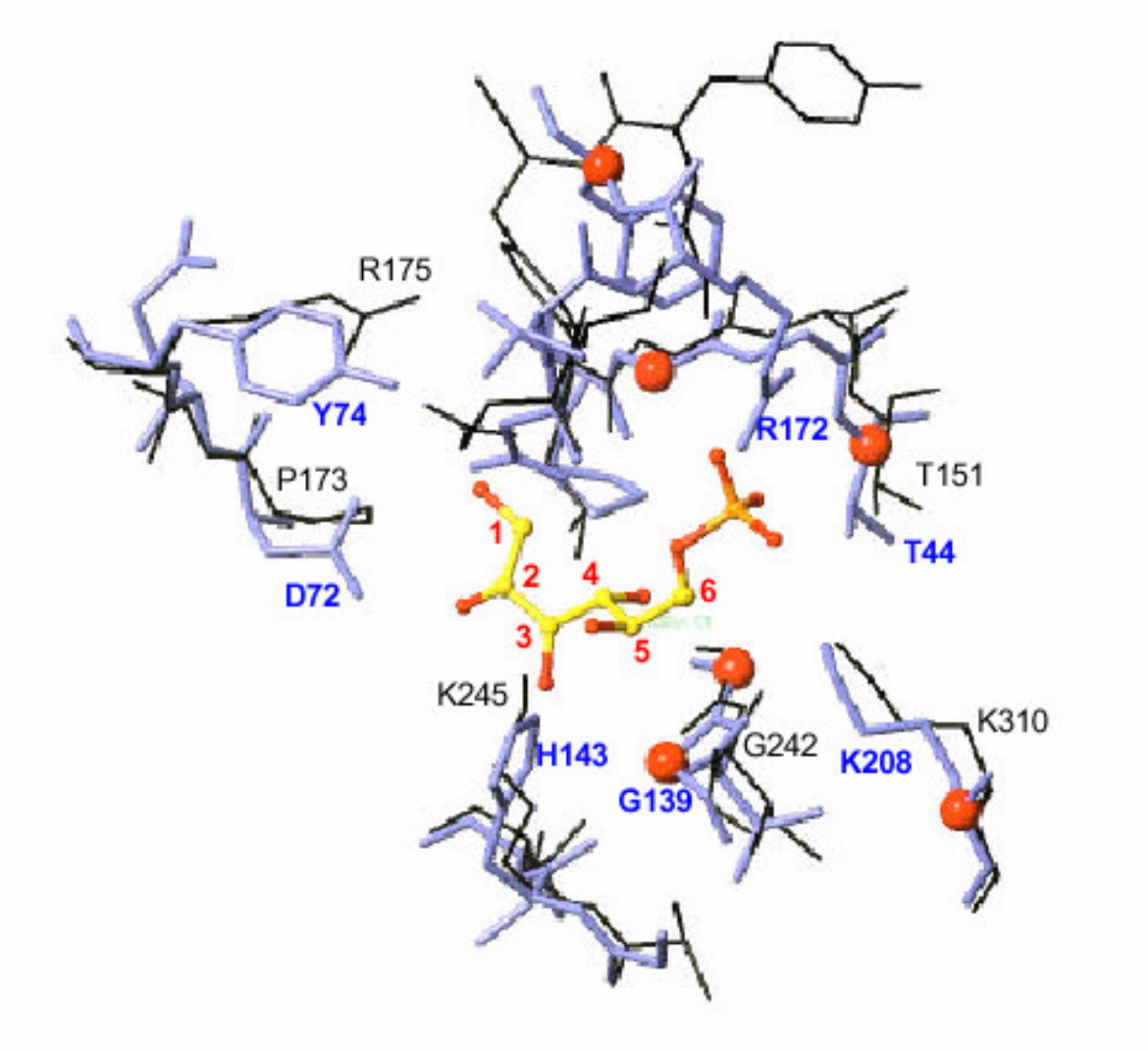


Figure 59. Modélisation du site actif de la glucosamine déaminase d'*Escherichia coli* (en bleu clair) montrant le F6P, superposé au site actif potentiel de CggR de *Bacillus subtilis* (en noir) modélisé à partir de 1fqo. Les sphères rouges indiquent les résidus très bien conservés. Lys310 est très conservée et interagirait avec le groupement phosphate en postiion 6 du F6P. Tyr74 de NagB est remplacée par Arg175 dans CggR uniquement qui pourrait interagir avec le groupement phosphate en position 1 du F6P.

c'est-à-dire des variants de CggR qui ne sont plus capables de répondre au signal FBP. Pour cela, il faudra mettre au point un crible satisfaisant qui permettra de différencier les mutants *cggR* ne répondant plus au signal FBP des autres mutants. La stratégie expérimentale envisagée est similaire à celle employée lors du test de complémentation d'un mutant *cggR* par *his*$_6$*-cggR* (Article n°1). Dans un premier temps, une banque de fragments PCR (comprenant la région régulatrice de l'opéron *gapA* et l'ORF *cggR*) mutés au hasard sera générée. Ces fragments seront insérés dans un vecteur permettant l'intégration au locus *amyE* du chromosome. Puis, cette banque plasmidique servira à obtenir une banque de transformants à partir d'une souche délétée de *cggR* et portant une fusion rapportrice de l'activité du promoteur de l'opéron *gapA*. Ces transformants seront criblés sur milieu minimum supplémenté en glucose et en X-Gal. Les clones blancs (ou bleus clair) seront retenus pour une analyse complémentaire.

Et le Fructose-1,6-Biphosphate *in vivo* ?

Sur le plan physiologique, le fait que le FBP est l'effecteur de CggR est cohérent avec le fait que l'expression de l'opéron *gapA* est stimulée lors de la glycolyse. Toutefois, deux résultats expérimentaux rendent le raisonnement plus compliqué. D'une part, une contradiction persiste encore dans notre compréhension de la modulation de l'activité de CggR *in vivo*. Lors de nos expériences de cinétique d'induction de l'expression de l'opéron *gapA*, nous avons montré que l'ajout de glycérol provoquait une plus forte induction de l'opéron *gapA* que le glucose (Figure 4, Article n°1), ce qui suggère que le métabolisme du glycérol conduit à une plus forte concentration de FBP intracellulaire que celui du glucose. Or Mijakovic et ses collaborateurs (2002) ont montré que la concentration du FBP varie de 6,2 mM à 14 mM suivant que *Bacillus subtilis* est cultivé en présence de glycérol ou de glucose. Certes, comme indiqué dans l'article, les conditions expérimentales utilisées par ces auteurs diffèrent des notres, notamment parce que nous procédons à un basculement brutal d'un métabolisme néoglucogénique à un métabolisme glycolytique alors que nos collègues travaillent dans des conditions de flux glycolytique constant. D'autre part, Ludwig et ses collaborateurs (2001) ont montré que l'induction de l'expression de l'opéron *gapA* était au moins deux fois plus forte lorsque les cellules sont cultivées en présence de glucose plus acides aminés que de glucose seul. Ces résultats suggèrent que CggR pourrait répondre à deux signaux distincts, l'un provenant de la glycolyse (le FBP !) et l'autre provenant du

métabolisme des acides aminés. Notre hypothèse est que le FBP est le signal perçu par CggR mais qu'il se trouve peut-être à plus forte concentration en présence d'acides aminés. En effet, dans ces conditions, le cycle de Krebs est peu actif ; le flux glycolytique pourrait alors être un peu plus rapide pour générer l'énergie qui n'est plus produite par le cycle de Krebs, ce qui aurait pour effet d'augmenter la concentration de FBP. Afin d'éclaircir les deux points précédemment évoqués, il sera intéressant de comparer, dans nos conditions expérimentales de croissance en conditions néoglucogéniques et d'induction de la glycolyse, la concentration de FBP après ajout de glycérol, de glucose ou de glucose plus acides aminés en utilisant la méthode mise au point par Mijakovic et ses collaborateurs (2002).

Intégration de la régulation de l'expression de l'opéron *gapA* médiée par CggR : une régulation fine de l'expression des gènes de la partie centrale de la glycolyse.

La démonstration du rôle du FBP dans la modulation directe de l'activité de CggR (Article n°1) est à mettre en relation avec celle de la démonstration du rôle indirect de CcpA dans la stimulation de l'expression de l'opéron *gapA* (Tobisch *et al.*, 1999b). Au laboratoire, des analyses génétiques ont montré que l'effet de CcpA était indirect et médié par CggR (Fillinger *et al.*, 2000). Enfin, plus récemment, l'équipe de Stülke a démontré pourquoi l'expression de l'opéron *gapA* n'était plus stimulée par le glucose dans un mutant *ccpA* (Ludwig *et al.*, 2002b). En effet, leurs analyses montrent que dans un tel mutant, HPr est séquestré de manière quasi totale sous la forme HPr-Ser46-P à cause soit d'une activité kinase très forte soit d'une activité phosphatase faible de l'enzyme bifonctionnelle HPrK/P. HPr n'est donc plus disponible pour le transfert du groupement phosphate du PEP vers le sucre transporté par une enzyme II du PTS. CcpA est donc nécessaire au transport des sucres transportés par le PTS, dont le glucose fait partie. Ces résultats sont cohérents avec le fait que la présence de glycérol induit toujours l'expression de l'opéron *gapA* dans un mutant *ccpA* (Fillinger *et al.*, 2000). En effet, le glycérol est transporté par un facilitateur indépendant du PTS. Ainsi, le FBP est le signal commun intégré par CcpA et CggR. Le modèle de la Figure 57 peut donc être complété pour prendre en compte le rôle indirect de CcpA (Figure 60).

Par ailleurs, il existe deux autres niveaux de régulation de l'expression de l'opéron *gapA* qui permettent de moduler l'expression relative des gènes de l'opéron. Le premier est

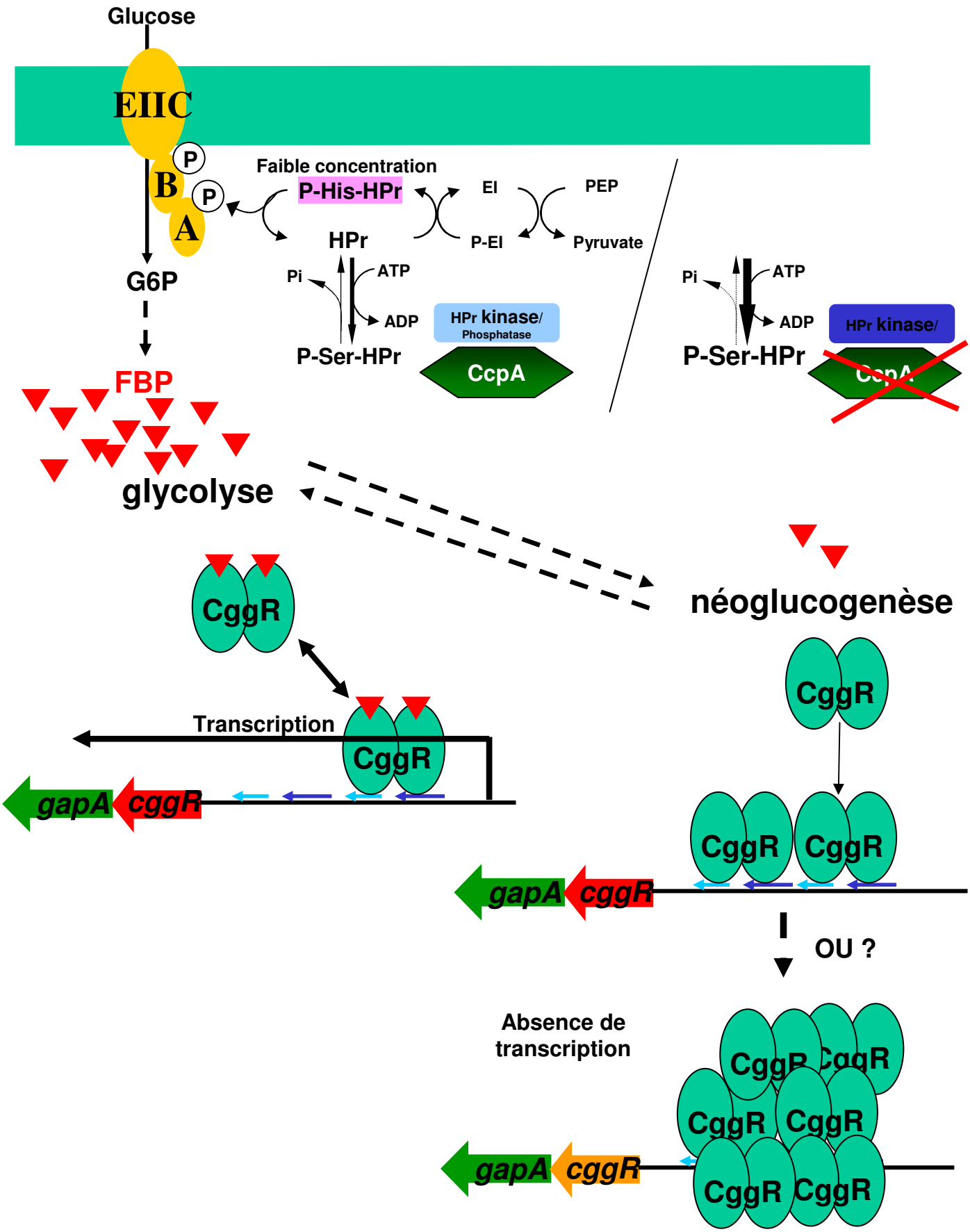


Figure 60. Modèle hypothétique de la régulation de l'expression de l'opéron *gapA* prenant en compte l'interaction entre CggR et l'ADN en fonction de la concentration en FBP et le rôle de CcpA dans le transport des sucres PTS et la formation de FBP. Les petites flèches en bleu foncé et bleu clair représentent la séquence cible de CggR.

une régulation au niveau post-transcriptionnelle. L'équipe de Stülke a mis en évidence un événement de maturation consistant en un clivage de l'ARNm au niveau d'une séquence riche en AU entourée de deux structures secondaires tige-boucle, dans la partie extrême 3'-terminale du transcrit *cggR*. Ce clivage entraîne une instabilité du transcrit *cggR* alors que le transcrit comprenant *gapA* est assez stable (Ludwig *et al.*, 2001). De plus, il a été montré que la synthèse de l'enzyme GapA est environ 100 fois supérieure à celle du régulateur CggR dans des conditions glycolytiques. Il semble que cette différence soit due, au moins partiellement, à ce processus de maturation (Meinken *et al.*, 2003 sous presse). Cette régulation permettrait de moduler la quantité de CggR, qui n'est *a priori* nécessaire qu'en très faible quantité, par rapport à celle des enzymes de la glycolyse qui doivent être en concentrations plus importantes. Le deuxième autre niveau de régulation de l'opéron *gapA* est autorisé par un deuxième promoteur constitutif situé en amont de la séquence codante du gène *pgk*. Ce promoteur permet une expression basale des enzymes de l'opéron à l'exception de *gapA*, puisque seuls ces premiers servent aussi à la néoglucogenèse. Le modèle de la Figure 59 peut donc (encore) être complété pour tenir compte de ces autres niveaux de régulation (Figure 61).

Ainsi, l'existence de plusieurs mécanismes de contrôle de la transcription suggère l'expression de l'opéron *gapA* nécessite d'être précisément régulée. Pourtant, la comparaison transcriptomique de la souche sauvage et d'un mutant *cggR* en conditions néoglucogéniques indique que la dérégulation de l'expression des gènes de la partie centrale de la glycolyse n'entraîne pas d'effets indirects suffisants pour être détectés par cette technique. Le sens physiologique d'une régulation si complexe reste donc ouvert !

Qu'en est-il chez les bactéries proches de *Bacillus subtilis* ?

Enfin, nous avons aussi montré que CggR de *Bacillus subtilis* et très probablement les régulateurs de la sous-famille qu'il définit et qui sont retrouvés chez presque toutes les bactéries à Gram positif et faible pourcentage en GC (Figure S1, Article n°1), sont spécifiquement dédiés au contrôle des gènes de la seconde partie de la glycolyse et que le FBP est probablement l'effecteur commun à ces régulateurs (Article n°1). Or, il faut se rappeler que *Bacillus subtilis* et seulement quelques autres bactéries présentent la particularité

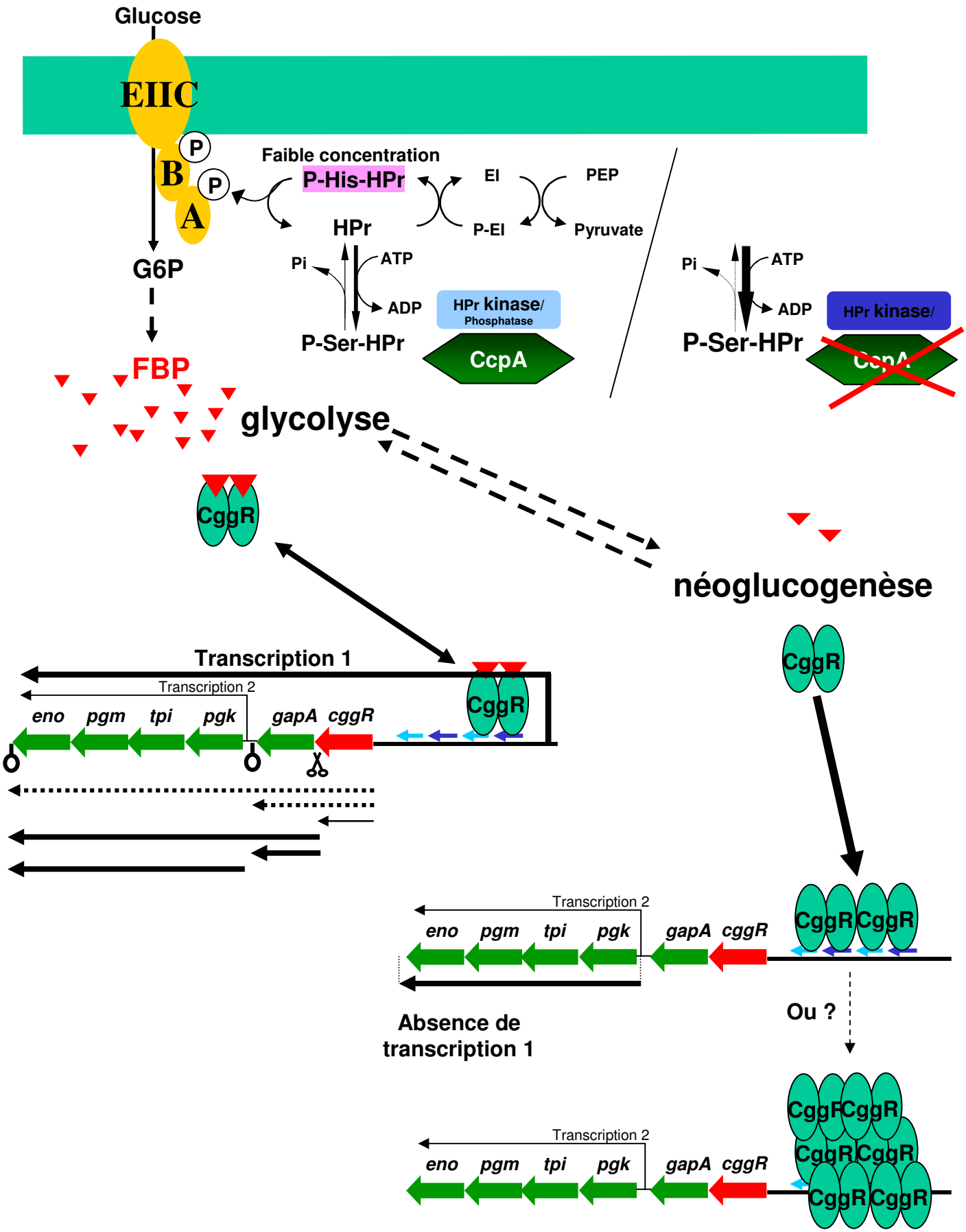


Figure 61. Modèle hypothétique de la régulation de l'expression de l'opéron *gapA* intégrant les différents niveaux de régulation décrits. Les petites flèches en bleu foncé et bleu clair représentent la séquence cible de CggR. Transcription 1 correspond à l'activité du promoteur régulé en amont de *cggR* et transcription 2 correspond à l'activité du promoteur constitutif en amont de *pgk*. Le site de coupure du transcrit *cggR* est aussi figuré par des ciseaux. Les terminateurs sont aussi figurés. Concernant la transcription de l'opéron *gapA*, les flèches en pointillés correspondent aux transcrits natifs du promoteur en amont de *cggR*. Les flèches pleines indiquent les transcrits présents après événement de maturation. L'épaisseur du trait indique l'abondance relative du transcrit dans la cellule.

de posséder deux gènes *gap*, l'un exprimé pour la glycolyse et l'autre exprimé pour la néoglucogenèse (Fillinger *et al.*, 2000). Sur la base des prédictions de séquences et lorsque le génome est disponible, les bactéries suivantes possèdent probablement, comme *Bacillus subtilis*, une deuxième GAPDH NADP-dépendante en plus d'une GAPDH NAD-dépendante : *Bacillus halodurans*, *Oceanobacillus iheyensis*, *Bacillus anthracis*, *Staphylococcus aureus*, *Staphylococcus epidermis* pour ce qui est des bactéries à Gram positif et *Neisseria meningitidis*, *Neisseria gonorrhoae* et *Helicobacter pylori* pour ce qui est des bactéries à Gram négatif. Chez ces trois dernières, rien n'est connu de la régulation du gène *gap*. Chez les bactéries à Gram positif mentionnées ci-dessus, la régulation du gène *gapA* par un répresseur de type CggR lors de la néoglucogenèse est cohérente avec son rôle physiologique et compatible avec le fait qu'une autre GAPDH fonctionne lors de la néoglucogenèse.

