

50376  
1991  
118

66316

50376  
1991  
118

N° d'ordre 733

Thèse de Doctorat de l'Université des  
Sciences et Techniques de  
Lille Flandre Artois

Spécialité : Science de la Vie et de la Santé  
Microbiologie

présentée par Alain VERTES  
pour obtenir le titre de Docteur de l'Université de Lille

Sujet

Etapas en Biotechnologie :  
Contributions à l'étude de *Clostridium acetobutylicum*,  
*Saccharomyces cerevisiae* et *Bacillus subtilis*



Soutenue le 18 Juin 1991

numéro d'ordre : 733

Devant le jury composé de :

|                                |            |
|--------------------------------|------------|
| Monsieur J. GUILLAUME.....     | Président  |
| Monsieur H.C. DUBOURGUIER..... | Rapporteur |
| Monsieur B. LUBOCHINSKY.....   | Rapporteur |
| Monsieur S. BALL.....          | Examineur  |
| Monsieur J.M. LEBEAULT.....    | Examineur  |
| Monsieur R. TAILLIEZ.....      | Examineur  |

Diga, mon grand, aquela rota  
 Que seguisses sens t'arrestar,  
 Diga, mon grand, vertat qu'es longa  
 E consi caminar ?

Filhet, passaràs las crèstas  
 Per d'autres òmes, d'autres camps  
 D'amics e plan de sèrs de fèsta  
 Es una joia avançar  
 Mas si s'amaga ton solelh  
 Si vas ton camin dins lo fred  
 Te faràs fòrt coma una pèira  
 E avança !

Traparas de rotas grandas  
 Ont los pòbles parlan fòrt  
 Passaràs per lors vilatges  
 Aparats dels vents de mòrt.  
 Mas veiras de malas sendas  
 Carrejar d'òmes plegats  
 Rosegats per d'autres òmes  
 E muts ...

Navegaràs dins los espacis  
 Saupràs totas las colors :  
 Lo negre qu'es lo silenci  
 Lo roge coma una flor.  
 Mas tornaràs cap a ton pòble  
 Que s'encamina el tamben.  
 Alara sera grand ton viatge  
 Amb el ...

Vaqui çò per la montanha  
 Contava un òme a un enfant  
 Vertat disiás e me remembri  
 La longa marcha a començat

Claude Marti

## REMERCIEMENTS

Je remercie Monsieur Jean GUILLAUME, Professeur, Monsieur Roger TAILLIEZ, Professeur, et Monsieur Henri-Charles DUBOURGUIER, Directeur de Recherches INRA, pour m'avoir accueilli au sein de leur équipe du Laboratoire de Microbiologie de l'Université de Lille Flandre Artois à l'issue de mon Master of Science à l'Université d'Illinois.

Dans le cadre de mon inscription en thèse à l'Université de Lille, j'ai également bénéficié d'un séjour de longue durée à l'Institut Pasteur dans l'Unité des Anaérobies et dans l'Unité de Biochimie Microbienne. Je tiens donc à remercier tout particulièrement Mademoiselle Madeleine SEBALD, Professeur, et Monsieur Georges RAPOPORT, Professeur, pour m'avoir fait bénéficier de leurs compétences dans les domaines de la biologie de *Bacteroides fragilis* et de *Bacillus subtilis*.

Je tiens en outre à remercier Messieurs les Professeurs Georges RAPOPORT, Raymond DEDONDER, et Antoine DANCHIN sans lesquels le projet du séquençage du génome de *Bacillus subtilis* n'aurait pu être possible.

Je remercie Monsieur André KLIER, Professeur, ainsi que Monsieur Frank KUNST et Mesdames Martine CRASNIER et Véronique RIBES pour leur disponibilité lors de la relecture de ce rapport.

Je remercie très chaleureusement Monsieur Bernard LUBOCHINSKY, Professeur, pour son aide précieuse. Ses conseils et ses critiques constructives m'ont été d'une grande aide lors de la rédaction de ce rapport.

Je remercie Monsieur Steven BALL, Professeur, et Monsieur Jean-Michel LEBEAULT, Professeur, pour avoir accepté de juger cette thèse.

Je remercie Monsieur Philippe GLASER pour son soutien constant..

Je remercie Messieurs Michel DEBARBOUILLE, Tarek MSADEK, et Johannes SCHWEITZER, ainsi que Mademoiselle Margarida SANTANA pour leur efficacité et leur bonne humeur sans égal.

Je remercie tous mes collègues de l'Unité des Anaérobies, de l'Unité de Biochimie Microbienne, et de l'Unité de Régulation Génétique pour avoir contribué par leur sympathie à faire de mon séjour à l'Institut Pasteur un moment inoubliable.

Que chacun trouve ici l'expression de ma plus profonde gratitude.

## SOMMAIRE

|   |      |
|---|------|
| PREFACE.....  | xv   |
| RESUME.....   | xiii |
| I) INTRODUCTION.....  | 1    |
| 1) LA FERMENTATION ACETONE BUTANOL.....   | 1    |
| 2) LA FERMENTATION ALCOOLIQUE EN OENOLOGIE.....   | 3    |
| 3) LE PROJET DE SEQUENCAGE DU GENOME DE <i>B. SUBTILIS</i> ...6   |      |
| a) INTERET BIOTECHNOLOGIQUE DE <i>B. SUBTILIS</i> .....6  |      |
| b) DEVELOPPEMENT DES PROGRAMMES DE SEQUENCAGE DES<br>GRANDS GENOMES : LE CAS DE <i>B. SUBTILIS</i> .....9                                     |      |
| II) ISOLEMENT DE LA FORME SIMPLE BRIN D'UN PHAGE<br>FILAMENTEUX DEFECTIF DE LA SOUCHE NCIB 6444 DE<br><i>CLOSTRIDIUM ACETOBUTYLICUM</i> ..... | 14   |
| A) BIOLOGIE DE <i>CLOSTRIDIUM ACETOBUTYLICUM</i> .....  | 14   |
| B) THEMES ET BUTS DE RECHERCHE.....   | 16   |
| C) MATERIELS ET METHODES.....   | 17   |
| D) RESULTATS ET DISCUSSION.....   | 19   |
| E) CONCLUSION.....  | 25   |
| F) ARTICLE 1.....   | 26   |
| G) ARTICLE 2.....   | 31   |
| III) CONTROLE DE L'EFFICACITE DU LEVURAGE PAR LA MISE EN<br>OEUVRE DE SOUCHES DE LEVURES MARQUEES.....  | 46   |
| A) LA PRATIQUE DU LEVURAGE. BUT DE L'ETUDE.....   | 46   |
| B) MATERIELS ET METHODES.....   | 47   |
| C) RESULTATS.....   | 47   |
| D) DISCUSSION.....  | 54   |
| E) ARTICLE 3.....   | 57   |
| IV) CONTRIBUTION AU SEQUENCAGE DE LA REGION <i>sacS-gerB</i> DU<br>GENOME DE <i>BACILLUS SUBTILIS</i> .....                                   | 66   |
| A) BIOLOGIE DE <i>BACILLUS SUBTILIS</i> .....   | 66   |
| B) THEMES ET BUTS DE RECHERCHE.....   | 68   |
| C) MATERIELS ET METHODES.....   | 69   |
| D) RESULTATS ET DISCUSSIONS.....  | 80   |
| 1) CONSTRUCTION DES OUTILS GENETIQUES.....80  |      |
| a) CONSTRUCTION D'UNE BANQUE DE <i>B. SUBTILIS</i><br>DANS LAMBDAFIXII.....80   |      |
| b) CONSTRUCTION D'UNE BANQUE DE <i>B. SUBTILIS</i><br>DANS UN PLASMIDE.....86   |      |
| c) CONSTRUCTION DE VECTEURS INTEGRATIFS.....86  |      |

|    |  |     |
|----|--|-----|
| 2) | PREPARATION DES SONDAS.....  | 88  |
| a) | CLONAGES PAR MOYENS GENETIQUES.....  | 91  |
| b) | PREPARATION DES SONDAS PAR PCR.....  | 96  |
| c) | CLONAGE D'UN FRAGMENT TERMINAL DU PHAGE<br>LAMBDA sacPT.....                             | 98  |
| d) | PREPARATION DES SONDAS ARN POUR LE PHAGE<br>LAMBDA narA4.0.....                          | 98  |
| 3) | CLONAGE DE GRANDS FRAGMENTS DU CHROMOSOME DE<br><i>B. SUBTILIS</i> .....                 | 98  |
| a) | LAMBDA narA.....   | 98  |
| b) | LAMBDA spoOE.....  | 105 |
| c) | LAMBDA spoIID.....   | 105 |
| d) | LAMBDA H.....  | 108 |
| e) | CLONAGE D'UNE PROTEINE BIOTINYLEE.....   | 110 |
| 4) | CARTOGRAPHIE DE LA REGION <i>sacS-sacA</i> .....   | 117 |
| 5) | SEQUENCAGE DU PHAGE LAMBDA narA.....   | 122 |
| a) | SOUS-CLONAGE EN SHOTGUN DU PHAGE LAMBDA<br>narA DANS LE PHAGE M13mp8.....                | 122 |
| b) | SEQUENCAGE.....  | 130 |
| 6) | ANALYSE DES SEQUENCES.....   | 130 |
| a) | TESTS RAPIDES POUR LA DETECTION D'ENZYMES<br>IDENTIFICATION D'UNE NOUVELLE PROTEASE..... | 130 |
| b) | ANALYSE DES SEQUENCES A L'AIDE DES BANQUES<br>PEPTIDIQUES.....                           | 131 |
| c) | OPERON ATP SYNTHASE DE <i>BACILLUS SUBTILIS</i> .....                                    | 133 |
| d) | IDENTIFICATION DE <i>SPOIID</i> ET ANALYSE DES<br>GENES ADJACENTS.....                   | 151 |
| e) | LA SOUS-UNITE GAMMA DE L'UREASE DE<br><i>BACILLUS SUBTILIS</i> .....                     | 157 |
| 7) | ANALYSE FONCTIONNELLE DES SEQUENCES OBTENUES..   | 162 |
|    | Analyse du gène de la tyrosine tRNA synthétase   |     |
| E) | DISCUSSION GENERALE.....   | 164 |
| F) | ARTICLE 4.....   | 180 |
|    | CONCLUSION.....  | 192 |
|    | ANNEXES.....   | 194 |
|    | REFERENCES.....  | 214 |

LISTE DES FIGURES

|  |     |
|--|-----|
| Figure 1 : Isolement d'ADN par électroporation de <i>C. acetobutylicum</i> NCIB 6444.....                | 20  |
| Figure 1bis : Identification de la forme simple brin du phage CAK1 par traitement à la nucléase S1.....  | 22  |
| Figure 2 : Isolement d'ADN par électroporation de <i>C. acetobutylicum</i> NCIB 6444 : D0260/ D0600..... | 24  |
| Figure 2bis : Cinétique de production de jus de goutte en macération carbonique.....                     | 4   |
| Figure 3 : Evolution de la température du moût en fermentation.....                                      | 48  |
| Figure 4 : Evolution de la densité du moût en fermentation.....  | 48  |
| Figure 5 : Vendange foulée en cuve rotative.....   | 50  |
| Figure 6 : Vendange foulée en cuve statique.....   | 51  |
| Figure 7 : Thermovinification.....   | 52  |
| Figure 8 : Macération carbonique.....  | 53  |
| Figure 10 : Carte génétique de <i>B. subtilis</i> .....  | 149 |
| Figure 11 : Carte physique du génome de <i>B. subtilis</i> .....   | 150 |
| Figure 12 : Carte du vecteur $\lambda$ FixII.....  | 81  |

|   |     |
|---|-----|
| Figure 13 : La technique de saut sur le chromosome.....   | 165 |
| Figure 16 : Hydrolyse partielle d'ADN chromosomique de <i>B. subtilis</i> .....   | 82  |
| Figure 17 : Construction d'une banque de <i>B. subtilis</i> .....   | 84  |
| Figure 18 : Carte des plasmides pDIA5304 et pDIA5305.....   | 87  |
| Figure 19 : Carte du plasmide pNAR2.....  | 92  |
| Figure 20 : Amplification de <i>spoOE</i> par PCR.....  | 94  |
| Figure 21 : Carte du phage $\lambda$ sacPT.....   | 97  |
| Figure 22 : Criblage d'une banque de phages.....  | 100 |
| Figure 23 : Hybridation des phages $\lambda$ narA avec le plasmide pNAR2.....   | 102 |
| Figure 24 : Carte des phages $\lambda$ narA1.0 et $\lambda$ narA4.0.....  | 104 |
| Figure 25 : Cartographie de la région <i>sacS-sacA</i> .....  | 109 |
| Figure 26 : Clonage d'une protéine biotinylée.....  | 111 |
| Figure 27 : Hybridation des phages $\lambda$ spoIID avec le fragment PCR <i>spoIID</i> .....  | 106 |
| Figure 28 : Clonage d'une protéine biotinylée : Analyse des phages recombinants.....  | 114 |
| Figure 29 : Cartographie de la région <i>sacS-sacA</i> : hybridation de l'ADN génomique avec le plasmide pthiC et la sonde "H"..... | 121 |

|  |     |
|--|-----|
| Figure 30 : Séquençage du phage $\lambda$ narA, sonication de l'ADN.....   | 123 |
| Figure 31 : Séquençage du phage $\lambda$ narA : sous-clonage dans le phage M13.....   | 125 |
| Figure 32 : Hybridation des phages M13mp8 recombinants avec l'insert chromosomique du phage $\lambda$ narA4.0.....                               | 128 |
| Figure 33 : L'opéron ATPase.....   | 135 |
| Figure 34a-34f : Alignement de la séquence peptidique des sous-unités de l'ATPase de <i>B. subtilis</i> avec celle de <i>B. megaterium</i> ..... | 137 |
| Figure 35 : Sites de fixation de l'ATP des sous-unités $\alpha$ et $\beta$ de l'ATP synthase.....  | 144 |
| Figure 36 : Termineurs des opérons <i>atp</i> de <i>B. subtilis</i> et <i>B. megaterium</i> .....  | 147 |
| Figure 37 : Carte des phages recombinants $\lambda$ sacS et $\lambda$ sacPT.....   | 176 |
| Figure 38 : Carte du phage $\lambda$ Eco-Sal et localisation du phage $\lambda$ narA.....  | 177 |
| Figure 39 : Les enzymes biotinylées dans le métabolisme cellulaire.....  | 169 |
| Figure 40 : Hybridation des phages $\lambda$ spoIID avec le phage $\lambda$ narA4.0.....   | 152 |
| Figure 41 : Carte de la région adjacente à <i>spoIID</i> .....   | 154 |
| Figure 43 : Alignement de la séquence peptidique de la protéine putative PROT R celle du <i>locus rev-4</i> .....                                | 156 |

Figure 44a-44b : Alignement de la séquence peptidique de la sous-unité  $\gamma$  de l'uréase de *B. subtilis* avec celle d'*U. urealyticum* et avec l'uréase du haricot.....158

**LISTE DES TABLEAUX**

|  |     |
|--|-----|
| Tableau 1 : Isolement d'ADN par électroporation de <i>C. acetobutylicum</i> NCIB 6444.....               | 19  |
| Tableau 13 : Produits et utilisations de <i>B. subtilis</i> .....  | 7   |
| Tableau 15 : Marqueurs génétiques de la région <i>sacS-gerB</i> ...                                      | 89  |
| Tableau 16 : Cartographie de la région <i>sacS-sacA</i> .....  | 119 |
| Tableau 17 : Homologies détectées dans le phage $\lambda$ narA4.0..                                      | 132 |
| Tableau 18 : L'opéron ATPsynthase de <i>B. subtilis</i> :<br>composition en acides aminés.....           | 143 |
| Tableau 19 : Organisation de l'opéron ATP synthase chez <i>B. subtilis</i> et <i>B. megaterium</i> ..... | 145 |
| Tableau 20 : L'opéron ATPase de <i>B. subtilis</i> :<br>usage du code.....                               | 148 |
| Tableau 22 : RBS des ORFs identifiées dans le phage<br>$\lambda$ sacPT.....                              | 178 |
| Tableau 23 : ORFs détectées dans le phage $\lambda$ narA :<br>composition en acides aminés.....          | 155 |

## **PREFACE**

Ayant une formation d'ingénieur centrée sur les aspects appliqués des technologies concernant les matériaux d'origine biologique, j'ai cherché à obtenir une formation en Microbiologie suffisamment diversifiée pour me permettre d'acquérir une connaissance approfondie des principaux systèmes utilisés en biotechnologie.

Je me suis d'abord intéressé à l'étude de Clostridium acetobutylicum étant donné que la fermentation Acétone-Butanol est l'une des plus anciennes bioconversions utilisées en industrie (Weizmann, 1914) encore opérationnelle aujourd'hui en Chine. J'ai tenu également à me familiariser avec l'étude des levures, étant donné que Saccharomyces cerevisiae est le microorganisme le plus important d'un point de vue industriel, tant au niveau des applications biotechnologiques qu'agroalimentaires. Par ailleurs, j'ai étayé ma connaissance des techniques actuelles d'analyse par l'étude de la Biologie Moléculaire de Bacillus subtilis, organisme d'importance primordiale étant donné que c'est, après Escherichia coli, le microorganisme dont la physiologie et la biologie moléculaire sont les mieux connues. B. subtilis est en outre l'organisme modèle pour l'étude de la régulation de l'expression génétique chez les bactéries à gram(+) ainsi que pour l'étude de la sporulation ; et surtout c'est avec E. coli, S. cerevisiae, et les cellules de mammifères, l'un des systèmes de choix pour diriger la sécrétion des protéines homologues (enzymes dégradatives) ou hétérologues.

La formation que j'ai reçue m'a donc permis d'utiliser les connaissances accumulées dans les domaines de la biologie de C. acetobutylicum, S. cerevisiae, et B. subtilis pour aborder des problèmes nouveaux et obtenir des résultats originaux. Cette formation se révèle en outre bien adaptée aux besoins de mon prochain poste de travail qui comprend une part importante de recherche fondamentale sur les Corynebactéries en vue de la production industrielle d'acides aminés.

## RESUME

Ce travail a été réalisé dans plusieurs laboratoires dont les thèmes de recherche incluent l'étude de trois organismes représentatifs des principaux systèmes utilisés en biotechnologie : Clostridium acetobutylicum, Saccharomyces cerevisiae, et Bacillus subtilis.

Au cours de cette étude, un phage filamenteux défectif de C. acetobutylicum qui existe sous forme simple brin a été isolé, et une méthode d'isolement d'ADN spécifique pour la forme simple brin a été mise au point à partir des techniques d'électroporation. Ceci devrait permettre la détermination d'un système efficace de transformation génétique pour C. acetobutylicum.

L'utilisation de levures marquées de S. cerevisiae pour le levurage de jus de fermentation a été vérifiée en conditions réelles de vinification (50 hl). Quatre technologies différentes de vinification ont été étudiées ; il apparait que seule la fermentation par macération carbonique pose le problème de l'établissement de la souche de levure sélectionnée.

L'intérêt du séquençage intégral du génome de B. subtilis a été illustré. Les résultats significatifs que j'ai obtenus lors de l'étude de B. subtilis incluent le clonage des gènes *narA* et *thiC* par des moyens génétiques et l'isolement de plusieurs phages  $\lambda$ FixII recombinants contenant des inserts chromosomiques de B. subtilis. L'un de ces phages a été séquencé qui contient les gènes *narA* et *spoIID*, une ORF identique à 50% à l'ORF R (*rev-4*) de B. subtilis, les sous-unités  $\delta$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\epsilon$  de l'ATPase, et un gène qui pourrait être celui de la sous-unité  $\gamma$  de l'uréase. Ce travail a également permis le clonage d'une protéine biotinylée de B. subtilis, la détection d'une nouvelle

protéase, et l'identification d'une tyrosine tRNA synthétase secondaire (*yrs*). Il est de plus intéressant de noter que les gènes *narA* et *spoIID* sont situés à proximité l'un de l'autre sur le chromosome de *B. subtilis*.

## 1) INTRODUCTION

### 1) LA FERMENTATION ACETONO-BUTANOLIQUE

La découverte que le butanol pouvait être produit par fermentation a été faite par Pasteur en 1861. Toutefois, ce n'est qu'en 1914 que Weizmann a isolé Clostridium acetobutylicum, un microorganisme capable de produire et d'accumuler de l'acétone, du butanol, et de l'éthanol.

Le début de la première guerre mondiale a considérablement accru l'intérêt pour la fermentation Acétono-Butanolique (ABE) pour la manufacture de cordite. Après l'armistice, l'acétone n'étant plus nécessaire en aussi grandes quantités, la plupart de usines de fermentation fermèrent. Néanmoins, le butanol a trouvé un regain d'intérêt avec l'extension de l'industrie automobile. De nouvelles usines de fermentation furent même construites et des souches capables d'utiliser les molasses sélectionnées.

Le début de la deuxième guerre mondiale a de nouveau accru l'intérêt pour la fermentation ABE destinée à couvrir les besoins en acétone pour la manufacture de munitions. Plus encore, la fin de la guerre a également sonné le déclin de la fermentation ABE. De plus, le développement rapide de l'industrie pétrochimique et l'utilisation de molasses pour l'élevage ont fait que cette bioconversion n'était plus rentable d'un point de vue économique. En 1960, la plupart de usines de fermentation fermèrent et seule celle d'Afrique du Sud était encore en activité du fait de l'abondance de molasses bon marché. Cette dernière cessa néanmoins ses activités en 1982 (Walton and Martin; 1979 ; Jones and Woods, 1986), mais la Chine produit de nos jours de l'acétone et du butanol à l'aide de cette bioconversion (communiqué au GIM Meeting 1990).

Néanmoins, le choc pétrolier de 1970 a induit un renouveau d'intérêt pour cette bioconversion. En effet, les

utilisations du butanol sont nombreuses et le marché du butanol a représenté en 1987 300 millions de dollars (Shamel and Chow, 1988). Il a en outre été rapporté que le mélange acétone-butanol-éthanol est un meilleur combustible que ne l'est l'éthanol (Daneman, 1981) ; mais la fermentation ABE connaît encore de sérieux obstacles qui doivent être surmontés pour rendre ce procédé compétitif d'un point de vue économique. L'une des principales voies de recherche de la fermentation ABE concerne l'extraction des solvants du milieu de fermentation. Des mutants capables de produire uniquement du butanol ont été isolés (Janati-Idrissi et al., 1987). D'autre part, étant donné que le butanol est un produit très toxique pour C. acetobutylicum, des mutants résistants au butanol ont également été isolés (Lin and Blaschek, 1983 ; Hermann et al., 1985 ; Blaschek, 1986). L'étude des facteurs inducteurs de la solvantogénèse a de même suscité une attention considérable (Bahl et al., 1986 ; Terraciano and Kashket, 1986 ; Bryant and Blaschek, 1988) ainsi que le potentiel de divers extractants utilisés durant la fermentation (Evans and Wang, 1988) ou la possibilité de recycler les cellules, de les immobiliser, ou de réaliser une culture continue (Bahl et al., 1982b, Monot and Engasser, 1983 ; Afshar et al., 1985).

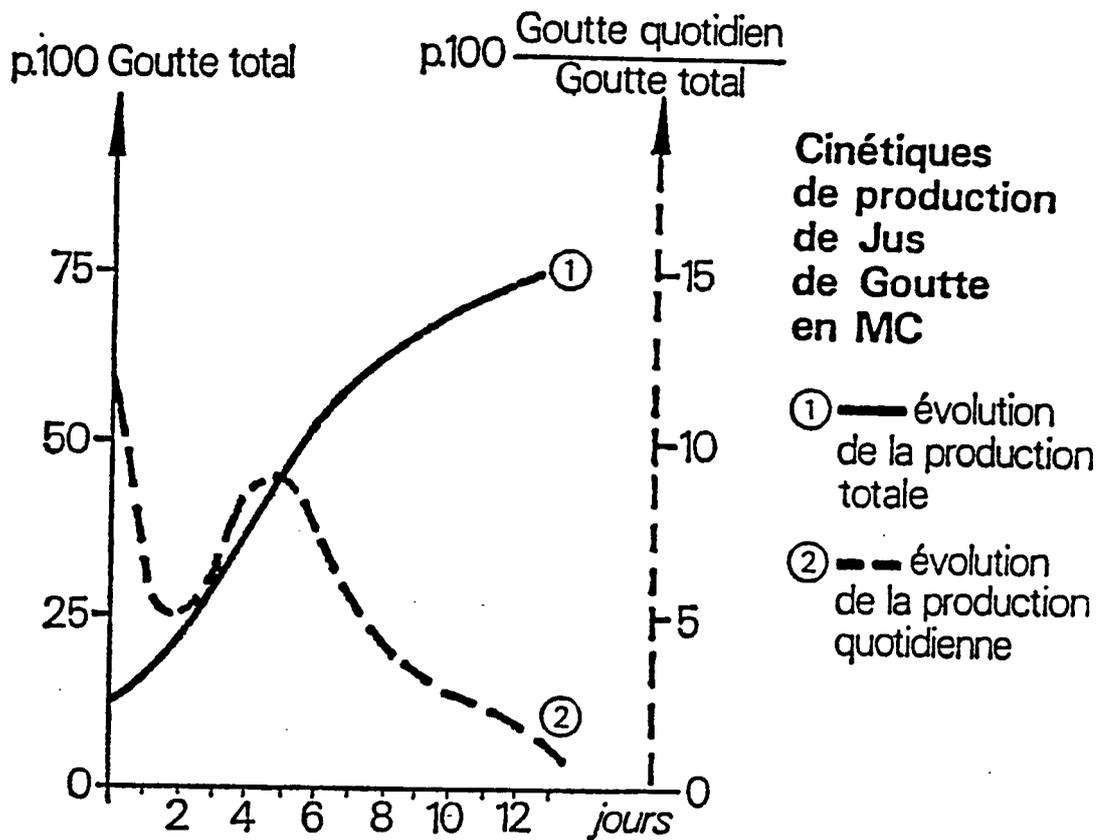
La deuxième voie de recherche concerne l'utilisation de substrats meilleur marché tels que les déchets des industries agroalimentaires ou de l'agriculture. Toutefois, il apparaît que les techniques de l'ADN recombinant doivent être appliquées à C. acetobutylicum pour rendre ce dernier capable d'utiliser un large spectre de substrats (Blaschek, 1986). Deux approches différentes sont utilisées pour accroître le pouvoir catabolique de C. acetobutylicum, à savoir la mutagénèse chimique et le clonage de gènes qui fait appel aux techniques de transformation génétique.

## 2) LA FERMENTATION ALCOOLIQUE EN OENOLOGIE

Les applications biotechnologiques des levures sont nombreuses et incluent en particulier la production d'éthanol, de lipides, l'utilisation d'amidon, de lactose, d'inuline, de xylose, de cellobiose, la dégradation de composés aromatiques et la sécrétion de protéines hétérologues, sans oublier l'utilisation traditionnelle des levures en brasserie, en oenologie et dans l'industrie boulangère (Verachtert and de Mot, 1990).

Comme souligné par Peynaud (1983), l'oenologie est une science microbiologique et le devient du reste de plus en plus. En effet, la production de vin s'effectue en deux phases principales, à savoir la fermentation alcoolique réalisée par des levures, et la fermentation malolactique réalisée principalement par des souches de Leuconostoc et de Lactobacilli. Cette dernière fermentation correspond à l'utilisation par des bactéries lactiques de l'acide malique produit par les levures lors de la fermentation alcoolique (Ribereau-Gayon et coll., 1975 ; Peynaud, 1983).

Il existe quatre différentes technologies de base pour la production des vins : le foulage, le pressurage, la thermovinification, et la macération carbonique. Le foulage consiste à rompre la pellicule du raisin de façon à en dégager la pulpe, et à en libérer le jus qui est légèrement aéré et simultanément mélangé avec les levures qui se trouvent à la surface des pellicules. Cette opération est le plus ancien traitement appliqué à la vendange rouge. La thermovinification a pour but d'améliorer la macération et l'extraction des matières colorantes, ainsi que de détruire les enzymes oxydatives. Cette technique consiste principalement en un traitement thermique à 70°C pendant quelques minutes après foulage et égrappage de la vendange. Le principe de la macération carbonique (cf. INRA, 1973) consiste au cours du premier stade de vinification à encuver sans foulage la vendange dans une atmosphère fortement enrichie en dioxyde de carbone afin de bénéficier des voies



(d'après La vinification par macération carbonique, INRA)

Figure 2bis

métaboliques qui ont lieu dans les baies placées en anaérobiose. D'un point de vue organoleptique, l'originalité des vins de macération carbonique provient de la composante "kirsch" de l'arôme (Ribereau-Gayon, 1975 ; Peynaud, 1983). La première phase fermentaire par les levures concerne le jus de goutte résultant de l'écrasement des baies sous le poids de la gravité. Il est particulièrement notable que le volume de ce jus de goutte augmente au cours du temps (cf. Figure 2bis). La deuxième phase fermentaire débute après l'opération de pressurage et implique une fermentation alcoolique et une fermentation malolactique.

Pour pallier l'irrégularité des départs et des durées de fermentation, le levurage est une technique qui est très largement employée de nos jours. En effet, comme souligné par Peynaud (1983), l'usage des levures sèches actives en vinification constitue un grand progrès dans la maîtrise des vinifications et dans la recherche de la qualité. Les principaux problèmes rencontrés lors de l'opération de levurage sont les problèmes de la contamination initiale, la compétition entre la souche sélectionnée et la population indigène, la préparation des levains.

La pratique des fermentations alcooliques en souche pure (Barre and Vezinhet, 1984) permet de plus de bénéficier des techniques de génétique (mutagénèse, cytoduction, transformation) (Snow, 1982) pour améliorer la qualité des levures oenologiques (Vezinhet, 1981). L'implantation des levures sélectionnées peut du reste être suivie par l'utilisation de levures marquées (Vezinhet et Lacroix, 1984 ; Vezinhet et Loiseau, 1985 ; Loiseau et coll., 1985).

### 3) LE PROJET DE SEQUENCAGE DU GENOME DE B. SUBTILIS

#### a) INTERET BIOTECHNOLOGIQUE DE BACILLUS SUBTILIS

B. subtilis est une bactérie dont le potentiel pour l'industrie des bioconversions est immense (Workman et al., 1986 ; Priest, 1989). En effet, B. subtilis est après E. coli le microorganisme dont la génétique et la physiologie sont les mieux connus, c'est aussi une bactérie non pathogène qui jouit du statut "GRAS" (*Generally Recognized As Safe*) aux Etats-Unis et dont l'usage en biotechnologie industrielle est déjà ancien, en particulier pour la production d'enzymes extracellulaires. On peut par exemple noter que la production d'amylase par B. subtilis était déjà étudiée en 1917 par Boidin et Effront, ou encore que la production de 2,3-butanediol à partir de B. subtilis était envisagée dès 1912 (Lemoigne, cité par Prescott et Dunn, 1949).

Les applications et les applications potentielles des enzymes dérivées de B. subtilis sont nombreuses. On peut en effet retrouver l'utilisation de ces enzymes (protéases, amylases, lipases) dans l'industrie agro-alimentaire, dans l'industrie textile, dans l'industrie du cuir ou dans la papetterie (Priest, 1984 ; Hammes, 1987 ; Kula, 1987), ainsi que dans la production de détergents (Bahn and Schmid, 1987). Il convient également de noter la production d'antibiotiques à partir de B. subtilis et d'autres Bacilli (Schaeffer, 1969 ; Hopwood and Merrick, 1977 ; Katz and Demain, 1977 ; Shoji, 1978 ; Nakano and Zuber, 1990) qui peuvent trouver leur utilité non seulement en médecine, bien que leur intérêt en galénique soit limité, mais encore dans l'agriculture (Phae et al., 1990). Une liste relativement exhaustive des produits et usages de B. subtilis a pu être dressée à partir de la compilation publiée par Edwards (1988) (cf. Tableau 13). Les procédés industriels relatifs à certaines de ces productions ont du reste été revus par Crueger et Crueger (1982).

Outre la production d'enzymes, d'antibiotiques, ou de nucléosides puriniques, l'intérêt industriel de B. subtilis

**PRODUITS ET UTILISATIONS DE BACILLUS SUBTILIS**

| Produits ou usage   | Souche ATCC         |
|---|---------------------|
| acide 5' adenylique   | 14617, 14618        |
| adhésifs  | 6984                |
| AICAR (riboside 5 amino-dazolecarboxamide)  | 15115, 15116        |
| amylase   | 15841, 35854        |
| $\alpha$ -amylase   | 21556, 21770        |
| L-arginine  | 21742, 31002, 31004 |
| bacillomycine B   | 10774               |
| citrulline  | 15561, 15562        |
| deacylation de tetrazole  | 31028               |
| de penicilline  |                     |
| degradation de cyanide (eaux)   | 21697               |
| degradation de nitriles (eaux)  | 21697               |
| detection de phenylketonuria  | 6051                |
| diacétinase   | 31954               |
| dosage et essai de :  | 6633                |
| amoxicilline, cephalaxine, cephaloglycine, cephalosidine, ceftrimide, dihydrostreptomycine, gentamicine, hexachlorophene, kanamycine, kanamycine B, lasalocide A lastanax Q, mytocide, monensine, neomycine, novobiocine, penicilline, rifampicine, |                     |
| dosage des sporocides   | 19659               |
| dosage de streptothricine   | 9524                |
| dosage tobramycine, vancomycine   | 6633, 11774         |
| acide diaminopimolique  | 19549, 19550        |
| difficidine   | 39320, 39374        |
| additifs  | 19385               |
| 5 fluorouracileriboside   | 19062               |
| $\alpha$ -1,6 glucosidase   | 31068, 31077        |
| acide L-glutamique  | 13062               |
| L-guanine   | 14660, 14662        |
| guanosine   | 19217, 19219, 19220 |
| acide 5' guanylique   | 19217, 19219, 19220 |
| L-histidine   | 21603               |
| hypoxanthine  | 15512               |

Tableau 13

| Produit ou usage   | Souche ATCC  |
|--|--|
| inosine  | 13952, 13956, 19162,<br>19163, 21005, 21008  |
| L-inosine  | 14660, 14662   |
| acide 5' inosinique  | 14617, 14618, 19162,<br>19163  |
| lactase  | 31382  |
| 6 mercaptopurineriboside                                     | 19062  |
| orotidine  | 15181, 15184   |
| pentosanases   | 12711  |
| protease   | 21556, 35854   |
| protease alcaline  | 21228, 21394   |
| purines et derives   | 15129  |
| Bsu1145  | 14593  |
| Bsu6633 (FnuDII)   | 6633   |
| D-ribose   | 31093, 31095, 21356,<br>21360, 21951, 21952,<br>31091, 31092, 31094,<br>31096, 31098 |
| SAICAR (riboside 5amino4imi-<br>dazole-n-succinocarboxamide) | 15476, 15477   |
| L-sorbose  | 27860  |
| subtiline  | 6633   |
| surfactant   | 39307  |
| L-tryptophane  | 21336, 21777, 21778  |
| urate oxidase (uricase)                                      | 21183  |
| xanthosine   | 15039, 15044   |

(d'après Edwards, 1988)

Tableau 13 (suite)

réside dans la sécrétion de protéines hétérologues, la production de surfactants biodégradables (Cooper et al., 1981), et la production d'acides aminés. Le potentiel de B. subtilis comprend également la production et la sécrétion d'enzymes d'autres Bacilli résistantes à des conditions extrêmes, telles que les enzymes des Bacilli thermophiles ou alcalophiles (Fritze et al., 1990). Il est encore remarquable que B. subtilis se prête relativement bien à la technologie des cellules immobilisées, vu que la plupart des produits de B. subtilis sont associés avec la fin de la phase exponentielle de croissance. Des cellules immobilisées et bloquées à ce stade pourraient donc se révéler très utiles pour l'élaboration de procédés continus. L'utilisation de cellules immobilisées de B. subtilis a été rapportée par Kokubu et coll. (1978) pour la production de bacitracine, et par Mosbach et coll. (1983) pour la production d'insuline. Cependant, comme souligné par Chibata et coll. (1986), les techniques d'immobilisation doivent être mises au point empiriquement pour chaque type de cellule ; de plus, la méthode d'immobilisation optimale dépend également des substrats ainsi que des produits impliqués. Bien que l'utilisation de cellules immobilisées de B. subtilis soit encore très limitée, le potentiel d'une telle technologie apparaît très grand, qui permettrait le passage d'une production par *batch* à une production en continu dont les avantages industriels sont bien établis.

## **b) DEVELOPPEMENT DES PROGRAMMES DE SEQUENCE DES GRANDS GENOMES : LE CAS DE BACILLUS SUBTILIS**

### Développement des programmes de séquence

Le premier projet international de séquençage d'un grand génome a été celui du génome humain en 1987. Cependant, une entreprise de cette envergure a rencontré et rencontre de nombreux opposants, tant pour des raisons d'ordre scientifique et éthique que pour des problèmes de

financement. Parmi les principaux arguments lancés à son encontre, on peut en particulier citer la très grande taille de ce génome -3 10<sup>6</sup> kb-, soit environ 1000 fois la taille d'E. coli (4750 kb, Bachmann, 1990) ou de B. subtilis (4250 kb, Amjad et al., 1990). De plus, il apparaît que 95% du génome des eucaryotes supérieurs sont constitués d'archives non fonctionnelles, ce qui pose un problème majeur de stockage, d'utilisation, et d'analyse des données. La présence d'introns au milieu des gènes exprimés complique également l'identification des séquences codantes et des protéines correspondantes. En outre, la détermination de la séquence de ces organismes est d'autant plus difficile à établir que le taux de polymorphisme qui y est observé est élevé.

En fait, tout programme de séquençage exhaustif du génome d'un organisme ne peut prendre sa véritable dimension que s'il est possible d'analyser, et surtout de comparer, les séquences obtenues avec celles provenant d'autres organismes afin de pouvoir tisser un réseau inter-espèces de correspondances fonctionnelles. Une telle réalisation permettrait d'identifier rapidement les produits des nombreuses phases ouvertes de lecture découvertes, et ce, principalement chez les eucaryotes où le problème des introns est un inconvénient majeur qui peut rendre extrêmement difficile l'analyse des séquences ; en outre, une telle réalisation serait d'un prix inestimable quant à la compréhension générale de l'évolution.

Il semble donc très important de pouvoir mener à bien de tels grands projets de séquençage pour chaque type d'organisme. Le choix d'E. coli et de B. subtilis est approprié vu que E. coli est une bactérie à gram(-) représentative du genre des entérobactéries, et que B. subtilis est une bactérie à gram(+) qui est le modèle de référence utilisé pour l'étude de la sporulation et l'étude des systèmes de régulation et d'expression des gènes chez les gram(+) en général.

Le projet de séquençage qui à l'heure actuelle est le plus avancé est celui du génome de S. cerevisiae, ce qui apparait du reste très justifié au vu des nombreux intérêts que revêt ce microorganisme. En outre, de très nombreux gènes de S. cerevisiae, tels que les gènes *RAS*, *CDC*, ... sont très fortement homologues à des gènes humains, ce qui explique en partie que la communauté scientifique qui s'intéresse à l'étude de cet organisme soit très large. De plus, S. cerevisiae est devenu l'organisme de choix pour l'étude des phénomènes qui régissent la vie de la cellule eucaryote, comme par exemple l'expression des gènes, la différenciation cellulaire, la sécrétion, ou la modulation de l'activité des protéines, etc ... Il n'est donc pas étonnant que l'un des attrait de la recherche fondamentale sur les levures réside dans les possibles applications médicales contre les grandes maladies telles que le cancer ou le SIDA, et que la connaissance suivie de l'analyse fonctionnelle de la séquence nucléotidique de ce génome apparaisse logiquement comme un préalable obligé au séquençage du génome humain. En fait, S. cerevisiae semble être le candidat idéal pour un tel projet à plusieurs points de vue : la levure est l'eucaryote le mieux connu, il est de plus très facile d'y procéder à l'étude des gènes par génétique inverse. S. cerevisiae a en outre un génome relativement petit de  $15 \cdot 10^6$  bp, soit 3 fois plus grand que celui d'E. coli, 10 fois plus petit que celui de la drosophile, et surtout 200 fois plus petit que celui de l'homme ou de la souris. Le plus important est peut-être l'observation que, tout comme chez les bactéries, il n'existe chez S. cerevisiae que très peu de séquences non codantes, ainsi que peu d'introns et peu de séquences répétées.

Les programmes de séquençage des grands génomes ne seraient cependant pas complets sans un représentant du règne végétal, étant donné l'importance primordiale de l'agriculture dans tous les systèmes économiques. Le choix s'est porté non pas sur une plante de première importance économique comme le maïs, mais sur Arabidopsis thaliana qui

est une "petite mauvaise herbe anodine". A. thaliana est en effet le végétal qui a le plus petit génome, composé "uniquement" de  $100 \cdot 10^6$  de paires de bases, et qui a l'avantage de ne comprendre que peu de séquences répétées, ce qui permet en particulier de corréliser carte génétique et carte physique. En outre, A. thaliana est une petite plante très prolifique pour laquelle un système de transformation existe. Les implications de la réalisation d'un tel programme de recherche incluent tous les domaines de la physiologie végétale, et en particulier le métabolisme, les réponses aux *stimuli* extérieurs, le développement, la différenciation, etc ...

#### Le cas de B. subtilis

L'idée de créer un programme international de recherche pour séquencer le génome de B. subtilis a pris naissance en Juin 1987. Basé sur une coopération internationale étroite, ce projet a débuté en 1989 en Europe avec 5 laboratoires auxquels l'une des régions suivantes a été assignée : *argA*(100°)-*nprE*(130°), *nprE-polC*(150°), *polC-dnaA*(165°), *ilvA*(195°)-*lys*(210°), *gerB*(315°)-*sacS*(334°). La stratégie utilisée pour mener à bien ce projet consiste à couvrir dans un premier temps chaque région à l'aide de clones chevauchants. Quatre banques différentes du génome de B. subtilis ont été construites dans 3 vecteurs différents :  $\lambda$ FixII,  $\Phi$ 105, et YACs (Yeast Artificial Chromosomes). Deux banques de  $\lambda$ FixII (Stratagene) ont été créées, une au Trinity College (Dublin) et une à l'Institut Pasteur (Paris) ; cette dernière a été distribuée aux autres laboratoires. Une banque de  $\Phi$ 105 a été créée à Oxford et une banque de YACs a été créée à Jouy-en-Josas (Serror et al., 1990). L'avantage du phage  $\lambda$ FixII en tant que vecteur de clonage est que des fragments allant de 15 à 23 kb en taille peuvent y être clonés. Des fragments de cette taille sont encore compatibles avec une stratégie de séquençage en *shotgun*. De plus, du fait qu'ils lysent les cellules qu'ils infectent, les vecteurs dérivés du phage  $\lambda$  sont capables de transporter certaines

séquences toxiques qui seraient délétées rapidement dans un autre type de vecteur.

Les standards de qualité retenus dans le cadre du programme de séquençage du génome de B. subtilis sont très stricts et impliquent, au détriment de la vitesse d'acquisition, l'obtention d'un taux idéal de zéro erreur. En outre, préalablement au séquençage, la carte physique des clones à séquencer est vérifiée par rapport à l'ADN de la souche de départ afin de remédier à d'éventuels remaniements, comme la présence sur le même clone de fragments non adjacents sur le chromosome. Enfin, les séquences sont soumises aux banques de données EMBL ou GenBank par segments de 10 à 20 kb.

L'intérêt direct d'un tel programme pour les applications biotechnologiques de B. subtilis en particulier inclue la détermination des gènes impliqués dans la sécrétion, l'identification des gènes régulateurs des enzymes dégradatives (système *deg*, ...), de nouvelles protéases, de nouveaux antibiotiques, de gènes impliqués dans la régulation catabolique, ou de terminateurs de transcription efficaces afin de permettre l'utilisation de promoteurs forts tout en évitant le danger d'une amplification éventuellement toxique de la transcription des gènes adjacents ; sans oublier la découverte de nouveaux gènes non détectables par les techniques classiques de génétique ou de séquences *enhancer* impliquées dans la régulation traductionnelle, et sans oublier l'étude des mécanismes de la régulation traductionnelle ou de la répression catabolique (Miwa and Fujita, 1990 ; Henkin et al., 1991), ou encore l'étude des séquences d'attachement des ribosomes, comme illustré par Sharp et coll. (1990). On pourra par exemple mettre en place tout au long de ce projet des tests systématiques simples pour détecter la présence, sur les phages recombinants isolés, de gènes particulièrement importants tels que les protéases.

## II) ISOLEMENT DE LA FORME SIMPLE BRIN D'UN PHAGE FILAMENTEUX DEFECTIF DE LA SOUCHE NCIB 6444 DE CLOSTRIDIUM ACETOBUTYLICUM

### A) BIOLOGIE DE CLOSTRIDIUM ACETOBUTYLICUM

*C. acetobutylicum* est, d'après le "Bergey's Manual of Systematic Bacteriology", une bactérie sporulante à gram(+), mobile, et de dimensions 0,5-0,9 x 1,6-6,4  $\mu\text{m}^2$ . La présence de granulose, un polysaccharide non branché contenant des molécules de D-glucopyranose liées par une liaison  $\alpha$ -1,4, est souvent observée dans les cellules de *C. acetobutylicum*. Ses produits de fermentation incluent l'acétate, le butyrate, le lactate, le butanol, l'acétone, l'éthanol, le dioxyde de carbone et l'hydrogène. Lors de la croissance en phase logarithmique, les produits principaux sont l'acétate et le butyrate, alors que durant la phase stationnaire, ce sont l'acétone et le butanol qui sont prédominants (O'Brien and Morris, 1971 ; Jones et al., 1982).

*C. acetobutylicum* est un organisme saccharolytique (Mead, 1971). La souche ATCC 824 de *C. acetobutylicum* (souche type) nécessite pour assurer sa croissance une source fermentable de carbone ainsi que de la biotine et de l'acide paraaminobenzoïque (Soni et al., 1987). Bien que *C. acetobutylicum* soit une bactérie anaérobie stricte, de courtes expositions à l'oxygène ne sont pas léthales (O'Brien and Morris, 1971). La fermentation de *C. acetobutylicum* s'effectue en deux phases : l'acidogénèse et la solvantogénèse. Les différentes voies métaboliques impliquées dans ces deux phases ont été résumées par Jones et Woods (1986). Le butyrate et l'acétate, produits de la phase d'acidogénèse, sont repris par les cellules lors de la solvantogénèse pour y être métabolisés en butanol et en acétone. Parmi les différents facteurs qui induisent la solvantogénèse, le pH externe joue un rôle central (Bahl et al., 1982 ; Nishio et al., 1983 ; Jones and Woods, 1986) car un faible pH coïncide avec la transition du stade

d'acidogénèse au stade de solvantogénèse, mais n'est pas pour autant le facteur inducteur (Long et al., 1984). Les mécanismes d'induction de la phase de solvantogénèse ont du reste suscité de nombreuses études mais sont encore mal connus.

Les solvants produits par C. acetobutylicum lui sont toxiques, mais il apparait que le butyrate est le produit de fermentation dont la toxicité est la plus grande. En effet, ce dernier détruit le gradient de protons et désénergétise la membrane (Bryant and Blaschek, 1988), la forme non dissociée de l'acide étant la forme capable de traverser la membrane et d'exercer l'effet inhibiteur (Wang and Wang, 1984). A faible pH, le butyrate est principalement sous la forme non dissociée, ce qui explique que la concentration inhibitrice décroît lorsque le pH décroît. Le début de la phase de solvantogénèse a pu être relié à une baisse du pH et il est généralement admis que la solvantogénèse représente un mécanisme de détoxification qui tend à réduire la concentration du butyrate et d'autres acides organiques.

Un autre aspect important de la biologie de C. acetobutylicum concerne le processus de sporulation. En effet, les cellules de C. acetobutylicum évoluent spontanément, par l'intermédiaire d'un processus de dégénérescence, de cellules capables de produire des solvants en cellules capables uniquement de produire des acides organiques. Une perte de la capacité à former des solvants est du reste liée à une perte de la capacité à sporuler, ce qui a permis de mettre au point des tests visuels simples pour détecter les colonies susceptibles d'être des producteurs efficaces de solvants (Adler and Crow, 1987).

Les modifications morphologiques qui surviennent lors de la solvantogénèse et lors de la sporulation ont été étudiées chez la souche P262 de C. acetobutylicum étant donné que chez cette dernière, les modifications cytologiques associées avec les différentes phases de la fermentation sont beaucoup plus marquées que chez la souche ATCC 824 (Long et al., 1984b).

Parmi celles-ci, on distingue l'accumulation de granulose, la production de la capsule, et la formation de l'endospore.

L'accumulation de granulose est initiée lorsque des conditions défavorables pour la croissance sont rencontrées (fortes concentrations d'acides organiques ou de solvants), en conjonction avec un excès de source de carbone. Le *pool* de granulose décroît lors de la maturation de la spore (Preiss, 1984 ; Woods and Jones, 1986). La production de la capsule est un autre évènement associé à la conversion de la cellule végétative en forme clostridiale. La capsule est principalement composée d'un polysaccharide extracellulaire ayant un haut degré d'acétylation. Ce dernier est produit durant la solvantogénèse et pourrait représenter un déversoir pour les composés oxydés (Haggstrom and Forsberg, 1986). Le début de la formation de l'endospore est généralement accompagnée de la production de solvants ; toutefois, l'isolement de mutants asporogènes capables de produire de l'acétone et du butanol a permis de démontrer que ces deux évènements sont indépendants, de même que l'apparition de la forme clostridiale, la production de granulose, ou de la capsule (Jones and Woods, 1986)

## **B) THEMES ET BUTS DE RECHERCHE**

Comme indiqué précédemment (paragraphe I), C. acetobutylicum est utilisé en biotechnologie depuis de nombreuses années. Cependant, certaines de ses caractéristiques doivent être modifiées pour que la fermentation Acétone-Butanol redevienne un procédé rentable d'un point de vue économique. Le clonage de gènes dans C. acetobutylicum est donc une étape critique dans la mise au point des procédés de production de solvants à partir de la biomasse.

L'un des outils nécessaires pour permettre le clonage de gènes dans une bactérie est, outre la disponibilité d'un système de transformation adéquat, l'existence d'une méthode efficace pour isoler l'ADN plasmidique. Chez les Clostridia,

la génétique est difficile du fait de la présence de nucléases très actives (Blaschek and Klacik, 1984 ; Lin and Blaschek, 1984 ; Boosermsuwong and Blaschek, 1986) et du fait de l'absence d'une méthode efficace pour isoler de l'ADN provenant d'éléments génétiques extrachromosomiques. Le but de ce travail était donc de développer une telle méthode.

## **C) MATERIELS ET METHODES**

### **1) SOUCHES BACTERIENNES**

La souche de C. acetobutylicum NCIB 6444 a été fournie par Rogers (Département de Microbiologie, Université du Minnesota).

Les plasmides contenus dans la souche d'E. coli V517 ont été utilisés comme marqueurs de poids moléculaires (Macrina et al., 1978).

### **2) MILIEUX ET CONDITIONS DE CULTURE**

C. acetobutylicum a été cultivé dans du milieu CBM (10 g/l glucose, 4 g/l hydrolysate de caséine, 0,2 g/l MgSO<sub>4</sub>.7H<sub>2</sub>O, 0,01 g/l MnSO<sub>4</sub>.4H<sub>2</sub>O, 0,01 FeSO<sub>4</sub>.7H<sub>2</sub>O, 10<sup>-3</sup> g/l acide paraaminobenzoïque, 2 10<sup>-6</sup> g/l biotine, 10<sup>-3</sup> g/l thiamine HCl, 0,5 g/l K<sub>2</sub>HPO<sub>4</sub>, 0,5 g/l KH<sub>2</sub>PO<sub>4</sub>) (O'Brien and Morris, 1971). C. acetobutylicum a été cultivé en anaérobiose (85% N<sub>2</sub>, 10% CO<sub>2</sub>, 5% H<sub>2</sub>) dans une chambre anaérobie (Coy Laboratory Products, Inc., Ann Harbor, Michigan). C. acetobutylicum a été stocké à 4°C à l'état de spores en suspension dans de l'eau distillée.

### **3) ISOLEMENT D'ADN PAR ELECTROPORATION**

Les électroporations ont été réalisées à l'aide d'un Bio-Rad Gene Pulser connecté à un Pulse Controller (Bio-Rad Laboratories, Richmond, Californie). Les cellules ont été cultivées jusqu'à une densité optique adéquate, lavées dans de l'eau distillée, puis resuspendues dans de l'eau distillée à raison de 1/25 à 1/100 du volume initial. Après

électroporation, les cellules ont été transférées quantitativement dans un tube Eppendorf et centrifugées 3 min. Le surnageant a alors été transféré dans un nouveau tube et de l'EDTA, du SDS, et de la protéinase-K ont été ajoutés de façon à obtenir une concentration finale de 12,5 mM, 0,1%, et 0,1 µg/ µl, respectivement. Une extraction au phénol et au chloroforme a été réalisée après 15 min d'incubation à 37°C. L'ADN ainsi purifié a alors été précipité dans de l'éthanol et dissous dans du tampon TE (10 mM Tris pH 8,0, 1 mM EDTA).

#### **D) RESULTATS ET DISCUSSION**

Comme indiqué Tableau 1, l'électroporation est un moyen très efficace pour isoler de l'ADN du phage déficient CAK1 (pDM6) de la souche NCIB 6444 de C. acetobutylicum (Kim and Blaschek, 1991). L'accroissement observé de la quantité d'ADN isolé après l'étape de traitement au SDS-EDTA-protéinase k (Tableau 1) peut s'expliquer par la nécessité de détruire l'enveloppe protéique des particules de phages pour libérer l'ADN.

L'ADN isolé par électroporation (cf. Figure 1) correspond à la forme simple brin du phage CAK1, comme démontré par la dégradation spécifique de ces molécules par la nucléase S1 (Figure 1bis). Cet ADN est de plus exempt de toute contamination par de l'ADN chromosomique. Une explication plausible pour interpréter cette observation est que l'électroperméabilisation facilite l'extrusion des particules de phages, alors que la forme répliquative double brin (forme m ou d, cf. Figure 1) demeure intracellulaire. En effet, le phage CAK1 est un phage remarquablement homologue au phage M13 d'E. coli et est libéré des cellules de C. acetobutylicum tout au long de la phase de croissance (cf. Figure 1, puits 4) (Kim and Blaschek, 1991). Cette possibilité semble en outre être confirmée par l'observation que divers plasmides de plusieurs souches de Clostridium

**ISOLEMENT D'ADN PAR ELECTROPORATION DE  
CLOSTRIDIUM ACETOBUTYLICUM NCIB 6444**

| Echantillon<br>(a)  | Constante de<br>temps | Voltage | Quantité d'ADN<br>µg/ml de cellules |
|---|-----------------------|---------|-------------------------------------|
| Electroporation<br>traitement au SDS,<br>EDTA, protéinase-k<br>extraction au phénol | 18.6 ms               | 2,49 kV | 2,10                                |
| Cellules témoin<br>traitement au SDS,<br>EDTA, protéinase-k<br>extraction au phénol |                       |         | 0,35                                |
| Electroporation<br>extraction au phénol   | 27,0 ms               | 2,50 kV | 0,57                                |
| Cellules témoin<br>extraction au phénol   |                       |         | 0,14                                |

Les cellules ont été concentrées 25 fois

Tableau 1



Figure 1 : Isolement d'ADN par électroporation de *Clostridium acetobutylicum* NCIB 6444

1) marqueur de poids moléculaires V517, 2) ADN du phage CAK1 isolé par une méthode de lyse alcaline et purifié en gradient de chlorure de césium, 3) ADN du phage CAK1 isolé par électroporation, 4) même cellules qu'en 3 sans électroporation (la différence d'intensité entre les bandes des puits 3 et 4 reflète la différence en quantité d'ADN isolé)

a) 2,1 kb, b) 2,8 kb, c) 3,2 kb, d) 4,1 kb, e) 5,3 kb  
 f) 5,8 kb, g) 7,6 kb, h) 56,4 kb  
 m) monomère du phage CAK1, ss) forme simple brin du phage CAK1, d) dimère du phage CAK1

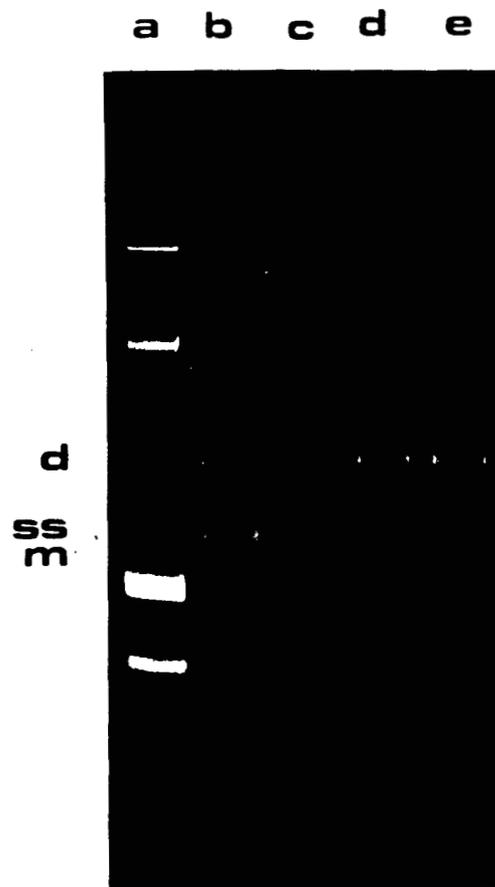


Figure 1bis : Identification de la forme simple brin du phage CAK1 par traitement à la nucléase S1  
 a) marqueur de poids moléculaires V517, b) ADN du phage CAK1 isolé par une méthode de lyse alcaline et purifié en gradient de chlorure de césium (témoin), c) idem incubé pendant 10 min à température ambiante en présence de  $14,3 \cdot 10^{-5}$  Unité Kunitz de nucléase S1 par ul, d) idem en présence de  $14,3 \cdot 10^{-4}$  U/ ul, e) idem en présence de  $14,3 \cdot 10^{-3}$  U/ ul  
 m) monomère, ss) forme simple brin, d) dimère

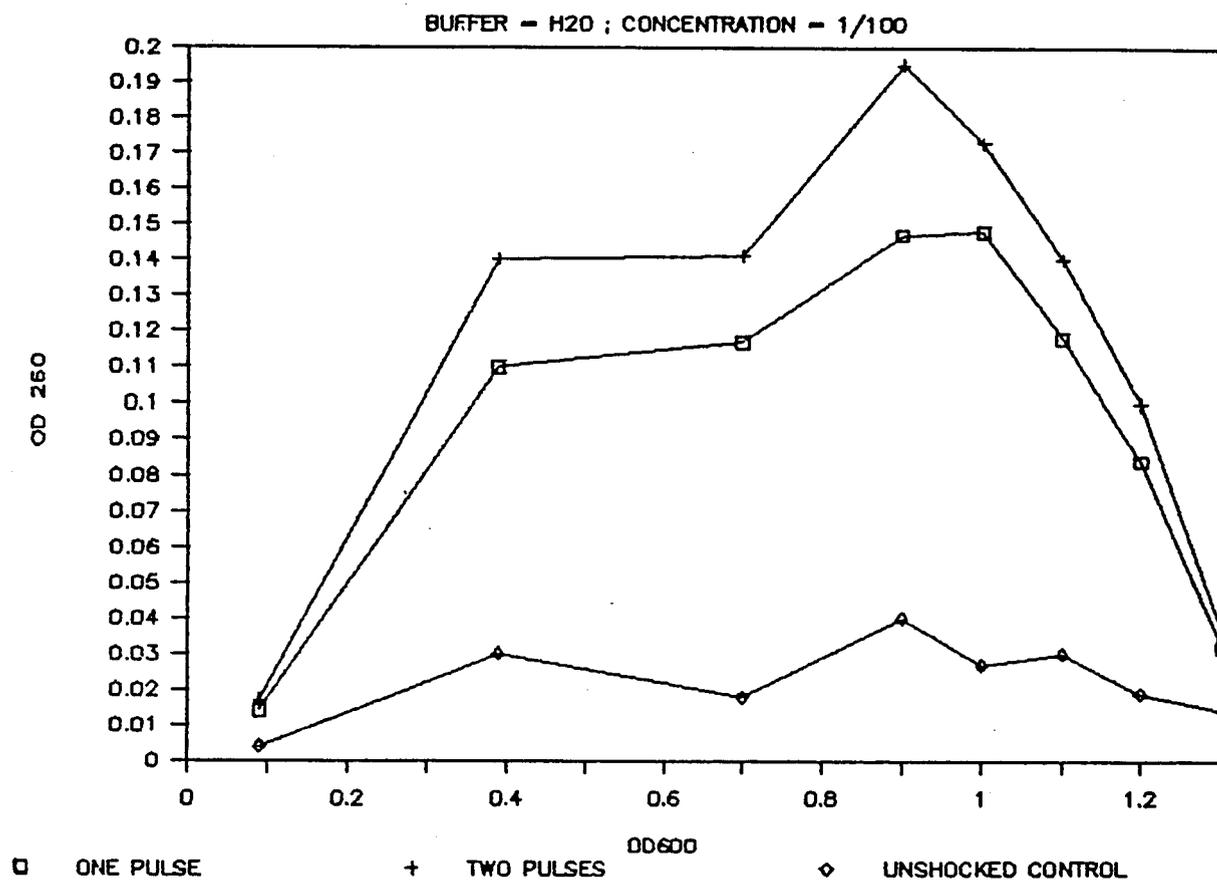


Figure 2 : Isolement d'ADN par électroporation de *Clostridium acetobutylicum* NCIB 6444

perfringens n'ont pas pu être isolés à partir des techniques d'électroperméabilisation.

Des résultats préliminaires reportés Tableau 1 ont montré qu'un choc électrique unique permet d'isoler 6 fois plus d'ADN du phage CAK1, comparativement à l'échantillon témoin. Des expériences complémentaires ont été conduites pour optimiser cette méthode. Les cellules ont été récoltées à différents moments tout au long de la phase de croissance, lavées et concentrées 100 fois dans de l'eau distillée stérile. La quantité d'ADN ainsi isolée a été estimée par la détermination de la densité optique à 260 nm de l'échantillon purifié. La courbe de la densité optique à 260 nm en fonction de la densité optique à 600 nm indique que les plus fortes concentrations d'ADN sont obtenues au moment de la croissance où la densité optique à 600 nm atteint la valeur de 0,9 (cf. Figure 2).

#### **E) CONCLUSION**

Il apparait que la méthode d'isolement de l'ADN par électroporation est particulièrement efficace étant donné qu'elle permet d'obtenir des rendements de l'ordre de 2  $\mu\text{g}$  d'ADN par ml de cellules, alors que les méthodes de lyse alcaline suivies d'une centrifugation en gradient de césium offrent des rendements de l'ordre de 0,01 à 0,08  $\mu\text{g}$  d'ADN par ml de cellules. Il est possible que le succès de cette méthode réside dans une plus faible exposition de l'ADN aux nucléases.

F) ARTICLE 1

A.A. Y. KIM, A.A. VERTES, and H.P. BLASCHEK. 1990.  
ISOLATION OF A SINGLE-STRANDED PLASMID FROM  
*CLOSTRIDIUM ACETOBUTYLICUM* NCIB 6444. App. Env.  
Microbiol. 56:1725-1728

## Isolation of a Single-Stranded Plasmid from *Clostridium acetobutylicum* NCIB 6444

AUGUSTINE Y. KIM,<sup>1</sup> ALAIN A. VERTES,<sup>2</sup> AND HANS P. BLASCHEK<sup>1\*</sup>

Department of Food Science, 580 Bevier Hall, University of Illinois, 905 S. Goodwin Avenue, Urbana, Illinois 61801,<sup>1</sup>  
and Unite des Anaerobes, Institut Pasteur, 75724 Paris Cedex 15, France<sup>2</sup>

Received 1 November 1989/Accepted 3 April 1990

The cryptic plasmid pDM6 was isolated from late exponential-phase cells of *Clostridium acetobutylicum* NCIB 6444 by either alkaline lysis or electroporation. The application of high voltage during electroporation resulted in higher DNA yield than did the alkaline lysis procedure. However, electroporation-induced plasmid release generated high amounts of single-stranded DNA compared with the alkaline lysis procedure, which generated both double-stranded DNA (monomer and dimer forms) and single-stranded DNA.

Interest in *Clostridium acetobutylicum* has been renewed because of the solvent-producing capability (acetone-butanol fermentation) of this microorganism. This has encouraged the development of genetic systems for this microorganism in order to understand the metabolic pathway for solvent formation and thereby increase solvent production during fermentation (9, 18). Numerous *C. acetobutylicum* genes have been cloned in *Escherichia coli*, and genetic analysis has been carried out in this latter microorganism (6, 9). In order to better understand the *C. acetobutylicum* genetic system, the development of vector and transformation systems continues to be an important area of research for this species. The development of an efficient means for isolation of plasmid DNA is also an important prerequisite for being able to carry out plasmid genetics in *C. acetobutylicum*.

Plasmid DNA has been isolated from several *Clostridium* species, including *C. perfringens*, *C. tetani*, *C. cochlearium*, *C. butylicum*, and *C. acetobutylicum* (18). However, most plasmid isolation procedures have been developed for *C. perfringens* (4, 18). Although indigenous plasmid DNA was previously isolated from *C. acetobutylicum* NCIB 6444 (19), the recovered plasmid was not further characterized.

The objective of this work was to develop an efficient plasmid DNA isolation procedure for *C. acetobutylicum*. This involved the preliminary characterization of the recovered pDM6 cryptic plasmid in preparation for the construction of a versatile *E. coli*-*C. acetobutylicum* shuttle vector. Because of the potential instability of clostridial shuttle vectors (10, 17), the developed vector should contain a host-related replication origin (10) from an indigenous *C. acetobutylicum* plasmid instead of from a wide-host-range plasmid (15).

### MATERIALS AND METHODS

**Bacterial strains, media, and culture conditions.** The bacterial strains used in this study were *C. acetobutylicum* ATCC 824, NCIB 6444, and P262. The latter two strains were provided by P. Rogers (Department of Microbiology, University of Minnesota, Minneapolis) and D. Woods (Department of Microbiology, University of Cape Town, South Africa), respectively. All *C. acetobutylicum* cultures were maintained in Cooked Meat Medium (10) (Difco Laboratories, Detroit, Mich.) or as a spore suspension in sterilized

distilled H<sub>2</sub>O at 4°C. For plasmid isolation, *C. acetobutylicum* was grown in either Trypticase Glucose Yeast Extract medium (10) or Clostridial Basal Medium (16) at 37°C under anaerobic conditions (85% N<sub>2</sub>, 10% CO<sub>2</sub>, 5% H<sub>2</sub>; Coy Laboratory Products, Inc., Ann Arbor, Mich.) and harvested at late exponential phase (optical density at 600 nm, 0.9 to 1.0). Growth was measured by monitoring the optical density at 600 nm by using a Spectronic 20 spectrophotometer (Bausch & Lomb, Inc., Rochester, N.Y.).

**Plasmid isolation. (i) Alkaline lysis procedure.** Plasmid DNA from *C. acetobutylicum* was isolated by using a modification of the alkaline lysis method previously reported (10). Cell cultures were washed with one-half volume of glucose-Tris-EDTA buffer (10) before lysozyme treatment. To reduce RNA contamination in the plasmid DNA preparation, DNase-free RNase (50 µg/ml; Sigma Chemical Co., St. Louis, Mo.) was added during the lysozyme treatment.

**(ii) Electroporation-induced procedure.** Electroporation was performed by using a Gene Pulser connected to a pulse controller (Bio-Rad Laboratories, Richmond, Calif.). Cells were grown to late exponential phase (optical density at 600 nm, 0.9) and centrifuged for 10 min at 5,000 × *g* by using an RC-2 Centrifuge (Ivan Sorvall, Inc., Norwalk, Conn.). The cell pellet was washed with distilled water and suspended in 1/25 to 1/100 of the original cell culture volume. The cell suspension was kept in an ice bath for 5 min before electroporation, and 0.8 ml was subjected to one or two electrical pulses with the Gene Pulser set at 2,500 V and 25 µF. The electroporated cells were quantitatively transferred to a 1.5-ml microcentrifuge tube and centrifuged for 3 min. The supernatant was transferred to a fresh tube, and EDTA, sodium dodecyl sulfate, and proteinase K (Sigma) were added to final concentrations of 12.5 mM, 0.1%, and 0.1 µg/ml, respectively. The electroporated cell suspension was incubated for 15 min at 37°C in an Incu-Block (Denville Scientific, Inc., Denville, N.J.) and subsequently purified by using phenol and chloroform extraction (10). Plasmid DNA was precipitated by adding 2 volumes of cold ethanol (99.9%), and the pellet was subsequently suspended in Tris-EDTA buffer (10) containing RNase (20 µg/ml).

**Detection and analysis of plasmid DNA.** Plasmid DNA was analyzed by 0.7% (wt/vol) agarose gel electrophoresis as described previously (10). Individual DNA bands in agarose were excised and recovered by using the procedure described by Benson (3). All restriction enzymes, as well as BAL 31 and S1 nucleases, were purchased from Bethesda

\* Corresponding author.



FIG. 1. Plasmid DNA recovery from *C. acetobutylicum* NCIB 6444 using various methods. Lanes: 1, molecular weight markers from *E. coli* V517 (13); 2, plasmid pDM6 recovered by alkaline lysis; 3, DNA recovered by electroporation; 4, DNA recovered without electroporation. d, Dimer form; ss, single-stranded DNA; m, monomer.