Chez les autres bactéries qui possèdent un homologue fonctionnel de CggR (régulant l'expression d'un gène *gap* et très souvent des autres gènes de la seconde partie de la glycolyse), le gène *gap* lié à *cggR* code très souvent une GAPDH de type GapA de *Bacillus subtilis*. Mais, si ces bactéries ne possèdent pas de gène *gap* codant une GAPDH capable de fonctionner dans le sens néoglucogénique, comment peuvent elles pousser sur des substrats néoglucogéniques alors que la GAPDH dont elles disposent est moins exprimée ? Le problème se pose surtout pour la GAPDH puisque les autres gènes de l'opéron pourraient être exprimés à un niveau suffisant *via* l'activité d'un promoteur constitutif en aval du gène *gap*, comme chez *Bacillus subtilis*. Plusieurs cas de figure sont possibles. Une telle bactérie pourrait tout simplement ne pas être capable de croître en présence de substrats néoglucogéniques comme seule source de carbone et d'énergie. Sinon, il faut aussi se rappeler que, chez *Bacillus subtilis*, la répression de l'expression de l'opéron *gapA* n'est pas totale et permet une synthèse basale de GapA dans des conditions néoglucogéniques (Fillinger *et al.*, 2000). Si en plus la GAPDH exprimée était enzymatiquement un peu moins favorable à la réaction dans le sens glycolytique, cela pourrait permettre à la bactérie de produire une quantité suffisante de G3P à partir de la néoglucogenèse pour la croissance.

Le deuxième volet de cette thèse a porté sur la caractérisation du gène *ywkA*, l'un des quatre gènes paralogues dont le produit ressemble à une enzyme malique. Notre étude a permis de démontrer qu'*ywkA* code bien une enzyme malique et que son expression est spécifiquement et directement activée par le système à deux composants YufL-YufM en réponse à la présence de malate extracellulaire. Toutefois, contrairement à ce que l'on pourrait attendre d'une enzyme malique, *ywkA* ne semble pas être utile à la croissance en présence de malate comme seule source de carbone. Par ailleurs, cette étude nécessitait aussi de s'intéresser à celles des gènes *malS* et *ytsJ*. Certains résultats spécifiques concernant l'expression et le rôle de ces gènes seront discutés.

ywkA, _malS_ et _ytsJ_ codent bien des enzymes maliques mais leurs spécificités de cofacteur diffèrent.

Les enzymes maliques bactériennes présentent de fortes homologies de séquence avec les enzymes malolactiques. Nos tests enzymologiques permettent de vérifier qu'YwkA est bien une enzyme malique et suggèrent que son cofacteur préférentiel est le NAD (Article n°2 et Table 6). De plus, ces tests ont démontré qu'YwkA ne présente pas d'activité enzyme malolactique (Figure 3, Article n°2). Par ailleurs, ces tests ont aussi permis de démontrer que les gènes paralogues *malS* et *ytsJ* codent aussi des enzymes maliques (Table 6). MalS exhibe des caractéristiques enzymatiques très proches de celles d'YwkA. En revanche, YtsJ semble être une enzyme malique NADP-dépendante. Ces résultats suggèrent qu'*ywkA* et *malS* d'un côté et *ytsJ* de l'autre pourraient jouer des rôles spécifiques suivant les conditions physiologiques.

ywkA et _malS_ ne sont pas impliqués dans l'utilisation du malate pour la croissance, contrairement à _ytsJ_.

Nos tests physiologiques avec un simple mutant *ywkA* et les différentes combinaisons de deux ou trois gènes mutés parmi les trois paralogues considérés ne nous ont pas permis de mettre en évidence un défaut significatif de croissance qui soit lié à la mutation d'*ywkA*, cela quelle que soit la source de carbone présente dans le milieu (Figure 2, Article n°2). De même, la délétion de *malS*, quel que soit le contexte génétique testé, n'a pas d'effet sur la croissance

de *Bacillus subtilis*, bien qu'il code une enzyme malique. En revanche, l'interruption d'*ytsJ* (seule ou en combinaison avec les mutations d'*ywkA* et de *malS*) provoque un défaut de croissance significatif uniquement lors de la croissance en présence de malate (Figure 2, Article n°2), de succinate ou de fumarate (Lerondel et Aymerich, non publié). Ces résultats suggèrent qu'*ytsJ* est requis pour la croissance efficace en présence d'acides dicarboxyliques en C_4. Par ailleurs, le flux métabolique passant à travers l'étape enzyme malique a été mesuré dans la souche sauvage et dans les différents mutants d'*ywkA*, de *malS* et d'*ytsJ* dont nous disposons. Les résultats de cette comparaison réalisée en collaboration par l'équipe d'Uwe Sauer sont globalement en accord avec nos observations : seule la mutation d'*ytsJ* entraîne une diminution du flux à travers cette étape dans des cultures en chémostat en présence de glucose comme source de carbone (Zamboni et Sauer, communication personnelle). Cependant, l'expression d'*ywkA* n'est sans doute pas induite dans leurs conditions expérimentales (absence de malate). Par conséquent, il n'est pas possible de tirer des conclusions sur le rôle d'*ywkA* d'après ces résultats.

L'expression d'*ywkA* est induite par le système à deux composants YufL-YufM spécifiquement en présence de malate extracellulaire.

En parallèle, afin de connaître les conditions de croissance dans lesquelles ces gènes sont exprimés et donc potentiellement utilisés, nous avons entrepris l'étude de leur régulation transcriptionnelle. Pour ma part, je me suis focalisé sur l'étude de l'expression d'*ywkA*. Les analyses génétiques (Tables 2 et 3, Article n°2) et transcriptomiques (Figure 1, Article n°2) (Kobayashi *et al.*, 2001) ont permis de démontrer que l'expression d'*ywkA* était induite spécifiquement en réponse à la présence de malate extracellulaire directement *via* le système à deux composants (TCS) YufL-YufM. Par ailleurs, dans le cadre de notre collaboration, nos collègues japonais de l'équipe de Yasutaro Fujita ont montré qu'*ywkA* est cotranscrit avec le gène en aval, *ywkB* qui code une protéine de fonction inconnue (Figure 4, Article n°2). S'agissant de l'expression de *malS*, aucune activité transcriptionnelle des régions régulatrices potentielles n'a pu être détectée (Lerondel et Aymerich, non publié ; cf résultats complémentaires partie 5). Ces résultats sont en contradiction avec ceux tirés de plusieurs analyses transcriptomiques qui suggèrent que *malS* est exprimé de manière significative et constitutive (Kobayashi *et al.*, 2001). Par ailleurs, des expériences qualitatives de RT-PCR ont été réalisées qui montrent l'existence d'un transcrit comprenant la séquence de *malS*. Une

explication possible de cette contradiction serait que les fusions transcriptionnelles construites ne montrent pas d'activité à cause d'une inhibition de la transcription par des structures secondaires générées spécifiquement par des régions régulatrices et le point de fusion choisi avec le gène *lacZ*. Une autre explication serait que les signaux obtenus dans les expériences de transcriptomes et par RT-PCR ne correspondent en fait qu'à une transcription basale très faible de *malS*.

Concernant l'étude d'*ytsJ*, nos résultats montrent qu'il est exprimé de façon forte et constitutive.

Recherche de la fonction physiologique d'*ywkA* par une approche génétique classique.

En ce qui concerne l'étude d'*ywkA*, outre le fait que nous devons vérifier que l'enzyme malique qu'il code est aussi capable de décarboxyler l'oxaloacétate pour pouvoir la classer parmi les enzymes de type EC 1.1.1.38, une grande question demeure quant à son rôle physiologique. En effet, il semble étrange que (d'après l'ensemble de nos résultats) la seule enzyme malique spécifiquement synthétisée en présence de malate ne soit pas utile à son utilisation comme substrat carboné pour la croissance. Pour tenter de lever ce mystère, nous nous sommes d'ores et déjà lancé dans deux voies différentes.

La première approche choisie est de type génétique et vise à compléter la première série d'analyses phénotypiques. Il se peut que l'inactivation d'*ywkA* et/ou de *malS* ne conduise à aucune différence de phénotype parce qu'il existe une deuxième voie d'intégration du malate dans le métabolisme du carbone. En effet, le malate peut être converti en oxaloacétate par la malate déshydrogénase (*mdh*) (Figure 46). L'oxaloacétate pourra alors être converti en PEP par la phosphoénolpyruvate carboxykinase (*pckA*). Dans ces conditions, le phénotype d'un mutant *ytsJ* serait plutôt du à la capacité de l'enzyme qu'il code à produire du NADPH. Cette seconde voie permettrait la complète compensation de la perte d'activité enzyme malique codée par *ywkA* et/ou *malS*. Le très léger défaut de croissance d'un mutant *mdh* en présence d'acides dicarboxyliques (Jin *et al.*, 1996) pourrait être expliqué par la compensation de la perte de l'activité malate déshydrogénase par l'enzyme malique (et l'activité pyruvate carboxylase). Il faudra tester l'effet des mutations d'*ywkA* et/ou de *malS*

dans des conditions où l'activité enzyme malique est absolument nécessaire à la cellule, c'est-à-dire dans un contexte *mdh* mutant ou *mdh* et *ytsJ* mutants. Pour cela, les souches portant les différentes combinaisons de mutations des gènes *mdh*, *ywkA*, *malS* et *ytsJ* sont en cours de construction et leur phénotype sera testé en présence de malate comme seule source de carbone.

Par ailleurs, pour déterminer si *in vivo* YwkA est une enzyme malique fonctionnelle, une étude de flux métabolique dans des cultures en chémostat en présence de glucose et en présence ou non de malate pourrait permettre de déterminer si l'expression d'*ywkA* conduit à une augmentation du flux de carbone allant du malate vers le pyruvate.

Dans notre seconde approche, nous tentons de déterminer à quelle grande fonction physiologique *ywkA* pourrait participer. Pour cela, plusieurs pistes ont été envisagées.

Analyse du régulon YufL-YufM.

Nous utilisons tout d'abord les différents résultats disponibles concernant l'étude du régulon YufL-YufM. Nous allons essayer d'analyser et de dégager des catégories fonctionnelles régulées par ce TCS puis déterminer si *ywkA* pourrait y participer.

Le premier résultat qui nous intéresse a été obtenu par l'équipe de Naotake Ogasawara au Japon. D'ailleurs, leurs résultats font l'objet d'un article qui a été soumis pour publication conjointe avec notre article concernant *ywkA* (Tanaka *et al.*, 2003). Nos collègues s'intéressent en particulier au fonctionnement du système YufL-YufM. Les résultats de leurs analyses génétiques et biochimiques montrent que l'expression des gènes *maeN* (*yufR*) et *yflS* est aussi induite en présence de malate par YufL-YufM. Chez *Bacillus subtilis*, *maeN* et *yufL-yufM* se trouvent dans la même région chromosomique mais séparés par les gènes *yufNOPQ* (Figure 50). De façon notable, chez d'autres bactéries à Gram positif, on retrouve ces gènes physiquement liés, entre eux, et à un gène codant une enzyme malique homologue à YtsJ mais dont il n'est pas possible de prédire la spécificité de cofacteur (Figure 50). De plus, il a été montré que *maeN* code un symporteur Na^+/malate dont l'expression est essentielle pour l'utilisation efficace du malate comme seule source de carbone (Tanaka *et al.*, 2003). Par

ailleurs, *yflS* code probablement un translocateur malate/oxoglutarate qui serait aussi impliqué dans le transport du malate bien qu'un mutant *yflS* ne présente pas de phénotype notable. L'expression coordonnée par YufL-YufM d'*ywkA*, de *maeN* et d'*yflS* en présence de malate est donc cohérente avec l'idée qu'YwkA sert à utiliser le malate transporté par MaeN et YflS. Alors, nous ne comprenons pas pourquoi nos résultats indiquent qu'YwkA n'est pas utile dans ces conditions de croissance.

L'équipe d'Ogasawara a aussi déterminé la séquence d'interaction du régulateur YufM dans les régions promotrices de *maeN* et d'*yflS*. Cependant, l'ensemble de leurs résultats et des notres ne permet pas de dégager une séquence d'interaction entre YufM et l'ADN qui soit consensuelle, ce qui suggère que YufM ne reconnaît pas une séquence primaire spécifique. Cette propriété n'est pas surprenante dans la mesure où d'autres régulateurs de TCS la partagent (Perego & Hoch, 2002). Quand bien même, nous avons réalisé une analyse bioinformatique des cibles potentielles de ce régulateur sur la base des différents motifs remarquées dans les régions régulatrices d'*ywkA*, de *maeN* et d'*yflS*. Hélas, cette recherche n'a pas permis pas d'obtenir de résultats encourageants.

Enfin, l'approche transcriptomique de la détermination du régulon YufL-YufM a permis de révéler 97 gènes dont l'expression est dérégulée dans un mutant *yufL* surexprimant YufM (Kobayashi *et al.*, 2001). Parmi ces gènes, beaucoup sont impliqués dans la compétence et plus globalement dans un état d'adaptation au stress appelé « K-state » (Berka *et al.*, 2002). Le système YufL-YufM contribuerait-il à définir cet état physiologique et/ou l'état de compétence ? Nous pouvons affirmer que la mutation de ce système (ou d'*ywkA*) n'entraîne pas une diminution notable de l'efficacité de transformation (Doan et Aymerich, non publié), ce qui suggère que ce TCS n'est pas directement impliqué dans la compétence. De plus, il convient d'être très prudent puisque, vraisemblablement, une seule comparaison (une seule microarray hybridée) a permis d'obtenir ces résultats et l'induction des gènes de compétence pourrait en fait être artéfactuelle. S'il s'avère qu'aucune autre piste n'est confirmée, nous pourrions peut-être envisager d'effectuer nous-mêmes une comparaison transcriptomique dans des conditions un peu différentes pour prendre en compte l'hypothèse du rôle spécifique du TCS YufL-YufM en réponse à la présence du malate extracellulaire. Il s'agirait alors de comparer les profils d'expression de la souche sauvage et d'une souche mutée pour *yufL-yufM* et surexprimant artificiellement le transporteur *maeN* de façon à ce que les deux souches aient des vitesses de croissance comparables en présence de malate.

Un lien avec la résistance au stress alcalin ?

Pour trouver une fonction physiologique qui impliquerait *ywkA*, nous nous basons aussi sur une étude qui montre que dans certaines conditions très spécifiques, le gène *maeN* pouvait être exprimé en même temps que les gènes *mrpABCDEF*, situés juste en aval. En effet, à l'exception de *mrpA*, la délétion de l'un des autres gènes de l'opéron *mrp* conduit à une surexpression de l'expression de *maeN* et de l'opéron *mrp* (comparée au niveau d'expression dans la souche sauvage) lors de la croissance dans un milieu contenant du malate (Ito *et al.*, 2000). Il convient de rester prudent puisque ce résultat ne suggère pas que l'expression de *maeN* et du cluster *mrp* puisse être coordonnée pour répondre à un besoin physiologique. De plus, on peut d'ores et déjà remarquer que l'organisation génétique *maeN-mrp* n'est pas conservée chez les autres Bacilli ou d'autres bactéries à Gram positif dont le génome est disponible (Figure 49), ni chez les bactéries à Gram négatif considérées lors de cette recherche. Malgré tout, cette piste mérite notre attention.

Les gènes *mrp* sont impliqués dans la résistance au stress alcalin et dans la régulation du pH intracellulaire par un système d'antiport Na^+/H^+ (Ito *et al.*, 2000). Les auteurs suggèrent que la coordination de l'expression de *maeN* et des gènes *mrp* pourrait être nécessaire pour accomplir un cycle complet dans lequel le Na^+ serait importé avec le malate *via* MaeN et réexpulsé en échange de H^+ *via* le système Mrp. Dans des conditions de pH alcalin, cela permettrait le transport de malate et la régulation du pH intracellulaire de façon conjointe. Ce cycle impliquerait aussi une activité NADH déshydrogénase, codée par les gènes *mrp*, pour fournir de l'énergie électrochimique au système de transport et ainsi augmenter son efficacité. YwkA pourrait convertir le malate en présence de NAD et fournir le NADH requis dans le cycle du Na^+. Pour tester cette hypothèse complexe, il sera intéressant de comparer la vitesse de croissance de la souche sauvage et d'un mutant *ywkA* en fonction du pH et de la concentration en sodium du milieu.

ywkA pourrait être impliqué dans la croissance en anaérobiose.

Bacillus subtilis est capable de croître en anaérobiose (Nakano & Zuber, 1998) (cf partie 1.3.4). Dans ces conditions, l'accepteur final d'électron de la chaîne respiratoire est le nitrate ou le nitrite. Mais chez *Escherichia coli*, l'accepteur final d'électron est le fumarate. Il est produit par le cycle de Krebs qui fonctionne alors dans le sens inverse (oxaloacétate → malate → fumarate). Dans le cas de *Bacillus subtilis*, la production de fumarate comme accepteur d'électron pourrait impliquer l'enzyme malique : le malate serait converti en pyruvate par YwkA, à son tour converti en oxaloacétate par PycA, à son tour converti en aspartate par AspB, enfin converti en fumarate par AsnB (Figure 7). Pour tester l'implication d'*ywkA* dans ce processus, il sera intéressant de comparer la croissance de la souche sauvage et d'un mutant *ywkA* en milieu anaérobie. Cependant, l'expression d'*ywkA* ne semble pas régulée par la présence ou l'absence d'oxygène : il n'y a pas de cible de FNR, le régulateur global de l'expression des gènes impliqués dans la croissance en anaérobiose, dans la région régulatrice de l'expression d'*ywkA* et il n'est pas fait mention de ce gène dans une comparaison transcriptomique d'une souche sauvage cultivée dans des conditions aérobies ou anaérobies.

Quelle est la fonction de la protéine codée par _ywkB_ ?

Notre dernière piste concerne l'étude du gène *ywkB*. Les résultats de l'étude transcriptionnelle effectuée par nos collaborateurs montrent qu'*ywkA* et *ywkB* forment un opéron transcrit en réponse à la présence de malate (Figure 4, Article n°2). Ce gène code une protéine de 319 acides aminés de fonction inconnue. Cependant, la recherche de motifs protéiques connus avec l'outil SMART (Simple Modular Architecture Research Tool) indique qu'YwkB possède un domaine homologue à un domaine membranaire de MdcE, un transporteur potentiel du malonate chez *Klebsiella pneumoniae*. De plus, la recherche d'orthologues dans la base de données SSDB *via* le serveur KEGG indique qu'YwkB présente des similitudes faibles (autour de 25 % d'identité au plus sur toute la séquence) avec des perméases bactériennes putatives du malonate, du malate et des acides dicarboxyliques en C_4. La recherche de la fonction de ce gène pourra nous aider à déterminer celle d'*ywkA*. Dans le mutant *ywkA* que nous avons utilisé lors de cette étude, l'expression du gène *ywkB* a été placée sous contrôle d'un promoteur inductible à l'IPTG. Nous avons donc comparé la

croissance de ce mutant avec ou sans induction de l'expression d'*ywkB* mais aucune différence de phénotype n'est apparue.

Et *malS* et *ytsJ* ?

L'étude du rôle physiologique d'*ywkA* nous a conduit à nous intéresser aux gènes *malS* et *ytsJ*. L'étude de ces deux gènes a été plus particulièrement confiée à Guillaume Lerondel qui débute sa thèse au laboratoire.

Concernant l'étude du gène *malS*, Guillaume a décidé de reconstruire des fusions transcriptionnelles différentes de celles déjà existantes à partir des régions régulatrices en amont de *malS* ou en amont d'*ytnP* (Figure 47), cela afin de savoir laquelle des deux hypothèses avancées pour expliquer la contradiction observée est la plus probable.

L'ensemble des résultats acquis sur *ytsJ* suggère que ce gène pourrait être requis pour catalyser une réaction anaplérotique qui permettrait le fonctionnement du cycle de Krebs et/ou pour régénérer le NADPH nécessaire pour certaines réactions anaboliques lorsque *Bacillus subtilis* utilise des acides dicarboxyliques en C_4 comme seule source de carbone. Pour tester ces hypothèses, une souche est en cours de construction dans laquelle le gène *ytsJ* est délété et l'expression du gène *ywkA* est placée sous contrôle d'un promoteur inductible à l'IPTG. Ainsi, l'activité enzyme malique NADP-dépendante d'YtsJ peut être remplacée, en présence d'IPTG, par l'activité enzyme malique NAD-dépendante d'YwkA. L'idée est donc de comparer la croissance de cette souche en présence de malate comme seule source de carbone avec ou sans l'expression d'*ywkA*. Si cette souche croît à une même vitesse en présence ou en absence de l'activité d'YwkA et que cette vitesse de croissance est celle du mutant *ytsJ* déjà testé, cela voudra dire qu'*ywkA* ne complémente pas la délétion d'*ytsJ*. Ce résultat indiquerait qu'YtsJ est très utile pour la production de NADPH requis dans ces conditions de croissance. Dans ce cas, il faudra s'assurer qu'YwkA est bien synthétisée (et bien active). La stratégie envisagée est de remplacer la copie sauvage d'*ywkA* (toujours sous contrôle de son promoteur inductible) par un gène codant une protéine fusionnée à une étiquette histidine en N-terminal pour pouvoir purifier cette enzyme et réaliser des tests enzymologiques comme indiqué dans l'article n°2. En revanche, si cette souche retrouve une croissance efficace en présence de malate lorsque l'expression d'*ywkA* est induite, cela suggèrera que c'est la production de pyruvate qui est importante pour la cellule. Et si la vitesse de croissance est intermédiaire

entre celle de la souche sauvage et celle du mutant *ytsJ*, cela suggèrera qu'*ytsJ* est requis à la fois pour le bon fonctionnement du cycle de Krebs et à la fois pour la production de NADPH.