Research Laboratories (Gaithersburg, Md.) or Sigma, and digestions were carried out as described by the manufacturer or Maniatis et al. (14). Southern transfer was performed as described by Maniatis et al. (14) with modifications (1). Plasmid DNA was transferred to nitrocellulose by diffusion with or without a prior denaturation step. Biotin-labeled and  $^{32}\text{P}$ -labeled probes were prepared by nick translation (Bethesda Research Laboratories) using either Biotin-7-dATP or  $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$  (Du Pont Chemicals, Inc., Wilmington, Del.), respectively. The filter was hybridized, and the DNA was detected by either the BluGENE Nonradioactive Nucleic Acid Detection System Kit (Bethesda Research Laboratories) or by autoradiography on X-ray film (Sigma).

## RESULTS

**Isolation of plasmid DNA from *C. acetobutylicum* by using the modified alkaline lysis procedure.** Plasmid DNA could not be isolated from *C. acetobutylicum* ATCC 824 or P262 by using the modified alkaline lysis procedure, which is in agreement with previously reported results (8, 12). The recovery of chromosomal DNA with the unmodified protocol may be due to an insufficiently alkaline pH during lysis. *C. acetobutylicum* produces acetate and butyrate during growth (9, 12, 18), which may reduce the pH of the alkaline lysis solution. A washing of the cell pellet with one-half volume of glucose-Tris-EDTA buffer reduced the amount of chromosomal DNA recovered. An agarose gel of plasmid DNA isolated from *C. acetobutylicum* NCIB 6444 using the modified alkaline lysis procedure showed three distinct bands (Fig. 1, lane 2). Truffaut and Sebald (19) suggested that *C. acetobutylicum* NCIB 6444 contained more than two plasmids. Restriction enzyme digestion analysis revealed that the upper band (Fig. 1, lane 2) corresponded to the double-stranded dimer form of plasmid pDM6, while the lower band corresponded to the double-stranded monomer form of plasmid pDM6 (data not shown). The intermediate band was initially believed to be a different species of plasmid.

**Electroporation-induced plasmid DNA recovery.** Electro-

TABLE 1. Electroporation-induced plasmid DNA release from *C. acetobutylicum* NCIB 6444

| Sample <sup>a</sup>                  | DNA recovered ( $\mu\text{g}$ of DNA/ml of cell input) <sup>b</sup> |
|--------------------------------------|---|
| Electroporated cells <sup>c</sup>    | 2.1   |
| Nonelectroporated cells <sup>c</sup> | 0.4   |
| Electroporated cells <sup>d</sup>    | 0.6   |
| Nonelectroporated cells <sup>d</sup> | 0.1   |

<sup>a</sup> Concentrated 25-fold.

<sup>b</sup> Determined by measuring the  $\text{OD}_{260}$ .

<sup>c</sup> Added to a solution containing 12.5 mM EDTA, 0.1% SDS, and 0.1  $\mu\text{g}$  of proteinase K per ml and solvent extracted.

<sup>d</sup> Solvent extracted only.

poration was a highly efficient method for the recovery of plasmid DNA from *C. acetobutylicum*. Plasmid DNA was recovered from *C. acetobutylicum* NCIB 6444 after electroporation (Fig. 1, lane 3) as well as in the absence of electroporation (Fig. 1, lane 4). The DNA band recovered with or without electroporation corresponded to the middle (intermediate) plasmid band recovered from *C. acetobutylicum* NCIB 6444 when the modified alkaline lysis procedure was used (Fig. 1, lane 2). The application of a single pulse resulted in a five- to six-fold increase in recovered DNA compared with the nonelectroporated control (Table 1). Incubation of the cell suspension with EDTA, sodium dodecyl sulfate, and proteinase K significantly increased the DNA yields from both electroporated and nonelectroporated cells. The effect of an additional high-voltage pulse on the recovery of plasmid DNA was not significant (Fig. 2). It is interesting that the peak for DNA recovery corresponded with the stationary phase and the point at which the culture switches from acidogenesis to solventogenesis (9, 18).

**Characterization of pDM6 single-stranded DNA.** In order to characterize the DNA band present between the dimer and monomer configurations of pDM6 (Fig. 1, lane 2), the lysate was subjected to digestion by BAL 31 exonuclease, which has a specificity for the linear form of double-stranded DNA (dsDNA), as well as linear and circular single-stranded DNA (ssDNA) (11). BAL 31 exonuclease digestion successfully removed the intermediate DNA band with covalently closed circular intact dimer and monomer dsDNA remaining (Fig.

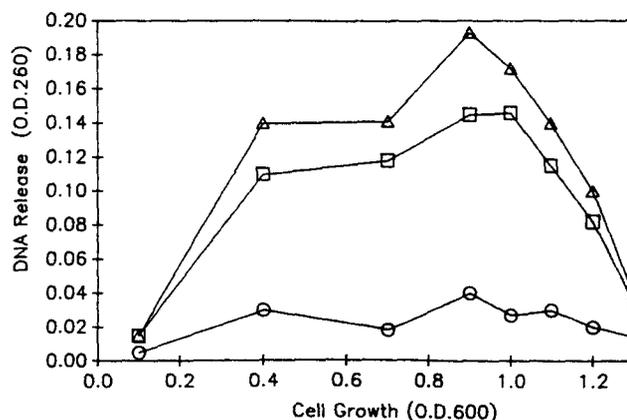


FIG. 2. Effect of zero (O), one (□), or two (Δ) high-voltage pulses on the release of DNA as measured by determining the optical density at 260 nm over the course of the growth curve of *C. acetobutylicum* 6444.

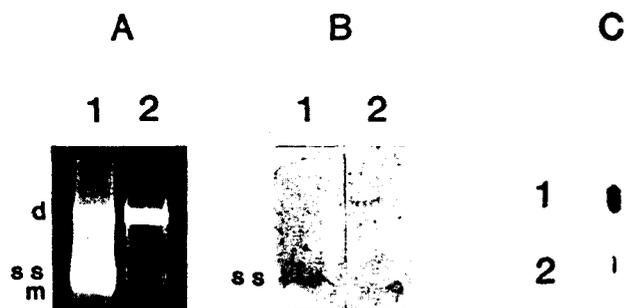


FIG. 3. (A) Agarose gel electrophoresis of plasmid pDM6 DNA recovered by using the modified alkaline lysis method without BAL 31 exonuclease digestion (lane 1) or with BAL 31 exonuclease digestion (lane 2). (B) Southern hybridization of DNA from an undenatured agarose gel probed with biotin-labeled pDM6 without BAL 31 exonuclease digestion (lane 1) or with BAL 31 digestion (lane 2). (C) Southern hybridization on the slot blot of undenatured DNAs recovered by either electroporation or cell leakage and probed with  $\alpha$ - $^{32}$ P-labeled pDM6. Lanes: 1, DNA recovered by electroporation; 2, DNA recovered from the supernatant of nonelectroporated cells.

3A, lane 2). A similar result was obtained when nuclease S1 (a ssDNA-specific endonuclease) digestion of electroporation-recovered DNA was used (data not shown). On the basis of these results, the intermediate DNA band appears to be a single-stranded form of the pDM6 plasmid. In order to confirm this, the DNA on the agarose gel was transferred to a nitrocellulose membrane without prior denaturation and subsequently probed with the dimer form of pDM6. Only the ssDNA band hybridized with the pDM6 probe (Fig. 3B, lane 1). No hybridization was observed after digestion with BAL 31 exonuclease (Fig. 3B, lane 2). In addition, the pDM6 DNA recovered by electroporation or from the supernatant of nonelectroporated cells was transferred to a slot blotter without a heat denaturation step and probed with pDM6. The results can be seen in Fig. 3C. These results suggest that *C. acetobutylicum* NCIB 6444 contains ssDNA which has homology with double-stranded pDM6. However, the structure and function of this ssDNA remains unknown.

## DISCUSSION

Plasmid pDM6, the only known plasmid recovered to date from *C. acetobutylicum*, appears to exist in both a double-stranded (dimer and monomer) as well as a single-stranded form. On the basis of results obtained elsewhere (7), it is possible that the single-stranded form of pDM6 is a replication intermediate. The ssDNA coliphage in *E. coli* replicates by a rolling-circle mechanism, and many small gram-positive plasmids (e.g., pC194, pT181, and pSL1) appear to replicate by a similar mechanism (7). However, in the experiments reported herein, the ssDNA form of pDM6 demonstrated some characteristics which differed from those of other ssDNAs. First, the dimer form of pDM6 predominates in the alkaline lysate. However, dimer formation is not commonly observed for other plasmids that generate ssDNA (2). Second, analysis of the molecular weight of the ssDNA formed by other plasmids demonstrated that it was half the size of dsDNA, with the migration of monomeric closed ssDNA being greater than that of the corresponding dsDNA monomer (2). When analyzed by 0.7% agarose gel electrophoresis, the pDM6 ssDNA band is located at a position above that of the monomer form. In addition, the ability of ethidium

bromide to bind with pDM6 ssDNA suggests that it contains some secondary structure (i.e., internal double-stranded structure). The structural characteristics and replication mechanism of pDM6 require additional investigation.

Electroporation involves the generation of transient pores in the outer membrane of the cell which allow for diffusion and exchange of intra- and extracellular components during the life span of the pore (5). Electroporation techniques were recently used for transformation of *C. perfringens* (1, 10), a microorganism which did not previously have an efficient means for genetic transformation. The successful isolation of plasmid DNA from *E. coli* by electroporation was reported by Calvin and Hanawalt (5). The recovery of plasmid DNA together with cellular DNA and RNA suggest that electroporation generated pores large enough to allow for the leakage of cellular materials. The greater recovery of plasmid DNA compared with cellular DNA (5) suggested that plasmid DNA may be weakly bound to the cell membrane, and weakly bound plasmid DNA may be preferentially released during pore formation. During electroporation-mediated plasmid DNA recovery from *C. acetobutylicum* NCIB 6444, only single-stranded pDM6 plasmid DNA rather than a mixture of double-stranded plasmid and genomic DNA were recovered (Fig. 1, lane 3). These results suggest that single-stranded pDM6 is not tightly bound to the cell membrane and may easily pass through the transient pore. Electroporation-induced transformation data suggest that pore formation for gram-positive microorganisms is less efficient at a given electrical field strength than it is for gram-negative microorganisms, such as *E. coli* (10). Extracellular plasmid DNA transfer through the transient pore may be limited by cell wall structure. Also, cell wall structure may limit cellular leakage during electroporation. However, ssDNA may have a configuration which allows for its passage during electroporation. The increased level of ssDNA release during late exponential phase (Fig. 2) may be a consequence of cell wall structural changes that occur during the transition of *C. acetobutylicum* from the acidogenic to the solventogenic phase. The structural properties of pDM6 ssDNA and the *C. acetobutylicum* NCIB 6444 cell wall require further examination.

The isolation of pDM6 from *C. acetobutylicum* NCIB 6444 makes it possible to construct an *E. coli*-*C. acetobutylicum* shuttle vector based on this plasmid. Shuttle vectors for gram-positive microorganisms which have been based on ssDNA-generated plasmids have not been stable in *E. coli* because of their recombinatory activities during replication (7). However, in *C. acetobutylicum* such vectors are reportedly stable (20). The vectors for *C. acetobutylicum* which have been constructed by using replication origins from heterogeneous species showed a high degree of segregational instability (20). On the other hand, the pAK201 *C. perfringens*-*E. coli* shuttle vector (10) which contains a replication origin from an indigenous *C. perfringens* plasmid demonstrates a high degree of stability during segregation. The identification of the replication origin of plasmid pDM6 and the construction of an *E. coli*-*C. acetobutylicum* shuttle vector based on this plasmid is currently under way in this laboratory.

## ACKNOWLEDGMENTS

This work was supported in part by Hatch Grant 50-0314 from the University of Illinois Agricultural Experiment Station, grant ICMB 86-0015-01 from the Illinois Corn Marketing Board, the University of Illinois Research Board Grant 1-2-69157, and State of Illinois Competitive Value-Added Grant 1-1-11963.

## LITERATURE CITED

1. Allen, S. P., and H. P. Blaschek. 1988. Electroporation induced transformation of intact cells of *Clostridium perfringens*. *Appl. Environ. Microbiol.* **54**:2322-2324.
2. Ballester, S., P. Lopez, M. Espinosa, J. C. Alonso, and S. A. Lacks. 1989. Plasmid structural instability associated with pC194 replication functions. *J. Bacteriol.* **171**:2271-2277.
3. Benson, S. 1984. A rapid procedure for isolation of DNA fragments from agarose gel electrophoresis. *Biotechniques* **2**: 66-67.
4. Blaschek, H. P. 1989. Genetic manipulation of the clostridia. *Dev. Indust. Microbiol.* **30**:35-42.
5. Calvin, N. M., and P. C. Hanawalt. 1988. High-efficiency transformation of bacterial cells by electroporation. *J. Bacteriol.* **170**:2796-2801.
6. Efstathiou, I., and N. Truffaut. 1986. Cloning of *Clostridium acetobutylicum* genes and their expression in *Escherichia coli* and *Bacillus subtilis*. *Mol. Gen. Genet.* **204**:317-321.
7. Gruss, A., and S. D. Ehrlich. 1989. The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. *Microbiol. Rev.* **53**:231-241.
8. Jones, D. T., W. A. Jones, and D. R. Woods. 1985. Production of recombinants after protoplasts fusion in *Clostridium acetobutylicum* P262. *J. Gen. Microbiol.* **131**:1213-1216.
9. Jones, D. T., and D. R. Woods. 1986. Acetone-butanol fermentation revisited. *Microbiol. Rev.* **50**:147-172.
10. Kim, A. Y., and H. P. Blaschek. 1989. Construction of an *Escherichia coli*-*Clostridium perfringens* shuttle vector and plasmid transformation of *Clostridium perfringens*. *Appl. Environ. Microbiol.* **55**:360-365.
11. Legerski, R. J., J. L. Hodnett, and H. B. Gray, Jr. 1978. Extracellular nucleases of *Pseudomonas* BAL31. III. Use of the double-stranded deoxyribonuclease activity as basis of a convenient method for the mapping of fragments of DNA produced by cleavage with restriction enzymes. *Nucleic Acids Res.* **5**:1445-1464.
12. Lin, Y., and H. P. Blaschek. 1989. Transformation of heat-treated *Clostridium acetobutylicum* with pUB110 plasmid DNA. *Appl. Environ. Microbiol.* **48**:737-742.
13. Macrina, F. L., D. J. Kopecko, K. R. Jones, D. J. Ayers, and S. M. McCowen. 1978. A multiple plasmid-containing *Escherichia coli* strain: convenient source of size reference plasmid molecules. *Plasmid* **1**:417-420.
14. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. Minton, N. P., and J. G. Morris. 1981. Isolation and partial characterization of three cryptic plasmids from strains of *Clostridium butylicum*. *J. Gen. Microbiol.* **127**:325-331.
16. O'Brien, R. W., and J. G. Morris. 1971. Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. *J. Gen. Microbiol.* **68**:307-318.
17. Roberts, I., W. M. Holmes, and P. P. Hylemon. 1988. Development of a new shuttle plasmid system for *Escherichia coli* and *Clostridium perfringens*. *Appl. Environ. Microbiol.* **54**:268-270.
18. Rogers, P. 1986. Genetics and biochemistry of Clostridia relevant to development of fermentation processes. *Adv. Appl. Microbiol.* **31**:1-60.
19. Truffaut, N., and M. Sebald. 1983. Plasmid detection and isolation in strains of *Clostridium acetobutylicum* and related species. *Mol. Gen. Genet.* **189**:178-180.
20. Young, M., N. P. Minton, and W. L. Staudenbauer. 1989. Recent advances in the genetics of the clostridia. *FEMS Microbiol. Rev.* **63**:301-326.

**G) ARTICLE 2**

Une seconde voie de recherche portant sur la fermentation acétone butanol concerne les mécanismes d'induction de la phase de solvantogénèse. Une telle voie de recherche est illustrée dans le manuscrit suivant, soumis pour publication en 1991.

**H.C. DUBOURGUIER and A.A. VERTES. 1991. PHYSIOLOGICAL CHANGES AND SOLVENT PRODUCTION IN CLOSTRIDIUM ACETOBUTYLICUM. Soumis pour publication**

PHYSIOLOGICAL CHANGES AND SOLVENT PRODUCTION IN  
*CLOSTRIDIUM ACETOBUTYLICUM*

Henri-Charles DUBOURGUIER and Alain VERTES

**SUMMARY**

The objective of this work was to study the shift from acidogenesis to solventogenesis and to relate cellular form to solventogenesis under a given set of fermentation conditions. The cellular forms studied were obtained from two distinct fermentations using pH regulation at different stages of the fermentation. Experiments were carried out using resting cells. The kinetics of granulose accumulation appears to be a potentially useful criterion to characterize the solventogenic cellular form.

**INTRODUCTION**

Acidogenesis and solventogenesis characterize the two main phases in the fermentation pattern of *Clostridium acetobutylicum* and have been intensively studied. Butyrate and acetate, end products of the acidogenesis phase, are taken up by the cells during the solventogenic phase and then further metabolized into butanol and acetone.

Among the different factors triggering solventogenesis, the external pH turns out to play a central role (Bahl et al., 1982 ; Jones and Wood, 1986 ; Nishio et al., 1983 ; Ross, 1961) but is not by itself the triggering agent as demonstrated by the early induction of solvent production by the addition of acetate and butyrate to cultures of *C. acetobutylicum* maintained at pH 5.0 (Gottschal and Morris, 1981) or the production of solvents by *C. acetobutylicum* cultures maintained at neutral pH (Holt et al., 1984). At low pH, butyrate will be mostly in the undissociated form and therefore will be able to cross the membrane (Wang and Wang, 1984), act as an uncoupler, and make the cells leaky to protons thereby collapsing the proton motive force (Herrero et al., 1985 ; Terraciano and Kashket, 1986). Butyrate being the product most inhibitory to *C. acetobutylicum* (Costa and Moreira, 1983), it is thought that solventogenesis is a detoxification mechanism that tends to reduce the level of butyrate or other inhibitory organic acids.

Solventogenesis appears nevertheless also linked to morphological and physiological changes (Long et al., 1984) that present three main stages, namely granulose accumulation, capsule production, and endospore formation. Characterization of the solventogenic cellular form under a given set of fermentation conditions appears to be a useful tool for a continuous fermentation process. The objective of this work was thus to study the shift from acidogenesis to solventogenesis and to determine a procedure for the identification of the cellular stage responsible for solventogenesis.

## MATERIALS AND METHODS

### 1) Bacterial strain

The bacterial strain used was *C. acetobutylicum* B6 which is a spontaneous butanol resistant mutant of strain 6XBD (soil isolate showing a better glucose utilization and a higher solvent produced/ glucose consumed ratio than *C. acetobutylicum* NCIB 8052 (Zygmunt, 1983). Strain B6 butanol resistance is characterized as follows : 11 g/l of butanol was necessary to induce a 50% decrease in the growth rate whereas 10 g/l of butanol was necessary to induce a 50% decrease in the growth rate of strain NCIB 8052.

### 2) Inoculum preparation

Cells were stored on Reinforced Clostridial Medium (3 g/l yeast extract, 10 g/l Lab-Lemco, 10 g/l peptone, 1 g/l soluble starch, 5 g/l dextrose, 0.5 g/l cystéine-HCl, 5 g/l NaCl, 3 g/l sodium acetate) and transferred to Fermentation Medium (40 g/l dextrose, 4.5 g/l NH<sub>4</sub>Cl, 0.5 g/l KH<sub>2</sub>PO<sub>4</sub>, 0.5 g/l K<sub>2</sub>HPO<sub>4</sub>, 0.01 g/l MnCl<sub>2</sub>.4H<sub>2</sub>O, 0.2 g/l MgSO<sub>4</sub>.7H<sub>2</sub>O, 0.08 g/l CaCl<sub>2</sub>.2H<sub>2</sub>O, 0.02 g/l NaCl, 0.01 g/l FeSO<sub>4</sub>.7H<sub>2</sub>O, 0.5 g/l asparagine.H<sub>2</sub>O, 5 g/l yeast extract, 1 ml/l of 0.2% resazurine solution, pH 6.2 with CaCO<sub>3</sub>) following heat shock treatment (100°C for 10 min) ; the culture was then transferred on the same medium every two days for several days. A flask of 160 ml of fermentation medium (pH 6.2 no CaCO<sub>3</sub>) was inoculated 12 hrs before inoculation of the fermentor.

### 3) Fermentation conditions

Fermentations were performed under the following conditions : fermentor SETRIC (2 l), 35°C, stirring at 125 rpm, bubbling with N<sub>2</sub>/CO<sub>2</sub> (85/15) 30 min before inoculation and throughout the fermentation, pH regulated by addition of 4N NaOH.

### 4) Resting cells preparation

A sample of 180 to 230 ml was centrifugated at 20,000 g (SORVALL RC-5), cells were washed in phosphate buffer (0.2 M KH<sub>2</sub>PO<sub>4</sub>, 0.2 M K<sub>2</sub>HPO<sub>4</sub>, 5 g/l NaCl, 1 ml/l of 0.2% resazurine solution, 2% (vol/vol) Na<sub>2</sub>S, pH of fermentation broth at the time of sampling) and resuspended in 40 ml of phosphate buffer. Glucose (6 g/l) or butyrate (2 g/l) was added aseptically. Resting cells were incubated under anaerobic atmosphere (75% N<sub>2</sub>, 15% CO<sub>2</sub>, 10% H<sub>2</sub>) at 35°C in an anaerobic chamber (FORMA SCIENTIFIC 1024).

### 5) Analytical method

Growth was monitored by following absorbance at 600 nm. The cell mass (dry weight) was extrapolated from a standard curve OD vs dry weight. Acetate, acetone, butanol, butyrate, and ethanol concentrations were measured by gas liquid chromatography using a VARIAN 370 gas chromatograph equipped with a flame ionization detector and a 2.5 m x 1/8 mm column packed with 100/200 mesh chromosorb / 25% NPGA, 2% H<sub>3</sub>PO<sub>4</sub>. Glucose concentrations were determined using a D glucose/ D fructose enzymatic determination kit, UV method (Boehringer-Mannheim).

## 6) Carbon recovery calculations

The carbon recovery  $\eta$  was determined as follows according to the respective stoichiometries.

$$\text{glucose assay : } \eta = 100(a+b/2+c)/(y_0-y_t)$$

$$\text{butyrate assay : } \eta = 100 \times 6(b/2+c-(a_0-a_t))$$

assuming that all butyrate consumed is reduced in butanol

with a = butyrate concentration

b = acetate concentration

c = butanol concentration

$y_0-y_t$  = initial (glucose) - final (glucose)

$a_0-a_t$  = initial (butyrate) - final (butyrate)

## RESULTS

### 1) Attainment of the different cellular forms

Phosphate buffer was chosen over citrate buffer (0.1 M citrate, 0.1 M trisodium citrate, 5 g/l NaCl, 1 ml/l of 0.2% resazurine solution, 2% (vol/vol) Na<sub>2</sub>S) since resting cells were observed to give higher yields of butanol in phosphate buffer (data not shown) in good accordance with the observation that phosphate concentration of the growth medium affects butanol production by *C. acetobutylicum* (Bahl et al., 1982 ; Bryant and Blaschek, 1988).

Five cellular forms were tested for their ability to incorporate glucose and butyrate, and their ability to produce solvents. These five forms were issued from two different fermentations and they were maintained at the pH of the fermentation medium at the time of sampling.

The first fermentation starts at pH 6.2 and is maintained at this level for 11 hrs after inoculation where the first sample is taken and processed to produce resting cells. These cells were blocked in acidogenesis and were showing a high motility as well as the highest growth rate.

The pH regulator is then set at pH 5.1 (pH where solvent production by strain B6 starts (Zygmunt, 1983)). The second sample (vegetative cells type I) is taken and processed when pH 5.1 was reached. At this stage, cells were poorly motile, but no clostridial form was observed.

The second fermentation starts in a medium at pH 6.2 but the pH regulator is set immediately at pH 5.1. The third sample (vegetative cells type II) is taken 11 hrs after inoculation. Cells were then poorly motile to not motile but no clostridial form was observed.

The fourth sample is taken when 70 to 80% of the cells are clostridial forms. No cell was motile at this stage. The fifth sample was taken 32 hrs after inoculation and all cells were sporulating.

Analysis of the fermentation broth (Table I) indicate that only the clostridial forms are in solventogenesis and that vegetative cells type I have a slight solventogenic activity (0.81 g/l of butanol).

## 2) Glucose and butyrate utilization by the resting cells

The utilization of glucose and butyrate by the five cellular forms were followed for 8 hrs. Data are shown in Table II and Figure 1-8.

Vegetative cells pH 6.2 are clearly in acidogenesis : all glucose is consumed in 3 hrs, acids are produced in large quantities while little butanol is produced (12% of carbon produced, CO<sub>2</sub> excluded). The carbon recovery indicates furthermore that all glucose consumed is utilized for product formation.

Vegetative cells type I - pH 5.1 show a slower kinetics of glucose utilization and utilize only 80% of available glucose, larger quantities of butanol are produced (48% of carbon produced, CO<sub>2</sub> excluded) and acid production follows a slower kinetics. Moreover, the carbon recovery indicates once again that all glucose consumed is used for product formation.

Vegetative cells type II - pH 5.1 show a very quick kinetics of glucose consumption during the first hour that reverses afterwards due to cell lysis. Butanol represents 20% of the carbon produced (CO<sub>2</sub> excluded) and carbon recovery determinations on glucose is accumulated in the cell as granulose (80%). Butyrate on the other hand is consumed at a quick pace but only 55% of the available butyrate is consumed. Interestingly, carbon recovery determinations on the butyrate assay suggest that granulose synthesized during fermentation in the fermentor is utilized.

Clostridial forms show on the contrary a glucose utilization kinetics 9 fold slower than vegetative cells type II and the carbon recovery is always lower than 20%. Very little to no acid is produced and butanol represents 58% of all carbon produced (CO<sub>2</sub> excluded). Butyrate is incorporated 3 fold slower but carbon recovery on butyrate exceeds 150% suggesting that granulose is, again, utilized.

The fifth sample (spores) was not showing, as expected, any kind of enzymatic activity.

## **DISCUSSION**

The relationships between physiological changes and solvent production ability of the five cellular forms studied are summarized in Table III.

Solventogenesis appears to be induced very early on since vegetative cells at pH 6.2 are able to produce low concentrations of butanol suggesting that the enzymes necessary for solvent production (butyraldehyde dehydrogenase, butanol dehydrogenase) are already induced in the fermentor. This hypothesis is furthermore supported by the findings of Ballongue et al. (1989) who reported that under their assay conditions the enzymes necessary for solvent production are present both under conditions for acid production (pH 6.8) and under conditions for solvent production (pH 4.8), but are correlated with solvent production only in the latter case. The inability of vegetative cells pH 6.2 and vegetative cells type I to reduce butyrate into butanol during the butyrate assay is likely due to the lack of an energy source and as a result of a low pool of reduced NAD. Moreover, at pH 6.2, very little butyrate will be protonated and be able to get into the cell to be reduced. The difference in enzymatic activities between those two cellular forms could thus be mainly ascribed to the difference in the level of undissociated butyric acid present in the medium and the difference in internal pH, in good

accordance with the observation that cells maintained at neutral pH can initiate solvent production (Holt et al., 1984).

Vegetative cells type II represent a more advanced stage more oriented towards the accumulation of granulose, that leads ultimately to morphological changes and clostridial shape formation. Both vegetative cells type II and clostridial forms are able to utilize exogenous butyrate by the requisition of stored energy (granulose) to replenish the pool of reducing equivalents. Clostridial forms show a rather low solventogenic activity that is maximal during their appearance and not when they represent the majority of the cell population.

As a result, pH regulation and resting cells formation appears to be a successful method to characterize the main solventogenic cellular form under a given set of fermentation conditions. Moreover, the kinetics of granulose accumulation (cf. carbon recovery determinations and Table III) seems to be a good indicator of the solventogenic ability of a particular cellular form. It would be of interest to quantify the amount of granulose accumulated by the cells along the fermentation as well as to follow the kinetics of its accumulation. On a practical basis, such a characterization could be of importance to develop more efficient continuous processes.

#### ACKNOWLEDGMENTS

We thank Dennis Bryant for reviewing the manuscript of this paper.

#### LITERATURE CITED

- 1) **BAHL H., W. ANDERSCH, and G. GOTTSCHALK.** 1982. Continuous production of acetone and butanol by *Clostridium acetobutylicum* in a two stage phosphate limited chemostat. *Appl. Microbiol. Biotechnol.* 15:201-203
- 2) **BAHL H., W. ANDERSCH, K. BRAUN, and G. GOTTSCHALK.** 1982. Effect of pH and butyrate concentration on the production of acetone and butanol by *Clostridium acetobutylicum* in continuous culture. *Eur. J. Appl. Microbiol. Biotechnol.* 14:17-20
- 3) **BALLONGUE J., R. JANATI-IDRISSI, H. PETITDEMANGE, and R. GAY.** 1989. Correlation between solvent production and level of solventogenic enzymes in *Clostridium acetobutylicum*. *J. Appl. Bact.* 67:611-617
- 4) **BRYANT D.L. and H.P. BLASCHEK.** 1988. Buffering as a means for increasing growth and butanol production by *Clostridium acetobutylicum*. *J. Ind. Microbiol.* 3:49-55
- 5) **COSTA J.M. and A.R. MOREIRA.** 1983. Growth inhibition kinetics for the acetone-butanol fermentation. In "Foundations of Biochemical Engineering", H.N. Blanch, E.T. Papoutsakis, and G. Stephanopoulos eds., p. 501-502. American Chemical Society, Washington D.C.
- 6) **GOTTSCHAL J.C. and J.G. MORRIS.** 1981. The induction of acetone and butanol production in cultures of *Clostridium acetobutylicum* by elevated concentrations of acetate and butyrate. *FEMS Microbiol. Letters* 12:385-389

- 7) HERRERO A.A., R.F. GOMEZ, B. SNEDECOR, C.J. TOLMAN, and M.F. ROBERTS. 1985. Growth inhibition of *Clostridium thermocellum* by carboxylic acids : a mechanism based on uncoupling by weak acids. Appl. Microbiol. Biotechnol. 22:53-62
- 8) HOLT R.A., G.M. STEPHENS, and J.G. MORRIS. 1984. Production of solvents by *Costridium acetobutylicum* cultures maintained at neutral pH. Appl. Env. Microbiol. 48:1166-1170
- 9) JONES D.T. and D.R. WOODS. 1986. Acetone butanol fermentation revisited. Microbiol. Rev. 50:484-524
- 10) LONG S., D.T. JONES, and D.R. WOODS. 1984. Initiation of solvent production, clostridial stage and endospore formation in *Clostridium acetobutylicum* P262. Appl. Microbiol. Biotechnol. 20:256-261
- 11) NISHIO N., H. BIEBL, and M. MEINERS. 1983. Effect of pH on the production of acetone and butanol by *Clostridium acetobutylicum* in a minimum medium. J. Ferm. Technol. 61:101-104
- 12) ROSS D. 1961. The acetone butanol fermentation. Prog. Ind. Microbiol. 3:73-85
- 13) TERRACIANO J.S. and E.R. KASHKET. 1986. Intracellular conditions required for initiation of solvent production by *Clostridium acetobutylicum*. Appl. Env. Microbiol. 52:86-91
- 14) WANG G. and D.I.C. WANG. 1984. Elucidation of growth inhibition and acetic acid production by *Clostridium thermoaceticum*. Appl. Env. Microbiol. 47:294-298
- 15) ZYGMUNT M. 1983. Thèse de 3<sup>ième</sup> cycle, Université des Sciences et Techniques de Lille

**Analysis of the fermentation broth  
at the time of sampling (g/l)**

| Cell type                   | Cell mass<br>(dry weight) | Glucose<br>consumed | Butanol | Acetone | Butyrate | Acetate |
|-----------------------------|---------------------------|---------------------|---------|---------|----------|---------|
| vegetative<br>cells pH 6.2  | 0.34                      | 7.7                 | 0.17    | -       | 0.60     | 2.00    |
| vegetative<br>cells type I  | 0.62                      | 9.0                 | 0.81    | 0.05    | 1.73     | 1.76    |
| vegetative<br>cells type II | 0.50                      | 8.5                 | traces  | -       | 1.71     | 1.46    |
| clostridial<br>forms        | 2.26                      | 13.7                | 7.32    | 0.86    | 1.36     | 0.72    |
| spores                      | 0.92                      | 31.1                | 10.32   | 2.20    | 1.02     | 0.81    |

Tableau 1

**Initial kinetics of utilization and production  
of main compounds : glucose assay  
(mM/ (g cell x hr))**

| Cell type                   | Glucose | Butyrate | Butanol | Acetate |
|-----------------------------|---------|----------|---------|---------|
| vegetative<br>cells pH 6.2  | -4.55   | 2.19     | 0.35    | 3.21    |
| vegetative<br>cells type I  | -3.57   | 0.55     | 0.99    | 0.54    |
| vegetative<br>cells type II | -5.06   | 0.26     | 0.21    | 0.69    |
| clostridial<br>forms        | -0.57   | 0.02     | 0.11    | 0.06    |
| spores                      | -       | 0.00     | 0.00    | 0.00    |

Tableau 2

**Initial kinetics of utilization and production  
of main compounds : butyrate assay  
(mM/ (g cell x hr)**

| Cell type                   | Glucose | Butyrate | Butanol | Acetate |
|-----------------------------|---------|----------|---------|---------|
| vegetative<br>cells pH 6.2  |         | -        | 0.00    | 0.00    |
| vegetative<br>cells type I  |         | -        | 0.00    | 0.00    |
| vegetative<br>cells type II |         | -2.24    | 5.25    | 3.69    |
| clostridial<br>forms        |         | -0.68    | 1.17    | 0.90    |
| spores                      |         | 0.00     | 0.00    | 0.00    |

Tableau 2

**Physiological Evolution of the different cellular  
forms of *Clostridium acetobutylicum***

| Cell type                     | Vegetative<br>pH 6.2 | Vegetative<br>type I | Vegetative<br>type II | Clostridial<br>forms | Spores |
|-------------------------------|----------------------|----------------------|-----------------------|----------------------|--------|
| Motility                      | +++                  | +                    | +/-                   | -                    | -      |
| Total activity<br>on glucose  | +++                  | ++                   | +++                   | +/-                  | -      |
| Acidogenesis                  | +++                  | ++                   | +                     | +/-                  | -      |
| Solventogenesis               | +/-                  | +                    | ++                    | +++                  | -      |
| Granulose<br>synthesis        | -                    | -                    | ++                    | +                    | -      |
| Total activity<br>on butyrate | -                    | -                    | +++                   | +                    | -      |
| Acidogenesis                  | -                    | -                    | ++                    | +                    | -      |
| Solventogenesis               | -                    | -                    | ++                    | +                    | -      |

Tableau 3

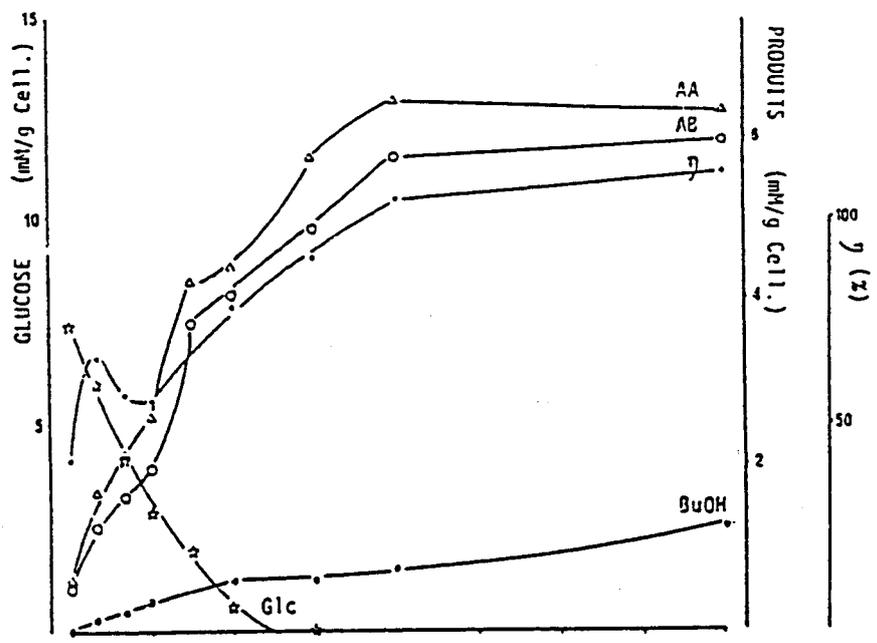


Figure 1 : Glucose assay. Vegetative Cells pH 6.2

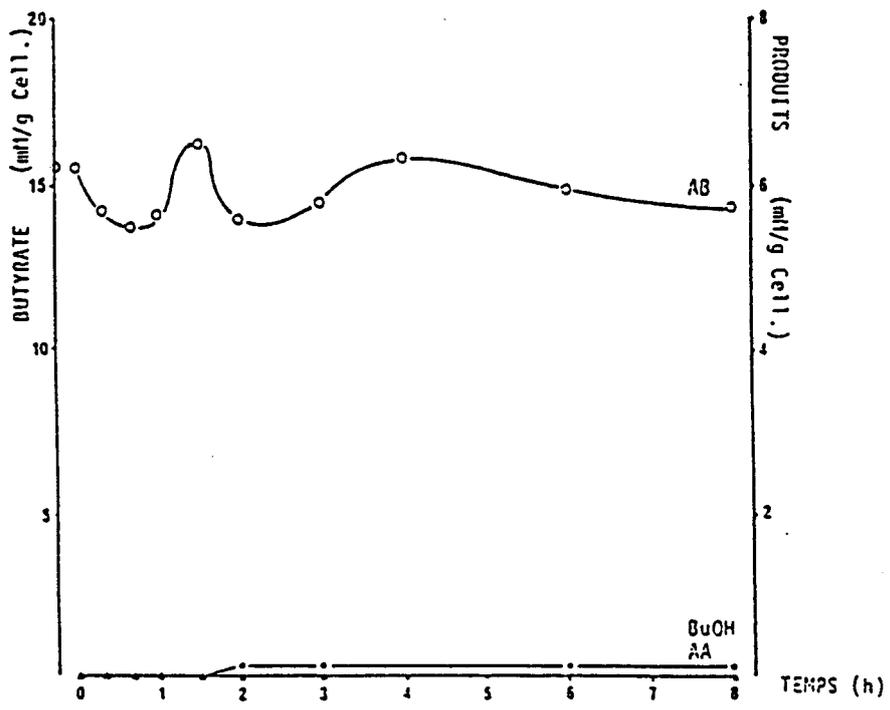


Figure 2: Butyrate assay. Vegetative Cells pH 6.2

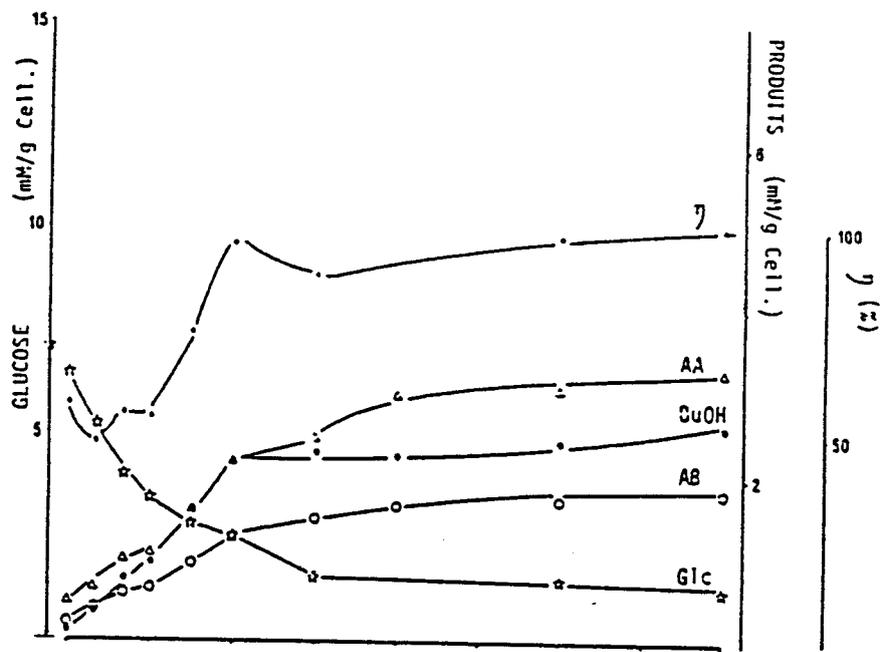


Figure 3 : Glucose assay. Vegetative Cells type I pH 5.1

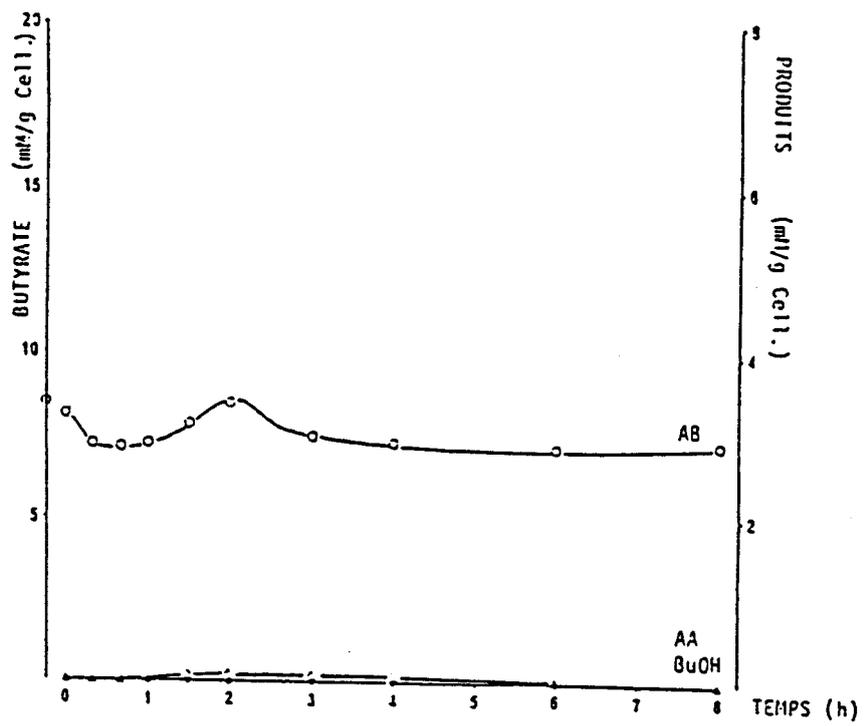


Figure 4: Butyrate assay. Vegetative Cells type I pH 5.1

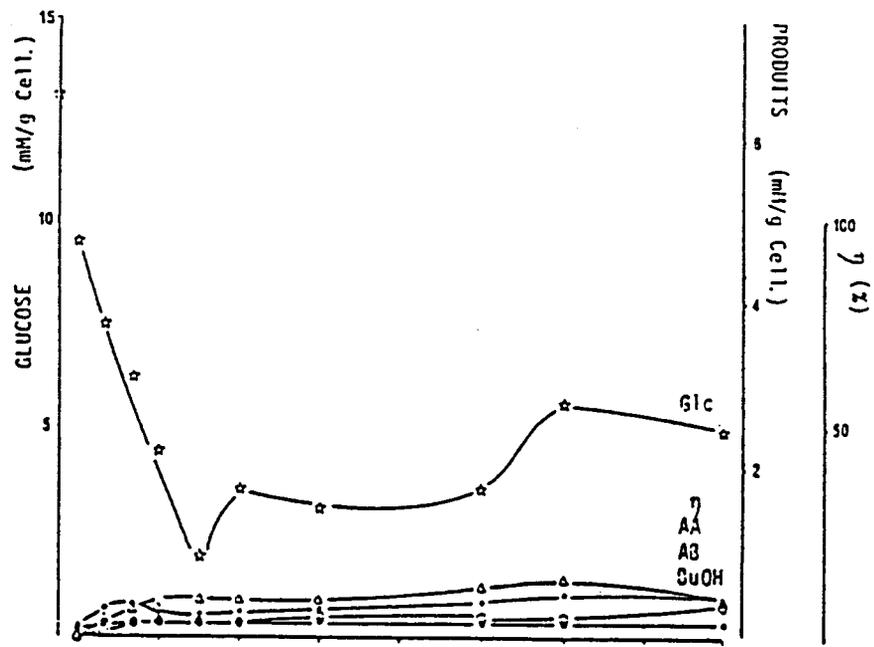


Figure 5 : Glucose assay. Vegetative Cells type II pH 5.1

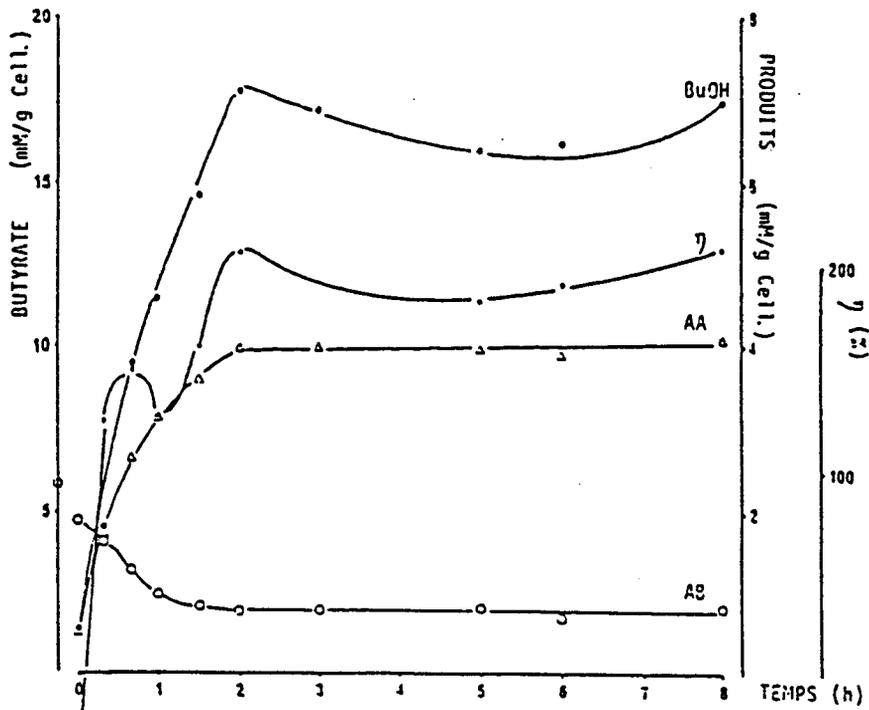


Figure 6: Butyrate assay. Vegetative Cells type II pH 5.1

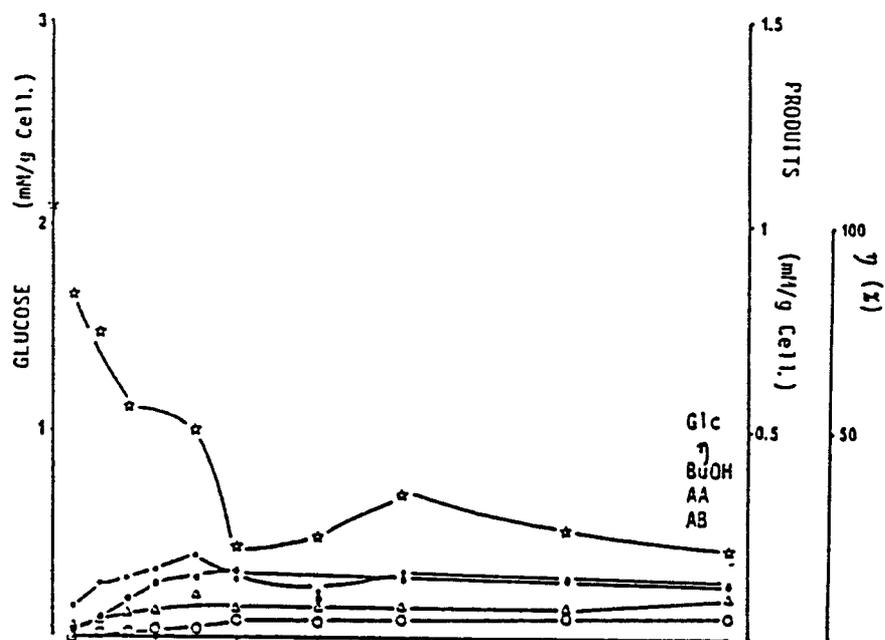


Figure 7 : Glucose assay. Clostridial forms pH 5.1

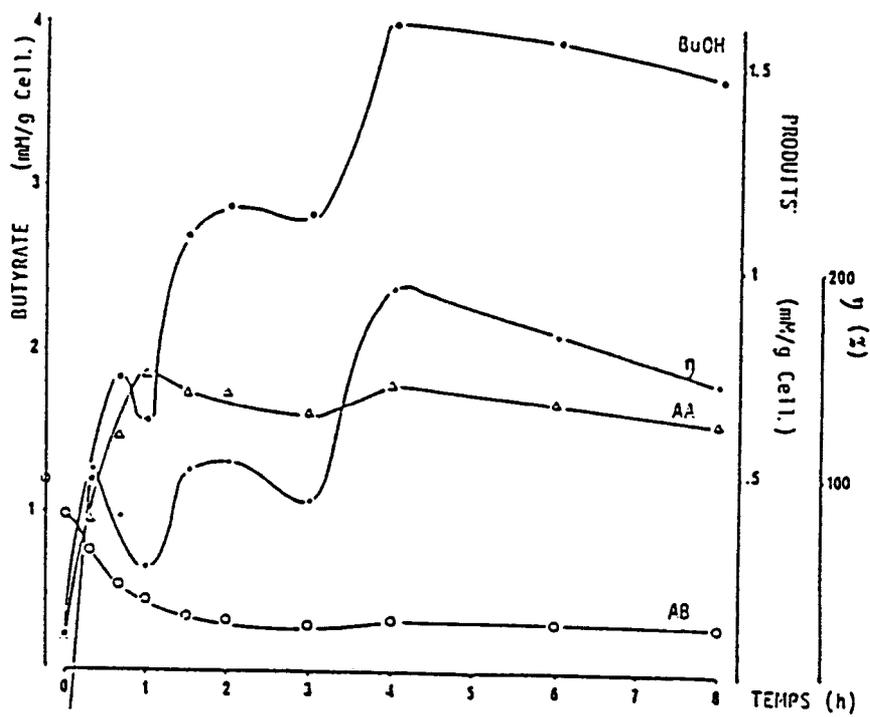


Figure 8: Butyrate assay. Clostridial forms pH 5.1

### III) CONTROLE DE L'EFFICACITE DU LEVURAGE PAR LA MISE EN OEUVRE DE SOUCHES DE LEVURES MARQUEES

#### A) LA PRATIQUE DU LEVURAGE. BUT DE L'ETUDE

La pratique du levurage s'impose comme une étape nécessaire étant donné les impératifs économiques et de maintien de standard de la qualité actuels (Moulin et Valade, 1983) ; mais la fermentation en souche pure est un objectif qui rencontre certains problèmes pratiques (Barre and Vezinhet, 1984).

L'efficacité du levurage, facilité par la mise en oeuvre de levures sèches actives (Bidan and Maugenet, 1981), doit donc être contrôlée à l'aide de marqueurs qui permettent de distinguer si les souches isolées dans le moût de fermentation sont identiques à celles apportées par le levurage ou si elles correspondent au contraire à des souches de levures indigènes (Loiseau et coll., 1985). Le marquage des souches de levures oenologiques au niveau des mitochondries (résistance au chloramphénicol, à l'oligomycine) (Vezinhet, 1985 ; Vezinhet et Loiseau, 1985) permet d'obtenir un marquage sélectif qui n'altère pas le potentiel oenologique de la souche et ne confère aucun désavantage sélectif à la souche mutée par rapport à la souche sauvage, le rôle des mitochondries dans la fermentation étant négligeable. La technique de l'utilisation de souches antibiorésistantes en fermentation alcoolique n'avait été testée qu'en laboratoire sur un mode unique de vinification. Il était donc indispensable de vérifier le comportement de ces souches en utilisant différents modes de vinification (vendange foulée, vendange foulée en cuve rotative, vendange chauffée, macération carbonique). D'autre part, il était indispensable de tester ces souches en conditions réelles d'utilisation (cave expérimentale et cuves de 50 hl).

## **B) MATERIELS ET METHODES**

### **1) SOUCHE DE LEVURE**

La souche utilisée pour le levurage était la souche K1 de *S. cerevisiae* (killer, résistante au chloramphénicol et à l'oligomycine)

### **2) MILIEUX DE CULTURE**

Les milieux de culture employés étaient les milieux YEPD (10 g/l yeast extract, 20 g/l peptone, 20 g/l glucose), ou N (10 g/l yeast extract, 10 g/l peptone, 20 g/l glycérol, 100 ml de tampon phosphate de sodium 0,5 M, pH 6,5). Le chloramphénicol est ajouté à raison de 4 mg/ml et l'oligomycine à raison de 3 µg/ml.

### **3) PROTOCOLE EXPERIMENTAL**

Le levain a été préparé en ensemençant 10 ml de YEPD à l'aide d'une colonie de la souche K1 de *S. cerevisiae*. Après une nuit d'incubation à 37°C, la culture a été transférée successivement dans 100 ml de YEPD puis 8 l de jus de raisin qui ont ensuite été utilisés pour ensemencher une cuve de 50 hl. Le levurage a été effectué dès l'encuvage avec des ensemençements comparables en population à ceux pratiqués avec des levures sèches actives. Les contrôles microbiologiques ont été effectués du début à la fin de la fermentation. En outre, la flore indigène présente dans la cuve ainsi que dans le levain a été contrôlée avant ensemençement. Les différentes numérations de la population de levure ont été effectués par comptage sur des boîtes YEPD (détermination de la population totale) et par *replica plating* sur des boîtes N supplémentées de chloramphénicol et d'oligomycine (détermination du nombre de levures marquées).

## **C) RESULTATS**

L'évolution de la température et de la densité des jus de fermentation a été reportée Figures 3 et 4. Le pourcentage

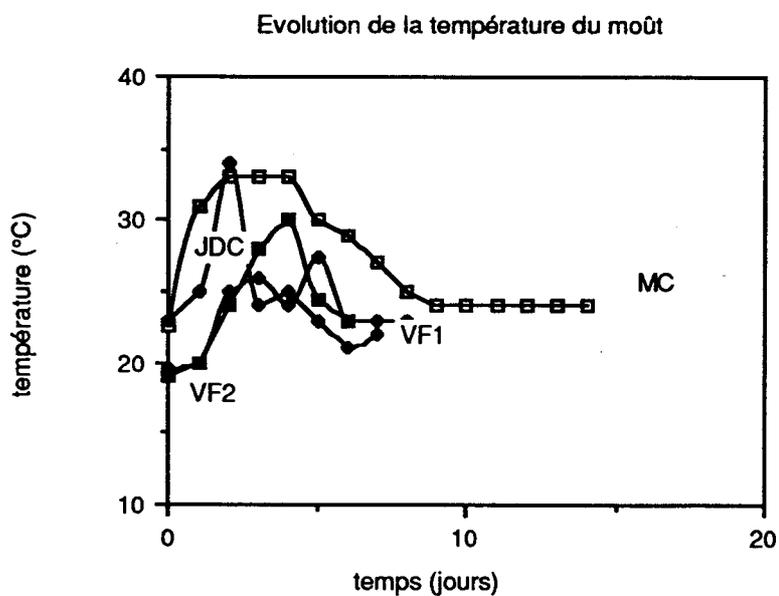


Figure 3

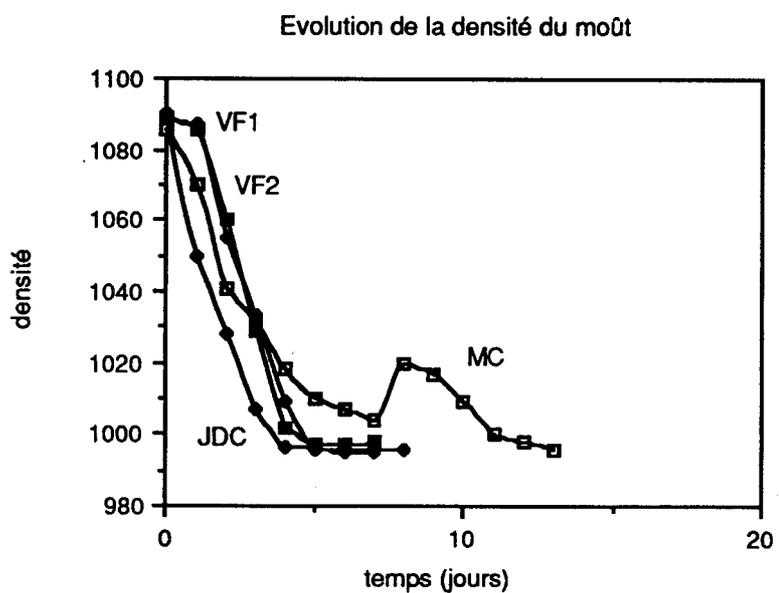


Figure 4

$z = 1$  - levures marquées/ levures indigènes est choisi comme indicateur pour apprécier l'implantation de la souche sélectionnée dans le jus de fermentation. L'expérience a été conduite pendant toute la durée de la fermentation, à partir d'une vendange foulée en cuve rotative (VF1), d'une vendange foulée en cuve statique (VF2), d'une vendange chauffée (JDC), et d'une macération carbonique (MC). La vendange utilisée était composée de 60% de Carignan et de 40% de Grenache.

Les valeurs de  $z$  pour les 4 essais VF1, VF2, JDC, et MC ont été reportées en fonction du temps dans les Figures 5, 6, 7, et 8, respectivement. Il apparaît que les essais VF1, VF2, et JDC ont un comportement relativement semblable durant toute l'expérience : le taux  $\mu$  de contamination par des levures indigènes est maximal au début de l'expérience puis décroît de façon exponentielle dès les premiers jours pour se stabiliser à une valeur faible. Les taux  $\mu$  moyens sont donc peu élevés :  $\mu_{VF1} = 0,3 \%$  ;  $\mu_{VF2} = 0,8 \%$ ,  $\mu_{JDC} = 1,8 \%$ . L'opération de pressurage survenue au 8<sup>ième</sup> jour de fermentation pour l'essai VF1 et au 7<sup>ième</sup> jour pour l'essai VF2 est restée sans conséquence sur l'évolution de l'indicateur  $z$ .

Par contre, la fermentation en macération carbonique n'est pas réalisée entièrement par la souche sélectionnée et utilisée pour réaliser le levurage. On peut en effet noter Figure 8 la présence de deux courbes de croissance bien distinctes. La souche sélectionnée s'est bien implantée dans le jus de fermentation pendant les 3 premiers jours mais la population totale s'effondre du 4<sup>ième</sup> au 6<sup>ième</sup> jour, ce qui résulte pratiquement en un arrêt de la fermentation. En fait, la fermentation ne reprend qu'après l'opération de pressurage, à partir d'une souche de levure indigène. Le taux  $\mu$  de contamination par des levures indigènes est égal à 2,5% pour les 5 premiers jours de fermentation et à 69,3% pour les 9 derniers.

Un échantillon de souches indigènes isolées au cours de la fermentation en macération carbonique a été stocké sur

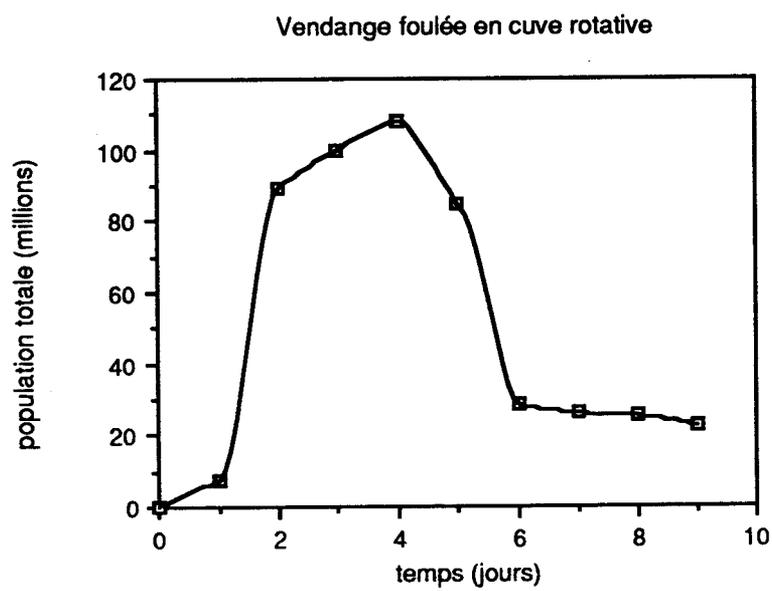
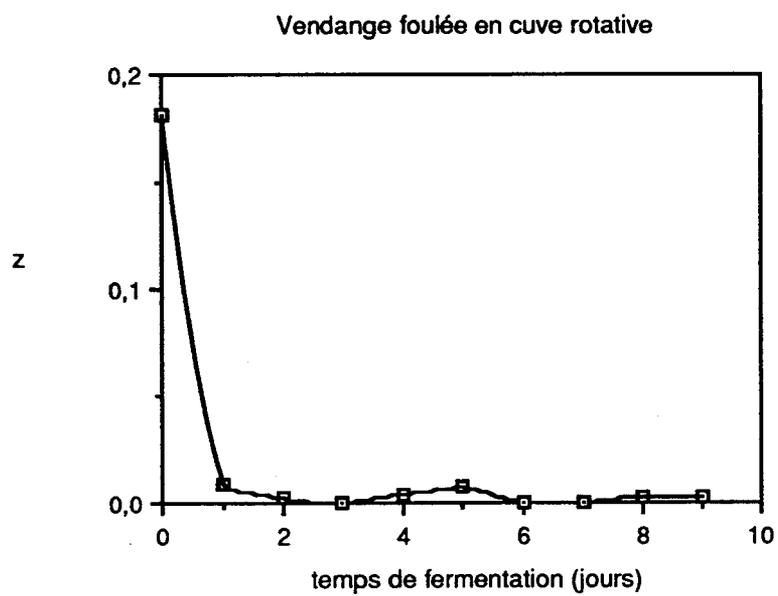


Figure 5

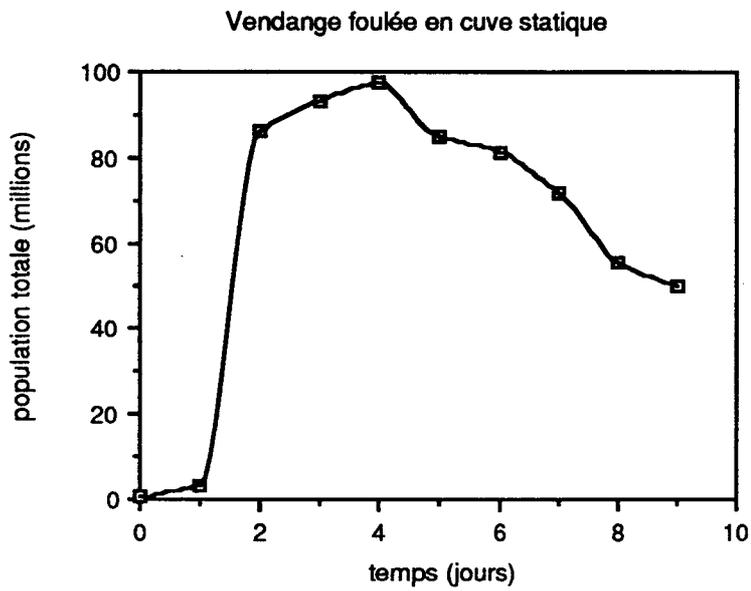
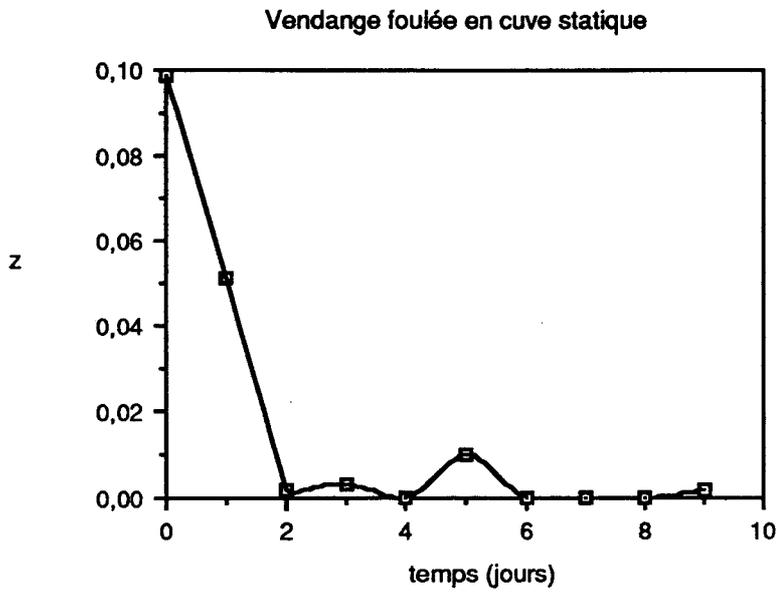


Figure 6

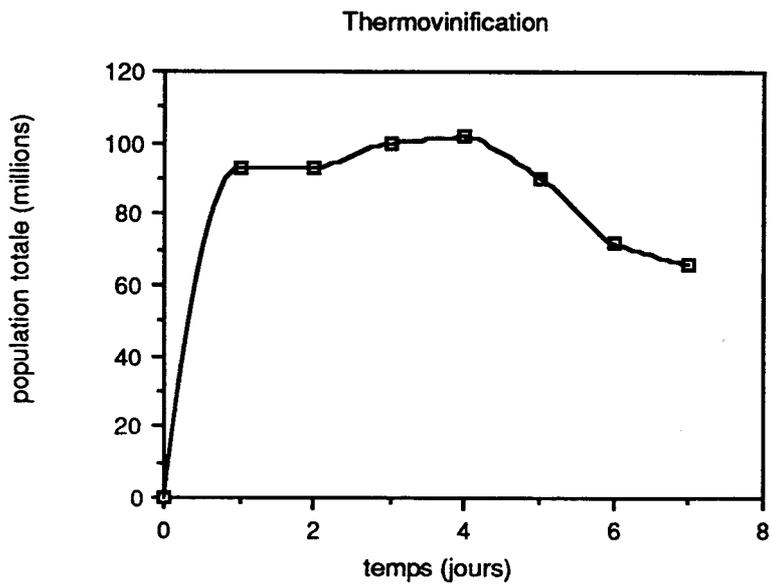
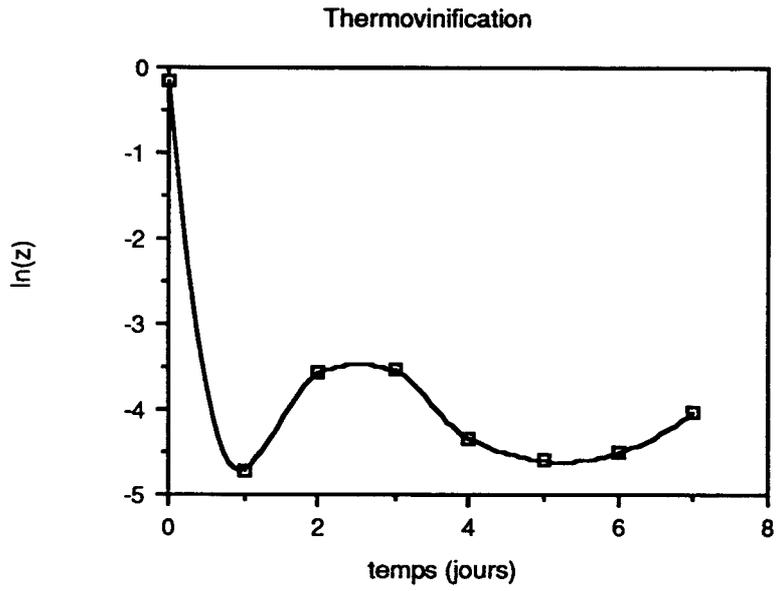


Figure 7

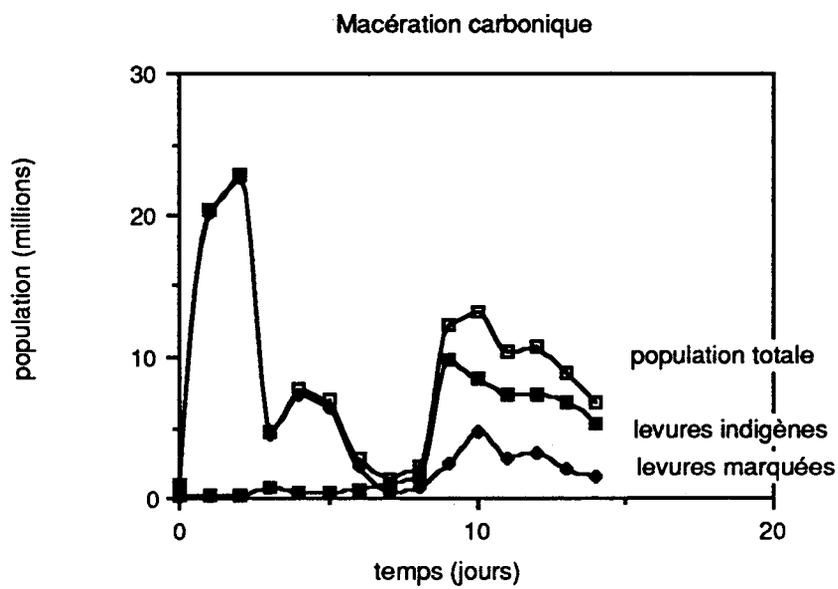
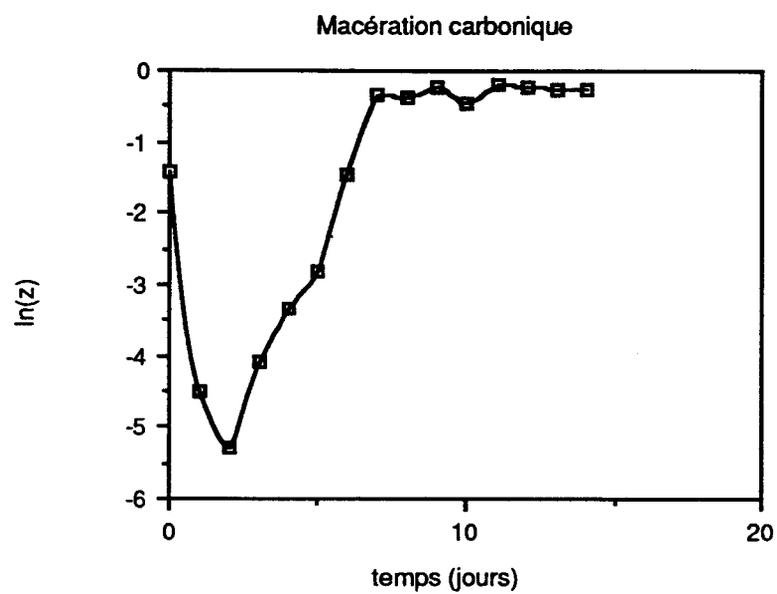


Figure 8

des géloses inclinées de YEPD. Le genre de ces diverses souches a été déterminé à partir de critères morphologiques (taille et forme des cellules, forme des spores, mode de bourgeonnement) et de critères biochimiques (fermentation du glucose, assimilation des nitrates, assimilation du galactose, résistance à la cyclohexamide). Avant l'opération de levurage ( $t_0$ ), 11 souches ont été analysées parmi lesquelles 8 Kloeckera, 2 Pichia, et 1 Metschnikowia ont été dénombrées. A l'issue du premier jour de fermentation, 1 Saccharomyces, 1 Kloeckera, et 1 Metschnikowia ont été isolées, tandis qu'à partir du deuxième jour de fermentation jusqu'au 14<sup>ième</sup>, seules des levures du genre Saccharomyces ont été isolées (1 à  $t_2$ , 4 à  $t_3$ , 22 à  $t_5$ , 19 à  $t_6$ , 30 à  $t_7$ , 36 à  $t_{11}$ ). Toutes ces souches de Saccharomyces ainsi isolées (131) sont capables d'assimiler le galactose.