Enfin, le gène *mleA* a été laissé à l'écart de cette étude parce qu'il a été montré que le gène *mleN* (*yqkI*), avec lequel il est apparemment cotranscrit, est un antiporteur malate/lactate et que donc il code très probablement une enzyme malolactique plutôt qu'une enzyme malique (Wei *et al.*, 2000). Il sera intéressant de confirmer ou d'infirmer cette hypothèse. Nous allons tenter de purifier la protéine MleA et de déterminer quel type d'activité enzymatique elle possède.

Les GAPDH et les enzymes maliques sont-elles impliquées dans la gestion de l'excès de NADPH produit chez *Bacillus subtilis* dans certaines conditions ?

Dans une vision plus large de la fonction des GAPDH et des enzymes maliques chez *Bacillus subtilis*, nous avons aussi entamé une collaboration avec l'équipe d'Uwe Sauer pour savoir si ces enzymes sont impliquées dans une fonction physiologique particulière : la gestion de l'excès de NADPH produit par *Bacillus subtilis* dans certaines conditions.

L'équipe de Uwe Sauer a montré que le flux de glucose qui passe dans la voie des pentoses phosphate n'excède pas 50 % dans des cultures en chémostat limitées par la quantité de glucose disponible (Dauner *et al.*, 2001a; Dauner *et al.*, 2001b). Cependant, lorsque le glucose se trouve en excès par rapport aux besoins anaboliques, le flux de carbone à travers la partie oxydative de la voie des pentoses phosphate peut excéder 50 % (Sauer *et al.*, 1997; Dauner *et al.*, 2001b). Dans ces conditions, les enzymes de la voie des pentoses phosphate produisent un excès de NADPH. Cet excès devrait être contrebalancé par la conversion du NADPH en NADH par une activité de type NAD transhydrogénase (Dauner *et al.*, 2001b). Effectivement, Sauer et Dauner (2001) ont mis en évidence une telle activité dans des cultures de *Bacillus subtilis* en chémostat. Chez *Escherichia coli*, il existe deux transhydrogénases de ce type codées par les gènes *pntA-pntB* (codant deux sous unités) (Clarke & Bragg, 1985; Clarke *et al.*, 1986) et *udhA* (Boonstra *et al.*, 1999; Canonaco *et al.*, 2001). Cependant, l'analyse de la séquence complète du génome de *Bacillus subtilis* n'a pas révélé l'existence d'un gène qui coderait une protéine homologue aux protéines d'*Escherichia coli* ou une protéine homologue à une transhydrogénase d'une autre classe connue pour cette fonction.

Chez l'Archée *Methanobacter thermoautotrophicum* qui ne possède pas de NAD transhydrogénase non plus (sur la base de la séquence), il a été suggéré que les deux malate déshydrogénases, l'une catalysant la réduction de l'oxaloacétate en présence de NADH ou de NADPH et l'autre catalysant la réaction inverse en présence de NAD, pourraient être impliqués dans le processus de transhydrogénation du NADP (Thompson *et al.*, 1998).

Comment l'excès de NADPH est-il converti en NADH chez *Bacillus subtilis* ? Sauer et ses collègues ont proposé que ce processus pouvait être réalisé à travers un flux de carbone cyclique impliquant des isoenzymes possédant des spécificités différentes de cofacteur (Dauner *et al.*, 2001b). Deux cycles possibles étaient alors envisagés, au niveau des glycéraldéhyde-3-phosphate déshydrogénases ou au niveau des enzymes maliques (puisque nous avons montré qu'YwkA et MalS sont NAD-dépendantes et qu'YtsJ est NAPD-dépendante), voire par une combinaison des deux types d'enzymes. Ces cycles ne sont quasiment pas détectables par analyse de flux métaboliques dans la souche sauvage. Dans ce cadre, nous collaborons avec l'équipe de Uwe Sauer pour déterminer si les GAPDH ou les enzymes maliques sont impliquées dans ce processus. Pour cela, la comparaison de la distribution des flux de carbone dans la souche sauvage et dans les souches portant les différentes combinaisons de mutations des gènes d'intérêt permettra de déterminer le rôle de chaque enzyme. Notre principale contribution a été de construire ou de fournir ces différents mutants. Concernant GapA et GapB, nos collègues suisses ont déjà déterminés qu'elles n'étaient pas impliquées dans la conversion du NADPH en NADH. Une analyse préliminaire semble montrer que les enzymes maliques ne sont pas impliquées dans ce processus. Si de nouvelles expériences confortaient ce résultat, il faudrait faire l'hypothèse que *Bacillus subtilis* possède une transhydrogénase d'un type nouveau ou met en oeuvre un processus complexe de régénération du NADH à partir du NADPH.

La conversion directe du pyruvate en PEP est-elle possible chez *Bacillus subtilis* 168CA ?

Le pyruvate est un métabolite situé au carrefour entre la glycolyse et le cycle de Krebs. Il est à la fois une porte de sortie de la glycolyse et une porte d'entrée dans la néoglucogenèse. L'enzyme malique est l'une des étapes qui permet de produire le pyruvate et de faire

fonctionner la néoglucogenèse *via* PckA. Mais il faut aussi considérer une potentielle activité phosphoénolpyruvate synthase.

Un mutant *pckA* présente un défaut de croissance très sévère dans un milieu minimum supplémenté en malate (Doan et Aymerich, non publié). Ce résultat suggère que, lorsque la bactérie est cultivée en présence de malate comme seule source de carbone, celui-ci est très majoritairement converti en oxaloacétate *via* la malate déshydrogénase (*mdh*) et/ou l'enzyme malique et la pyruvate carboxylase (*pycA*) puis en PEP *via* la phosphoénolpyruvate carboxykinase (*pckA*) pour entrer dans la néoglucogenèse (Figure 46). Mais la croissance résiduelle du mutant *pckA* dans ces conditions suggère aussi qu'il existe une autre voie de conversion du malate en PEP. Notre hypothèse est que le pyruvate produit par l'enzyme malique pourrait être directement converti en PEP *via* une phosphoénolpyruvate synthase ou une pyruvate dikinase. En effet, l'analyse de la séquence complète du génome révèle l'existence de deux gènes, respectivement *pps* et *yvkC*, dont les produits sont homologues à ces enzymes. Aucune étude détaillée n'est disponible concernant ces gènes. Dans un premier temps, nous allons regarder si ces deux gènes sont exprimés dans des conditions néoglucogéniques. Puis, nous essaierons de détecter l'activité enzymatique des protéines purifiées Pps et YvkC. Par ailleurs, il sera intéressant de savoir si la surexpression de l'un de ces gènes permet de restaurer ou non la croissance efficace d'un mutant *pckA* en présence de malate comme seule source de carbone.

Etude transcriptomique.

La troisième partie de cette thèse porte sur la mise au point de l'outil transcriptomique et son utilisation pour la recherche de nouveaux gènes impliqués dans le métabolisme carboné. Jusqu'à présent, notre principal critère pour déterminer si un gène est significativement différentiellement exprimé est la reproductibilité stricte d'un résultat dans plusieurs expériences identiques. Certes, cette méthode d'analyse évite probablement de retenir un grand nombre de « faux positifs » mais elle élimine très certainement des gènes intéressants. Nous avons employé cet outil avec succès pour confirmer la spécificité du répresseur CggR (Article n°1). Mais dans le cadre de notre projet qui vise à comparer le profil d'expression du génome en fonction de six sources de carbone, la masse de données à analyser sera trop grande et trop complexe. Il nous faut maintenant acquérir une expertise pour l'analyse statistique de nos données afin de pouvoir comparer en même temps un grand nombre d'expériences réalisées à partir de plus de deux sources de carbone différentes.

Deux méthodes statistiques sont envisagées pour analyser notre échantillon de plusieurs expériences répétées (réplicats, triplicats, voire plus). La première approche est basée sur un test de Student (*t*-test) : pour chaque gène, les intensités d'expression sont normalisées en tenant compte de toutes les expériences comparées et un *t*-test est effectué en utilisant les ratios des log des intensités des réplicats pour déterminer un score de probabilité, *p*-value. Pour le gène X, l'hypothèse nulle Ho est qu'il n'y pas de différence d'expression entre les deux conditions comparées. Le *p*-value sera donc de 1 pour un gène dont l'expression est identique dans les deux conditions et tendra vers 0 pour un gène dont l'expression sera très différente entre les deux conditions. Ce test permet donc de déterminer les gènes significativement différentiellement exprimés par leur faible valeur de *p*-value. Ainsi, l'utilisation d'un intervalle de confiance de 5 % par exemple (p-value < 0,05) permet de sélectionner un certain nombre de gènes différentiellement exprimés avec une probabilité de 5 % d'erreur de sélectionner un faux positif. Si l'on prend en compte tous les gènes de *Bacillus subtilis*, cela représente tout de même un taux d'erreur au hasard de près de 200 gènes. Il existe des méthodes pour rectifier la valeur du *p*-value. Celle qui semble la plus appropriée est une méthode appelée SAM (Significance Analysis of Microarrays) décrite par Tusher et ses collaborateurs (Tusher *et al.*, 2001; Pan, 2002). Sans entrer dans les détails, un pourcentage de gènes identifiés par chance (False Discovery Rate) est calculé à partir de permutations des mesures répétées pour un intervalle de confiance donné. La seconde

approche est celle communément nommée ANOVA (Analysis of Variance). De façon simplifiée, une valeur est associée à la variation dûe aux erreurs inhérentes à la manipulation (hétérogénéité des membranes, différences de marquage, etc...) pour permettre à l'expérimentateur d'estimer la signification d'un ratio d'expression entre deux conditions.

Par ailleurs, le laboratoire s'est doté d'un lecteur de fluorescence sur lames de verre. Aussi, nous pourrons comparer nos résultats obtenus à partir d'hybridations de membrane de nylon à des expériences futures sur support de verre.

Comment exploiter les résultats ?

Les résultats préliminaires que nous avons obtenus en comparant l'expression des gènes en conditions glycolytiques ou néoglucogéniques sont encourageant dans le sens où ils sont en accord avec d'autres résultats publiés (Moreno *et al.*, 2001; Yoshida *et al.*, 2001; Blencke *et al.*, 2003 sous presse) (Article n°2) (Tables 9 et 10). A cet égard, ces listes de gènes pourront être exploitées lorsque nous nous serons assurés de la reproductibilité de nos résultats (en parallèle de la poursuite de l'étude transcriptomique avec d'autres sources de carbone). La première piste qui nous intéresse concerne les gènes codant des protéines de fonction inconnue que nous avons trouvés régulés en fonction de la glycolyse/néoglucogenèse et qui pourraient être impliqués dans le métabolisme carboné. Dans un premier temps, il s'agira de vérifier expérimentalement que l'expression des gènes candidats est bien régulée en fonction de la présence de glucose ou de succinate. Puis, nous envisagerons les études génétiques et biochimiques qui pourraient permettre de déterminer la fonction de ces gènes.

De façon intéressante, nos résultats (Tables 9 et 10) et ceux de nos confrères (Blencke *et al.*, 2003 sous presse) suggèrent que les gènes de la voie des pentoses phosphate sont constitutivement exprimés à l'exception du gène *zwf* (glucose-6-phosphate 1-déshydrogénase). Au laboratoire, une autre analyse réalisée à partir de 4 comparaisons du transcriptome de la souche sauvage et d'un mutant YqzB (répresseur de *gapB* et *pckA*) en conditions glycolytiques montre que l'expression de *zwf* est significativement plus exprimé dans le mutant *yqzB*. Ces résultats suggère une coordination entre l'activité de la néoglucogenèse et celle de la voie des pentoses phosphate. L'induction de l'expression de *zwf* servirait peut-être à permettre l'entrée du carbone en quantité suffisante dans la voie des pentoses pour fournir les précurseurs anaboliques en C_5. Actuellement, l'étude expérimentale

de cette dérégulation est cours. Si elle confirme que l'expression de *zwf* est induite par la néoglucogenèse, il sera intéressant d'étudier le mécanisme moléculaire de cette régulation. Aussi, à plus long terme, nous allons focaliser une partie de nos efforts sur la régulation des gènes de la voie des pentoses phosphate.

Intégration des données dans le contexte général du métabolisme carboné chez *Bacillus subtilis.*

Nos recherches ont permis d'acquérir de nouvelles informations concernant la régulation des gènes du métabolisme central du carbone. Il convient donc de replacer ces données dans leur contexte. En introduction, l'intégration des données dans un modèle global à contrainte avait été évoquée. Evidemment, nous n'en sommes pas là ! Cependant, dans cette optique, la Figure 62 tente de rassembler les principales données de régulation de l'expression des gènes impliqués dans le transport des glucides (à partir quelques exemples) et le métabolisme central du carbone chez *Bacillus subtilis*. Ce schéma intègre nos récentes découvertes : l'étude du mécanisme de régulation de l'expression de l'opéron *gapA* (Fillinger *et al.*, 2000; Ludwig *et al.*, 2001; Ludwig *et al.*, 2002b) (Article n°1) ; l'analyse de la régulation de l'expression des gènes clés de la néoglucogenèse, *gapB* et *pckA* (Servant et Aymerich, manuscrit en cours de préparation) et l'étude de quatre gènes paralogues pouvant coder des enzymes maliques (Article n°2, soumis ; Lerondel, Doan et Aymerich, non publié). CcpA joue un rôle majeur dans la régulation du métabolisme carboné. Mais ce rôle est exercé principalement sur la régulation de l'expression des gènes impliqués dans le transport et le catabolisme spécifique des sources de carbone avant leur entrée dans les voies générales de la glycolyse, des pentoses phosphate et du cycle de Krebs (Figure 61). Si l'on considère uniquement ces trois voies *sensu stricto*, on constate que CcpA n'exerce son rôle direct de régulateur transcriptionnel qu'au niveau du cycle de Krebs, sur l'expression du gène *citZ* (Figures 28 et 61). Les rôles de CcpA dans la régulation de l'expression des opérons *gapA* et *pdhABCD*, des gènes *gapB* et *pckA*, et du gène *citB* sont indirects (Fillinger *et al.*, 2000; Moreno *et al.*, 2001; Kim *et al.*, 2002a; Ludwig *et al.*, 2002b; Blencke *et al.*, 2003 sous presse) (Servant et Aymerich, manuscrit en préparation ; Article n°1). Par ailleurs, les résultats des expériences globales de transcriptome citées suggèrent que les opérons *sucCD* et

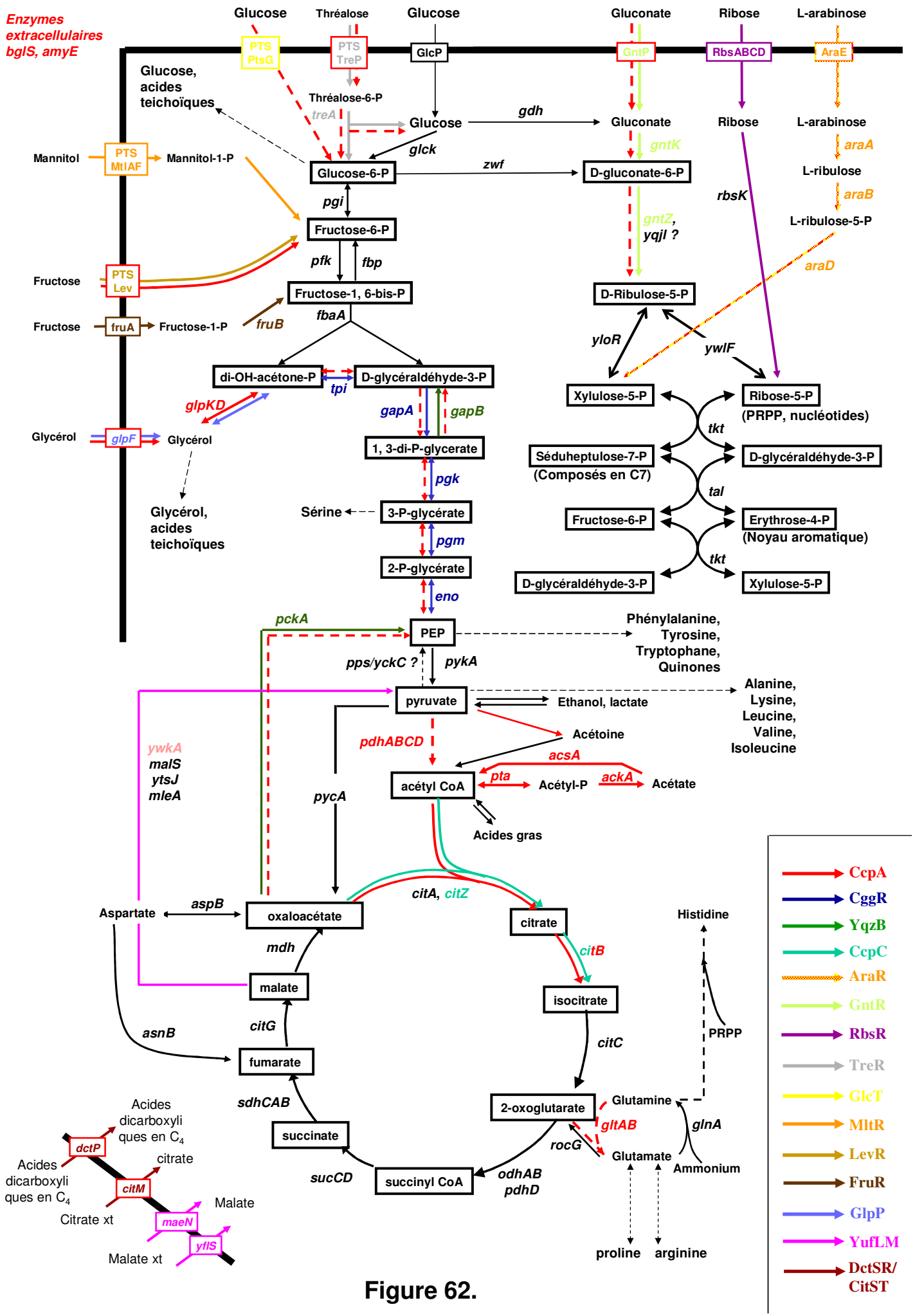

Enzymes extracellulaires bglS, amyE
Glucose
Thréalose
Glucose
Gluconate
Ribose
L-arabinose
PTS PtsG
PTS TreP
GlcP
GntP
RbsABCD
AraE
Glucose, acides teichoïques
Thréalose-6-P
treA
Glucose
gdh
Gluconate
Ribose
L-arabinose
glck
gntK
araA
Mannitol
PTS MtlAF
Mannitol-1-P
Glucose-6-P
zwf
D-gluconate-6-P
L-ribulose
rbsK
araB
pgi
gntZ, yqjI ?
L-ribulose-5-P
Fructose-6-P
pfk
fbp
araD
Fructose
PTS Lev
Fructose-1, 6-bis-P
D-Ribulose-5-P
Fructose
fruA
Fructose-1-P
fruB
fbaA
yloR
ywlF
di-OH-acétone-P
tpi
D-glycéraldéhyde-3-P
Xylulose-5-P
Ribose-5-P
(PRPP, nucléotides)
glpKD
gapA
gapB
tkt
Glycérol
glpF
Glycérol
1, 3-di-P-glycerate
Séduheptulose-7-P
(Composés en C7)
D-glycéraldéhyde-3-P
pgk
tal
Glycérol, acides teichoïques
Sérine
3-P-glycérate
Fructose-6-P
Erythrose-4-P
(Noyau aromatique)
pgm
tkt
2-P-glycérate
D-glycéraldéhyde-3-P
Xylulose-5-P
eno
pckA
PEP
Phénylalanine, Tyrosine, Tryptophane, Quinones
pps/yckC ?
pykA
pyruvate
Ethanol, lactate
Alanine, Lysine, Leucine, Valine, Isoleucine
Acétoine
ywkA
malS
ytsJ
mleA
pdhABCD
acsA
acétyl CoA
pta
Acétyl-P
ackA
Acétate
pycA
Acides gras
citA, citZ
Aspartate
aspB
oxaloacétate
citrate
Histidine
mdh
citB
malate
isocitrate
asnB
citG
citC
PRPP
fumarate
sdhCAB
2-oxoglutarate
Glutamine
Acides dicarboxyliques en C_4
succinate
gltAB
glnA
dctP
rocG
Glutamate
citrate
Acides dicarboxyliques en C_4
sucCD
succinyl CoA
odhAB
pdhD
Ammonium
citM
Malate
Citrate xt
maeN
yflS
Malate xt
proline
arginine
CcpA
CggR
YqzB
CcpC
AraR
GntR
RbsR
TreR
GlcT
MltR
LevR
FruR
GlpP
YufLM
DctSR/ CitST

Figure 62.

odhAB sont aussi soumis à la répression catabolique mais le mécanisme sous-jacent n'est pas connu. D'ailleurs, si l'on compare *Escherichia coli* (Figure 38) et *Bacillus subtilis* (Figure 61), on s'aperçoit que le nombre de gènes régulés des voies centrales du métabolisme carboné semble plus important chez cette première. Cette observation est peut-être à mettre en relation avec la capacité d'*Escherichia coli* à s'adapter à des conditions physiologiques très variées. Chez *Bacillus subtilis*, bien que CcpA, CcpC, CggR et YqzB régulent directement l'expression d'un faible nombre de gènes des voies centrales du métabolisme carboné, on peut remarquer que ces derniers codent des enzymes catalysant des étapes clés du métabolisme : l'opéron *gapA* code les enzymes constituant la voie unique de conversion entre le G3P et le PEP ; le gène *citZ* code l'enzyme majeure catalysant l'entrée du carbone (glycolytique) dans le cycle de Krebs ; *gapB* et *pckA* codent deux enzymes nécessaires à la néoglucogenèse.

Bibliographie

Alen, C. & Sonenshein, A. L. (1999). *Bacillus subtilis* aconitase is an RNA-binding protein. *Proc Natl Acad Sci U S A* **96**, 10412-10417.

Alexeeva, S., de Kort, B., Sawers, G., Hellingwerf, K. J. & de Mattos, M. J. (2000). Effects of limited aeration and of the ArcAB system on intermediary pyruvate catabolism in *Escherichia coli*. *J Bacteriol* **182**, 4934-4940.