#### **D) DISCUSSION**

Cet essai réalisé sur plusieurs mois a permis de mettre en évidence l'utilisation du levurage ainsi que la possibilité pratique de réaliser une fermentation en souche pure de levure. En outre, les résultats obtenus pour les essais qui concernent les vendanges foulées et la thermovinification confirme en grandeur réelle que le marquage mitochondrial des levures est conservé.

La technologie de la macération carbonique est cependant très différente de ces 3 autres types de vinification. En effet, il apparaît que la souche sélectionnée s'implante bien dans le milieu au début de la fermentation, mais que celle-ci est terminée par une souche indigène gal<sup>+</sup> de S. cerevisiae.

La première observation importante est que la population totale chute très rapidement et très bas ( $2 \cdot 10^6$  cellules/ml) au 5<sup>ième</sup> jour de fermentation alors que le jus de goutte n'est pas entièrement fermenté (densité 1010), ce qui est synonyme d'un possible arrêt de fermentation, toujours préjudiciable à la qualité du vin dû à la prolifération des bactéries

lactiques. L'augmentation progressive du volume du jus de goutte permet d'écarter l'hypothèse d'une carence nutritive pour rendre compte de cette observation. On peut en outre remarquer que la population totale de levures a toujours été très faible :  $25 \cdot 10^6$  cellules/ ml au maximum contre  $10^8$  cellules/ ml dans les autres types de vinification. Une hypothèse plausible est donc que les conditions d'anaérobiose sont directement responsables de ces différences entre les diverses techniques de vinification. En effet, d'après les travaux de Cochin (1880), Lafourcade (1954), et Cantarelli (1956) (cités par Ribereau-Gayon et coll., 1975) la multiplication des levures est impossible dans des conditions d'anaérobiose parfaite. Ceci s'explique par l'observation que la biosynthèse des stérols des levures (ergostérol, zymostérol, ...) à partir du squalène ainsi que la désaturation des acides gras à longue chaîne nécessitent de l'oxygène (accepteur final d'électrons). Le rôle physiologique des acides gras et des stérols est triple : ce sont des constituants cellulaires, des supports d'enzymes et des facteurs de perméabilité cellulaire. Néanmoins, l'ergostérol, le cholestérol, quelques stéroïdes et acides gras à longue chaîne dont certains se trouvent être des constituants de la pruine des raisins, permettant de pallier ces carences dues à une insuffisance en oxygène. De plus, on trouve dans la pellicule des raisins des activateurs spécifiques des levures en anaérobiose. Il est cependant possible que les concentrations de ces molécules ne sont pas suffisamment élevées dans le jus de goutte de la macération carbonique pour permettre une croissance complète. L'hypothèse que le manque d'oxygène ou d'un accepteur final d'électrons est la cause de la chute de la population totale des levures est par ailleurs soutenue par l'observation que la population totale ne parvient à s'accroître qu'après l'opération de pressurage. Outre l'apport massif de sucres dans le milieu, cette dernière opération permet d'oxygéner le moût et de dissoudre les substances contenues dans la baie,

la pellicule, et la pruine par augmentation des surfaces de contact de macération ; ce qui aurait pour effet de rendre le milieu à nouveau propice à la multiplication des levures.

Le deuxième problème important concerne l'implantation dans le milieu de la souche sélectionnée et utilisée dans le levain. Il est clair que cette dernière, bien qu'adaptée aux autres types de vinification, ne parvient pas à s'établir durablement dans le moût en macération carbonique, et que la courbe de population observée pour cette fermentation est typiquement composée de la juxtaposition de deux courbes de croissance. Ce phénomène peut être expliqué par l'observation que le jus de goutte produit chaque jour apporte massivement dans le jus de fermentation des levures fermentaires non marquées (le levurage du jus de goutte a été effectué par le bas de la cuve pour des raisons techniques). Ces levures fermentaires indigènes ont bénéficié de tous les éléments nutritifs présents dans le jus de goutte ainsi que des activateurs spécifiques qui se trouvent dans la pellicule ou la pruine des raisins. Il est en outre possible que les conditions d'anaérobiose sont moins rigoureuses dans la moitié supérieure de la cuve (cuve non close). Une explication plausible est donc que ces souches de levures indigènes étaient moins carencées (grâce à la proximité de sources d'activateurs, d'oxygène, ...) que les levures marquées. Ceci peut suffire pour leur conférer un avantage compétitif important, étant donné que les levures possèdent des enzymes respiratoires à un stade élevé d'oxydation et qui constituent une réserve d'oxygène suffisante pour assurer en anaérobiose quelques générations avant épuisement total (Cantarelli, cité par Ribéreau-Gayon, 1975). Il est également possible que la souche de *S. cerevisiae* isolée est mieux adaptée aux conditions de la macération carbonique que ne l'est la souche sélectionnée. Des expériences complémentaires sont nécessaires pour établir ce dernier point.

E) ARTICLE 3

G. LOISEAU, F. VEZINHET, M. VALADE, A. VERTES, C. CUINIER, D. DELTEIL. 1987. CONTROLE DE L'EFFICACITE DU LEVURAGE PAR LA MISE EN OEUVRE DE SOUCHES DE LEVURES MARQUEES. Revue Française d'Oenologie 27:29-36

# CONTROLE DE L'EFFICACITE DU LEVURAGE PAR LA MISE EN OEUVRE DE SOUCHES DE LEVURES OENOLOGIQUES MARQUEES

G. LOISEAU (1), F. VEZINHET (2), M. VALADE (1), A. VERTES (3), C. CUINIER (4) et D. DELTEIL (5)

## RESUME

*Ce travail montre au travers de différentes expérimentations réalisées en Champagne, Beaujolais, Touraine et Languedoc-Roussillon, comment l'utilisation de souches œnologiques marquées permet d'étudier différents aspects du levurage : la concurrence entre souche introduite et flore indigène, le positionnement du levurage par rapport au sulfitage, le levurage dans différents modes de vinification, le levurage avec une association de souches.*

*Dans tous les cas, la composition des populations levuriennes a pu être suivie tout au long de la fermentation avec une précision de quelques pour-cent.*

**Mots clés :** Levures - Levurage - Fermentation - Marquage génétique.

## SUMMARY

*Several trials were performed in Champagne, Beaujolais, Touraine and Languedoc-Roussillon, with enological yeast strains carrying genetic markers. These experimentations allowed the study of different aspects of inoculation : competition between the inoculated strain and the wild population, timing of inoculation in relation with addition of SO<sub>2</sub>, inoculation related with wine making processes, inoculation with an association of strains.*

*In all cases, yeast population composition could be followed throughout the fermentation with a few percent precision.*

**Key Words :** Yeast - Inoculation - Fermentation - Genetic markers.

(1) Comité Interprofessionnel du Vin de Champagne (C.I.V.C.) - 5, rue Henri Martin, BP 135 51204 Epernay Cedex.

(2) I.N.R.A. - Institut des Produits de la Vigne - Laboratoire de Microbiologie et Technologie des Fermentations - Place Viala 34060 Montpellier Cedex.

(3) I.N.R.A. - Institut des Produits de la Vigne - Station expérimentale de Foch Rouge 11430 Gruissan

(4) ITV - CETEX TOURS - 12, rue E. Pallu, 37033 Tours Cedex

(5) Institut Coopératif du Vin - Hameau de la Jasse - Maunin 34970 Lattes

L'intérêt du levurage dans de très nombreuses situations œnologiques n'est plus à démontrer et on constate dans les faits un développement régulier de l'emploi des levures sèches actives (KRAUS et coll., 1984).

La diversité des souches proposées sur le marché permet à l'œnologue de choisir celle qui est le mieux adaptée au type de vin qu'il souhaite élaborer (VALADE et MOULIN, 1982 ; CUINIER et coll., 1985).

Toutefois, s'il veut faire une optimisation des conditions de levurage, il est gêné par le manque d'un moyen de contrôle simple et fiable de l'implantation de la souche ensemencée.

Le marquage génétique de souches de levures œnologiques, mis au point par l'I.N.R.A. à l'Institut des Produits de la Vigne, peut apporter une réponse à ce problème (VEZINHET et LACROIX, 1984 ; VEZINHET, 1985).

Dans le présent article, nous décrivons un certain nombre d'expérimentations réalisées avec des levures œnologiques identifiables et qui avaient pour objectif d'étudier différents aspects du levurage : la concurrence entre souche introduite et flore indigène, le positionnement du levurage par rapport au sulfitage, le levurage dans différents modes de vinification, la possibilité de réaliser une fermentation avec une association de souches.

## EXPERIMENTATIONS ET RESULTATS

Avant de réaliser des expérimentations avec une souche œnologique marquée, nous avons vérifié que les LSA (levures sèches actives), produites à partir de cette souche, donnaient les mêmes résultats que les LSA produites à partir de la souche parentale d'origine.

### 1. COMPARAISON DE DEUX PREPARATIONS DE LSA

La souche parentale utilisée est la LSA Levuline CHP (souche CIVC n° 8130) et la souche marquée est un clone 8130 OE1 identifiable grâce à l'acquisition d'une

double mutation au niveau de l'ADN mitochondrial lui conférant un phénotype de résistance à l'oligomycine et à l'érythromycine (G. Loiseau, 1986). Une vérification préalable des poudres de LSA a montré qu'elles contenaient environ  $25 \cdot 10^9$  cellules revivifiables par gramme. Dans les deux pou-

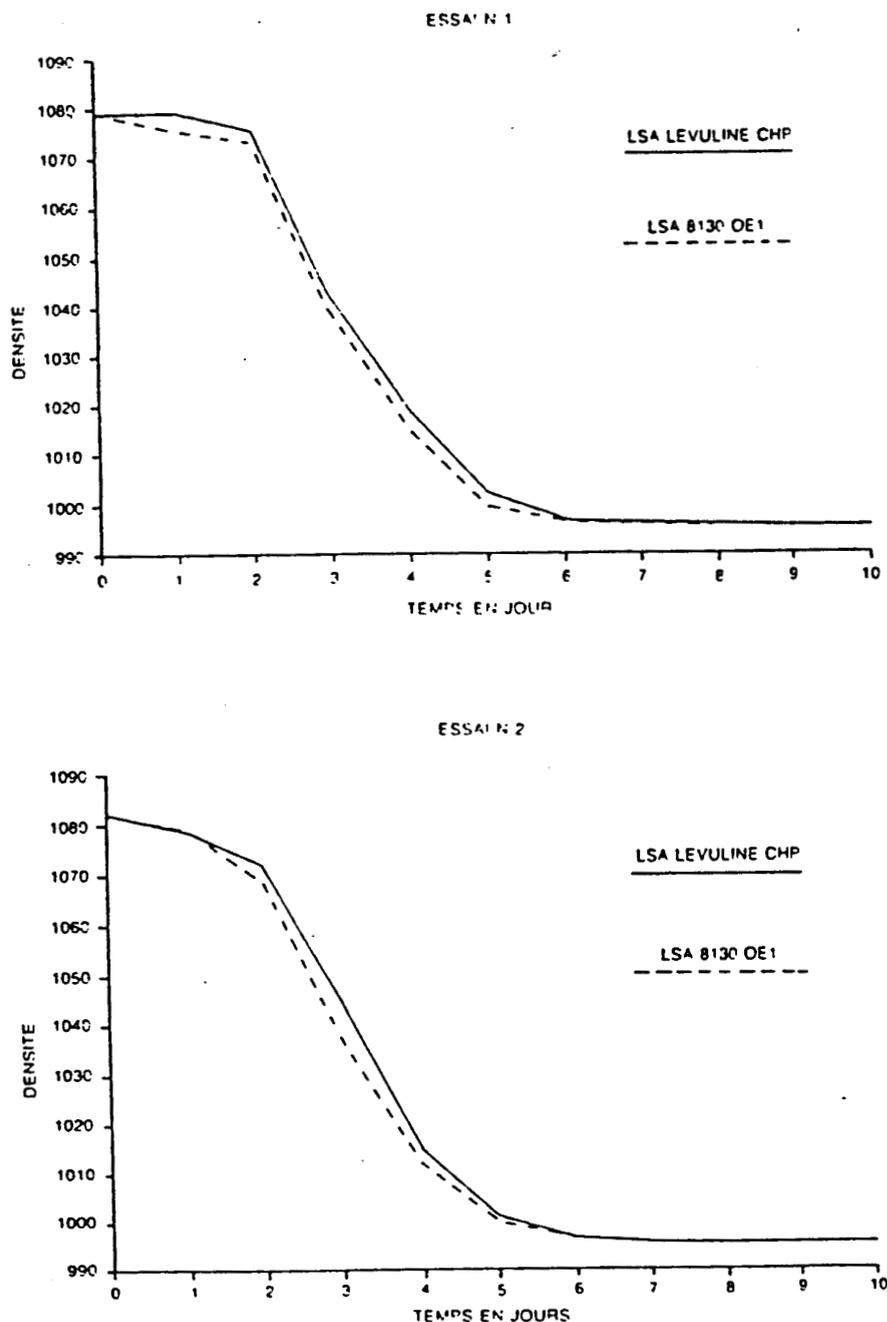


Fig n° 1 - Cinétiques de fermentation

dres, cent pour cent des cellules avaient une activité killer (phénotype de la souche parentale) et dans le cas de la souche 8130 OE1, cent pour cent des cellules étaient résistantes à l'oligomycine et à l'érythromycine.

La mise en œuvre de ces deux préparations a été faite sur des moûts de Champagne dans deux essais réalisés dans des conditions identiques : un même assemblage de moût a été réparti dans deux cuves semblables thermorégulées.

Essai n° 1 Cuves de 80 l - Régulation de la température à 20°C.

Essai n° 2 Cuves de 12 hl - Régulation de la température à 18°C.

Dans chaque essai les cuves ont étéensemencées à raison de 10 g/hl, l'une avec la levuline CHP et l'autre avec la souche 8130 OE1.

La flore indigène des moûts était faible :  $7,7 \cdot 10^1$  cellules par ml pour l'essai n° 1 et  $8,5 \cdot 10^1$  cellules par ml pour l'essai n° 2.

La population levurienne a atteint  $300 \cdot 10^6$  cellules par ml entre le troisième et le quatrième jour de fermentation. Dans le cas où cela a pu être vérifié (souche 8130 OE1), cette population était constituée, tout au long de la fermentation, à 100 % par la souche introduite lors du levurage. Les cinétiques de fermentation ont été établies par un suivi journalier de la densité du moût (figure n° 1). Il n'y a aucune différence significative entre les cuvesensemencées avec la Levuline CHP et cellesensemencées avec la souche 8130 OE1.

Les résultats analytiques en fin de fermentation (tableau n° 1) montrent que les vins obtenus avec l'une ou l'autre souche sont très semblables.

Les tests de dégustation (tableau n° 2) n'ont pas davantage permis de différencier les vins obtenus. En conséquence, la préparation de LSA à partir de la souche modifiée 8130 OE1 a donné des résultats tout

Tableau n° 1 - Résultats analytiques en fin de fermentation alcoolique

| NOM DE L'ECHANTILLON  | ESSAI N° 1       |              | ESSAI N° 2       |              |
|---|------------------|--------------|------------------|--------------|
|   | LSA Levuline CHP | LSA 8130 OE1 | LSA Levuline CHP | LSA 8130 OE1 |
| <b>ANALYSE OENOLOGIQUE</b>                                  |                  |              |                  |              |
| - Densité à 20° C   | 993,5            | 993,5        | 993,45           | 993,55       |
| - Alcool à 20° Aréomètre % V                                | 11° 10           | 11° 10       | 11° 30           | 11° 40       |
| - Glucose (g/l)   | 0,15             | 0,15         | 0,15             | 0,28         |
| - Fructose (g/l)  | 0,15             | 0,15         | 0,32             | 0,39         |
| - Glucose + Fructose (g/l)                                  | -                | -            | 0,47             | 0,67         |
| - Glycérol (g/l)  | 6,33             | 6,22         | 6,54             | 6,75         |
| - pH  | 3,12             | 3,12         | 3,18             | 3,17         |
| - Acidité Totale (H <sub>2</sub> SO <sub>4</sub> g/l)       | -                | -            | 7,20             | 7,30         |
| - Ac. Totale (Méthier) (H <sub>2</sub> SO <sub>4</sub> g/l) | 7,25             | 7,25         | 6,95             | 7,05         |
| - Ac. Volatile Dist (H <sub>2</sub> SO <sub>4</sub> g/l)    | 0,25             | 0,22         | 0,24             | 0,17         |
| - SO <sub>2</sub> Libre (mg/l)                              | 2                | 2            | 8                | 8            |
| - SO <sub>2</sub> Total (mg/l)                              | 8                | 13           | 28               | 33           |
| <b>ACIDES ORGANIQUES</b>                                    |                  |              |                  |              |
| - Acide Tartrique (g/l)                                     | 4,30             | 4,60         | 5,20             | 5,20         |
| - Acide Malique (g/l)                                       | 6,05             | 6,05         | 5,60             | 5,80         |
| <b>ABSORPTION ATOMIQUE</b>                                  |                  |              |                  |              |
| - Potassium (mg/l)  | 760              | 790          | 1120             | 1115         |
| <b>AZOTE</b>  |                  |              |                  |              |
| - Total (mg N/l)  | 430              | 445          | 395              | 395          |
| - Ammoniacal (mg N/l)                                       | 4                | 4            | Traces           | Traces       |
| - Assimilable (Soerensen) (mg N/l)                          | 130              | 122          | 104              | 106          |

à fait comparables à ceux obtenus avec la souche parentale Levuline CHP.

La modification de la souche lors du marquage n'a donc introduit aucune perturbation apparente dans son comportement technologique et cela confirme tous les contrôles qui avaient été réalisés à l'échelle du laboratoire et au niveau pilote (G. LOISEAU et coll., 1984).

## 2. EXPERIMENTATIONS POUR L'OPTIMISATION DU LEVURAGE

### 2.1. Concurrence entre la souche introduite et la flore indigène

L'efficacité d'un levurage était jusqu'alors impossible à contrôler par les méthodes microbiologiques classiques. Tout au plus pouvait-on

Tableau n° 2 - Tests de dégustation pour la comparaison des souches LSA levuline CHP et LSA 8130 OE1

| Test de dégustation   | LSA Levuline CHP                            | LSA 8130 OE1 |
|---|---|--------------|
| <b>Essai n° 1</b><br>- Test triangulaire<br>22 dégustateurs<br>Seuil à 5 % : 12 | 9 réponses exactes<br>Test non significatif |              |
| - Test par paire<br>22 dégustateurs<br>Seuil à 5 % : 17                         | 12  | 10           |
| Test non significatif   |   |              |
| <b>Essai n° 2</b><br>- Test par paire<br>25 dégustateurs<br>Seuil à 5 % : 18    | 14  | 11           |
| Test non significatif   |   |              |

apprécier sa probabilité par le dénombrement de la population indigène du moût au moment du levurage.

Au moyen de la souche marquée 8130 OE1 nous avons pu déterminer de façon fine et précise sa capacité de concurrence vis-à-vis d'une flore indigène plus ou moins importante.

Nous avons modulé le niveau de la flore indigène par l'introduction dans les cuves d'expérimentation d'un volume variable d'un pied de cuve dont la fermentation avait démarré spontanément.

L'expérimentation a été réalisée en Champagne sur un assemblage de moût de Chardonnay simplement débourbé statiquement pendant 15-16 heures. Il a été réparti dans quatre cuves de 12 hectolitres ; la température de fermentation a été réglée à 18°C.

Tableau n° 3 - Estimation des populations indigènes et des populations apportées par le levain

|  | Lot n° 1 | Lot n° 2 | Lot n° 3 | Lot n° 4 |
|--|----------|----------|----------|----------|
| Population indigène 10 <sup>7</sup> /ml moût + levures apportées par le pied de cuve | 0.0085   | 0.2      | 0.6      | 3.4      |
| Population levurienne apportée par le levain 8130 OE1. 10 <sup>7</sup> /ml           | 2.4      | 2.4      | 0.7      | 1.3      |
| R*   | 282      | 12,1     | 1,2      | 0,4      |

\* R =  $\frac{\text{Nombre de levures apportées par le levurage}}{\text{Nombre de levures indigènes}}$

Dans le tableau n° 3 sont indiquées la population apportée par le levurage et la population indigène ainsi que le rapport entre ces populations. Ce dernier varie de 280 à 0,4.

Les populations levuriennes ont été analysées quotidiennement vis-à-vis de la résistance aux deux anti-

biotiques oligomycine et érythromycine dans des conditions similaires à celles décrites précédemment (LOISEAU et coll., 1985) et ces observations ont permis de déterminer avec une précision de l'ordre de 1 % le pourcentage de la souche introduite par le levurage, dans la population totale (figure n° 2).

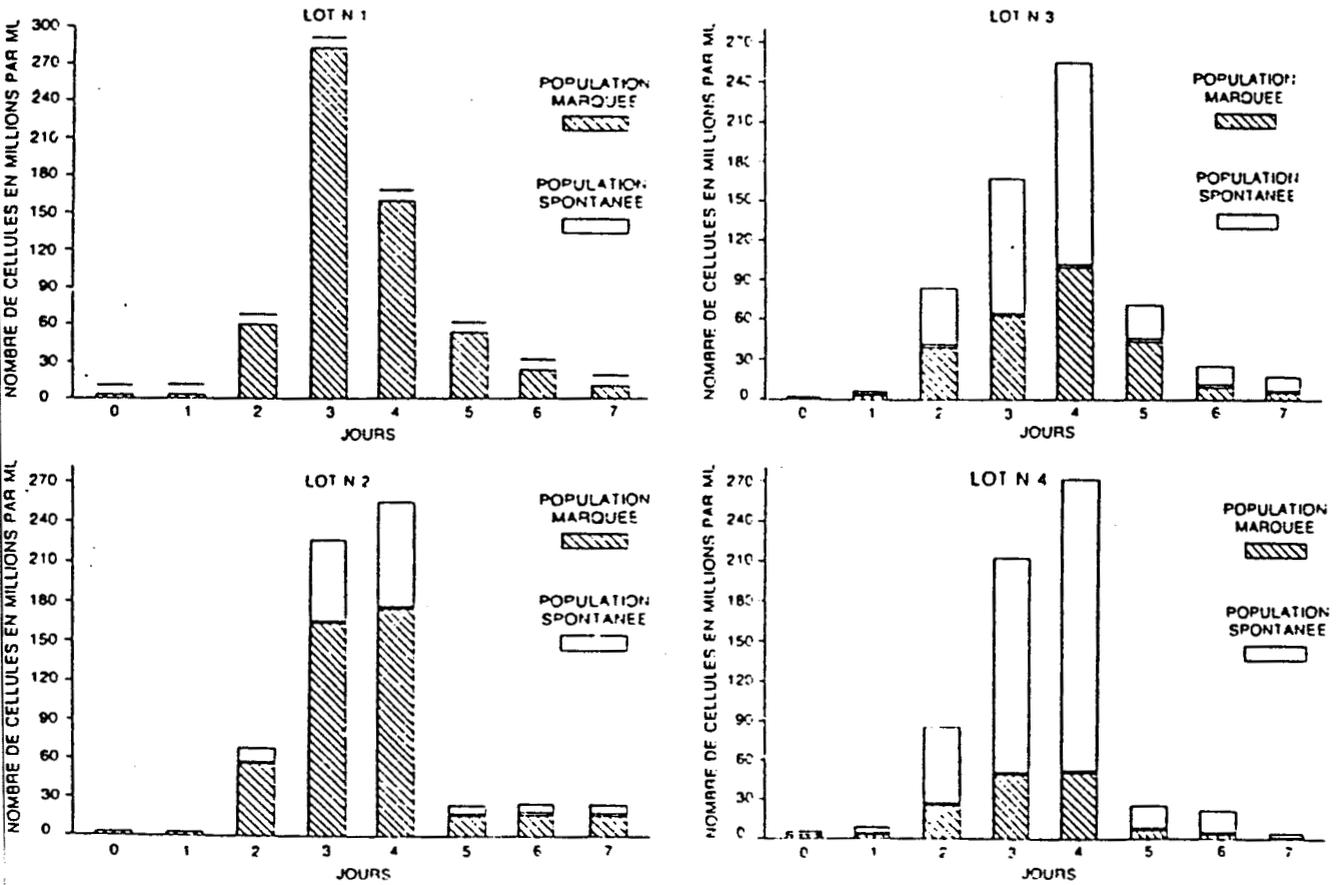


Fig n° 2 - Evolution des populations levuriennes en fonction du taux de levurage

Nous voyons que l'efficacité de l'ensemencement dépend étroitement du rapport de population entre flore apportée par le levurage et flore indigène.

Dans l'expérimentation décrite ici, le levurage n'a conduit à une fermentation en souche pure que dans le cas de la cuve n° 1 où le rapport entre les deux flores est proche de 300. Pour un rapport de l'ordre de 10 la souche introduite est majoritaire ; pour un rapport de 1 ou inférieur à 1, moins de 50 % des clones isolés sont issus de la souche d'ensemencement.

## 2.2. Positionnement du levurage par rapport au sulfitage

L'expérimentation menée en Touraine et décrite ici avait pour objectif de répondre à la question suivante : dans le cas du sulfitage d'un moût à l'encuvage, faut-il réaliser le levurage simultanément et ce dans l'optique de limiter le développement de la flore indigène, ou bien est-il préférable de différer le levurage de quelques heures afin d'éviter une inhibition potentielle des levures par l'anhydride sulfureux.

Un moût de Pineau de la Loire contenant 18 mg de  $SO_2$  total près débouillage a été réparti en cinq récipients de 5 l. :

- Témoin non levuré, non sulfité.
- $SO_2$  3 g/hl levurage immédiat.
- $SO_2$  3 g/hl levurage différé de 8 heures.
- $SO_2$  5 g/hl levurage immédiat.
- $SO_2$  5 g/hl levurage différé de 8 heures.

La fermentation a été conduite à 18-20°C. Le levurage a été réalisé avec la souche 8130 OE2 identifiable par sa double résistance à l'oligomycine et à l'érythromycine.

La cinétique de fermentation a été suivie par mesure de la densité du milieu.

La fermentation du témoin avait pas démarré au bout de 8

jours et l'addition d'un pied de cuve de levures indigènes a dû être réalisée pour conduire la fermentation à son terme en 35 jours.

Dans les 4 autres lots la durée de la fermentation a été de 10 jours avec la dose de  $SO_2$  de 5 g/hl et de 8 jours avec la dose de 3 g/hl.

La nécessité du levurage dans ce cas est donc évidente. Une analyse des populations levuriennes vis-à-vis des résistances aux deux antibiotiques a montré que la fermentation a été réalisée dans les quatre lots exclusivement par la souche introduite lors du levurage. Celui-ci a donc été efficace et dans les conditions de cet essai, l'implantation de la souche du levain s'est bien faite quel que soit le positionnement du levurage : immédiatement après le sulfitage ou 18 heures après ce dernier (CUINIER et coll., 1985).

## 2.3. Le levurage dans différents modes de vinification

Les essais décrits ci-dessous visaient à savoir si le levurage a la même efficacité selon le mode de vinification mis en œuvre.

Dans un premier essai nous avons comparé une vinification avec foulage de la vendange, une vinification avec chauffage de la vendange (20 min à 70°C puis refroidissement à 20°C) et enfin une vinification avec macération carbonique (macération 8 jours à 35°C suivie d'un pressurage et d'une phase de fermentation de 8 jours à 20-25°C).

Le levurage a été réalisé dans les trois cas dès l'encuvage avec un levain liquide d'une souche œnologique DB52 (dérivée de la souche ICV-K1) résistante au chloramphénicol et à l'oligomycine à un taux de  $5 \cdot 10^7$  cellules par ml.

Le suivi de la composition de la population levurienne a été réalisé en dénombrant quotidiennement sur 100 à 1 000 clones, le nombre de clones résistant à l'oligomycine et au chloramphénicol et donc appartenant à la souche inoculée (tableau n° 4).

Pour les vinifications en vendange foulée et en vendange chauffée, la souche inoculée s'implante bien. Aux tous premiers jours de la fermentation dans le cas de la ther-

Tableau n° 4 - Evolution des populations levuriennes dans différents modes de vinification

| Temps en jours | Vendange foulée |           | Thermovinification |       | Macération carbonique |       |
|----------------|-----------------|-----------|--------------------|-------|-----------------------|-------|
|                | N (1)           | Z (2) (%) | N                  | Z (%) | N                     | Z (%) |
| 0              | 118             | 90        | ND                 | ND    | 516                   | 76    |
| 1              | 120             | 95        | 100                | 13    | 693                   | 99    |
| 2              | 469             | 99        | 457                | 99    | 492                   | 99    |
| 3              | 467             | 99        | 381                | 97    | 237                   | 98    |
| 4              | 421             | 100       | 601                | 97    | 355                   | 97    |
| 5              | 389             | 99        | 574                | 99    | 416                   | 94    |
| 6              | 489             | 100       | ND                 | ND    | 466                   | 77    |
| 7              | 359             | 100       | 359                | 99    | 359                   | 29    |
| 8              | 333             | 100       | 395                | 98    | 676                   | 30    |
| 9              | 381             | 99        | ND                 | ND    | 569                   | 20    |
| 10             | ND              | ND        | ND                 | ND    | 98                    | 36    |
| 11             | ND              | ND        | ND                 | ND    | 976                   | 18    |
| 12             | ND              | ND        | ND                 | ND    | 1 002                 | 20    |
| 13             | ND              | ND        | ND                 | ND    | 531                   | 23    |
| 14             | ND              | ND        | ND                 | ND    | 403                   | 23    |

(1) N = Nombre de clones testés

(2) Z = Pourcentage de clones appartenant à la souche introduite par rapport à la flore totale

movinification, elle peut se trouver en présence d'une flore indigène importante mais dans notre essai, cette dernière est mal adaptée à la fermentation alcoolique et est rapidement éliminée.

Les résultats sont assez différents dans le cas de la vinification par macération carbonique. Si la souche du levain s'implante bien dans le jus du fond de cuve pendant les premiers jours de la macération ; par contre, elle se trouve ensuite éliminée en fin de macération alors que la fermentation du jus de goutte n'est pas achevée. L'évolution de la population totale et sa composition (figure n° 3) montre nettement qu'un relais de flores s'établit. La souche introduite en début de macération sans doute mal adaptée aux conditions de cette macération (température, déficience en oxygène) est relayée par une nouvelle flore qui réalise la fermentation alcoolique après pressurage.

Il est bien évident que de telles informations qui n'ont pu être obtenues que grâce au marquage de la souche utilisée, devront être confirmées par des répétitions de cet essai.

Si l'élimination de la souche DB52 était systématiquement observée, le problème de la sélection de souches mieux adaptées à la macération carbonique serait posé.

En effet, dans l'essai décrit ci-dessus un accident par arrêt de fermentation aurait pu se produire si une flore n'avait pu prendre le relais de la souche initiale. Pour pallier l'éventualité de tels accidents, il serait indispensable de sélectionner des souches résistantes aux conditions drastiques de la macération carbonique sur un plan microbiologique.

Dans une série d'essais réalisés par CUINIER et coll. (1984) avec la souche 8130 OE2, l'efficacité du

levurage a été appréciée dans le cas de la vinification beaujolaise et dans le cas d'une vinification d'une vendange de Cabernet récoltée mécaniquement en Touraine. Dans ces deux cas, le levurage s'est révélé efficace à 100 % environ, que l'inoculation ait lieu au moment de l'encuvage ou douze heures après.

Dans un autre essai concernant l'implantation de la souche en fonction du mode de vinification, nous avons testé un système de fermentation de type continu, comprenant trois cuves (1, 2 et 3) de 800 hl et deux cuves (4 et 5) de 4 000 hl et 5 000 hl. La cuve 1 est alimentée en moût frais ; son trop plein coule dans la cuve 2, le trop plein de la cuve 2 dans la cuve 3 et le trop plein de cette dernière dans la cuve 4 puis dans la cuve 5 après remplissage de la cuve 4.

Le levurage avec la souche DB52 a été réalisé au moyen d'un levain de 2 hl développé sur moût stérile. Des contrôles de la composition

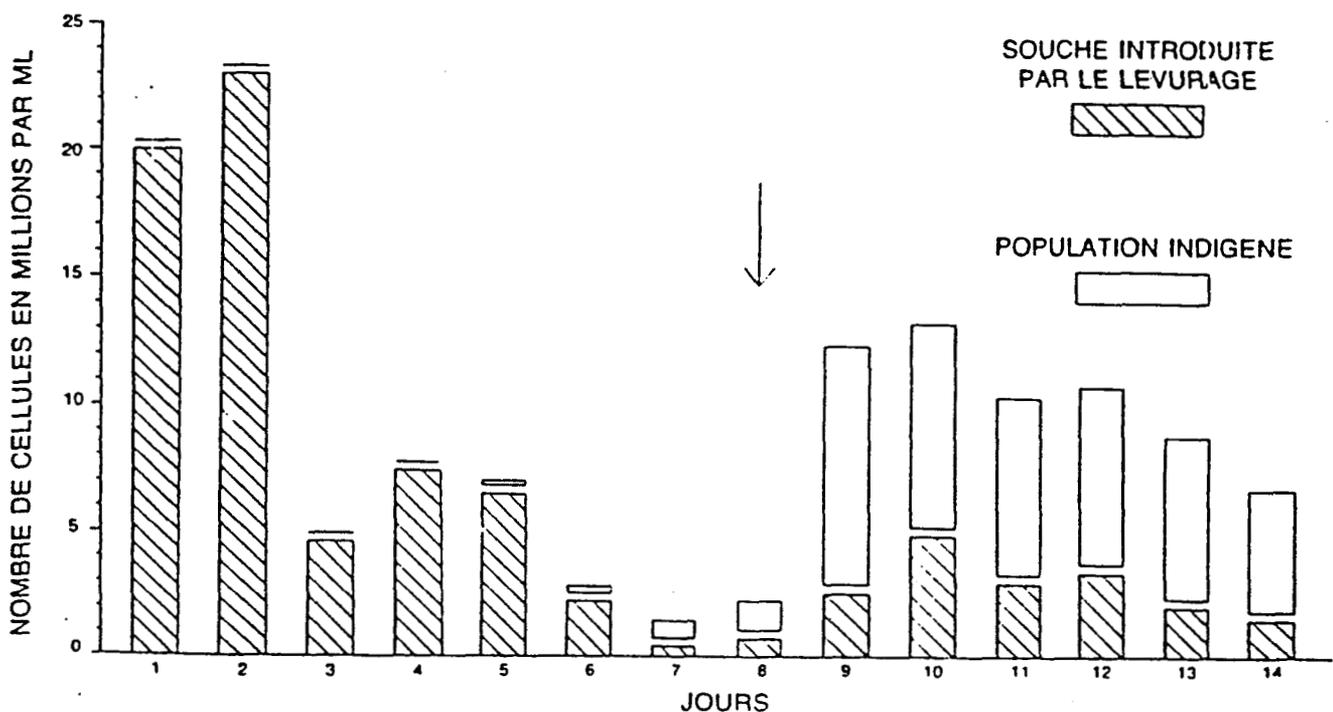


Fig n° 3 - Evolution des populations au cours d'une vinification en macération carbonique

↓ Pressurage de la vendange : les jus de goutte et les jus de presse sont réunis

Tableau n° 5 - Evolution de la composition de la population levurienne dans une chaîne de fermentation continue

| Volume de moût fermenté (hl) | Population viable totale nombre de cellules/ml | % souches indigènes |
|------------------------------|--|---------------------|
| 800                          | 40-48 10 <sup>6</sup>                          | 2-5                 |
| 3 000                        | 42-44 10 <sup>6</sup>                          | 10-19               |
| 8 000                        | 14-38 10 <sup>6</sup>                          | 5-22                |

Les résultats sont donnés en valeurs maximales et minimales observées sur les différentes cuves.

levurienne à différents stades du fonctionnement de cette chaîne ont été réalisés (tableau n° 5). Nous constatons qu'après passage de 800 hl dans l'installation, le taux de flore indigène est négligeable mais par contre après passage de 3 000 hl à 8 000 hl, une « contamination » significative par une flore indigène se produit. Les souches « contaminantes » ont un phénotype killer et le niveau de leur population semble se stabiliser autour de 20 % de la population totale.

Il est certain qu'avec une vinification de ce type, les volumes de moût traités sont tels qu'il est difficile de travailler en souche pure sans une technologie très appropriée.

#### 2.4. Levurage avec association de souches

Avec l'apparition sur le marché d'un nombre croissant de levures œnologiques ayant des spécificités identifiées, il pourrait être intéressant d'associer plusieurs souches au cours d'une vinification et de bénéficier ainsi des propriétés des unes et des autres.

Toutefois, les cultures mixtes étant difficiles à réaliser, pour juger du résultat obtenu, il est absolument indispensable de pouvoir analyser l'état d'équilibre des populations en présence.

Ceci a été réalisé avec deux souches marquées de façon différente lors d'une expérimentation en

Champagne qui a déjà fait l'objet d'une publication (LOISEAU et coll., 1985).

Nous ne décrivons donc pas cette expérimentation mais nous rappellerons simplement les principaux résultats obtenus : un ensemencement simultané en quantités égales (en population) de deux souches de *Saccharomyces cerevisiae* l'une galactose (+), l'autre galactose (-) (anciennement appelée *Saccharomyces bayanus*) a abouti à une population mixte tout au long de la fermentation sans modification sensible du rapport entre les populations des deux souches.

L'expérimentation n'avait pas pour but de juger de l'intérêt de la culture mixte par rapport à la culture pure de l'une ou l'autre des deux souches mais de vérifier la possibilité de réaliser une fermentation par une population mixte, ce qui a été fait.

### DISCUSSION ET CONCLUSIONS

Les expérimentations décrites dans ce travail n'avaient pas pour objet d'apporter des conclusions définitives sur les conditions optimales d'un levurage en œnologie.

En effet, les résultats obtenus doivent être modulés en fonction des conditions de l'expérimentation.

• Dans le cas de la concurrence entre souche ensemencée et flore indigène, nous avons observé la

nécessité d'un rapport élevé entre ces populations pour implanter totalement la souche du levain. Mais il faut noter que la « contamination » créée par des levures d'un pied de cuve en fermentation ne peut être assimilée entièrement à une flore spontanée d'un milieu frais, en général moins concurrentielle.

Il est donc possible que dans la pratique courante la levure apportée puisse s'imposer dans le moût si le rapport des populations souche du levain sur flore indigène est supérieur à 10, voire même à 5 et ce en vinification en blanc.

• Dans le cas du positionnement du levurage par rapport au sulfitage, il est bien évident que les conclusions dépendent de plusieurs facteurs : la dose de SO<sub>2</sub> utilisée, le taux de « contamination » initiale du moût par une flore spontanée, la sensibilité de la souche introduite à l'anhydride sulfureux. Tous ces facteurs n'ont pu être pris en compte dans la seule expérimentation décrite. Les conclusions de celle-ci sont donc partielles.

• Dans le cas de l'étude des différents modes de vinification, nous avons pu voir qu'en vinification classique (vendange récoltée mécaniquement, vendange foulée, vinification beaujolaise, thermovinification) l'implantation de la souche introduite se fait bien. Par contre, avec d'autres modes de vinifications comme macération carbonique ou fermentation de type continu, des problèmes microbiologiques ont pu être mis en évidence et devront être confirmés avant d'étudier des solutions pour y pallier.

Plus que les conditions précises de réussite d'un levurage, nous voyons au travers des différents exemples décrits dans ce travail comment l'outil « souches œnologiques marquées » peut être utilisé par l'œnologue.

• Cette technique peut être pour lui un moyen de contrôler les pou-

dres de LSA qui lui sont fournies et la pureté d'un levain lors de sa propagation.

- Le « marquage » peut lui servir à tester les capacités de différentes préparations de LSA en conditions contrôlées. En effet, si l'on ne peut s'assurer de la participation majoritaire d'une souche à la fermentation, comment tirer des conclusions sur les effets de celle-ci sur le produit final obtenu.

- Les levures marquées sont aussi utilisables pour juger de

l'opportunité d'associations de souches en ayant la possibilité de définir avec précision (de l'ordre de quelques pourcents) la participation des différentes flores à la fermentation.

- Enfin, nous avons illustré ici l'utilisation de souches œnologiques marquées pour montrer les limites ou les possibilités d'accidents microbiologiques avec certains modes de vinification.

L'ensemble de ce travail montre comment les souches œnologiques

marquées rentrent dans la panoplie des moyens dont l'œnologue peut se doter pour accroître toujours plus la maîtrise de la transformation biologique qu'est la fermentation alcoolique.

Ce travail a été réalisé avec l'aide d'une ATP-INRA n°4366, d'une aide à l'innovation de l'ANVAR et d'une aide dans le cadre du contrat de Plan Etat-Région Languedoc-Roussillon. Il a bénéficié de la collaboration du laboratoire départemental et régional d'analyses et de recherches de Tours.

#### BIBLIOGRAPHIE

CUINIER C., GROS C., LACOSTE J., PUISAIS J., 1985 - Comparaison de levures sèches actives utilisées en œnologie. *Rev. Fr. Oenol.*, 97, 41-46.

CUINIER C., GROS C., LACOSTE J., BOISSON R., BERGER J.-L., PUISAIS J., 1984 - Utilisation d'une souche marquée de levure pour comparer l'efficacité de techniques de levurage : vinification en rouge, en Beaujolais et en Touraine. *C.R. Trav., ITV*, 12-14.

CUINIER C., GROS C., LACOSTE J., BOISSON R., BERGER J.-L., PUISAIS J., 1985 - Incidence du délai entre le sulfitage et le levurage sur l'implantation d'une souche marquée de levure. *C.R. Trav. ITV*, 80-83.

KRAUS J.K., REED G., VILLETAZ J.C., 1984 - Levures sèches actives de vinification : utilisation et évaluation. *Conn. Vigne Vin*, 18, 1-26.

LOISEAU G., 1986 - Obtention d'un marquage mitochondrial de levures œnologiques sélectionnées en Champagne. Application au contrôle microbiologique en œnologie. Thèse 3<sup>e</sup> cycle, Université de Montpellier.

LOISEAU G., VALADE M., MOULIN J.-P., 1984 - Vérification technologique de souches porteuses de marqueurs d'antibiorésistances. *C.R. Trav. ITV*, 24-25.

LOISEAU G., VALADE M., MOULIN J.-P., 1985 - Utilisation de souches anti-

bio-résistantes en fermentation alcoolique. Mise en évidence de l'efficacité d'un levurage et étude d'une association de souches. *Le Vigneron champenois*, 4, 186-200.

VALADE M., MOULIN J.-P., 1982 - Premiers résultats concernant l'utilisation en Champagne des levures sèches actives en fermentation alcoolique. *Le Vigneron champenois*, 9, 354-362.

VEZINHET F., 1985 - Le marquage génétique de souches de levures œnologiques. *Rev. Fr. Oenol.*, 97, 47-51.

VEZINHET F., LACROIX S., 1984 - Marquage génétique de levures : outil de contrôle des fermentations en souche pure. *Bull. O.I.V.*, 643-644, 759-773.

#### IV) CONTRIBUTION AU SEQUENCAGE DE LA REGION SACS-GERB DU GENOME DE BACILLUS SUBTILIS

##### A) BIOLOGIE DE DE BACILLUS SUBTILIS

B. subtilis est un bacille sporulant aérobie, mobile, de dimension  $0,7-0,8 \times 2-3 \mu\text{m}^2$ . Grâce à la résistance de son endospore à différents types de stress, B. subtilis peut être isolé à partir des sources les plus variées. Les cellules végétatives sont impliquées dans la dégradation de divers matériaux d'origine végétale et animale. Les colonies de B. subtilis sur milieu solide sont généralement rondes et peuvent devenir marron lors de la sporulation. Le milieu minimum pour la croissance végétative de B. subtilis contient du glucose et des sels d'ammonium comme seules sources de carbone et d'azote. B. subtilis pousse entre pH 5,5 et 8,5 à une température pouvant varier de 10°C à 50°C (Sneath et al., 1986). La teneur en GC de l'ADN de la souche type de B. subtilis (ATCC 6051) est de 42,9% (Fahnyet et al., 1985).

Le cycle cellulaire de B. subtilis comprend 3 phases principales, à savoir la croissance végétative, la sporulation et la germination.

Lors de la phase végétative, la division cellulaire est caractérisée par la formation d'un septum au milieu de la membrane cytoplasmique (Bulla and Hoch, 1985). Durant la croissance végétative, les acides organiques s'accumulent dans le milieu de culture (pyruvate), ce qui résulte en l'abaissement du pH jusqu'à une valeur de 5 à 6. Lors de la transition de la cellule végétative à la cellule sporulante, les acides organiques accumulés dans le milieu de culture sont repris et métabolisés par les cellules. Le catabolisme des acides organiques accumulés se traduit alors par une élévation du pH du milieu de culture (jusqu'à pH 7), et par

l'accumulation d'ATP et de NADH. L'initiation de la sporulation de B. subtilis est induite par une carence en source métabolisable de carbone, d'azote, ou de composés phosphorés (GTP) (Schaeffer et al., 1965b ; Sonenshein, 1985). Toutefois, d'autres réponses sont également induites par de tels stress comme une mobilité accrue vers les plus hautes concentrations de substrats (Ordal et al., 1985), l'induction des enzymes du cycle de Krebs (Hanson et al., 1963) et d'autres enzymes du catabolisme du carbone (Nihashi and Fujita, 1984 ; Doi, 1989), ou comme la dérégulation des gènes codant pour les enzymes dégradatives extracellulaires telles que les protéases, les nucléases, les amylases, les phosphatases etc ... afin de pouvoir utiliser les macromolécules présentes dans le milieu de culture (Schaeffer, 1969).

On distingue 7 stades différents dans la sporulation : les stades 0, II, III, IV, V, VI, et VII (Schaeffer et al., 1965a ; Ryter et al., 1966). La sporulation est un processus qui dure 6 à 8 heures, où chaque stade dure environ 1 heure (Doi, 1989). Il est également remarquable que certaines fonctions (compétence -cf. Dubnau, 1991-, ...) et enzymes (protéases, ...) qui apparaissent tout au long du processus de sporulation ne sont pas nécessaires à la sporulation (Warren, 1968 ; Kay and Warren, 1968 ; Waites et al., 1970 ; Kawamura and Doi, 1984 ; Yang et al., 1984). En outre, B. subtilis utilise de nombreux facteurs  $\sigma$  (holoenzymes de l'ARN polymérase) durant la phase végétative, mais aussi et surtout durant la sporulation où des "cascades" de facteurs  $\sigma$  régulent l'expression de différents promoteurs (Losick and Pero, 1981 ; Moran et al., 1981a ; 1981b ; 1982 ; Doi, 1982a ; Johnson et al., 1983 ; Chamberlin et al., 1985 ; Doi et al., 1985 ; Doi and Wang 1986 ; Stragier and Losick, 1990).

Le processus de germination se divise en 3 phases principales : l'activation, la germination et la croissance (*outgrowth*) (Keynan and Halvorson, 1965). L'activation peut

être induite par une courte exposition à une température élevée dans un milieu adéquat. La germination est le processus de conversion de la spore en une cellule active, et la croissance est le phénomène d'élongation de la cellule jusqu'à la formation d'une cellule végétative complète (Bulla et al., 1980 ; Setlow, 1981). L'importance des protéases au cours de la germination a été démontrée par Foster et Johnstone (1990) qui ont observé que des inhibiteurs de protéases pouvaient stopper la germination de Bacillus cereus ainsi que celle de Bacillus megaterium à un stade précoce.

## B) THEMES ET BUTS DE RECHERCHE

B. subtilis est un microorganisme d'intérêt biotechnologique bien établi. En particulier, B. subtilis est l'un des systèmes de choix avec E. coli et S. cerevisiae pour diriger la sécrétion de protéines homologues ou hétérologues. D'autre part, certaines de ses enzymes, dont les protéases, ont un intérêt industriel de tout premier plan.

L'importance de l'étude de B. subtilis ne se borne pas aux seuls intérêts industriels. En effet, B. subtilis est, après E. coli, le microorganisme le mieux connu du fait de l'intérêt porté à ses systèmes de recombinaison et de transformation génétique particuliers. En outre, cette bactérie est devenue le modèle principal pour l'étude de la sporulation.

L'avènement et la banalisation de techniques efficaces permettant le clonage et le séquençage de l'ADN, ainsi que d'une technologie permettant le traitement des informations récoltées (informatique) rendent possible aujourd'hui l'étude du chromosome d'un organisme dans son entier. Ce changement d'échelle d'étude -de l'opéron au génome- s'avèrera certainement comme une étape importante de la Biologie.

## C) MATERIELS ET METHODES

### 1) SOUCHES BACTERIENNES

La souche type de *B. subtilis* utilisée dans cette étude est la souche de *B. subtilis* 168 (*trpC2*) (Spizizen, 1958) utilisée dans la plupart des laboratoires ; cette souche a été fournie par Anagnostopoulos (INRA, Jouy-en-Josas). Les souches de *B. subtilis* QB692 (*sacA321, narA1*), QB6001 (*trpC2, sacT::aphA3*) (Debarbouillé et al., 1990), QB25 (*sacS49, trpC2*) (Lepesant et al., 1972), QB4238 (*trpC2, Δ(degSdegU)::aphA3*) (Msadek et al., 1990), et QB899 (*thiC*) proviennent de la collection du laboratoire, la souche BGSC 1A486 (*leuA8, rodC1*) a été fournie par le centre BGSC (*Bacillus Genetic Stock Center*) (Ohio State University, Columbus, Ohio). Les souches d'*E. coli* XL1-blue (*supE44, hsdR17, recA1, endA1, gyrA46, thi, relA1, lac<sup>-</sup>, F'[proAB<sup>+</sup> lacI<sup>q</sup> lacZΔM15 Tn10(tet<sup>R</sup>)]*) (Bullock et al., 1987), TG1 (*supE, hsd5, thi, Δ(lac<sup>-</sup>proAB), F'[traD36 proAB<sup>+</sup> lacI<sup>q</sup> lacZΔM15]*) (Gibson T.J., 1984, PhD thesis, Cambridge University, Angleterre), et FF283 (*F<sup>-</sup>, lacΔx74, araD139, [araABO1C<sup>-</sup>leu] Δ7679, galU, galK, rpsL/F'lac<sup>-</sup>prolacI<sup>q</sup>*) (Brown and Fournier, 1984) proviennent de la collection du laboratoire. Les souches d'*E. coli* LE392 (*supE44, supF58, hsdR514(=r<sub>k</sub><sup>-</sup>m<sub>k</sub><sup>-</sup>), galK2, galT22, metB1, trpR55, lacY1(=ΔlacIZY)*) (Borck et al., 1976 ; Murray et al., 1977) et son lysogène pour le phage P2 (P2392) ont été fournies par Stratagene (La Jolla, Californie).

### 2) PHAGES ET PLASMIDES

Le phage λFixII a été fourni par Stratagene et le phage M13mp8 (Messing and Vieira, 1982) par Pharmacia (Upsalla, Suède). Le plasmide pBluescript (Amp<sup>R</sup>) a été fourni par Stratagene ; le plasmide pJH101-NotI (Amp<sup>R</sup>Cm<sup>R</sup>Tet<sup>R</sup>) a été construit par Glaser (Unité de Régulation de l'Expression Génétique) par insertion d'un site NotI dans le plasmide pJH101 (Ferrari et al., 1983), le plasmide pAC5 (Amp<sup>R</sup>Cm<sup>R</sup>) a été construit par Martin-Verstraete (Unité de Biochimie Microbienne) ; le plasmide pMTL22 (Amp<sup>R</sup>) (Chambers et al., 1988), le plasmide pSU21 (Cm<sup>R</sup>) (Martinez et al., 1988) et le plasmide pMK4 (Amp<sup>R</sup>Cm<sup>R</sup>) (Sullivan et al., 1984) proviennent de la collection du laboratoire ; le plasmide pTCaphA3 (Kan<sup>R</sup>) (Trieu-Cuot and Courvalin, 1983) a été fourni par Trieu-Cuot (Unité des Agents Antibactériens) ; le plasmide pMF2 (Amp<sup>R</sup>Cm<sup>R</sup>) a été construit par Kunst et Débarbouillé (Unité de Biochimie Microbienne) à partir du plasmide pMF1 (Débarbouillé et al., 1987) en procédant à un échange allélique (Iglesias et al., 1981) du gène *sacY* dans la souche type de *B. subtilis* 168.

### 3) MILIEUX ET CONDITIONS DE CULTURE

E. coli a été cultivé sous agitation à 37°C ou 42°C, B. subtilis a été cultivé sous agitation à 37°C ou 45°C. Les milieux de culture utilisés pour E. coli sont : le milieu "LB" (10 g/l tryptone, 5 g/l Yeast Extract, 5 g/l NaCl, 1 g/l glucose), le milieu "SOB" (20 g/l tryptone, 5 g/l Yeast Extract, 0,5 g/l NaCl), et le milieu "2YT" (16 g/l tryptone, 10 g/l Yeast Extract, 5 g/l NaCl). B. subtilis a été cultivé soit dans du milieu "LB", soit dans du milieu "Penassay Antibiotic Medium 3" (Difco) (1,5 g/l Beef Extract, 1,5 g/l Yeast Extract, 5 g/l peptone, 1 g/l dextrose, 3,5 g/l NaCl, 3,7 g/l  $K_2HPO_4$ , 1,3 g/l  $KH_2PO_4$ , pH 7,0), soit dans du milieu minimum "MM" (Anagnostopoulos and Spizizen, 1961) (15 mM  $(NH_4)_2SO_4$ , 60 mM  $K_2HPO_4$ , 45 mM  $KH_2PO_4$ , 1,5 mM  $MgSO_4$ , 1 g/l citrate de sodium tetrasodique, 10 g/l source de carbone, 100 mg/l facteur d'auxotrophie), soit dans du milieu de sporulation "SP" (8 g/l Nutrient Broth, 1 mM  $MgSO_4$ , 22 mg/l citrate ammonioferrique, 10 mM KCl, 0,5 mM  $CaCl_2$ , 10  $\mu M$   $MnCl_2$ ). Les milieux solides sont obtenus en ajoutant au milieu liquide 15 g/l d'agar (Difco). Les colonies résistantes ont été sélectionnées chez E. coli en ajoutant selon les cas soit 100  $\mu g$ / ml d'ampicilline, soit 5  $\mu g$ / ml de chloramphénicol, soit 100  $\mu g$ / ml d'ampicilline et 2,5  $\mu g$ / ml de chloramphénicol. Les niveaux de sélection utilisés chez B. subtilis étaient de 5  $\mu g$ / ml de chloramphénicol, ou 5  $\mu g$ / ml de kanamycine, ou 1  $\mu g$ / ml d'erythromycine et 25  $\mu g$ / ml de lincomycine.

### 4) STOCKAGE DES SOUCHES BACTERIENNES

Les souches bactériennes ont été conservées à -70°C dans du milieu de culture additionné de 15% de glycérol.

### 5) CULTURE DU PHAGE $\lambda$

Le phage  $\lambda$  a été cultivé sur les souches d'E. coli LE392 ou P2392 ; cette dernière est utilisée pour bénéficier du phénomène d'inhibition de fertilité des phages  $\lambda$  non recombinants par le phage P2. Le phage  $\lambda$  a été dilué de façon adéquate dans du tampon SM (50 mM Tris, 10 mM  $MgSO_4$ , 100 mM NaCl, 0,01% gélatine, pH 8,0) ou dans du milieu LB additionné de 10 mM  $MgSO_4$  (LBM), puis des cellules d'E. coli LE392 ou P2392 cultivées et diluées dans du milieu LBM à une densité optique à 600 nm de 0,4 ont été infectées à la multiplicité d'infection choisie (*moi*) pendant 15 min à 37°C. Les cellules infectées ont été diluées dans 5 ml de gélose molle fondue (10 g/l tryptone, 5 g/l Yeast Extract, 5 g/l NaCl, 2 g/l  $MgSO_4$ , 5 g/l agarose), et le mélange a été déposé sur des

boîtes  $\lambda$  (10 g/l tryptone, 5 g/l NaCl, 2 g/l MgSO<sub>4</sub>, 10 g/l agarose) ou RMg (10 g/l tryptone, 1 g/l Yeast Extract, 8 g/l NaCl, 2 g/l MgSO<sub>4</sub>, 15 g/l agar). Après solidification de la gélose molle, les boîtes ont été incubées une nuit à 37°C.

#### **6) PREPARATION D'UN STOCK DE PHAGES $\lambda$**

Les stocks de phages  $\lambda$  ont été préparés en déposant 5 ml de tampon SM sur une boîte  $\lambda$  ou RMg qui a été incubée à 37°C toute une nuit dans des conditions de lyse confluente. Après élution des particules de phages (de 1 à 24 h à 4°C), le tampon SM a été récupéré, puis quelques gouttes de chloroforme y ont été ajoutées afin de tuer toute cellule bactérienne. La suspension a ensuite été stockée à 4°C jusqu'à utilisation.

#### **7) TRANSFECTION D'*E. COLI* AVEC LE PHAGE M13**

La souche XL1-Blue d'*E. coli* a été transfectée avec le mélange de ligation du phage M13mp8 (clonage en *shotgun*) par la méthode de Hanahan (1983). Les cellules transfectées additionnées de 200  $\mu$ l d'une culture de nuit ont été diluées dans 3 ml de RTop fondu (10 g/l tryptone, 1 g/l Yeast Exwtract, 8 g/l NaCl, 8 g/l agar, 2 ml/l d'une solution molaire de CaCl<sub>2</sub>, 5 ml/l d'une solution de glucose à 20%), puis le mélange a été déposé sur des boîtes LB. Après solidification du RTop, les boîtes ont été incubées une nuit à 37°C.

#### **8) DETECTION DE L'ACTIVITE DE LA $\beta$ -GALACTOSIDASE**

L'activité de la  $\beta$ -galactosidase (couleur bleue = positif) a été détectée chez *E. coli* en ajoutant au milieu de culture utilisé 40  $\mu$ l/ boîte d'une solution de 20 mg/ ml de Xgal dans de la diméthylformamide et 20  $\mu$ l/ boîte d'une solution aqueuse de 100 mM d'IPTG.

#### **9) TRANSFORMATION DE CELLULES COMPETENTES D'*E. COLI***

Les expériences de transformation des cellules d'*E. coli* ont été effectuées selon la méthode de Hanahan (1983) et selon la méthode décrite par Dagert et Ehrlich (1979)

#### **10) TRANSFORMATION DE CELLULES COMPETENTES DE *B. SUBTILIS***

Les cellules compétentes de *B. subtilis* ont été obtenues et transformées selon les protocoles décrits par Anagnostopoulos et Spizizen (1961) et Kunst et coll. (1988). La période d'expression phénotypique a été réalisée en présence de quantités sous-inhibitrices d'antibiotique afin de lever les phénomènes d'atténuation traductionnelle (Bechhofer, 1990 ; Lovett, 1990 ; Rogers et al., 1990).

## 11) ISOLEMENT D'ADN

### a) ADN chromosomique

L'ADN chromosomique a été préparé selon la méthode décrite par Msadek et coll. (1990).

### b) ADN plasmidique

L'ADN plasmidique a été préparé en grand volume en utilisant des colonnes Qiagen selon le protocole décrit par le fournisseur. Les minipréparations d'ADN plasmidique ont été effectuées selon la méthode décrite par Sambrook et coll. (1989).

### c) ADN du phage $\lambda$

Les préparations en grand volume (1 l) d'ADN du phage  $\lambda$  ont été réalisées en utilisant la méthode des colonnes Qiagen selon le protocole décrit par le fournisseur. Les lysats de phage utilisés ont été obtenus en milieu liquide comme décrit par Sambrook et coll. (1989), ou en milieu solide sur boîte  $\lambda$  comme décrit précédemment pour l'obtention des stocks de phages.

Les minipréparations de phages  $\lambda$  ont été réalisées en suivant une modification du protocole décrit par Sambrook et coll. (1989). Une plaque de phage  $\lambda$  recombinant a été mise à éluer dans 200  $\mu$ l de tampon SM pendant 4 à 6 h à 4°C. Une culture d'*E. coli* LE392 diluée à une densité optique à 600 nm de 0,4 à 0,5 a été infectée (200  $\mu$ l) avec 50  $\mu$ l de la suspension de phages ; les cellules infectées ont alors été diluées dans 4 ml de gélose molle fondue, et le mélange a été déposé sur une boîte  $\lambda$  (obtention d'une lyse confluyente). Après une nuit d'incubation à 37°C, 3 ml de tampon SM ont été déposés sur la gélose et la boîte de Pétri a été agitée à 4°C afin de permettre l'élution du phage. Le stock de phages ainsi obtenu (2 ml) a été transféré dans un tube Eppendorf puis centrifugé 10 s afin d'éliminer les bactéries. Une solution (2  $\mu$ l) de DNase (0,5 mg/ml) et RNase (0,5 mg/ml) a été ajoutée au surnageant, puis le mélange a été incubé 15 min à 37°C. Après addition d'un volume d'une solution de 20% (w/v) de PEG6000 et de 2 M NaCl, les phages ont été précipités par incubation pendant 1 h à 4°C. Le surnageant a été éliminé quantitativement après une centrifugation pendant 10 min à 4°C, 10000 rpm. Le culot de phages a alors été resuspendu dans 250  $\mu$ l de TE (10 mM Tris, 1 mM EDTA, pH 8,0), puis 5  $\mu$ l d'une solution de SDS à 10% ont été ajoutés. Après 5 min d'incubation à 68°C, 10  $\mu$ l d'une solution de NaCl 5 M ont été ajoutés et une extraction au phénol suivie d'une extraction au chloroforme ont été réalisées. L'ADN obtenu après

précipitation à l'éthanol a finalement été dissous dans 50 µl de TE et conservé à 4°C.

#### d) ADN simple brin du phage M13

Les minipréparations des matrices simple brin provenant des phages recombinants M13mp8 ont été réalisées à l'aide de microplaques de 96 puits (Nunc) et d'une centrifugeuse GPR (Beckman) en utilisant le protocole décrit par Eperon (1986) et Eperon et coll. (1988) et modifié par Smith et coll. (1990). Les matrices double brin provenant des phages recombinants M13mp8 ont été cultivées pendant 6 h dans la souche d'E. coli XL1-Blue (1 ml d'une dilution au 1/100 d'une culture de nuit dans du 2YT) et isolées selon le protocole de Birnboim et Doly (1979) dont l'extraction au phénol et au chloroforme a été remplacée par une précipitation au chlorure de lithium à une concentration finale de 2 M dont le rôle est de précipiter ARN et protéines.

#### e) Extraction d'ADN à partir d'un gel d'agarose

Les extractions d'ADN à partir de fragments d'agarose ont été conduites en utilisant le kit GeneClean (Bio101) en suivant le protocole décrit par le fournisseur.

#### f) Electroélué

Les fragments à électroéluer ont été placés dans des cupules à électroélué placées dans un bac à électroélué contenant du TAE 0,25 X. L'électroélué a été effectuée à 80 V pendant 1 h 30, puis la polarité a été inversée pendant 30 s afin de détacher l'ADN de la membrane à dialyse. Après extraction du bromure d'éthidium à l'aide de butanol, l'ADN a été purifié par extraction au phénol et au chloroforme puis resuspendu dans du TE après précipitation dans de l'éthanol à 100%.

### 12) SYNTHÈSE D'OLIGONUCLEOTIDES

Les oligonucléotides utilisés ont été synthétisés à l'aide d'un synthétiseur d'ADN "Cyclone Plus DNA Synthetiseur" (Milligen/Biosearch, Division of Millipore, Bedford, Massachussets).

### 13) PCR (POLYMERASE CHAIN REACTION)

Les réactions de PCR ont été conduites en utilisant l'ADN polymérase de Thermus aquaticus (Taq polymérase) (New England Biolabs) selon les conditions décrites par Msadek et coll. (1990) en utilisant un bain sec programmé ("Hybaid, Teddington, Angleterre).

### 14) ENZYMES

Les enzymes de restriction ont été fournies par Pharmacia, New England Biolabs, et Boehringer-Mannheim. La ligase du phage T4 a été fournie par New England Biolabs ou

par Pharmacia (clonage en *shotgun* dans M13), le fragment Klenow de l'ADN polymérase I a été fourni par Boehringer-Mannheim, l'ADN polymérase du phage T4 et l'ADN polymérase du phage T7 ont été fournies par Pharmacia. Toutes les enzymes ont été utilisées selon les recommandations du fournisseur.

#### **15) MARQUAGE DE L'ADN PAR TRANSLATION DE CÉSURE**

Les marquages par translation de césure (Rigby, 1977) ont été effectués à l'aide d'un kit de translation de césure (Boehringer-Mannheim) en ce qui concerne les sondes radioactives, et à l'aide du kit "Bionick Labeling System" (BRL) en ce qui concerne les sondes froides, en suivant les indications du fournisseur. Après 40 min d'incubation à température ambiante (sondes radioactives) ou 1 h d'incubation à 16°C (sondes froides), les réactions de translation de césure ont été stoppées par addition de 4 µl d'EDTA 250 mM pour 20 µl de mélange réactionnel. La sonde a été précipitée par addition (pour 20 µl de mélange réactionnel) de 20 µl d'acétate d'ammonium 5 M, 10 µg de Dextran T40 et 110 µl d'éthanol 100%, suivie d'une incubation pendant 15 min dans un bain de glace pour permettre à l'ADN de précipiter.

#### **16) TRANSFERT DE L'ADN SUR DES MEMBRANES**

Les *Southern blot* ont été réalisés comme décrit par Sambrook et coll. (1989) sur des membranes de nylon "Hybond N" type RPN203N (Amersham). L'ADN a été fixé à la membrane par exposition aux rayons UV en utilisant un "Fluolink" (Vilber Lourmat, Torcy, France) réglé pour délivrer 3 J/cm<sup>2</sup>.

Les transferts de l'ADN de phages M13 ou de phages λ ont été réalisés selon un protocole standard. Des membranes de nitrocellulose (Référence 401144, Schleicher and Schuell, Dassel, Allemagne ; ou Hybond C type RPN82C, Amersham) ont été utilisées pour l'hybridation avec des sondes radioactives, et des membranes de nylon (Hybond N type RPN82N, Amersham) ont été utilisées pour l'hybridation avec des sondes froides. L'utilisation de membranes de nylon apparaît nécessaire pour les détections des sondes froides, sans doute du fait de la faculté du nylon de modifier la longueur d'onde du spectre de chemiluminescence (Tizard et al., 1990). Les membranes ont été déposées pendant 5 min à température ambiante sur les boîtes de Pétri, préalablement incubées à 4°C pendant 30 min afin de raffermir la gélose. Elles ont ensuite été déposées pendant 5 min à température ambiante sur du papier Whatman imprégné de tampon de dénaturation (0,5 N NaOH, 1,5 M NaCl) puis neutralisées sur du papier Whatman imbibé de tampon de neutralisation (1 M Tris, 0,75 M NaCl, pH 7,0) pendant 5 min à température ambiante. Des repères (1 ng de la sonde dénaturée) ont ensuite été déposés (1 µl) afin de pouvoir déterminer la

position relative de la membrane par rapport à la boîte de Pétri. L'ADN a ensuite été fixé sur les membranes par cuisson pendant 2 h à 80°C.