Alpert, C. A., Dorschug, M., Saffen, D., Frank, R., Deutscher, J. & Hengstenberg, W. (1985). The bacterial phosphoenolpyruvate-dependent phosphotransferase system. Isolation of active site peptides by reversed-phase high-performance liquid chromatography and determination of their primary structure. *J Chromatogr* **326**, 363-371.

Amster-Choder, O. & Wright, A. (1990). Regulation of activity of a transcriptional anti-terminator in *E. coli* by phosphorylation *in vivo*. *Science* **249**, 540-542.

Amster-Choder, O. & Wright, A. (1992). Modulation of the dimerization of a transcriptional antiterminator protein by phosphorylation. *Science* **257**, 1395-1398.

Amster-Choder, O. & Wright, A. (1993). Transcriptional regulation of the bgl operon of *Escherichia coli* involves phosphotransferase system-mediated phosphorylation of a transcriptional antiterminator. *J Cell Biochem* **51**, 83-90.

Amster-Choder, O., Houman, F. & Wright, A. (1989). Protein phosphorylation regulates transcription of the beta-glucoside utilization operon in *E. coli*. *Cell* **58**, 847-855.

Andersen, H. W., Solem, C., Hammer, K. & Jensen, P. R. (2001). Twofold reduction of phosphofructokinase activity in *Lactococcus lactis* results in strong decreases in growth rate and in glycolytic flux. *J Bacteriol* **183**, 3458-3467.

Anderson, A. & Cooper, R. A. (1969). Gluconeogenesis in *Escherichia coli* The role of triose phosphate isomerase. *FEBS Lett* **4**, 19-20.

Arnaud, M., Debarbouille, M., Rapoport, G., Saier, M. H., Jr. & Reizer, J. (1996). In vitro reconstitution of transcriptional antitermination by the SacT and SacY proteins of *Bacillus subtilis*. *J Biol Chem* **271**, 18966-18972.

Attwood, P. V. (1995). The structure and the mechanism of action of pyruvate carboxylase. *Int J Biochem Cell Biol* **27**, 231-249.

Auzat, I., Le Bras, G. & Garel, J. R. (1997). Allosteric activation increases the maximum velocity of *E. coli* phosphofructokinase. *J Mol Biol* **267**, 476-480.

Aymerich, S. & Steinmetz, M. (1992). Specificity determinants and structural features in the RNA target of the bacterial antiterminator proteins of the BglG/SacY family. *Proc Natl Acad Sci U S A* **89**, 10410-10414.

Aymerich, S., Gonzy-Treboul, G. & Steinmetz, M. (1986). 5'-noncoding region sacR is the target of all identified regulation affecting the levansucrase gene in *Bacillus subtilis*. *J Bacteriol* **166**, 993-998.

Babul, J. & Guixe, V. (1983). Fructose bisphosphatase from *Escherichia coli*. Purification and characterization. *Arch Biochem Biophys* **225**, 944-949.

Bachem, S. & Stülke, J. (1998). Regulation of the *Bacillus subtilis* GlcT antiterminator protein by components of the phosphotransferase system. *J Bacteriol* **180**, 5319-5326.

Bachem, S., Faires, N. & Stülke, J. (1997). Characterization of the presumptive phosphorylation sites of the *Bacillus subtilis* glucose permease by site-directed mutagenesis: implication in glucose transport and catabolite repression. *FEMS Microbiol Lett* **156**, 233-238.

Baldwin, S. A., Perham, R. N. & Stribling, D. (1978). Purification and characterization of the class-II D-fructose 1,6-bisphosphate aldolase from *Escherichia coli* (Crookes' strain). *Biochem J* **169**, 633-641.

Beijer, L. & Rutberg, L. (1992). Utilisation of glycerol and glycerol 3-phosphate is differently affected by the phosphotransferase system in *Bacillus subtilis*. *FEMS Microbiol Lett* **79**, 217-220.

Bergmann, S., Rohde, M., Chhatwal, G. S. & Hammerschmidt, S. (2001). alpha-Enolase of *Streptococcus pneumoniae* is a plasmin(ogen)-binding protein displayed on the bacterial cell surface. *Mol Microbiol* **40**, 1273-1287.

Berka, R. M., Hahn, J., Albano, M., Draskovic, I., Persuh, M., Cui, X., Sloma, A., Widner, W. & Dubnau, D. (2002). Microarray analysis of the *Bacillus subtilis* K-state: genome-wide expression changes dependent on ComK. *Mol Microbiol* **43**, 1331-1345.

Blangy, D., Buc, H. & Monod, J. (1968). Kinetics of the allosteric interactions of phosphofructokinase from *Escherichia coli*. *J Mol Biol* **31**, 13-35.

Blencke, H. M., Homuth, G., Ludwig, H., Mäder, U., Hecker, M. & Stülke, J. (2003 sous presse). Transcriptional profiling of gene expression in response to glucose in *Bacillus subtilis*: regulation of the central metabolic pathways. *Metabolic Engineering*.

Boiteux, A., Markus, M., Plesser, T., Hess, B. & Malcovati, M. (1983). Analysis of progress curves. Interaction of pyruvate kinase from *Escherichia coli* with fructose 1,6-bisphosphate and calcium ions. *Biochem J* **211**, 631-640.

Boonstra, B., French, C. E., Wainwright, I. & Bruce, N. C. (1999). The udhA gene of *Escherichia coli* encodes a soluble pyridine nucleotide transhydrogenase. *J Bacteriol* **181**, 1030-1034.

Boschi-Muller, S. & Branlant, G. (1999). The active site of phosphorylating glyceraldehyde-3-phosphate dehydrogenase is not designed to increase the nucleophilicity of a serine residue. *Arch Biochem Biophys* **363**, 259-266.

Boschi-Muller, S., Azza, S., Pollastro, D., Corbier, C. & 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.

Brown, C. K., Kuhlman, P. L., Mattingly, S., Slates, K., Calie, P. J. & Farrar, W. W. (1998). A model of the quaternary structure of enolases, based on structural and evolutionary analysis of the octameric enolase from *Bacillus subtilis*. *J Protein Chem* **17**, 855-866.

Bruckner, R. & Titgemeyer, F. (2002). Carbon catabolite repression in bacteria: choice of the carbon source and autoregulatory limitation of sugar utilization. *FEMS Microbiol Lett* **209**, 141-148.

Bulatnikov, I. G., Polyakova, O. V., Asryants, R. A., Nagradova, N. K. & Muronetz, V. I. (1999). Participation of chaperonin GroEL in the folding of D-glyceraldehyde-3-phosphate dehydrogenase. An approach based on the use of different oligomeric forms of the enzyme immobilized on sepharose. *J Protein Chem* **18**, 79-87.

Burklen, L., Schock, F. & Dahl, M. K. (1998). Molecular analysis of the interaction between the *Bacillus subtilis* trehalose repressor TreR and the tre operator. *Mol Gen Genet* **260**, 48-55.

Cancilla, M. R., Davidson, B. E., Hillier, A. J., Nguyen, N. Y. & Thompson, J. (1995). The *Lactococcus lactis* triosephosphate isomerase gene, *tpi*, is monocistronic. *Microbiology* **141 (Pt 1)**, 229-238.

Canonaco, F., Hess, T. A., Heri, S., Wang, T., Szyperski, T. & Sauer, U. (2001). Metabolic flux response to phosphoglucose isomerase knock-out in *Escherichia coli* and impact of overexpression of the soluble transhydrogenase UdhA. *FEMS Microbiol Lett* **204**, 247-252.

Carls, R. A. & Hanson, R. S. (1971). Isolation and characterization of tricarboxylic acid cycle mutants of *Bacillus subtilis*. *J Bacteriol* **106**, 848-855.

Carpousis, A. J. (2002). The *Escherichia coli* RNA degradosome: structure, function and relationship in other ribonucleolytic multienzyme complexes. *Biochem Soc Trans* **30**, 150-155.

Chander, M., Setlow, B. & Setlow, P. (1998). The enzymatic activity of phosphoglycerate mutase from gram-positive endospore-forming bacteria requires Mn2+ and is pH sensitive. *Can J Microbiol* **44**, 759-767.

Chao, Y. P., Patnaik, R., Roof, W. D., Young, R. F. & Liao, J. C. (1993). Control of gluconeogenic growth by *pps* and *pck* in *Escherichia coli*. *J Bacteriol* **175**, 6939-6944.

Charrier, V., Deutscher, J., Galinier, A. & Martin-Verstraete, I. (1997). Protein phosphorylation chain of a *Bacillus subtilis* fructose-specific phosphotransferase system and its participation in regulation of the expression of the *lev* operon. *Biochemistry* **36**, 1163-1172.

Chatterjee, S., Zhou, Y. N., Roy, S. & Adhya, S. (1997). Interaction of Gal repressor with inducer and operator: induction of gal transcription from repressor-bound DNA. *Proc Natl Acad Sci U S A* **94**, 2957-2962.

Chauvaux, S., Paulsen, I. T. & Saier, M. H., Jr. (1998). CcpB, a novel transcription factor implicated in catabolite repression in *Bacillus subtilis*. *J Bacteriol* **180**, 491-497.

Chauvin, F., Brand, L. & Roseman, S. (1994a). Sugar transport by the bacterial phosphotransferase system. Characterization of the *Escherichia coli* enzyme I monomer/dimer equilibrium by fluorescence anisotropy. *J Biol Chem* **269**, 20263-20269.

Chauvin, F., Brand, L. & Roseman, S. (1994b). Sugar transport by the bacterial phosphotransferase system. Characterization of the *Escherichia coli* enzyme I monomer/dimer transition kinetics by fluorescence anisotropy. *J Biol Chem* **269**, 20270-20274.

Chen, C. J., Chin, J. E., Ueda, K., Clark, D. P., Pastan, I., Gottesman, M. M. & Roninson, I. B. (1986). Internal duplication and homology with bacterial transport proteins in the mdr1 (P-glycoprotein) gene from multidrug-resistant human cells. *Cell* **47**, 381-389.

Chen, J., Sharma, S., Quiocho, F. A. & Davidson, A. L. (2001). Trapping the transition state of an ATP-binding cassette transporter: evidence for a concerted mechanism of maltose transport. *Proc Natl Acad Sci U S A* **98**, 1525-1530.

Chin, A. M., Feucht, B. U. & Saier, M. H. J. (1987). Evidence for regulation of gluconeogenesis by the fructose phosphotransferase system in *Salmonella typhimurium*. *J Bacteriol* **169**.

Chin, A. M., Feldheim, D. A. & Saier, M. H. J. (1989). Altered transcriptional patterns affecting several metabolic pathways in strains os *Salmonella typhimurium* which overexpress the fructose regulon. *J Bacteriol* **171**, 2424-2434.

Choy, H. E., Hanger, R. R., Aki, T., Mahoney, M., Murakami, K., Ishihama, A. & Adhya, S. (1997). Repression and activation of promoter-bound RNA polymerase activity by Gal repressor. *J Mol Biol* **272**, 293-300.

Clarke, D. M. & Bragg, P. D. (1985). Cloning and expression of the transhydrogenase gene of *Escherichia coli*. *J Bacteriol* **162**, 367-373.

Clarke, D. M., Loo, T. W., Gillam, S. & Bragg, P. D. (1986). Nucleotide sequence of the *pntA* and *pntB* genes encoding the pyridine nucleotide transhydrogenase of *Escherichia coli*. *Eur J Biochem* **158**, 647-653.

Clermont, S., Corbier, C., Mely, Y., Gerard, D., Wonacott, A. & 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**, 10178-10184.

Cocaign-Bousquet, M., Garrigues, C., Loubiere, P. & Lindley, N. D. (1996). Physiology of pyruvate metabolism in *Lactococcus lactis*. *Antonie Van Leeuwenhoek* **70**, 253-267.

Cooper, R. A. (1984). Metabolism of methylglyoxal in microorganisms. *Annu Rev Microbiol* **38**, 49-68.

Corbier, C., Michels, S., Wonacott, A. J. & Branlant, G. (1994). Characterization of the two anion-recognition sites of glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* by site-directed mutagenesis and chemical modification. *Biochemistry* **33**, 3260-3265.

Corbier, C., Clermont, S., Billard, P., Skarzynski, T., Branlant, C., Wonacott, A. & Branlant, G. (1990a). Probing the coenzyme specificity of glyceraldehyde-3-phosphate dehydrogenases by site-directed mutagenesis. *Biochemistry* **29**, 7101-7106.

Corbier, C., Mougin, A., Mely, Y., Adolph, H. W., Zeppezauer, M., Gerard, D., Wonacott, A. & Branlant, G. (1990b). The nicotinamide subsite of glyceraldehyde-3-phosphate dehydrogenase studied by site-directed mutagenesis. *Biochimie* **72**, 545-554.

Covert, M. W. & Palsson, B. O. (2002). Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*. *J Biol Chem* **277**, 28058-28064.

Cozzone, A. J. (1998). Regulation of acetate metabolism by protein phosphorylation in enteric bacteria. *Annu Rev Microbiol* **52**, 127-164.

Craig, J. E., Ford, M. J., Blaydon, D. C. & Sonenshein, A. L. (1997). A null mutation in the *Bacillus subtilis* aconitase gene causes a block in Spo0A-phosphate-dependent gene expression. *J Bacteriol* **179**, 7351-7359.

Cruz-Ramos, H., Glaser, P., Wray, L. V., Jr. & Fisher, S. H. (1997). The *Bacillus subtilis* ureABC operon. *J Bacteriol* **179**, 3371-3373.

Curtis, S. J. & Epstein, W. (1975). Phosphorylation of D-glucose in *Escherichia coli* mutants defective in glucosephosphotransferase, mannosephosphotransferase, and glucokinase. *J Bacteriol* **122**, 1189-1199.

Dahl, M. K. (2002). CcpA-independent carbon catabolite repression in *Bacillus subtilis*. *J Mol Microbiol Biotechnol* **4**, 315-321.

Dahl, M. K., Degenkolb, J. & Hillen, W. (1994). Transcription of the *xyl* operon is controlled in *Bacillus subtilis* by tandem overlapping operators spaced by four base-pairs. *J Mol Biol* **243**, 413-424.

Dahl, M. K., Schmiedel, D. & Hillen, W. (1995). Glucose and glucose-6-phosphate interaction with Xyl repressor proteins from *Bacillus* spp. may contribute to regulation of xylose utilization. *J Bacteriol* **177**, 5467-5472.

Daldal, F. (1983). Molecular cloning of the gene for phosphofructokinase-2 of *Escherichia coli* and the nature of a mutation, pfkB1, causing a high level of the enzyme. *J Mol Biol* **168**, 285-305.

Dandekar, T., Schuster, S., Snel, B., Huynen, M. & Bork, P. (1999). Pathway alignment: application to the comparative analysis of glycolytic enzymes. *Biochem J* **343 Pt 1**, 115-124.

Darbon, E., Servant, P., Poncet, S. & Deutscher, J. (2002). Antitermination by GlpP, catabolite repression via CcpA and inducer exclusion triggered by P-GlpK dephosphorylation control *Bacillus subtilis* glpFK expression. *Mol Microbiol* **43**, 1039-1052.

Dauner, M., Bailey, J. E. & Sauer, U. (2001a). Metabolic flux analysis with a comprehensive isotopomer model in *Bacillus subtilis*. *Biotechnol Bioeng* **76**, 144-156.

Dauner, M., Storni, T. & Sauer, U. (2001b). *Bacillus subtilis* metabolism and energetics in carbon-limited and excess-carbon chemostat culture. *J Bacteriol* **183**, 7308-7317.

Davidson, A. L. & Nikaido, H. (1990). Overproduction, solubilization, and reconstitution of the maltose transport system from *Escherichia coli*. *J Biol Chem* **265**, 4254-4260.

Davidson, A. L., Shuman, H. A. & Nikaido, H. (1992). Mechanism of maltose transport in *Escherichia coli*: transmembrane signaling by periplasmic binding proteins. *Proc Natl Acad Sci U S A* **89**, 2360-2364.

Davidson, A. L., Laghaeian, S. S. & Mannering, D. E. (1996). The maltose transport system of *Escherichia coli* displays positive cooperativity in ATP hydrolysis. *J Biol Chem* **271**, 4858-4863.

D'Costa, S. S., Romer, T. G. & Boyle, M. D. (2000). Analysis of expression of a cytosolic enzyme on the surface of *Streptococcus pyogenes*. *Biochem Biophys Res Commun* **278**, 826-832.

de Graef, M. R., Alexeeva, S., Snoep, J. L. & Teixeira de Mattos, M. J. (1999). The steady-state internal redox state (NADH/NAD) reflects the external redox state and is correlated with catabolic adaptation in *Escherichia coli*. *J Bacteriol* **181**, 2351-2357.

De Wulf, P. (1998). Presence of the ribulose monophosphate pathway in *Bacillus subtilis*. *Microbiology* **144**, 596-597.

Dean, D. A., Davidson, A. L. & Nikaido, H. (1990a). The role of ATP as the energy source for maltose transport in *Escherichia coli*. *Res Microbiol* **141**, 348-352.

Dean, D. A., Reizer, J., Nikaido, H. & Saier, M. H., Jr. (1990b). Regulation of the maltose transport system of *Escherichia coli* by the glucose-specific enzyme III of the phosphoenolpyruvate-sugar phosphotransferase system. Characterization of inducer exclusion-resistant mutants and reconstitution of inducer exclusion in proteoliposomes. *J Biol Chem* **265**, 21005-21010.

Débarbouillé, M., Martin-Verstraete, I., Klier, A. & Rapoport, G. (1991). The transcriptional regulator LevR of *Bacillus subtilis* has domains homologous to both sigma 54- and phosphotransferase system-dependent regulators. *Proc Natl Acad Sci U S A* **88**, 2212-2216.

Débarbouillé, M., Arnaud, M., Fouet, A., Klier, A. & Rapoport, G. (1990). The *sacT* gene regulating the *sacPA* operon in *Bacillus subtilis* shares strong homology with transcriptional antiterminators. *J Bacteriol* **172**, 3966-3973.

Declerck, N., Vincent, F., Hoh, F., Aymerich, S. & van Tilbeurgh, H. (1999). RNA recognition by transcriptional antiterminators of the BglG/SacY family: functional and structural comparison of the CAT domain from SacY and LicT. *J Mol Biol* **294**, 389-402.

Declerck, N., Dutartre, H., Receveur, V., Dubois, V., Royer, C., Aymerich, S. & van Tilbeurgh, H. (2001). Dimer stabilization upon activation of the transcriptional antiterminator LicT. *J Mol Biol* **314**, 671-681.

Dervyn, E., Suski, C., Daniel, R., Bruand, C., Chapuis, J., Errington, J., Janniere, L. & Ehrlich, S. D. (2001). Two essential DNA polymerases at the bacterial replication fork. *Science* **294**, 1716-1719.

Deutscher, J. & Saier, M. H., Jr. (1983). ATP-dependent protein kinase-catalyzed phosphorylation of a seryl residue in HPr, a phosphate carrier protein of the phosphotransferase system in *Streptococcus pyogenes*. *Proc Natl Acad Sci U S A* **80**, 6790-6794.

Deutscher, J., Sossna, G. & Gonzy-Treboul, G. (1989). Regulatory functions of the phosphocarrier protein HPr of the phosphoenol pyruvate-dependent phosphotransferase system in gram-positive bacteria. *FEMS Microbiol Rev* **5**, 167-174.

Deutscher, J., Galinier, A. & Martin-Verstraete, I. (2002). Carbohydrate Uptake and Metabolism. In *Bacillus subtilis and its Closest Relatives: from Genes to Cells*, pp. 129-150. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology Press.

Deutscher, J., Kessler, U., Alpert, C. A. & Hengstenberg, W. (1984). Bacterial phosphoenolpyruvate-dependent phosphotransferase system: P-ser-HPr and its possible regulatory function. *Biochemistry* **23**, 4455-4460.

Deutscher, J., Pevec, B., Beyreuther, K., Kiltz, H. H. & Hengstenberg, W. (1986). Streptococcal phosphoenolpyruvate-sugar phosphotransferase system: amino acid sequence and site of ATP-dependent phosphorylation of HPr. *Biochemistry* **25**, 6543-6551.

Deutscher, J., Kuster, E., Bergstedt, U., Charrier, V. & Hillen, W. (1995). Protein kinase-dependent HPr/CcpA interaction links glycolytic activity to carbon catabolite repression in gram-positive bacteria. *Mol Microbiol* **15**, 1049-1053.

Deutscher, J., Reizer, J., Fischer, C., Galinier, A., Saier, M. H., Jr. & Steinmetz, M. (1994). Loss of protein kinase-catalyzed phosphorylation of HPr, a phosphocarrier protein of the phosphotransferase system, by mutation of the ptsH gene confers catabolite repression resistance to several catabolic genes of Bacillus subtilis. *J Bacteriol* **176**, 3336-3344.

Deutscher, J., Fischer, C., Charrier, V., Galinier, A., Lindner, C., Darbon, E. & Dossonnet, V. (1997). Regulation of carbon metabolism in gram-positive bacteria by protein phosphorylation. *Folia Microbiol (Praha)* **42**, 171-178.

Didierjean, C., Rahuel-Clermont, S., Vitoux, B., Dideberg, O., Branlant, G. & Aubry, A. (1997). A crystallographic comparison between mutated glyceraldehyde-3-phosphate dehydrogenases from *Bacillus stearothermophilus* complexed with either NAD+ or NADP+. *J Mol Biol* **268**, 739-759.

Didierjean, C., Corbier, C., Fatih, M., Favier, F., Boschi-Muller, S., Branlant, G. & 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.*

Diesterhaft, M. D. & Freese, E. (1973). Role of pyruvate carboxylase, phosphoenolpyruvate carboxykinase, and malic enzyme during growth and sporulation of *Bacillus subtilis*. *J Biol Chem* **248**, 6062-6070.