### 17) HYBRIDATIONS

Les membranes ont été préhybridées pendant 15 min à 42°C dans du tampon de préhybridation (5XSSPE, 50% formamide) (20XSSPE = 174 g/l NaCl, 27,5 g/l  $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ , 7,5 g/l EDTA, pH du mélange dans 800 ml d'eau ajusté à 6,55 puis complété à 1 l). Les sondes d'ADN ont été dénaturées dans un bain-marie pendant 5 min à 100°C ou dans un bain sec à 100°C pendant 10 min. La sonde dénaturée (300000 à 350000 cpm/ ml pour les sondes radioactives, et de 0,5 à 1 µg d'ADN pour les sondes froides) a ensuite été rajoutée au tampon d'hybridation (5XSSPE, 50% formamide, 1X Denhardt) (50X Denhardt = 5 g/l ficoll, 5 g/l polyvinylpyrrolidone, 5 g/l albumine bovine - BSA-) et pouvait être conservée dans ce tampon à 4°C et réutilisée plusieurs fois sans autre étape de dénaturation. L'hybridation a été réalisée à 42°C pendant 15 h (éventuellement jusqu'à 65 h). Trois lavages ont ensuite été effectués à 42°C pendant 15 min dans du tampon de préhybridation, du 2XSSPE-0,1% SDS, et du 0,5XSSPE-0,1% SDS. Les membranes qui ont été hybridées avec une sonde froide marquée à la biotine ont été traitées suivant une modification du protocole fourni par BRL pour le kit "Photogene" (BRL). Les membranes (10 membranes de diamètre 85 mm ou 1 blot de 180x160 mm<sup>2</sup>) hybridées et lavées ont été rincées pendant 1 min dans 100 ml d'une solution de TBS-Tween20 (100 mM Tris, 150 mM NaCl, 0,05% w/v Tween20 = polyoxyéthylène sorbitan monolaurate, pH 7,5 ajusté à l'aide d'acide chlorhydrique 4 M) (Tween20 Sigma) puis incubées à 65°C pendant 1 h dans 200 ml d'une solution d'albumine bovine (6 g de BSA/ 200 ml de TBS-Tween20) (BSA référence 735086, Boehringer-Mannheim). La solution de BSA peut être réutilisée de nombreuses fois et stockée à 4°C après addition de 0,01% d'azide de sodium dont le rôle est d'empêcher toute prolifération bactérienne. Après centrifugation pendant 4 min afin d'éviter l'addition de molécules précipitées, le conjugué streptavidine-phosphatase alcaline (SA-AP) est ajouté à du tampon TBS-Tween20 (50 ml total et 15 µl de SA-AP) dans lequel les membranes sont incubées 10 min à température ambiante, puis lavées deux fois dans 250 ml de TBS-Tween20 pendant 15 min à température ambiante. Un dernier lavage est effectué à température ambiante dans du "Final Wash Buffer" 1X (fourni avec le kit) pendant 1 h 45 à 2 h. Après avoir enlevé l'excès de liquide des membranes à l'aide de papier buvard, le réactif de détection (300 µl/ membrane ou 1 à 2 ml par blot) est réparti de façon homogène sur chaque membrane qui est ensuite enveloppée dans une feuille de saran puis placée dans une cassette pour autoradiographie. Le signal obtenu est stable pendant 24 h après une période de 3 à 5 h durant laquelle il se développe.

### **18) AUTORADIOGRAPHIE**

Les films utilisés pour l'autoradiographie étaient des films Kodak X-OMAT AR ou DEF5 (Eastman Kodak), ou hyperfilm (Amersham). Les autoradiographies des gels de séquence ont été réalisées à l'aide de films Fuji X-Ray (Fuji). Les films ont été développés à l'aide d'une machine à développer Kodak X-OMAT "Processor Model ME3" (Eastman Kodak).

### **19) ELECTROPHORESE**

Les électrophorèses analytiques ont été réalisées sur gel d'agarose à 0,7% (1,5% pour les produits de PCR) dans du tampon TAE (4 mM Tris acétate, 1 mM EDTA, pH 8,0) ajusté avec de l'acide acétique. Des solutions de bleu de bromophénol (50% glycérol, 0,1 M EDTA, 0,3% bleu de bromophénol ; ou 25% glycérol, 5% SDS, 0,025% bleu de bromophénol) ont été ajoutées aux échantillons comme marqueur de migration.

L'ADN a été visualisé à l'aide de bromure d'éthidium (soit 1 µg/ ml dans le tampon de migration, soit après migration dans des solutions de 1 µg/ ml de bromure d'éthidium) par irradiation aux rayons UV. Les photographies ont été réalisées en utilisant des films Polaroid type 665 à l'aide d'un filtre rouge Kodak Wratten # 23A.

Les électrophorèses en champs pulsé ont été effectuées dans du tampon TBE (25 mM Tris, 90 mM acide borique, 2,5 mM EDTA, pH 8,5) avec un FA Rotor (Rotaphor type IV, Biometra, Göttingen, Allemagne). L'angle du rotor a varié de 120 à 95°, le temps de pulsation a varié de façon linéaire de 30 à 3 s, la tension de migration utilisée était de 200 V. La migration s'est effectuée à 10°C sur gel d'agarose horizontal à 1% pendant 16 h.

### **20) DETECTION DE L'ACTIVITE PROTEOLYTIQUE**

L'activité protéolytique des phages recombinants isolés a été détectée en infectant des cellules d'E. coli LE392 avec les phages à tester, et en déposant sur des boîtes à les cellules infectées diluées dans du TYNM agarose (10 g/l tryptone, 5 g/l Yeast Extract, 5 g/l NaCl, 2 g/l MgSO<sub>4</sub>, 4 g/l agarose) supplémenté de lait écrémé (80 g/l) et de 2 mM de chlorure de calcium. Un résultat positif est caractérisé après 16 h d'incubation à 37°C par une clarification du lait écrémé.

### **21) DETECTION DU CARACTERE CONSTITUTIF DE LA SACCHARASE**

Les colonies à tester ont été recouvertes d'un mélange d'agar fondu à 7,5 g/l, de 25% de sucrose, de 5 mg/ ml de lysozyme, et de 250 µg/ ml de chloramphénicol. Après incubation pendant 15 min à 37°C, la présence de glucose a été détectée à l'aide du kit "God-Perid" de détection du

glucose (Boehringer-Mannheim). Une coloration verte indique la présence de glucose et donc la présence de saccharase.

## **22) ENCAPSIDATION DES PHAGES $\lambda$ FIXII RECOMBINANTS**

L'encapsidation des phages  $\lambda$ FixII recombinants a été réalisée en utilisant un kit d'encapsidation "Gigapack XL" de Stratagene selon les instructions du fournisseur.

## **23) SONICATION DE L'ADN**

L'ADN des phages  $\lambda$  recombinants a été soniqué à l'aide d'un sonicateur Vibra-Cell (Sonics and Materials, Danbury, Connecticut).

## **24) REPARATION DES EXTREMITES 5' ET 3' DES FRAGMENTS D'ADN**

Les extrémités 5' et 3' des fragments d'ADN ont été réparés à l'aide du fragment Klenow de l'ADN polymérase I (Boehringer-Mannheim) ou à l'aide de l'ADN polymérase du phage T4 (Pharmacia) soit pour obtenir un remplissage partiel des sites de restriction, soit pour obtenir des extrémités franches.

## **25) CLONAGE EN SHOTGUN DANS LE PHAGE M13mp8**

Le phage M13mp8 coupé par SmaI et déphosphorylé a été fourni par Pharmacia. L'ADN à cloner a été soniqué, les extrémités 5' et 3' ont ensuite été réparées à l'aide du fragment Klenow de la polymérase I, puis l'ADN a été séparé par électrophorèse sur gel d'agarose à 1% en fragments de 500 à 800 paires de bases (bp), 800 à 1000 bp, 1000 à 1500 bp. Les fragments ont ensuite été purifiés du gel d'agarose à l'aide d'un kit GeneClean (Bio 101) puis une deuxième réparation a été opérée à l'aide de l'ADN polymérase du phage T4 (Pharmacia). Après avoir inactivé l'ADN polymérase par incubation pendant 15 min à 70°C, la ligation des fragments avec le phage M13mp8 a été effectuée à l'aide de la ligase du phage T4 (Pharmacia) en utilisant un rapport vecteur/ insert de 2/1. La souche d'E. coli XL1-Blue a ensuite été transfectée par le mélange de ligation obtenu en utilisant le protocole décrit par Hanahan (1983) et les phages recombinants ont été détectés à l'aide d'IPTG et de Xgal. Les plages blanches ont été repiquées sur des boîtes LB fraîches, et les phages contenant un insert d'ADN chromosomique ont été détectées par hybridation des phages transférées sur des membranes de nitrocellulose, hybridées avec une sonde radioactive obtenue par marquage au <sup>32</sup>P de l'insert d'ADN chromosomique de B. subtilis séparé de l'ADN du phage  $\lambda$  par digestion avec l'enzyme SalI et électrophorèse sur gel d'agarose.

## **26) SEQUENCAGE**

Les séquences ont été obtenues par la méthode des didéoxynucléotides développée par Sanger et coll. (1977) en utilisant l'ADN polymérase du phage T7 (Tabor and Richardson, 1977), fournie par Pharmacia. Les amorces utilisées étaient CAGCACTGACCCTTTTG (*primer* -40) et AACAGCTATGACCATG (*reverse primer*), pour la forme simple brin et la forme répliquative du phage M13, respectivement. Les réactions de séquence ont été réalisées automatiquement en utilisant un robot "Dilutor 401" (Gilson Medical Electronics, Middleton, Wisconsin). Les gels d'acrylamide ont été réalisés en conditions dénaturantes (7 M urée) et en gradient de tampon comme décrit par Biggin et coll. (1983) afin de permettre une répartition des bandes relativement uniforme le long du gel. Les réactions de séquence ont été réalisées avec du 7-deaza-dGTP (Pharmacia, afin d'éviter les problèmes de compression et d'artefacts dans les gels.

## **27) TRAITEMENT DES SEQUENCES**

Les séquences d'ADN obtenues ont été compilées et analysées à l'aide du programme écrit par Staden (1979, 1980) et modifié par Caudron (Unité d'Informatique Scientifique) pour l'utilisation dans l'ordinateur central MV10000 de l'Institut Pasteur.

La recherche d'identité entre deux protéines a été effectuée selon le test de Lipman et Pearson (1985) en utilisant la matrice Pam250 (ou *Mutation Data Matrix*) qui permet de tenir compte des mutations survenues au cours de l'évolution des protéines. La signification d'une relation trouvée lors d'une recherche dans les banques de données dépend de la grandeur relative du score initial, de la variation de ce score après optimisation et de l'alignement entre la séquence test et la séquence de la banque de données.

Une estimation de la signification statistique des relations trouvées peut être obtenue en comparant la séquence test avec la séquence résultant de la recherche, permutée plusieurs fois de façon aléatoire. Chaque séquence ainsi générée garde la longueur exacte et la composition en acides aminés de la séquence originale de la banque de données. L'écart type  $z$  des scores optimisés de la comparaison de la séquence test avec la séquence connue permutée plusieurs fois est la grandeur utilisée pour apprécier si la relation entre les deux protéines est significative. Avec  $z > 3$  le résultat est peut-être significatif, avec  $z > 6$  le résultat est probablement significatif, avec  $z > 10$  le résultat est significatif (Lipman and Pearson, 1985).

La recherche des profils hydropathiques des protéines a été effectuée selon la méthode décrite par Kyte et Doolittle (1982) et Hopp et Woods (1981).

### **28) CONCENTRATIONS D'ADN**

Les concentrations d'ADN ont été estimées à l'aide d'un spectrophotomètre "Ultrospec II" (Pharmacia).

### **29) PREPARATION DES SONDAS ARN**

L'ADN du phage  $\lambda$  recombinant a été hydrolysé à l'aide d'une enzyme de restriction à 4 bases. L'ADN a ensuite été traité à la protéinase-k (50  $\mu$ g/ ml) pendant 30 min à 37°C dans du tampon STE (0,1 M NaCl, 20 mM Tris, 10 mM EDTA, pH 7,5) puis purifié deux fois à l'aide de phénol et de chloroforme. Après précipitation de l'ADN et resuspension dans du tampon TE fait avec de l'eau traité au DEPC (diethylpyrocarbonate) (le DEPC a été ajouté à raison de 0,1%, chauffé à 37°C pendant 8 h puis autoclavé). La transcription a été réalisée en utilisant un kit de transcription (Stratagene) selon les instructions du fabricant. Le nucléotide marqué qui a été incorporé était la biotine-21-UTP (Clontech) ; un inhibiteur des ribonucléases (RNaseBlockII, Stratagene) a également été ajouté dans les proportions recommandées par le fabricant. Les ARN ainsi synthétisés ont été précipités à -20°C en ajoutant à 25  $\mu$ l de mélange réactionnel : 6  $\mu$ l d'acétate ammonium 5 M, 7  $\mu$ l de Dextran T40 à 2 mg/ ml et 300  $\mu$ l d'éthanol, puis repris dans du TE fait avec de l'eau traitée au DEPC.

### **30) HYBRIDATIONS DES SONDAS ARN**

Les hybridations des sondes ARN ont été conduites pendant 16 h à 42°C dans du 5XSSPE, 50% formamide réalisé avec de l'eau traitée au DEPC. Trois lavages ont été réalisés dans du 1XSSPE à température ambiante pendant 10 min et un lavage dans du 0,1XSSPE, 0,1% SDS à 50°C pendant 30 min. Les hybridations ont alors été révélées à l'aide du kit "Photogene" de BRL selon le protocole décrit au paragraphe 17.

### **31) ELECTROPHORESE DE PROTEINES ET WESTERN BLOTS**

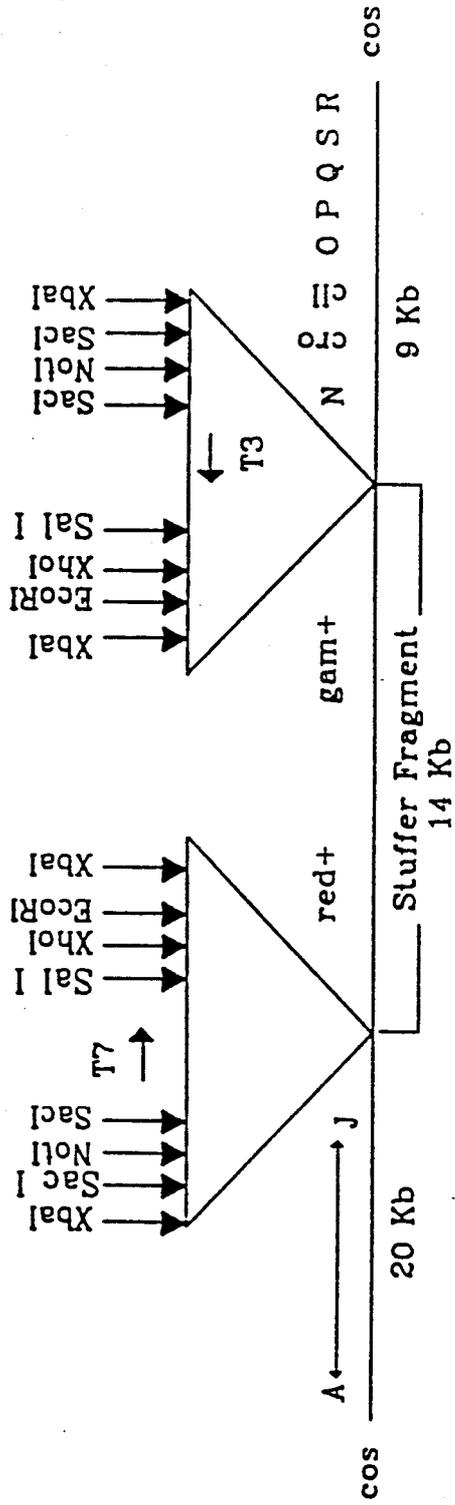
Les protéines ont été séparées sur des gels de SDS-polyacrylamide à 8% comme décrit par Sambrook et coll. (1989). Les extraits bruts de cellules d'E. coli et de B. subtilis ont été obtenus par rupture à l'aide de billes de verre, selon le protocole décrit par Ribes et coll. (1990). Après migration, les protéines ont été transférées sur une membrane Immobilon (Millipore) à l'aide d'un "Transblot-Cell" (Bio-Rad) en suivant le protocole décrit par le fournisseur.

## D) RESULTATS ET DISCUSSIONS

### 1) CONSTRUCTION DES OUTILS GENETIQUE

#### a) CONSTRUCTION D'UNE BANQUE DE *B. SUBTILIS* DANS $\lambda$ FIXII

La banque de phages  $\lambda$ FixII recombinants a été obtenue à partir d'ADN chromosomique de *B. subtilis* hydrolysé par l'enzyme de restriction Sau3A. Les conditions optimales de digestion partielle ont été déterminées par des essais de digestion (cf. Figure 16) comme étant celles de l'échantillon # 2, soit 48  $\mu$ g d'ADN chromosomique dans 80  $\mu$ l de TE, 10 unités de Sau3A, dans un volume final de 240  $\mu$ l obtenu par addition de 140  $\mu$ l de tampon 1X High (50 mM Tris, 10 mM MgSO<sub>4</sub>, 100 mM NaCl, pH 7,4). Après incubation pendant 1 h à 37°C, la réaction a été arrêtée en ajoutant 10  $\mu$ l d'EDTA 0,5 M, et en réalisant une extraction au phénol/ chloroforme suivie d'une précipitation à l'isopropanol. L'ADN chromosomique a été mis en suspension dans du TE et a ensuite été traité avec le fragment Klenow de la polymérase I et les déoxyribonucléotides dATP et dGTP afin de procéder au remplissage partiel des sites Sau3A. L'ADN chromosomique a alors été déposé sur gel d'agarose à 0,7% et le *pool* correspondant aux fragments de taille 9,5 à 23 kb a été prélevé, puis l'ADN extrait du gel d'agarose par électroélution. Le *pool* obtenu a ensuite été caractérisé comme comprenant des fragments allant de 8,5 à environ 30 kb (cf. Figure 17) et dont la majorité a une taille d'environ 20 kb. La ligation avec le vecteur  $\lambda$ FixII digéré par XhoI et dont les sites ont été partiellement remplis avec les déoxyribonucléotides dCTP et dTTP a été réalisée par Kunst (Unité de Biochimie Microbienne) selon le protocole fourni par Stratagene (*Predigested  $\lambda$ FixII vector cloning kit*). La mixture de ligation a ensuite été encapsidée par Kunst (Unité de Biochimie Microbienne) dans des enveloppes de phage  $\lambda$  selon le protocole fourni par Stratagene en utilisant le kit *Gigapack II XL* qui permet d'encapsider des phages



**Lambda FIX™ II**

**Figure 12 :** Carte du vecteur lambda Fix II  
(d'après Stratagene)

1 2 3 4 5 6



Figure 16 : Digestion partielle d'ADN chromosomique de *B. subtilis*.

1) marqueur  $\lambda$ HindIII, 2) ADN chromosomique de *B. subtilis* non hydrolysé (0,6  $\mu$ g), 3) idem hydrolysé par 0,05 unité de Sau3A pendant 1h à 37°C dans du tampon "high", 4) idem hydrolysé par 0,1 unité de Sau3A, 5) idem hydrolysé par 0,2 unité de Sau3A, 6) idem hydrolysé par 0,4 unité de Sau3A

1 2



Figure 17 : Construction d'une banque de *B. subtilis*  
1) marqueur  $\lambda$ HindIII, 2) pool d'ADN après purification

recombinants de 47 à 51 kb. Des cellules d'E. coli P2392 ont alors été infectées avec la banque ainsi obtenue afin de l'amplifier pour pouvoir obtenir un titre suffisant et permettre sa distribution aux divers laboratoires impliqués dans le projet. Le titre des 5 ml de lysat obtenu était de  $10^9$  phages/ ml.

#### b) CONSTRUCTION D'UNE BANQUE DE B. subtilis DANS UN PLASMIDE

Une banque de B. subtilis QB25 (*sacS49*, *trpC2*) avait été construite par Débarbouillé (Unité de Biochimie Microbienne) (Débarbouillé et al., 1987) dans le vecteur navette E. coli - B. subtilis pMK4 en ligaturant de l'ADN chromosomique de B. subtilis coupé par Sau3A et dont le pool correspondant aux poids moléculaires de 3 à 6 kb avait été séparé par électrophorèse en gel d'agarose puis récupéré de l'agarose par électroélution. Les fragments ainsi obtenus avaient été ligaturés au site BamHI déphosphorylé du plasmide pMK4, la souche d'E. coli TG1 avait ensuite été transformée avec le mélange de ligation et l'ADN plasmidique d'un pool de 3000 clones avait été extrait. Un échantillon de cet ADN avait ensuite été utilisé pour transformer à nouveau la souche TG1 afin d'amplifier la banque obtenue. Les clonages des gènes *thiC*, *nara* et *rodC* ont été conduits en utilisant de l'ADN plasmidique de cette banque ainsi amplifiée plusieurs fois.

#### c) CONSTRUCTION DE VECTEURS INTEGRATIFS

Les plasmides intégratifs disponibles pour l'étude de B. subtilis n'offrant pas de *polylinker* suffisamment riche en sites de restriction, il a été nécessaire de construire deux nouveaux vecteurs. Les plasmides pDIA5304 (4,3 kb, Amp<sup>R</sup>, Cm<sup>R</sup>) et pDIA5305 (3,8 kb, Amp<sup>R</sup>, Cm<sup>R</sup>) (cf. Figure 18) ont été construits à partir des plasmides pBluescript et pMTL22, respectivement. Le gène de la résistance au chloramphénicol (CAT) a été extrait du plasmide pAC5 par une digestion EcoRI-SalI ; ce gène est dérivé de pC194 et s'exprime donc chez E. coli et B. subtilis (Ballester et al., 1990). Après

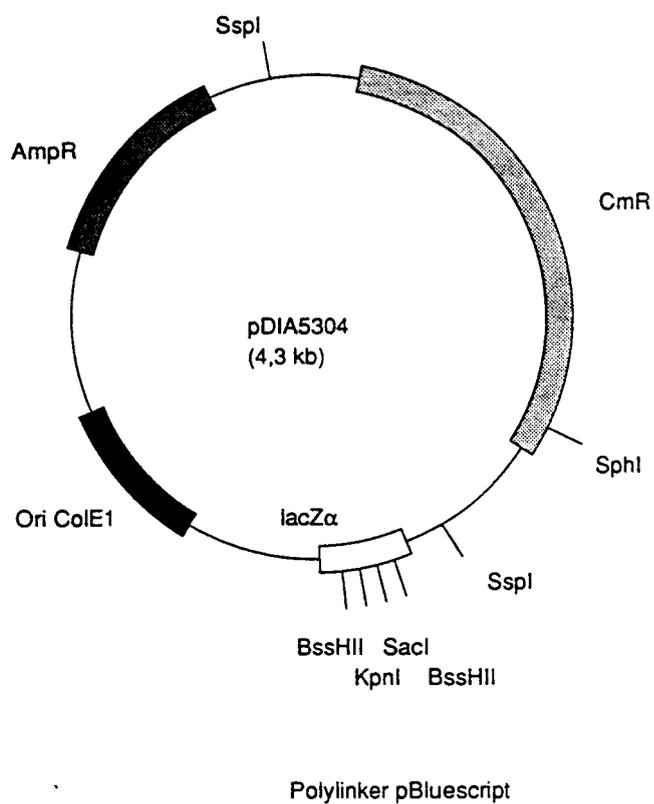


Figure 18a

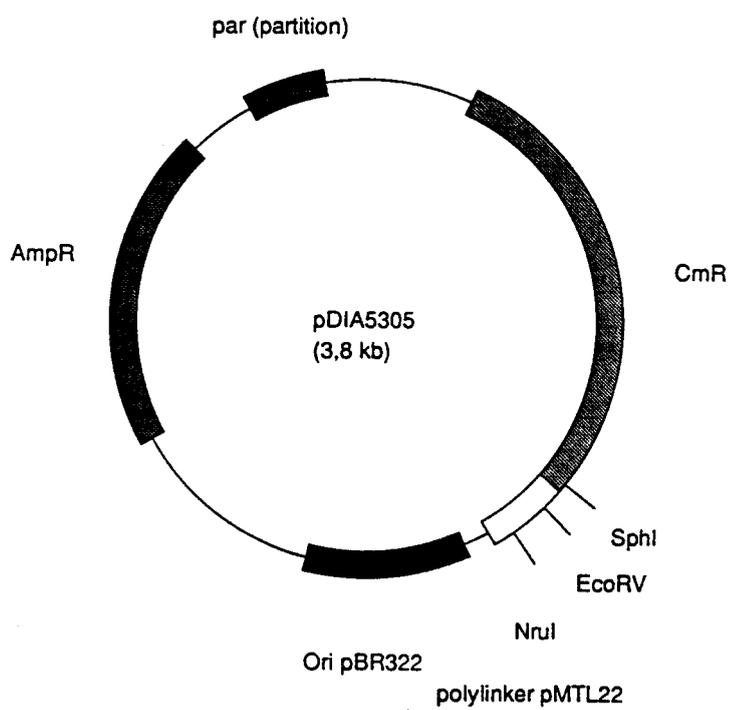


Figure 18b

séparation du gène CAT par électrophorèse en gel d'agarose, les extrémités 5' et 3' ont été remplies à l'aide du fragment Klenow de la polymérase I et des déoxyribonucléotides dATP, dCTP, dGTP, et dTTP. Le vecteur (pMTL22, pBluescript) a alors été linéarisé à l'aide de l'enzyme NaeI puis ligaturé au gène CAT ; le mélange de ligation ainsi obtenu a alors été utilisé pour transformer la souche d'E. coli TG1.

Le fragment dérivé de pAC5 qui porte le gène CAT contient un segment EcoRI-SalI de 26 bp issu de pBR322 (coordonnées 4361-4323), un segment HpaII-Sau3A de 1036 bp issu de pC194, et un fragment BamHI-SalI de 276 bp issu de pBR322 (coordonnées 375-651). Les plasmides pDIA5304 et pDIA5305 contiennent dans leur *polylinker* (cf. Annexes) une majorité de sites utiles pour l'étude de B. subtilis ; en outre, pDIA5304 a l'avantage de comprendre une copie fonctionnelle du gène *lacZ $\alpha$* .

## **2) PREPARATIONS DES SONDÉS**

L'un des préliminaires à la construction d'une collection ordonnée de phages  $\lambda$  recombinants est l'obtention de sondes adéquates qui permettent de recouvrir la région étudiée.

Les marqueurs connus de la région *gerB-sacS* sont dans l'ordre (cf. Piggot, 1989) : *sacS*, *epr*, *thiC*, *sacA*, *sacP*, *azpB*, *ahrB*, *bac*, *ebr*, *furC*, *furE*, *abrA*, *rpoE*, *rev-4*, *spoOF*, *tsr*, *ctrA*, *narA*, *glyC*, *alsR*, *divII*, *outF*, *spoIID*, *lssF*, *rodC*, *gerB* (cf. Tableau 15) parmi lesquels seuls *sacS*, *epr*, *sacA*, *sacP*, *sacT*, *rpoE*, *rev-4*, *spoOF*, *tsr*, *ctrA*, *spoIID*, *rodC*, et *gerB* ont été clonés ou séquencés.

Les gènes qui permettent la production de sondes par PCR (*rpoE*, *spoOF*, *spoIID*, *rodC*, *sacAPT*, *sacS*, *gerB*) et les gènes clonés constituent des points de départs utiles pour la construction d'une collection ordonnée de phages recombinants. Toutefois, il est clair que les sondes disponibles actuellement sont très largement insuffisantes en nombre pour espérer cloner entièrement sur des phages

## MARQUEURS GENETIQUES DE LA REGION SACS-GERB

| Gène         | Position<br>(degré) | Phénotype                                      | Références<br>(gènes clonés ou séquencés)                  |
|--------------|---------------------|--|--|
| <i>sacS</i>  | 333                 | régulation de :<br>$\beta$ -fructofuranosidase | Aymerich et Steinmetz, 1987<br>Débarbouillé et coll., 1987 |
| <i>epr</i>   |                     | protéase                                       | Sloma et coll., 1990a                                      |
| <i>thiC</i>  | 331                 | auxotrophie pour<br>la thiamine                |  |
| <i>comM</i>  | 330                 | compétence                                     |  |
| <i>sacA</i>  | 330                 | sucrase  | Fouet et coll., 1987                                       |
| <i>sacP</i>  | 330                 | enzyme II du<br>système PTS                    | Fouet et coll., 1986<br>Fouet et coll., 1987               |
| <i>sacT</i>  | 330                 | régulation <i>sacA</i>                         | Débarbouillé et coll., 1990                                |
| <i>azpB</i>  | 329                 | résistance aux<br>azopyrimidines               |  |
| <i>ahrB</i>  | 328                 | résistance à<br>l'arginine hydroxymate         |  |
| <i>bac</i>   | 327                 | bacilysine                                     |  |
| <i>ebr</i>   | 326                 | résistance au BET                              |  |
| <i>furC</i>  | 326                 | résistance au<br>5-fluorouracile               |  |
| <i>furE</i>  | 326                 | idem   |  |
| <i>abrA</i>  | 326                 | suppression de<br>phénotypes <i>spo0</i>       |  |
| <i>rpoE</i>  | 325                 | sous-unité de<br>l'ARN polymérase              | Lampe et coll., 1988                                       |
| <i>rev-4</i> | 324                 | identique à <i>abrA</i>                        | Trach et coll., 1985 ; 1988                                |

Tableau 15

| Gène          | Position<br>(degré) | Phénotype                               | Références<br>(gènes clonés ou séquencés)     |
|---------------|---------------------|---|---|
| <i>spoOF</i>  | 324                 | sporulation<br>stade 0                  | Trach et coll., 1988                          |
| <i>tsr</i>    | 324                 | synthèse d'ARN<br>thermosensible        | Trach et coll., 1988                          |
| <i>ctrA</i>   | 324                 | auxotrophie pour<br>la cytidine         | Trach et coll., 1985                          |
| <i>narA</i>   | 320                 | utilisation du<br>nitrate               |   |
| <i>glyC</i>   | 320                 | auxotrophie pour<br>la glycine          |   |
| <i>alsR</i>   | 317                 | synthèse constitutive<br>d'acétolactate |   |
| <i>divII</i>  | 317                 | division cellulaire<br>thermosensible   |   |
| <i>outF</i>   | 316                 | stade d'élongation                      |   |
| <i>spoIID</i> | 316                 | sporulation<br>stade II                 | Lopez-Diaz et coll., 1986                     |
| <i>lssF</i>   | 316                 | lyse thermosensible                     |   |
| <i>rodC</i>   | 314                 | paroi cellulaire                        | Honeyman et Stewart, 1989                     |
| <i>gerB</i>   | 314                 | germination                             | Karamata, Smith,<br>communication personnelle |

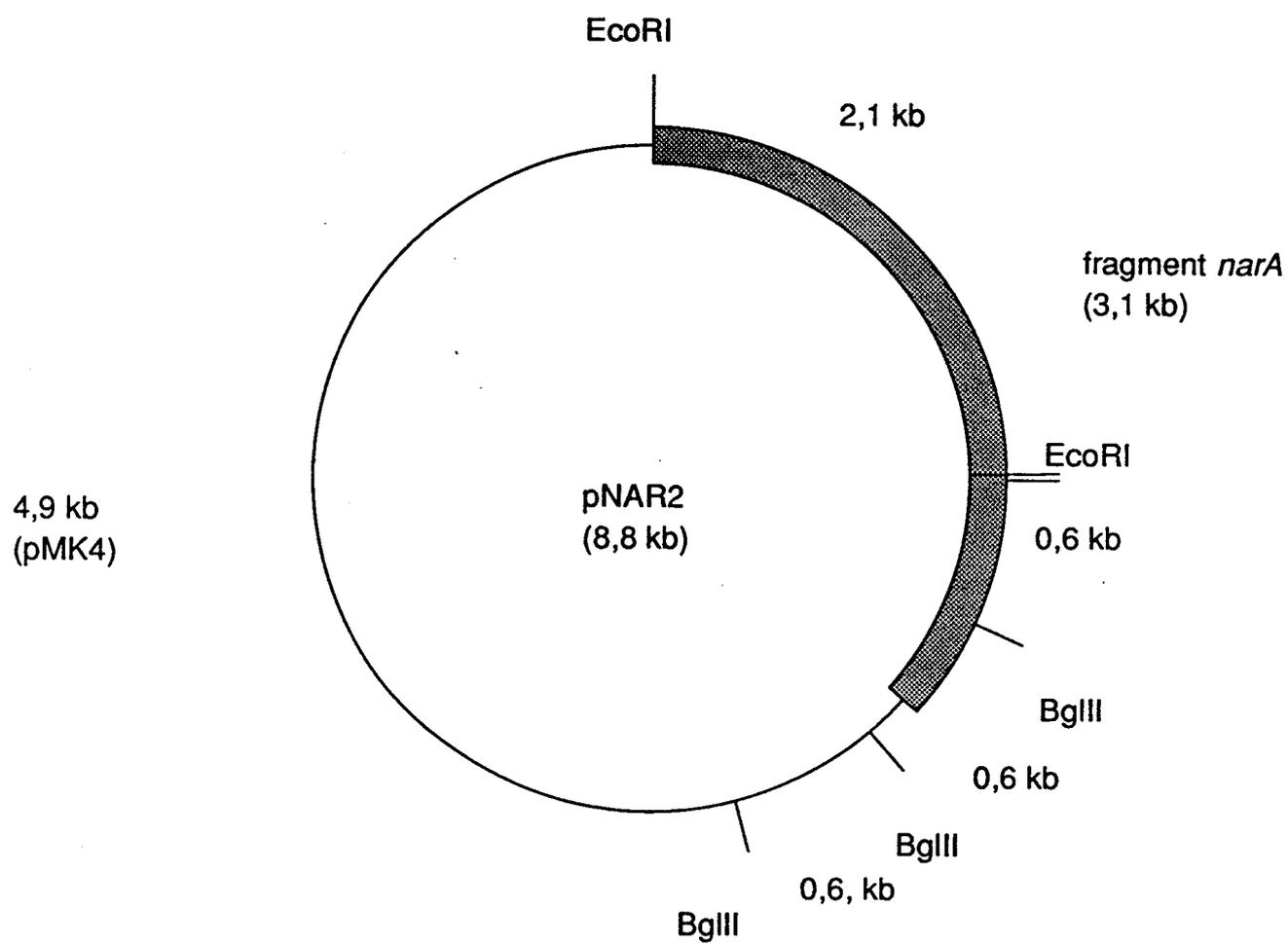
Tableau 15 (suite)

chevauchants la région allant de *sacS* à *gerB*. Il est donc nécessaire dans un premier temps de cloner d'autres loci de cette région et de mettre au point des techniques de marche ou de saut sur le chromosome. La marche sur le chromosome à partir d'un phage préalablement isolé est possible grâce à la possibilité de synthétiser des sondes ARN par l'intermédiaire des promoteurs T3 et T7 de l'ARN polymérase qui se trouvent de part et d'autre de l'insert chromosomique (cf. carte de  $\lambda$ FixII). Il est également possible de cloner sur un plasmide une extrémité du phage considéré, à condition que la carte de restriction de ce phage soit connue. Les techniques de saut sur le chromosome ont été décrites précédemment (cf. Figure 13) et ont été utilisées dans ce projet par Schneider et Santana (Unité de Régulation de l'Expression Génétique) afin de progresser sur le chromosome à partir du locus *sacA*. Une autre alternative consiste à cloner de nouveaux loci à l'aide des techniques classiques de génétique.

#### a) CLONAGE PAR MOYENS GENETIQUES

##### *narA*

Des cellules compétentes de la souche de *B. subtilis* QB692 (*sacA21*, *narA1*) ont été transformées à l'aide de l'ADN de la banque de *B. subtilis* QB25 dans le plasmide pMK4 et ont été sélectionnées pour le phénotype *narA*<sup>+</sup> en étalant les cellules transformées sur des boîtes MMnitrate (44 mM KH<sub>2</sub>PO<sub>4</sub>, 60 mM K<sub>2</sub>HPO<sub>4</sub>, 2,9 mM trisodium citrate, 20 mM KNO<sub>3</sub>, 0,1% glucose, 1,6 mM MgSO<sub>4</sub>, 4 ml/l d'une solution de citrate de fer ammoniacal à 2,2 mg/ml). Du milieu PAB (9 ml) supplémenté de 0,2% Yeast Extract a été ajouté à chaque boîte de Pétri et les cellules présentes ont été récupérées en grattant la surface de la gélose à l'aide d'un rateau stérile. La suspension de cellules a alors été supplémentée de chloramphénicol puis mise à incuber à 37°C jusqu'à ce qu'une densité optique à 600 nm de 1 à 1,5 ait été atteinte. L'ADN plasmidique de ce pool de cellules a été extrait, puis utilisé pour retransformer des cellules compétentes de *B.*



**Figure 19**

subtilis QB692, et les étapes décrites précédemment ont été répétées. L'ADN plasmidique obtenu à la fin du deuxième cycle de sélection a finalement été utilisé pour transformer des cellules compétentes d'E. coli TG1, la sélection des transformants ayant été réalisée en supplémentant les milieux de culture de 100 µg/ ml d'ampicilline. L'ADN plasmidique de quelques clones choisis au hasard a alors été extrait et utilisé pour retransformer des cellules de QB692. Une cinquantaine de colonies isolées ainsi obtenues ont alors été repiquées sur du milieu SP, incubée à 37°C pendant toute une nuit, puis repiquées à nouveau sur du milieu SP supplémenté de chloramphénicol et sur du milieu MMnitrate. L'obtention de clones *narA*<sup>+</sup>/Cm<sup>s</sup> permet de témoigner de l'intégration du fragment cloné au locus *narA* par l'intermédiaire d'un évènement de double recombinaison ; ce qui permet de vérifier que la complémentation observée est bien due à l'allèle sauvage du gène *narA* et non à un quelconque gène régulateur qui confère le phénotype *narA*<sup>+</sup> lorsqu'il est présent en multicopies. Deux plasmides différents, pNAR2 et pNAR4, ont ainsi été isolés qui confèrent le phénotype *narA*<sup>+</sup> à la souche QB692. Le plasmide pNAR2 est le plasmide qui a été choisi pour la suite de cette étude. La carte de restriction de ce plasmide est indiquée Figure 19. On peut en particulier noter que l'insert d'ADN chromosomique de ce plasmide est de 3,1 kb.

### thiC

Le clonage du gène *thiC* de B. subtilis a été réalisé selon le même schéma que celui établi pour cloner le gène *narA*. Le milieu sélectif utilisé était le milieu SP. La majeure partie de ce clonage a été effectuée par Kunst et Rapoport (Unité de Biochimie Microbienne). Un plasmide, p*thiC*, a été isolé qui confère à la souche de B. subtilis QB899 (*thiC*) le phénotype *thiC*<sup>+</sup>, l'insert chromosomique de ce plasmide a été caractérisé comme étant de 3 kb.

1 2 3 4 5

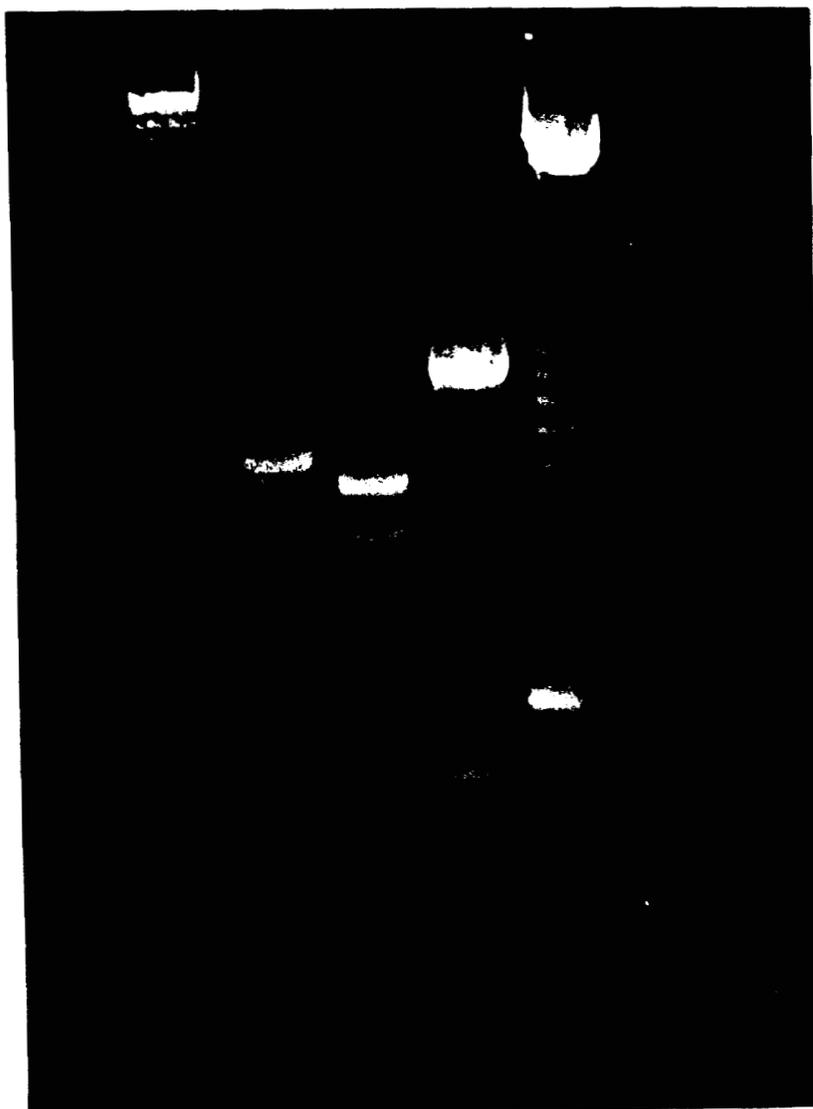


Figure 20 : Amplification de spo0E par PCR.

1) marqueur lambdaHindIII, 2) fragment amplifié hydrolysé par StuI, 3) fragment amplifié hydrolysé par SspI, 4) fragment amplifié non hydrolysé, 5) échelle de 123 bp

rodC

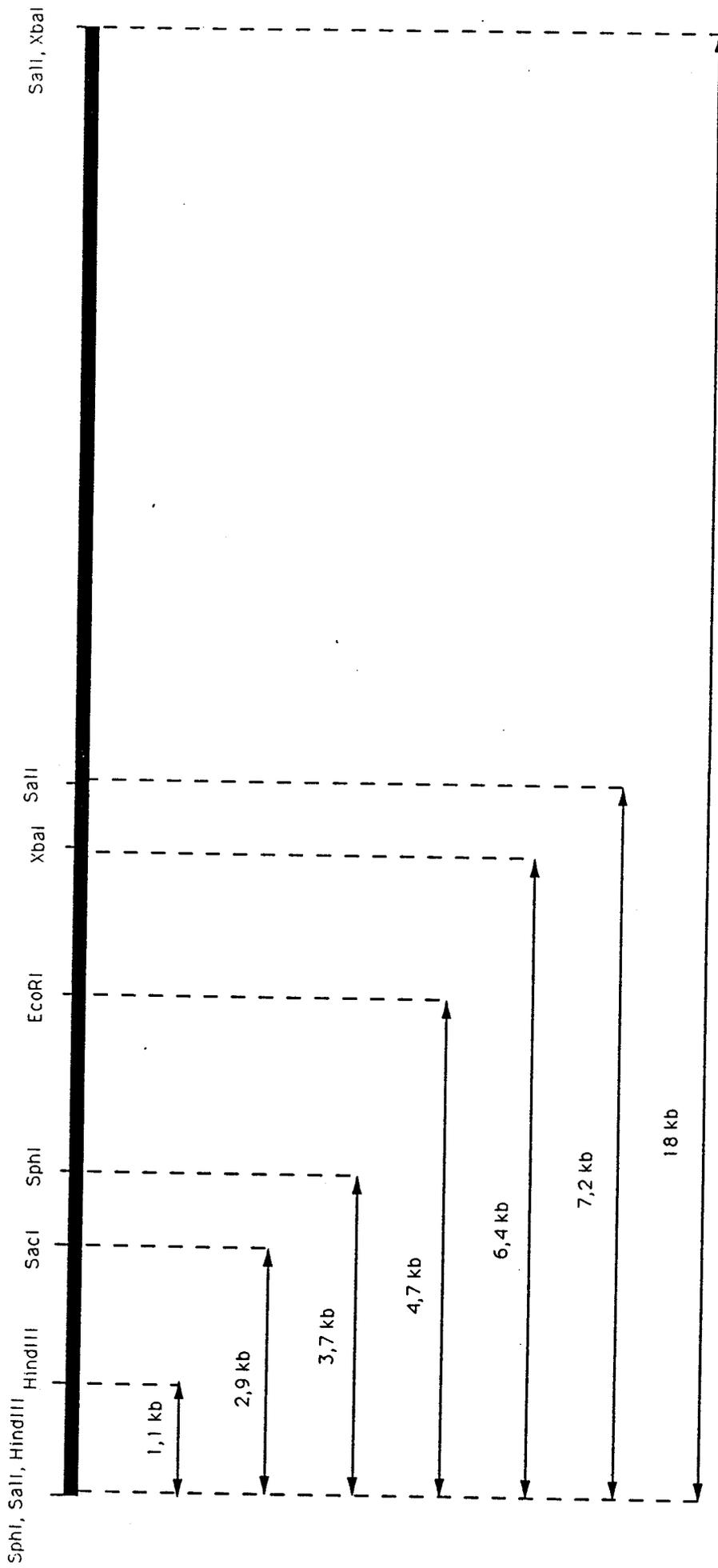
La souche BGSC 1A486 (leuA8, rodC1) qui montre un phénotype thermosensible a été utilisée pour cloner le locus *rodC*. Des cellules compétentes de cette souche ont été transformées et traitées selon le protocole décrit pour le clonage du gène *narA*. Les conditions de sélection relatives à l'identification des clones *rodC*<sup>+</sup> étaient déterminées par une croissance à 45°C. Les plasmides obtenus n'ont cependant pas été utilisés vu que la séquence nucléotidique de ce gène a été déterminée entre-temps par Honeyman et Stewart (1989).

b) PREPARATION DES SONDAS PAR PCR

L'ADN chromosomique de la souche de *B. subtilis* QB4332 (souche 168 dont le gène *yrs* a été interrompu avec une cassette kanamycine, voir paragraphe 7) a été utilisé comme matrice de PCR.

Un fragment de 1 kb de la région codante de *spo0E* séquencé par Perego et Hoch (1987) a été synthétisé par PCR à partir des oligonucléotides : 5'CTTCACCTGGGTATTGTTCTTCTAATCC<sup>3'</sup> et 5'GGCACAGAGGTGACGATCACCTCCCCG<sup>3'</sup>. Les produits de PCR ont été vérifiés comme appartenant à la séquence codante de *spo0E* par hydrolyse à l'aide des enzymes de restriction *StuI* et *SspI* qui ont fourni des fragments de 333 et 675 paires de bases, et 579 et 429 paires de bases, respectivement (cf. Figure 20).

Un fragment de 1,4 kb de la région séquencée par Lopez-Diaz et coll. (1986) a été amplifiée à partir des oligonucléotides suivants : 5'GCCGCTCTGGGCGCAGACATCG<sup>3'</sup> et 5'GACGGTCAGAAGCTCCTCTAAACCG<sup>3'</sup>. Le fragment de 1,4 kb obtenu a été vérifié comme correspondant à la région choisie de *spoIID* par hydrolyse à l'aide des enzymes de restriction *PvuII* et *HindIII* qui ont fourni des fragments de 279 et 1122 paires de bases, 860, 113 et 428 paires de bases, respectivement.



Carte du phage  $\lambda$  sacPT

Figure 21

### c) CLONAGE D'UN FRAGMENT TERMINAL DU PHAGE $\lambda$ SACPT

Un fragment de 6 kb du phage recombinant  $\lambda$ sacPT (cf. Figure 21) comprenant les loci *sacP* et *sacT*, isolé par Kunst (Unité de Biochimie Microbienne) a été cloné (pSal7sacP1) dans pSU21 par Schneider (Unité de Régulation de l'Expression Génétique). L'établissement d'une carte de restriction de ce fragment (cf. Figure 21) a permis l'identification d'un fragment HindIII de 1 kb qui permet de marcher sur le chromosome du locus *sacT* vers le locus *sacS*. Le fragment HindIII de 1 kb ainsi identifié a été cloné par Glaser (Unité de Régulation de l'Expression Génétique) dans pJH101-Not1.

### d) PREPARATIONS DES SONDAS ARN POUR LE PHAGE $\lambda$ narA4.0

L'ADN du phage  $\lambda$ narA4.0 a été hydrolysé par les enzymes de restriction RsaI ou AluI (Sau3A ne convient pas étant donné qu'il existe sur les phages recombinants de la banque de B. subtilis des sites Sau3A entre l'insert d'ADN chromosomique et les promoteurs des ARN polymérases T3 et T7). Les ARN obtenus ont été analysés sur gels d'agarose à 1% et la taille de l'ARN obtenu à partir de l'ADN hydrolysé par AluI a été estimé par comparaison avec l'échelle d'ARN 0,24-9 kb (BRL, Gaithersbrug, Maryland) comme étant d'environ 1,5 kb.

## 3) CLONAGE DE GRANDS FRAGMENTS DU CHROMOSOME DE B. SUBTILIS

### a) $\lambda$ narA

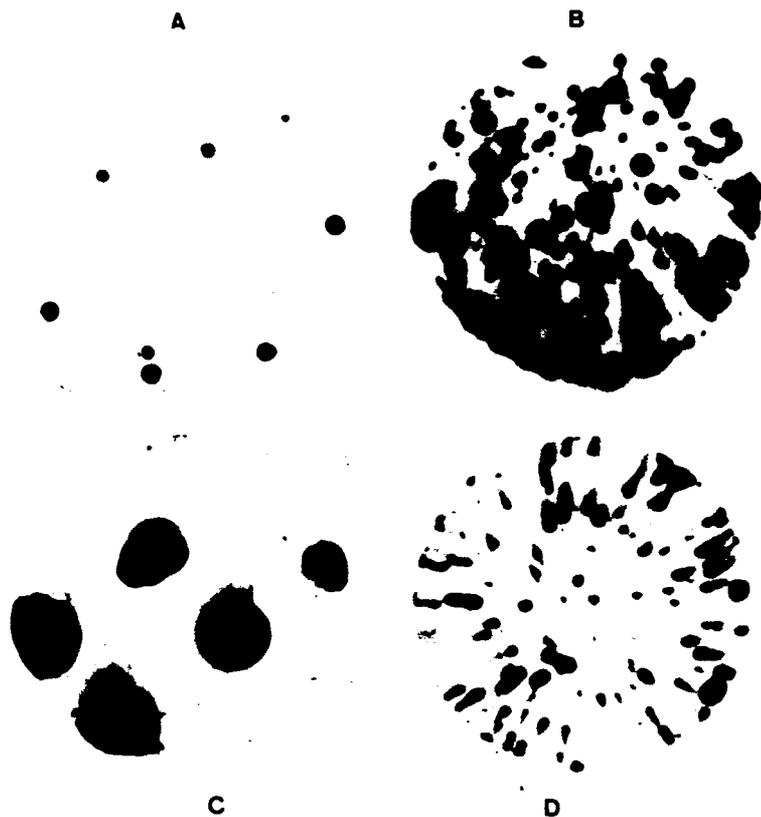
Les clonages des différents phages recombinants ont été réalisés selon le même protocole. La banque de B. subtilis dans le phage  $\lambda$ FixII a été étalée de façon à obtenir environ 10000 phages/ boîte, ce qui d'après la formule de Clarke et Carbon (1976) devrait fournir environ 7 phages recombinants positifs pour une sonde donnée et par boîte testée :

$$N = \frac{\ln(1-p)}{\ln(1-f)}$$

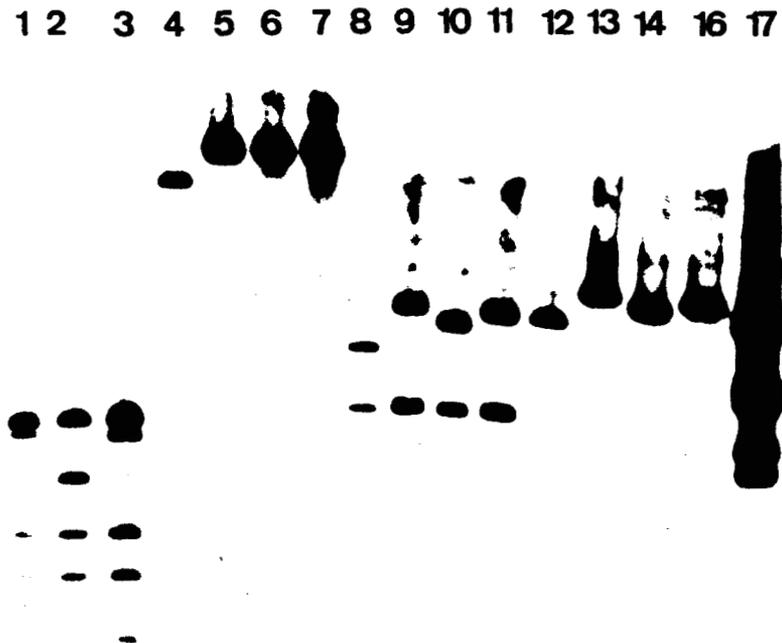


avec  $p$  représentant la probabilité de trouver une séquence unique d'ADN dans une collection de  $N$  phages recombinants contenant chacun un insert d'ADN chromosomique représentant une fraction  $f$  du génome total. En fixant la valeur de  $p$  à  $p = 0,99$  et la valeur de  $f$  à  $f = 15/4250$ , il vient que  $N \approx 1300$  phages. Chaque criblage a été réalisé en utilisant 10 boîtes différentes. Les sondes ont été purifiées sur gel d'agarose et extraites du gel à l'aide du kit Gene Clean (Bio 101, La Jolla, Californie). A l'issue du premier criblage, les plages hybridant avec la sonde ont été prélevées par carottage d'un fragment de gélose d'environ 5 mm de diamètre qui a ensuite été mis à éluer dans du tampon SM et soumis à 3 autres étapes de purification. La deuxième purification a consisté à étaler les plages prélevées de façon à obtenir environ 1000 phages par boîte ; les plages positives ont alors été prélevées à l'aide d'une pipette Pasteur par carottage d'un fragment de gélose d'environ 1 à 2 mm de diamètre, puis soumise à un troisième criblage qui consistait à hybrider des plages d'environ 10 mm de diamètre obtenues en déposant 5  $\mu$ l d'une suspension de phages sur des boîtes  $\lambda$  ensemencées avec la souche P2392 d'E. coli. La dernière étape de purification consistait à étaler les phages positifs ainsi identifiés de façon à obtenir environ 100 phages par boîte ; les phages positifs formant des plages isolées étaient alors amplifiés afin d'obtenir un titre suffisamment élevé pour constituer des stocks. Des autoradiographies typiques pour chacune de ces étapes sont présentées Figure 22.

Le plasmide pNAR2 a été marqué au  $^{32}\text{P}$  par translation de césure à raison de  $2,59 \cdot 10^7$  cpm/  $\mu$ g d'ADN pour les deux premiers criblages, et  $1,3 \cdot 10^7$  cpm/  $\mu$ g d'ADN pour les deux derniers. La concentration de sonde utilisée était de 300000

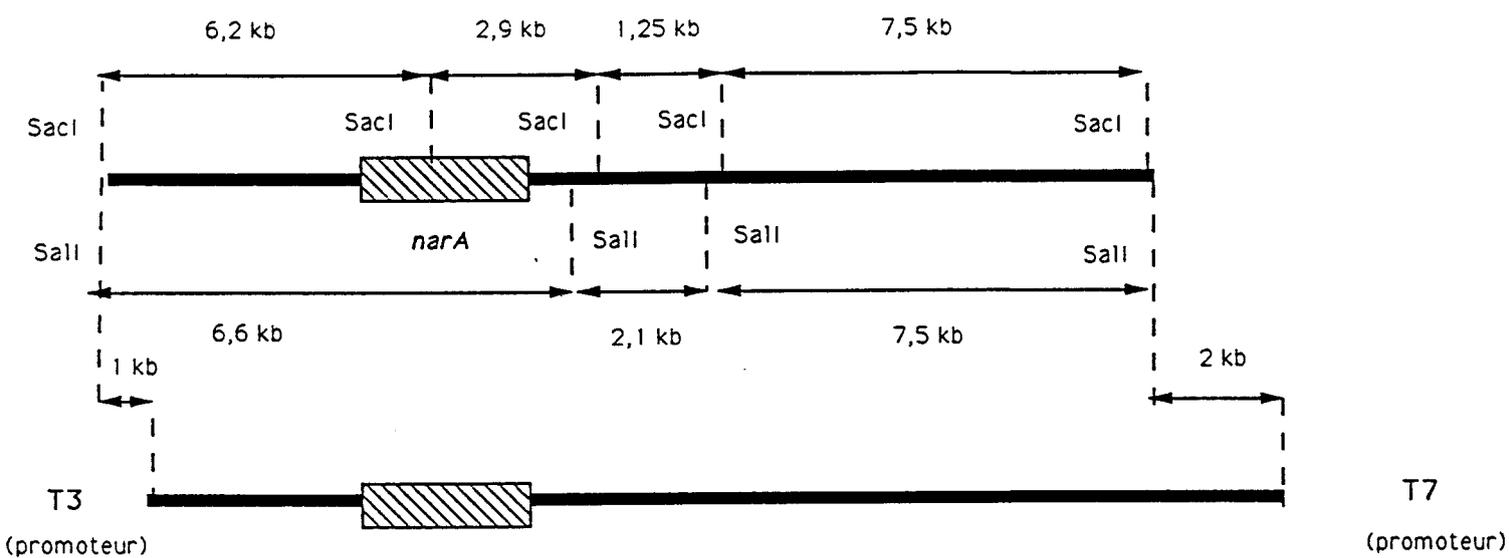


**Figure 22** : Criblage d'une banque de phages  
A) 1<sup>er</sup> criblage, B) 2<sup>ième</sup> criblage, C) 3<sup>ième</sup> criblage,  
D) 4<sup>ième</sup> criblage



**Figure 23** : Hybridation des phages *lambdanarA* avec le plasmide pNAR2

- 1) *lambdanarA3* hydrolysé par *SalI* et *EcoRI*,
- 2) *lambdanarA2.0* hydrolysé par *SalI* et *EcoRI*,
- 3) *lambdanarA1.0* hydrolysé par *SalI* et *EcoRI*,
- 4) *lambdanarA4.0* hydrolysé par *BglII*, 5) *lambdanarA3* hydrolysé par *BglII*, 6) *lambdanarA2.0* hydrolysé par *BglII*,
- 7) *lambdanarA1.0* hydrolysé par *BglII*, 8) *lambdanarA4.0* hydrolysé par *SacI*, 9) *lambdanarA3* hydrolysé par *SacI*,
- 10) *lambdanarA2.0* hydrolysé par *SacI*, 11) *lambdanarA1.0* hydrolysé par *SacI*, 12) *lambdanarA4.0* hydrolysé par *SalI*,
- 13) *lambdanarA3* hydrolysé par *SalI*, 14) *lambdanarA2.0* hydrolysé par *SalI*, 16) *lambdanarA1.0* hydrolysé par *SalI*,
- 17) marqueur de poids moléculaires réalisé par hydrolyse de pMK4 : 5 bandes de 1,9 kb, 2,3 kb, 3,3 kb, 3,5 kb, et 5,8 kb



Carte des phages  $\lambda narA1.0$  (haut) et  $\lambda narA4.0$  (bas)

Figure 24

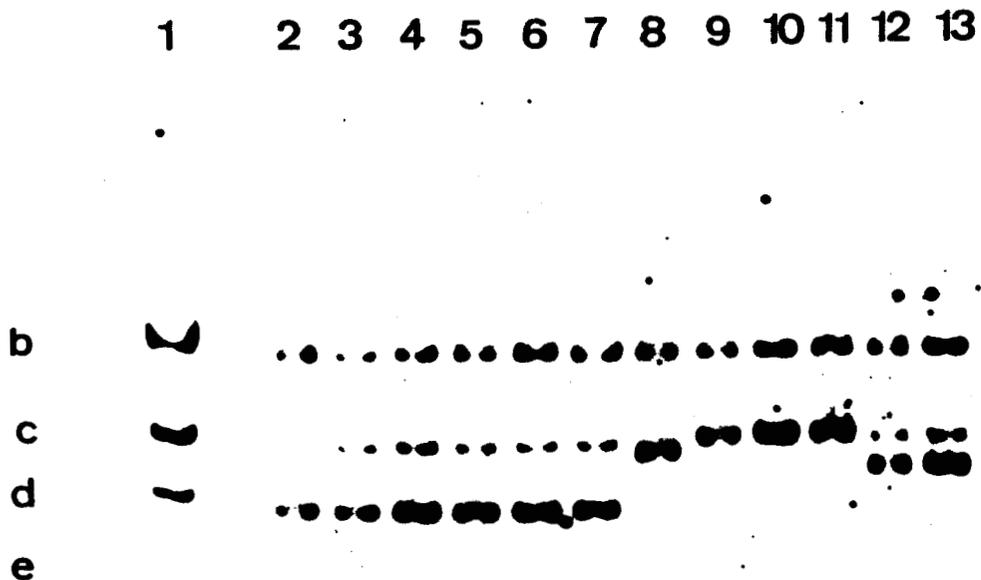
cpm/ ml de tampon d'hybridation. A l'issue du premier criblage, 56 signaux potentiels ont été identifiés et 4 phages différents ont été purifiés :  $\lambda$ narA1.0,  $\lambda$ narA2.0,  $\lambda$ narA3,  $\lambda$ narA4.0, parmi lesquels les phages  $\lambda$ narA1.0,  $\lambda$ narA2.0 et  $\lambda$ narA3 sont particulièrement chevauchants. Comme indiqué Figure 23, les 4 phages isolés sont tous différents bien que relativement semblables, ce qui suggère que les inserts chromosomiques obtenus n'ont pas subi de remaniement pendant ou après l'étape de ligation. Une carte de restriction des deux phages recombinants ayant la plus petite région de chevauchement est présentée Figure 24. Les promoteurs T3 et T7 du phage  $\lambda$ narA4.0 ont pu être placés relativement à cette carte d'après les résultats d'une expérience d'hybridation entre les différents phages  $\lambda$ narA et une sonde ARN obtenue par transcription à partir de l'ADN du phage  $\lambda$ narA4.0 hydrolysé par l'enzyme AluI. On peut également noter que le gène *narA* est situé sur une extrémité dans les 4 phages  $\lambda$ narA, ce qui suggère la présence sur le chromosome de *B. subtilis* d'une région plus difficile à cloner aux alentours du locus *narA*. En outre, ce résultat donne une idée de la représentativité de la banque étant donné que tous les phages obtenus sont différents les uns des autres.

#### b) $\lambda$ spo0E

La sonde de *spo0E* obtenue par PCR a été marquée (de 0,5 à 1  $\mu$ g d'ADN) à la biotine par translation de césure. Les hybridations ont été révélées à l'aide du kit Photogene de BRL (Gaithersburg, Maryland). Un total de 20 plages positives ont été identifiées dont 4 ont été purifiées. Ces phages se situent à l'extérieur de la région *sacS-gerB* et ont donc été transmis à Devine (Trinity College, Dublin, Irlande) pour être séquencés.

#### c) $\lambda$ spoIID

La sonde de *spoIID* obtenue par PCR a été marquée (de 0,5 à 1  $\mu$ g) à la biotine par translation de césure ; les



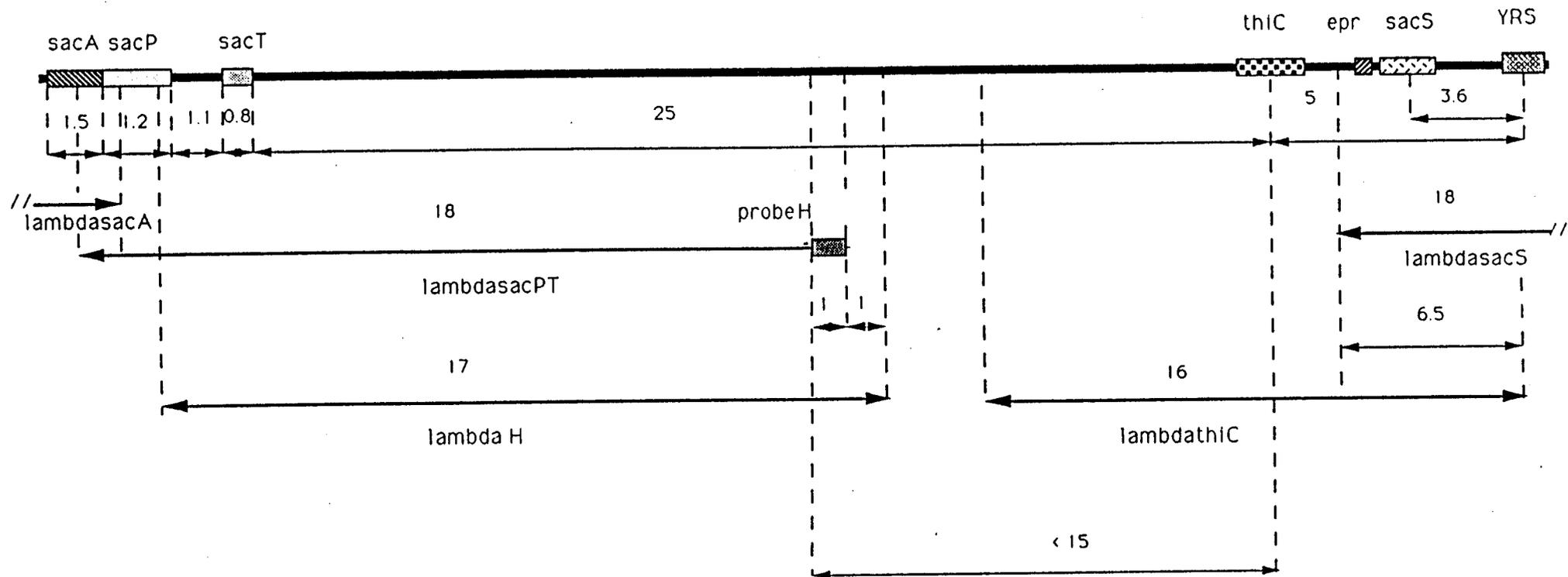
**Figure 27** : Hybridation des phages lambda spoIID avec le fragment PCR spoIID

1) marqueur lambdaHindIII, 2) lambda spoIID3 hydrolysé par SalI, 3) lambda spoIID hydrolysé par SalI, 4) lambda spoIID10 hydrolysé par SalI, 5) lambda spoIID14 hydrolysé par SalI, 6) lambda spoIID17 hydrolysé par SalI, 7) lambda spoIID hydrolysé par SalI, 8) lambda spoIID3 hydrolysé par SacI, 9) lambda spoIID6 hydrolysé par SacI, 10) lambda spoIID10 hydrolysé par SacI, 11) lambda spoIID14 hydrolysé par SacI, 12) lambda spoIID17 hydrolysé par SacI, 13) lambda spoIID19 hydrolysé par SacI

hybridations ont été révélées à l'aide du kit Photogene de BRL. Un total de 20 plages positives ont été identifiées dont 6 ont été purifiées. L'hybridation de ces 6 phages avec la sonde *spoIID* (cf. Figure 27) suggère qu'ils sont très semblables, bien que différents.

#### d) $\lambda$ H

La première région étudiée dans ce projet a été la région située entre les loci *sacS* et *sacA*. Les phages comprenant les loci *sacS*, *thiC*, *sacPT* et *sacA* ont été clonés (Kunst, Unité de Biochimie Microbienne), et les chevauchements entre les phages  $\lambda$ sacS et  $\lambda$ thiC d'une part, et  $\lambda$ sacA et  $\lambda$ sacPT d'autre part ont été mis en évidence. Une expérience d'hybridation entre la sonde provenant de l'extrémité la plus proche de *thiC* du phage  $\lambda$ sacPT, avec le phage  $\lambda$ thiC a par ailleurs permis de démontrer que les phages  $\lambda$ thiC et  $\lambda$ sacPT ne sont pas chevauchant. Afin de cloner le fragment d'ADN chromosomique qui sépare  $\lambda$ thiC et  $\lambda$ sacPT, le fragment HindIII de 1,1 kb provenant de l'extrémité du phage  $\lambda$ sacPT (sonde "H", voir Figure 25) a été marqué au  $^{32}\text{P}$  à raison de  $9,6 \cdot 10^6$  cpm/  $\mu\text{g}$  d'ADN. Un total de 22 plages positives ont été identifiées dont 8 ont été purifiées. L'hybridation avec les différents  $\lambda$ H du phage  $\lambda$ thiC marqué à la biotine (Southern blot d'un gel d'agarose contenant les 8  $\lambda$ H hydrolysés par SalI et SacI) a cependant permis de mettre en évidence le fait qu'aucun des 8 phages  $\lambda$ H ne comporte une région commune avec le phage  $\lambda$ thiC. Une deuxième hybridation entre les différents  $\lambda$ H et le fragment HindIII de 1,1 kb (sonde "H") révèle en outre que les 8 phages isolés sont particulièrement semblables et que tous chevauchent très largement le phage  $\lambda$ sacPT. En effet, la sonde "H" apparaît située sur un fragment SalI de 7 kb et sur des fragments SacI de 3,0 à 4,7 kb. Ces résultats ainsi que la carte de restriction du phage  $\lambda$ sacPT (Figure 21) indiquent donc que les 8 phages  $\lambda$ H isolés n'ont permis de marcher sur le chromosome que de  $4,7 - 2,9 = 1,8$  kb. Il semble donc que la



**Figure 25** : Cartographie de la région *sacS-sacA*

Les distances sont données en kb

région située entre *thiC* et *sacPT* soit plus difficile à cloner dans le phage  $\lambda$  que les régions alentours, ce qui suggère qu'une autre approche de clonage, telle que le clonage dans un plasmide intégratif de *B. subtilis* ou dans le phage  $\Phi$ 105, doive être envisagée. Il apparaît également nécessaire d'établir une carte physique précise de cette région afin de sélectionner la stratégie la plus adéquate.