Dills, S. S., Apperson, A., Schmidt, M. R. & Saier, M. H., Jr. (1980). Carbohydrate transport in bacteria. *Microbiol Rev* **44**, 385-418.

Dingman, D. W. & Sonenshein, A. L. (1987). Purification of aconitase from *Bacillus subtilis* and correlation of its N-terminal amino acid sequence with the sequence of the *citB* gene. *J Bacteriol* **169**, 3062-3067.

Dingman, D. W., Rosenkrantz, M. S. & Sonenshein, A. L. (1987). Relationship between aconitase gene expression and sporulation in *Bacillus subtilis*. *J Bacteriol* **169**, 3068-3075.

Donahue, J. L., Bownas, J. L., Niehaus, W. G. & Larson, T. J. (2000). Purification and characterization of *glpX*-encoded fructose 1, 6-bisphosphatase, a new enzyme of the glycerol 3-phosphate regulon of *Escherichia coli*. *J Bacteriol* **182**, 5624-5627.

Dossonnet, V., Monedero, V., Zagorec, M., Galinier, A., Perez-Martinez, G. & Deutscher, J. (2000). Phosphorylation of HPr by the bifunctional HPr Kinase/P-ser-HPr phosphatase from *Lactobacillus casei* controls catabolite repression and inducer exclusion but not inducer expulsion. *J Bacteriol* **182**, 2582-2590.

Edlin, J. D. & Sundaram, T. K. (1989). Regulation of isocitrate dehydrogenase by phosphorylation in *Escherichia coli* K-12 and a simple method for determining the amount of inactive phosphoenzyme. *J Bacteriol* **171**, 2634-2638.

Edwards, J. S. & Palsson, B. O. (2000). Metabolic flux balance analysis and the in silico analysis of *Escherichia coli* K-12 gene deletions. *BMC Bioinformatics* **1**, 1.

Eisermann, R., Deutscher, J., Gonzy-Treboul, G. & Hengstenberg, W. (1988). Site-directed mutagenesis with the ptsH gene of *Bacillus subtilis*. Isolation and characterization of heat-stable proteins altered at the ATP-dependent regulatory phosphorylation site. *J Biol Chem* **263**, 17050-17054.

el-Kabbani, O. A., Waygood, E. B. & Delbaere, L. T. (1987). Tertiary structure of histidine-containing protein of the phosphoenolpyruvate:sugar phosphotransferase system of *Escherichia coli*. *J Biol Chem* **262**, 12926-12929.

Essenberg, M. K. & Cooper, R. A. (1975). Two ribose-5-phosphate isomerases from *Escherichia coli* K12: partial characterisation of the enzymes and consideration of their possible physiological roles. *Eur J Biochem* **55**, 323-332.

Even, S., Garrigues, C., Loubiere, P., Lindley, N. D. & Cocaign-Bousquet, M. (1999). Pyruvate metabolism in *Lactococcus lactis* is dependent upon glyceraldehyde-3-phosphate dehydrogenase activity. *Metab Eng* **1**, 198-205.

Eyschen, J., Vitoux, B., Rahuel-Clermont, S., Marraud, M., Branlant, G. & Cung, M. T. (1996). Phosphorus-31 nuclear magnetic resonance studies on coenzyme binding and specificity in glyceraldehyde-3-phosphate dehydrogenase. *Biochemistry* **35**, 6064-6072.

Faires, N., Tobisch, S., Bachem, S., Martin-Verstraete, I., Hecker, M. & Stulke, J. (1999). The catabolite control protein CcpA controls ammonium assimilation in *Bacillus subtilis*. *J Mol Microbiol Biotechnol* **1**, 141-148.

Fiegler, H., Bassias, J., Jankovic, I. & Bruckner, R. (1999). Identification of a gene in *Staphylococcus xylosus* encoding a novel glucose uptake protein. *J Bacteriol* **181**, 4929-4936.

Fieulaine, S., Morera, S., Poncet, S., Mijakovic, I., Galinier, A., Janin, J., Deutscher, J. & Nessler, S. (2002). X-ray structure of a bifunctional protein kinase in complex with its protein substrate HPr. *Proc Natl Acad Sci U S A* **99**, 13437-13441.

Figge, R. M., Schubert, M., Brinkmann, H. & Cerff, R. (1999). Glyceraldehyde-3-phosphate dehydrogenase gene diversity in eubacteria and eukaryotes: evidence for intra- and inter-kingdom gene transfer. *Mol Biol Evol* **16**, 429-440.

Fillinger, S., Boschi-Muller, S., Azza, S., Dervyn, E., Branlant, G. & Aymerich, S. (2000). Two glyceraldehyde-3-phosphate dehydrogenases with opposite physiological roles in a nonphotosynthetic bacterium. *J Biol Chem* **275**, 14031-14037.

Fisher, S. H. (1993). Utilization of amino acids and other nitrogen-containing compounds. In *Bacillus subtilis and Ohter Gram-Positive Bacteria*, pp. 221-228. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology Press.

Fisher, S. H. & Magasanik, B. (1984). Synthesis of oxaloacetate in *Bacillus subtilis* mutants lacking the 2-ketoglutarate dehydrogenase enzymatic complex. *J Bacteriol* **158**, 55-62.

Fisher, S. H. & Débarbouillé, M. (2002). Nitrogen source utilization and its regulation. In *Bacillus subtilis and its Closest Relatives: from Genes to Cells*, pp. 181-192. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology Press.

Flechtner, V. R. & Hanson, R. S. (1970). Regulation of the tricarboxylic acid cycle in bacteria. A comparison of citrate synthases from different bacteria. *Biochim Biophys Acta* **222**, 253-264.

Fortnagel, P. (1993). Glycolysis. In *Bacillus subtilis and Ohter Gram-Positive Bacteria*, pp. 171-180. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology Press.

Fortnagel, P. & Freese, E. (1968a). Analysis of sporulation mutants. II. Mutants blocked in the citric acid cycle. *J Bacteriol* **95**, 1431-1438.

Fortnagel, P. & Freese, E. (1968b). Inhibition of aconitase by chelation of transition metals causing inhibition of sporulation in *Bacillus subtilis*. *J Biol Chem* **243**, 5289-5295.

Fraenkel, D. G. (1968a). The accumulation of glucose 6-phosphate from glucose and its effect in an *Escherichia coli* mutant lacking phosphoglucose isomerase and glucose 6-phosphate dehydrogenase. *J Biol Chem* **243**, 6451-6457.

Fraenkel, D. G. (1968b). Selection of *Escherichia coli* mutants lacking glucose-6-phosphate dehydrogenase or gluconate-6-phosphate dehydrogenase. *J Bacteriol* **95**, 1267-1271.

Fraenkel, D. G. (1987). Glycolysis, Pentose Phosphate Pathway, and Entner-Doudoroff Pathway. In *Escherichia coli and Salmonella typhimurium: cellular and molecular biology.*, pp. 142-150. Edited by F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, S. M. & H. E. Umbarger. Washington D.C.: American Society for Microbiology.

Fraenkel, D. G. & Horecker, B. L. (1965). Fructose-1, 6-diphosphatase and acid hexose phosphatase of *Escherichia coli*. *J Bacteriol* **90**, 837-842.

Fraenkel, D. G., Falcoz-Kelly, F. & B.L., H. (1964). The utilization of glucose-6-phosphate by glucokinasesless and wild type strains of *Escherichia coli*. *Proc Natl Acad Sci U S A* **52**, 1207-1213.

Freese, E. & Fortnagel, U. (1969). Growth and sporulation of *Bacillus subtilis* mutants blocked in the pyruvate dehydrogenase complex. *J Bacteriol* **99**, 745-756.

Freese, E. & Heinze, J. E. (1984). Metabolic and genetic control of bacterial sporulation. In *The bacterial Spore*, pp. 101-173. Edited by A. Hurst, G. Gould & J. Dring. New York: Academic Press, Inc.

Freese, E., Ichikawa, T., O, Y. K., Freese, E. B. & Prasad, C. (1974). Deficiencies or excesses of metabolites interfering with differentiation. *Proc Natl Acad Sci U S A* **71**, 4188-4193.

Freese, E. B. & Marks, C. L. (1973). Developmental block in citric acid cycle mutants of *Bacillus subtilis*. *J Bacteriol* **116**, 1466-1468.

Fry, B., Zhu, T., Domach, M. M., Koepsel, R. R., Phalakornkule, C. & Ataai, M. M. (2000). Characterization of growth and acid formation in a *Bacillus subtilis* pyruvate kinase mutant. *Appl Environ Microbiol* **66**, 4045-4049.

Fu, D., Libson, A., Miercke, L. J., Weitzman, C., Nollert, P., Krucinski, J. & Stroud, R. M. (2000). Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* **290**, 481-486.

Fujita, Y. & Freese, E. (1981). Isolation and properties of a *Bacillus subtilis* mutant unable to produce fructose-bisphosphatase. *J Bacteriol* **145**, 760-767.

Fujita, Y. & Fujita, T. (1983). Genetic analysis of a pleiotropic deletion mutation (delta igf) in *Bacillus subtilis*. *J Bacteriol* **154**, 864-869.

Fujita, Y. & Miwa, Y. (1994). Catabolite repression of the *Bacillus subtilis gnt* operon mediated by the CcpA protein. *J Bacteriol* **176**, 511-513.

Fujita, Y., Miwa, Y., Galinier, A. & Deutscher, J. (1995). Specific recognition of the *Bacillus subtilis gnt* cis-acting catabolite-responsive element by a protein complex formed between CcpA and seryl-phosphorylated HPr. *Mol Microbiol* **17**, 953-960.

Fujita, Y., Yoshida, K., Miwa, Y., Yanai, N., Nagakawa, E. & Kasahara, Y. (1998). Identification and expression of the *Bacillus subtilis* fructose-1, 6-bisphosphatase gene (*fbp*). *J Bacteriol* **180**, 4309-4313.

Galinier, A., Deutscher, J. & Martin-Verstraete, I. (1999). Phosphorylation of either crh or HPr mediates binding of CcpA to the *Bacillus subtilis xyn cre* and catabolite repression of the xyn operon. *J Mol Biol* **286**, 307-314.

Galinier, A., Haiech, J., Kilhoffer, M. C., Jaquinod, M., Stulke, J., Deutscher, J. & Martin-Verstraete, I. (1997). The Bacillus subtilis crh gene encodes a HPr-like protein involved in carbon catabolite repression. *Proc Natl Acad Sci U S A* **94**, 8439-8444.

Galinier, A., Kravanja, M., Engelmann, R., Hengstenberg, W., Kilhoffer, M. C., Deutscher, J. & Haiech, J. (1998). New protein kinase and protein phosphatase families mediate signal transduction in bacterial catabolite repression. *Proc Natl Acad Sci U S A* **95**, 1823-1828.

Gao, H., Jiang, X., Pogliano, K. & Aronson, A. I. (2002). The E1beta and E2 subunits of the *Bacillus subtilis* pyruvate dehydrogenase complex are involved in regulation of sporulation. *J Bacteriol* **184**, 2780-2788.

Garrido-Pertierra, A. & Cooper, R. A. (1983). Evidence for two distinct pyruvate kinase genes in *Escherichia coli* K-12. *FEBS Lett* **162**, 420-422.

Gay, P. (1979). Le métabolisme vectoriel des glucides et le catabolisme du fructose chez *Bacillus subtilis* Marburg : étude génétique et biochimique. In *Sciences Naturelles*, pp. 175. Paris: Université Pierre et Marie Curie Paris VI.

Geerse, R. H., Izzo, F. & Postma, P. W. (1989). The PEP: fructose phosphotransferase system in *Salmonella typhimurium*: FPr combines enzyme IIIFru and pseudo-HPr activities. *Mol Gen Genet* **216**, 517-525.

Geerse, R. H., Ruig, C. R., Schuitema, A. R. & Postma, P. W. (1986). Relationship between pseudo-HPr and the PEP: fructose phosphotransferase system in *Salmonella typhimurium* and *Escherichia coli*. *Mol Gen Genet* **203**, 435-444.

Giammarinaro, P. & Paton, J. C. (2002). Role of RegM, a homologue of the catabolite repressor protein CcpA, in the virulence of *Streptococcus pneumoniae*. *Infect Immun* **70**, 5454-5461.

Gonzy-Treboul, G. & Steinmetz, M. (1987). Phosphoenolpyruvate:sugar phosphotransferase system of Bacillus subtilis: cloning of the region containing the ptsH and ptsI genes and evidence for a crr-like gene. *J Bacteriol* **169**, 2287-2290.

Gonzy-Treboul, G., Zagorec, M., Rain-Guion, M. C. & Steinmetz, M. (1989). Phosphoenolpyruvate:sugar phosphotransferase system of Bacillus subtilis: nucleotide sequence of ptsX, ptsH and the 5'-end of ptsI and evidence for a ptsHI operon. *Mol Microbiol* **3**, 103-112.

Gonzy-Treboul, G., de Waard, J. H., Zagorec, M. & Postma, P. W. (1991). The glucose permease of the phosphotransferase system of Bacillus subtilis: evidence for IIGlc and IIIGlc domains. *Mol Microbiol* **5**, 1241-1249.

Gosseringer, R., Kuster, E., Galinier, A., Deutscher, J. & Hillen, W. (1997). Cooperative and non-cooperative DNA binding modes of catabolite control protein CcpA from *Bacillus megaterium* result from sensing two different signals. *J Mol Biol* **266**, 665-676.

Greenberg, D. B., Stulke, J. & Saier, M. H., Jr. (2002). Domain analysis of transcriptional regulators bearing PTS regulatory domains. *Res Microbiol* **153**, 519-526.

Groisillier, A. & Lonvaud-Funel, A. (1999). Comparison of partial malolactic enzyme gene sequences for phylogenetic analysis of some lactic acid bacteria species and relationships with the malic enzyme. *Int J Syst Bacteriol* **49 Pt 4**, 1417-1428.

Grundy, F. J., Turinsky, A. J. & Henkin, T. M. (1994). Catabolite regulation of *Bacillus subtilis* acetate and acetoin utilization genes by CcpA. *J Bacteriol* **176**, 4527-4533.

Guérout-Fleury, A. M., Frandsen, N. & Stragier, P. (1996). Plasmids for ectopic integration in *Bacillus subtilis*. *Gene* **180**, 57-61.

Gunnewijk, M. G. & Poolman, B. (2000a). HPr(His approximately P)-mediated phosphorylation differently affects counterflow and proton motive force-driven uptake via the lactose transport protein of Streptococcus thermophilus. *J Biol Chem* **275**, 34080-34085.

Gunnewijk, M. G. & Poolman, B. (2000b). Phosphorylation state of HPr determines the level of expression and the extent of phosphorylation of the lactose transport protein of *Streptococcus thermophilus*. *J Biol Chem* **275**, 34073-34079.

Gunsalus, I. C. & Shuster, C. W. (1961). Energy-yielding metabolism in Bacteria. *The Bacteria* **2**, 1-28.

Hansen, T., Reichstein, B., Schmid, R. & Schonheit, P. (2002). The first archaeal ATP-dependent glucokinase, from the hyperthermophilic crenarchaeon *Aeropyrum pernix*, represents a monomeric, extremely thermophilic ROK glucokinase with broad hexose specificity. *J Bacteriol* **184**, 5955-5965.

Hederstedt, L. (1993). The Krebs Citric Acid Cycle. In *Bacillus subtilis and Ohter Gram-Positive Bacteria*, pp. 181-198. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology Press.

Heller, K. B., Lin, E. C. & Wilson, T. H. (1980). Substrate specificity and transport properties of the glycerol facilitator of *Escherichia coli*. *J Bacteriol* **144**, 274-278.

Hemila, H., Palva, A., Paulin, L., Arvidson, S. & Palva, I. (1990). Secretory S complex of *Bacillus subtilis*: sequence analysis and identity to pyruvate dehydrogenase. *J Bacteriol* **172**, 5052-5063.

Henkin, T. M. & Yanofsky, C. (2002). Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. *Bioessays* **24**, 700-707.

Henkin, T. M., Grundy, F. J., Nicholson, W. L. & Chambliss, G. H. (1991). Catabolite repression of alpha-amylase gene expression in *Bacillus subtilis* involves a trans-acting gene product homologous to the *Escherichia coli lacI* and *galR* repressors. *Mol Microbiol* **5**, 575-584.

Henstra, S. A., Duurkens, R. H. & Robillard, G. T. (2000). Multiple phosphorylation events regulate the activity of the mannitol transcriptional regulator MtlR of the *Bacillus stearothermophilus* phosphoenolpyruvate-dependent mannitol phosphotransferase system. *J Biol Chem* **275**, 7037-7044.

Henstra, S. A., Tuinhof, M., Duurkens, R. H. & Robillard, G. T. (1999). The *Bacillus stearothermophilus* mannitol regulator, MtlR, of the phosphotransferase system. A DNA-binding protein, regulated by HPr and iicbmtl-dependent phosphorylation. *J Biol Chem* **274**, 4754-4763.

Herzberg, O., Reddy, P., Sutrina, S., Saier, M. H., Jr., Reizer, J. & Kapadia, G. (1992). Structure of the histidine-containing phosphocarrier protein HPr from *Bacillus subtilis* at 2.0-A resolution. *Proc Natl Acad Sci U S A* **89**, 2499-2503.

Hodgson, J. A., Lowe, P. N. & Perham, R. N. (1983). Wild-type and mutant forms of the pyruvate dehydrogenase multienzyme complex from *Bacillus subtilis*. *Biochem J* **211**, 463-472.

Hogema, B. M., Arents, J. C., Bader, R., Eijkemans, K., Yoshida, H., Takahashi, H., Aiba, H. & Postma, P. W. (1998). Inducer exclusion in *Escherichia coli* by non-PTS substrates: the role of the PEP to pyruvate ratio in determining the phosphorylation state of enzyme IIAGlc. *Mol Microbiol* **30**, 487-498.

Holms, W. H. & Bennett, P. M. (1971). Regulation of isocitrate dehydrogenase activity in *Escherichia coli* on adaptation to acetate. *J Gen Microbiol* **65**, 57-68.

Hopper, D. J. & Cooper, R. A. (1971). The regulation of *Escherichia coli* methylglyoxal synthase; a new control site in glycolysis? *FEBS Lett* **13**, 213-216.

Houman, F., Diaz-Torres, M. R. & Wright, A. (1990). Transcriptional antitermination in the *bgl* operon of *E. coli* is modulated by a specific RNA binding protein. *Cell* **62**, 1153-1163.

Hove-Jensen, B. & Maigaard, M. (1993). *Escherichia coli rpiA* gene encoding ribose phosphate isomerase A. *J Bacteriol* **175**, 5628-5635.

Hsu, R. Y. (1982). Pigeon liver malic enzyme. *Mol Cell Biochem* **43**, 3-26.

Hueck, C. J. & Hillen, W. (1995). Catabolite repression in *Bacillus subtilis*: a global regulatory mechanism for the gram-positive bacteria? *Mol Microbiol* **15**, 395-401.

Hueck, C. J., Hillen, W. & Saier, M. H., Jr. (1994). Analysis of a cis-active sequence mediating catabolite repression in gram-positive bacteria. *Res Microbiol* **145**, 503-518.

Hueck, C. J., Kraus, A., Schmiedel, D. & Hillen, W. (1995). Cloning, expression and functional analyses of the catabolite control protein CcpA from *Bacillus megaterium*. *Mol Microbiol* **16**, 855-864.

Hughes, M. J., Moore, J. C., Lane, J. D., Wilson, R., Pribul, P. K., Younes, Z. N., Dobson, R. J., Everest, P., Reason, A. J., Redfern, J. M., Greer, F. M., Paxton, T., Panico, M., Morris, H. R., Feldman, R. G. & Santangelo, J. D. (2002). Identification of major outer surface proteins of *Streptococcus agalactiae*. *Infect Immun* **70**, 1254-1259.

Hurley, J. H., Faber, H. R., Worthylake, D., Meadow, N. D., Roseman, S., Pettigrew, D. W. & Remington, S. J. (1993). Structure of the regulatory complex of *Escherichia coli* IIIGlc with glycerol kinase. *Science* **259**, 673-677.

Huynen, M. A., Dandekar, T. & Bork, P. (1999). Variation and evolution of the citric-acid cycle: a genomic perspective. *Trends Microbiol* **7**, 281-291.

Iddar, A., Serrano, A. & Soukri, A. (2002). A phosphate-stimulated NAD(P)+-dependent glyceraldehyde-3-phosphate dehydrogenase in *Bacillus cereus*. *FEMS Microbiol Lett* **211**, 29-35.

Idelson, M. & Amster-Choder, O. (1998). SacY, a transcriptional antiterminator from *Bacillus subtilis*, is regulated by phosphorylation *in vivo*. *J Bacteriol* **180**, 660-666.

Irani, M. H. & Maitra, P. K. (1977). Properties of *Escherichia coli* mutants deficient in enzymes of glycolysis. *J Bacteriol* **132**, 398-410.

Ireton, K., Jin, S., Grossman, A. D. & Sonenshein, A. L. (1995). Krebs cycle function is required for activation of the Spo0A transcription factor in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **92**, 2845-2849.

Ito, M., Guffanti, A. A., Wang, W. & Krulwich, T. A. (2000). Effects of nonpolar mutations in each of the seven *Bacillus subtilis mrp* genes suggest complex interactions among the gene products in support of Na(+) and alkali but not cholate resistance. *J Bacteriol* **182**, 5663-5670.

Iverson, T. M., Luna-Chavez, C., Cecchini, G. & Rees, D. C. (1999). Structure of the *Escherichia coli* fumarate reductase respiratory complex. *Science* **284**, 1961-1966.

Izu, H., Adachi, O. & Yamada, M. (1996). Purification and characterization of the *Escherichia coli* thermoresistant glucokinase encoded by the gntK gene. *FEBS Lett* **394**, 14-16.