#### e) CLONAGE D'UNE PROTEINE BIOTINYLEE

L'inconvénient du kit Photogene développé par BRL est que les hybridations de criblage de la banque réalisées avec des séquences rares, voire absentes de la banque, conduit à l'isolement de phages recombinants dont les plages de lyse révèlent un signal positif avec différentes sondes (quoique les signaux obtenus soient plus faibles et plus flous que ceux des "vrais" positifs) mais dont l'ADN (après hydrolyse et transfert par la méthode de Southern) ne présente pas d'homologie avec les sondes utilisées. Cette observation a donc motivé une étude approfondie des phénomènes moléculaires qui ont lieu lors des expériences d'hybridation. De prime abord, il est particulièrement remarquable que les motifs de restriction de tous ces phages sont singulièrement semblables (cf. Figure 26), ce qui suggère que les phages recombinants isolés comportent un gène commun dont le produit réagit avec les composants utilisés pour la révélation des hybridations. Il était donc envisageable que soit le gène de la biotine, soit le gène de la phosphatase alcaline, soit l'équivalent chez *B. subtilis* du gène de la streptavidine ou le gène d'une autre protéine biotinylée avait été cloné.

En collaboration avec Ribes (Unité de Régulation de l'Expression Génétique), un gel de protéines en conditions dénaturantes (SDS-polyacrylamide à 8%) a donc été réalisé afin de déterminer la taille de la protéine qui produit le signal détecté par autoradiographie. Des extraits protéiques totaux (30  $\mu$ l) de lysats d'*E. coli* P2392 infecté par 2 phages recombinants ainsi isolés ( $\lambda$ T15,  $\lambda$ T22) ont donc été

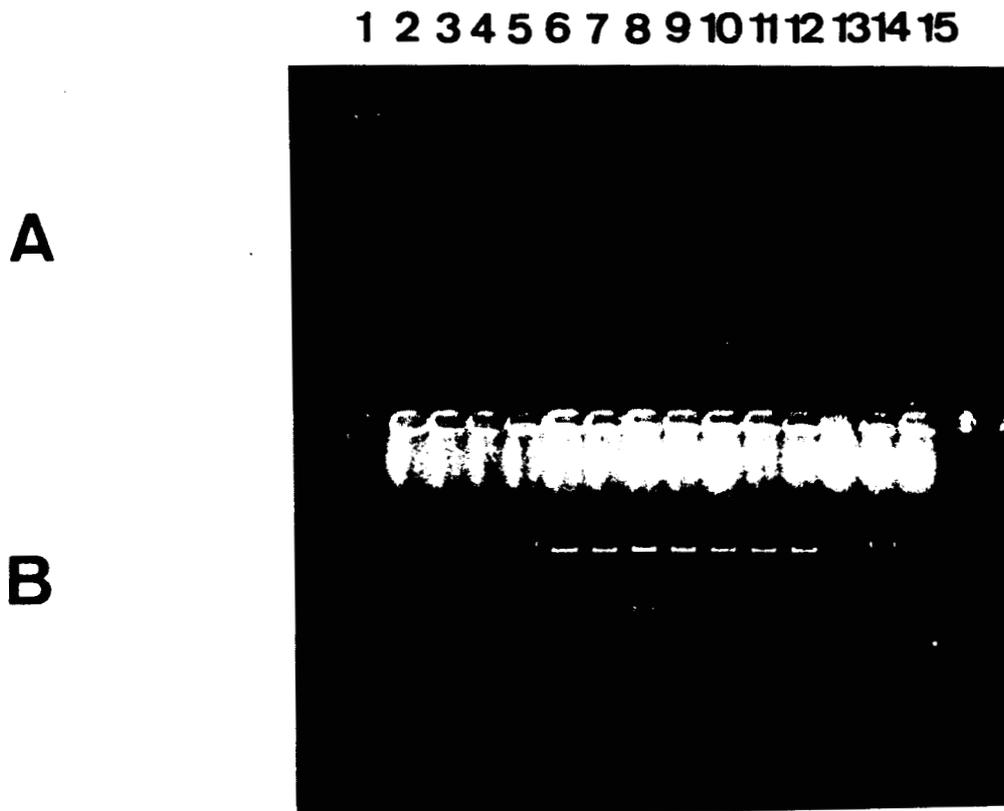
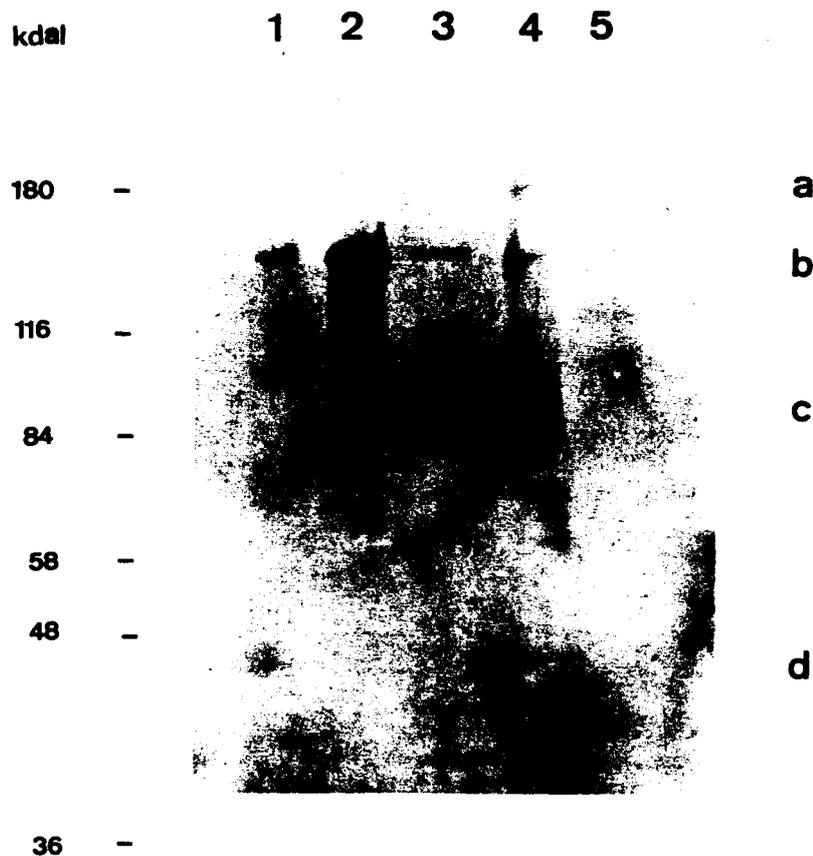


Figure 26 : Clonage d'une protéine biotinylée  
 (A) 1) marqueur  $\lambda$ HindIII, 2) à 8)  $\lambda$ SF3 à  $\lambda$ SF10 hydrolysé par SacI, 9) à 15)  $\lambda$ SF3 à  $\lambda$ SF10 hydrolysé par BglII  
 (B) 5) marqueur  $\lambda$ HindIII, 6) à 12)  $\lambda$ SF3 à  $\lambda$ SF10 hydrolysé par SacI et BglII  
 14) marqueur  $\lambda$ HindIII

analysés, ainsi qu'un extrait protéique total (équivalent de 3 unités de D0 à 600 nm) d'un lysat d'E. coli FF283 et de B. subtilis 168, et d'E. coli P2392 lysé par un phage témoin  $\lambda$ spoIID (témoin négatif). Après transfert des protéines par *Western blot* sur une membrane d'Immobilon, la membrane a été incubée pendant 10 min en présence du conjugué SA-AP, puis traitée à partir de cette étape dans les mêmes conditions que celles décrites pour l'ADN ou l'ARN, les signaux ainsi obtenus ont alors été détectés par autoradiographie. Une telle procédure permet d'éliminer, d'après les poids moléculaires apparents, la possibilité que le gène cloné corresponde au gène de la biotine. La possibilité qu'il s'agisse d'un gène de phosphatase a également été écartée par l'observation qu'aucun signal n'était détecté par le réactif PPD si l'étape de fixation du conjugué SA-AP était omise. Ces résultats confirment donc l'hypothèse que le gène cloné est celui d'une protéine biotinylée, ce qui est du reste en accord avec les méthodes utilisées chez d'autres organismes pour l'isolement de ce type d'enzyme (Bramwell, 1987 ; Haneji and Koide, 1989 ; Daunert et al., 1990). La possibilité que le gène cloné corresponde à un analogue chez B. subtilis de la streptavidine est de même peu probable attendu qu'il n'est pas nécessaire de fournir à ces protéines un intermédiaire biotinylé pour devenir visible après addition de PPD. De plus, il paraît peu vraisemblable que deux protéines de type avidine puissent se fixer à une même molécule de biotine, la biotine étant une petite molécule. Par contre, il est bien établi qu'une protéine qui est biotinylée de façon covalente et une molécule d'avidine peuvent être fixées simultanément à une même molécule de biotine, étant donné que la biotine contient un site d'attachement pour chacun de ces types de protéines (Samols et al., 1988 ; Korpela et al., 1984 ; Hunt et al., 1989).

Il apparaît donc (cf. Figure 28) que, excepté le lysat d'E. coli P2392 infecté par le phage  $\lambda$ spoIID17, tous les échantillons testés contiennent au moins une protéine



**Figure 28** : Clonage d'une protéine biotinylée :

Analyse des phages recombinants.

- 1) lysat de 3 unités de DO à 600 nm d'E. coli FF283,
- 2) lysat de 3 unités de DO à 600 nm de B. subtilis 168,
- 3) lysat d'E. coli P2392 infectées par le phage lambdaT15,
- 4) lysat d'E. coli P2392 infectées par le phage lambdaT22,
- 5) lysat d'E. coli P2392 infectées par le phage lambda spoIID17  
(30 µl de lysat ont été chargés dans les puits 3,4 et 5)

biotinylée qui donne un signal visible sur l'autoradiogramme. Il est donc clair à ce stade que le bruit de fond précédemment observé est dû à cette protéine exprimée par les phages recombinants qui en contiennent le gène. Il est intéressant de noter qu'une (même ?) protéine biotinylée de même taille est présente chez E. coli et B. subtilis. D'un point de vue quantitatif, la protéine de B. subtilis semble être exprimée à un niveau beaucoup plus fort que celle d'E. coli. Le fait que cette protéine ne soit pas visible dans les différents lysats du phage s'explique par la différence en quantité de matériel chargé : l'équivalent de 3 unités de DO à 600 nm a été chargé dans les puits correspondants aux extraits bruts d'E. coli et B. subtilis, tandis que 30  $\mu$ l seulement d'un lysat de phages ont été chargés dans les puits correspondants aux phages  $\lambda$ T15,  $\lambda$ T22, et  $\lambda$ spoIID17.

Au vu de la Figure 28, la bande principale b correspond à un poids moléculaire de 130 kdal, on peut encore noter la présence de bandes secondaires de 44 kdal, 80 kdal, et 170 kdal, ce qui pourrait indiquer que la protéine considérée se compose de 4 sous-unités de 44 kdal ; les autres bandes mineures pouvant être associées à des produits de dégradation. Il est également possible, par analogie avec les poids moléculaire des autres enzymes biotinylées isolées jusqu'à ce jour (Bramwell, 1987) que la protéine biotinylée dont le gène a été cloné est une protéine de 130 kdal, et que les autres bandes observées en sont des produits de dégradation. Des expériences complémentaires sont cependant nécessaires pour éclaircir ce point.

Il est en outre très vraisemblable que la protéine clonée soit une carboxylase. En effet, la famille des enzymes biotinylées connues à ce jour est composée de carboxylases (carboxylases, transcarboxylases, décarboxylases) (Moss and Lane, 1971 ; Wood and Barden, 1977 ; Schwarz et al., 1988). La biotine est ajoutée aux apoenzymes dans une réaction catalysée par la biotine holoenzyme synthétase (Lane et al., 1964a, 1964b ; Wood et al., 1980) qui lie de façon covalente

la biotine par un lien amide à une lysine spécifique (Moss and Lane, 1971). Cette réaction est du reste assez générale dans le sens que la synthétase d'un organisme peut biotinyler des apoenzymes hétérologues (cf. Murtif and Samols, 1987). Comme souligné par Bramwell (1987), un examen de la littérature suggère qu'il est envisageable d'identifier les enzymes biotinylées provenant des vertébrés d'après le poids moléculaire apparent d'une sous-unité caractéristique : les acetylcoA carboxylases ont une sous-unité caractéristique dont la taille est de l'ordre de 220 kdal, les pyruvates carboxylases de l'ordre de 130 kdal, les  $\beta$ -methylcrotonylcoA carboxylases de l'ordre de 75 kdal, et les propionylcoA carboxylases de l'ordre de 70 kdal. Il est également connu que les pyruvates carboxylases provenant des vertébrés sont organisées en tétramères d'environ 500 kdal de la sous-unité biotinylée d'environ 130 kdal (Rylatt et al., 1977). En outre, l'acetylcoA carboxylase d'E. coli est composée de 3 sous-unités de poids moléculaire 45 kdal, 130 kdal, et 98 kdal (Wood and Barden, 1977) mais l'étude de l'enzyme native est rendue difficile par sa dissociation spontanée *in vitro* (Sutton et al., 1977). Il est encore remarquable que les enzymes biotinylées provenant de microorganismes sont encore très mal connues, ce qui rend hasardeux l'extrapolation aux microorganismes des résultats obtenus chez les vertébrés, malgré la forte conservation de ces enzymes sur une large distance évolutive. Parmi les carboxylases identifiées chez les procaryotes, on peut citer la pyruvate carboxylase de B. subtilis (*pycA*), localisée en position 149 sur la carte génétique (Piggot, 1989).

#### **4) CARTOGRAPHIE DE LA REGION SACS-SACA**

L'établissement d'une carte physique de la région *sacS-sacA* est facilitée par la connaissance de la séquence nucléotidique de la région *sacA-sacPT* (Fouet et al., 1986, 1987 ; Débarbouillé et al., 1990). Les distances entre les différents *loci* ont été estimées par transformation de

cellules compétentes de B. subtilis avec de l'ADN chromosomique de B. subtilis et par extrapolation des distances physiques à partir des fréquences de cotransfert de deux marqueurs indépendants. Les cellules compétentes ont été transformées en condition saturante d'ADN transformant, c'est pourquoi la fréquence de double cross over entre deux marqueurs séparés de plus de 100 kb (*sacS* -333°- et *degU* -314°-) a été déterminée et prise en compte dans l'estimation des distances physiques. La souche de B. subtilis QB899 (*thiC*) a été transformée avec de l'ADN chromosomique de la souche QB4332 (*yrs::aphA3*) ou de la souche QB6001 (*trpC2, sacT::aphA3*). Les transformants *thiC*<sup>+</sup> ont été sélectionnés sur du milieu SP et la fréquence de cotransfert a été déterminée par repiquage des transformants ainsi obtenus sur du milieu SP et sur du milieu SP supplémenté de kanamycine. La souche QB25 (*trpC2, sacS49*) a été transformée avec de l'ADN chromosomique de la souche QB4238 (*trpC2, degS/degU::aphA3*) ou de la souche QB4332. Les transformants ont été sélectionnés sur du milieu SP supplémenté de kanamycine ; la fréquence de cotransfert a été déterminée par repiquage des transformants ainsi obtenus sur du milieu SP supplémenté de kanamycine, et par vérification du caractère constitutif de la saccharase.

Les fréquences de cotransfert sont reportées Tableau 16. Les distances physiques indiquées Figure 25 sont issues de la résolution de l'équation de Kemper (1974) à partir des fréquences de cotransfert auxquelles la fréquence de double cross over a été soustraite (% de cotransfert corrigé) afin de tenir compte des phénomènes de double transformation. La taille D des fragments transformants a été obtenue à partir de la fréquence de cotransfert des gènes *sacY* et *yrs* (tyrosine tRNA synthétase secondaire), et de la distance physique d entre ces deux marqueurs comme établi par la séquence nucléotidique du fragment *sacS-yrs* (d = 3,565 kb) ; la valeur de D utilisée (39,8 kb) est le résultat de 20 itérations de l'équation (1) (valeur initiale D<sub>0</sub> = 50 kb). Les

**CARTOGRAPHIE DE LA REGION SACS-SACA**

| marqueurs            | % cotransfert | % cotransfert corrigée | distance (kb) |
|----------------------|---------------|------------------------|---------------|
| <i>sacY/degSdegU</i> | 4             | 0                      |               |
| <i>thiC/sacT</i>     | 14            | 10                     | 23,4          |
| <i>thiC/yrs</i>      | 65            | 61                     | 5,1           |
| <i>sacY/yrs</i>      | 73            | 69                     | 3,6           |

avec D = 39,8 kb

Tableau 16

distances physiques  $d$  indiquées sont les résultats de 20 itérations de l'équation (2) (valeur initiale  $d_0 = 1$  kb).

$$(1) D_n = \frac{d(\ln d - \ln D_{n-1})}{c-1}$$

$$(2) d_n = \frac{D(c-1)}{\ln D - 1 - \ln d_{n-1}}$$

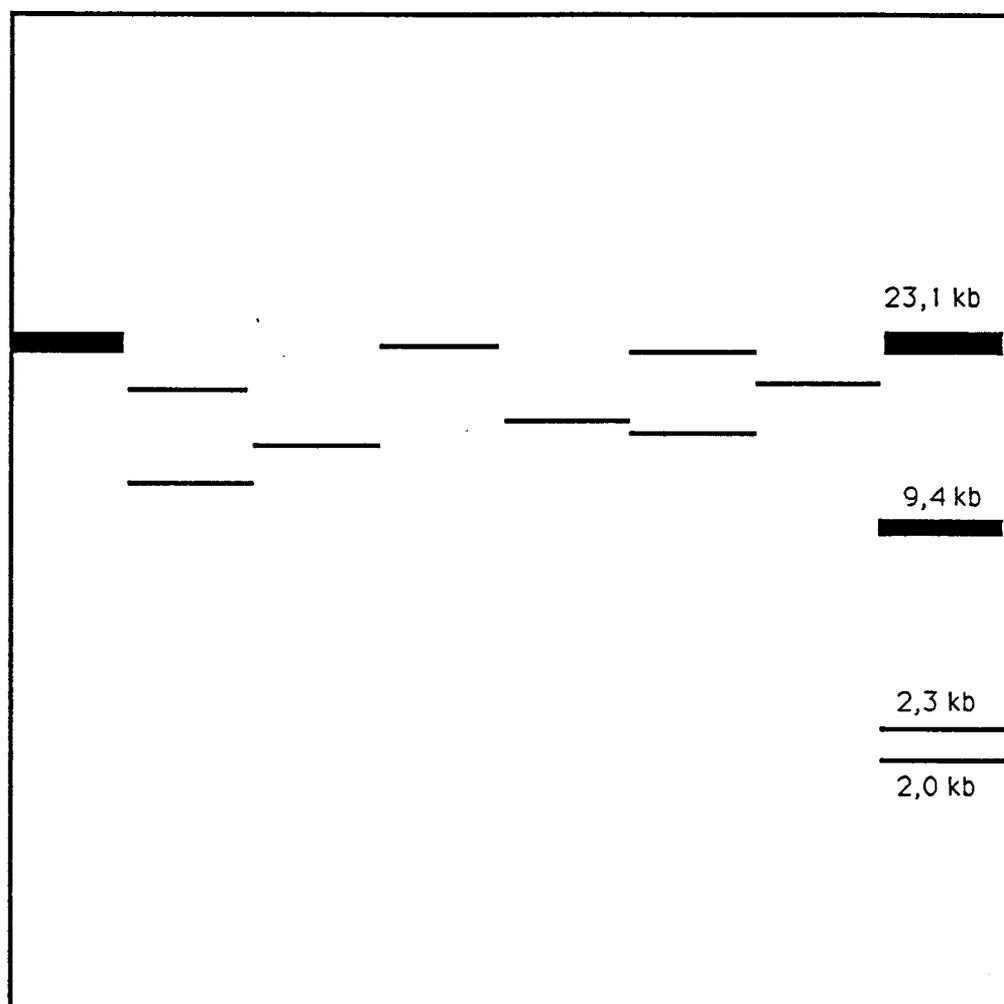
$D$  = taille moyenne des fragments d'ADN transformant

$d$  = distance physique entre les deux marqueurs

$c$  = fréquence de cotransfert

La région qui s'étend des marqueurs *thiC* à la tyrosine tRNA synthétase secondaire a en outre été caractérisée par hybridation : de l'ADN chromosomique hydrolysé par les enzymes BamHI, SalI, XbaI, SacI, BglIII, et EcoRI a été soumis à une électrophorèse en champs pulsé, puis transféré sur une membrane de nylon et hybridé successivement avec une sonde pour le gène de la thiamine et une sonde pour le fragment terminal HindIII de 1 kb du phage recombinant  $\lambda$ sacPT (cf. Figure 25 : sonde "H"). Le résultat de ces hybridations est représenté Figure 29. Le fragment le plus petit issu d'une hydrolyse totale et qui contient les deux séquences test est le fragment SacI dont la taille estimée est d'environ 13 kb ; ce qui implique que l'espace séparant les phages  $\lambda$ thiC et  $\lambda$ H est au plus égal à  $13 - 2 = 11$  kb. Il est également

$\lambda$  BamHI Sall XbaI SacI BglII EcoRI  $\lambda$  HindIII/  
EcoRI



Cartographie de la région *sacS-sacA* : Hybridation de l'ADN génomique avec le plasmide pthiC et la sonde "H"

Figure 29

intéressant de noter que la région entre le gène *thiC* et la sonde "H" comporte au moins un site BamHI et un site BglII, mais ne semble pas comporter de site Sall, SacI, EcoRI, ou XbaI, ce qui implique que ces enzymes sont adéquates pour cloner le fragment manquant dans un plasmide intégratif de B. subtilis. Une telle stratégie a été utilisée avec succès par Glaser (Unité de Régulation de l'Expression Génétique) pour cloner la région séparant le marqueur *thiC* de la sonde "H". Il a donc été possible de cloner sur divers vecteurs la totalité de la région *sacS-sacA*, l'originalité de cette réalisation vient du fait qu'il est maintenant possible d'avoir accès à l'information nucléotidique d'un grand fragment ( $\approx 60\text{kb}$ ) du chromosome de B. subtilis.

## **5) SEQUENCAGE DU PHAGE $\lambda$ narA**

### **a) CLONAGE EN SHOTGUN DU PHAGE $\lambda$ narA DANS LE PHAGE M13mp8**

L'ADN du phage  $\lambda$ narA4.0 a été soniqué ( $6 \times 5 \mu\text{g}$ ) pendant  $2 \times 20\text{s}$ ,  $3 \times 20$ , et  $4 \times 20\text{s}$  à l'aide d'un sonicateur VIBRACELL (Sonics and Materials, Danbury, Connecticut) réglé sur l'énergie maximum. Les fragments obtenus ont été séparés sur gel d'agarose à 1% (Figure 30) et 3 *pool* différents ont été prélevés comprenant des fragments de 500 à 800 paires de bases, 800 à 1000 paires de bases, et 1000 à 1500 paires de bases. L'ADN de ces 3 *pool* a été extrait de l'agarose à l'aide du kit Gene Clean (Figure 31) puis ligaturé au phage M13mp8 suivant les opérations décrites dans la section "Matériels et Méthodes". La sonde XL1Blue d'E. coli a été transformée avec la mixture de ligation ainsi obtenue selon la méthode de Hanahan, les clones pour lesquels le gène de la  $\beta$ -galactosidase a été interrompu par un insert (blancs sur du milieu de culture contenant de l'IPTG et de l'XGal) ont été repiqués sur des boîtes LB, à raison de 100 clones par boîte. Un total de  $1350 + 1300$  clones blancs ont pu être ainsi isolés. Les clones comprenant un fragment de l'insert chromosomique de B. subtilis du phage  $\lambda$ narA4.0 ont été détectés par hybridation avec une sonde de l'insert

1 2 3 4 5 6 7 8

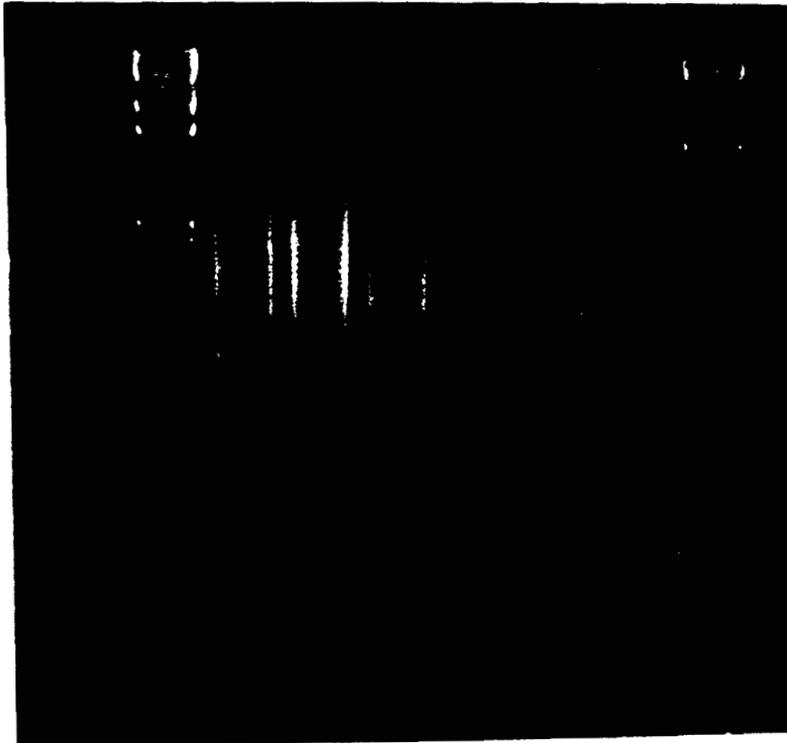


Figure 30 : Séquençage du phage lambda narA.

Sonication de l'ADN

1) marqueur lambdaHindIII, 2) et 3) ADN soniqué pendant 2 x 20 s, 4) et 5) ADN soniqué pendant 3 x 20 s, 6) et 7) ADN soniqué pendant 4 x 20 s, 8) marqueur lambdaHindIII/ EcoRI

1 2 3 4 5

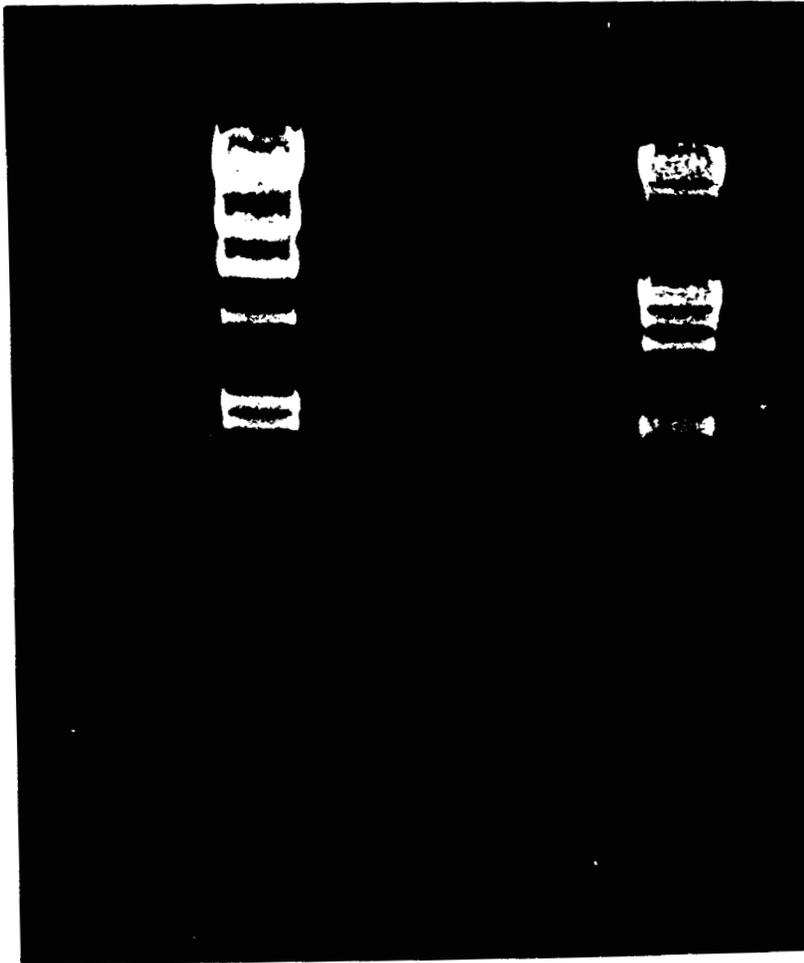


Figure 31 : Séquençage du phage  $\lambda$ danarA.

Sous-clonage dans le phage M13

1) marqueur  $\lambda$ HindIII, 2) fragments de 500 à 800 bp, 3) fragments de 800 à 1000 bp, 4) fragments de 1000 à 1500 bp,

5) marqueur  $\lambda$ HindIII/ EcoRI

chromosomique marquée au  $^{32}\text{P}$  à raison de  $10^7$  cpm/  $\mu\text{g}$  d'ADN dans une expérience concernant 1350 clones blancs. La sonde de l'insert chromosomique a été obtenue par hydrolyse du phage recombinant  $\lambda\text{narA4.0}$  par l'enzyme *SalI*, suivie d'une séparation des bandes sur gel d'agarose à 0,7% et purification de l'ADN des blocs d'agarose à l'aide du kit Gene Clean. Un total de 102 plages contenant un insert chromosomique ont ainsi été obtenues (cf. Figure 32). Par ailleurs, le nombre de plages blanches comprenant un insert de l'ADN du phage  $\lambda\text{FixII}$  ont été détectées par hybridation avec de l'ADN du phage  $\lambda\text{FixII}$  marqué au  $^{32}\text{P}$  à raison de  $9,8 \cdot 10^6$  cpm/  $\mu\text{g}$  d'ADN ; 308 clones réagissant avec la sonde du phage  $\lambda\text{FixII}$  ont ainsi été identifiées. Il est intéressant de noter que les clones contenant un insert du phage  $\lambda\text{FixII}$  sont environ 3 fois plus nombreux que ceux contenant un fragment de l'ADN chromosomique de *B. subtilis*, ce qui est en bonne corrélation avec les tailles respectives de ces deux séquences : 28 kb pour l'ADN de  $\lambda\text{FixII}$  et environ 17 kb pour l'insert chromosomique de *B. subtilis*. Le bruit de fond (plages blanches formées du vecteur M13mp8 religaturé sur lui-même mais avec un *frameshift* dans le gène de la  $\beta$ -galactosidase) apparaît également très important et doit être pris en compte afin d'estimer le nombre de plages comprenant un insert chromosomique du génome de *B. subtilis* à partir du nombre total de plages blanches. Une telle estimation a pu être obtenue avec une certaine précision en utilisant une équation empirique :  $R = 0,3Br$ , avec R représentant le nombre de phages M13mp8 contenant un insert chromosomique du gène de *B. subtilis*, B représentant le nombre total de plages blanches, r étant fixé à 0,43, 0,27 et 0,20 pour les pool d'ADN soniqué contenant des fragments de 500 à 800, 800 à 1000, 1000 à 1500 paires de bases, respectivement. Un deuxième criblage concernant 1300 clones a été réalisé en marquant environ 1  $\mu\text{g}$  d'ADN de l'insert chromosomique du phage  $\lambda\text{narA4.0}$  au  $^{32}\text{P}$  à raison de  $1,7 \cdot 10^7$  cpm/  $\mu\text{g}$  d'ADN. Un total de 231 plages hybridant avec la sonde ont ainsi été



Figure 32 : Hybridation des phages M13mp8 recombinants avec l'insert chromosomique de *B. subtilis* du phage  $\lambda$ danerA4.0

identifiées, ce qui est en accord avec la valeur attendue :  
 $0,3 \times 1300 \times 0,47 = 183$ .

#### b) SEQUENCAGE

Comme décrit dans le paragraphe "Matériels et Méthodes", le phage  $\lambda$ narA4.0 a été séquencé en gradients de tampon afin d'obtenir une meilleure répartition des bandes le long du gel. Un autre choix stratégique important a été l'utilisation systématique de 7-deaza-dGTP, analogue du dGTP dont l'avantage est de permettre de réduire les problèmes liés aux compressions. Un total de 408 clones ont été séquencés à partir de la forme simple brin du phage M13, et pour environ une centaine de ces clones, la forme répliquative double brin a également été utilisée comme matrice de séquençage. Un total de 120000 paires de bases ont été lues résultant en la formation de 55 *contig* différents. La somme de ces différents *contig* représente 34 kb, ce qui signifie que malgré l'étape d'hybridation pour cribler les phages M13 qui contiennent un insert chromosomique de B. subtilis (cf. Figure 32), de nombreux phages M13 contenant un fragment d'ADN provenant du phage  $\lambda$  ont été prélevés. Ceci peut s'expliquer par l'observation que les résultats du criblage par hybridation (Figure 32) sont relativement difficiles à interpréter, du moins dans certains cas, sans doute du fait de la présence dans l'ADN de la sonde utilisée de quelques fragments résiduels d'ADN du phage  $\lambda$ . L'ordonnancement final des différents *contig* doit s'effectuer à l'aide d'une autre stratégie, telle que l'utilisation d'oligonucléotides spécifiques comme amorces pour les réactions de séquence.

### 6) ANALYSE DES SEQUENCES

#### a) TESTS RAPIDES POUR LA DETECTION D'ENZYMES :

##### IDENTIFICATION D'UNE NOUVELLE PROTEASE

Dans le but d'identifier rapidement de nouvelles enzymes à fort potentiel biotechnologique, les phages isolés peuvent être soumis à un criblage systématique. Une telle approche a

été envisagée pour détecter la présence de protéases sur les phages  $\lambda$  recombinants purifiés ; une nouvelle protéase a ainsi pu être localisée sur le phage  $\lambda$  sacPT dont les plages montrent une forte zone de protéolyse (clarification du lait écrémé sur des boîtes  $\lambda$  supplémentées de lait écrémé et de  $\text{CaCl}_2$ ).

La séquence de cette nouvelle protéase a été réalisée par Santana (Unité de Régulation de l'Expression Génétique) et Arnaud (Unité de Biochimie Microbienne) (ORF XXIX, cf. Figure 37, paragraphe V), et le gène a été cloné par Kunst (Unité de Biochimie Microbienne). L'interruption de ce gène (réalisée par Kunst) dans une souche de B. subtilis dont les gènes des 6 protéases connues sont interrompus (Sloma et al., 1990) indique qu'il n'est pas indispensable étant donné que la souche sporule et croît normalement. La souche ainsi obtenue est donc défective pour 7 protéases différentes et devrait se révéler un hôte utile pour la production et la sécrétion de protéines hétérologues.

#### b) ANALYSE DES SEQUENCES A L'AIDE DE BANQUES PEPTIDIQUES

Les séquences brutes obtenues à partir du phage  $\lambda$  narA ont été traduites dans les 6 différentes phases de lecture et l'ensemble a été comparé aux données de la banque peptidique NBRF selon le test de Lipman et Pearson (matrice Pam250) (1985). Les résultats de ce test sont reportés Tableau 17. D'après ce test, il apparaît donc que le phage  $\lambda$  narA4.0 contient des gènes déjà connus. En particulier, il apparaît que le gène *spoIID* et le gène *narA* sont situés au voisinage l'un de l'autre sur le chromosome de B. subtilis, alors que la carte génétique (Piggot, 1989) les situe à 4° l'un de l'autre. En outre, les gènes de la sous-unité b de la sous-unité Fo et des sous-unités  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$  de la sous-unité F1 de l'ATP synthase de B. subtilis sont présents sur ce segment de chromosome. Cependant, ces gènes sont situés à une extrémité du fragment cloné, ce qui explique l'absence des gènes des sous-unités c, a et i de la sous-unité Fo. De même,

**HOMOLOGIES DETECTEES DANS LE PHAGE  
 $\lambda$ NARA**

| Gène                          | Organisme                          | Score optimisé |
|-------------------------------|------------------------------------|----------------|
| spoIID                        | <i>B. subtilis</i>                 | 924            |
| ATPase Fo b                   | <i>B. megaterium</i>               | 226            |
| ATPase F1 $\alpha$            | <i>B. megaterium</i>               | 1168           |
| ATPase F1 $\beta$             | <i>B. megaterium</i>               | 626            |
| ATPase F1 $\gamma$            | <i>B. megaterium</i>               | 346            |
| ATPase F1 $\delta$            | <i>B. megaterium</i>               | 335            |
| ATPase F1 $\epsilon$          | <i>B. megaterium</i>               | 209            |
| protéine R<br>(près de spoOF) | <i>B. subtilis</i>                 | 365            |
| 3-déhydroquinone              | <i>E. coli</i>                     | 308            |
| uréase                        | haricot                            | 180            |
| protéine PSS2<br>(nodulation) | <i>Rhizobium<br/>leguminosarum</i> | 176            |

Test de Lipman et Pearson (1985). Les homologies ont été déduites à partir de la séquence consensus contenant les différents contigs.

Tableau 17

une protéine localisée 196 paires de bases en amont du gène *spoIID* est 50% identique à la protéine R (*orfR*) découverte par Trach et coll. (1988) qui ont suggéré que *orfR* est le site de la mutation *rev-4*. Le phénotype de cette mutation a été caractérisé par Sharrock et Leighton (1982) comme étant une suppression pléiotrope des déficiences de sporulation de certaines mutations ribosomales induites par la résistance à l'érythromycine. En outre, le produit de *orfR* ne paraît pas indispensable pour la croissance ou la sporulation chez *B. subtilis* (Trach et al., 1988). On peut encore noter la possible présence du gène de la 3-dehydroquinase synthase de *B. subtilis* détecté par son homologie avec le produit du gène de la 3-dehydroquinase synthase d'*E. coli* (Millar and Coggins, 1986). La 3-dehydroquinase synthase de *B. subtilis* a du reste été purifiée et étudiée par Hasan et Nester (1978). L'intérêt de cette enzyme réside dans son implication dans la biosynthèse des chaînes aromatiques, et donc dans la biosynthèse des acides aminés (Umbarger and Davis, 1962). On peut finalement noter la présence d'un gène homologue à l'uréease du haricot (*jack bean*).

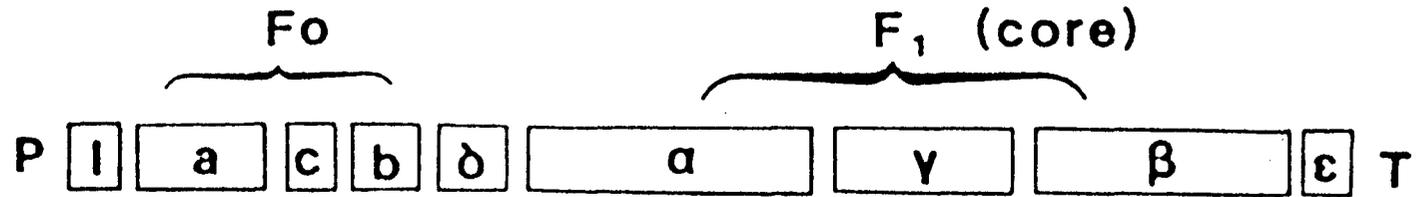
### c) OPERON ATP SYNTHASE DE *BACILLUS SUBTILIS*

Les ATPases sont des enzymes qui sont étudiées de façon intensive. La principale raison de cet intérêt est que ce complexe enzymatique catalyse l'une des réactions les plus importantes pour la cellule, à savoir la synthèse de l'ATP en utilisant l'énergie électrochimique d'un gradient de protons, gradient qui peut être généré par une chaîne respiratoire ou par l'utilisation de la lumière chez les bactéries photosynthétiques. Le mécanisme par lequel le gradient de protons est utilisé pour conduire la synthèse de l'ATP est l'un des problèmes principaux de Biochimie non encore résolus. Il est donc clair que les connaissances fondamentales qui concernent la structure moléculaire et le mécanisme d'assemblage du complexe enzymatique ATP synthase

sont d'importance majeure en regard de la biologie cellulaire (Marzuki, 1989).

L'ATP synthase est une structure membranaire très complexe qui contient deux sous-unités principales (F<sub>1</sub>, F<sub>o</sub>) et plusieurs sous-unités. La sous-unité F<sub>o</sub> est une protéine transmembranaire insoluble dans l'eau composée de 4 à 5 types de sous-unités (i, a, c, b chez B. megaterium, cf. Brusilow et coll., 1989) et contient un canal pour permettre la translocation des protons. La sous-unité F<sub>1</sub> est une protéine soluble dans l'eau composée de 5 types de sous-unités ( $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ) assemblés en un complexe de la forme  $\alpha_3\beta_3\gamma\delta\epsilon$ . Les sous-unités  $\delta$ ,  $\gamma$ , et  $\epsilon$  interagissent avec la sous-unité F<sub>o</sub>, les sous-unités  $\delta$  et  $\epsilon$  étant nécessaires à la fixation de F<sub>1</sub> sur le "socle membranaire" F<sub>o</sub> ; tandis que la sous-unité  $\gamma$  pourrait agir comme une porte régulant le flux de protons. La sous-unité  $\beta$  est responsable de l'activité catalytique de la synthèse de l'ATP ; la sous-unité  $\alpha$  pourrait être impliquée dans la régulation de l'activité ATP synthase (Fillingame, 1980 ; Cross, 1981 ; Futai and Kanazawa, 1983 ; Voet and Voet, 1990).

Les séquences des gènes du complexe ATPase de divers procaryotes ont été déterminées. Bien que ces gènes apparaissent bien conservés sur une large distance évolutive, leur organisation varie relativement. En effet, il apparait que chez Rhodopseudomonas blastica et Rhodopseudomonas rubrum les gènes de la sous-unité F<sub>1</sub> et de la sous-unité F<sub>o</sub> ne sont pas adjacents, tandis que chez Synechococcus PCC6301 et Anabaena PCC7120 ils ne sont pas tous contigus. La séquence nucléotidique des gènes *atpF*, *atpH*, *atpA*, *atpG*, *atpD*, et *atpC* de B. subtilis est présentée en Annexes. En outre, ces gènes sont organisés en opéron de façon virtuellement identique chez B. subtilis, B. megaterium (cf. Figure 33) (Hawthorne and Brusilow, 1988 ; Brusilow et al., 1989), E. coli (McCarthy et al., 1985), et la bactérie thermophile PS3 (Ohta et al., 1988), du moins en ce qui concerne les gènes des sous-unités b,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$  (cf. Figure 33). Les tailles en



|                        |     |     |    |     |     |     |     |     |     |
|------------------------|-----|-----|----|-----|-----|-----|-----|-----|-----|
| <i>B. megaterium</i>   | 123 | 236 | 70 | 172 | 181 | 502 | 285 | 472 | 134 |
| <i>thermophile PS3</i> | 127 | 210 | 72 | 163 | 179 | 502 | 286 | 473 | 132 |
| <i>B. subtilis</i>     |     |     |    | 170 | 181 | 502 | 287 | 473 | 132 |
| <i>E. coli</i>         | 127 | 271 | 79 | 156 | 177 | 514 | 288 | 460 | 139 |

**Figure 33** : L'opéron ATP synthase

Les tailles des protéines sont indiquées en codons

(d'après Kanazawa et coll., 1981 ; Ohta et coll., 1988 ; Brusilow et coll., 1989)

(codons) de ces différentes sous-unités sont de plus très semblables chez B. subtilis, B. megaterium, et la bactérie thermophile PS3 (cf. Figure 33).

L'alignement des séquences peptidiques des différentes sous-unités du complexe ATPase de B. subtilis avec celles de B. megaterium (cf. Figure 34a-34f) et de la bactérie thermophile PS3 (voir Annexes) indique la très forte conservation de ces gènes chez ces trois microorganismes. Il est en outre remarquable que les sous-unités  $\alpha$  et  $\beta$  sont particulièrement conservées (respectivement, 85,3% et 86,3% d'identité entre B. subtilis et B. megaterium) alors que les sous-unités  $\gamma$ ,  $\delta$ , et  $\epsilon$  le sont moins (respectivement, 74,6%, 48,3%, 68,9% d'identité chez B. subtilis et B. megaterium). Ceci peut s'expliquer par l'observation que, d'une manière générale, les sous-unités de Fo sont peu conservées, ce qui implique que les sous-unités  $\gamma$ ,  $\delta$  et  $\epsilon$  qui interagissent avec elles ont dû évoluer en parallèle (Walker et al., 1985). La composition en acides aminés des sous-unités séquencées de l'ATPase de B. subtilis (cf. Tableau 18) est bien entendu très proche de celle des sous-unités de B. megaterium (Hawthorne and Brusilow, 1988 ; Brusilow et al., 1989). Les sites de fixation de l'ATP des sous-unités  $\alpha$  et  $\beta$  ont également été comparés (cf. Figure 35) ; il est clair que ces sites sont très conservés sur une large distance évolutive.

L'organisation de l'opéron ATP synthase de B. subtilis a été comparée en détails avec celle de B. megaterium (Tableau 19). Certaines caractéristiques intéressantes peuvent être issues de cette comparaison. On peut tout d'abord remarquer que, comme chez B. megaterium, le gène *atpG* (sous-unité  $\gamma$ ) débute par le codon UUG, ce qui pourrait représenter un mécanisme pour éviter la surproduction de la sous-unité  $\gamma$ , cette sous-unité étant présente à raison d'une seule copie par complexe ATPase. L'utilisation du codon de démarrage UUG a en effet été démontrée chez E. coli comme résultant en une expression de gène nettement moindre, comparée à l'expression

```

031482          522, 532
65.3% IDENTITY IN 160 AA OVERLAP

      10      20      30      40      50
F8D      MSQPLELGLSFYGGDILFOLLAMLILLALLKKYALGPLLNIMKOREDHIAGEITSAE
      ..X::::::::: .::::::::: .::::::::: .::::::::: .::::::::: .:::::::::
D31482    MAVSNMFVLGAAGINGGDILFQLVMFLILLALLOKFAFGPVMGIMKKREEHIAGEIDEAE
      10      20      30      40      50      60

      60      70      80      90      100     110
F8D      EKNKEAQQLEEDRVLLKEARQESOTLIENAKKLGEKOKEEIIQAARAESEFLKEAARTE
      ..::::::::: .::::::::: .::::::::: .::::::::: .::::::::: .::::::::: .
D31482    KQNEEAKKLVEEQREILKQSRQEVQVMENARKSAEDKKEEIVAAAREESEFLKAAAKQE
      70      80      90      100     110     120

      120     130     140     150     160     170
F8D      IVKEKEQAVSALREQVASLSVMIASKVIEKELDEQAQEKLIODYLKEVGESR
      : .::::::::: .::::::::: .::::::::: .::::::::: .::::::::: .X. :
D31482    IEQOKDQAVAAALREQVASLSVLIASKVIEKELSEQDQEKLIHEYIQEVGDVR
      130     140     150     160     170

```

**Figure 34a** : Alignement de la séquence peptidique de la sous-unité b de Fo de *B. subtilis* avec la sous-unité b de Fo de *B. megaterium* (D31482)

```

E31482          450, 468
48.3% IDENTITY IN 176 AA DVER_AP

      10      20      30      40      50      60
DP1      MSGSAVSKRYASALF DIANESAQLNQVEEEL IYVKOVFONEKALNDVLNHPKVPAAKKKE
X:      ::::: ::::: ::::: ::::: ::::: :::::
E31482      MSQPAVAKRYALALF QLATEKOMIDEMODDQI VEEVFAKTPELMDVLTHPKITIERKKQ
      10      20      30      40      50      60

      70      80      90      100     110     120
DP1      LIONAFGSL SOSVLNTIFLLIDRHRAAIVPELT DEFIKLANVARQTEDAIVYSVKPLTDA
      .. ::::: ::::: ::::: ::::: ::::: :::::
E31482      FVSEAFaelSPTVQHTVLLLLERHRIGIVSEMVKRYRFLANEVARGVADATVYSVKPLSAD
      70      80      90      100     110     120

      130     140     150     160     170     180
DP1      EMLPLSOVFAKKAGVAsLRIRNEVQD LIGGIKVRIGNRIYDGSVSGKLORIERQLAGEN
      :. ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
E31482      EKRAISQSFASKVg<HTLNISNIVDKTVIGGVKLRIGNRIYDGSISSKLETIHRGLLAHR
      130     140     150     160     170     180

DP1          R
E31482       S

```

**Figure 34b** : Alignement de la séquence peptidique de la sous-unité  $\delta$  de F1 (DP1) de *B. subtilis* avec la sous-unité  $\delta$  de F1 de *B. megaterium* (E31482)



```

G31482          700, 1026
74.6% IDENTITY IN 287 AA OVERLAP

      10      20      30      40      50      60
GP1    LASLRDIKSRIITSTKKT SQITKAMOMVSAAKLNRAENNAKSFVPYMDKIOEVVSNVGRVS
G31482 .X::::::::::: :
      10      20      30      40      50      60
      70      80      90      100     110     120
GP1    GNVKHPMLLSREVKK TAYL VITSDRGLAGAFNS SVLRSAYOAMQERHOSKDEYAVIAIGR
G31482 ..::::: .:.....: :
      70      80      90      100     110     120
      130     140     150     160     170     180
GP1    VGRDFFKKREIPII SELTGLGDEVTFTEIKDLARQTIQMFIDGAFDELHLYVNHVSAIT
G31482 ::::::: :.....: :
      130     140     150     160     170     180
      190     200     210     220     230     240
GP1    QEVTEKKLLPLSDLGSGGGKRTASYEFEPSEEEVLEVLLPQYAESLIFGALLDSKASEHA
G31482 ::::::: :.....: :
      190     200     210     220     230
      250     260     270     280
GP1    ARMTAMKNATDNAKELIDSLSLSYNRARQAAITQEITEIVGGAAALE
G31482 ::::::: :.....: :
      240     250     260     270     280

```

Figure 34d : Alignement de la séquence peptidique de la sous-unité  $\gamma$  de F1 de *B. subtilis* (GP1) avec la sous-unité  $\gamma$  de F1 de *B. megaterium* (G31482)



```

PWBSEM          444, 451
68.9% IDENTITY IN 132 AA OVERLAP

      10      20      30      40      50      60
EPI      MKTVKVNIVTPDGPVYDADIEMVSVRAESGDLGILPGHIPTVAPLKIGAVRLKKOGTEM
X::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
PWBSEM    MKT IHVS VVTPDGPVYESEVEMVSTRAOSGELGILHGHPMVAPLQIGAVRLKKASSTEL
      10      20      30      40      50      60

      70      80      90      100     110     120
EPI      YAVSGGFVEVRPDHVTILAQAAETAEGIDKERAEEAARORAQERLNSQSDDTDIRRAELAL
::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
PWBSEM    YAVSGGFLEVRPDKVTILAQAAETAEEIDVARAEEAKKRAEMRLDSKQDDVDVKRAEIAL
      70      80      90      100     110     120

      130
EPI      QRALNRLDVAGK
:::::::::X..
PWBSEM    KRAVNRLDISQRF
      130

```

**Figure 34f** : Alignement de la séquence peptidique de la sous-unité  $\epsilon$  de F1 de *B. subtilis* (EPI) avec la sous-unité  $\epsilon$  de F1 de *B. megaterium* (PWBSEM)

**L'OPERON ATPase DE BACILLUS SUBTILIS :  
COMPOSITION EN ACIDES AMINES**

| sous-unité<br>acide aminé | F <sub>0b</sub> | F <sub>1δ</sub> | F <sub>1α</sub> | F <sub>1γ</sub> | F <sub>1β</sub> | F <sub>1ε</sub> |
|---------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Alanine                   | 18              | 19              | 46              | 31              | 41              | 17              |
| Cystéine                  | 0               | 0               | 2               | 0               | 0               | 0               |
| Aspartate                 | 4               | 8               | 32              | 13              | 28              | 11              |
| Glutamate                 | 27              | 13              | 33              | 24              | 40              | 10              |
| Phénylalanine             | 2               | 6               | 15              | 10              | 19              | 1               |
| Glycine                   | 7               | 9               | 44              | 16              | 44              | 10              |
| Histidine                 | 1               | 2               | 6               | 5               | 10              | 2               |
| Isoleucine                | 13              | 14              | 39              | 19              | 30              | 8               |
| Lysine                    | 17              | 13              | 26              | 20              | 19              | 7               |
| Leucine                   | 26              | 19              | 49              | 27              | 38              | 10              |
| Méthionine                | 4               | 2               | 9               | 8               | 13              | 3               |
| Asparagine                | 4               | 11              | 15              | 10              | 14              | 3               |
| Proline                   | 2               | 5               | 24              | 6               | 22              | 6               |
| Glutamine                 | 15              | 11              | 26              | 12              | 23              | 6               |
| Arginine                  | 8               | 11              | 27              | 16              | 23              | 11              |
| Serine                    | 10              | 13              | 29              | 26              | 21              | 5               |
| Threonine                 | 3               | 5               | 23              | 17              | 29              | 7               |
| Valine                    | 7               | 17              | 41              | 19              | 48              | 14              |
| Tryptophane               | 0               | 0               | 0               | 0               | 0               | 0               |
| Tyrosine                  | 2               | 3               | 16              | 8               | 11              | 1               |

Tableau 18

SITES DE FIXATION DE L'ATP DES SOUS-UNITES  
 $\alpha$  ET  $\beta$  DE L'ATP SYNTHASE

sous-unité

F1 $\alpha$

|                         |  |
|-------------------------|--|
| <i>E. coli</i>          | G-D-R-Q-T-G-K-T-N5-A-I-N69-G-E-D-A-L-I-I-Y-D-D |
| <i>B. subtilis</i>      | G-D-R-Q-T-G-K-T-N5-A-I-N70-G-K-H-V-L-V-V-Y-D-D |
| <i>B. megaterium</i>    | G-D-R-Q-T-G-K-T-N5-T-I-N70-G-K-H-V-L-V-I-Y-D-D |
| thermophile             | G-D-R-Q-T-G-K-T-N5-T-I-N70-G-K-H-V-L-V-V-I-D-D |
| Boeuf<br>(mitochondrie) | G-D-R-Q-T-G-K-T-N5-T-I-N76-G-K-H-A-L-I-I-Y-D-D |

sous-unité

F1 $\beta$

|                         |  |
|-------------------------|--|
| <i>E. coli</i>          | G-G-A-G-G-V-K-T-N5-L-I-N71-G-R-D-V-L-L-F-V-D |
| <i>B. subtilis</i>      | G-G-A-G-V-G-K-T-N5-L-I-N71-G-Q-D-V-L-F-F-I-D |
| <i>B. megaterium</i>    | G-G-A-G-V-G-K-T-N5-L-I-N71-G-Q-D-V-L-F-F-I-D |
| thermophile             | G-G-A-G-V-G-K-T-N5-L-I-N71-G-Q-D-G-L-L-F-I-D |
| Boeuf<br>(mitochondrie) | G-G-A-G-G-V-K-T-N5-L-I-N77-G-Q-D-V-L-L-F-I-D |

Les sites de fixation de l'ATP des séquences d'*E. coli* et des mitochondries de boeuf proviennent de Fry et coll., 1986 et Walker et coll., 1985. Les séquences de *B. megaterium* et de la bactérie thermophile PS3 proviennent de la banque peptidique NBRF.

Figure 35

ORGANISATION DE L'OPERON ATP SYNTHASE  
CHEZ *B. SUBTILIS* ET *B. MEGATERIUM*

| unité             | caractère                     | <i>B. subtilis</i> | <i>B. megaterium</i> |
|-------------------|-------------------------------|--------------------|----------------------|
| Fo (b)            | taille                        | 170 aa             | 172 aa               |
|                   | % d'identité                  | 66,3               | 100                  |
|                   | RBS                           | AAGGGAG            | AAGGGAG              |
|                   | Start                         | AUG                | AUG                  |
|                   | Stop                          | UGA                | UGA                  |
|                   | distance b- $\delta$          | -1                 | -1                   |
| F1 ( $\delta$ )   | taille                        | 181 aa             | 181 aa               |
|                   | % d'identité                  | 48,3               | 100                  |
|                   | RBS                           | AGGAGA             | AGGGGA               |
|                   | Start                         | AUG                | AUG                  |
|                   | Stop                          | UAG                | UAG                  |
|                   | distance $\delta$ - $\alpha$  | 16                 | 15                   |
| F1 ( $\alpha$ )   | taille                        | 502 aa             | 502 aa               |
|                   | % d'identité                  | 83,9               | 100                  |
|                   | RBS                           | AGGGGTG            | AGGGGTG              |
|                   | Start                         | GUG                | AUG                  |
|                   | Stop                          | UAA                | UAA                  |
|                   | distance $\alpha$ - $\gamma$  | 76 bp              | 105 bp               |
| F1 ( $\gamma$ )   | taille                        | 287 aa             | 285 aa               |
|                   | % d'identité                  | 74,6               | 100                  |
|                   | RBS                           | AAGGTG             | AAGGTG               |
|                   | Start                         | UUG                | UUG                  |
|                   | Stop                          | UAG                | UAG                  |
|                   | distance $\gamma$ - $\beta$   | 25 bp              | 36 bp                |
| F1 ( $\beta$ )    | taille                        | 473 aa             | 472 aa               |
|                   | % d'identité                  | 86,3               | 100                  |
|                   | RBS                           | AGGAGG             | AGGAGG               |
|                   | Start                         | AUG                | AUG                  |
|                   | Stop                          | UAA                | UAA                  |
|                   | distance $\beta$ - $\epsilon$ | 24 bp              | 26 bp                |
| F1 ( $\epsilon$ ) | taille                        | 132 aa             | 134 aa               |
|                   | % d'identité                  | 68,9               | 100                  |
|                   | RBS                           | AGGAGGGT           | AGGAGGT              |
|                   | Start                         | AUG                | AUG                  |
|                   | Stop                          | UAA                | UAA                  |

Tableau 19

à partir des codons de démarrage AUG et GUG (Reddy et al., 1985). Les espacements entre les divers gènes sont du même ordre chez B. subtilis, B. megaterium, et E. coli (cf. Tableau 19) quoique plus petits chez E. coli (Brusilow et al., 1989). En outre, à l'instar de B. megaterium (Brusilow et al., 1989) et Anabaena PCC7120 (McCarn et al., 1988), les gènes des sous-unités  $\beta$  et  $\delta$  de B. subtilis se chevauchent de quatre bases. Comme souligné par Brusilow et coll. (1989) l'un des modèles pour expliquer l'assemblage de l'ATPase chez E. coli implique une interaction entre les sous-unités  $\beta$  et  $\delta$  qui pourrait être importante dans le processus de fixation de la sous-unité F<sub>1</sub> à la sous-unité F<sub>o</sub>. Un tel chevauchement entre ces deux gènes pourrait donc faciliter soit le couplage traductionnel de l'expression de ces deux gènes, soit une quelconque étape d'assemblage entre ces deux sous-unités. On peut également remarquer que les codons stop utilisés chez B. subtilis et B. megaterium sont identiques pour chacun des gènes  $b$ ,  $\delta$ ,  $\alpha$ ,  $\gamma$ ,  $\beta$ ,  $\epsilon$  ainsi que les séquences de Shine-Dalgarno, sauf dans le cas des gènes  $\delta$  et  $\epsilon$  où la différence est de une base. Les opérons de B. subtilis et B. megaterium se terminent après le codon stop de la sous-unité  $\epsilon$  par un terminateur représenté Figure 36.

L'usage du code des gènes séquencés de l'ATPase de B. subtilis a été déterminé par le calcul du rapport de la fréquence d'utilisation du codon considéré et de la fréquence d'utilisation du codon le plus utilisé de la même classe (Tableau 20). L'analyse des valeurs obtenues semble indiquer qu'il n'y a pas sensiblement de biais contre l'usage des codons "rares" de B. subtilis. Ce résultat suggère donc que le gène de l'ATPase est exprimé à un niveau moyen. Ceci n'est du reste pas étonnant étant donné que la surproduction d'ATP synthase affecte la division cellulaire chez E. coli et implique une inhibition de la croissance (von Meyenburg et al., 1984).

La présence d'un site NotI dans le gène de la sous-unité  $\alpha$  permet de relier la position de cet opéron par rapport à la



**L'OPERON ATPase DE BACILLUS SUBTILIS :  
USAGE DU CODE**

| acide aminé | codon | w     | acide aminé | codon | w     |
|-------------|-------|-------|-------------|-------|-------|
| Phe         | TTT   | 0,780 | Ala         | GCA   | 0,767 |
| Phe         | TTC   | 1,000 | Ala         | GCG   | 1,000 |
| Leu         | TTA   | 0,451 | Tyr         | TAT   | 1,000 |
| Leu         | TTG   | 0,354 | Tyr         | TAC   | 0,770 |
| Leu         | CTT   | 1,000 | His         | CAT   | 0,811 |
| Leu         | CTC   | 0,162 | His         | CAC   | 1,000 |
| Leu         | CTA   | 0,079 | Gln         | CAA   | 1,000 |
| Leu         | CTG   | 0,515 | Gln         | CAG   | 0,838 |
| Ile         | ATT   | 0,914 | Asn         | AAT   | 0,527 |
| Ile         | ATC   | 1,000 | Asn         | AAC   | 1,000 |
| Ile         | ATA   | 0,136 | Lys         | AAA   | 1,000 |
| Met         | ATG   | 1,000 | Lys         | AAG   | 0,311 |
| Val         | GTT   | 1,000 | Asp         | GAT   | 1,000 |
| Val         | GTC   | 0,467 | Asp         | GAC   | 0,795 |
| Val         | GTA   | 1,000 | Glu         | GAA   | 1,000 |
| Val         | GTG   | 0,488 | Glu         | GAG   | 0,499 |
| Ser         | TCT   | 1,000 | Cys         | TGT   | 1,000 |
| Ser         | TCC   | 0,412 | Cys         | TGC   | 0,859 |
| Ser         | TCA   | 0,969 | Trp         | TGG   | 0,000 |
| Ser         | TCG   | 0,175 | Arg         | CGT   | 1,000 |
| Ser         | AGT   | 0,354 | Arg         | CGC   | 0,386 |
| Ser         | AGG   | 0,077 | Arg         | CGA   | 0,192 |
| Pro         | CCT   | 0,563 | Arg         | CGG   | 0,192 |
| Pro         | CCC   | 0,031 | Arg         | AGA   | 0,308 |
| Pro         | CCA   | 0,468 | Arg         | AGG   | 0,077 |
| Pro         | CCG   | 1,000 | Gly         | GGT   | 0,931 |
| Thr         | ACT   | 0,275 | Gly         | GGC   | 1,000 |
| Thr         | ACC   | 0,751 | Gly         | GGA   | 0,908 |
| Thr         | ACA   | 1,000 | Gly         | GGG   | 0,137 |
| Thr         | ACG   | 0,625 | Stop        | TAA   | 0,572 |
| Ala         | GCT   | 0,677 | Stop        | TAG   | 0,286 |
| Ala         | GCC   | 0,465 | Stop        | TGA   | 1,000 |

w représente le rapport entre la fréquence d'utilisation du codon considéré et la fréquence d'utilisation du codon le plus utilisé de la même classe.

Tableau 20

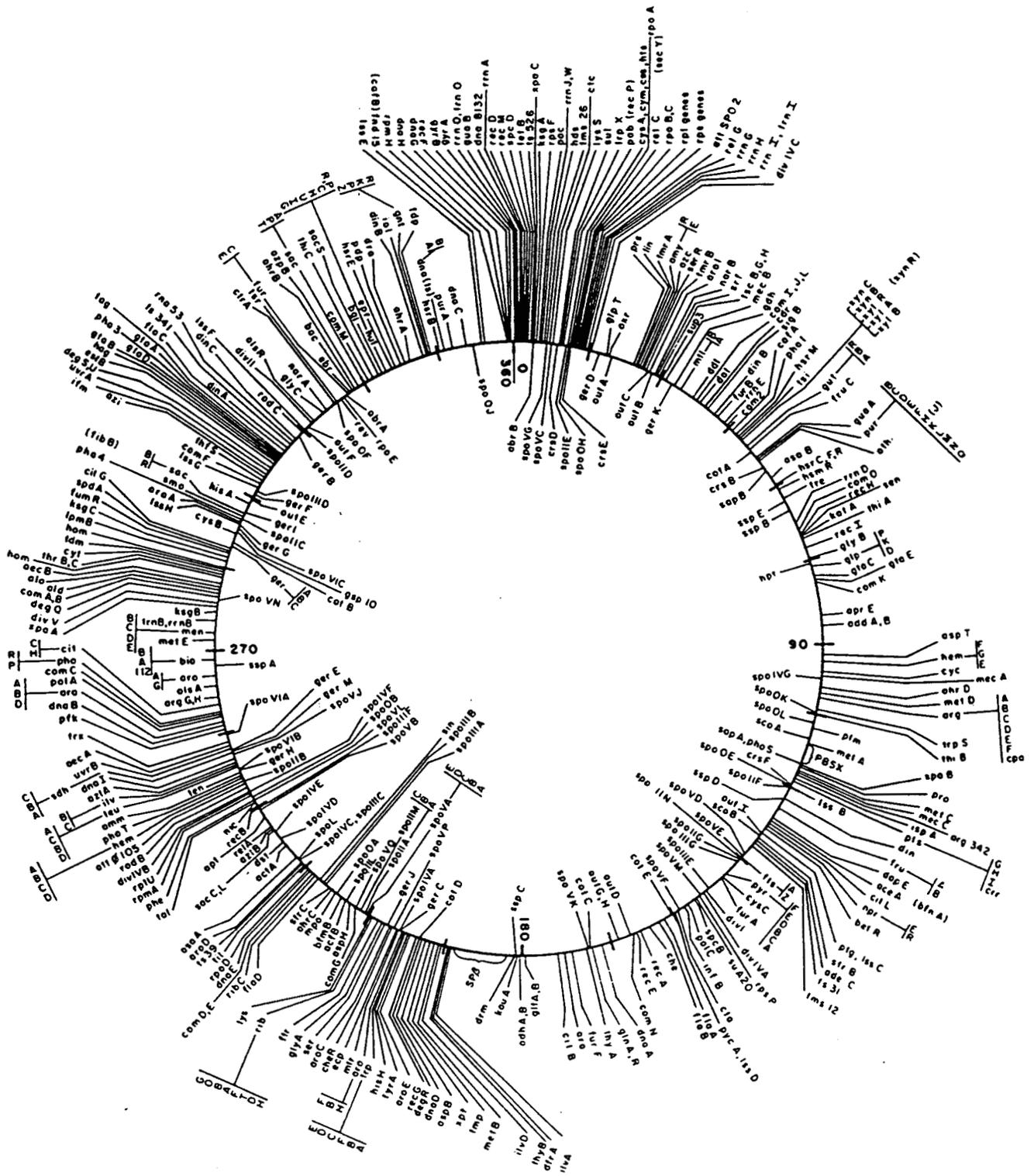
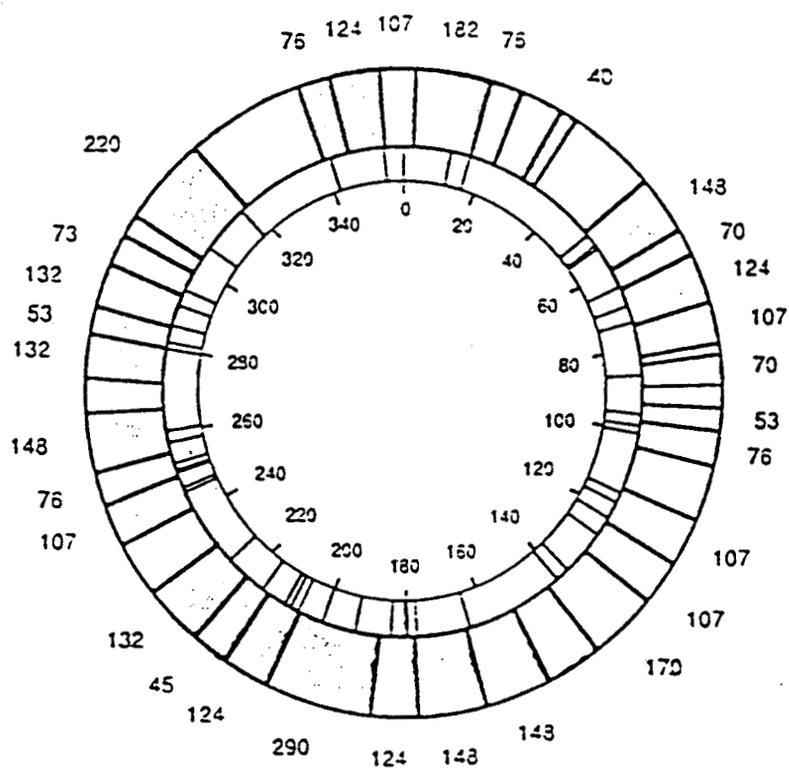


Figure 10 : Carte génétique de *B. subtilis* 168 (d'après Piggot, 1989)



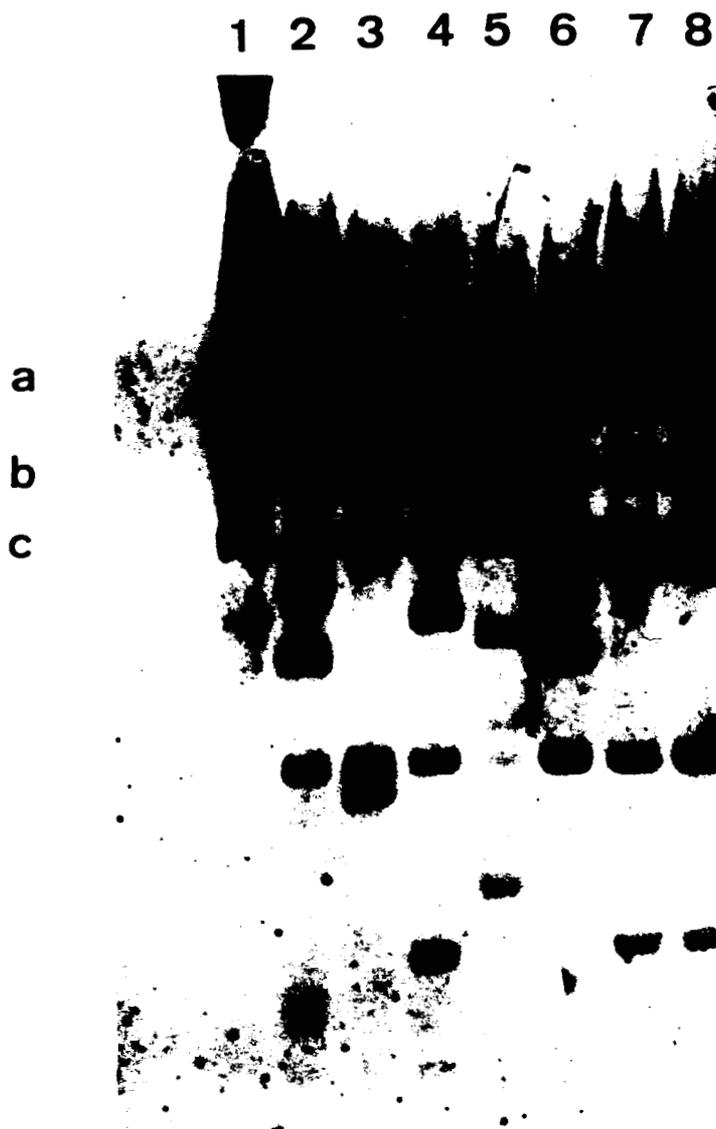
**Figure 11** : Carte physique du chromosome de *B. subtilis* 168. Les tailles des différents fragments NotI sont indiquées en kb. Les coordonnées de la carte génétique sont reportées à l'intérieur de la Figure. (d'après Ventura et Weiss, 1989)

carte physique (cf. Figure 11) et à la carte génétique (cf. Figure 10). D'après les résultats obtenus par Ventra et Weiss (1989), l'opéron ATP synthase est situé entre 317° et 320°.

#### d) IDENTIFICATION DE *SPOIID* ET ANALYSE DES GENES ADJACENTS

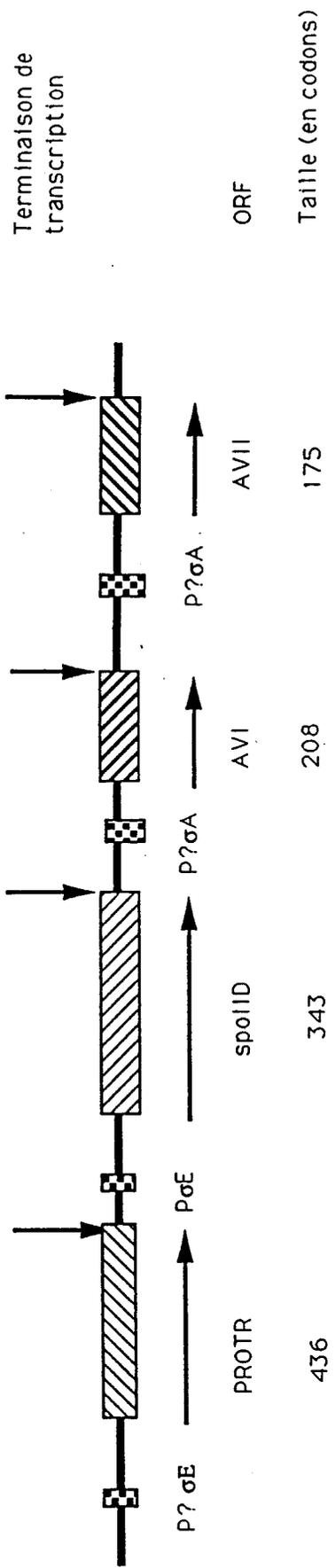
Comme indiqué Tableau 17, l'analyse globale de la séquence nucléotidique de l'insert chromosomique du phage  $\lambda$ narA révèle la présence du gène *spoIID* (séquencé par Lopez-Diaz et coll., 1986). Ce résultat a en outre été confirmé par la séquence peptidique déduite de la séquence nucléotidique (cf. Annexes). Le chevauchement des phages  $\lambda$ spoIID avec les phages  $\lambda$ narA a également été vérifié. Les ADN du phage  $\lambda$ narA4.0 et des 6 phages  $\lambda$ spoIID ont été hydrolysés par SallI et transférés sur une membrane de nylon après électrophorèse. L'hybridation avec une sonde constituée de l'ADN du phage  $\lambda$ narA4.0 marqué à la biotine révèle (cf. Figure 40) que toutes les bandes des 6  $\lambda$ spoIID différentes réagissent avec les séquences contenues dans le phage  $\lambda$ narA4.0, ce qui démontre clairement que les phages  $\lambda$ narA et  $\lambda$ spoIID constituent une même famille de phages recombinants. La position de l'un des marqueurs *narA* (320°) ou *spoIID* (316°) et donc erronée.

L'analyse de la séquence adjacente au locus *spoIID* (cf. Annexes) révèle la présence de 3 phases ouvertes de lecture (ORF : PROTR, AVI, AVII) (cf. Figure 41). L'ORF PROTR pourrait être contrôlée par un promoteur de type  $\sigma^E$  (région "-35" GAACAGTT et "-10" GATATT), le RBS semble être GGAGGGG, le codon d'initiation UUG, et le codon stop UAA. La protéine AVI pourrait être contrôlée par un promoteur de type  $\sigma^A$  (région "-35" TTGTCG et "-10" TATAAT), le RBS semble être AAAGGGGT, le codon de démarrage AUG, et le codon stop UAA. La protéine AVII pourrait être contrôlée par un promoteur de type  $\sigma^A$  (région "-35" TTGCA et "-10" TATTCT), le RBS semble être AAAGGAAGT, le codon de démarrage AUG, et le codon stop UAA. La composition en acides aminés de ces 3 protéines putatives est indiquée Tableau 23. Leur profil hydrophatique



**Figure 40** : Hybridation des phages lambdaspoIID avec le phage lambdaDanarA4.0

- 1) marqueur lambdaHindIII, 2) lambdaDanarA4.0 hydrolysé par Sall, 3) lambdaspoIID3 hydrolysé par Sall, 4) lambdaspoIID6 hydrolysé par Sall, 5) lambdaspoIID10 hydrolysé par Sall, 6) lambdaspoIID14 hydrolysé par Sall, 7) lambdaspoIID17 hydrolysé par Sall, 8) lambdaspoIID19 hydrolysé par Sall



P : promoteur  
P? : promoteur putatif

Figure 41 : Carte de la région adjacente à spoIID

**ORFs DETECTEES DANS LE PHAGE  $\lambda$ NARA :**  
**COMPOSITION EN ACIDES AMINES**

| ORF<br>acide aminé | ProtR | AVI | AVII | Uréase |
|--------------------|-------|-----|------|--------|
| Alanine            | 48    | 18  | 16   | 12     |
| Cystéine           | 2     | 3   | 2    | 1      |
| Aspartate          | 17    | 5   | 10   | 4      |
| Glutamate          | 41    | 13  | 10   | 12     |
| Phénylalanine      | 11    | 10  | 7    | 3      |
| Glycine            | 44    | 44  | 16   | 8      |
| Histidine          | 13    | 1   | 3    | 2      |
| Isoleucine         | 32    | 7   | 9    | 5      |
| Lysine             | 25    | 19  | 17   | 8      |
| Leucine            | 39    | 13  | 19   | 10     |
| Méthionine         | 15    | 5   | 4    | 5      |
| Asparagine         | 18    | 11  | 6    | 1      |
| Proline            | 16    | 7   | 5    | 5      |
| Glutamine          | 10    | 6   | 3    | 4      |
| Arginine           | 20    | 4   | 6    | 4      |
| Serine             | 16    | 18  | 16   | 3      |
| Thréonine          | 24    | 11  | 12   | 5      |
| Valine             | 39    | 9   | 9    | 11     |
| Tryptophane        | 0     | 1   | 2    | 0      |
| Tyrosine           | 6     | 3   | 3    | 2      |

Tableau 23

```

G32354          394, 917
50.0% IDENTITY IN 352 AA OVERLAP

      10      20      30      40      50      60
PROTR      LEKIIYRGGQKLNGT VKVEGAKNAVLPVIAASLLASEEKSIVICDVPTLSDVYTINEVLRH
G32354      MEKLNIA GGDSLNGT VTHISGAKNSAVALIPATILANSEVTIE-GLPEISDIETLRLLKE
      10      20      30      40      50

      70      80      90      100     110     120
PROTR      LGADVHFENNEVTVNASYALQTEAPFEYVRKMRASVLMGPELLARTGHARVALPGGCAIG
G32354      IGGNVHFEENGEHVVDPTSMIS MPLPNGKVKKLRASYLMGAMLGRFKQAVIGLPGGCHLG
      60      70      80      90      100     110

      130     140     150     160     170     180
PROTR      SRPIDQHLKGFEAMGAEIKVGNFGIEAEVKGRLGAKIYLOFPSVYGATENLIMAAALAE
G32354      PRPIDQHKGFEALGAEVTNEQGAIYLR-ERLRGARIYLDVVSVGATINIMLAALVLAEG
      120     130     140     150     160     170

      190     200     210     220     230     240
PROTR      TTLENVAKEPEIIDLANYINGMGGKIRGAGGTIKIEGVEKLGHVKHIIIPDRIEAGTF
G32354      KTIIEAAKEPEIIDVATLLTSMGAKIKGAGTNVIRIDGVKELHGCKHTIIPDRIEAGTF
      180     190     200     210     220     230

      250     260     270     280     290
PROTR      MVAAAITEGNVYLVKGAPEHLTSLIAKMEEMGVTIKDEGEGLRVI-GPKELKPIDIKTMP
G32354      MIAGAAMGKEVIIIDNVIPHTLESHTAKLREMGYHIETSDQQLIVGGQKNLKPVDVKTLY
      240     250     260     270     280     290

      300     310     320     330     340     350
PROTR      HPGFPTDMOSQMMALLRASGTSMITETVFENRFMHAEFFRRMNGDIKIEGRSVIINGPV
G32354      YPGFPTDLQOPMTA_LTRAKGTSVYTDIYSARFKHIDELRRMGANMKVEGR
      300     310     320     330     340     350

      360     370     380     390     400     410
PROTR      QLQGAEVAATDLRAS AALILAGLVAEGHTRYTELKHLDRGYVDFHOKLAALGADIERNVD
      360     370     380     390     400     410

      420     430
PROTR      ESASEQENKEVYVSDLNA

```

**Figure 43** : Alignement de la séquence peptidique de la protéine putative protR avec la séquence du locus rev-4 (G32354)

comparé avec le profil hydrophatique d'une protéine typiquement membranaire (cf. en Annexes le profil hydrophatique de la sous-unité b de l'ATPase de B. subtilis) indique que ces dernières devraient être des protéines cytoplasmiques. Leur masse molaire ont été déterminées comme étant de 46629 pour PROTR, 19234 pour AVI, et 18861 pour AVII. La recherche dans la banque peptidique NBRF de protéines homologues à ces trois protéines putatives indique que les ORF AVI et AVII ne sont homologues à aucun gène connu, tandis que l'ORF PROTR apparait identique à 50% à l'ORF R située à environ 2,2 kb en aval de *spoOF* (Trach et al., 1988) (Figure 43).

#### e) LA SOUS-UNITE $\gamma$ DE L'UREASE DE BACILLUS SUBTILIS

Le test de Lipman et Pearson conduit sur les produits déduits de la séquence consensus des différents *contig* révèle (cf. Tableau 17) la présence d'un gène identique à 55,6% à la sous-unité  $\gamma$  de l'uréase d'un Mycoplasme (Ureaplasma urealyticum) (cf. Figure 44a), et identique à 60,2% aux 103 premiers acides aminés de l'uréase du haricot (cf. Figure 44b).

Une estimation de la signification statistique de ce test a été réalisée (cf. Matériels et Méthodes) afin de vérifier cette correspondance ; le score obtenu de  $z = 34,6$  confirme que le résultat est significatif.

La détermination de la séquence nucléotidique de l'uréase de B. subtilis est importante à plusieurs niveaux. Les uréases sont en effet des enzymes d'intérêt majeur étant donné qu'elles sont impliquées dans la pathogénicité de certaines bactéries (U. urealyticum, Proteus mirabilis, certaines souches de Pseudomonas, de Klebsiella, de Corynebacteria, de Campylobacter, de Staphylococci, etc ...). La présence de bactéries uréolytiques telles que Staphylococcus saprophyticus, Micrococcus varians, Lactobacillus casei, Klebsiella aerogenes, Bacteroides

```

S10030          257, 270
55.6% IDENTITY IN 99 AA OVERLAP

      10      20      30      40      50      60
JREASEP      MKLTPVEOEKLLIFAGELAKQRKARGVLLNYPEAAAYITCFIMEGARDGKGVAELMEAG
:.:. : .X:. .:..:..: :.:. :.:.: .: :. :.:.:.:.: :.:.:.:.
S10030      MNLSLREIQKLLVTVAADVARRRLARGLKLNYSERVALITDHSVVEGARDGKLVADLMQSA
      10      20      30      40      50      60

      70      80      90      100
UREASEP      RHLVTEKDVMEGVPEMLDSIQVEATFPDGVKLVTVHQPISAEVKS
:.:. :.:.:.: :.:. :.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.X
S10030      REVLRVDDQVMEGVDTMVGIIQVEVTFPDGTKLVSVHDPIYK
      70      80      90      100

```

Figure 44a : Alignement de la séquence peptidique de la sous-unité  $\gamma$  de l'uréase de *B. subtilis* avec la sous-unité  $\delta$  de l'uréase d'*Ureaplasma urealyticum* (S10030)

URJB 157, 270  
50.2% IDENTITY IN 103 AA OVERLAP

|         |                  |  |   |              |      |
|---------|------------------|--|---|--------------|------|
|         | 10               | 20   | 30  | 40           | 50   |
| UREASEP | MKLT             | PVEQEKLLIFA  | AGELAKQRKARGVLLNYPEAAAYITCFIMEGARDG-KGVAELMEA |              |      |
|         | :::.             | :::.   | :::.  | :::.         | :::. |
| URJB    | MKLS             | PREVEKLGHLNAGYLAQKRLARGVRLNYTEAVALIASOIMEYARDGEKTVAQLMCL |   |              |      |
|         | 10               | 20   | 30  | 40           | 50   |
|         | 60               | 70   | 80  | 90           | 100  |
| UREASEP | GRHVL            | TEKDVMEGVPEMLDSIQVEATFPDGVKLVTVHQPISAEVKS                |   |              |      |
|         | :::.             | :::.   | :::.  | :::.         | :::. |
| JRJB    | GQHLL            | GRROVLPVPHLLNAVQVEATFPDGTKLVTVDHPISRENGELOEALFGSLLPVPS   |   |              |      |
|         | 70               | 80   | 90  | 100          | 110  |
|         | 130              | 140  | 150   | 160          | 170  |
| JRJB    | LDKFAETKEDNRI    | PGEILCEDECLT   | LNIGRKAVILKVTSKGRPIQVGS                       | SHYHFIEVNPYL |      |
|         | 130              | 140  | 150   | 160          | 170  |
| URJB    | TFDRRKAYGMRLNIA  | AGTAVRFEPGDCKSVTLVSIEGNKVIRGGNAIADGPVNETNLEAA            |   |              |      |
|         | 190              | 200  | 210   | 220          | 230  |
| URJB    | MHAVRSR          | GFGHEEEKDAPEGFTKEDPNCSFNTFIHRKEYANKYGP                   | TTGDKIRLGD                                    | TNLLA        |      |
|         | 250              | 260  | 270   | 280          | 290  |
| URJB    | EIEKDYALYGDECVF  | GGKQVIRDGMGQSCGHPPAISLDTVI                               | TNAVIIDYTGIIKADIGIK                           |              |      |
|         | 310              | 320  | 330   | 340          | 350  |
| URJB    | DGLIASIGKAGNPDI  | MNGVFSNMIIGANTEVIA                                       | GEGLIVTAGAIDCHVHYICPQLVYEA                    |              |      |
|         | 370              | 380  | 390   | 400          | 410  |
| URJB    | SSGITTLVGGGTGPA  | AGTRATTCTPSP   | TQMLMLOSTDOLPLNFGFTGKGS                       | SSKPDELHE    |      |
|         | 430              | 440  | 450   | 460          | 470  |
| URJB    | IIKAGAMGLKLHEDWG | STPAAIDNCLTIAEHHD  | IQINIHTDTLNEAGFVEHSIAAFKGR                    |              |      |
|         | 490              | 500  | 510   | 520          | 530  |
|         | 550              | 560  | 570   | 580          | 590  |
| URJB    | IHTYHSEGAGGGHAP  | DIKVCYGIKNVLPSS  | TNPTRPLTSNTIDEHLDMLMVCHHL                     | DREIP        |      |
|         | 550              | 560  | 570   | 580          | 590  |
| URJB    | EDLAFASRIRKKTIA  | AEDVLDIGAI   | SISSDSQAMGRVGEVISRHWQADKMKAAQTGP              |              |      |
|         | 610              | 620  | 630   | 640          | 650  |
| URJB    | LKCDSSDNDNFRIR   | RYIAKYTINPAIANGFSQYVGSVEYVKLADLV                         | MKPSFFGT                                      | KPEMV        |      |
|         | 670              | 680  | 690   | 700          | 710  |
| URJB    | IKGGMVAWADIGD    | PNASIPTPEPVKMRPMYGT                                      | LKAGGALSIAFVSKAALDORVNVLYGL                   |              |      |
|         | 730              | 740  | 750   | 760          | 770  |
| URJB    | NKRVEAVSNVRKL    | TKLDMLKLNALPEITVD  | PESYTVKADGKLLCYSEATT                          | VPLSRNYFLF   |      |
|         | 790              | 800  | 810   | 820          | 830  |

**Figure 44b** : Alignement de la séquence peptidique de la sous-unité  $\gamma$  de l'uréase de *B. subtilis* avec l'uréase du haricot (URJB)

ruminicola, Ruminococcus bromii, Selenomonas ruminantium, etc ... ont été mises en évidence dans l'estomac des bovins et des ovins. De même chez l'homme, Peptostreptococcus productus et d'autres espèces uréolytiques sont impliquées dans la dégradation de l'urée transféré du sang vers l'intestin par diffusion. Les uréases jouent en outre un rôle prédominant dans l'agriculture de par leur implication dans le cycle de l'azote (Mobley and Hausinger, 1989).

L'organisation structurale des uréases varie selon les organismes. Par exemple, l'uréase du haricot est un hexamère d'une sous-unité de masse molaire 90770 (Takishima et al., 1988), celle de Bacillus pasteurii est un tétramère d'une sous-unité de masse molaire 65500 (Christians and Kaltwasser, 1986), alors que l'uréase de Klebsiella aerogenes, de P. mirabilis, de Providencia stuartii, et de S. ruminantium ont trois types de sous-unités distinctes dont les masses molaires varient de 68000 à 73000 (sous-unité  $\alpha$ ), de 8000 à 12000 (sous-unité  $\beta$ ), et 8000 à 10000 (sous-unité  $\gamma$ ) (Mobley and Hausinger, 1989). L'uréase d'U. urealyticum est également composée de 3 types de sous-unités dont les masses molaires sont 11,2 kdal (sous-unité  $\gamma$ ), 13,6 kdal (sous-unité  $\beta$ ), et 66,6 kdal (sous-unité  $\alpha$ ) (Blanchard, 1990). Les gènes de ces 3 sous-unités constituent un opéron chez U. urealyticum dont l'organisation est : Promoteur- $\gamma$ - $\beta$ - $\alpha$ . La distance entre le gène de la sous-unité  $\gamma$  et le gène de la sous-unité  $\beta$  est de 47 paires de bases. Le segment séquencé et identifié de l'uréase de B. subtilis comprend toute la région en amont du gène de la sous-unité  $\gamma$  et 53 codons en aval (cf. Annexes). Chez B. subtilis, l'uréase semble donc être également organisée en opéron dont le premier gène est celui de la sous-unité  $\gamma$  et le second celui de la sous-unité  $\beta$  (on peut noter une certaine identité entre la séquence de la sous-unité  $\beta$  d'U. urealyticum et les 30 premiers codons de la deuxième ORF). On peut encore noter que ces deux phases ouvertes de lecture se chevauchent de quatre bases. A l'instar de ce qui a été proposé au sujet de l'ATP synthase,

il est possible qu'un tel chevauchement pourrait soit faciliter le couplage traductionnel de l'expression de ces deux gènes, soit faciliter une quelconque étape d'assemblage. La séquence nucléotidique révèle en outre la présence d'un promoteur putatif  $\sigma^A$  (région "-35" TTGTCG et "-10" TATTAT). La composition en acides aminés de la protéine déduite de la séquence nucléotidique est reportée Tableau 23. La masse molaire calculée de la sous-unité  $\gamma$  de l'uréase de B. subtilis est de 11442, ce qui est très semblable à la masse molaire des sous-unités  $\gamma$  des autres microorganismes. Le profil hydrophatique indique sa localisation cytoplasmique, en accord avec la localisation des autres uréases (Mobley and Hausinger, 1989). L'alignement de la séquence nucléotidique de la sous-unité  $\gamma$  de l'uréase de B. subtilis avec celle d'U. urealyticum révèle 168 résidus identiques entre les deux séquences (55,1%), 78 transitions (25,6%) et 59 transversions (19,3%), et souligne donc la parenté entre ces deux enzymes.

Chez P. stuartii et chez P. mirabilis, l'expression de l'uréase est induite par l'urée et réprimée par la présence de meilleures sources d'azote dans le milieu de culture (Jones and Mobley, 1988 ; Mulrooney et al., 1988). De même chez Helicobacter pylori, l'expression de l'uréase est probablement régulée par le facteur  $\sigma^{54}$  (Kustu et al., 1989 ; Labigne et al., 1991). En ce qui concerne la régulation de l'uréase chez B. subtilis, il apparaît que cette enzyme est principalement soumise à un système de régulation général impliquant une copie sauvage du gène *glnA* et contrôlant de nombreuses enzymes cataboliques intervenant dans le métabolisme de l'azote (Atkinson and Fisher, 1991). Toutefois, malgré la présence d'un facteur de type  $\sigma^{54}$  chez B. subtilis (Débarbouillé et al., 1991), le gène de l'uréase n'est pas régulé par son intermédiaire (Débarbouillé et Kunst, communication personnelle).

Un examen de la séquence nucléotidique (cf. Annexes) révèle la séquence TTGAAAG 117 paires de bases (bp) en amont

du codon de démarrage de la traduction. Une telle séquence suggère un éventuel mécanisme de répression catabolique (Nicholson et al., 1987 ; Miwa and Fujita, 1990 ; Weickert and Chambliss, 1990 ; Henkin et al., 1991). En outre, on peut noter 65 bp en amont du gène de la sous-unité  $\gamma$  et dans une autre phase de lecture la présence d'un oligopeptide putatif (LTVPQIRSSRICFI) dont le codon de démarrage serait CUG et la séquence de Shine-Dalgarno AAGGGAGGTGAT. La probabilité d'occurrence d'une telle séquence très complémentaire de l'extrémité 16S des ARN ribosomiaux (11 bases sur 12) est très faible (1 fois tous les  $4^{11}$  nucléotides). Il serait donc particulièrement intéressant de vérifier si cet oligopeptide est exprimé, et d'évaluer la signification des signaux génétiques détectés afin de vérifier s'ils ne sont pas impliqués dans la régulation de l'expression génétique de l'uréase.

## **7) ANALYSE FONCTIONNELLE DES SEQUENCES OBTENUES**

### **Analyse du gène de la tyrosine tRNA synthétase (yrs)**

La détection de gènes homologues à partir de la comparaison des séquences obtenues avec les séquences compilées dans une banque de données est un outil d'analyse très puissant, mais comme, indiqué précédemment, soumis à certaines limitations.

L'usage de la génétique inverse est donc une étape obligée de l'analyse des séquences. Il est cependant clair qu'une analyse exhaustive, en particulier des gènes qui ne montrent aucune homologie notable avec des gènes déjà étudiés, est nécessaire pour exploiter totalement un tel travail.

La détermination de la séquence adjacente au locus *sacS* (Glaser et al., 1991) a permis de mettre en évidence une phase ouverte de lecture codant pour une protéine ayant une homologie significative avec les gènes de la tyrosine tRNA synthétase de plusieurs organismes. La fonctionnalité de cette phase ouverte de lecture a été vérifiée par complémentation

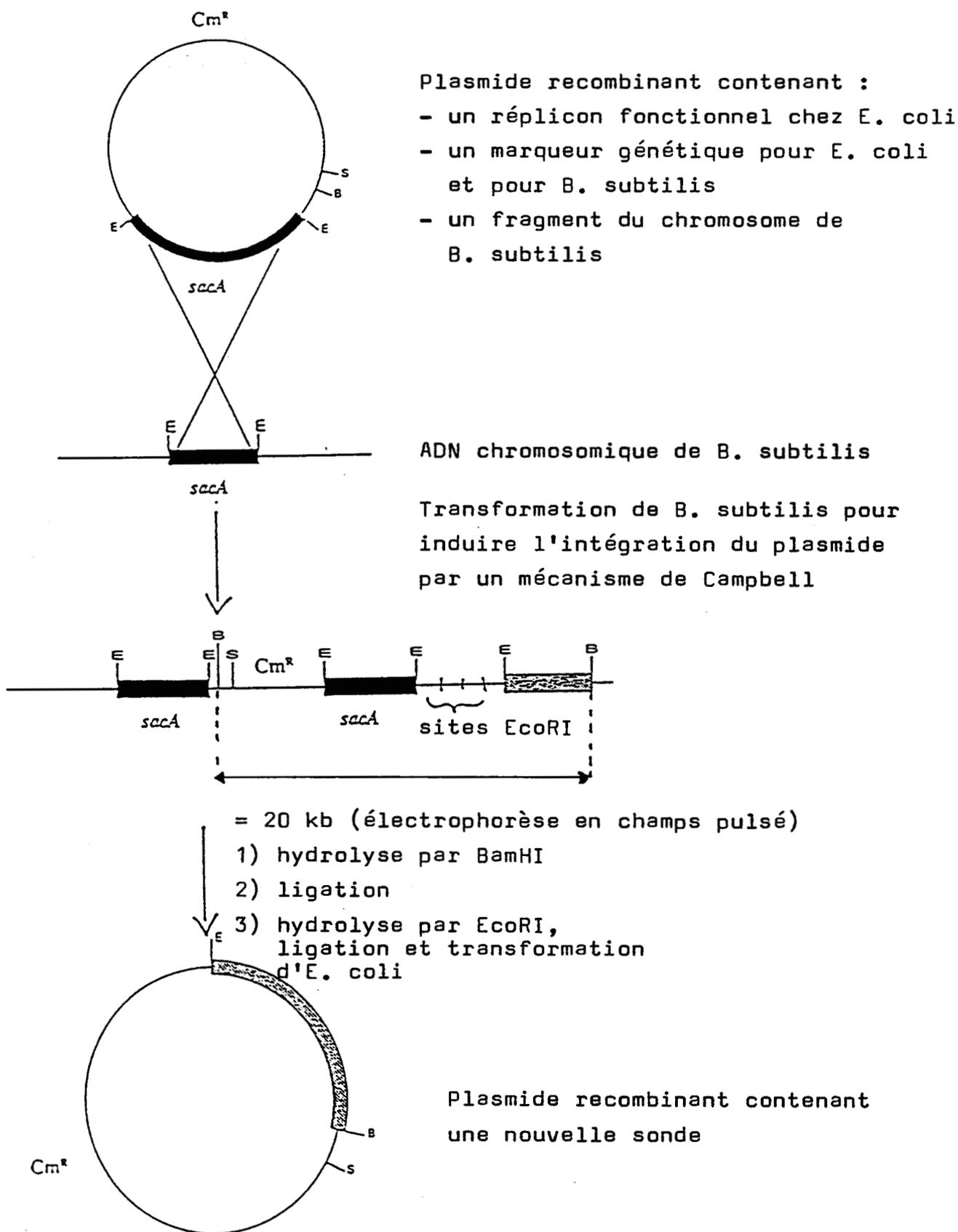
d'une souche thermosensible d'E. coli dont le phénotype est dû à une déficience en tyrosine tRNA synthétase à 44°C (Glaser, Unité de Régulation de l'Expression Génétique). L'hypothèse avancée par Aymerich et Steinmetz (1987) que cette région du chromosome de B. subtilis n'est pas indispensable a été vérifiée en ce qui concerne le gène *yrs*. Le plasmide pMF2 a été linéarisé au site unique XbaI situé dans la séquence codante du gène *yrs*. Un fragment de 1,5 kb contenant le déterminant aphA3 (Trieu-Cuot et al., 1983) de la résistance à la kanamycine a été purifié par électrophorèse en gel d'agarose suivie d'une purification du fragment à l'aide du kit Gene Clean. Les extrémités des sites XbaI et ClaI ont été converties en bouts francs à l'aide du fragment Klenow de l'ADN polymérase, puis le fragment aphA3 a été ligaturé au plasmide pMF2 en utilisant de la ligase du phage T4. Des cellules compétentes de la souche d'E. coli TG1 ont été transformées avec le mélange de ligation ainsi obtenu, les transformants ont été sélectionnés sur des boîtes LB supplémentées d'ampicilline et de kanamycine. L'un de ces transformants, pMF3, a été linéarisé au site SmaI et utilisé pour transformer des cellules compétentes de B. subtilis 168. Les transformants obtenus par un évènement de double cross over ( $Kan^R$ ,  $Cm^S$ ) ont le gène *yrs* interrompu par la cassette kanamycine. Le phénotype de l'un d'eux, QB4332, a été caractérisé ; il en résulte que le gène *yrs* n'est pas indispensable dans des conditions standard de laboratoire étant donné qu'aucun phénotype relatif à une thermosensibilité quelconque, à la sporulation, à la compétence, ou à des auxotrophies n'a pu être mis en évidence.

## **E) DISCUSSION GENERALE**

Un projet de séquençage exhaustif d'un grand fragment de chromosome d'un organisme rencontre un obstacle (le coût) ainsi que deux principales étapes critiques, à savoir le sous-clonage ordonné de la région considérée en des fragments dont la taille est compatible avec une approche de séquençage en *shotgun*, et la détermination de la séquence nucléotidique.

En ce qui concerne le sous-clonage en fragments ordonnés, le choix du vecteur  $\lambda$ FixII apparaît approprié étant donné la facilité de criblage d'une banque de phages ainsi que la taille des fragments obtenus (de 15 à 20 kb) qui permettent de limiter le nombre de sous-clonages nécessaires. Dans un souci de réduire au minimum les coûts d'un projet de cette envergure, il apparaît fondamental d'optimiser chaque étape. Le gain de rentabilité le plus évident semble consister en une simple réduction du temps nécessaire à l'obtention des résultats par économie du temps de travail. L'utilisation de la technologie de sondes froides permet d'optimiser l'étape de criblage étant donné la rapidité avec laquelle le résultat d'une expérience d'hybridation peut être obtenu : seulement 30 minutes d'exposition sont nécessaires en utilisant le système des sondes froides développés par BRL contre plus de 3 à 4 jours dans le cas d'une sonde marquée au  $^{32}\text{P}$ , et ce en conservant un niveau de sensibilité comparable.

Dans l'état actuel des techniques et en ayant recours à la technologie des sondes froides, un nouveau phage recombinant peut être aisément isolé en 2 à 3 semaines, moyennant la disponibilité d'une sonde de la région correspondante. Dans ces conditions, l'étape limitante d'un tel projet est la vitesse d'acquisition de la séquence nucléotidique elle-même. Etant donné le petit nombre de marqueurs connus sur le chromosome de *B. subtilis*, plusieurs approches complémentaires sont



**Figure 13** : La technique de saut sur le chromosome

utilisables pour obtenir des sondes : d'une part les techniques de saut sur le chromosome (cf. Figure 13), et la possibilité de cloner un fragment terminal d'un fragment cloné ou de synthétiser une sonde ARN à partir d'un phage isolé, et d'autre part l'utilisation des YACs identifiés par le groupe de Ehrlich (INRA, Jouy-en-Josas). Il est par exemple envisageable de cribler la banque de phages à l'aide d'une sonde réalisée à partir d'un YAC recouvrant la région étudiée, puis d'ordonner les phages ainsi identifiés et isolés d'après leur carte de restriction et les résultats d'expériences d'hybridation entre ces phages et différentes sondes ARN synthétisées à partir des promoteurs des ARN polymérases T3 et T7. Une telle approche est cependant assujettie à une bonne maîtrise de la technologie des YACs ainsi qu'à une méthode efficace d'isolement de l'ADN des YACs.

Une réserve doit être émise cependant en ce qui concerne l'utilisation des systèmes de sondes froides basés sur l'interaction avidine/ biotine : la forte interaction entre la streptavidine (ou l'avidine) et les protéines biotinylées (biotine liée de façon covalente) est dommageable étant donné l'existence d'un "bruit de fond" lors de l'étape de criblage de la banque dû à la présence de phages recombinants contenant des gènes de protéines biotinylées de B. subtilis. Ces phages sont d'autant plus apparents que la séquence recherchée est mal représentée, voire absente, de la banque. Il est en effet connu que, pour détecter chez les plantes ou les eucaryotes supérieurs, des carboxylases qui sont des protéines appartenant au groupe des protéines biotinylées, l'un des moyens les plus efficaces est l'utilisation de la très forte interaction entre la biotine et l'avidine (Bramwell, 1987 ; Haneji and Koide, 1989). En outre, certains tests de dosage de la biotine ont été construits sur la base de la très forte compétition pour la biotine entre les molécules d'avidine et les molécules de pyruvate carboxylase (Daunert et al., 1990). Une alternative consiste donc soit à

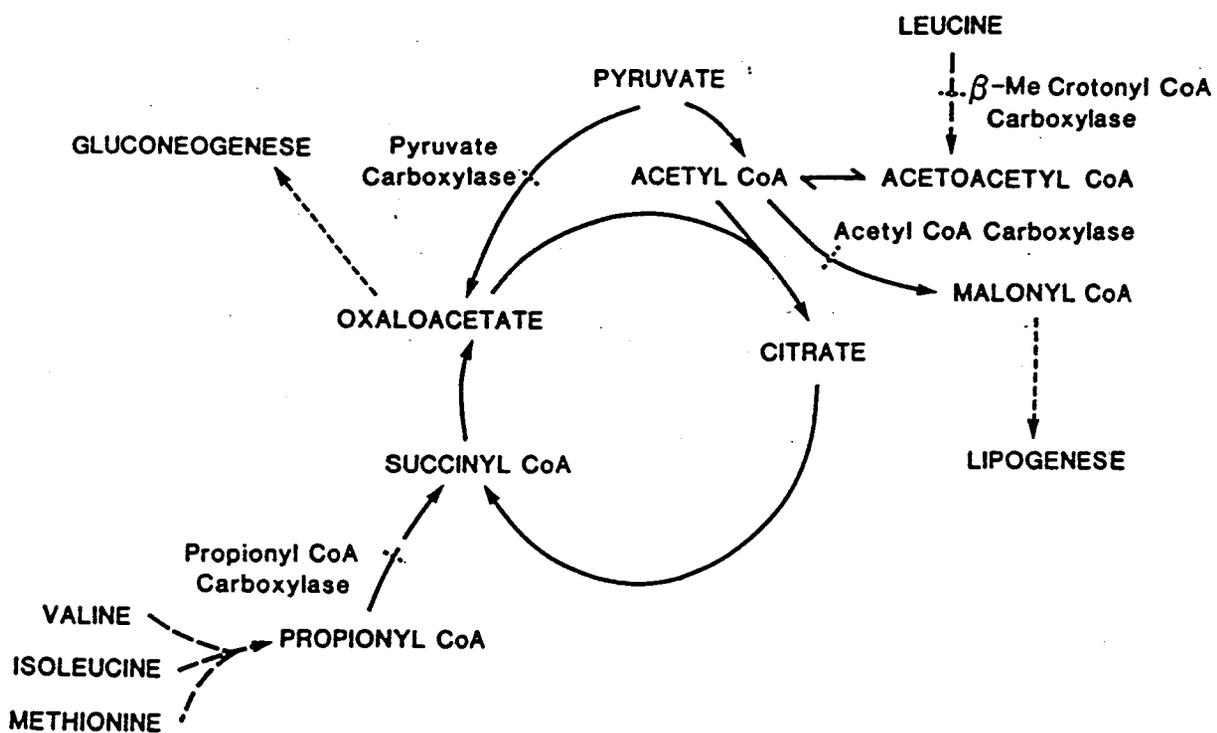
inclure un traitement à base de protéases (Zeph et al., 1991), soit à utiliser un système de sondes froides basé sur un autre type d'interactions. Un tel système est commercialisé par Boehringer Mannheim : la différence majeure réside dans le type d'interactions utilisées. Ce dernier système est en effet basé sur des interactions du type antigène/ anticorps. Les hybridations sont détectées dans les deux cas à l'aide d'une solution de PPD (4-methoxy-4(3-phosphatényl-spiro(1,2-dioxetane-3,2'-adamantane)). Il a du reste été rapporté qu'un autre composé, l'AMPPD (3-(2'spiroadamantane)-4-methoxy-4-(3''-phosphoryloxy)phenyl-1,2-dioxetane) (Tizard et al., 1990) permettrait d'obtenir un niveau de sensibilité 10 fois supérieur à celui offert par le PPD.

Il est également notable que la marche sur le chromosome est en pratique assez laborieuse, comme illustré par les difficultés rencontrées pour tenter de cloner un phage chevauchant les phages  $\lambda$ thiC et  $\lambda$ sacPT. En effet, il apparaît que dans tous les cas où de 5 à 10 phages réagissant avec la sonde employée ont été purifiés, les profils de restriction de ces phages étaient relativement identiques. Ceci suggère qu'un biais existe dans la majorité des cas pour la sélection d'un même "type" de phage contenant un *locus* particulier. Une explication plausible pour rendre compte de ce phénomène est que lors de l'amplification de la banque après l'étape de ligation, certains phages se répliquant moins bien du fait de la présence dans l'insert chromosomique de B. subtilis de gènes toxiques deviennent sous-représentés dans la banque. L'utilisation de YACs ainsi que la pratique des sauts sur le chromosome semble donc être des stratégies nettement plus prometteuses que la marche sur le chromosome. Les problèmes causés par les régions plus difficiles à cloner devraient être résolus à l'aide de vecteurs endogènes à B. subtilis, tels que des plasmides réplicatifs ou des phages comme  $\Phi$ 105. Une autre possibilité consiste à utiliser des techniques de PCR telles que les méthodes dites de "chimie génétique" qui

ne nécessitent la connaissance que d'un seul oligonucléotide (Copley et al., 1991). Ce type d'approche est de plus facilité par la possibilité d'utiliser une enzyme thermostable de Thermococcus litoralis (Neuner et al., 1990) dont la particularité est une fonction d'exonuclease 3'-5' qui rend cette enzyme plus fidèle que la polymérase Taq (cf. New England Biolabs).

Par ailleurs, le séquençage en *shotgun* de la totalité de l'insert chromosomique d'un phage recombinant (de 15 à 20 kb) sans autre sous-clonage que le sous-clonage dans le phage M13 apparaît réalisable dans un délai relativement court ( $\approx$  6 mois). Il est clair qu'à ce stade le problème posé par la lecture des gels et surtout par la correction des erreurs sont les problèmes majeurs et constituent en fait l'étape limitante car la plus laborieuse du processus d'acquisition d'une séquence nucléotidique. C'est donc particulièrement à ce niveau que la technologie doit évoluer afin de minimiser les coûts des grands programmes de séquençage. Le développement de lecteurs automatiques de gels devrait rapidement offrir une telle possibilité, mais il semble également indispensable de pouvoir disposer d'une banque d'images (vidéodisque) reliée à un programme d'exploitation permettant de visualiser rapidement les différents gels et de localiser automatiquement sur ceux-ci les régions à vérifier.

Comme indiqué précédemment, le clonage du gène d'une protéine biotinylée de B. subtilis a permis de mettre en évidence les limites des systèmes de sondes froides basés sur l'interaction biotine/ avidine quant au criblage de banques. Il est en outre intéressant de constater la présence d'une telle protéine chez B. subtilis, en particulier pour l'étude de l'Evolution. Il est en effet bien établi que les enzymes biotinylées se rencontrent chez de très nombreux organismes (Murtif and Samols, 1987 ; Samols et al., 1988) où elles jouent des rôles essentiels dans le métabolisme (cf. Figure 39) (Murtif and Samols, 1987 ; Knowles, 1989). Le site



**Figure 39** : Les enzymes biotinylées (carboxylases) dans le métabolisme cellulaire.

(d'après Dakshinamurti et coll., 1986)

d'attachement de la biotine (Met-Lys-Met) est invariant dans toutes les enzymes étudiées à ce jour ; en outre, Browner (1989) a observé une conservation absolue de la séquence Gly-X5-Gly-X7-Ala-Met-Lys-Met-X9-Gly-X-Val dans 8 enzymes biotinylées sur 10 provenant de bactéries, de plantes, de levures, et de mammifères (Maloy et al., 1979 ; Takai et al., 1987 ; Browner et al., 1989). Une telle conservation des séquences sur une large distance évolutive, conjuguée avec d'autres homologues frappantes observées dans d'autres régions que celles jouxtant la lysine biotinylée (Samols et al., 1988), suggère que ces enzymes dérivent d'un ancêtre commun qui s'est adapté à une variété de fonctions métaboliques dont l'invariance du mécanisme d'action a résulté en la conservation de nombreuses séquences critiques sur de larges distances évolutives (Samols et al., 1988).

La découverte d'une nouvelle protéase est riche de potentialités biotechnologiques, mais l'étude des propriétés physiques de cette protéine est nécessaires pour établir les possibles usages de cette protéase. En tous les cas, la souche de B. subtilis dont tous les gènes connus de protéase ont été délétés y compris celui-ci (7) trouvera certainement des applications industrielles et de recherche fondamentale.

La découverte d'un gène codant pour une tyrosine tRNA synthétase secondaire est importante d'un point de vue fondamental étant donné que très peu de tRNA synthétases de B. subtilis ont été clonées (Chow and Wong, 1988 ; Brakhage et al., 1989 ; 1990 ; Putzer et al., 1990) ; ce résultat permet de confirmer que la présence de deux gènes différents codant pour un même type de tRNA synthétase n'est pas un phénomène unique chez les procaryotes.

En effet, deux gènes indépendants codant pour un même type de tRNA synthétase ont été isolés dans

plusieurs cas tels que B. subtilis où 2 threonyl tRNA synthétases (Putzer et al., 1990), ou chez E. coli où 2 lysyl tRNA synthétases (Hirshfield et al., 1984) ont été identifiées. De plus, la présence dans la souche de B. subtilis 168 d'un supprimeur qui décode le codon stop UGA en tryptophane suggère l'existence d'une tryptophanyl tRNA synthétase secondaire qui décoderait les codons UGA et UGG (Lovett et al., 1991). La présence du gène secondaire de la lysyl tRNA synthétase chez E. coli a pu être reliée à une réponse au stress d'une élévation brutale de température ; par contre, une telle relation n'a pu être mise en évidence en ce qui concerne le gène secondaire de la threonyl tRNA synthétase de B. subtilis. Une hypothèse envisageable pour l'utilité physiologique de ce gène se situe peut-être au niveau de la phase de différenciation de B. subtilis. Il est également intéressant de noter que simultanément le gène principal de la tyrosine tRNA synthétase de B. subtilis a été séquencé par Grundy et Henkin (1990).

On peut encore remarquer que l'utilisation des techniques d'hybridation à l'aide de sondes hétérologues (Putzer et al., 1990) ou de séquençage exhaustif (Glaser et al., 1991) constituent une approche radicalement différente des techniques classiques de génétique, et offrent par là-même des résultats plus complets car moins ciblés (biaisés) que ceux fournis par ces dernières. Il est donc possible d'identifier des gènes qui ont échappé au crible des techniques de complémentation.

La localisation sur la carte génétique ainsi que la détermination de l'intégralité de la séquence nucléotidique de la sous-unité F<sub>1</sub> du complexe enzymatique ATP synthase de B. subtilis est un résultat important non seulement d'un point de vue de la recherche fondamentale, mais aussi d'un point de vue biotechnologique. En effet, l'ATPase est une enzyme qui catalyse l'une des réactions les plus importantes pour la

cellule, à savoir la synthèse de l'ATP. En ce qui concerne les bioconversions basées sur l'utilisation de B. subtilis, on peut encore noter que le contrôle de la lyse cellulaire et l'énergétisation de la membrane sont intimement couplés (Joliffe et al., 1981). Comme illustré au paragraphe précédent, les gènes des sous-unités  $b$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$  des ATPases de B. subtilis et B. megaterium apparaissent très fortement homologues, non seulement au niveau de la séquence peptidique, mais encore au niveau de l'organisation structurale de ces deux opérons. La détermination de la séquence nucléotidique des autres sous-unités de Fo ( $i$ ,  $a$ ,  $c$ ) et du promoteur (clonés par Glaser, cf. Figure 38) permettra de compléter cette comparaison. En outre, cet opéron offre la possibilité d'étudier la régulation traductionnelle chez B. subtilis. En effet, le taux de synthèse de chaque sous-unité de l'ATPase apparaît chez E. coli corrélé à la stoichiométrie par des mécanismes qui contrôlent l'efficacité de l'initiation de la traduction et de la réinitiation de la traduction (Brusilow et al., 1982 ; McCarthy et al., 1985). En particulier, une séquence en amont du gène de la sous-unité  $c$  de l'ATPase Fo d'E. coli ( $5'$ TTTACCAACACTACTACGTT-TTAAACTGAAACAAA $3'$ ) permet d'accroître l'expression du gène *atpE* (sous-unité  $c$ ) par l'accroissement de la fréquence d'initiation de la traduction de ce gène (séquence *enhancer*) (McCarthy et al., 1985). Une séquence similaire ( $5'$ TTTGGACA-AAACAAACTTAGTTAAACAAA $3'$ ) a du reste été identifiée en amont du gène *atpE* de B. megaterium (Brusilow et al., 1989). Il est d'ailleurs connu que l'altération de la région d'initiation de la traduction peut influencer significativement le taux de synthèse d'une protéine, et que les séquences *enhancer* se rencontrent fréquemment dans les régions d'initiation de la traduction des gènes fortement exprimés (cité par McCarthy et al., 1986). En ce qui concerne l'utilisation de tels signaux, l'addition de la séquence *enhancer* du gène *atpE* d'E. coli dans un vecteur d'expression d'E. coli a permis d'augmenter la production d'interleukine-2 et d'interféron- $\beta$  humains par

un facteur de 6 à 10, alors que l'usage du code de ces deux gènes est similaire à l'usage du code des gènes peu exprimés chez E. coli (McCarthy et al., 1986). Il serait donc très utile pour la biotechnologie de B. subtilis de déterminer la séquence *enhancer* du gène *atpE* de B. subtilis, si elle existe, afin de l'insérer dans un vecteur de sécrétion. Parallèlement, il serait intéressant de déterminer les mécanismes de régulation traductionnelle qui existent chez B. subtilis. La connaissance de la séquence nucléotidique de l'ATPase de B. subtilis permettra en outre de construire et d'étudier l'assemblage et le fonctionnement de molécules hybrides d'ATPases en testant séparément chaque région individuelle de chaque sous-unité. Ce résultat permettra également de faciliter les études biochimiques en autorisant la surproduction individuelle de chaque sous-unité, quoique de bonnes connaissances biochimiques concernant l'ATPase de B. subtilis soient déjà disponibles (Serrahima-Zieger and Monteil, 1978 ; Hicks and Krulwich, 1987). Les masses molaires obtenues au moyen des données biochimiques sont du même ordre que les masses molaires déduites de la séquence nucléotidique (sous-unité  $\alpha$  : de 49300 à 51100 et 54504, sous-unité  $\beta$  : de 49300 à 54500 et 51362, sous-unité  $\gamma$  : de 30700 à 35300 et 31601, sous-unité  $\delta$  : de 21100 à 24000 et 19959, sous-unité  $\epsilon$  : de 18600 à 21400 et 14193, respectivement) (Hicks and Krulwich, 1987). Il a en outre été démontré que les sous-unités  $\alpha$  et  $\beta$  de B. megaterium sont capables de compléter les mutants *uncA* et *uncD* d'E. coli (Hawthorne and Brusilow, 1986). Il serait également intéressant de cartographier les différents mutants *unc* de B. subtilis (Guffanti et al., 1987 ; Hicks and Krulwich, 1987) et de comparer leur position avec la position de l'opéron *atp* (entre 320° et 317°). Les séquences nucléotidiques de l'ATPase de B. subtilis ont en outre l'avantage de faciliter l'étude des ATPases d'autres bactéries proches telles que Bacillus firmus (Hicks and Krulwich, 1986 ; Ivey et al., 1990) ou Streptococcus mutans (Quivey et al., 1991), sans

compter que les ATPases sont des outils aussi puissants que les ARN 16S pour une étude globale de l'évolution, étant donné leur fort degré de conservation.

Le fait que le gène *spoIID* soit situé à proximité du gène *narA* souligne le fait que la carte génétique peut comporter des erreurs importantes et qu'il est donc nécessaire de vérifier la position des phages isolés par rapport à plusieurs marqueurs différents, afin d'éviter de séquencer par mégarde plusieurs fois le même segment de chromosome.

L'identification d'une ORF (PROT R) dont le produit est identique à 50% à l'ORF R située à proximité de *spoOF* est également un résultat intéressant. En effet, cette dernière phase ouverte de lecture a été identifiée comme étant le locus *rev-4* qui est un suppresseur pléiotrope pour les déficiences de sporulation de certaines mutations ribosomales induites par l'erythromycine (Sharrock and Leighton, 1982 ; Trach et al., 1988). L'interruption de l'ORF R n'a pas d'effet phénotypique, ce qui suggère que son produit n'est pas nécessaire pour la croissance ou la sporulation. Il serait intéressant de vérifier si l'ORF PROT R peut de même être interrompue sans dommage majeur pour la cellule. En outre, il serait utile de déterminer la fonction de tels gènes, et éventuellement d'identifier leurs promoteurs et leur expression au cours des différentes phases de croissance.

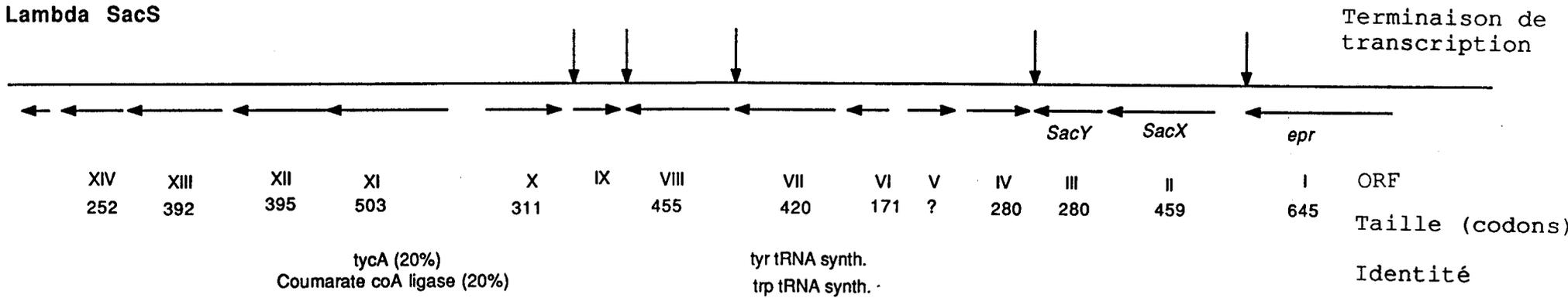
Le séquençage de la sous-unité  $\gamma$  et la localisation de l'opéron uréase de *B. subtilis* sont aussi des résultats nouveaux. En effet, comme décrit précédemment, les uréases sont des enzymes importantes pour la médecine, l'élevage et l'agriculture. En outre, la forte identité qui existe entre la sous-unité  $\gamma$  de l'uréase de *B. subtilis* et celle d'*U. urealyticum* souligne le fort degré de parenté qui existe

entre ces deux organismes et le fait que la connaissance de la séquence nucléotidique du génome de B. subtilis devrait faciliter grandement l'étude et la compréhension de la biologie des Mycoplasmes, comme illustré par la découverte d'un suppresseur du codon stop UGA chez B. subtilis 168 (Lovett et al., 1991) et le décodage de ce codon en tryptophane chez U. urealyticum (Blanchard, 1991).

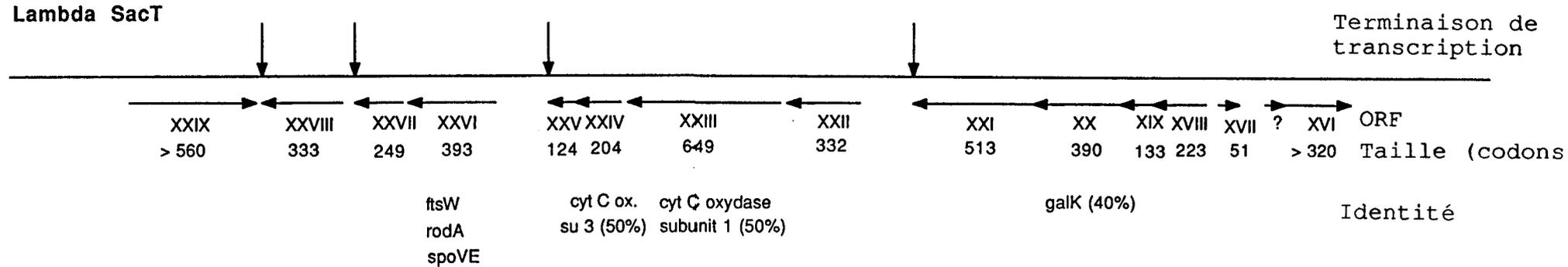
La région située aux alentours de 320° apparaît de plus très riche en gènes impliqués dans le métabolisme de l'azote, à savoir le gène *narA*, le gène de l'uréase, et le gène *nrg29* (Atkinson and Fisher, 1991). Le séquençage de cette région semblerait donc être prometteur pour l'étude de la régulation du métabolisme de l'azote chez B. subtilis.

Les caractéristiques des phases ouvertes de lecture (ORF) codant pour les protéines putatives déduites des séquences nucléotidiques obtenues jusqu'à présent dans l'Unité de Biochimie Microbienne et dans l'Unité de Régulation de l'Expression Génétique, dans le cadre du projet de séquençage du génome de B. subtilis, sont reportées Figure 37 et 38. Il est intéressant de souligner que parmi les 44 ORFs identifiées dans les phages  $\lambda$ sacS,  $\lambda$ sacPT, et  $\lambda$ Eco-SalI (séquencés par Hullo, Marcellino, Santana, Schneider, Schweitzer, et Glaser), 4 correspondent à des gènes connus, 13 correspondent à des gènes homologues à d'autres gènes déjà séquencés, et surtout 28 (65%) ne correspondent à aucun gène connu. Parmi ces phases ouvertes de lecture, certaines ont des sites RBS qui semblent insuffisants pour assurer la traduction chez B. subtilis, tels que les sites RBS des ORFs XIX et XXII (cf. Tableau 22). Il devrait donc être particulièrement instructif d'étudier l'expression de ces ORFs, à la fois chez E. coli et chez B. subtilis. D'autre part, 17 terminateurs de transcription ont été identifiés au travers des premiers résultats du programme de séquençage du génome de B. subtilis, ce qui devrait permettre, en conjonction avec les données déjà accumulées,

**Lambda SacS**



**Lambda SacT**



**Figure 37 : Carte des phages recombinants  
lambdasacS et lambdasacT**

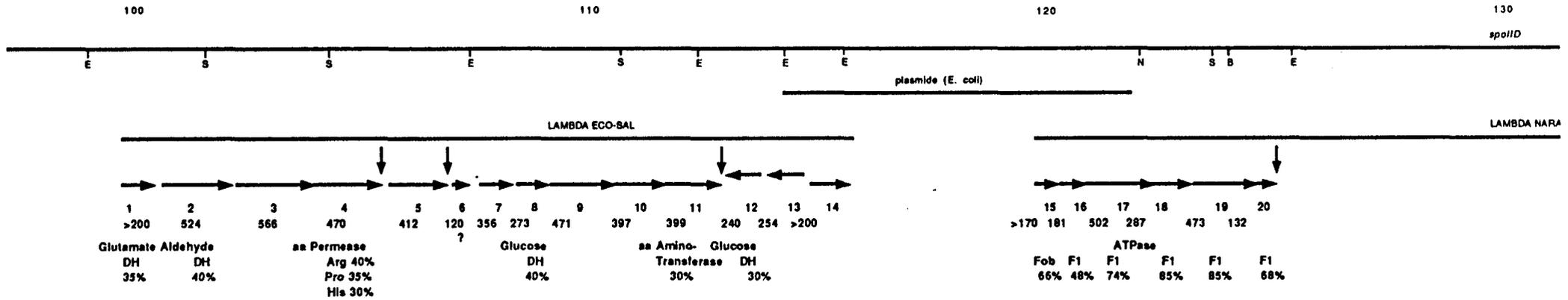


Figure 38 : Carte du phage lambdaEco-Sal et localisation du phage lambda narA

RBS DES ORFs IDENTIFIEES DANS LE PHAGE  $\lambda$ sacPT

| ORF    | Séquence nucléotidique des ARNm                |
|--------|--|
| XVI    | CUUU <u>AAAAGAGG</u> AUCUGAAGGGUGAAC           |
| ?      | GGGAAU <u>AAAAGGAGGA</u> AAGCCAUGUAC           |
| XVII   | UUUAUACA <u>AUUGGAGGUU</u> CUAUGAAA            |
| XVIII  | CAUUACCGCC <u>AGGAGAAA</u> AUAUGAAA            |
| XIX    | GUGAAGU <u>AUAGAG</u> AAUAAUUAUGU <u>AU</u>    |
| XX     | AAACCAG <u>AAAGGUGAU</u> UGCACAUGCGG           |
| XXI    | GAACU <u>AAAGGGGGAG</u> UGACAU <u>AUGAGU</u>   |
| XXII   | GAGAAAAGAUUUUG <u>GAGGAAGU</u> AUGCAC          |
| XXIII  | UCAGG <u>AAAAGGAGGA</u> AAGCCUGCAUGAAA         |
| XXIV   | CCGAAU <u>AAGGAGGCG</u> UGAGUUAUGGAA           |
| XXV    | UGGG <u>AUUGGGGGG</u> UCUGUAAA <u>AUGGCA</u>   |
| XXVI   | GAAA <u>AUGUAAGGCCGG</u> GAUAGAA <u>AUGAGU</u> |
| XXVII  | AGUGAU <u>AAGGAGGU</u> UUUUGUGAUGAAU           |
| XXVIII | GAUGAGG <u>AUGAGGUG</u> GUAGGG <u>AUGGUU</u>   |

Tableau 22

de mieux comprendre les mécanismes de la terminaison de la transcription chez B. subtilis. Un autre fait important est qu'environ 75% de la séquence obtenue est codante. Le nombre élevé des phases ouvertes de lecture identifiées implique donc par lui-même que l'exploitation totale de la séquence du génome de B. subtilis nécessitera un temps notablement plus long que le temps nécessaire à l'accumulation des données brutes (séquençage). D'autre part, la taille moyenne d'un gène chez B. subtilis est de 1 kb avec un écart type de 0,4 kb, ce qui confirme bien que la taille moyenne d'un gène chez les procaryotes est d'un millier de paires de bases ; l'originalité de ce résultat étant qu'il a été acquis à partir d'un ensemble de gènes (presque) contigus sur le chromosome.

**F) ARTICLE 4**

P. GLASER, F. KUNST, M. DEBARBOUILLE, A. VERTES, A. DANCHIN, and R. DEDONDER. 1991. A GENE ENCODING A TYROSINE tRNA SYNTHETASE IS LOCATED NEAR SACS IN *BACILLUS SUBTILIS*. DNA Sequence 1:251-261

# A gene encoding a tyrosine tRNA synthetase is located near *sacS* in *Bacillus subtilis*

P. GLASER†, F. KUNST††, M. DÉBARBOUILLÉ††, A. VERTÈS††,  
A. DANCHIN† and R. DEDONDER††

† Unité de Régulation de l'Expression Génétique, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France

†† Unité de Biochimie Microbienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France

EMBL Data Library Accession No. X52480.

Within the frame of an attempt to sequence the whole *Bacillus subtilis* genome, a region of 5.5 kbp of the *B. subtilis* chromosome near the *sacS* locus has been sequenced. It contains five complete coding sequences, including the sequence of *sacY*, three unknown CDS and a sequence coding for a tyrosine tRNA synthetase. That the corresponding CDS encodes a functional synthetase has been demonstrated by complementation of an *Escherichia coli* mutant possessing a thermosensitive tRNA synthetase. Insertion of a kanamycin resistance cassette in the *B. subtilis* chromosome at the corresponding locus resulted, however, in no apparent phenotype, demonstrating that this synthetase is dispensable. Finally phylogenetic relationships between known tyrosine and tryptophan tRNA synthetases are discussed.

**Key words:** large genome sequencing, tryptophan tRNA synthetase, aminoacyl adenylate, unknown reading frames

## INTRODUCTION

Autonomy of living systems is the consequence of internal self consistency of their genetic program. Taken together the rules specified by the DNA sequence decide of survival and reproduction of all organisms. This corresponds to a program which is finite in length. All "rewriting" rules that interpret the genetic program (transcription and translation) and fix the actual structure of all effectors of metabolism are totally included in the corresponding nucleotide and amino acid sequences. Up to now the self consistent feature of genomes has remained an inaccessible information, and it has not been possible to explore exhaustively the very nature of signals that limit either

the diachronic or structural building, or the actual functioning of the cell.

New techniques in DNA sequencing, as well as improvements in information handling already permit to reconsider this situation. Total sequencing of bacterial genomes is certainly accessible to experimentation although DNA sequencing of whole genomes will require a vast amount of work, especially if one hopes to investigate the structure of higher eukaryotic genomes. It seems therefore of the utmost importance to organize and make proper choices for a productive strategy. It is clear that, as such, the sequence of a genome of a single entity will be of very high basic value. Indeed, at the simplest level, sequencing will make available all signals that are important for the core machinery, responsible for gene expression and replication. A further important result will be that we shall have access to information corresponding to the ecological niche of the organisms considered. Indeed, the DNA sequence information that remains present after the core sequences have been subtracted, represents information specific to each type of organism.

*B. subtilis* is an organism of choice for total genome sequencing and several laboratories in Europe have agreed upon a cooperation meant to yield the total sequence in a decade or so. Here we report the first 5.5 kb of the *gerB-sacS* DNA region located at 315°–335° on the *B. subtilis* chromosomal map (Piggot, 1989). In order to start from a known region, a clone containing the *sacS* locus was first isolated and partially sequenced (Débarbouillé *et al.*, 1987; Zukowski *et al.*, 1988). This locus is composed of two genes, *sacX* and *sacY*. Both genes encode regulatory

Abbreviations: CDS, Coding Sequence; ORF, Open Reading Frame; WRS, Tryptophan tRNA Synthetase; YRS, Tyrosine tRNA Synthetase

proteins involved in the induction mechanism of levansucrase (Steinmetz *et al.*, 1988, Zukowski *et al.*, 1988). As shown below, this DNA region harboured a gene encoding a protein having all features of a tyrosine tRNA synthetase.

## RESULTS

### Analysis of the region adjacent to *sacY*

Figure 1 displays the nucleotide sequence of the region located downstream from the *sacY* gene. ORFs that might encode proteins have been analyzed using the RNY best fit (Shepherd, 1981), and checked for enrichment in A+T third base composition as compared to overall base composition. This revealed the presence of five complete ORFs containing potential coding sequences (CDS) and two partial ORFs present at the extremities of the fragment. The region upstream from the *sacY* ORF (393–1232) (identified as identical with the published sequence of the gene (Zukowski *et al.*, 1988)) corresponds to a truncated ORF, which is part of *sacX* (1–336). These ORFs are followed by an inverted repeat (1243 to 1277) that may indicate transcription termination. More downstream two ORFs, ORF1 (from 2969 (start codon TTG) to 2259) and ORF2 (from 2128 (start codon ATG) to 1289), are prominent on the complementary strand. Then follows ORF3 having the same orientation as *sacY*. It corresponds to a putative protein of 171 amino acid residues starting at position 3120 and ending at position 3632. The polypeptide sequences predicted from these three ORFs have no feature that could be related to known gene products present in the EMBL data library (release 20).

From position 3887 to 5122, CDS4 encodes a protein having significant similarities with tyrosine tRNA synthetase from several organisms (as discussed below). It is bordered downstream by a hairpin structure that may behave as a transcription termination signal. Further downstream a CDS starts at position 5282. It codes for a protein starting with a typical signal sequence (Watson, 1984). The upstream DNA sequence (termed *tyrS1*) does not seem to display highly significant transcription initiation regions, at least in the present state of knowledge of *B. subtilis* promoters, but only a well preserved ribosome binding site at position 3875.

It may be significant that from position 4190 to 4211 is a sequence that is complementary to tRNA<sup>tyr</sup> from *E. coli*, just downstream from a region that

seems to be duplicated both at the nucleotide and aminoacid residues levels (GDFTGK/GDPTGK).

### CDS4 Encodes a dispensable protein endowed with YRS activity

The *KpnI-PstI* (fig 1) DNA fragment from plasmid pMF2 was subcloned into the expression vector pTZ18 (Mead *et al.* 1986). Overexpression of CDS4 allowed the growth of strain HB2109 at 44°C. Strain HB2109 has a thermosensitive growth phenotype due to a defective YRS activity at high temperature (Bedouelle and Winter, 1986). This demonstrated that CDS4 encoded an active YRS. Aymerich and Steinmetz have suggested that the corresponding region of the *Bacillus subtilis* chromosome was dispensable (Aymerich and Steinmetz, 1987). Since there is an absolute requirement of a YRS activity for growth it was important to substantiate their preliminary data. The DNA segment carrying *tyrS1* was interrupted by a Kanamycin resistance cartridge at the unique *XbaI* site (fig 1). The modified strain displayed no identifiable phenotype with respect to thermosensitive growth, auxotrophies, sporulation or competence. These results show that CDS4 product has a YRS activity which is dispensable for *B. subtilis* under usual laboratory growth conditions.