Janausch, I. G., Zientz, E., Tran, Q. H., Kroger, A. & Unden, G. (2002). C4-dicarboxylate carriers and sensors in bacteria. *Biochim Biophys Acta* **1553**, 39-56.

Jault, J. M., Fieulaine, S., Nessler, S., Gonzalo, P., Di Pietro, A., Deutscher, J. & Galinier, A. (2000). The HPr kinase from *Bacillus subtilis* is a homo-oligomeric enzyme which exhibits strong positive cooperativity for nucleotide and fructose 1,6-bisphosphate binding. *J Biol Chem* **275**, 1773-1780.

Jedrzejas, M. J. (2000). Structure, function, and evolution of phosphoglycerate mutases: comparison with fructose-2,6-bisphosphatase, acid phosphatase, and alkaline phosphatase. *Prog Biophys Mol Biol* **73**, 263-287.

Jedrzejas, M. J., Chander, M., Setlow, P. & Krishnasamy, G. (2000). Structure and mechanism of action of a novel phosphoglycerate mutase from *Bacillus stearothermophilus*. *Embo J* **19**, 1419-1431.

Jensen, M. O., Tajkhorshid, E. & Schulten, K. (2001). The mechanism of glycerol conduction in aquaglyceroporins. *Structure (Camb)* **9**, 1083-1093.

Jia, Z., Vandonselaar, M., Hengstenberg, W., Quail, J. W. & Delbaere, L. T. (1994). The 1.6 A structure of histidine-containing phosphotransfer protein HPr from *Streptococcus faecalis*. *J Mol Biol* **236**, 1341-1355.

Jiang, X. & Doyle, M. P. (2000). Growth supplements for *Helicobacter pylori*. *J Clin Microbiol* **38**, 1984-1987.

Jin, S. & Sonenshein, A. L. (1994). Identification of two distinct *Bacillus subtilis* citrate synthase genes. *J Bacteriol* **176**, 4669-4679.

Jin, S. & Sonenshein, A. L. (1996). Characterization of the major citrate synthase of *Bacillus subtilis*. *J Bacteriol* **178**, 3658-3660.

Jin, S., De Jesus-Berrios, M. & Sonenshein, A. L. (1996). A *Bacillus subtilis* malate dehydrogenase gene. *J Bacteriol* **178**, 560-563.

Jin, S., Levin, P. A., Matsuno, K., Grossman, A. D. & Sonenshein, A. L. (1997). Deletion of the *Bacillus subtilis* isocitrate dehydrogenase gene causes a block at stage I of sporulation. *J Bacteriol* **179**, 4725-4732.

Jones, B. E., Dossonnet, V., Kuster, E., Hillen, W., Deutscher, J. & Klevit, R. E. (1997). Binding of the catabolite repressor protein CcpA to its DNA target is regulated by phosphorylation of its corepressor HPr. *J Biol Chem* **272**, 26530-26535.

Jordan, P. A., Tang, Y., Bradbury, A. J., Thomson, A. J. & Guest, J. R. (1999). Biochemical and spectroscopic characterization of *Escherichia coli* aconitases (AcnA and AcnB). *Biochem J* **344 Pt 3**, 739-746.

Josephson, B. L. & Fraenkel, D. G. (1969). Transketolase mutants of *Escherichia coli*. *J Bacteriol* **100**, 1289-1295.

Josephson, B. L. & Fraenkel, D. G. (1974). Sugar metabolism in transketolase mutants of *Escherichia coli*. *J Bacteriol* **118**, 1082-1089.

Jourlin-Castelli, C., Mani, N., Nakano, M. M. & Sonenshein, A. L. (2000). CcpC, a novel regulator of the LysR family required for glucose repression of the *citB* gene in *Bacillus subtilis*. *J Mol Biol* **295**, 865-878.

Kalbitzer, H. R. & Hengstenberg, W. (1993). The solution structure of the histidine-containing protein (HPr) from *Staphylococcus aureus* as determined by two-dimensional 1H-NMR spectroscopy. *Eur J Biochem* **216**, 205-214.

Kalbitzer, H. R., Hengstenberg, W., Rosch, P., Muss, P., Bernsmann, P., Engelmann, R., Dorschug, M. & Deutscher, J. (1982). HPr proteins of different microorganisms studied by hydrogen-1 high-resolution nuclear magnetic resonance: similarities of structures and mechanisms. *Biochemistry* **21**, 2879-2885.

Kashket, E. R. (1983). Stoichiometry of the H+-ATPase of *Escherichia coli* cells during anaerobic growth. *FEBS Lett* **154**, 343-346.

Kataoka, M., Sri Rohani, L. P., Wada, M., Kita, K., Yanase, H., Urabe, I. & Shimizu, S. (1998). *Escherichia coli* transformant expressing the glucose dehydrogenase gene from *Bacillus megaterium* as a cofactor regenerator in a chiral alcohol production system. *Biosci Biotechnol Biochem* **62**, 167-169.

Kim, H. J., Roux, A. & Sonenshein, A. L. (2002a). Direct and indirect roles of CcpA in regulation of *Bacillus subtilis* Krebs cycle genes. *Mol Microbiol* **45**, 179-190.

Kim, H. J., Jourlin-Castelli, C., Kim, S. I. & Sonenshein, A. L. (2002b). Regulation of the *Bacillus subtilis ccpC* gene by *ccpA* and *ccpC*. *Mol Microbiol* **43**, 399-410.

Kim, J. H., Voskuil, M. I. & Chambliss, G. H. (1998). NADP, corepressor for the *Bacillus* catabolite control protein CcpA. *Proc Natl Acad Sci U S A* **95**, 9590-9595.

Kim, J. H., Guvener, Z. T., Cho, J. Y., Chung, K. C. & Chambliss, G. H. (1995). Specificity of DNA binding activity of the *Bacillus subtilis* catabolite control protein CcpA. *J Bacteriol* **177**, 5129-5134.

Klein, M., Sprenger, G. A. & Freudl, R. (1996). Cloning, nucleotide sequence, and functional expression of the *Escherichia coli* enolase (*eno*) gene in a temperature-sensitive eno mutant strain. *DNA Seq* **6**, 351-355.

Klevit, R. E. & Waygood, E. B. (1986). Two-dimensional 1H NMR studies of histidine-containing protein from *Escherichia coli*. 3. Secondary and tertiary structure as determined by NMR. *Biochemistry* **25**, 7774-7781.

Knezevic, I., Bachem, S., Sickmann, A., Meyer, H. E., Stulke, J. & Hengstenberg, W. (2000). Regulation of the glucose-specific phosphotransferase system (PTS) of *Staphylococcus carnosus* by the antiterminator protein GlcT. *Microbiology* **146 (Pt 9)**, 2333-2342.

Knowles, V. L., Smith, C. S., Smith, C. R. & Plaxton, W. C. (2001). Structural and regulatory properties of pyruvate kinase from the *Cyanobacterium synechococcus* PCC 6301. *J Biol Chem* **276**, 20966-20972.

Kobayashi, K., Ogura, M., Yamaguchi, H., Yoshida, K., Ogasawara, N., Tanaka, T. & Fujita, Y. (2001). Comprehensive DNA microarray analysis of *Bacillus subtilis* two-component regulatory systems. *J Bacteriol* **183**, 7365-7370.

Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessières, P., Boland, F., Brignell, S. C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L. C., Danchin, A., Débarbouillé, M., Dervyn, E., Deuerling, E., Devine, K., Devine, S. K., Dreesen, O., Errington, J., Fillinger, S., Foster, S. J., Fujita, Y., Galizzi, A., Gardan, R., Eschevins, C., Fukushima, T., Haga, K., Harwood, C. R., Hecker, M., Hosoya, D., Hullo, M. F., Kakeshita, H., Karamata, D., Kasahara, Y., Kawamura, F., Koga, K., Koski, P., Kuwana, R., Imamura, D., Ishimaru, M., Ishikawa, S., Ishio, I., Le Coq, D., Mason, A., Mauël, C., Meima, R., Mellado, R. P., Moir, A., Moriya, S., Nagakawa, E., Nanamiya, H., Nakai, S., Nygaard, P., Ogura, M., Ohanan, T., O'Reilly, M., O'Rourke, M., Pragai, Z., Pooley, H. M., Rapoport, G., Rawlins, J. P., Rivas, L. A., Rivolta, C., Sadaie, A., Sadaie, Y., Sarvas, M., Sato, T., Saxild, H. H., Scanlan, E., Schumann, W., Seegers, J. F. M. L., Sekiguchi, J., Sekowska, A., Séror, S., Simon, M., Stragier, P., Studer, R., Takamatsu, H., Tanaka, T., Takeuchi, M., Thomaides, H. B., Vagner, V., van Dijl, J. M., Watabe, K., Wipat, A., Yamamoto, H., Yamamoto, M., Yamamoto, Y., Yamane, K., Yata, K., Yoshida, K., Yoshikawa, H., Zuber, U. & Ogasawara, N. (2003 sous presse). Essential *Bacillus subtilis* genes. *Proc Natl Acad Sci U S A*.

Koksharova, O., Schubert, M., Shestakov, S. & 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**, 183-194.

Kornberg, H., Watts, P. D. & Brown, K. (1980). Mechanisms of 'inducer exclusion' by glucose. *FEBS Lett* **117 Suppl**, K28-36.

Kraft, L., Sprenger, G. A. & Lindqvist, Y. (2001). Crystallization and preliminary X-ray crystallographic studies of recombinant thermoresistant gluconate kinase GntK from *Escherichia coli*. *Acta Crystallogr D Biol Crystallogr* **57**, 1159-1161.

Kraft, L., Sprenger, G. A. & Lindqvist, Y. (2002). Conformational changes during the catalytic cycle of gluconate kinase as revealed by X-ray crystallography. *J Mol Biol* **318**, 1057-1069.

Kraus, A., Hueck, C., Gartner, D. & Hillen, W. (1994). Catabolite repression of the *Bacillus subtilis xyl* operon involves a cis element functional in the context of an unrelated sequence, and glucose exerts additional *xylR*-dependent repression. *J Bacteriol* **176**, 1738-1745.

Kraus, A., Kuster, E., Wagner, A., Hoffmann, K. & Hillen, W. (1998). Identification of a co-repressor binding site in catabolite control protein CcpA. *Mol Microbiol* **30**, 955-963.

Kravanja, M., Engelmann, R., Dossonnet, V., Bluggel, M., Meyer, H. E., Frank, R., Galinier, A., Deutscher, J., Schnell, N. & Hengstenberg, W. (1999). The *hprK* gene of *Enterococcus faecalis* encodes a novel bifunctional enzyme: the HPr kinase/phosphatase. *Mol Microbiol* **31**, 59-66.

Kruger, S. & Hecker, M. (1995). Regulation of the putative bglPH operon for aryl-beta-glucoside utilization in Bacillus subtilis. *J Bacteriol* **177**, 5590-5597.

Krüger, S. & Hecker, M. (1995). Regulation of the putative *bglPH* operon for aryl-beta-glucoside utilization in *Bacillus subtilis*. *J Bacteriol* **177**, 5590-5597.

Kuhnel, K. & Luisi, B. F. (2001). Crystal structure of the *Escherichia coli* RNA degradosome component enolase. *J Mol Biol* **313**, 583-592.

Kundig, W., Kundig, F. D., Anderson, B. & Roseman, S. (1966). Restoration of active transport of glycosides in *Escherichia coli* by a component of a phosphotransferase system. *J Biol Chem* **241**, 3243-3246.

Kupor, S. R. & Fraenkel, D. G. (1969). 6-phosphogluconolactonase mutants of *Escherichia coli* and a maltose blue gene. *J Bacteriol* **100**, 1296-1301.

Kupor, S. R. & Fraenkel, D. G. (1972). Glucose metabolism in 6 phosphogluconolactonase mutants of *Escherichia coli*. *J Biol Chem* **247**, 1904-1910.

Kuster, E., Luesink, E. J., de Vos, W. M. & Hillen, W. (1996). Immunological crossreactivity to the catabolite control protein CcpA *Bacillus megaterium* is found in many gram-positive bacteria. *FEMS Microbiol Lett* **139**, 109-115.

Lavergne, J. P., Jault, J. M. & Galinier, A. (2002). Insights into the functioning of *Bacillus subtilis* HPr kinase/phosphatase: affinity for its protein substrates and role of cations and phosphate. *Biochemistry* **41**, 6218-6225.

le Coq, D., Fillinger, S. & Aymerich, S. (1999). Histidinol phosphate phosphatase, catalyzing the penultimate step of the histidine biosynthesis pathway, is encoded by *ytvP* (*hisJ*) in *Bacillus subtilis*. *J Bacteriol* **181**, 3277-3280.

le Coq, D., Lindner, C., Kruger, S., Steinmetz, M. & Stulke, J. (1995). New beta-glucoside (*bgl*) genes in *Bacillus subtilis*: the bglP gene product has both transport and regulatory functions similar to those of BglF, its *Escherichia coli* homolog. *J Bacteriol* **177**, 1527-1535.

Leboeuf, C., Auffray, Y. & Hartke, A. (2000a). Cloning, sequencing and characterization of the *ccpA* gene from *Enterococcus faecalis*. *Int J Food Microbiol* **55**, 109-113.

Leboeuf, C., Leblanc, L., Auffray, Y. & Hartke, A. (2000b). Characterization of the *ccpA* gene of *Enterococcus faecalis*: identification of starvation-inducible proteins regulated by ccpA. *J Bacteriol* **182**, 5799-5806.

Lengeler, J. & Jahreis, K. (1996). Phosphotransferase systems pr PTSs as carbohydrate transport and as signal transduction systems. In *Handbook of Biological Physics*, pp. 573-598. Edited by W. Konings, H. Kaback & J. Lolkema.

Lengeler, J. W., Jahreis, K. & Wehmeier, U. F. (1994). Enzymes II of the phospho enol pyruvate-dependent phosphotransferase systems: their structure and function in carbohydrate transport. *Biochim Biophys Acta* **1188**, 1-28.

Lengeler, J. W., Titgemeyer, F., Vogler, A. P. & Wohrl, B. M. (1990). Structures and homologies of carbohydrate: phosphotransferase system (PTS) proteins. *Philos Trans R Soc Lond B Biol Sci* **326**, 489-504.

Leong-Morgenthaler, P., Zwahlen, M. C. & Hottinger, H. (1991). Lactose metabolism in *Lactobacillus bulgaricus*: analysis of the primary structure and expression of the genes involved. *J Bacteriol* **173**, 1951-1957.

Levashov, P., Orlov, V., Boschi-Muller, S., Talfournier, F., Asryants, R., Bulatnikov, I., Muronetz, V., Branlant, G. & Nagradova, N. (1999). Thermal unfolding of phosphorylating D-glyceraldehyde-3-phosphate dehydrogenase studied by differential scanning calorimetry. *Biochim Biophys Acta* **1433**, 294-306.

Leyva-Vazquez, M. A. & Setlow, P. (1994). Cloning and nucleotide sequences of the genes encoding triose phosphate isomerase, phosphoglycerate mutase, and enolase from *Bacillus subtilis*. *J Bacteriol* **176**, 3903-3910.

Lindner, C., Galinier, A., Hecker, M. & Deutscher, J. (1999). Regulation of the activity of the *Bacillus subtilis* antiterminator LicT by multiple PEP-dependent, enzyme I- and HPr-catalysed phosphorylation. *Mol Microbiol* **31**, 995-1006.

Lovell, S. C., Mullick, A. H. & Muirhead, H. (1998). Cooperativity in *Bacillus stearothermophilus* pyruvate kinase. *J Mol Biol* **276**, 839-851.

Ludwig, H., Meinken, C., Matin, A. & Stulke, J. (2002a). Insufficient Expression of the *ilv-leu* Operon Encoding Enzymes of Branched-Chain Amino Acid Biosynthesis Limits Growth of a *Bacillus subtilis ccpA* Mutant. *J Bacteriol* **184**, 5174-5178.

Ludwig, H., Rebhan, N., Blencke, H. M., Merzbacher, M. & Stulke, J. (2002b). Control of the glycolytic *gapA* operon by the catabolite control protein A in *Bacillus subtilis*: a novel mechanism of CcpA-mediated regulation. *Mol Microbiol* **45**, 543-553.

Ludwig, H., Homuth, G., Schmalisch, M., Dyka, F. M., Hecker, M. & 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**, 409-422.

Luesink, E. J., Beumer, C. M., Kuipers, O. P. & De Vos, W. M. (1999). Molecular characterization of the *Lactococcus lactis* ptsHI operon and analysis of the regulatory role of HPr. *J Bacteriol* **181**, 764-771.

Luesink, E. J., van Herpen, R. E., Grossiord, B. P., Kuipers, O. P. & de Vos, W. M. (1998). Transcriptional activation of the glycolytic las operon and catabolite repression of the gal operon in *Lactococcus lactis* are mediated by the catabolite control protein CcpA. *Mol Microbiol* **30**, 789-798.

Lyngstadaas, A., Sprenger, G. A. & Boye, E. (1998). Impaired growth of an *Escherichia coli rpe* mutant lacking ribulose-5-phosphate epimerase activity. *Biochim Biophys Acta* **1381**, 319-330.

Mahadevan, S. & Wright, A. (1987). A bacterial gene involved in transcription antitermination: regulation at a rho-independent terminator in the *bgl* operon of *E. coli*. *Cell* **50**, 485-494.

Mahadevan, S., Reynolds, A. E. & Wright, A. (1987). Positive and negative regulation of the *bgl* operon in *Escherichia coli*. *J Bacteriol* **169**, 2570-2578.

Mahr, K., Hillen, W. & Titgemeyer, F. (2000). Carbon catabolite repression in *Lactobacillus pentosus*: analysis of the ccpA region. *Appl Environ Microbiol* **66**, 277-283.

Manival, X., Yang, Y., Strub, M. P., Kochoyan, M., Steinmetz, M. & Aymerich, S. (1997). From genetic to structural characterization of a new class of RNA-binding domain within the SacY/BglG family of antiterminator proteins. *Embo J* **16**, 5019-5029.

Martin-Verstraete, I., Deutscher, J. & Galinier, A. (1999). Phosphorylation of HPr and Crh by HprK, early steps in the catabolite repression signalling pathway for the *Bacillus subtilis* levanase operon. *J Bacteriol* **181**, 2966-2969.

Martin-Verstraete, I., Debarbouille, M., Klier, A. & Rapoport, G. (1990). Levanase operon of *Bacillus subtilis* includes a fructose-specific phosphotransferase system regulating the expression of the operon. *J Mol Biol* **214**, 657-671.

Martin-Verstraete, I., Debarbouille, M., Klier, A. & Rapoport, G. (1992). Mutagenesis of the *Bacillus subtilis* "-12, -24" promoter of the levanase operon and evidence for the existence of an upstream activating sequence. *J Mol Biol* **226**, 85-99.

Martin-Verstraete, I., Debarbouille, M., Klier, A. & Rapoport, G. (1994). Interactions of wild-type and truncated LevR of *Bacillus subtilis* with the upstream activating sequence of the levanase operon. *J Mol Biol* **241**, 178-192.

Martin-Verstraete, I., Stulke, J., Klier, A. & Rapoport, G. (1995). Two different mechanisms mediate catabolite repression of the *Bacillus subtilis* levanase operon. *J Bacteriol* **177**, 6919-6927.

Martin-Verstraete, I., Charrier, V., Stulke, J., Galinier, A., Erni, B., Rapoport, G. & Deutscher, J. (1998). Antagonistic effects of dual PTS-catalysed phosphorylation on the *Bacillus subtilis* transcriptional activator LevR. *Mol Microbiol* **28**, 293-303.

Matsuno, K., Blais, T., Serio, A. W., Conway, T., Henkin, T. M. & Sonenshein, A. L. (1999). Metabolic imbalance and sporulation in an isocitrate dehydrogenase mutant of *Bacillus subtilis*. *J Bacteriol* **181**, 3382-3391.

Meinken, C., Blencke, H. M., Ludwig, H. & Stülke, J. (2003 sous presse). Expression of the glycolytic *gapA* operon in *Bacillus subtilis*: mRNA processing contributes to differential synthesis of proteins encoded by the operon. *Microbiology*.

Meyer, D., Schneider-Fresenius, C., Horlacher, R., Peist, R. & Boos, W. (1997). Molecular characterization of glucokinase from *Escherichia coli* K-12. *J Bacteriol* **179**, 1298-1306.

Mijakovic, I., Poncet, S., Galinier, A., Monedero, V., Fieulaine, S., Janin, J., Nessler, S., Marquez, J. A., Scheffzek, K., Hasenbein, S., Hengstenberg, W. & Deutscher, J. (2002). Pyrophosphate-producing protein dephosphorylation by HPr kinase/phosphorylase: A relic of early life? *Proc Natl Acad Sci U S A*.

Miles, J. S. & Guest, J. R. (1987). Molecular genetic aspects of the citric acid cycle of *Escherichia coli*. *Biochem Soc Symp* **54**, 45-65.

Miller, S. P., Karschnia, E. J., Ikeda, T. P. & LaPorte, D. C. (1996). Isocitrate dehydrogenase kinase/phosphatase. Kinetic characteristics of the wild-type and two mutant proteins. *J Biol Chem* **271**, 19124-19128.

Miller, S. P., Chen, R., Karschnia, E. J., Romfo, C., Dean, A. & LaPorte, D. C. (2000). Locations of the regulatory sites for isocitrate dehydrogenase kinase/phosphatase. *J Biol Chem* **275**, 833-839.

Misset, O., Brouwer, M. & Robillard, G. T. (1980). *Escherichia coli* phosphoenolpyruvate-dependent phosphotransferase system. Evidence that the dimer is the active form of enzyme I. *Biochemistry* **19**, 883-890.