### Tyrosine tRNA synthetases and related synthetases

The sequence of CDS4 (*tyrS1*) was used to scan the protein data libraries present at the Unité d'Informatique Scientifique, using the FASTP software. This permitted us to discover that it was related to tyrosine tRNA synthetases (YRS) from several organisms, including *E. coli* and *Bacillus stearothermophilus*. Alignment of all YRS present in the banks was then performed manually using the consensus KMSKS amino-acyl-AMP binding site as an anchoring site. A second consensus, HIGH, was used as a secondary anchor to align the aminoterminal of the protein. This permitted identification of several identity regions. Namely, the NH<sub>2</sub> terminal domain is well preserved, whereas the COOH distal part of the protein is less preserved (see fig 2). In particular, it is worth noting that the *B. subtilis* putative YRS is more similar to that of *E. coli* than to that of known YRS from *Bacilli*. It was then possible to scan the sequence of this synthetase against those of all synthetases present in the libraries. Another noteworthy observation was that *B. subtilis* tryptophan tRNA synthetase (WRS) seemed clearly related to all YRSs. Figure 2 displays the sequence of the *tyrS1* gene product as compared with other tyrosine tRNA synthetases. Similarities are

```

V F L K T K Q S S L K K I A L P A S L T A F L G I V E P I V F G V N L K L I R P
GTATTCCTCAAACGAAACAATCAAGCCTGAAGAAAATTGCGCTCCCCGCTTCTCTGACCGCGTTTTAGGCATTGTTGAGCCGATTGTATTCGGGGTAAATCTCAAATGATCCGTC
10      20      30      40      50      60      70      80      90      100     110     120

F I G A A I G G A I G G A Y V V A V Q V V A N S Y G L T G I P M I S I V L P F G
TTTATCGGAGCAGCCATCGGCGGTGCTATTGGCGGAGCATTACGTGGTTGCGGTACAAGTTGTGGCAAATTCCTACGGACTGACAGGCATCCGATGATTTCAATCGTGCTGCCGTTGGC
130     140     150     160     170     180     190     200     210     220     230     240

A A N F V H Y M I G F L I A A V S A F I A T L F L G F K E E T E *
GCCGCTAATTTGTTCAATATATGATCGGTTTCTTGATCGCAGCCGCTCTGCTTTTATAGCTACATGTTTCTCGGGTTTAAAGAAGAGACAGAATAACTGGATTTATCGATTTCAAT
250     260     270     280     290     300     310     320     330     340     350     360

M K I K R I L N H H A I V V K D Q N E E K I L L G A G I A F
CATAAAACGGGGGATGAAAGGACAAAAAGCTATGAAAATTAAGAATCTTAAATCATAATGCTATCGTCGTAAGGATCAAATGAAGAGAAGATTCTCTGGGTGCAGGAATTGCGT
370     380     390     400     410     420     430     440     450     460     470     480

N K K K N D I V D P S K I E K T F I R K D T P D Y K Q F E E I L E T L P E D H I
TTAACAAAAGAAGAATGATATTGTCGATCCGTCAAAAATAGAAAAACCTTTATCAGAAAAGATACACCTGACTATAAGCAGTTCGAAGAGATTTAGAAACATTGCCGTAAGACCACA
490     500     510     520     530     540     550     560     570     580     590     600

Q I S E Q I I S H A E K E L N I K I N E R I H V A F S D H L S F A I E R L S N G
TTCAGATTTCTGAGCAAATTATCTCTCATGCCGAAAAAGAGCTGAACATCAAATCAACGAGCGCATTTCATGTCGCTTTTTCAGACCATCTTTCTTTGCAATTGAACGCTGAGCAATG
610     620     630     640     650     660     670     680     690     700     710     720

M V I K N P L L H E I K V L Y P K E F Q I G L W A R A L I K D K L G I H I P D D
GGATGGTTATCAAATCCGCTGCTGAATGAAATCAAAGTCTTTATCCAAAGGAGTTCAGATCGGCTTATGGGCCAGAGCACTGATTAAGATAAACTGGGGATTACATTCCTGATG
730     740     750     760     770     780     790     800     810     820     830     840

E I G H I A M H I H T A R N H A G D M T Q T L D I T T M I R D I I E I I E I Q L
ATGAAATCGGCAATATCGCCATGCATATCCACACAGCAAGAAACAATGCCGCGGATATGACACAACCGTTGATATTACAACAATGATCCGTGATATTATCGAGATTATCGAAATTCAC
850     860     870     880     890     900     910     920     930     940     950     960

S I N I V E D T I S Y E R L V T H L R F A I Q H I K A G E S I Y E L D A E M I D
TGCTATTAATATCGTTGAAGATACCATCTCTTATGAAAGGCTCGTGACCCATCCCGCTTTGCCATTGACATATCAAAGCAGGCGAATCCATTTACGAGCTGGACGCAGAAATGATTG
970     980     990     1000    1010    1020    1030    1040    1050    1060    1070    1080

```

Figure 1 Nucleotide sequence of fragment BSIP1. The translation of the putative CDS are displayed above the nucleotide sequence when this corresponds to translation of a messenger RNA having the same orientation. When translation proceeds from a messenger RNA corresponding to the complementary strand the aminoacid residues are placed beneath the nucleotide sequence and they are shown as underlined single letters. The first partial polypeptide sequence corresponds to the end of *sacX*. It is followed by *sacY*. Two CDS are then expressed, corresponding to the complementary strand. Then follows a CDS of unknown function, in the same orientation as *sacY*. The *YRS* CDS that has been described in the text is located further downstream, followed by the start of a protein likely to be secreted (a putative signal peptide is shown by arrows: two possible cleavage sites are indicated by brackets). A few other features are also indicated: a putative transcription termination structure (1242–1277); the RBS of *YRS* (likely to start at the second methionine codon); a short duplication (4150–4190) and a region complementary to tyrosine tRNA (4190–4211). In the absence of experimental data we have not shown sequences that are likely to be involved in transcription control. It is noteworthy, however, that the intergenic regions are rich in sequences of the form TGTG or CACA that are known to be present in protein-DNA binding regions.

I I K E K F K D A F L C A L S I G T F V K K E Y G F E F P E K E L C Y I A M H I  
 ACATCATTAAAGAGAAGTTAAGGATGCCTTCTGTGTGCCCTAAGCATCGGCACCTTTGTGAAGAAGGAATACGGCTTTGAGTTTCTGAAAAAGAATTGGCTACATCGCCATGCATA  
 1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200

Q R F Y Q R S V A R \*  
 TTCAGCGGTTCTACCAACGGTCAGTCGCACGCTGAGACAAAACAAAAACGCTTTTGATCATCTCAAAGCGTTTTTTTATCTGATTTATTGGTTGATCGCCGGATTCCCAATCGTATTCC  
 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320

ACTTCGTATAGGACATGTAATGGAAATACTCATCGCGATACGGATGTTTTGTATTGGAGTTCCAAGGCTTTTCTCCGCCGAGAAGTGCACGATGCGGGGTTTTCTCTGTTTCATTAT  
 K I Y S M Y H E Y E D R Y P H K I N S N W P K E G G C F H V I A P N E R I E N Y  
 1330 1340 1350 1360 1370 1380 1390 1400 1410 1420 1430 1440

ATTGCTCCGCCCTAACAGAGTCGATGGGGTTTTTCAGCTTAAACATGATATAGGTTTGAGCGTTCCAGCGGGATGAAGTTCATACCACTGATCATACAAGATCGCATTAGTGCATCTT  
 Q K R G L L I S P I K L K L M I Y I Q A N W R P H L E Y W Q D Y L I A N L A D Q  
 1450 1460 1470 1480 1490 1500 1510 1520 1530 1540 1550 1560

GGTCATGCAGCACAAGAAAGTTCCTGCTGGGTGTTTCATTGATAAAGTTGATGACTTTTTCTGTGATGTTTTGCTTTCTCCAAGACTCAAAATCAATAATCATGATACCTGAGTTAAAT  
 D H L V L E D E D P H E H I F H I V K E I I H Q K R W S E F D I I M I G S N F Y  
 1570 1580 1590 1600 1610 1620 1630 1640 1650 1660 1670 1680

ACTTCCCTGTATCAGTGACGTTCACTTCTTTTCAGGCTTCATGCTGCCCGCATCTCAACAGCGGCTACGGTGTATGGCGCAATGTCTAAGTCCCATAGCTTAGAAAATATCCTCTAGGA  
 K G I D I V H M E K L R E H Q G A D E V A A V I Y P A I D L D W L K S I D E L V  
 1690 1700 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800

CAAGCGCATCACAATCGATGTAATCATTCTGTTGATGCTTTCATCCTTAATTAAGTCGGGAATCGAAATGCGGTAAATACGGCGCTTTTGTAAATATGGCTGCTCAACCGCATGTTTCA  
 L A D C D I Y I M R K I S E D K I L D P I S I R Y Y A A K I I H S S E V A H E Y  
 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900 1910 1920

ACATGTTGGTATCCACTTCTAAAACTCAATCGGCACCTCCGAATTTCAAAGTGGTTTCTCCAGTCTTTTTTGTATCAGGCTTAATTCGCCATCAATGACGTATAATTTTACCTCTC  
 M H I D V E L F E I P V G F K L I I E E L R K K N D P K I G G D I V Y L K V E R  
 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020 2030 2040

TCTCTGATCCATGTTAGTCAATAAAGAAACGAACATCCACCTAAATGACGAGCATAATTATCATCTGCGCATGATACGATATGCATGATTTTCATCTTTCTCAATGGTATTCATCTCC  
 E Q D M N I L L S V F M G G L H R A Y N D D A C S V I H M  
 2050 2060 2070 2080 2090 2100 2110 2120 2130 2140 2150 2160

CTTTTGTGTACAAGTTACATACCCTTTTTCAAAAAATAAACACGTTCTCCGACATTTTTTTAAACAGAATGTGTATACAACAAAAAGGTGTCATTTTTTGAAGATGCGATAAAGC  
 2170 2180 2190 2200 2210 2220 2230 2240 2250 2260 2270 2280

CAATAAAGAAATGAAAAAGCTTGGGATTGTGATGGCTTCAGCCGAGACGATATACCAGGCCATGGCCCAGAAACATCTAAGAGAGAGGAACCGCCCGGTTTTTCTGTCAAATACATATAA  
 W Y L I F F S A I I I A E A S V I Y W P W P G L V D L L S S G G P K K I L Y M Y  
 2290 2300 2310 2320 2330 2340 2350 2360 2370 2380 2390 2400

TTCCGCCAAGCCAGCGATCAATCAGAAAAATACAGACCGCTACACATTAAACAAGCAGGACTGTACCACATAACGAGCGCTGGCCATTCTGTAGGTGCCAATCACAGCCATCAGCAGA  
 H A G L W R D I L F I C V G Y V H V L L V T V W L S R O G M R Y T G I V A M L L  
 2410 2420 2430 2440 2450 2460 2470 2480 2490 2500 2510 2520

Figure 1 continued

CAGGATAAAAACACAGACCCGTGAGAAATAAAAAAGAGGATATATCGGATATGCGAAATGAAAACATACCGAGGTCAGGAGTGGCAAGAGCCTGAATCGAACTGCCGATGCCGACAAAA  
C S L F V S G H S I F F L I Y R I H P F S F M G L D P T A L A O I S S G I G V F  
 2530 2540 2550 2560 2570 2580 2590 2600 2610 2620 2630 2640

AACAAAAACACAACAGCTTTTGCTTTTGTCTACTAACATAATAGCTGACAAATACACCGACAGGTCACTAAGCTGGAGAGGCAATGAAGTGGCACCACACCTGTCCGTGGCGACC  
F L F V F L K R S K I V L M I A S L Y V S L D S L Q L P L S T R V S W R D I A V  
 2650 2660 2670 2680 2690 2700 2710 2720 2730 2740 2750 2760

ATCCATATTTGATAGCCGATTTGCGAACCAAGCAAAAGAAACACAACAGAGAACGCAAAAACCGGCTGGCCCTCGGCTGCTTGACCTCATCTTGGAAACAAAACACAGAATGGCCAAA  
M H I Q Y G I O S G L L L F V F L S R L F R S A R P Q K V E D Q F L F L L I A L  
 2770 2780 2790 2800 2810 2820 2830 2840 2850 2860 2870 2880

AGGGAGATAATCGCTAACGCTACTACATGCTCTGTGAAAACAGATGAAACGGATCATGCTTATAGTCGGATTGAACATATTTTGCAAAACTGCACCCCGTTTTTCGGGTTTTTTGTTA  
L S I I A L I V V H E I S F L H F P D H K Y D S Q V Y K Q M  
 2890 2900 2910 2920 2930 2940 2950 2960 2970 2980 2990 3000

TTTTGGCGTTGTTTGGGCTTTTTTATTCAACCGATATATGGGCTTGAAGTTCACACCAGACTCTGTTAAAATTCGTTTCATAGCTAATGATTCGTATAATTTAAGGCAGGCCAAAAAA  
 3010 3020 3030 3040 3050 3060 3070 3080 3090 3100 3110 3120

H E R S H Q H Q Q L R K E E H D T L S K L K Q M P V E S L H L E A I S I A T H L  
 TGGAACGTTACATCAACATCAGCAGCTTCGGAAGGAAGAACACGATACACTTTCTAAGCTCAAACAAATGCCCGTCGAGTCACTCAACCTGGAAGCCATTCGATCGCTACGAATCTGT  
 3130 3140 3150 3160 3170 3180 3190 3200 3210 3220 3230 3240

Y R S A Q R L R V K M E T E V L S T Y N L S W T A F S I L Y D L W V H G A L E T  
 ATCGGCTGCACAAAGGCTCCGTGTCAAAAAGGAGACTGAGGTTCTTTCTACCTATAACCTTTCTTGGACCGCTTTCTCCATTCTTTATGATCTATGGGTATGGGGAGCGCTCGAAACGA  
 3250 3260 3270 3280 3290 3300 3310 3320 3330 3340 3350 3360

R K I A E L S G I S T A T A S H V I K T L E K K S F C R K S I D T R D R R L V F  
 GAAAAATCGCTGAGCTGTGAGGATATCTACAGCTACAGCAAGCAACGTGATCAAAACGCTCGAAAAAAGAGTTTTTGCCGTAAGCAATCGACACAAAGAGATCGGGACTCGTGTTTG  
 3370 3380 3390 3400 3410 3420 3430 3440 3450 3460 3470 3480

V S I T D S G K Q A I E E L Y P E F H K G E T E L I A G M T K D E Q K I L T G L  
 TCTCAATCACCGATTCCGGCAACAAGCGATCGAAGAATCTACCCGGAGTTTCACAAAGGGGAAACAGAGCTCATTGCGGGAATGACGAAGGACGAACAGAAAAATTGACAGGGCTGC  
 3490 3500 3510 3520 3530 3540 3550 3560 3570 3580 3590 3600

L R K V A D N L H T T \*  
 TTAGAAAGGTAGCTGACAATCTTACATAACATAAGCTTTGACGGGAATCCAGTAGCGGTTTGTCTCCCTGTTTCAGAGAGCCGGCGGTAGGTGCGAGCCGGCCACAGGCAACCGTGAA  
 3610 3620 3630 3640 3650 3660 3670 3680 3690 3700 3710 3720

TTACCTCCCAAGCTTTCTTTGTGAACCTTAGTAGCAAAAGACGGGAAACCGTTATTCGAATGAGCGGAAGACAATGCTTCAACCAGGGTGGTACCGCGTGATGAGCCACGCTCCCTT  
 3730 3740 3750 3760 3770 3780 3790 3800 3810 <KpnI> 3830 3840

RBS

———— (M) M R T F E Q L T A S Q Q K E V E R Q L Q L Y M T G  
 ATTGGGATGGGCTCTTTTGTGTTTGTAAAAAGGAGGAAAAATGATGAGAACATTTGAGCAGCTCACAGCGTCACAACAAAAAGAGGTAGAAGACAGCTTACGCTATACATGACAGG  
 3850 3860 3870 3880 3890 3900 3910 3920 3930 3940 3950 3960

A H E V I P P E E L K A K L V K S I S T G T P L K I K L G L D P S A P D V H L G  
 CGCCACGAAGTCATACCCCGGAGGAATTAAGGCAAGCTCGTGAATCAATTTCCACGGGCACGCCGCTTAAAATTAAGCTCGGATTAGATCCGCTGACCCGGATGACATTTGGG  
 3970 3980 3990 4000 4010 4020 4030 4040 4050 4060 4070 4080

Figure 1 continued

(—————) (—————)

H T V V L N K L R Q F Q E N G H I V Q L L I G D F T G K I G D P T G K S A A R K  
 CCATACGGTGTGTTAAATAAGCTTCGCCAATTTCAAGAAAACGGCCACATTTCCAGCTGTTAATGGGGATTCACAGGAAAAATGGTGATCCGACCGAAAAATCGGCAGCCAGAAA  
 4090 4100 4110 4120 4130 4140 4150 4160 4170 4180 4190 4200

Q L T D E E V Q H H A K T Y F E Q F G K V L D P E K V E L H Y N S K W L K T L N  
 GCAACTGACTGATGAAGAAGTTCAGCACAAATGCCAAAACCTACTTTGAGCAATTCGGAAAAGTCTTGATCCAGAAAAAGTCGAGCTTCACTATAACTCAAAATGGCTGAAAACATTGAA  
 4210 4220 4230 4240 4250 4260 4270 4280 4290 4300 4310 4320

L E D V I E L A G K I T V A R L M E R D D F E E R I A M Q K P I S L H E F F Y P  
 TCTAGAAGATGTCATTGAATTAGCAGGGAAAAAACGGTAGCCCGCTGATGGAGCGGACGACTTTGAAGAACGCATCGCCATGCAAAAACCAATCTCACTGCATGAATTCTTTTACCC  
 <XbaI>4330 4340 4350 4360 4370 4380 4390 4400 4410 4420 4430 4440

L M Q G Y D S V V L E S D I E L G G T D Q H F H V L M G R H F Q E R Y N K E K Q  
 ATTGATGCAGGGCTATGATTCTGTCGTTCTCGAAAAGTGATATTGAATTAGGCGGAACGGATCAGCATTTCATGTCCATGGGACGGCATTTCGAAGAACGATACAACAAAGAAAAACA  
 4450 4460 4470 4480 4490 4500 4510 4520 4530 4540 4550 4560

V V I L M P L L E G L D G V E K M S K S K H N Y I G I N E H P H D M Y G K T M S  
 AGTCGTCATCCTTATGCCGCTCTTGAAGGCTTGGATGGCGTCGAGAAAATGTCGAAGTCGAAAACACAACACTACATTGGCATTAAACGAACATCCAACGACATGTACGGAAAAACGATGTC  
 4570 4580 4590 4600 4610 4620 4630 4640 4650 4660 4670 4680

L P D S L M K K Y I H L A T D L E L E E K K Q L V K D L E T G A V H P R D A K M  
 ATGCCCCGACGCCTGATGAAAAAGTACATCCACTTGGCGACAGACTTAGAGCTTGAAGAGAAAAACAGCTCGTAAAAGACTTAGAAAACCGCCCGCTTATCCCGTGATGCCAAAAT  
 4690 4700 4710 4720 4730 4740 4750 4760 4770 4780 4790 4800

L L A R T I V R M Y H G E K A A E A A E H S F K T V F Q E N S L P E D I P A V N  
 GCTTTTAGCCAGAACAATCGTCCGAATGTATCAGGAGAGAAAGCAGCAGAAGCTGCCGAACACTCGTTTAAAACAGTCTTTCAGGAAAACAGCTGCCGGAAGATATACCGCCGTAAA  
 4810 4820 4830 4840 4850 4860 4870 4880 4890 4900 4910 4920

W K G E K T J A M I D L L V K L K L L S S K S E A R R H I Q N G G V R I D G E K  
 CTGGAAGGCGAAAAACGATAGCGATGATTGATCTGCTCGTCAAGCTAAAGCTCCTCTCTTGAAGAGCGAGGCGCGCGCATGATTCAAACGGAGGTGTACGCATCGACGGAGAGAA  
 4930 4940 4950 4960 4970 4980 4990 5000 5010 5020 5030 5040

V T D V H A K A E I R E N M I I Q V G K R K F L K L Q \*  
 AGTAACAGATGTTTACGCCAAAGCAGAGATAAGAGAGAATATGATCATCAAGTCGGCAAACGCAAGTTTTTAAAGCTCCAATAATCTGATATTCGAAACGCTTATGACCCCTTCATTCA  
 5050 5060 5070 5080 5090 5100 5110 5120 5130 5140 5150 5160

TAAGCGTTTTTTTGCAGGTATAAAGCACTGCTTTTTCAAAAACGCTCTCCATCTTGCAGGAAACTCCTTTAGCATCCCGAAATATCCCAACCATACATTTAGGGGAGGTTTTTTTTT  
 5170 5180 5190 5200 5210 5220 5230 5240 5250 5260 5270 5280

—————>

M K K L L T V M T M A V L T A G T L L L P A Q S) V T P A A H A) V Q I S N S E R E  
 GATGAAAAAGCTTTTACTGTCATGACGATGGCTGTTTTAACTGCCGGCACACTGCTTGGCCGACAGAGTGTACCCCTGCCGCGCAGCTGTACAAATATCAAATAGCGAGCGGA  
 5290 5300 5310 5320 5330 5340 5350 5360 5370 5380 5390 5400

L P F K A K H A Y S T I S Q L S E A I G P R I A G T A  
 GCTTCCATTCAAAGCAAAACATGCGTACTCTACCATTTCTCAGTAAGTGAAGCAATCGGCCCCAGAAATAGCCGGAACGTCAGCT  
 5410 5420 5430 5440 5450 5460 5470 <PstI>

Figure 1 continued

obvious especially in regions known to be important for recognition of the amino acid, the tRNA, and ATP (Rould *et al.*, 1989). In addition there is weak, but significant, similarity with *B. subtilis* tryptophan tRNA synthetase. This is prominent in several regions as shown in the figure and indicates that these synthetases may derive from a common ancestor.

The regions of similarity between the YRSs span the first 300 aminoacid residues. Several consensus can be found in this domain, but it appears that the YRS corresponding to the *tyrS1* gene departs significantly from the consensus of all YRSs presently known, including the recently identified gene product for the mitochondrial YRS from *Neurospora crassa* (Akins and Lambowitz, 1987). The first well defined consensus is G(F/L/I/A)(D/Q)P(S/T)(A/G) (region I) sequence preceding the H(L/I/V)G(H/N) (region II) consensus. It is followed by a long sequence, well preserved between all YRSs, with some significant variations in the *N. crassa* or the present *B. subtilis* enzymes, down to the KMSKS (region III) consensus (which is KFGK(T/S) in all other YRSs, except the one presented here). After this point, divergence between the different synthetases increases, the *tyrS1* gene product is showing particularly strong variation.

Comparison with the tryptophan tRNA synthetases shows that the proteins may have evolved from a common ancestor and permit to identify several consensus regions common both to YRS and WRS. The HIGH (region II) region is poorly preserved in WRS. Immediately downstream from this region are several well preserved regions present both in YRSs and WRSs: LR(R/Q)(F/W)QQ(A/S)G, G(L/I)..FSYPL (M/L)(Q/A) or G-DQ-(G/Q)(N/H)I. Then follows a region which is very well preserved in all WRSs and absent in YRSs. Downstream from the consensus sequence KMSKS, several regions of identity are present in both families, centered around basic aminoacids, as expected for RNA binding proteins. Finally it should be noted that the present YRS is devoid of cysteine residues, although it contains many methionine residues.

## DISCUSSION

Two main features can be emphasized after this first step of "blind" sequencing of the *B. subtilis* genome. First the region that has been sequenced is densely packed with ORFs. And these ORFs have features permitting to suppose that they code for proteins. Among the five complete ORFs found, three would

code for proteins having no counterparts in the presently available DNA libraries (EMBL data library release 20). Although this cannot be used as statistically significant, this demonstrates that sequencing can permit identification of genes that would be, otherwise, extremely difficult to discover. The second observation is the identification of a tyrosine tRNA synthetase that departs from the known features of cognate enzymes. The fact that the region harbouring the gene is dispensable shows that there exists at least one other copy of a functional YRS, presumably similar to the enzymes already known from other organisms, especially from *Bacilli*. Southern hybridization studies using the corresponding DNA fragment did not reveal other homologous genes, even at low stringency (data not shown), thus showing that the dispensable *tyrS1* gene is substantially different from the normal, yet unknown, *tyrS* housekeeping gene. The study of the corresponding aminoacid residues sequence does not suggest that it is a pseudo-gene or a degenerate form of YRS. Indeed it is difficult to see how, for instance, the tyrosyl-AMP binding consensus KFGKT, could have evolved towards the more usual consensus KMSKS by random mutations. The same is true in other parts of the sequence, as well as in the fact that it has significant similarities with tryptophan tRNA synthetases, thus providing a missing link for the evolution of these aromatic aminoacid tRNA synthetases, and suggesting that they come from a common descent. It should be noted here that the alignment presented is consistent both with the alignment proposed by Akins and Lambowitz (1987) for YRSs and by Jones *et al.* (1986) for YRSs and WRSs, although use of an alignment software proposed by Saurin and Marlière (1986) has suggested a different alignment of the amino-terminal end of the proteins centered on a conserved region, GIDPTA (not shown). Careful analysis of the constraints underlying this latter alignment suggests that the amino-terminal end of the protein derives probably from a repeated modular structure, consistent with a Rossmann-fold tertiary structure (not shown).

*tyrS1* seems dispensable, because the gene can be deleted without much damage for the cells. This suggests that the corresponding YRS could have a function different from the usual function in protein synthesis. It is known that tRNAs are sometimes used in intermediary metabolism reactions (e.g. Li *et al.*, 1989). This could be the case with a tyrosine residue used as a precursor of some dispensable metabolic intermediate or cofactor. Another possibility is that, like the newly discovered second lysyl-tRNA synthetase in *E. coli*, this synthetase could be part of a heat-



III

|  |     |        |
|--|-----|--------|
| ITSGLEVKAARESEPDQERKYVTPKTALDECVGF.TVPLLT..DSSGAKFGKSAGNAI   | 333 | NCYRS  |
| ITAGLELIRKTKGEARA.F.....GL.TIPLVTKADGT..KFGKTESGTI           | 239 | BSTYRS |
| ITSGIDLTRRLHQVQV..F.....GL.TVPLITKADGT..KFGKTEGGAV           | 244 | ECYRS  |
| ITAGLELIRKTKGEARA.F.....GL.TIPLVTKADGT..KFGKTESGTI           | 239 | BCYRS  |
| VLGRHFQERYNKEKQ.....VVI.LMPLLEGLDVE.KHKSCKHNYI               | 251 | BSYRS1 |
| IELTROLAERFNKRYNKRYGELFTIPEARIPK.VGARIMSLV...DPTK.KHKSODPNP. | 207 | BSWRS  |
| LELTRHLAEKFNMYKKNF...F..PKPVTLAQTKKVLSL...STPEK.KHKSODPNH.   | 252 | SCWRS  |
| LELSRDIAQRFNALYGEIF..K.V.PEFPFKS.GARVMSLL...EPTK.KHKSODNR.   | 203 | ECWRS  |

I GR L R N F P G T P L T DGT KHKS N I  
FG

|   |     |        |
|---|-----|--------|
| WLDPYQTSVDFYGYFVRRSDQEVENLLKLFTHP ISEITKMEEHKDPKSRVAQHTLA     | 393 | NCYRS  |
| WLDKEKTSPEYFYQFWINTDORVIRYLKYFTFLSKEEIEALEQELREAPEKRAAQKTLA   | 299 | BSTYRS |
| WLDPKKTSPEYFYQFWINTADADVYRFLKFFTFMSIEEINALEEDKNSGKAPRAQYVLA   | 304 | ECYRS  |
| WLDKEKTSPEYFYQFWINTDORVIRYLKYFTFLSKEEIEALEQELREAPEKRAAQKALA   | 199 | BCYRS  |
| GIN.EH...PNDMYGKTH.S.LPDS..LMKKYIHLATD....LE..L...EEKK..QLVKD | 293 | BSYRS1 |
| KAYITLLDDAKTIEKKIKSAVTDSEGTIRYL.....DKEAKPGIS..NLLNI          | 252 | BSWRS  |
| DSVIFLNDEPKAIQKRLA.ALTDISISDRFY.....DPVERPGVS..MLINI          | 296 | SCWRS  |
| NNVIGLLEDPKSVVKKIKRAVTDSEPPVRY.....DVQNKAGVS..NLLDI           | 249 | ECWRS  |

WLD TSP KFYQK I AD DV LKYF F EI LE E P K AQL LA  
S

|  |     |        |
|--|-----|--------|
| REVTLVHGKQEAASAEQ.QHRMYTGQMT.IPOVSRAKDAATGGDQYKTSIDQPVTLNN   | 451 | NCYRS  |
| EEVTKLVHGEEALRQAIRISEA.LFSGDIANLTAAEI.EQGFKDVPSFV.....HEG..  | 349 | BSTYRS |
| EQVTRLVHGEEGLQAAKRITEC.LFSGLSALSSEADF.EQLAQDGVPMVEM.EK...G.. | 356 | ECYRS  |
| EEVTKLVHGEEALRQAIRISEA.LFSGDIANLTAAEI.EQGFKDVPSFV.....HEG..  | 349 | BCYRS  |
| LE.TGAVHPRDAKMLLARTI.VRMYHGE..K..AAEAHESFKTVFQENSLPEDI.PAVN  | 346 | BSYRS1 |
| ..YSTL..SGQSIIEELER..Q...YEG...K.....GYGVFKADLAQWVI.....     | 285 | BSWRS  |
| ..VSGI..QRKSIEDVVE..DVSFRN...N.....YRDFKDYVSEVII.....        | 330 | SCWRS  |
| ..LSAV..TGQSIPELEK..Q...FEG...KH.....YGHKGEVADAVS.....       | 282 | ECWRS  |

EVT LVHG A R F G A E FKD V

|   |     |        |
|---|-----|--------|
| ..A...PRIDHILPESLINGKS.....I..GRILSPP.VLPAAPPRATS.SRLPRAATS   | 496 | NCYRS  |
| ..GD..VPLVELLVSAGISPSKRQAREDIQNGAIYVNGERLQDVGAILTAEHRLEGRFTV  | 405 | BSTYRS |
| ..AD...LMQALVDSSELQPSRQARKTIASNAITINGEKOSDPEYFFKEEDRLFGRFTL   | 410 | ECYRS  |
| ..GD..VPLVELLVSAGISPSKRQAREDIQNGAIYVNGERLQDVGAILTAEHRLEGRFTV  | 405 | BCYRS  |
| WKGEKTIAMIDLKLLKLLSSKSEARRNIQNGGVRIDGKVTDVHA..KAEIR.E.NM.I    | 401 | BSYRS1 |
| ...ETLRPIQERY..HHMWESE.ELDRVLDEGA.....EKANRV.....ASEMVRKHEQA  | 329 | BSWRS  |
| ...EELKGPRTF...EKYINEPTYLHSVVEGSH.....RKARE.....KAAKNL..PTFIK | 374 | SCWRS  |
| ...GMLTELQERY..HRFRNDEAFLQVMDGA.....EKASA.....HASRTLKAVYQA    | 327 | ECWRS  |

PL E LV SK AR I GAI GEK V AE RL FT

|                                     |     |        |
|-------------------------------------|-----|--------|
| ADAQSRRRERHHEPGPHLFYARQAVVPRRDPALPD | 531 | NCYRS  |
| I..RRGKKKYY.....LIRYA               | 419 | BSTYRS |
| L..RRGKKKYYC.....LICHK              | 424 | ECYRS  |
| I..RRGKKKYY.....LIRYA               | 419 | BCYRS  |
| I..QVGKRFK.....LKLQ                 | 413 | BSYRS1 |
| M..GLGRRR                           | 336 | BSWRS  |
|                                     |     | SCWRS  |
| IGFVAKR                             | 334 | ECWRS  |

I GK

Figure 2 continued

shock response set of proteins (Hirshfield et al., 1984). Further experiments are required to clarify this point.

## MATERIALS AND METHODS

### Bacterial Strains and Growth Media

*E. coli* strain TG1 (Sambrook et al., 1989) carrying plasmid pMF2 was grown in L-broth supplemented with ampicillin (50 µg/ml) and chloramphenicol (5 µg/ml). M13 phage stocks were prepared using strain TG1 grown in 2YT medium (Sambrook et al., 1989). Complementation of YRS activity was performed in the *E. coli* thermosensitive strain HB2109, a F- derivative from strain HB2111 (Bédouelle and Winter, 1986).

*B. subtilis*, strain 168 was grown in Penassay antibiotic medium No. 3 (Difco).

### Isolation of Recombinant Plasmid pMF2

The isolation of a recombinant plasmid (pMF1) containing a *sacY49* mutation leading to constitutive levansucrase production has been described (Débarbouillé et al., 1987; Débarbouillé et al., 1990, in press). Plasmid pMF2 was obtained from pMF1 by *in vivo* allelic exchange in the wild-type strain *B. subtilis* 168 replacing the plasmid-borne mutant allele *sacY49* by the wild-type *sacY* allele.

### DNA Cloning and Sequencing

Random fragments of sonicated DNA from plasmid pMF2 were shotgun cloned into the *Sma*I site of M13mp8 (Messing et al., 1982; Deininger, 1983). Recombinant clones carrying DNA from *Bacillus subtilis* insert were identified by plaque hybridization (Benton and Davis, 1977) and sequenced by the dideoxy method (Sanger et al., 1977) using the modified T7 DNA polymerase Sequenase<sup>®</sup> USB (Tabor and Richardson, 1987). In a first step 112  $\neq$  M13 phages from the shotgun cloning were sequenced using the single strand DNA as template. After a first compilation some clones were sequenced in the other orientation using the double strand DNA as template. This permitted to read each base of the fragment an average of 5.5 times. All base pairs have been read at least once on both strands. DNA sequences were compiled and analysed by the programs of Staden (1979, 1980) as modified by B. Caudron for the Data General Computer MV8000 at the Unité d'Informatique Scientifique. The 5' end of the fragment sequenced overlapped genes *sacX* and *sacY*, the sequence of which has already been published (Zukowski et al., 1988). The 3' end of the fragment will be overlapped by sequencing the adjacent region in the near future.

### Disruption of the Tyrosine tRNA Synthetase Gene

An *Xba*I site is present within the tyrosine tRNA synthetase coding sequence. Plasmid pMF2 was linearised at its single *Xba*I site. The 1.5 kb *Clal* fragment carrying the kanamycin resistance determinant *aphA3* was purified (Trieu-Cuot et al., 1983). The *Xba*I and *Clal* protruding ends were repaired using Klenow DNA polymerase. The linear pMF2 DNA and the DNA fragment containing *aphA3* were ligated using T4 DNA ligase. The ligation mixture was used to transform *E. coli* strain TG1. Clones containing recombinant plasmids were selected on two plates containing ampicillin (50 µg/ml) and kanamycin (5 µg/ml). These recombinant plasmids confer to *B. subtilis* resistance to chloramphenicol (5 µg/ml) and kanamycin (5 µg/ml). One of those, pMF3, was linearised with *Sma*I and used to transform *B. subtilis* 168. Kanamycin resistant,

chloramphenicol sensitive integrants were identified, which carry a tyrosine tRNA gene disrupted by the *aphA3* gene.

## ACKNOWLEDGMENTS

We are grateful to T. Garnier for fruitful discussions and technical advices, to F. Solari and F. Ausseil for their collaboration and to H. Bédouelle for constructive comments as well as for the gift of strain HB2109. We also thank B. George for efficient secretarial help. This work benefited from Institut Pasteur, CNRS (UA1129) and EEC (SCI-0211-C-EDB) grants. The sequence has been deposited with the EMBL Data Library under the accession number X52480.

(Received 2 March 1990)

## NOTE ADDED IN PROOF

Further analysis of unknown reading frames has revealed that CDS3 is weakly related to another *B. subtilis* protein that regulates protease production and sporulation (Perego, M. and Hoch, J.A. (1988) *J. Bacteriol.* 170, 2560–2567). In addition CDS1 is significantly related to mitochondrial NADH-ubiquinone oxidoreductase (chain 4) (e.g. Bibb, M.J. et al. (1981) *Cell* 26, 167–180).

## REFERENCES

- Akins, R.A. and Lambowitz, A.M. (1987). A protein required for splicing group I introns in *Neurospora* mitochondria is mitochondrial tyrosyl-tRNA synthetase or a derivative thereof (Sequence translated from the DNA sequence). *Cell* 50, 331–345.
- Aymerich, S. and Steinmetz, M. (1987). Cloning and preliminary characterization of the *sacS* locus from *Bacillus subtilis* which controls the regulation of the exoenzyme levansucrase. *Mol. Gen. Genet.* 208, 114–120.
- Bédouelle, H. and Winter, G. (1986). A model of synthetase/transfer RNA interaction as deduced by protein engineering. *Nature* 320, 371–373.
- Benton, W.D. and Davis, R.W. (1977). Screening of gt recombinant clones by hybridization to single plaques *in situ*. *Science* 196, 180–182.
- Débarbouillé, M., Kunst, F., Klier, A. and Rapoport, G. (1987). Cloning of the *sacS* gene encoding a positive regulator of the sucrose regulon in *Bacillus subtilis*. *FEMS Microbiology* 41, 137–140.
- Débarbouillé, M., Arnaud, M., Fouet, A., Klier, A. and Rapoport, G. (1990). The *sacT* gene regulating the *sacPA* operon in *Bacillus subtilis*, shares strong homology with transcriptional antiterminators, in press.
- Deininger, P.L. (1983). Random subcloned of sonicated DNA: application to DNA sequence analysis. *Analytical Biochem.* 129, 216–223.

- Hirshfield, I.N., Tenreiro, R., VanBogelen, R.A. and Neidhardt, F.C. (1984). *Escherichia coli* K-12 lysyl-tRNA synthetase mutant with a novel reversion pattern. *J. Bact.* 158, 615-620.
- Jones, M.D., Lowe, D.M., Borgford, T. and Fersht, A.R. (1986). Natural variation of tyrosyl-tRNA synthetase and comparison with engineered mutants. *Biochemistry* 25, 1887-1891.
- Li, J.-M., Brathwaite, O., Cosloy, S.D. and Russell, C.S. (1989). 5-aminolevulinic acid synthesis in *Escherichia coli*. *J. Bact.* 171, 2547-2552.
- Mead, D.A., Szczesna-Skorupa, E. and Kemper, B. (1986). Single-stranded DNA "Blue" T7 promoter plasmids: a versatile tandem promoter system for cloning and protein engineering. *Protein Eng.* 1, 67-74.
- Messing, J. and Vieiza, J. (1982). A new pair of M13 vectors for selecting either DNA strand of double digest restriction fragments. *Gene* 19, 269-276.
- Piggot, P.J. (1989). Revised Genetic Map of *Bacillus subtilis* 168. In Smith, I., Slepecky, R.A., Setlow, P. (eds). *Regulation of prokaryotic development. Structural and functional analysis of bacterial sporulation and germination* (Washington DC, American Society for Microbiology), pp. 1-41.
- Rould, M.A., Perona, J.J., Söll, D. and Steitz, T.A. (1989). Structure of *E. coli* glutamyl-tRNA synthetase complexed with tRNA<sup>Gln</sup> and ATP at 2.8 Å resolution. *Science* 246, 1135-1142.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual*. Second edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
- Saurin, W. and Marlière, P. (1986). Comparaison de plusieurs séquences protéiques par reconnaissance de blocs conservés. *C. R. Acad. Sc. Paris*, tome 303, série II, N° 13, pp. 541-546.
- Shepherd, J.C.W. (1981). Method to determine the reading frame of the protein from the purine/pyrimidine genome sequence and its possible evolutionary justification. *Proc. Natl. Acad. Sci. USA* 78, 1596-1600.
- Staden, R. (1979). A strategy of DNA sequencing employing computer programs. *Nucleic Acids Res.* 6, 2601-2610.
- Staden, R. (1980). A new computer method for the storage and manipulation of DNA gel reading data. *Nucleic Acids Res.* 8, 3673-3694.
- Steinmetz, M., Aymerich, S., Gonzy-Tréboul, G. and LeCoq, D. (1988). Levansucrase induction by sucrose in *Bacillus subtilis* involves an antiterminator homology with the *Escherichia coli* *bgl* operon. In Ganesan A.T. and Hoch J.A. (eds), *Genetics and Biotechnology of Bacilli*, Vol. 2. San Diego, Academic Press, pp. 11-15.
- Tabor, S. and Richardson, C.C. (1987). DNA sequence analysis with a modified T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA*, 84, 4767-4771.
- Trieu-Cuot, P. and Courvalin, P. (1983). Nucleotide sequence of the *Streptococcus faecalis* plasmid gene encoding the 3'5'-aminoglycoside phosphotransferase type III. *Gene* 23, 331-341.
- Watson, M.E.E. (1984). Compilation of published signal sequences. *Nucleic Acids Res.* 12, 5145-5164.
- Zukowski, M., Miller, L., Cogswell, P. and Chen, K. (1988). Inducible expression system based on sucrose metabolism genes of *Bacillus subtilis*. In Ganesan A.T. and Hoch J.A. (eds). *Genetics and Biotechnology of Bacilli*, vol. 2. San Diego, Academic Press, pp. 17-22.

## V) CONCLUSION

Ce travail a été réalisé dans plusieurs laboratoires dont les thèmes de recherche incluent l'étude de trois organismes représentatifs des principaux systèmes utilisés en Biotechnologie : Clostridium acetobutylicum, Saccharomyces cerevisiae, et Bacillus subtilis. L'étude de ces trois microorganismes m'a permis d'utiliser les connaissances accumulées dans ces différents domaines de la Biologie pour aborder des problèmes nouveaux et obtenir des résultats originaux.

Au cours de cette étude, un phage filamenteux défectif de C. acetobutylicum qui existe sous forme simple brin a été isolé, et une méthode d'isolement d'ADN spécifique pour la forme simple brin a été mise au point à partir des techniques d'électroporation. Ceci devrait permettre la détermination d'un système efficace de transformation génétique pour C. acetobutylicum.

L'utilisation de levures marquées de S. cerevisiae pour le levurage de jus de fermentation a été vérifiée en conditions réelles de vinification (50 hl). Quatre technologies différentes de vinification ont été étudiées ; il apparait que seule la fermentation par macération carbonique pose le problème de l'établissement de la souche de levure sélectionnée.

L'intérêt du séquençage intégral du génome de B. subtilis a été illustré. Les résultats significatifs que j'ai obtenus lors de l'étude de B. subtilis incluent le clonage des gènes *narA* et *thiC* par des moyens génétiques et l'isolement de plusieurs phages  $\lambda$ FixII recombinants contenant des inserts chromosomiques de B. subtilis. L'un de ces phages a été séquencé qui contient les gènes *narA* et *spoIID*, une ORF identique à 50% à l'ORF R (*rev-4*) de B.

subtilis, les sous-unités  $b$ ,  $\delta$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\epsilon$  de l'ATPase, et un gène qui pourrait être celui de la sous-unité  $\gamma$  de l'uréase. Ce travail a également permis le clonage d'une protéine biotinylée de B. subtilis, la détection d'une nouvelle protéase, et l'identification d'une tyrosine tRNA synthétase secondaire (*yrs*). Il est de plus intéressant de noter que les gènes *narA* et *spoIID* sont situés à proximité l'un de l'autre sur le chromosome de B. subtilis.

Ces trois thèmes de recherche différents m'ont permis d'obtenir une formation en Microbiologie suffisamment diversifiée pour acquérir une connaissance approfondie des principaux systèmes utilisés en Biotechnologie. Bien que les résultats obtenus lors de l'étude du fragment *sacS-gerB* du chromosome de B. subtilis soient les plus nombreux, ceux-ci ne constituent dans cette optique qu'un apport équivalent à celui des autres méthodologies et matériaux présentés.

# ANNEXES

## SOMMAIRE DES ANNEXES

Annexe 0 : Polylinker de pBluescript et pMTL22

Annexe 1 à 8 : Alignement de la séquence peptidique de l'ATPase de *B. subtilis* avec celle de la bactérie thermophile PS3

Annexe 9 : Séquence nucléotidique de la sous-unité F1 et de la sous-unité b de la sous-unité Fo de l'ATP synthase de *B. subtilis*

Annexe 10 : Séquence nucléotidique de la région adjacente à *spoIID*

Annexe 11 : Séquence nucléotidique de la sous-unité  $\gamma$  de l'uréase de *B. subtilis*

Annexe 12 : Profil hydrophatique de la sous-unité b de la sous-unité Fo de l'ATP synthase de *B. subtilis*

Annexe 13 : Alignement de la séquence nucléotidique de la sous-unité  $\gamma$  de l'uréase de *B. subtilis* avec celle d'*U. urealyticum*

Polylinker de pBluescript :

BssHII KpnI ApaI XhoI Sali ClaI HindIII EcoRV EcoRI PstI  
DraII AccI  
HincII

SmaI BamHI SpeI XbaI NotI EagI BstXI SacII SacI BssHII

Polylinker de pMTL22 :

EcoRV BamHI NdeI AatII AccI Sali MluI PstI HindIII XbaI

EcoRI SstI SmaI KpnI NcoI SphI ClaI BglII XhoI StuI NruI

Annexe 0

S01399 448, 456  
62.4% IDENTITY IN 149 AA OVERLAP

```

      10      20      30      40      50
F80      MSQLPLELGLSFNGGDILFQLLAMLILLALLKKYALGPLLNIMKOREDHIAGEIT
      : . . . X : . . . : . . . : . . . : . . . : . . . : . . . :
S01399    MLWKANVWVLGEAAHGISGGTIIYOLLMFIILLALLRKFANOPLMNIMKOREEHIATKST
      10      20      30      40      50      60

      60      70      80      90      100     110
F80      SAEKKNK EADOLIEEORVLLKEARDESOTLIENAKKLGEKOEI IQAARAESERLKEAA
      . . . . . : . . . : . . . : . . . : . . . : . . . : . . . :
S01399    RRKNDROEA EKLL EORELMKOSRQEAQALIENAA SLAEEQKEI VASARAEAE RVKEAA
      70      80      90      100     110     120

      120     130     140     150     160     170
F80      RTEIVKEKEQAVSALREOVASLSVMIASKVIEKELDEQAQEKLIQDYLKEVGSR
      . . . . . : . . . : . . . : . . . : . . . : . . . : . . . :
S01399    KKEIEREKEQAMALREOVASLSVLIASKVIEKELTEQQAAS
      130     140     150     160

```

Annexe 1 : Alignement de la séquence peptidique de la sous-unité b de Fo de *B. subtilis* (F80) avec la sous-unité b de Fo de la bactérie thermophile PS3 (S01399)

S01400 219, 436  
46.4% IDENTITY IN 179 AA OVERLAP

```

      10      20      30      40      50      60
DP1  MSGSAVSKRYASALFDIANESAQLNQVEEELIVVKQVFQNEKALNDVNLNHPKVPAAKKKE
S01400 MNQEVIAKRYASALFQIALEGGOLDRIEEDVRAVRQALAENGEFLSLLSYPKLSLDQKKA
      10      20      30      40      50      60

      70      80      90      100     110
DP1  LIONAFGSLSQSVLVITIFLLIDRHRAAIVPELTDEFIK-LANVARQTEDAIVYSVKPLTD
S01400 LIREAFAGVSTPVONTLLLLLRRHFRGLVPELAGTVSRPRSTTARGIAKAVAYSGAASTD
      70      80      90      100     110     120

      120     130     140     150     160     170
DP1  AEMLPLSQVFAKKAGVASLRIRNEVOTDLIGGIKVRIGNRIYDGSVSGKLORIERQLAGE
S01400 EELRALSDVFAQKVGKOTLEIENIIDPELIGGVNVRIGNRIYDGSVSGQLERIRRLQIG
      130     140     150     160     170

      180
DP1  NR

```

Annexe 2 : Alignement de la séquence peptidique de la sous-unité  $\delta$  de F1 de *B. subtilis* (DP1) avec la sous-unité  $\delta$  de F1 de la bactérie thermophile PS3 (S01400)



S01401 2008, 2012  
81.9% IDENTITY IN 502 AA OVERLAP

|        |           |          |          |         |         |        |
|--------|-----------|----------|----------|---------|---------|--------|
|        | 10        | 20       | 30       | 40      | 50      | 60     |
| AP1    | VSIKAEF   | ISTLIKDD | IONYSO   | SDIEVQ  | DDVGTV  | IOVGDG |
| S01401 | .X:       | .....    | .....    | .....   | .....   | .....  |
|        | MSIRAE    | ISALIKDD | IENYES   | QIQVSD  | QVSDVGT | IOVGDG |
|        | 10        | 20       | 30       | 40      | 50      | 60     |
|        | 70        | 80       | 90       | 100     | 110     | 120    |
| AP1    | GMAONLE   | ESNVGI   | VILGPF   | SEIREG  | DEVKRT  | GRIMEV |
| S01401 | :::       | .....    | .....    | .....   | .....   | .....  |
|        | GMALNLE   | FENVGI   | VILGPY   | TGIKEG  | DEVRR   | TGRIMF |
|        | 70        | 80       | 90       | 100     | 110     | 120    |
|        | 130       | 140      | 150      | 160     | 170     | 180    |
| AP1    | ILTSKTR   | PIESPAP  | GVMDRKS  | VHEPLOT | GIKAI   | DALIP  |
| S01401 | . . . . . | .....    | .....    | .....   | .....   | .....  |
|        | VETTET    | RPIESRA  | PGVMDRRS | VHEPLOT | GIKAI   | DALVPI |
|        | 130       | 140      | 150      | 160     | 170     | 180    |
|        | 190       | 200      | 210      | 220     | 230     | 240    |
| AP1    | DAILNOK   | DDMICV   | YVAIGO   | KESTV   | RGVVET  | LRKHGA |
| S01401 | . . . . . | .....    | .....    | .....   | .....   | .....  |
|        | DTIINQ    | KDONMI   | CVVAIG   | KESTVA  | TVVFLA  | KHGAPD |
|        | 190       | 200      | 210      | 220     | 230     | 240    |
|        | 250       | 260      | 270      | 280     | 290     | 300    |
| AP1    | AGVTHAE   | EFHYNG   | KHVLVY   | YDDLSK  | QAAAY   | RELSLL |
| S01401 | .....     | .....    | .....    | .....   | .....   | .....  |
|        | AGVAMG    | EYFMIM   | GKHVLV   | YIDDL   | SKQAA   | AYRQL  |
|        | 250       | 260      | 270      | 280     | 290     | 300    |
|        | 310       | 320      | 330      | 340     | 350     | 360    |
| AP1    | AAKLSDA   | KGAGSI   | TALPFV   | ETQAGD  | ISAYIP  | TNVISI |
| S01401 | .....     | .....    | .....    | .....   | .....   | .....  |
|        | AAKLSDA   | KGGGSL   | TALPFV   | ETQAGD  | ISAYIP  | TNVISI |
|        | 310       | 320      | 330      | 340     | 350     | 360    |
|        | 370       | 380      | 390      | 400     | 410     | 420    |
| AP1    | LSVSRV    | GGAOIK   | AMKKV    | SGTLRL  | DLASY   | RELEA  |
| S01401 | .....     | .....    | .....    | .....   | .....   | .....  |
|        | LSVSRV    | GGAOIK   | AMKKV    | AGTLRL  | DLAAY   | RELEA  |
|        | 370       | 380      | 390      | 400     | 410     | 420    |
|        | 430       | 440      | 450      | 460     | 470     | 480    |
| AP1    | KQDLNK    | PLPVEK   | QVALY    | ALT     | KG YL   | DJIPVA |
| S01401 | .....     | .....    | .....    | .....   | .....   | .....  |
|        | KQDLHQ    | PIPVEK   | QVLI     | IYAL    | TRGFL   | DDIPVE |
|        | 430       | 440      | 450      | 460     | 470     | 480    |
|        | 490       | 500      |          |         |         |        |
| AP1    | LPADED    | FKA      | IEG      | F       | KRT     | FAPSN  |
| S01401 | .....     | .....    | .....    | .....   | .....   | .....  |
|        | LPNEDD    | L        | NAI      | E       | A       | F      |
|        | 490       | 500      |          |         |         |        |

Annexe 3 : Alignement de la séquence peptidique de la sous-unité  $\alpha$  de F1 de B. subtilis (AP1) avec la sous-unité  $\alpha$  de F1 de la bactérie thermophile PS3 (S01401)



S01402 481, 957  
68.6% IDENTITY IN 287 AA OVERLAP

```

GP1          10      20      30      40      50
      LASLRDIKSRITSTKKTSQITKAMOMVSAAKLNR AENNAKSFVPYMDKIQEVVSNVG
S01402      MKPLASLRDIKTRIVATKKTSQITKAMEMVLSKLNRAEKR-EIVRPYMEKIQEVVANVA
              10      20      30      40      50

GP1          60      70      80      90     100     110
      RVSGNVKHPMLLSREVKKTAYLVITSDRGLAGAFNSSVLRSAEQAMQERHOSKDEYAVIA
S01402      -LAARASHPMLVSRPVKKTGYLVIITSDRGLAGAYNSNVLRLVYQTIQKRHASPDEYAIIV
              60      70      80      90     100     110

GP1          120     130     140     150     160     170
      IGRVGRDFFKKREIP II SELTGLGDEVTFTIKDLAROTIQMFI DGA FDELHLVYNHFVS
S01402      IGRVGLSFFRKRNPVILDITRLPDQPSFADIKEIARKTVGLFADGTFDELVMYNNHYVS
              120     130     140     150     160     170

GP1          180     190     200     210     220     230
      AITQEVTEKLLPLSLDLSGGGKRTASYEFEPSEEEVLEVLLPOYAESLIFGALLDSKAS

S01402      :: ::::::::::::::::::::. . :::::::::::::::::::::
      AIQQEVTERKLLPLTDL--AENKQRTVYEFEPSEEEILDVLLPOYAESLIYGALLDAKAS
              180     190     200     210     220     230

GP1          240     250     260     270     280
      EHAARMTAMKNATDN AKELIDLSLSYNRRARQAAITOEITEIVGGAAALE
S01402      EHAARMTAMKNATDNANELIRTLTL SYNRRARQAAITOEITEIVAGANALO
              240     250     260     270     280

```

Annexe 5 : Alignement de la séquence peptidique de la sous-unité  $\gamma$  de F1 de *B. subtilis* (GP1) avec la sous-unité  $\gamma$  de F1 de la bactérie thermophile PS3 (S01402)



Annexe 9 : Séquence de la sous-unité F1 et de la sous-unité b de la sous-unité Fo de l'ATP synthase de *B. subtilis*.

Les séquences de Shine-Dalgarno, les codons start et les codons stop sont soulignés.

## ATPASE

CCGGCTGAAA AGAGACAAGT GCGGCTGCTG GATCAGCAAA GAAAAACCTG CAGAAGGGAG TTGCCGTAGT AGATGTCTCA 379  
 ATTACCACTT GAACTAGGAT TGTCGTTTAA CGGCGGAGAT ATCCTGTTCC AACTGTTAGC TATGTTAATC TTATTAGCGC 459  
 TTCTGAAGAA ATACGCTTTA GGGCCGCTAT TAAACATAAT GAAACAGCGT GAAGACCACA TCGCTGGAGA AATTACGTCT 539  
 GCTGAASAAA AAAATAAAGA AGCGCAGCAG CTGATTGAAG AGCAGCGCGT TCTTTTAAAA SAAGCAAGAC AGGAATCCCA 619  
 AACTCTTATC GAAAAACGCA AGAAACTGGG AGAGAAGCAA AAAGAAGAGA TTATTCAGGC TGCACGTGCA GAATCTGAAC 699  
 GTCTGAAAGA AGCAGCAAGA ACTGAAATCG TGAAGGAAAA GGAACAGGCG GTTCTGCTC TCCGTGAGCA AGTAGCGTCT 779  
 CTTTCTGTCA TGATTCGTC GAAAGTGATC GAAAAAGAAC TGGATGAACA AGCGCAAGAG AAATTGATCC AGGACTATCT 859  
 TAAAGAGGTA GGAGAAAGCC GATGAGTGGA TCAGCTGTCT CTAACGATA TGCATCAGCT CTTTTGATA TAGCCAATGA 939  
 GTCCGCTCAG CTGAATCAAG TAGAAGAAGA GCTAATTGTT GTAAAAAAG TATTTCAAAA TGA AAAAGCG CTTAATGATG 1019  
 TGTTGAACCA TCCGAAGGTG CCGGCTGCGA AGAAAAAAGA GCTGATTCAA AATGCATTG GCTCTTTGTC ACAGTCCGTA 1099  
 CTCAATACGA TTTTCTTTT GATTACCAGC CATCGTGCCG CGATTGTCCC TGAGCTCACA GATGAGTTTA TCAAACCTGC 1179  
 AAATGTGGCC CGTCAAACAG AAGACGCAAT CGTATATTCA GTGAAACCGC TGACGATGC AGAAATGTTA CCATTATCAC 1259  
 AAGTGTGTC AAAAAAGCC GGAGTCGCTT CACTGAGAAT CAGAAATGAA GTGCAGACGG ATTTAATAGG CGGTATTA 1339  
 GTCCGATTG GAAACCGGAT TTATGACGGC AGCGTAAGCG GGAAGCTTCA GCGCATTGAA CGTCAATTAG CCGGGGAAAA 1419  
 TCGATAGAAG GGGTGAAACC TAAGT GAGCA TCAAAGCTGA AGAGATTAGC ACGCTGATAA AACAGCAAAT AAAAAATTAT 1499  
 CAATCTGATA TTGAAGTTCA AGACGTAGGT ACAGTCATCC AAGTCGGTGA CCGTATTGCA CGTGTGCAGC GCCTTGACAA 1579  
 CTGTATGGCC GGTGAACCTG TCGAATTTTC AAACGGTGTT TTGGGTATGG CTCAAAACCT TGAGGAATCA AACGTAGGTA 1659  
 TCGTCATCTT AGGACCTTTC AGTGAGATCC GTGAGGGAGA CGAAGTAAAA AGAACAGGCC GCATCATGGA GGTTCTGT 1739  
 GGTGAAGAGT TAATCGGCCG TATTSTAAAC CCGCTAGGCC AGCCGGTTGA CCGACTAGGC CCGATTCTGA CAAGCAAAAC 1819  
 TCGTCCGATC GAAAGCCCTG CACCAGGCGT TATGGACCGT AAATCCGTTT ATGAACCGCT TCAAAACCGST ATCAAAGCGA 1899  
 TCGATGCACT GATTCCAATC GGCCGCGGCC AGCGTGACT GATCATCGGT GACCGTCAAA CAGGTAAAAC ATCTGTTGCG 1979  
 ATCGATGGGA TCCTGAACCA AAAAGACCAA GACATGATCT GTGTATATGT TCGGATCGGC CAAAAAGAAT CAACAGTCCG 2059  
 CGGCGTAGTA GAAACATTGC GTAACACGG CCGCTTGAT TATACAATTG TTGTAACGGC GTCTGCGTCA CAGCCGGCAC 2139  
 CGCTTCTGTA CCTGGCACCG TATGCTGGGG TTACAATGGC AGAAGAATTT ATGTACAACG GCAAGCACST TCTTGTGTA 2219  
 TACGATGATC TTTCTAAACA AGCGSCCGCT TACCGTGAGC TGTCCTTGCT TCTTCGCGCT CCGCCAGGCC GTGAAGCGTT 2299  
 CCTGSGGAT GTATTCTATC TTCATCCCG TCTGCTTGAG CGTGCAGCAA AGCTTAGCGA CCGGAAAGGC GCAGGATCAA 2379  
 TTACAGCTCT GCCGTTGTA GAAACACAAG CCGGAGATAT CTCTGCTTAT ATTCCGACGA ACGTCATTTT CATCACCGAC 2459  
 GGACAGATCT TCCTGCAATC TGATTGTTT TTCTCAGGCG TACGTCCAGC GATCAATGCC GGATTGTCTG TATCCCGTGT 2539

CGCGGGCTCA GCGCAAAATCA AAGCGATGAA AAAAGTATCA GGTACTTTGC GTCTTGACCT TCGGTCAATC CGTGAGCTGG 2619  
AAGCATTCCG TCAATTCCGT TCTGACCTCG ACCAAGCGAC TCAGGCAAAA CTGAACCCGG GTGCGCGTAC AGTTGAAAGT 2699  
CTGAAGCAGG ATCTGAADAA GCCCCTCCG GTTGAAAAGC AAGTAGCTAT TCTTTATGCG CTGACAAAAG GATATCTCGA 2779  
TGATATTCTT GTGGCGGATA TCAGACGTTT TGAAGAAGAG TACTACATGT ACCITGACCA AAACCATAAA GACCTGCTTG 2859  
ACGGAATTCC GAAAACAGGA AACCTTCCG CTGATGAAGA CTTCAAAGCT GCAATCGAAG GCTTCAAAAG CACATTTGCA 2939  
CCAAGCAACT AACTCGAATG CTGATGAGAG AAAAAGGTTT TCTTTTTCTC TCTTTTTACG CAGATGAGA GAAAAGGTGG 3019  
TGAAATCTTT GSCCTCATTG CGCATATTA AGTCAAGGAT CACGTCAACG AAAAAFACAA GTCAGATTAC AAAAGCCATG 3099  
CAGATGGTAT CTGCGGCTAA GCTGATCGT GCTGAAAAA ATGCAAAATC ATTTGTGCCA TATATGGATA AAATCCAAGA 3179  
GGTTGTGTC AAGCTCGGAA GAGTTCCGG CAACGTGAAS CACCCGATGC TTCTCAGCAG AGAAGTGAAG AAAACGGCAT 3259  
ACCTTGTCTT TACGTCTGAC CGCGTCTTG CCGCGGCTTT TAACAGTTCC GTTTTACGGA GTGCTTATCA GGCCATGCAA 3339  
GAACGTCTAT AGTCTAAGGA TGASTATGCG GTGATTGCCA TCGGAAGAGT GGGCCGTGAT TTCTTTAAGA AACGGGAGAT 3419  
TCCGATCATT TCCGASTTAA CAGGACTTGG AGATGAAGTA ACGTTTACAG AAATTAAGA TCTTGCCCGT CAAACGATTC 3499  
AAATGTTTAT AGACGGTCCG TTTGATGAAT TGCACCTTGT TTATAACCAT TTTGTCAGCG CCATTACTCA AGAAGTAACG 3579  
GAGAAAAAAC TTCTGCCGTT ATCTGATTTG GGCAGCGGCG GCGGAAAAG AACGGCGTCT TATGAATTTG AACCATCTGA 3659  
AGAGGAGGTT CTGGAGGTTT TECTTCTCA ATATGCAGAA AGCTTAATCT TCGGTGCGCT TCTTGACAGT AAAGCAAGTG 3739  
AGCACGCTGC AAGAATGACG GCGATGAAAA ACGCGACAGA CAACCGGAAG GAACTTATCG ATTCACTTTC GCTTTCTTAC 3819  
AACCGCGCTC GCCAAGCAGC CATCAACAA GAAATTACGG AAATTGTCCG CGGAGCAGCC GCTTTAGAA AGAAAGATTT 3899  
TGT CAGGAGG GATAGCGATG AAGAAAGGAC GCGTTAGCCA GGTATTAGGA CCGGTGTCG ACGTGCCTTT TGAAGACGGT 3979  
CACTTGCTCG AAATTTATAA TCGGATTAAG ATTTACAGC CAGCTGCAAG TGAATAACGAA GTAGGTATTG ATTTAACGCT 4059  
TGAGGTGCTT CTTCATTTAG GTGATGATAC AGTCCGTACA ATCGCAATGG CATCTACAGA TGGTGTTCAG CGCGGTATGG 4139  
AAGCTGTAGA TACAGGAGCG CCAATCTCAG TACCGGTTGG TGATGTAACA CTTGACGCTG TATTTAACGT TCTCGGAGAA 4219  
AATATTGATT TGAATGAGCC GGTTCCTGCG GATGCGAAAA AAGGATCCGAT TCACAGACAG GCGCCTTCAT TCGATCAGCT 4299  
TTCAACAGAA GTTGAAATC TTGAAACAGG TATTAAGTT GTTGATTTGC TTGCTCCTTA CATTAAAGGGC GGTAATAACG 4379  
GATTGTTCCG TGGTGCCCGT GTAGTAAAA CCGTATTAAT CCAGGAATTA ATCAACAACA TCGCGCAAGA GCACGGCGGT 4459  
ATCTCTGTAT TCGCCGCGGT AGGASAGCGT ACTCGTGAAG GGAACGACCT TTTCTACGAA ATGAGTGACT CTGGCGTAAT 4539  
CAACAACAACA GCCATGGTAT TCGGACAAAT GAACGAGCCG CCGGCGCAC GTATGCGTGT TGCTTTGACA GCCTTACAA 4619  
TGGCTGAGCA CTTCCTGAT GTACAAGGAC AGGACGTAAT GTTCTTATC GATAACATTT TCCGTTTAC ACAAGCGGGT 4699  
TCAGAGGTTT CAGCCCTTCT TGGCCGTATG CCTTCAGCGG TTGGTTATCA GCCGACGCTT GCAACTSAGA TGGGTACGCT 4779  
CCAAGAGCGT ATCAGCTCTA CGAACGTTGG ATCAATTAACA TCTATCCAGG CGATCTACGT GCTTCCCGAT GACTACACTG 4859  
ACCCGGCGCC GCGGACAACG TTCGCTCACT TGGATGCGAC AACAAACCTT GAGCGTAAAT TAACTGAAAT GGGTATTTAC 4939  
CCTGCGGTTG ATCGGTTGCG ATCTACATCA CGCGCCCTTG CTCTGAAAT TGTTGGAGAA GAGCACTATG CGGTTGCGCG 5019  
TGAASTACAG TCAACGCTTC AGCGTACAA AGAGCTTCA GATATCATTG CGATTCTCGG TATGGATGAA TTAGCGGAGG 5099

AAGACAAACT TGTCGTTAC CGCGCACGTC GTATCCAGTT CTTCC7TTCT CAGAACTTCC ACGTGGCTGA ACAGTTCACT 5179  
 GGACAAAAAG GTTCTTACGT GCCTSTAAAA GAGACGGTAC AAGGCTTCAA AGAAATCTTA GCCGGTAAAT ATGACCATCT 5259  
 TCCAGAAGAT GCATTCCGTC TTGTAGGCCG TATCGAAGAA GTTGTGAGA AAGCAAAAGA AATGGGTGA GAAGTTTAAA 5339  
 TCTGGTCTTA GGAGGGTAAA AGCATGAAGA CCGTTAAAGT CAATATCGTT ACTCCCGACG GCCCAGTATA CGATGCGGAT 5419  
 ATCGAAATGG TGAGTGTGAG AGCCSAAAGC GCGGATCTCG GTATTTTGCC AGGCCATATT CCAACCGTGG CTCCTCTTAA 5499  
 AATCGGGCT GTCCGTCTGA AAAAAACCGG GCAGACTGAA ATGGTTGCCG TCAACGGCGG TTTTGTAGAA GTCCGTCTTG 5579  
 ATCATGTAC CATCCTTGCC CAGGCTGCCG AGACAGCGGA AAGCATCGAT AAAGAGCGCG CTGAAGCTGC ACCCCAGCGG 5659  
 GCGCAGGAC GTTTGAATTC TCAATCASAT GATACTGACA TTCGTGGGC TGAGCTTGGC TTACAGCGGG CTTTGAACAG 5739  
 ATTGGATGA GCAGGSAAT AAGAAAAAT CCTTCTCTT ATGAGAAGGA TTTTTTATG AACGCACAAG AGAATCAALC 5819  
 AGCGACAGCA ACGGGCGATC CTCTTTTGGG CGCAGGTATA AACCATCACA GGCATTATT GTATGCCAG ATTTGCGATG 5899  
 TCTTCTAAAG CCAAGTGTAT AAGCCCGTC TTTATAGGGA TAAGGTTTA CAACCCATTC AGCATCTTCA TCGCGGGCAA 5979  
 TCACACAATC TATAASATAA AACTGATCA TTTTCTCGAA TTCTTCAATA GGAAGATTCC AATCAGCAAG CAAAAGTTCT 6059  
 TGTATACTTT GAAAATAGAG GTTTAAGCT CTTCCAAGGC TTTCTCTCGT AATAGSTAAG AGAAATGTTA AAAAATGA

Annexe 10 : Séquence de la région adjacente à spoIID.  
Les principaux signaux génétiques sont soulignés.  
La phase codante de protR s'étend des nucléotides 458 à 1688, la phase codante de spoIID s'étend des nucléotides 1964 à 2993, la phase codante de AVI s'étend des nucléotides 3380 à 3917, la phase codante de AVII s'étend des nucléotides 4388 à 4913.

## SPOIID

CTTTtagact GCAGCTTGTC ACAACCCTCA AAAACACAA TTCCA~~ACTT~~C GTATTTATTA TATGAGCAAA TGAGTCTAGA 80  
 GACACCGGAT AGCTGGAATG ATACATATGA ACAGTTTGAA CGGGAGACAC TTGGGATATT CCAAGAAAAG TAGTTATTTT 160  
 TACTTGCTA AATGGTCACT TAGATGATAA TATGAATATT GTTTTGCAA AAAGCAATCA GTTATTAAT GAATTAAGC 240  
 AAGATCATT A AACATGTAG TTGAGCCAAA TTTGGTTTCT ATTTCTGGCT TTACAGATGA GTGGGAAGAG TACATTATGA 320  
 CCTCGAACA TAAAATGAAT TTGCA AATTG CACTTASAAG TGCAGGAATG GCGGAAAAC ATACCGTTAC GGTGGCACA 400  
 CCAATCGTTA GACTGAATA TTAATATAGA GAATTTAGGG ACGGGAGGG GAATACCTTG GAAAAAATCA TCGTCCGGC 480  
 CGGTCAGAAG TTAACGGCA CAGTCAAAGT TGAAGGCGCT AAAAATGCCG TTTTACCTGT TATCGCTGCA TCTTTATTAG 560  
 CAAGTGAAGA AAAAAGCGTA ATTTGTGATG TACCTACGCT CTCCGATGTA TATACAATTA ACGAAGTGT GCGTCATTTA 640  
 GGAGCAGATG TGCAATTTGA AAATAATGAA GTGACTGTAA ATGCTTCATA CGCTTTGCAA ACTGAAGCAC CTTTGAATA 720  
 TGTTGTA AAA ATGCGTGGCT CTGTGCTTGT CATGGGGCCG CTTCTGGCG GTACAGGTCA TGCAAGAGTT GCACTTCCGG 800  
 GCGGATGGC AATTGGTCC AGACCGATTG ATCAGCATT AAAAGGTTTT GAAGCAATGG GCGCAGAAAT CAAAGTCGGT 880  
 AATGGCTTCA TTGAAGCTGA AGTAAAGGC GACTGCAAG GCGCAAAAAT TTATCTGGAC TTCCAAGTG TAGGAGCTAC 960  
 AGAGAACCTG ATTATGGCAG CCGCTCTAGC TGAAGSACA ACAACGCTGG AAAACGTGG AAAAGAACC GAAATCGTTG 1040  
 ATTAGCAAAA CTATATCAAC GGCATGGCG GAAAAATCCG CGGAGCTGG ACCGGCACCA TCAAAATGA AGGAGTCGAA 1120  
 AAGTTCACG GCGTAAAACA CCATATTATT CCTGACGTA TTGAA GCGGG CACATTTATG GTTGCTGCTG CAATCACTGA 1200  
 AGGAAACGTA TTAGTAAAAG GAGCGTTC TGAGCACCTC ACCTCTTTAA TTGCAAAAAT GGAAGAGATG GGTGTAACAA 1280  
 TTAAGGATGA AGGTGAAGT CTGCGTGCA TCGGCCGAA AGAGCTTAAA CCGATTGACA TCAAAACAAT GCCTCACCCG 1360  
 GCGTCCCGA CTGATATGCA GTCACAAATG ATGGCGCTT TGCTTCGTGC AAGCGGCACA AGCATGATTA CAGAAACCGT 1440  
 TTTTGA AAC CGTTTTATGC ATGCGGAAGA ATTCGCGCT ATGAATGGTG ATATCAAGAT TGAAGGACGT TCTGTATCA 1520  
 TTAACGGTCC TGTAAGCTT CAGGGAGCTG AAGTTGCAGC GACTGATTTG CGTGCAGGTG CAGCGCTGAT TCTTGGGGG 1600  
 TTAGTGGCTG AAGGTCACAC ACGTCTTACT GAATTGAAGC ACTTAGACCG CGGTACGTT GATTTCCATC AGAAGCTTGC 1680  
 CGCTCTGGC GCAGACATCG AACGTGTAAT TGATGAGTCT GCTTCTGAGC AAGAGAATAA AGAAGTCGTT TCTGACTTAA 1760  
 ATGCATAAAC TACTTGAAA ATCAGTATGT CACTCTGGCA TACTGGTTTT TTATTTTTTG TGTATTATCA TCATTATGGC 1840  
 CAAAATAGCA AAAAGAATA CGTAAATGAC AAATAAAGTT TCTGTCCAAA ACGAGAGTCA TATTAGCTTG TCCCTGCCCA 1920  
TAGACTAGAC TAGAGTCGAA TCCCGAGCAG GAGGCAGCTG AATATGAAAC AATTCGCAAT CACTATCC GACTATGTG 2000  
 CATTGATCT CTTGGTTC CACTCTTGG TTATACCATT CCAGCATAAT AAGGAAGCGG GGGCCAGCGT AGAATCAGAA 2080  
 AAGACAGCAG TCAGCACGAA ACCACATCG AAAGGAGCAG AAACATTGAA AGCATCACCT GTTCTATT CCGTCTATCG 2160  
 AACCGCAAAT CAATCCGTAG AAAAATTCC GCTTGAAGAG TATGTGATTG GAGTCGTGCG CTCCGAAATG CCGCAACCT 2240

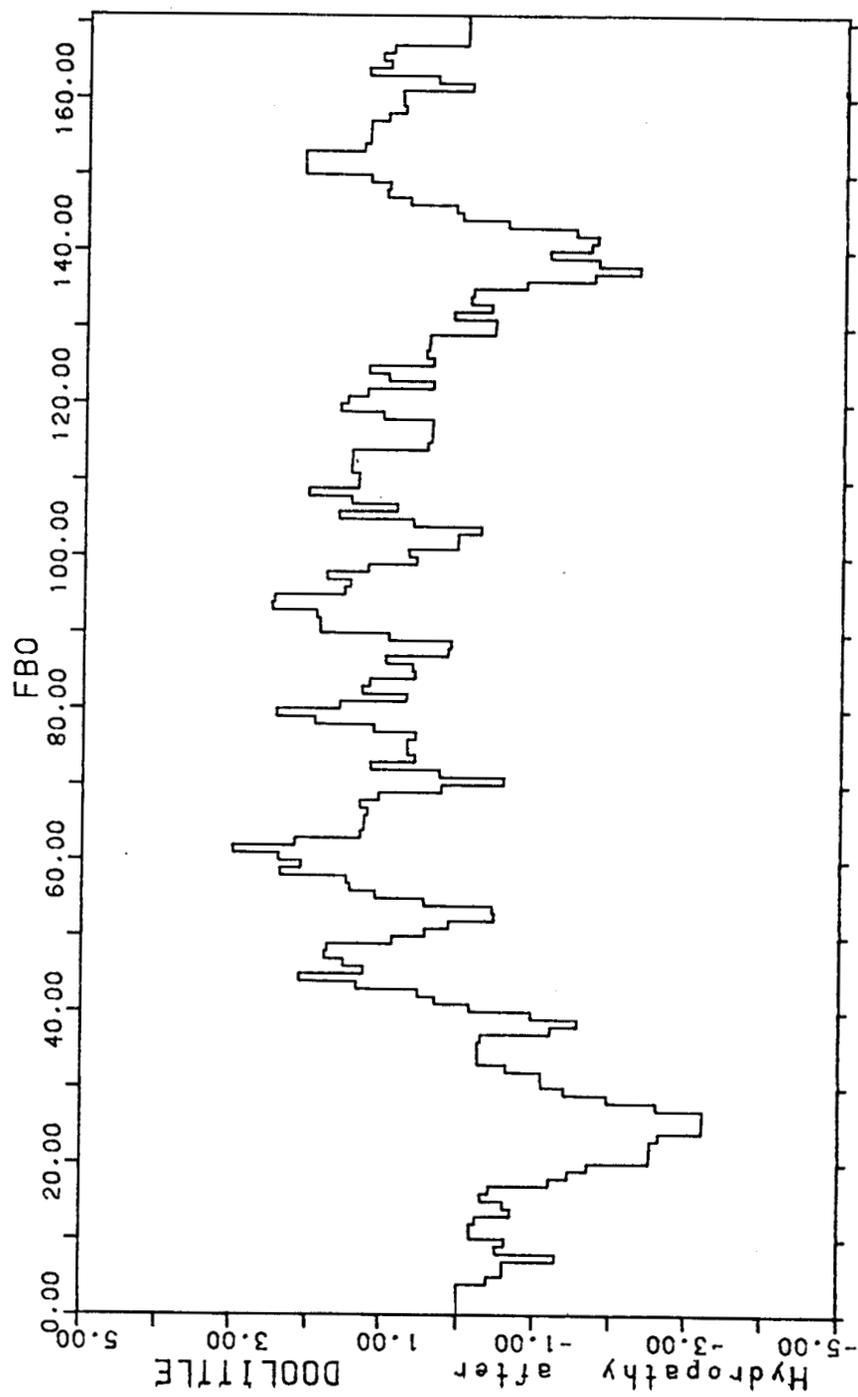
TTAAACCTGA AGCGCTGAAA GCCCAGGCGC TTGCCGCCAG AACATTTATT GTCAGACTGA TGGTGTCAAA TTCAGCGGTA 2320  
 GAAGCTCCTA AAGGCTCACT GGTGGATGAT ACACAGATGT TCCAGGTGTA TAAAAGCAAA GCGGAGCTGA AAAAAACAGTG 2400  
 GGGCACAAGC TATGAGACAA AACTG AAGAA AATCACAGAT GCGGTAGCCA GTACGCAAGG CAAAATCTTA ACGTACAACA 2480  
 ACCAGCCGAT TGAAGCATCC TTTTCTCCA CAAGCAACGG CTACACAGAG AATGCAGAAG CTTATTGGAC AAGTGCTATC 2560  
 CCATATTTAA AAAGCGTCAA AAGCCCATGG GATAAAAAAT CTCCGAAATA TAAGSCAAGG AAAACCTTTA CAGCGGCAGA 2640  
 ATTTAGCAA AAGCTTGGCG TCAAGCTGGA TGGATCTAGC GCAGTAGGGA AGATTACCGG CGAGACACCG GGCCACCAGG 2720  
 TCGCAACTGC CGTCATTAAC GGCAAGACGC TGAAGGGAAG AGACATACGT GAAAAGTTGG GTCTCAACTC CGCCGATTTT 2800  
 GAATGGAAGC GAAATGGAGA CACAATTACA GTCACGACGA GAGGATTTGG CCACGGTGTG GGGATGAGCC AATA CGGAGC 2880  
 GAATTTTATG GCCAAAGAAG GCAAAACGGT TGATGACATT GTAAAGTACT ATTACCAAGG CACACAAATT TCTGAGGCAG 2960  
 ATGCGTTTTT GAATAAATAT ATGGCGAAAA AGTAGAAAA GACGCTCATT GGCGTCTTT TCTGTTCCT GAATATCTTT 3040  
 TATCTAAGAG TAGCTGAAAC GGTTAGAGG AGCTTCTGAC CGTCTTTAAG GATCAGTCTT TTTTCTCGA TTTACGGCTT 3120  
 TGAGATTGTG GTAAAGTCAT CCACATTTGT CGACAGACTT CGACAGACTC ATATAACGAA AGCCGAAAGG CGACAATCTT 3200  
 CTCTTTTCAT AGGGGCGTTA GGTGSAATC ACTAATAAAA ACATAGGTTT CTATGCAAAA GTAAACAACA GCAATGGGGA 3280  
 AGCCCGCCCT ATTTATTTCT CGGTTTTTGT TGATTACGAT AGATTATGTT CAGCGTTGAA CATCTGCAGC CAGAAACAAA 3360  
 TTGGTAAAGG GGTACGAACA TGAAGAAAAG ATTTTCACTG ATCATGATGA CAGGATTGCT TTTTGGATTA ACTTCACCTG 3440  
 CTTTTGCAGC TGAAGAGACA GAAACGGAGG CTAAAGCGCC GCGGAACGTG GCGGTCTTGC TTGATGCCAG CGGAAGTATG 3520  
 GCGAAAAGAA TAGATGGTGT ATCTAAATTT AATTCAGCCA AAAAAAGAAAT TTCAAATTTT GCGAGCTCAT TGCCAGAAGG 3600  
 AACTCAGGTG AAAATGAGCG TGTTTGGCTC TGAAGGAAAC AATAAAAATT CCGGAAAAGT TCAGTCATGT GAAGCCATTC 3680  
 GCAACGTATA CGGCTTCCAA AGCTTAAACG AGCAAAGCTT CCTCAATTCT CTCAATACAA TTGGGCGGAC TGCTGGACG 3760  
 CCAATTGCCA AAGCGCTGAA CGAGGCGAAA TCTTCTTTTG ACCAGCTTGA TGCAAAAGGG GAAAAAGTGG TGATCTGTT 3840  
 GACAGACGGT GAGGAAACGT GCGGAGGAAA CCCS ATAAA ACAGCAAAGG AACCTGCAAA AAGACAATAT CACTGTAAAT 3920  
 GTGATCGGCT TTGATTATAA AGAGGATAC AAAGGCCASC TGAACGCAAT TGCGAAAGTA GCGGCGGTG AATACTTCCC 4000  
 GGCTTATACC CAAAAAGATG TTGAGAAAAT TTTCACTCAG CAATCACTAA TGTTATCTAA ATAAGAAAA AGAGGCTGGT 4080  
 GATGGTACG CTCTTTTTTGT ATGTAACAAA AACTTCAAAA CTAATTAGCT GTCTGCCTGT AGCATAGTAG CAGATTTCCG 4160  
 TTTTCTGTC ATAAATCAAGG CGAAAAGA ATGAATGATC ACACAAGCTG TTTATGCTTT CTGGCGGAA ATGGCTAAA 4240  
 ATGAACGGCA TTTCAAGCGA TTTTGACAG ATCGTGAAAG CAGCACAAA GGCAGCATGT CCTATATCTT TTGCATTTT 4320  
CGTATTCTAG CATAAAGAAG AATCCACAT ATAGTTGGAA AATATGAAAA AGAAAGGAAG TAATCATATG AAAAAATTGC 4400  
 TGGCTCCGG TATCATTGGA TTGTTGACTG TTTCCATTGC CTCCCCTCT TTTGCCGCTG AGAAACAGGC TGATACAAAT 4480  
 GTAGCTGTTT TGTTGACGG AAGCGAAGC ATGTTTCAA AAACAGGAGG AGAGCGCAA ATAGACATTG CCAAAAAATC 4560  
 AGTAAAATCT TTTGCCGAAC TTCTTCCAAA GGATACAAAT CTTATGCTT GCCTTTTCGG GCATGCAGGG AACAATAAAC 4640  
 TGTCCGGTAA AGCCCTTTCG TGCAATACGA CGGAAACCAT TTACGGCTT CATCCATATG AAGGGTCTCT GTTTGACAAT 4720  
 TCATTGAGTS AGCTCAAGCC GACTGGATGG ACACCGATTG CCAAAGCACT TGCAGACACA AGAAAAGAA TTGAAGCATT 4800  
 TGACCGACAG GGGAAAACG TGGTSTATT AATCACTGAC GGTGAAGAAA CGTCCGCGCG AGACCTCCG CGAAAATCGA 4880  
 AAAGCTCCGT GCATCAAATG TGGATACGAT TGTAACATC ATTTGCTTCA ACTTTGATGT CAAGGGTAAT G 4960

Annexe 11 : Séquence nucléotidique de la sous-unité  $\gamma$   
de l'uréase de *B. subtilis*.

Les principaux signaux génétiques putatifs sont  
soulignés. La phase codante de la sous-unité  $\gamma$   
s'étend des nucléotides 847 à 1163.

## UREASE

ATGATCACAT AAGTTGTAAG CTTTTATAAC ATCATTGAA GAAATCAAAC AAAGGGGTTT TGATGACGGA ATGTACTGCA 80  
 CCGAAGAAAG TGTCGGAAGC CACGGGAGTG CGAGAAAACA GTTGATATCG ACACAAGCGG GGATTCAGTG AAGGCCGGAA 160  
 TTCTTCACTC ATTAAGAGGA ACGGTGGCGA TCAGCGAAGT GACCGTCCAT GATGCGGAGT TAATGGCAAT AGTCGGGTTT 240  
 TCGGTATCGT GGAGGTGGC GCCTAGACT GGCCTCCATG CCATGCTTCC GGTGCGGAC AAAACAACGG GCTGCCATCA 320  
 GGAATGTAAC GTCTGGTTAT GATCAATCGA TGTTTTGCA CAGCCTTAGC TTCGCCCGG GCAGATTGTC GCTGTATTAG 400  
 GCCGACTCCG GCTAAACAGA GGGGGCTATT TGTTTTAATC AAACGTTTGT TCAAAACGAA AAGGGGTCA ACAGCAGCAG 480  
 CTTGCAATCA CCAGAGCGCT GATGGTAAAT CTGAAAAATA TTTACTTAA TAAACAATTG GAAGGAATTC AGATTTAAT 560  
TGTCGAACTA GTCAGACAAG TTATTATAGA AATTCCAGA AAAAAAGACA TTTGTTATCT TAGTCGAGCG CAGTCTGGAA 640  
 AACCGTTAA TCTGATCTT CCGTSATCGA CAGGGGAACG GTTGTGACC ATGGCAAGGC TGATGAAGAA ACGGCATGCC 720  
GCYTTGAAAG GCATCTTACC GTATSACCC CATTCTCGAA GAAAGGGAGG TGATGTCCC CCTGACTGTT CCGCAAATTC 800  
 GTAGTAGCCG CATCTGCTT ATATGACAAA AGAAGGAGSA CTACAGATGA AACTGACACC AGTTGAACAA GAGAAATTGC 880  
 TCATTTTGC AGCGGGGAA TTAGCCAAAC AGCGAAAGGC GCGGGCGTT CTGCTGAAC ATCCTGAGGC TGCTGCTTAT 960  
 ATCACCTGCT TTATTATGGA AGGCSACGT GATGGGAAGG GAGTAGCAGA GCTGATGAA GCCGGCCGCC ATGTATTAAC 1040  
 GGAAAAAGAT GTGATGGAGG GTGTCCCTGA GATGCTGGAC AGCATTGAGG TGGAGGCCAC ATTTCCGGAT GGTGTAAGC 1120  
 TTGTTACGCT ACATCAGCCA ATTTCTGCGG AGGTGAAGTC ATGAAGCCGG GAGCATTCA AATTGCTGAG GGAACCATTA 1200  
 CGATTAATGA AGGCCGTGAG ATACGGGAGG TAACGGTAAA AAACACCGGA TCGAGTCGAC TCCCTTAGT GAGGGTTAAT 1280  
 TTGGGGCCG GAGCTCTAGA GATCCACTCG TTATTCTCGG ACGA



Annexe 12 : Profil hydropathique de la sous-unité b de la sous-unité Fo de l'ATP synthase de *B. subtilis*

URMYC 346, 408  
56.2% IDENTITY IN 297 NT OVERLAP

```

      850      860      870      880      890      900
UREASE CTACAGATGAAACTGACACCAGTTGAACAAGAGAAATGCTCATTTTGTCAGCGGGGAA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
URMYC   TGAATCTATCATTAAAGAGAAATCCAAAAGTTATTGGTAACAGTAGCTGCTGAC
      10      20      30      40      50

      910      920      930      940      950      960
UREASE TTAGCCAAACAGCGAAAGGCGGGGCGTTCTGCTGAACTATCCTGAGGCTGCTGCTTAT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
URMYC   GTTCAAGAAGACGT TTAGCTAGAGGTTTAAAATTAACACTACTCAGAACGTGTCGCTTA
      60      70      80      90      100     110

      970      980      990      1000     1010     1020
UREASE ATCACCTGCTTTATTATGGAAGGCGCACGTGATGGGAAGGGAGTAGCAGAGCTGATGGAA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
URMYC   ATTACTGACCACGTA GTGGAAGGGGCAAGAGATGGTAAGTTAGTTGCTGACTTAATGCAA
      120     130     140     150     160     170

      1030     1040     1050     1060     1070     1080
UREASE GCCGGCCGCCATGTA TTAACGGAAAAAGATGTGATGGAGGGTGTGCTGAGATGCTGGAC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
URMYC   TCTGCTCGTGAAGTATTACGTTGATCAAGTTATGGAAGGTGTAGATAAATGGTTGGT
      180     190     200     210     220     230

      1090     1100     1110     1120     1130     1140
UREASE AGCATTGAGGTGGAGGCCACATTTCCGGATGGTGTAAGCTTGTACGGTACATCAGCCA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
URMYC   ATAATCCAAGTTGAA GTTACTTTCCAGATGGTACTAAACTAGTTTCTGTACAGACCCA
      240     250     260     270     280     290

      1150     1160     1170     1180     1190     1200
UREASE ATTTCTGGGAGGTGAAGTCATGAAGCCGGGAGCATTCAAATTGCTGAGGGAACCATTA
      : : : X
URMYC   ATTTACAATAA
      300

```

Annexe 13 : Alignement de la séquence  
nucléotidique de la sous-unité  $\gamma$  de l'uréase  
de *B. subtilis* (UREASE) avec la sous-unité  
 $\gamma$  de l'uréase d'*U. urealyticum* (URMYC)

# REFERENCES

- 1) **ADLER H.I. and W. CROW.** 1987. A technique for predicting the solvent producing ability of *Clostridium acetobutylicum*. *App. Env. Microbiol.* 53:2496-2499
- 2) **AFSHAR A.S., H. BIEBL, K. SCHALLER, K. SCHUEGERL.** 1985. Production of acetone and butanol by *Clostridium acetobutylicum* in continuous culture with cell recycle. *App. Microbiol. Biotechnol.* 22:394-398
- 3) **AKAMATSU T. and J. SEKIGUCHI.** 1987. Genetic mapping by means of protoplast fusion in *Bacillus subtilis*. *Mol. Gen. Genet.* 208:254-262
- 4) **AMJAD M., J. CASTRO, J.J. WU, H. SANDOVAL, and P.J. PIGGOT.** 1990. Physical map of the *Bacillus subtilis* 168 chromosome. International Conference on the *Bacillus subtilis* Genome. Institut Pasteur. Poster P69
- 5) **AMZEL L.M. and P.L. PEDERSEN.** 1983. Proton ATPases : structure and mechanism. *Ann. Rev. Biochem.* 52:801-824
- 6) **ANAGNOSTOPOULOS C. and J. SPIZIZEN.** 1961. Requirements for transformation in *Bacillus subtilis*. *J. Bact.* 81:741-746
- 7) **ATKINSON M.R. and S.H. FISHER.** 1991. Identification of genes and gene products whose expression is activated during nitrogen-limited growth in *Bacillus subtilis*. *J. Bact.* 173:23-27
- 8) **AYMERICH S. and M. STEINMETZ.** 1987. Cloning and preliminary characterization of the *sacS* locus from *Bacillus subtilis* which controls the regulation of the exoenzyme levansucrase. *Mol. Gen. Genet.* 208:114-120
- 9) **BACHMANN B.J.** 1990. Linkage map of *Escherichia coli* K12, edition 8. *Microbiol. Rev.* 54:130-197
- 10) **BAHL H., W. ANDERSCH, K. BRAUN, and G. GOTTSCHALK.** 1982. Effect of pH and butyrate concentration on the production of acetone and butanol by *Clostridium acetobutylicum* grown in continuous culture. *Eur. J. Appl. Microbiol. Biotechnol.* 14:17-20
- 11) **BAHL H., W. ANDERSCH, and G. GOTTSCHALK.** 1982b. Continuous production of acetone and butanol by *Clostridium acetobutylicum* in a two stage phosphate limited chemostat. *Eur. J. Appl. Microbiol. Biotechnol.* 15:201-205
- 12) **BAHL H., M. GOTTWALD, A. KUHN, V. RALE, and W. ANDERSCH.** 1986. Nutritional factors affecting the ratio of solvents produced by *Clostridium acetobutylicum*. *App. Env. Microbiol.* 52:169-172

- 13) **BAHN M. and R.D. SCHMID.** 1987. Enzymes for detergents. *Biotec* 1:119-130
- 14) **BALLESTER S. J.C. ALONSO, P. LOPEZ, and M. ESPINOSA.** 1990. Comparative expression of pC194 cat gene in *Streptococcus pneumoniae*, *Bacillus subtilis*, and *Escherichia coli*. *Gene* 86:71-79
- 15) **BARRE P. and F. VEZINHET.** 1984. Evolution towards fermentation with pure culture yeasts in wine making. *Microbiological Sciences* 1:159-162
- 16) **BASEMAN J.B. and R.L. QUACKENBUSH.** 1990. Preliminary assessment of AIDS-associated *Mycoplasma*. *ASM News* 56:319-323
- 17) **BEECHHOFFER D.H.** 1990. Triple post-translational control. *Mol. Microbiol.* 4:1419-1423
- 18) **BIDAN P. et J. MAUGENET.** 1981. Informations récentes sur l'emploi des levures sèches actives. Leur influence sur la qualité des vins. *Bulletin de l'O.I.V.* 601:241-254
- 19) **BIGGIN M.D., T. GIBSON, and G.F. HONG.** 1983. Buffer gradient gels and <sup>35</sup>S label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA* 80:3963-3965
- 20) **BIRNBOIM H.C. and J. DOLY.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7:1513-1523
- 21) **BISHOP M.J., C.J. RAWLINGS.** 1987. Nucleic acid and protein sequence analysis : a practical approach. IRL Press. Oxford. England
- 22) **BLANCHARD A.** 1990. *Ureaplasma urealyticum* urease genes, use of a UGA tryptophan codon. *Mol. Microbiol.* 4:669-676
- 22bis) **BLASCHEK H.P. and M.A. KLACIK.** 1984. Role of DNases in recovery of plasmid DNA from *Clostridium perfringens*. *App. Env. Microbiol.* 48:178-181
- 23) **BLASCHEK H.P.** 1986. Genetic manipulation of *Clostridium acetobutylicum* for production of butanol. *Food Technology*. Oct.84-86
- 24) **BOIDIN A. and J. EFFRONT.** 1917. US Patent # 1227374, May 22
- 24bis) **BOOSERMSUWONG A. and H.P. BLASCHEK.** 1986. Localization and inactivation of DNase activity in *Clostridium pasteurianum* NRRL-B598. *J. Ind. Microbiol.* 1:265-270

- 25) **BORCK K., J.D. BEGGS, W.J. BRAMMAR, A.S. HOPKINS, and N.E. MURRAY.** 1976. The construction in vitro of transducing derivatives of phage  $\lambda$ . *Mol. Gen. Genet.* 146:199-207
- 26) **BRAKHAGE A.A., H. PUTZER, K. SHAZAND, R.J. ROSCHENTHALER, and M. GRUNBERG-MANAGO.** 1989. *Bacillus subtilis* phenylalanyl tRNA synthetase genes : cloning and expression in *Escherichia coli* and *Bacillus subtilis*. *J. Bact.* 171:1228-1232
- 27) **BRAKHAGE A.A., M. WOZNY, and H. PUTZER.** 1990. Structure and nucleotide sequence of the *Bacillus subtilis* phenylalanyl-tRNA synthetase genes. *Biochimie* 72:725-734
- 28) **BRAMWELL M.E.** 1987. Characterization of biotinylated proteins in mammalian cells using  $^{125}$ I-Streptavidin. *J. Biochem. Biophys. Meth.* 15:125-132
- 29) **BROWN S. and M.J. FOURNIER.** 1984. The 4.5S RNA gene of *Escherichia coli* is essential for cell growth. *J. Cell. Biol.* 178:533-550
- 30) **BROWNER M.F., F. TARONI, E. SZTUL, and L.E. ROSENBERG.** 1989. Sequence analysis, biogenesis, and mitochondrial import of the  $\alpha$ -subunit of rat liver propionyl-coA carboxylase. *J. Biol. Chem.* 264:12680-12685
- 31) **BRUSILOW W.S.A., D.J. KLIONSKY, and R.D. SIMONI.** 1982. Differential polypeptide synthesis of the proton translocating ATPase of *Escherichia coli*. *J. Bact.* 151:1363-1371
- 32) **BRUSILOW W.S.A, M.A. SCARPETA, C.A. HAWTHORNE, and W.P. CLARK.** 1989. Organization and sequence of the genes coding for the proton translocating ATPase of *Bacillus megaterium*. *J. Biol. Chem.* 264:1528-1538
- 33) **BRYANT D.L. and H.P. BLASCHEK.** 1988. Buffering as a means for increasing growth and butanol production by *Clostridium acetobutylicum*. *J. Ind. Microbiol.* 3:49-55
- 34) **BULLOCK W.O., J.M. FERNANDEZ, and J.M. SHORT.** 1987. XL1-Blue, a high efficiency plasmid transforming *recA* *Escherichia coli* strain with  $\beta$ -galactosidase selection. *Biotechniques* 5:376
- 35) **CHAMBERS S.P., S.E. PRIOR, D.A. BARSTOW, and N.P. MINTON.** 1988. The pMTLnic- cloning vectors.I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. *Gene* 68:139-149

- 36) **CHIBATA I., T. TOSA, and T. SATO.** 1986. Methods of cell immobilization. In "Manual of Industrial Microbiology and Biotechnology". A.L. Demain and N.A. Solomon eds. p.217-229. American Society for Microbiology, Washington D.C., USA
- 37) **CHOW K.C. and T.F. WONG.** 1988. Cloning and nucleotide sequence of the structural gene coding for *Bacillus subtilis* tryptophanyl-tRNA synthetase. *Gene* 173:537-543
- 38) **CHRISTIANS S. and H. KALTWASSER.** 1986. Nickel-content of urease from *Bacillus pasteurii*. *Arch. Microbiol.* 145:51-55
- 39) **CLARKE L. and J. CARBON.** 1976. A colony bank containing synthetic colEI hybrids plasmids representative of the entire *Escherichia coli* genome. *Cell* 9:91-99
- 40) **COLLINS F.S.** 1988. Chromosome jumping. In "Genome Analysis. A practical approach". K.E. Davies ed. p. 73-94. IRL Press, Oxford, England
- 41) **COLLINS F.S.** 1990. Genome project can benefit search for disease genes. *Human Genome News*, September p. 5
- 42) **CONDEMINE G. and C.L. SMITH.** 1990. Genetic mapping using large DNA technology : alignment of SfiI and AvrII sites with the NotI genomic restriction map of *Escherichia coli* K-12. In "The Bacterial Chromosome". K. Drlica and M. Riley eds. p. 53-60. American Society for Microbiology, Washington D.C., USA
- 43) **COOPER D.G., C.R. McDONALD, and S.J.B. DUFF.** 1981. Enhanced production of surfactin from *Bacillus subtilis* by continuous product removal and metal cations addition. *Appl. Env. Microbiol.* 42:408-412
- 44) **COPLEY C.G., C. BOOT, K. BUNDELL, and W.L. MCPHEAT.** 1991. Unknown sequence amplification : application to *in vitro* genome walking in *Chlamydia trachomatis* L2. *Biotechnology* 9:74-79
- 45) **CROSS R.L.** 1981. The mechanism and regulation of ATP synthesis by F<sub>1</sub> ATPases. *Ann. Rev. Biochem.* 50:681-714
- 46) **CRUEGER W. and A. CRUEGER.** 1982. *Biotechnology. A textbook of industrial microbiology.* Science Tech. Inc. Madison, Wisconsin, USA
- 47) **DAKSHINAMURTI K., L. CHALIFOUR, and R.P. BHULLAR.** 1986. Requirement for biotin and the function of biotin in cells in culture. *Ann Rev. N. Y. Acad. Sci. Biotin*:38

- 48) **DAGERT M. and S.D. EHRLICH.** 1979. Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. *Gene* 6:23-28
- 49) **DANCHIN A., R. DEDONDER, P. GLASER, M.F. HULLO, F. KUNST, L. MARCELINO, I. MOSZER, M. SANTANA, E. SCHNEIDER, and A. VERTES.** 1990. Cloning, mapping, and partial sequencing of the *sacS-sacA* region from *Bacillus subtilis* strain 168 chromosome. International Conference on the *Bacillus subtilis* Genome, Institut Pasteur, Paris, Poster P72
- 50) **DANEMAN H.L.** 1981. Baelene. HLD Associates. Santa Fe, New Mexico
- 51) **DAUNERT S., B.R. PAYNE, and L.G. BACHAS.** 1989. Pyruvate carboxylase as a model for oligosubstituted enzyme-ligand conjugates in homogenous enzyme immunoassays. *Anal. Chem.* 61:2160-2164
- 52) **DEBARBOUILLE M., F. KUNST, A. KLIER, and G. RAPOPORT.** 1987. Cloning of the *sacS* gene encoding a positive regulator of the sucrase regulon in *Bacillus subtilis*. *FEMS Microbiol. Letters* 41:137-140
- 53) **DEBARBOUILLE M., M. ARNAUD, A. FOUET, A. KLIER, and G. RAPOPORT.** 1990. The *sacT* gene regulating the *sacPA* operon in *Bacillus subtilis* shares strong homology with transcriptional antiterminators. *J. Bact.* 172:3966-3973
- 54) **DEBARBOUILLE M. I. MARTIN-VERSTRAETE, A. KLIER, and G. RAPOPORT.** 1991. The transcriptional regulator of *levR* of *Bacillus subtilis* has domains homologous to both  $\sigma^{54}$  and phosphotransferase system dependent regulators. *Proc. Natl. Acad. Sci. USA* 88:2212-2216
- 55) **DEDONDER R.A., J.A. LEPESANT, J. LEPESANT-KEIJLAROVA, A. BILLAUTLT, M. STEINMETZ, and F. KUNST.** 1977. Construction of a kit of reference strains for rapid genetic mapping in *Bacillus subtilis* 168. *Appl. Env. Microbiol.* 33:989-993
- 56) **DRLICA K. and M. RILEY.** 1990. A historical introduction to the bacterial chromosome. In "The Bacterial Chromosome" K. Drlica and M. Riley eds. p. 3-13. American Society for Microbiology, Washington D.C.
- 57) **DYBVIK K.** 1990. Mycoplasmal genetics. *Ann. Rev. Microbiol.* 44:81-104
- 58) **EDWARDS M.J.** 1988. ATCC Microbes and cells at work. American Type Culture Collection, Rockville, Maryland

- 59) **ELDER J.K. and E.M. SOUTHERN.** 1987. Automatic reading of DNA sequencing gel autoradiographs. In "Nucleic Acid and Protein Sequence Analysis : a Practical Approach". M.J. Bishop and C.J. Rawling eds. p.219-229. IRL Press, Oxford, England
- 60) **EPERON I.C.** 1986. Rapid preparation of bacteriophage DNA sequence analysis in sets of 96 clones using filtration. Anal. Biochem. 156:406-412
- 61) **EPERON L.P., I.R. GRAHAM, A.D. GRIFFITHS, and I.C. EPERON.** 1988. Effects of RNA secondary structure on alternate splicing of pre-mRNA : is folding limited to a region behind the transcribing RNA polymerase? Cell 54:393-401
- 62) **EVANS P.J. and H.Y. WANG.** 1988. Enhancement of butanol formation by *Clostridium acetobutylicum* in the presence of decanololeylalcohol mixed extractants. App. Env. Microbiol. 54:1662-1667
- 63) **FERRARI F.A., A. NGUYEN, D. LANG, and J.A. HOCH.** 1983. Construction and properties of an integrable plasmid for *Bacillus subtilis*. J. Bact. 154:1513-1515
- 64) **FERRARI E. and J.A. HOCH.** 1989. Genetics. In "Biotechnology Handbooks : *Bacillus*", p. 57-72. C.R. Harwood ed. Plenum Press, New York
- 65) **FILLINGAME R.H.** 1980. The proton translocating pumps of oxidative phosphorylation. Ann. Rev. Biochem. 49:1079-1113
- 66) **FOUET A., A. KLIER, and G. RAPOPORT.** 1986. Nucleotide sequence of the sucrase gene of *Bacillus subtilis*. Gene 45:221-225
- 67) **FOUET A., M. ARNAUD, A. KLIER, and G. RAPOPORT.** 1987. *Bacillus subtilis* sucrose specific enzyme II of the phosphotransferase system : expression in *Escherichia coli* and homology to enzyme II from enteric bacteria. Proc. Natl. Acad. Sci. USA 84:8773-8777
- 68) **FRICTE D., J. FLOSSDORF, and D. CLAUS.** 1990. Taxonomy of alkaliphilic *Bacillus* strains. Intern. J. System. Bact. 40:92-97
- 69) **FRY D.C., S.A. KUBY, and A.S. MILDVAN.** 1986. ATP-binding site of adenylate kinase : mechanistic implications of its homology with ras-encoded p21, F<sub>1</sub> ATPase and other nucleotide binding proteins. Proc. Natl. Acad. Sci. USA 83:907-911
- 70) **FUTAI M. and H. KANAZAWA.** 1983. Structure and function of proton-translocating adenosine triphosphatase

(F<sub>o</sub>F<sub>1</sub>) : biochemical and molecular biological approaches.  
Microbiol. Rev. 47:285-312

- 71) GLASER P., F. KUNST, M. DEBARBOUILLE, A. VERTES, A. DANCHIN, and R. DEDONDER. 1991. A gene encoding a tyrosine tRNA synthetase is located near *sacS* in *Bacillus subtilis*. DNA Sequence 1:251-261
- 72) GRUNDY F.J. and T.M. HENKIN. 1990. Cloning and analysis of the *Bacillus subtilis* *rpsD* gene encoding ribosomal protein S4. International Conference on the *Bacillus subtilis* genome, Institut Pasteur, Poster P65
- 73) GUFFANTI A.A., S. CLEJAN, L.H. FALK, D.B. HICKS, and T.A. KRULWICH. 1987. Isolation and characterization of uncoupler resistant mutants of *Bacillus subtilis*. J. Bact. 169:4469-4478
- 74) HAGGSTROM L. and C. FORSBERG. 1986. Significance of an extracellular polymer for the energy metabolism in *Clostridium acetobutylicum* : a hypothesis. App. Env. Biotechnol. 23:234-239
- 75) HALL B., S. YOKOYAMA, and D. CALHOUN. 1983. Role of cryptic genes in microbial evolution. Mol. Biol. Evol. 1:109-124
- 76) HAMMES W.P. 1987. Biotechnology, biochemical engineering and food technology. In "Biochemical Engineering", p. 11-35, H. Chmiel, W.P. Hammes, and J.E. Bailey eds. Gustav Verlag, Stuttgart, Germany
- 77) HANAHAN D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166:557-580
- 78) HANEJI T. and S.S. KOIDE. 1989. Transblot identification of biotin-containing proteins in rat liver. Anal. Biochem. 177:57-61
- 79) HASAN N. and E.W. NESTER. 1978. Dehydroquinase synthase in *Bacillus subtilis*. J. Biol. Chem. 253:4999-5003
- 80) HATEWAY C.L. 1990. Toxigenic *Clostridia*. Clin. Microbiol. Rev. 3:66-98
- 81) HAWTHORNE C.A. and W.S.A. BRUSILOW. 1986. Complementation of mutants in the *Escherichia coli* proton translocating ATPase by cloned DNA from *Bacillus megaterium*. J. Biol. Chem. 261:5245-5248
- 82) HAWTHORNE C.A. and W.S.A. BRUSILOW. 1988. Sequence of the genes for the  $\beta$  and  $\epsilon$  subunits of the ATP synthase of

*Bacillus megaterium* QMB1551. Biochem. Biophys. Res. Comm. 151:926-931

83) HENKIN T.M., F.J. GRUNDY, W.L. NICHOLSON, and G.H. CHAMBLISS. 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

84) HERMANN M., F. FAYOLLE, R. MARCHAL, L. PODVIN, M. SEBALD, and J.P. van de CASTELLE. 1985. Isolation and characterization of butanol resistant mutants of *Clostridium acetobutylicum*. App. Env. Microbiol. 50:1238-1243

85) HICKS D.B. and T.A. KRULWICH. 1986. The membrane ATPase of alkalophilic *Bacillus firmus* RAB is an F<sub>1</sub>-type ATPase. J. Biol. Chem. 261:12896-12902

86) HICKS D.B. and T.A. KRULWICH. 1987. Purification and characterization of the F<sub>1</sub> ATPase from *Bacillus subtilis* and its uncoupler resistant mutant derivatives. J. Bact. 169:4743-4749

87) HIRSHFIELD I.N., R. TENREIRO, R.A. van BOGELEN, and F.C. NEIDHARDT. 1984. *Escherichia coli* K12 lysyl tRNA synthetase mutant with a novel reversion pattern. J. Bact. 158:615-620

88) HONEYMAN A.L. and G.C. STEWART. 1989. The nucleotide sequence of the *rodC* operon of *Bacillus subtilis*. Mol. Microbiol. 3:1257-1268

89) HOPP T.P. and K.R. WOODS. 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA 78:3824-3828

90) HOPWOOD D.A. and M.J. MERRICK. 1977. Genetics of antibiotic production. Bacteriol. Rev. 41:595-635

91) HOTCHKISS R. and M. GABOR. 1980. Biparental products of bacterial protoplast fusion showing unequal parental expression. Proc. Natl. Acad. Sci. USA 77:3553-3557

92) HUNT L.T. and W.C. BARKER. 1989. Avidin-like domain in an epidermal growth factor homolog from a sea urchin. FASEB J. 3:1760-1764

93) IGLESIAS A., G. BENSI, U. CANOSI, and T.A. TRAUTNER. 1981. Plasmid transformation in *Bacillus subtilis*. Alterations introduced into the recipient. Homologous DNA of hybrid plasmids can be corrected in transformants. Mol. Gen. Genet. 184:405-409

- 94) **INRA**. 1973. La vinification par macération carbonique. SEI CNRA, Versailles
- 95) **IVEY D.M., J. ABOUDI, and T.A. KRULWICH**. 1990. Molecular characterization of the ATP synthase genes of *Bacillus firmus* OF4 and RAB. Ann. Meeting of the American Society for Microbiology. Poster K109
- 96) **JANATI-IDRISSI R., A.M. JUELLES, A. EL KANOUNI, H. PETITDEMANGE, and R. GAY**. 1987. Sélection de mutants de *Clostridium acetobutylicum* défectifs dans la production d'acétone. Ann. Inst. Pasteur/ Microbiol. 138:313-328
- 97) **JOLIFFE L.K., R.J. DOYLE, and U.N. STREIPS**. 1981. The energized membrane and cellular autolysis in *Bacillus subtilis*. Cell 25:753-763
- 98) **JONES B.D. and H.L.T. MOBLEY**. 1988. *Proteus mirabilis* urease, genetic organization, regulation, and expression of structural genes. J. Bact. 170:3342-3349
- 99) **JONES D.T., A. van der WESTHUIZEN, S. LONG, R. ALLCOCK, S.J. REID, and D.R. WOODS**. 1982. Solvent production and morphological changes in *Clostridium acetobutylicum*. App. Env. Microbiol. 43:1434-1439
- 100) **JONES D.T. and D.R. WOODS**. 1986. Acetone butaol fermentation revisited. Microbiol. Rev. Dec:482-524
- 101) **KANAZAWA H., K. MABUCHI, T. KAYANO, T. NOUMI, T. SEKIYA, and M. FUTAI**. 1981. Nucleotide sequence of the genes for Fo components of the proton translocating ATPase from *Escherichia coli*, prediction of the primary structure of Fo subunits. Biochem. Biophys. Res. Comm. 103:613-620
- 102) **KATZ E. and A.L. DEMAIN**. 1977. The peptide antibiotics of *Bacillus* : chemistry, biogenesis, and possible functions. Bact. Rev. 41:449-474
- 103) **KEMPER J.** 1974. Gene order and cotransduction in the *leu-ara-fol-pyrA* region of the *Salmonella typhimurium* linkage map. J. Bact. 117:94-99
- 104) **KIM A.Y. and H.P. BLASCHEK**. 1991. Isolation and characterization of a filamentous viruslike particle from *Clostridium acetobutylicum* NCIB 6444. J. Bact. 173:530-535
- 105) **KNOWLES J.R.** 1989. The mechanism of biotin-dependent enzymes. Ann. Rev. Biochem. 58:195-221
- 106) **KOHARA Y, K. AKIYAMA, K. ISONO**. 1987. The physical map of the whole *Escherichia coli* chromosome : applications of a new strategy for rapid analysis and sorting of a large genomic library. Cell 50:495-508

- 107) KOKUBU T., I. KRUBE, and S. SUZUKI. 1978.  $\alpha$ -amylase production by immobilized whole cells of *Bacillus subtilis*. Eur. J. Appl. Microbiol. Biotechnol. 5:233-240
- 108) KOLSTO A.B., A. GRONSTAD, and H. OPPEGAARD. 1990. Physical map of the *Bacillus cereus* chromosome. J. Bact. 172:3821-3825
- 109) KORPELA J. 1984. Avidin, a high affinity biotin binding protein as a tool and subject of biological research. Med. Biol. 62:5-26
- 110) KULA M.R. 1987. Trends in Enzymology. Biotec 1:77-83
- 111) KUNST F., M. DEBARBOUILLE, T. MSADEK, M. YOUNG, C. MAUEL, D. KARAMATA, A. KLIER, G. RAPOPORT, and R. DEDONDER. 1988. Deduced polypeptides encoded by the *Bacillus subtilis* *sacU* locus share homology with two component sensor regulator systems. J. Bact. 170:5093-5101
- 112) KUNST K. and K. DEVINE. 1991. The project of sequencing the entire *Bacillus subtilis* genome. Res. Microbiol. Res. Microbiol. (in press)
- 113) KUSTU S., E. SANTERO, J. KEENER, D. POPHAM, and D. WEISS. 1989. Expression of  $\sigma^{54}$  (*ntrA*) dependent genes is probably united by a common mechanism. Microbiol. Rev. 53:367-376
- 114) KYTE J. and R.F. DOOLITTLE. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132
- 115) LABIGNE A., V. CUSSAC, and P. COURCOUX. 1991. Shuttle cloning and nucleotide sequences of *Helicobacter pylori* genes responsible for urease activity. J. Bact. 173:1920-1931
- 116) LAMPE M., C. BINNIE, R. SCHMIDT, and R. LOSICK. 1988. Cloned gene encoding the delta subunit of *Bacillus subtilis* RNA polymerase. Gene 67:13-19
- 117) LANE D., D.L. YOUNG, and F. LYNEN. 1964a. The enzymatic synthesis of holotranscarboxylase from apotranscarboxylase and (+)biotin. I. Purification of the apoenzyme and synthetase characteristics of the reaction. J. Biol. Chem. 239:2858-2864
- 118) LANE D., K.L. ROMINGER, D.L. YOUNG, and F. LYNEN. 1964b. The enzymatic synthesis of holotranscarboxylase from apotranscarboxylase and (+)biotin. II. Investigation of the reaction mechanism. J. Biol. Chem. 239:2865-2871

- 119) LEPESANT J.A., F. KUNST, J. LEPESANT-KEJZLAROVA, and R. DEDONDER. 1972. Chromosomal location of mutations affecting sucrose metabolism in *Bacillus subtilis* Marburg. *Mol. Gen. Genet.* 118:135-160
- 120) LIN Y.L. and H.P. BLASCHEK. 1983. Butanol production by a butanol tolerant strain of *Clostridium acetobutylicum* in extruded corn broth. *App. Env. Microbiol.* 45:966-973
- 121) LIN Y.L. and H.P. BLASCHEK. 1984. Transformation of heat treated protoplasts with pUB110 plasmid DNA. *App. Env. Microbiol.* 48:737-742
- 122) LIPMAN D.J. and W.R. PEARSON. 1985. Rapid and sensitive protein similarity searches. *Science* 227:1435-1440
- 123) LO S.C., M.S. DAWSON, D.M. WONG, P.B. NEWTON III, M.A. SONODA, W.F. ENGLER, R.Y.H. WANG, J.W.K. SHIH, H.J. ALTER, and D.J. WEAR. 1989. Identification of *Mycoplasma incognitus* infection with AIDS patients : an immunohistochemical in situ hybridization and ultrastructure study. *Ann. J. Trop. Med. Hyg.* 41:601-616
- 124) LOISEAU G., M. VALADE, and J.P. MOULIN. 1985. Utilisation des souches antibiorésistantes en fermentation alcoolique. *Le Vigneron Champenois* 4:186-199
- 125) LONG S., D.T. JONES, and D.R. WOODS. 1984. The relationship between sporulation and solvent production in *Clostridium acetobutylicum* P262. *Biotechnol. Letters* 6:529-534
- 126) LONG S., D.T. JONES, and D.R. WOODS. 1984b. Initiation of solvent production, clostridial stage and endospore formation in *Clostridium acetobutylicum* P262. *App. Microbiol. Biotechnol.* 20:256-261
- 127) LOPEZ-DIAZ I., S. CLARKE, and J. MANDELSTAM. 1986. *SpoIID* operon of *Bacillus subtilis* : cloning and sequence. *J. Gen. Microbiol.* 132:341-354
- 128) LOVETT P.S. 1990. Translational attenuation as the regulator of inducible cat genes. *J. Bact.* 172:1-6
- 129) LOVETT P.S., N.P. AMBULOS Jr., W. MULBRY, N. NOGUCHI, and E.J. ROGERS. 1991. UGA can be decoded as tryptophan at low efficiency in *Bacillus subtilis*. *J. Bact.* 173:1810-1812

- 130) **LOVITT R.W., B.H. KIM, G.J. SHEN, and J.G. ZEIKUS.** 1988. Solvent production by microorganisms. *CRC Crit. Rev. Biotechnol.* 7:107-186
- 131) **McCARN D.F., R.A. WHITAKER, J. ALAM, J.M. VRBA, and S.E. CURTIS.** 1988. Genes encoding the  $\alpha$ ,  $\gamma$ ,  $\delta$  and four Fo subunits of ATP synthase constitute an operon in the *Cyanobacterium Anabaena* sp. strain PCC7120. *J. Bact.* 170:3448-3458
- 132) **MCCARTHY J.E.G., H.U. SCHAIRER, and W. SEBALD.** 1985. Translational initiation frequency of *atp* genes from *Escherichia coli*: identification of an intercistronic sequence that enhances translation. *EMBO J.* 4:519-526
- 133) **MCCARTHY J.E.G., W. SEBALD, G. GROSS, and R. LAMMERS.** 1986. Enhancement of translational efficiency by the *Escherichia coli atpE* translational initiation region: its fusion with two human genes. *Gene* 41:201-206
- 134) **McCLELLAN J.A., P. BOUBLIKOVA, E. PALECEK, and D.M.J. LILLEY.** 1990. Superhelicity torsion in cellular DNA responds directly to environmental and genetic factors. *Proc. Natl. Acad. Sci. USA* 87:8373-8377
- 135) **McCLELLAND M., R. JONES, Y. PATEL, and M. NELSON.** 1987. Restriction endonucleases for pulsed field mapping of bacterial genomes. *Nucl. Acids Res.* 15:5985-6005
- 136) **MACRINA F.L., D.J. KOPECKO; K.R. JONES, D.J. AYERS, and S.M. McCOWEN.** 1978. A multiple plasmid containing *Escherichia coli* strain: convenient source of size reference plasmid molecules. *Plasmid* 1:417-420
- 137) **MALOY W.L., B.U. BOWIEN, G.K. ZWOLINSKI, K.G. KUMAR, H.G. WOOD, L.H. ERICSSON, and K.A. WALSCH.** 1979. Amino acid sequence of the biotinyl subunit from transcarboxylase. *J. Biol. Chem.* 254:11615-11622
- 138) **MARTINEZ E., B. BARTOLOME, and F. de la CRUZ.** 1988. pACYC184-derived cloning vectors containing the multiple cloning site and *lacZ* $\alpha$  reporter gene of pUC8/9 and pUC18/19 plasmids. *Gene* 68:159-162
- 139) **MEAD G.C.** 1971. The amino acid fermenting *Clostridia*. *J. Gen. Microbiol.* 67:47-56
- 140) **MEDIGUE C., J.P. BOUCHE, A. HENAUT, and A. DANCHIN.** 1990a. Mapping of sequenced genes (700kb) in the restriction map of the *Escherichia coli* chromosome. *Mol. Microbiol.* 4:169-187

- 141) **MEDIGUE C., A. HENAUT, and A. DANCHIN.** 1990b. *Escherichia coli* molecular genetic map (1000kb) : update I. *Mol. Microbiol.* 4:1443-1454
- 142) **MESSING J. and J. VIEIRA.** 1982. A new pair of M13 vectors for selecting either DNA strand of double digest restriction fragments. *Gene* 19:269-276
- 143) **MILLAR G. and J.R. COGGINS.** 1986. The complete amino acid sequence of 3-dehydroquinate synthase of *Escherichia coli* K12. *FEBS Letters* 200:11-17
- 144) **MIWA Y. and Y. FUJITA.** 1990. Determination of the cis sequence involved in catabolite repression of the *Bacillus subtilis* *gnt* operon ; implication of a consensus sequence in catabolite repression in the genus *Bacillus*. *Nucl. Acids. Res.* 18:7049-7053
- 145) **MOBLEY H.T.L. and R.P. HAUSINGER.** 1989. Microbial ureases : significance, regulation, and molecular characterization. *Microbiol. Rev.* 53:85-108
- 146) **MONOT F. and J.M. ENGASSER.** 1983. Production of acetone and butanol by batch and continuous culture of *Clostridium acetobutylicum* under nitrogen limitation. *Biotechnol. Letters* 5:213-218
- 147) **MOORE J.F., D. BENTON, and C. BURKS.** 1990. The GenBank nucleic acid data bank. *Focus* 11:69-72
- 148) **MOSBACH K., S. BIRNBAUM, K. HARDY, J. DAVIES, and L. BULOW.** 1983. Formation of proinsulin by immobilized *Bacillus subtilis*. *Nature* 302:453-455
- 149) **MOSS J. and M.D. LANE.** 1971. The biotin dependent enzymes. *Adv. Enzymol.* 35:321-442
- 150) **MOULIN J.P. and M. VALADE.** 1983. Le levurage lors de l'élaboration des vins de base destinés à la champagnisation. *Le Vigneron Champenois* 5:237-249
- 151) **MSADEK T., F. KUSNT, D. HENNER, A. KLIER, G. RAPOPORT, and R. DEDONDER.** 1990. Signal transducing pathway controlling synthesis of a class of degradative enzymes in *Bacillus subtilis* : expression of the regulatory genes and analysis of mutations in *degS* and *degU*. *J. Bact.* 172:824-834
- 152) **MULLIGAN M.E. and W.R. McCLURE.** 1986. Analysis of the occurrence of promoter sites in DNA. *Nucl. Acids Res.* 14:109-126
- 153) **MULROONEY S.B., M.J. LYNCH, H.L.T. MOBLEY, and R.P. HAUSINGER.** 1988. Purification, characterization, and

genetic organization of recombinant *Providencia stuartii* urease expressed by *Escherichia coli*. J. Bact. 170:2202-2207

154) MURRAY N.E., W.J. BRAMMAR, and K. MURRAY. 1977. Lambdoid phages that simplify the recovery of *in vitro* recombinants. Mol. Gen. Genet. 150:53-61

155) NAKANO M. and P. ZUBER. 1990. Molecular biology of antibiotic production in *Bacillus*. CRC Crit. Rev. Biotechnol. 10:223-240

156) NEUNER A., H.W. JANNASCH, S. BELKIN, and K.O. STETTER. 1990. *Thermococcus litoralis* sp. nov. : a new species of extremely thermophilic marine archaeobacteria. Arch. Microbiol. 153:205-207

157) NICHOLSON W.L., Y.K. PARK, T.M. HENKIN, M. WON, M.J. WEICKERT, J.A. GASKELL, and G.H. CHAMBLISS. 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

158) NICHOLSON W.L., D. SUN, and P. SETLOW. 1990. Studies of DNA topology during *Bacillus subtilis* sporulation. In "Genetics and Biotechnology of Bacilli. Vol. 3", p. 339-347, Academic Press, Boca Raton, Florida

159) NISHIO H., H. BIEBL, and M. MEINERS. 1983. Effect of pH on the production of acetone and butanol by *Clostridium acetobutylicum* in a minimum medium. J. Ferm. Technol. 61:101-104

160) O'BRIEN R.W. and J.G. MORRIS. 1971. Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. J. Gen. Microbiol. 68:307-318

161) OHTA S., M. YOHDA, M. ISHIZUKA, H. HIRATA, T. HAMAMOTO, H. OTAWARA-HAMAMOTO, K. MATSUDA, and Y. KAGAWA. 1988. Sequence and over expression of subunits of adenosine triphosphate synthase in thermophilic bacterium PS3. Biochem. Biophys. Acta 933:141-155

162) PEREGO M. and J.A. HOCH. 1987. Isolation and sequence of the *spoOE* gene : its role in initiation of sporulation in *Bacillus subtilis*. Mol. Microbiol. 1:125-132

163) PEYNAUD E. 1983. Connaissance et travail du vin. Dunod, Paris

164) PHAE C.G., M. SHODA, and H. KUBOTA. 1990. Suppressive effect of *Bacillus subtilis* and its products on phytopathogenic microorganisms. J. Ferm. Bioeng. 69:1-7

- 165) **PIGGOT P.J.** 1989. Revised genetic map of *Bacillus subtilis* 168. In "Regulation of Prokaryotic Development", p. 1-41, American Society for Microbiology, Washington D.C.
- 166) **PREISS J.** 1984. Bacterial glycogen synthesis and its regulation. *Ann. Rev. Microbiol.* 38:419-458
- 167) **PRESCOTT S.C. and C.G. DUNN.** 1949. *Industrial Microbiology.* McGraw and Hill Book Company, Inc., New York
- 168) **PRIEST F.G.** 1984. Aspects of Microbiology : extracellular enzymes. American Society for Microbiology, Washington D.C.
- 169) **PRIEST F.G.** 1989. Products and applications of *Bacilli*. In "Biotechnology Handbooks : *Bacillus*", p. 293-320, C.R. Harwood ed., Plenum Press, New York
- 170) **PUTZER H., A.A. BRAKHAGE, and M. GRUNBERG-MANAGO.** 1990. Independent genes for two threonyl-tRNA synthetases in *Bacillus subtilis*. *J. Bact.* 172:4593-4602
- 171) **QUIVEY R.G., Jr., R.C. FAUSTOFERRI, W.A. BELLI, and J.S. FLORES.** 1991. Polymerase chain reaction amplification, cloning, sequence determination and homologies of streptococcal ATPase-encoding DNAs. *Gene* 97:63-68
- 178) **REDDY P., A. PETERKOFESKY, and K. MCKENNEY.** 1985. Translational efficiency of the *Escherichia coli* adenylate cyclase gene : mutating the UUG initiation codon to GUG or AUG results in increased gene expression. *Proc. Natl. Acad. Sci. USA* 82:5656-5660
- 179) **RIBEREAU-GAYON J., E. PEYNAUD, P. SUDRAUD, and P. RIBEREAU-GAYON.** 1975. *Sciences et Techniques du vin.* Dunod, Paris
- 180) **RIBES V., K. ROMISCH, A. GINER, B. DOBBERSTEIN, and D. TOLLERVEY.** 1990. *Escherichia coli* 4.5s RNA is part of a ribonucleoprotein particle that has properties related to signal recognition particle. *Cell* 63:591-600
- 181) **RIGBY P.W.J., M. DIECKMANN, C. RHODES, and P. BERG.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J. Mol. Biol.* 113:237-251
- 182) **ROGERS E.J., N.P. AMBULOS, Jr., and P.S. LOVETT.** 1990. Complementarity of *Bacillus subtilis* 16s RNA with sites of antibiotic-dependent ribosome stalling in *cat* and *erm* leaders. *J. Bact.* 172:6282-6290
- 183) **RYLATT D.B., D.B. KEECH, and J.C. WALLACE.** 1977. Pyruvate carboxylase : isolation of the biotin containing

tryptic peptide and the determination of its primary sequence. Arch. Biochem. Biophys. 183:113-122

184) **SAMBROOK J., E.T. FRITSCH, and T. MANIATIS.** 1989. Molecular Cloning Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

185) **SAMOLS D., C.G. THORNTON, V.L. MURTIF, G.K. KUMAR, F.C. HAASE, and H.G. WOOD.** 1988. Evolutionary conservation among biotin enzymes. J. Biol. Chem. 262:6464-6464

186) **SANGER F., S. NICKLEN, and A.R. COULSON.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467

187) **SCHAEFFER P.** 1969. Sporulation and the production of antibiotics, exoenzymes and endotoxins. Bact. Rev. 33:48-71

188) **SCHMID M.** 1988. Structure and function of the bacterial chromosome. Trends Biochem. Sci. 13:131-135

189) **SCHWARZ E., D. OESTERHELT, H. REINKE, K. BEYREUTHER, and P. DIMROTH.** 1988. The sodium ion translocating oxaloacetate decarboxylase of *Klebsiella pneumoniae*. J. Biol. Chem. 263:9640-9645

190) **SERRAHIMA-ZIEGER M. and H. MONTEIL.** 1978. Membrane ATPase of *Bacillus subtilis*. I. Purification and properties. Biochim Biophys. Acta 502:445-457

191) **SERROR P., G. DAMIANI, E. ALVAREZ, V. AZEVEDO, V. SGARAMELLA, and S.D. EHRLICH.** 1990. Construction of a *Bacillus subtilis* YAC library for physical mapping. International Conference on the *B. subtilis* genome, Institut Pasteur, Paris, Poster P71

192) **SHAMEL R.E. and J.J. CHOW.** 1988. Biotech's potential impact on the chemical industry. Biotechnology 6:681-682

193) **SHARP P.M., D.G. HIGGINS, D.C. SHIELDS, K.M. DEVINE, and J.A. HOCH.** 1990. *Bacillus subtilis* gene sequences. In "Genetics and Biotechnology of Bacilli. Vol. 3.", p. 89-93, Academic Press, Boca Raton, Florida

194) **SHARROCK R.A. and T. LEIGHTON.** 1982. Suppression of defective sporulation phenotypes by the *Bacillus subtilis* mutation rev-4. Mol Gen. Genet. 186:432-438

195) **SHOJI J.** 1978. Recent chemical studies on peptide antibiotics from the genus *Bacillus*. Adv. Appl. Microbiol. 24:187-214

- 196) **SLOMA A., C.F. RUDOLPH, G.A. RUFO, Jr., B.J. SULLIVAN, K.A. THERIAULT, D. ALLY, and J. PERO.** 1990a. Gene encoding a novel extracellular metalloprotease in *Bacillus subtilis*. *J. Bact.* 172:1024-1029
- 197) **SLOMA A., G.A. RUFO, Jr., C.F. RUDOLPH, B.J. SULLIVAN, K.A. THERIAULT, and J. PERO.** 1990c. Cloning and deletion of the gens for three minor extracellular proteases of *Bacillus subtilis*. In "Genetics and Biotechnology of *Bacilli*. Vol 3", p. 295-302, Academic Press, Boca Raton, Florida
- 198) **SMITH C.L., J. ECONOME, A. SCHUTT, S. KLCO, and C.R. CANTOR.** 1987. A physical map of the *Escherichia coli* K12 genome. *Science* 236:1448-1453
- 199) **SMITH V., C.M. BROWN, A.T. BANKIER, and B.G. BARRELL.** 1990. Semiautomated preparation of DNA templates for large scale sequencing projects. *DNA Sequence* 1:73-78
- 200) **SNOW R.** 1982. Genetic improvement of wine yeast. In "Yeasts genetics, fundamental and applied aspects", J.F.T. Spencer, M. Spencer, and A.R.W. Smith eds., Springer Verlag, New York, New York
- 201) **SONI B.K., P. SOUCAILLE, and G. GOMA.** 1987. Continuous acetone butanol fermentation : influence of vitamins on the metabolic activity of *Clostridium acetobutylicum*. *App. Env. Biotechnol.* 25:1-5
- 202) **SPIZIZEN J.** 1958. Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. *Proc. Natl. Acad. Sci. USA* 44:1072-1078
- 203) **STADEN R.** 1979. A strategy of DNA sequencing employing computer programs. *Nucl. Acids Res.* 6:2601-2610
- 204) **STADEN R.** 1980. A new computer method for the storage and manipulation of DNA gel reading data. *Nucl. Acids Res.* 8:3673-3694
- 205) **STADEN R.** 1987. Computer handling of DNA sequencing projects. In "Nucleic Acids and Protein Sequence Analysis : a practical approach", p. 173-217, M.J. Bishop and C.J. Rawlings eds., IRL Press, Oxford, England
- 206) **STUCKLE E.E., C. EMMRICH, U. GROB, and P.J. NIELSEN.** 1990. Statistical analysis of nucleotide sequences. *Nucl. Acids Res.* 18:6641-6647
- 207) **SULLIVAN M.A., R.E. YASBIN, and F.E. YOUNG.** 1984. New shuttle vectors for *Bacillus subtilis* and *Escherichia coli* which allow rapid detection of inserted fragments. *Gene* 29:21-26

- 208) SUTTON M.R., R.R. FALL, A.M. NERVI, A.W. ALBERTS, P.R. VAGELOS, and R.A. BRADSHAW. 1977. Amino acid sequence of *Escherichia coli* biotin carboxyl carrier protein (9100). J. Biol. Chem. 252:3934-3940
- 209) TABATA S., A. HIGASHITANI, M. TAKANAMI, K. AKIYAMA, Y. KOHARA, Y. NISHIMURA, A. NISHIMURA, S. YASUDA, and Y. HIROTA. 1989. Construction of an ordered cosmid collection of the *Escherichia coli* K12-W3110 chromosome. J. Bact. 171:1214-1218
- 210) TABOR S. and C.C. RICHARDSON. 1987. DNA sequence analysis with a modified T7 DNA polymerase. Anal. Chem. 84:4767-4771
- 211) TAKAI T., K. WADA, and T. TANABE. 1987. Primary structure of the biotin-binding site of chicken liver acetyl coA carboxylase. FEBS Letters 212:98-102
- 212) TAKISHIMA K., T. SUGA, and G. MAMIYA. 1988. The structure of jack bean urease. Eur. J. Biochem. 175:151-165
- 213) TERRACIANO J.S. and E.R. KASHKET. 1986. Intracellular conditions required for initiation of solvent production by *Clostridium acetobutylicum*. App. Env. Microbiol. 52:86-91
- 214) TIZARD R., R.L. CATE, K.L. RAMACHANDRAN, M. WYSK, J.C. VOYTA, O.J. MURPHY, and I. BRONSTEIN. 1990. Imaging of DNA sequences with chemiluminescence. Proc. Natl. Acad. Sci. USA 87:4515-4518
- 215) TRACH K., J.W. CHAPMAN, P. PIGGOT, D. LECOQ, and J.A. HOCH. 1985. Deduced product of the stage 0 sporulation gene *spo0F* shares homology with the *spo0A*, *ompR*, and *sfrA* proteins. Proc. Natl. Acad. Sci. USA 82:7260-7264
- 216) TRACH K., J.W. CHAPMAN, P. PIGGOT, D. LECOQ, and J.A. HOCH. 1988. Complete sequence and transcriptional analysis of the *spo0F* region of the *Bacillus subtilis* chromosome. J. Bact. 170:4194-4208
- 217) TRIEU-CUOT P. and P. COURVALIN. 1983. Nucleotide sequence of the *Streptococcus faecalis* plasmid gene encoding the 3'5"-aminoglycoside phosphotransferase type III. Gene 23:331-341
- 218) UMBARGER E. and B.D. DAVIS. 1962. Pathways of amino acid biosynthesis. In "The bacteria, a treatise on structure and function. Vol 3 : Biosynthesis", p. 167-251. I.C. Gunsalus, and R.Y. Stanier eds., Academic Press, New York, New York

- 219) **VAGNER V. and S.D. EHRLICH.** 1988. Efficiency of homologous DNA recombination varies along the *Bacillus subtilis* chromosome. *J. Bact.* 170:3978-3982
- 220) **VANDEYAR M.A. and S.A. ZAHLER.** 1986. Chromosomal insertions of Tn917 in *Bacillus subtilis*. *J. Bact.* 167:530-534
- 221) **VENTRA L. and A.S. WEISS.** 1989. Transposon mediated restriction mapping of the *Bacillus subtilis* chromosome. *Gene* 78:29-36
- 222) **VERACHTERT H. and R. de MOT.** 1990. *Yeast. Biotechnology and biocatalysis.* Marcel Dekker Inc., New York, New York
- 223) **VEZINHET F.** 1981. Quelques applications de la génétique des levures en oenologie. *Bulletin de l'O.I.V.* 608:831-843
- 224) **VEZINHET F. et S. LACROIX.** 1984. Marquage génétique des levures : outil de contrôle des fermentations en souche pure. *Bulletin de l'O.I.V.* 643/644:759-773
- 225) **VEZINHET F. et G. LOISEAU.** 1985. Le marquage des souches de levures sélectionnées, pourquoi, comment ? *Revue des Oenologues et des techniques viticoles et oenologiques* 37:19-20
- 226) **VEZINHET F.** 1985. Le marquage génétique de souches de levures oenologiques. *Revue Française d'Oenologie* 97:47-51
- 227) **VOET D. and J.G. VOET.** 1990. *Biochemistry,* p. 544-560, John Wiley and Sons, New York, New York
- 228) **von HEIJNE G.** 1987. *Sequence Analysis in Molecular Biology. Treasure Trove or Trivial Pursuit.* Academic Press, San Diego, California
- 229) **von MEYENBURG K., B.B. JORGENSEN, and B. van DEURS.** 1984. Physiological and morphological effects of overproduction of membrane bound ATP synthase in *Escherichia coli* K12. *EMBO J.* 3:1791-1797
- 230) **WALKER J.E., I.M. FEARNLEY, N.J. GAY, B.W. GIBSON, F.D. NORTHROP, S.J. POWELL, M.J. RUNSWICK, M. SARASTE, and V.L.J. TYBULEWICZ.** 1985. Primary structure and subunit stoichiometry of F<sub>1</sub> ATPase from bovine mitochondria. *J. Mol. Biol.* 184:677-701
- 231) **WALTON M.T. and J.L. MARTIN.** 1979. Production of butanol/ acetone by fermentation. *Microbial Technology* 2<sup>nd</sup> ed. Vol. I

- 232) **WANG G. and D.I.C. WANG.** 1984. Elucidation of growth inhibition and acetic acid production by *Clostridium thermoaceticum*. *App. Env. Microbiol.*
- 233) **WEICKERT M.J. and G.H. CHAMBLISS.** 1990. Site directed mutagenesis of a catabolite repression operator sequence in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* 87:6238-6242
- 234) **WENZEL R. and R. HERRMANN.** 1988. Physical mapping of the *Mycoplasma pneumoniae* genome. *Nucl. Acids Res.* 16:8323-8336
- 235) **WOESE C.R.** 1987. Bacterial evolution. *Microbiol. Rev.* 51:221-271
- 236) **WOOD H.G. and R.E. BARDEN.** 1977. Biotin enzymes. *Ann. Rev. Biochem.* 46:385-413
- 237) **WOOD H.G., F.R. HARMON, B. WUHR, K. HUBNER, and F. LYNEN.** 1980. Comparison of the biotination of apotranscarboxylase and its aposubunits. *J. Biol. Chem.* 255:7397-7409
- 238) **WOODS D.R. and D.T. JONES.** 1986. Physiological responses of *Bacteroides* and *Clostridia* strains to environmental stress factors. *Adv. Microbiol. Phys.* 27:1-64
- 239) **WORKMAN W.E., J.H. McLINDEN, and D.H. DEAN.** 1986. Genetic engineering applications to biotechnology in the genus *Bacillus*. *CRC Crit. Rev. Biotechnol.* 3:199-234
- 240) **ZEPH L.R., X. LIN, and G. STOTZKY.** 1991. Comparison of three nonradioactive and a radioactive DNA probes for the detection of target DNA by DNA hybridization. *Curr. Microbiol.* 22:79-84