Mitchell, C., Morris, P. W., Lum, L., Spiegelman, G. & Vary, J. C. (1992). The amino acid sequence of a Bacillus subtilis phosphoprotein that matches an orfY-tsr coding sequence. *Mol Microbiol* **6**, 1345-1349.

Miwa, Y. & Fujita, Y. (2001). Involvement of two distinct catabolite-responsive elements in catabolite repression of the *Bacillus subtilis* myo-inositol (*iol*) operon. *J Bacteriol* **183**, 5877-5884.

Miwa, Y., Saikawa, M. & Fujita, Y. (1994). Possible function and some properties of the CcpA protein of *Bacillus subtilis*. *Microbiology* **140 (Pt 10)**, 2567-2575.

Miwa, Y., Nakata, A., Ogiwara, A., Yamamoto, M. & Fujita, Y. (2000). Evaluation and characterization of catabolite-responsive elements (cre) of *Bacillus subtilis*. *Nucleic Acids Res* **28**, 1206-1210.

Miwa, Y., Nagura, K., Eguchi, S., Fukuda, H., Deutscher, J. & Fujita, Y. (1997). Catabolite repression of the *Bacillus subtilis gnt* operon exerted by two catabolite-responsive elements. *Mol Microbiol* **23**, 1203-1213.

Monedero, V., Gosalbes, M. J. & Perez-Martinez, G. (1997). Catabolite repression in *Lactobacillus casei* ATCC 393 is mediated by CcpA. *J Bacteriol* **179**, 6657-6664.

Monedero, V., Poncet, S., Mijakovic, I., Fieulaine, S., Dossonnet, V., Martin-Verstraete, I., Nessler, S. & Deutscher, J. (2001). Mutations lowering the phosphatase activity of HPr kinase/phosphatase switch off carbon metabolism. *Embo J* **20**, 3928-3937.

Monod, J. (1942). Recherches sur la croissance des cultures bactériennes.: Université de Paris, France.

Moreno, M. S., Schneider, B. L., Maile, R. R., Weyler, W. & Saier, M. H., Jr. (2001). Catabolite repression mediated by the CcpA protein in *Bacillus subtilis*: novel modes of regulation revealed by whole-genome analyses. *Mol Microbiol* **39**, 1366-1381.

Moszer, I., Glaser, P. & Danchin, A. (1995). SubtiList: a relational database for the *Bacillus subtilis* genome. *Microbiology* **141 (Pt 2)**, 261-268.

Muir, M., Williams, L. & Ferenci, T. (1985). Influence of transport energization on the growth yield of *Escherichia coli*. *J Bacteriol* **163**, 1237-1242.

Musrati, R. A., Kollarova, M., Mernik, N. & Mikulasova, D. (1998). Malate dehydrogenase: distribution, function and properties. *Gen Physiol Biophys* **17**, 193-210.

Nagy, E., Henics, T., Eckert, M., Miseta, A., Lightowlers, R. N. & Kellermayer, M. (2000). Identification of the NAD(+)-binding fold of glyceraldehyde-3-phosphate dehydrogenase as a novel RNA-binding domain. *Biochem Biophys Res Commun* **275**, 253-260.

Nakano, M. M. & Hulett, F. M. (1997). Adaptation of *Bacillus subtilis* to oxygen limitation. *FEMS Microbiol Lett* **157**, 1-7.

Nakano, M. M. & Zuber, P. (1998). Anaerobic growth of a "strict aerobe" (*Bacillus subtilis*). *Annu Rev Microbiol* **52**, 165-190.

Nakano, M. M., Zuber, P. & Sonenshein, A. L. (1998). Anaerobic regulation of *Bacillus subtilis* Krebs cycle genes. *J Bacteriol* **180**, 3304-3311.

Nakano, M. M., Dailly, Y. P., Zuber, P. & Clark, D. P. (1997). Characterization of anaerobic fermentative growth of *Bacillus subtilis*: identification of fermentation end products and genes required for growth. *J Bacteriol* **179**, 6749-6755.

Nakano, M. M., Zhu, Y., Haga, K., Yoshikawa, H., Sonenshein, A. L. & Zuber, P. (1999). A mutation in the 3-phosphoglycerate kinase gene allows anaerobic growth of *Bacillus subtilis* in the absence of ResE kinase. *J Bacteriol* **181**, 7087-7097.

Napper, S., Anderson, J. W., Georges, F., Quail, J. W., Delbaere, L. T. & Waygood, E. B. (1996). Mutation of serine-46 to aspartate in the histidine-containing protein of *Escherichia coli* mimics the inactivation by phosphorylation of serine-46 in HPrs from gram-positive bacteria. *Biochemistry* **35**, 11260-11267.

Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., M., S. & Umbarger, H. E. (1987). *Escherichia coli and Salmonella typhimurium: cellular and molecular biology.* Washington D.C.: American Society for Microbiology.

Nelson, S. O., Lengeler, J. & Postma, P. W. (1984). Role of IIIGlc of the phosphoenolpyruvate-glucose phosphotransferase system in inducer exclusion in *Escherichia coli*. *J Bacteriol* **160**, 360-364.

Neves, A. R., Ramos, A., Nunes, M. C., Kleerebezem, M., Hugenholtz, J., de Vos, W. M., Almeida, J. & Santos, H. (1999). In vivo nuclear magnetic resonance studies of glycolytic kinetics in *Lactococcus lactis*. *Biotechnol Bioeng* **64**, 200-212.

Nicholson, W. L. & Chambliss, G. H. (1985). Isolation and characterization of a cis-acting mutation conferring catabolite repression resistance to alpha-amylase synthesis in *Bacillus subtilis*. *J Bacteriol* **161**, 875-881.

Nicholson, W. L. & Chambliss, G. H. (1986). Molecular cloning of cis-acting regulatory alleles of the *Bacillus subtilis amyR* region by using gene conversion transformation. *J Bacteriol* **165**, 663-670.

Nicholson, W. L., Park, Y. K., Henkin, T. M., Won, M., Weickert, M. J., Gaskell, J. A. & Chambliss, G. H. (1987). Catabolite repression-resistant mutations of the *Bacillus subtilis* alpha-amylase promoter affect transcription levels and are in an operator-like sequence. *J Mol Biol* **198**, 609-618.

Nimmo, H. G. (1987). The tricarboxylic acid cycle and anaplerotic reactions. In *Escherichia coli and Salmonella typhimurium: cellular and molecular biology.*, pp. 156-169. Edited by F. C. Neidhardt, J. L.

Ingraham, K. B. Low, B. Magasanik, S. M. & H. E. Umbarger. Washington D.C.: American Society for Microbiology.

Nollert, P., Harries, W. E., Fu, D., Miercke, L. J. & Stroud, R. M. (2001). Atomic structure of a glycerol channel and implications for substrate permeation in aqua(glycero)porins. *FEBS Lett* **504**, 112-117.

Novotny, M. J., Frederickson, W. L., Waygood, E. B. & Saier, M. H., Jr. (1985). Allosteric regulation of glycerol kinase by enzyme IIIglc of the phosphotransferase system in *Escherichia coli* and *Salmonella typhimurium*. *J Bacteriol* **162**, 810-816.

Ohne, M. (1975). Regulation of the dicarboxylic acid part of the citric acid cycle in *Bacillus subtilis*. *J Bacteriol* **122**, 224-234.

Osteras, M., Driscoll, B. T. & Finan, T. M. (1997). Increased pyruvate orthophosphate dikinase activity results in an alternative gluconeogenic pathway in *Rhizobium* (*Sinorhizobium*) meliloti. *Microbiology* **143 (Pt 5)**, 1639-1648.

Osumi, T. & Saier, M. H., Jr. (1982a). Mechanism of regulation of the lactose permease by the phosphotransferase system in *Escherichia coli*: evidence for protein-protein interaction. *Ann Microbiol (Paris)* **133**, 269-273.

Osumi, T. & Saier, M. H., Jr. (1982b). 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**, 1457-1461.

Pan, W. (2002). A comparative review of statistical methods for discovering differentially expressed genes in replicated microarray experiments. *Bioinformatics* **18**, 546-554.

Paulsen, I. T., Chauvaux, S., Choi, P. & Saier, M. H., Jr. (1998). Characterization of glucose-specific catabolite repression-resistant mutants of *Bacillus subtilis*: identification of a novel hexose:H+ symporter. *J Bacteriol* **180**, 498-504.

Paulsen, I. T., Nguyen, L., Sliwinski, M. K., Rabus, R. & Saier, M. H., Jr. (2000). Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes. *J Mol Biol* **301**, 75-100.

Pearson, C. L., Loshon, C. A., Pedersen, L. B., Setlow, B. & Setlow, P. (2000). Analysis of the function of a putative 2,3-diphosphoglyceric acid-dependent phosphoglycerate mutase from *Bacillus subtilis*. *J Bacteriol* **182**, 4121-4123.

Perego, M. & Hoch, J. A. (2002). Two-component systems, phosphorelays, and regulation of their activities by phosphatases. In *Bacillus subtilis and its Closest Relatives: from Genes to Cells*, pp. 423-481. Edited by R. Losick. Washington, DC: American Society for Microbiology Press.

Perham, R. N. (1991). Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes: a paradigm in the design of a multifunctional protein. *Biochemistry* **30**, 8501-8512.

Phalakornkule, C., Fry, B., Zhu, T., Kopesel, R., Ataai, M. M. & Domach, M. M. (2000). 13C NMR evidence for pyruvate kinase flux attenuation underlying suppressed acid formation in Bacillus subtilis. *Biotechnol Prog* **16**, 169-175.

Pitson, S. M., Mendz, G. L., Srinivasan, S. & Hazell, S. L. (1999). The tricarboxylic acid cycle of *Helicobacter pylori*. *Eur J Biochem* **260**, 258-267.

Ponce, E., Flores, N., Martinez, A., Valle, F. & Bolivar, F. (1995). Cloning of the two pyruvate kinase isoenzyme structural genes from *Escherichia coli*: the relative roles of these enzymes in pyruvate biosynthesis. *J Bacteriol* **177**, 5719-5722.

Poolman, B., Royer, T. J., Mainzer, S. E. & Schmidt, B. F. (1989). Lactose transport system of *Streptococcus thermophilus*: a hybrid protein with homology to the melibiose carrier and enzyme III of phosphoenolpyruvate-dependent phosphotransferase systems. *J Bacteriol* **171**, 244-253.

Posthuma, C. C., Bader, R., Engelmann, R., Postma, P. W., Hengstenberg, W. & Pouwels, P. H. (2002). Expression of the xylulose 5-phosphate phosphoketolase gene, xpkA, from *Lactobacillus pentosus* MD363 is

induced by sugars that are fermented via the phosphoketolase pathway and is repressed by glucose mediated by CcpA and the mannose phosphoenolpyruvate phosphotransferase system. *Appl Environ Microbiol* **68**, 831-837.

Postma, P. W., Lengeler, J. W. & Jacobson, G. R. (1993). Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol Rev* **57**, 543-594.

Postma, P. W., Epstein, W., Schuitema, A. R. & Nelson, S. O. (1984). Interaction between IIIGlc of the phosphoenolpyruvate:sugar phosphotransferase system and glycerol kinase of *Salmonella typhimurium*. *J Bacteriol* **158**, 351-353.

Presecan-Siedel, E., Galinier, A., Longin, R., Deutscher, J., Danchin, A., Glaser, P. & Martin-Verstraete, I. (1999). Catabolite regulation of the pta gene as part of carbon flow pathways in *Bacillus subtilis*. *J Bacteriol* **181**, 6889-6897.

Priest, F. G. (1977). Extracellular enzyme synthesis in the genus *Bacillus*. *Bacteriol Rev* **41**, 711-753.

Ramstrom, H., Sanglier, S., Leize-Wagner, E., Philippe, C., Van Dorsselaer, A. & Haiech, J. (2003). Properties and regulation of the bifunctional enzyme HPr kinase/phosphatase in Bacillus subtilis. *J Biol Chem* **278**, 1174-1185.

Rangarajan, E. S., Sivaraman, J., Matte, A. & Cygler, M. (2002). Crystal structure of D-ribose-5-phosphate isomerase (RpiA) from *Escherichia coli*. *Proteins* **48**, 737-740.

Rasmussen, L. J., Moller, P. L. & Atlung, T. (1991). Carbon metabolism regulates expression of the *pfl* (pyruvate formate-lyase) gene in *Escherichia coli*. *J Bacteriol* **173**, 6390-6397.

Reitzer, L. J. & Magasanik, B. (1987). Ammonia assimilation and biosynthesis of glutamine, glutamate, aspartate, asparagine, L-alanine. In *Escherichia coli and Salmonella typhimurium: cellular and molecular biology*, pp. 302-320. Edited by F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, S. M. & H. E. Umbarger. Washington D.C.: American Society for Microbiology.

Reizer, J. & Panos, C. (1980). Regulation of beta-galactoside phosphate accumulation in *Streptococcus pyogenes* by an expulsion mechanism. *Proc Natl Acad Sci U S A* **77**, 5497-5501.

Reizer, J., Peterkofsky, A. & Romano, A. H. (1988). Evidence for the presence of heat-stable protein (HPr) and ATP-dependent HPr kinase in heterofermentative lactobacilli lacking phosphoenolpyruvate:glycose phosphotransferase activity. *Proc Natl Acad Sci U S A* **85**, 2041-2045.

Reizer, J., Novotny, M. J., Panos, C. & Saier, M. H., Jr. (1983). Mechanism of inducer expulsion in *Streptococcus pyogenes*: a two-step process activated by ATP. *J Bacteriol* **156**, 354-361.

Reizer, J., Novotny, M. J., Stuiver, I. & Saier, M. H., Jr. (1984). Regulation of glycerol uptake by the phosphoenolpyruvate-sugar phosphotransferase system in *Bacillus subtilis*. *J Bacteriol* **159**, 243-250.

Reizer, J., Hoischen, C., Reizer, A., Pham, T. N. & Saier, M. H., Jr. (1993). Sequence analyses and evolutionary relationships among the energy-coupling proteins Enzyme I and HPr of the bacterial phosphoenolpyruvate: sugar phosphotransferase system. *Protein Sci* **2**, 506-521.

Reizer, J., Bachem, S., Reizer, A., Arnaud, M., Saier, M. H., Jr. & Stulke, J. (1999). Novel phosphotransferase system genes revealed by genome analysis - the complete complement of PTS proteins encoded within the genome of *Bacillus subtilis*. *Microbiology* **145 (Pt 12)**, 3419-3429.

Reizer, J., Bergstedt, U., Galinier, A., Kuster, E., Saier, M. H., Jr., Hillen, W., Steinmetz, M. & Deutscher, J. (1996). Catabolite repression resistance of gnt operon expression in *Bacillus subtilis* conferred by mutation of His-15, the site of phosphoenolpyruvate-dependent phosphorylation of the phosphocarrier protein HPr. *J Bacteriol* **178**, 5480-5486.

Reizer, J., Hoischen, C., Titgemeyer, F., Rivolta, C., Rabus, R., Stulke, J., Karamata, D., Saier, M. H., Jr. & Hillen, W. (1998). A novel protein kinase that controls carbon catabolite repression in bacteria. *Mol Microbiol* **27**, 1157-1169.

Renna, M. C., Najimudin, N., Winik, L. R. & Zahler, S. A. (1993). Regulation of the *Bacillus subtilis* alsS, alsD, and alsR genes involved in post-exponential-phase production of acetoin. *J Bacteriol* **175**, 3863-3875.

Rippa, M., Giovannini, P. P., Barrett, M. P., Dallocchio, F. & Hanau, S. (1998). 6-Phosphogluconate dehydrogenase: the mechanism of action investigated by a comparison of the enzyme from different species. *Biochim Biophys Acta* **1429**, 83-92.

Robillard, G. T. & Broos, J. (1999). Structure/function studies on the bacterial carbohydrate transporters, enzymes II, of the phosphoenolpyruvate-dependent phosphotransferase system. *Biochim Biophys Acta* **1422**, 73-104.

Robinson, M. S., Danson, M. J. & Weitzman, P. D. (1983). Citrate synthase from a Gram-positive bacterium. Purification and characterization of the *Bacillus megaterium* enzyme. *Biochem J* **213**, 53-59.

Rohwer, J. M., Bader, R., Westerhoff, H. V. & Postma, P. W. (1998). Limits to inducer exclusion: inhibition of the bacterial phosphotransferase system by glycerol kinase. *Mol Microbiol* **29**, 641-652.

Roitel, O., Sergienko, E. & Branlant, G. (1999). Dimers generated from tetrameric phosphorylating glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* are inactive but exhibit cooperativity in NAD binding. *Biochemistry* **38**, 16084-16091.

Roitel, O., Vachette, P., Azza, S. & Branlant, G. (2003). P but not R-axis Interface is Involved in Cooperative Binding of NAD on Tetrameric Phosphorylating Glyceraldehyde-3-phosphate Dehydrogenase from *Bacillus stearothermophilus*. *J Mol Biol* **326**, 1513-1522.

Ronimus, R. S. & Morgan, H. W. (2001). The biochemical properties and phylogenies of phosphofructokinases from extremophiles. *Extremophiles* **5**, 357-373.

Rose, M. & Entian, K. D. (1996). New genes in the 170 degrees region of the Bacillus subtilis genome encode DNA gyrase subunits, a thioredoxin, a xylanase and an amino acid transporter. *Microbiology* **142 (Pt 11)**, 3097-3101.

Rowley, D. L. & Wolf, R. E., Jr. (1991). Molecular characterization of the *Escherichia coli* K-12 *zwf* gene encoding glucose 6-phosphate dehydrogenase. *J Bacteriol* **173**, 968-977.

Saier, M. H., Jr. (2000a). A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol Mol Biol Rev* **64**, 354-411.

Saier, M. H., Jr. (2000b). Families of transmembrane sugar transport proteins. *Mol Microbiol* **35**, 699-710.

Saier, M. H., Jr. & Roseman, S. (1976). Sugar transport. Inducer exclusion and regulation of the melibiose, maltose, glycerol, and lactose transport systems by the phosphoenolpyruvate:sugar phosphotransferase system. *J Biol Chem* **251**, 6606-6615.

Saier, M. H., Jr. & Reizer, J. (1994). The bacterial phosphotransferase system: new frontiers 30 years later. *Mol Microbiol* **13**, 755-764.

Saier, M. H., Jr. & Crasnier, M. (1996). Inducer exclusion and the regulation of sugar transport. *Res Microbiol* **147**, 482-489.

Saier, M. H., Jr. & Ramseier, T. M. (1996). The catabolite repressor/activator (Cra) protein of enteric bacteria. *J Bacteriol* **178**, 3411-3417.

Saier, M. H., Jr., Grenier, F. C., Lee, C. A. & Waygood, E. B. (1985). Evidence for the evolutionary relatedness of the proteins of the bacterial phosphoenolpyruvate:sugar phosphotransferase system. *J Cell Biochem* **27**, 43-56.

Saier, M. H., Jr., Ye, J. J., Klinke, S. & Nino, E. (1996). Identification of an anaerobically induced phosphoenolpyruvate-dependent fructose-specific phosphotransferase system and evidence for the Embden-Meyerhof glycolytic pathway in the heterofermentative bacterium Lactobacillus brevis. *J Bacteriol* **178**, 314-316.

Saier, M. H., Jr., Goldman, S. R., Maile, R. R., Moreno, M. S., Weyler, W., Yang, N. & Paulsen, I. T. (2002). Overall transport capabilities of *Bacillus subtilis*. In *Bacillus subtilis and its Closest Relatives: from*

Genes to Cells, pp. 113-128. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology Press.

Saier, M. H., Jr., Yamada, M., Erni, B., Suda, K., Lengeler, J., Ebner, R., Argos, P., Rak, B., Schnetz, K., Lee, C. A. & et al. (1988). Sugar permeases of the bacterial phosphoenolpyruvate-dependent phosphotransferase system: sequence comparisons. *Faseb J* **2**, 199-208.

Sauer, U., Hatzimanikatis, V., Bailey, J. E., Hochuli, M., Szyperski, T. & Wuthrich, K. (1997). Metabolic fluxes in riboflavin-producing *Bacillus subtilis*. *Nat Biotechnol* **15**, 448-452.

Schlapfer, B. S. & 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.

Schnetz, K. & Rak, B. (1988). Regulation of the *bgl* operon of *Escherichia coli* by transcriptional antitermination. *Embo J* **7**, 3271-3277.

Schnetz, K., Stulke, J., Gertz, S., Kruger, S., Krieg, M., Hecker, M. & Rak, B. (1996). LicT, a *Bacillus subtilis* transcriptional antiterminator protein of the BglG family. *J Bacteriol* **178**, 1971-1979.

Schonert, S., Buder, T. & Dahl, M. K. (1998). Identification and enzymatic characterization of the maltose-inducible alpha-glucosidase MalL (sucrase-isomaltase-maltase) of *Bacillus subtilis*. *J Bacteriol* **180**, 2574-2578.

Schurig, H., Beaucamp, N., Ostendorp, R., Jaenicke, R., Adler, E. & Knowles, J. R. (1995). Phosphoglycerate kinase and triosephosphate isomerase from the hyperthermophilic bacterium *Thermotoga maritima* form a covalent bifunctional enzyme complex. *Embo J* **14**, 442-451.

Seok, Y. J., Zhu, P. P., Koo, B. M. & Peterkofsky, A. (1998). Autophosphorylation of enzyme I of the *Escherichia coli* phosphoenolpyruvate:sugar phosphotransferase system requires dimerization. *Biochem Biophys Res Commun* **250**, 381-384.

Shikata, S., Ozaki, K., Kawai, S., Ito, S. & Okamoto, K. (1988). Purification and characterization of NADP+-linked isocitrate dehydrogenase from an alkalophilic *Bacillus*. *Biochim Biophys Acta* **952**, 282-289.

Shin, B. S., Choi, S. K. & Park, S. H. (1999). Regulation of the *Bacillus subtilis* phosphotransacetylase gene. *J Biochem (Tokyo)* **126**, 333-339.

Siebers, B., Brinkmann, H., Dorr, C., Tjaden, B., Lilie, H., van der Oost, J. & Verhees, C. H. (2001). Archaeal fructose-1,6-bisphosphate aldolases constitute a new family of archaeal type class I aldolase. *J Biol Chem* **276**, 28710-28718.

Simpson, C. L. & Russell, R. R. (1998). Identification of a homolog of CcpA catabolite repressor protein in *Streptococcus mutans*. *Infect Immun* **66**, 2085-2092.

Singer, M., Walter, W. A., Cali, B. M., Rouviere, P., Liebke, H. H., Gourse, R. L. & Gross, C. A. (1991). Physiological effects of the fructose-1,6-diphosphate aldolase ts8 mutation on stable RNA synthesis in *Escherichia coli*. *J Bacteriol* **173**, 6249-6257.

Singh, R. P. & Setlow, P. (1978). Enolase from spores and cells of *Bacillus megaterium*: two-step purification of the enzyme and some of its properties. *J Bacteriol* **134**, 353-355.

Singh, S. K., Matsuno, K., LaPorte, D. C. & Banaszak, L. J. (2001). Crystal structure of *Bacillus subtilis* isocitrate dehydrogenase at 1.55 A. Insights into the nature of substrate specificity exhibited by *Escherichia coli* isocitrate dehydrogenase kinase/phosphatase. *J Biol Chem* **276**, 26154-26163.

Skarlatos, P. & Dahl, M. K. (1998). The glucose kinase of *Bacillus subtilis*. *J Bacteriol* **180**, 3222-3226.

Skinner, A. J. & Cooper, R. A. (1971). The regulation of ribose-5-phosphate isomerisation in *Escherichia coli* K12. *FEBS Lett* **12**, 293-296.

Skinner, A. J. & Cooper, R. A. (1974). Genetic studies on ribose 5-phosphate isomerase mutants of *Escherichia coli* K-12. *J Bacteriol* **118**, 1183-1185.

Sonenshein, A. L. (2002). The Krebs citric acid cycle. In *Bacillus subtilis and its Closest Relatives: from Genes to Cells*, pp. 151-162. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology Press.

Sonenshein, A. L., Hoch, J. A. & Losick, R. (2002). *Bacillus subtilis and Its Closest Relatives: from genes to cells*. Washington, DC: American Society for Microbiology Press.

Sorensen, K. I. & Hove-Jensen, B. (1996). Ribose catabolism of *Escherichia coli*: characterization of the *rpiB* gene encoding ribose phosphate isomerase B and of the *rpiR* gene, which is involved in regulation of *rpiB* expression. *J Bacteriol* **178**, 1003-1011.

Spath, C., Kraus, A. & Hillen, W. (1997). Contribution of glucose kinase to glucose repression of xylose utilization in *Bacillus megaterium*. *J Bacteriol* **179**, 7603-7605.

Sprenger, G. A. (1995). Genetics of pentose-phosphate pathway enzymes of *Escherichia coli* K-12. *Arch Microbiol* **164**, 324-330.

Steinmetz, M. (1993). Carbohydrate Catabolism: Pathways, Enzymes, Genetic regulation, and Evolution. In *Bacillus subtilis and Ohter Gram-Positive Bacteria*, pp. 157-170. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology Press.

Steinmetz, M., Le Coq, D. & Aymerich, S. (1989). Induction of saccharolytic enzymes by sucrose in *Bacillus subtilis*: evidence for two partially interchangeable regulatory pathways. *J Bacteriol* **171**, 1519-1523.

Stulke, J. & Hillen, W. (1999). Carbon catabolite repression in bacteria. *Curr Opin Microbiol* **2**, 195-201.

Stulke, J. & Hillen, W. (2000). Regulation of carbon catabolism in Bacillus species. *Annu Rev Microbiol* **54**, 849-880.

Stulke, J., Martin-Verstraete, I., Charrier, V., Klier, A., Deutscher, J. & Rapoport, G. (1995). The HPr protein of the phosphotransferase system links induction and catabolite repression of the Bacillus subtilis levanase operon. *J Bacteriol* **177**, 6928-6936.

Stülke, J. (2002). Control of transcription termination in bacteria by RNA-binding proteins that modulate RNA structures. *Arch Microbiol* **177**, 433-440.

Stülke, J., Arnaud, M., Rapoport, G. & Martin-Verstraete, I. (1998). PRD--a protein domain involved in PTS-dependent induction and carbon catabolite repression of catabolic operons in bacteria. *Mol Microbiol* **28**, 865-874.

Stülke, J., Martin-Verstraete, I., Zagorec, M., Rose, M., Klier, A. & Rapoport, G. (1997). Induction of the *Bacillus subtilis ptsGHI* operon by glucose is controlled by a novel antiterminator, GlcT. *Mol Microbiol* **25**, 65-78.

Sulek, K. (1969). [Nobel prize for S. Ochoa and A. Kornberg in 1959 for the discovery of biosynthesis of ribo- and deoxyribonucleic acids]. *Wiad Lek* **22**, 789-790.

Takahashi, H., Inada, T., Postma, P. & Aiba, H. (1998). CRP down-regulates adenylate cyclase activity by reducing the level of phosphorylated IIA(Glc), the glucose-specific phosphotransferase protein, in *Escherichia coli*. *Mol Gen Genet* **259**, 317-326.

Tanaka, K., Kobayashi, K. & Ogasawara, N. (2003). The *Bacillus subtilis* YufLM two-component system regulates the expression of the malate transporters, MaeN (YufR) and YflS, and is essential for utilization of malate in minimal medium. *Microbiology* **149**, 2317-2329.

Tanaka, K., Komiyama, A., Sonomoto, K., Ishizaki, A., Hall, S. J. & Stanbury, P. F. (2002). Two different pathways for D-xylose metabolism and the effect of xylose concentration on the yield coefficient of L-lactate in mixed-acid fermentation by the lactic acid bacterium *Lactococcus lactis* IO-1. *Appl Microbiol Biotechnol* **60**, 160-167.

Thompson, H., Tersteegen, A., Thauer, R. K. & Hedderich, R. (1998). Two malate dehydrogenases in *Methanobacterium thermoautotrophicum*. *Arch Microbiol* **170**, 38-42.

Thompson, J. & Saier, M. H., Jr. (1981). Regulation of methyl-beta-d-thiogalactopyranoside-6-phosphate accumulation in *Streptococcus lactis* by exclusion and expulsion mechanisms. *J Bacteriol* **146**, 885-894.

Thompson, J. & Torchia, D. A. (1984). Use of 31P nuclear magnetic resonance spectroscopy and 14C fluorography in studies of glycolysis and regulation of pyruvate kinase in *Streptococcus lactis*. *J Bacteriol* **158**, 791-800.

Thompson, J. & Gentry-Weeks, C. R. (1994). Métabolisme des sucres par les bactéries lactiques. In *Bactéries lactiques*, pp. 239-289. Edited by H. D. r.: Uriage:Lorica.

Thomson, G. J., Howlett, G. J., Ashcroft, A. E. & Berry, A. (1998). The *dhnA* gene of *Escherichia coli* encodes a class I fructose bisphosphate aldolase. *Biochem J* **331 (Pt 2)**, 437-445.

Titgemeyer, F., Mason, R. E. & Saier, M. H., Jr. (1994). Regulation of the raffinose permease of *Escherichia coli* by the glucose-specific enzyme IIA of the phosphoenolpyruvate:sugar phosphotransferase system. *J Bacteriol* **176**, 543-546.

Tobisch, S., Stulke, J. & Hecker, M. (1999a). Regulation of the lic operon of *Bacillus subtilis* and characterization of potential phosphorylation sites of the LicR regulator protein by site-directed mutagenesis. *J Bacteriol* **181**, 4995-5003.

Tobisch, S., Glaser, P., Kruger, S. & Hecker, M. (1997). Identification and characterization of a new beta-glucoside utilization system in *Bacillus subtilis*. *J Bacteriol* **179**, 496-506.

Tobisch, S., Zuhlke, D., Bernhardt, J., Stulke, J. & Hecker, M. (1999b). Role of CcpA in regulation of the central pathways of carbon catabolism in *Bacillus subtilis*. *J Bacteriol* **181**, 6996-7004.

Tomb, J. F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A., Klenk, H. P., Gill, S., Dougherty, B. A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E. F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H. G., Glodek, A., McKenney, K., Fitzegerald, L. M., Lee, N., Adams, M. D., Venter, J. C. & et al. (1997). The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388**, 539-547.

Tortosa, P., Aymerich, S., Lindner, C., Saier, M. H., Jr., Reizer, J. & Le Coq, D. (1997). Multiple phosphorylation of SacY, a *Bacillus subtilis* transcriptional antiterminator negatively controlled by the phosphotransferase system. *J Biol Chem* **272**, 17230-17237.

Toyoshima, C., Nakasako, M., Nomura, H. & Ogawa, H. (2000). Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 A resolution. *Nature* **405**, 647-655.

Trach, K., Chapman, J. W., Piggot, P., LeCoq, D. & Hoch, J. A. (1988). Complete sequence and transcriptional analysis of the spo0F region of the *Bacillus subtilis* chromosome. *J Bacteriol* **170**, 4194-4208.

Tseng, C. P. (1997). Regulation of fumarase (*fumB*) gene expression in *Escherichia coli* in response to oxygen, iron and heme availability: role of the *arcA*, *fur*, and *hemA* gene products. *FEMS Microbiol Lett* **157**, 67-72.

Turinsky, A. J., Grundy, F. J., Kim, J. H., Chambliss, G. H. & Henkin, T. M. (1998). Transcriptional activation of the *Bacillus subtilis ackA* gene requires sequences upstream of the promoter. *J Bacteriol* **180**, 5961-5967.

Tusher, V. G., Tibshirani, R. & Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* **98**, 5116-5121.

Vagner, V., Dervyn, E. & Ehrlich, S. D. (1998). A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology* **144 (Pt 11)**, 3097-3104.

Valentini, G., Chiarelli, L., Fortin, R., Speranza, M. L., Galizzi, A. & Mattevi, A. (2000). The allosteric regulation of pyruvate kinase. *J Biol Chem* **275**, 18145-18152.

Valverde, F., Losada, M. & Serrano, A. (1997). Functional complementation of an *Escherichia coli gap* mutant supports an amphibolic role for NAD(P)-dependent glyceraldehyde-3-phosphate dehydrogenase of *Synechocystis* sp. strain PCC 6803. *J Bacteriol* **179**, 4513-4522.

van der Oost, J., Huynen, M. A. & Verhees, C. H. (2002). Molecular characterization of phosphoglycerate mutase in archaea. *FEMS Microbiol Lett* **212**, 111-120.

van der Rest, M. E., Frank, C. & Molenaar, D. (2000). Functions of the membrane-associated and cytoplasmic malate dehydrogenases in the citric acid cycle of *Escherichia coli*. *J Bacteriol* **182**, 6892-6899.

van Tilbeurgh, H. & Declerck, N. (2001). Structural insights into the regulation of bacterial signalling proteins containing PRDs. *Curr Opin Struct Biol* **11**, 685-693.

Varghese, S., Tang, Y. & Imlay, J. A. (2003). Contrasting sensitivities of *Escherichia coli* aconitases A and B to oxidation and iron depletion. *J Bacteriol* **185**, 221-230.

Vaughan, E. E., David, S. & de Vos, W. M. (1996). The lactose transporter in *Leuconostoc lactis* is a new member of the LacS subfamily of galactoside-pentose-hexuronide translocators. *Appl Environ Microbiol* **62**, 1574-1582.

Veiga-da-Cunha, M., Santos, H. & Van Schaftingen, E. (1993). Pathway and regulation of erythritol formation in *Leuconostoc oenos*. *J Bacteriol* **175**, 3941-3948.

Velayudhan, J. & Kelly, D. J. (2002). Analysis of gluconeogenic and anaplerotic enzymes in Campylobacter jejuni: an essential role for phosphoenolpyruvate carboxykinase. *Microbiology* **148**, 685-694.

Verhees, C. H., Tuininga, J. E., Kengen, S. W., Stams, A. J., van der Oost, J. & de Vos, W. M. (2001). ADP-dependent phosphofructokinases in mesophilic and thermophilic methanogenic archaea. *J Bacteriol* **183**, 7145-7153.

Viaene, A. & Dhaese, P. (1989). Sequence of the glyceraldehyde-3-phosphate dehydrogenase gene from *Bacillus subtilis*. *Nucleic Acids Res* **17**, 1251.

Viana, R., Monedero, V., Dossonnet, V., Vadeboncoeur, C., Perez-Martinez, G. & Deutscher, J. (2000). Enzyme I and HPr from *Lactobacillus casei*: their role in sugar transport, carbon catabolite repression and inducer exclusion. *Mol Microbiol* **36**, 570-584.

Vihinen, M. & Mantsala, P. (1990). Characterization of a thermostable *Bacillus stearothermophilus* alpha-amylase. *Biotechnol Appl Biochem* **12**, 427-435.

Viljoen, M., Subden, R. E., Krizus, A. & Van Vuuren, H. J. (1994). Molecular analysis of the malic enzyme gene (*mae2*) of *Schizosaccharomyces pombe*. *Yeast* **10**, 613-624.

Vinopal, R. T. & Fraenkel, D. G. (1974). Phenotypic suppression of phosphofructokinase mutations in *Escherichia coli* by constitutive expression of the glyoxylate shunt. *J Bacteriol* **118**, 1090-1100.

von Wachenfeldt, C. & Hederstedt, L. (2002). Respiratory cytochromes, other heme proteins, and heme biosynthesis. In *Bacillus subtilis and its Closest Relatives: from Genes to Cells*, pp. 163-179. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology Press.

Voskuil, M. I. & Chambliss, G. H. (1996). Significance of HPr in catabolite repression of alpha-amylase. *J Bacteriol* **178**, 7014-7015.

Watabe, K. & Freese, E. (1979). Purification and properties of the manganese-dependent phosphoglycerate mutase of *Bacillus subtilis*. *J Bacteriol* **137**, 773-778.

Waygood, E. B., Rayman, M. K. & Sanwal, B. D. (1975). The control of pyruvate kinases of *Escherichia coli*. II. Effectors and regulatory properties of the enzyme activated by ribose 5-phosphate. *Can J Biochem* **53**, 444-454.

Waygood, E. B., Mort, J. S. & Sanwal, B. D. (1976). The control of pyruvate kinase of *Escherichia coli*. Binding of substrate and allosteric effectors to the enzyme activated by fructose 1,6-bisphosphate. *Biochemistry* **15**, 277-282.

Weaver, T. M., Levitt, D. G. & Banaszak, L. J. (1993). Purification and crystallization of fumarase C from *Escherichia coli*. *J Mol Biol* **231**, 141-144.

Weaver, T. M., Levitt, D. G., Donnelly, M. I., Stevens, P. P. & Banaszak, L. J. (1995). The multisubunit active site of fumarase C from *Escherichia coli*. *Nat Struct Biol* **2**, 654-662.

Wehtje, C., Beijer, L., Nilsson, R. P. & Rutberg, B. (1995). Mutations in the glycerol kinase gene restore the ability of a *ptsGHI* mutant of *Bacillus subtilis* to grow on glycerol. *Microbiology* **141 (Pt 5)**, 1193-1198.

Wei, Y., Guffanti, A. A., Ito, M. & Krulwich, T. A. (2000). *Bacillus subtilis* YqkI is a novel malic/Na+-lactate antiporter that enhances growth on malate at low protonmotive force. *J Biol Chem* **275**, 30287-30292.

Weickert, M. J. & Chambliss, G. H. (1990). Site-directed mutagenesis of a catabolite repression operator sequence in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **87**, 6238-6242.

Weigel, N., Powers, D. A. & Roseman, S. (1982). Sugar transport by the bacterial phosphotransferase system. Primary structure and active site of a general phosphocarrier protein (HPr) from *Salmonella typhimurium*. *J Biol Chem* **257**, 14499-14509.

Weitzman, P. D. (1981). Unity and diversity in some bacterial citric acid-cycle enzymes. *Adv Microb Physiol* **22**, 185-244.

Weitzman, P. D. & Jaskowska-Hodges, H. (1982). Patterns of nucleotide utilisation in bacterial succinate thiokinases. *FEBS Lett* **143**, 237-240.

Wiechert, W. (2001). 13C metabolic flux analysis. *Metab Eng* **3**, 195-206.

Wiegand, G. & Remington, S. J. (1986). Citrate synthase: structure, control, and mechanism. *Annu Rev Biophys Biophys Chem* **15**, 97-117.

Wittekind, M., Reizer, J. & Klevit, R. E. (1990). Sequence-specific 1H NMR resonance assignments of Bacillus subtilis HPr: use of spectra obtained from mutants to resolve spectral overlap. *Biochemistry* **29**, 7191-7200.

Wittekind, M., Reizer, J., Deutscher, J., Saier, M. H. & Klevit, R. E. (1989). Common structural changes accompany the functional inactivation of HPr by seryl phosphorylation or by serine to aspartate substitution. *Biochemistry* **28**, 9908-9912.

Wittekind, M., Rajagopal, P., Branchini, B. R., Reizer, J., Saier, M. H., Jr. & Klevit, R. E. (1992). Solution structure of the phosphocarrier protein HPr from Bacillus subtilis by two-dimensional NMR spectroscopy. *Protein Sci* **1**, 1363-1376.

Wray, L. V., Jr., Pettengill, F. K. & Fisher, S. H. (1994). Catabolite repression of the *Bacillus subtilis hut* operon requires a cis-acting site located downstream of the transcription initiation site. *J Bacteriol* **176**, 1894-1902.

Wu, L. F., Tomich, J. M. & Saier, M. H., Jr. (1990). Structure and evolution of a multidomain multiphosphoryl transfer protein. Nucleotide sequence of the *fruB*(*HI*) gene in *Rhodobacter capsulatus* and comparisons with homologous genes from other organisms. *J Mol Biol* **213**, 687-703.

Ye, J. J. & Saier, M. H., Jr. (1995a). Cooperative binding of lactose and the phosphorylated phosphocarrier protein HPr(Ser-P) to the lactose/H+ symport permease of *Lactobacillus brevis*. *Proc Natl Acad Sci U S A* **92**, 417-421.

Ye, J. J. & Saier, M. H., Jr. (1995b). Allosteric regulation of the glucose:H+ symporter of *Lactobacillus brevis*: cooperative binding of glucose and HPr(ser-P). *J Bacteriol* **177**, 1900-1902.

Ye, J. J. & Saier, M. H., Jr. (1996). Regulation of sugar uptake via the phosphoenolpyruvate-dependent phosphotransferase systems in *Bacillus subtilis* and *Lactococcus lactis* is mediated by ATP-dependent phosphorylation of seryl residue 46 in HPr. *J Bacteriol* **178**, 3557-3563.

Ye, J. J., Reizer, J., Cui, X. & Saier, M. H., Jr. (1994a). ATP-dependent phosphorylation of serine-46 in the phosphocarrier protein HPr regulates lactose/H+ symport in *Lactobacillus brevis*. *Proc Natl Acad Sci U S A* **91**, 3102-3106.

Ye, J. J., Reizer, J., Cui, X. & Saier, M. H., Jr. (1994b). Inhibition of the phosphoenolpyruvate:lactose phosphotransferase system and activation of a cytoplasmic sugar-phosphate phosphatase in *Lactococcus lactis* by ATP-dependent metabolite-activated phosphorylation of serine 46 in the phosphocarrier protein HPr. *J Biol Chem* **269**, 11837-11844.

Ye, R. W., Tao, W., Bedzyk, L., Young, T., Chen, M. & Li, L. (2000). Global gene expression profiles of *Bacillus subtilis* grown under anaerobic conditions. *J Bacteriol* **182**, 4458-4465.

Yoshida, K., Kobayashi, K., Miwa, Y., Kang, C. M., Matsunaga, M., Yamaguchi, H., Tojo, S., Yamamoto, M., Nishi, R., Ogasawara, N., Nakayama, T. & Fujita, Y. (2001). Combined transcriptome and proteome analysis as a powerful approach to study genes under glucose repression in *Bacillus subtilis*. *Nucleic Acids Res* **29**, 683-692.

Yousten, A. A. & Hanson, R. S. (1972). Sporulation of tricarboxylic acid cycle mutants of *Bacillus subtilis*. *J Bacteriol* **109**, 886-894.

Yu, M. T., Kaney, A. R. & Atwood, K. C. (1965). Genetic mapping off fructose-1,6-diphosphatase in *Escherichia coli*. *J Bacteriol* **90**, 1150-1152.

Yumoto, N. & Tokushige, M. (1988). Characterization of multiple fumarase proteins in *Escherichia coli*. *Biochem Biophys Res Commun* **153**, 1236-1243.

Zalieckas, J. M., Wray, L. V., Jr. & Fisher, S. H. (1998a). Expression of the *Bacillus subtilis acsA* gene: position and sequence context affect cre-mediated carbon catabolite repression. *J Bacteriol* **180**, 6649-6654.

Zalieckas, J. M., Wray, L. V., Jr. & Fisher, S. H. (1999). trans-acting factors affecting carbon catabolite repression of the hut operon in *Bacillus subtilis*. *J Bacteriol* **181**, 2883-2888.

Zalieckas, J. M., Wray, L. V., Jr., Ferson, A. E. & Fisher, S. H. (1998b). Transcription-repair coupling factor is involved in carbon catabolite repression of the *Bacillus subtilis hut* and *gnt* operons. *Mol Microbiol* **27**, 1031-1038.

Zhao, G., Pease, A. J., Bharani, N. & 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.

Zhu, T., Phalakornkule, C., Koepsel, R. R., Domach, M. M. & Ataai, M. M. (2001). Cell growth and by-product formation in a pyruvate kinase mutant of *E. coli*. *Biotechnol Prog* **17**, 624-628.