



**HAL**  
open science

# Système protéolytique de surface de "streptococcus thermophilus": variabilité des capacités d'hydrolyse des caséines: caractérisation d'un nouveau système de transport de peptides

Nawara Jameh Galia

► **To cite this version:**

Nawara Jameh Galia. Système protéolytique de surface de "streptococcus thermophilus": variabilité des capacités d'hydrolyse des caséines: caractérisation d'un nouveau système de transport de peptides. Alimentation et Nutrition. Université de Lorraine, 2012. Français. NNT: 2012LORR0197. tel-01749262

**HAL Id: tel-01749262**

**<https://hal.univ-lorraine.fr/tel-01749262>**

Submitted on 29 Mar 2018

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

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



## AVERTISSEMENT

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

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

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

Contact : [ddoc-theses-contact@univ-lorraine.fr](mailto:ddoc-theses-contact@univ-lorraine.fr)

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

[http://www.cfcopies.com/V2/leg/leg\\_droi.php](http://www.cfcopies.com/V2/leg/leg_droi.php)

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



**UNIVERSITÉ  
DE LORRAINE**



**INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE**  
Ecole Nationale Supérieure d'Agronomie et des Industries Alimentaires  
Ecole doctorale Science et Ingénierie des Ressources, Procédés, Produits, Environnement

## **THESE**

Présentée pour l'obtention du grade de

**Docteur de l'Institut National Polytechnique de Lorraine**

**Discipline : Sciences Agronomiques**

Soutenue le 20 juin 2012

Par

**Nawara JAMEH**

**Système protéolytique de surface de *Streptococcus thermophilus* :**  
- **variabilité des capacités d'hydrolyse des caséines**  
- **caractérisation d'un nouveau système de transport de peptides**

### **Membres du jury**

#### Rapporteurs

M. Jean-François CAVIN, professeur, université de Bourgogne, Dijon

M. Vincent JUILLARD, directeur de recherche, INRA, Jouy en Josas

#### Examineurs

Mme Véronique MONNET, directrice de recherche, INRA, Jouy en Josas

Mme Annie DARY, professeur, Université de Lorraine, (Directrice de thèse)

Mme Clarisse PERRIN, maître de conférences, Université de Lorraine, (Co-directrice de thèse)

Unité de Recherche Animal et Fonctionnalités des Produits Animaux  
Equipe Protéolyse et Biofonctionnalité des Protéines et Peptides  
Faculté des Sciences et Techniques, Université de Lorraine, 54500 Vandoeuvre-lès-Nancy



## Avant-Propos

Je tiens à exprimer mes plus sincères remerciements aux différentes personnes qui m'ont permis de mener à bien mon travail de thèse.

En premier lieu, mes remerciements s'adressent aux rapporteurs et membres de la commission d'examen qui m'ont fait l'honneur d'accepter de juger ce travail :

- M. Jean-François CAVIN, professeur, université de Bourgogne, Dijon
- M. Vincent JUILLARD, directeur de recherche, INRA, Jouy en Josas
- Mme Véronique MONNET, directrice de recherche, INRA, Jouy en Josas

Je tiens à adresser un immense merci au Pr. Guido RYCHEN, directeur de l'Unité de Recherche Animal et Fonctionnalité des Produits Animaux, pour m'avoir permis de réaliser mes travaux de recherche au sein de l'URAFPA.

Je tiens à remercier ma directrice de thèse, Pr. Annie DARY-MOUROT, pour la confiance qu'elle m'a accordée en acceptant d'encadrer ce travail. Son dynamisme, sa persévérance, sa passion pour la recherche ont fait d'elle une directrice idéale de thèse. Merci pour ses multiples conseils et pour les nombreuses heures, soirées et samedis qu'elle a consacrés à mes réunions de travail et à la correction de mon manuscrit. Merci pour sa patience et sa grande disponibilité. Enfin, j'ai été extrêmement sensible à ses qualités humaines d'écoute et de compréhension tout au long de ce travail.

Je remercie vivement le Dr. Clarisse PERRIN, maître de conférences, pour avoir co-encadré mon travail de doctorat. Je tiens à la remercier pour m'avoir guidée dans cette thèse, pour tous les conseils et les informations précieuses qu'elle m'a donnés au cours des discussions. Je la remercie de m'avoir soutenue et aidée pour mener à bien mon travail. Qu'elle soit assurée de ma reconnaissance la plus vive.

Je tiens à remercier les Dr. Magali GENAY, Émeline ROUX et Laurent MICLO, maîtres de conférences à l'URAFPA et le Dr. Gérard GUEDON, maître de conférences au LGM (Laboratoire de Génétique Microbienne, Université de Lorraine), pour m'avoir fait profiter de leur expertise afin de mener à bien mon travail.

Je tiens aussi à remercier les membres de l'équipe PB2P "Protéolyse-Biofonctionnalité des Protéines et Peptides", à savoir Céline CAKIR-KIEFER, Jean-Michel GIRARDET, Chantal POIRSON, Katy POIRET, Anne LAGRANGE, Catherine CORBIER, Alain DRIOU, Gérard HUMBERT, Yves LE ROUX, Emilie LORSON, Yvonne ROUSSEL qui ont toujours été de bons conseils avec moi et pour leur amitié forte.

Je tiens spécialement à remercier Delphine DUFOUR de m'avoir guidée dès mon arrivée. Merci à Faïza ZIDANE, Aurélie MATÉOS, Maira JUNJUA, Xavier LECOMTE, Nazir AHMAD, François BAGLINIERE, Zeeshan HAFEEZ, et Oun KI CHANG pour votre amitié et pour votre bonne humeur. A ceux qui ne sont pas encore Docteur, je vous souhaite bon courage, beaucoup de résultats et de publications.

Je remercie mon cher époux Wessam pour son soutien quotidien indéfectible et son enthousiasme communicatif tant sur le plan du travail que de la vie en général.

Je dédie cette thèse aux âmes de mes parents, mon père trop tôt disparu, et ma mère qui m'a quittée avant de connaître la fin de cette aventure. J'espère que vous êtes fiers de moi là où vous êtes.

Enfin, à mes sœurs et mes frères, j'aurais aimé vous avoir à mes côtés et je vous remercie pour votre soutien à distance.

<b>Abréviations</b> .....	<b>9</b>
<b>Liste des publications et communications</b> .....	<b>11</b>
<b>Table des Figures</b> .....	<b>13</b>
<b>Table des Tableaux</b> .....	<b>15</b>
<b>I. Synthèse Bibliographique</b> .....	<b>17</b>
<b>1. Les bactéries lactiques</b> .....	<b>19</b>
1. 1. Intérêt général des bactéries lactiques en industrie alimentaire .....	19
1. 2. Les bactéries lactiques : un ensemble hétérogène.....	19
1. 3. Classification phylogénétique des bactéries lactiques.....	20
<b>2. <i>Streptococcus thermophilus</i> : une espèce représentative de la diversité des bactéries lactiques</b> .....	<b>22</b>
2. 1. Positionnement de <i>Streptococcus thermophilus</i> au sein des streptocoques .....	22
2. 2. Positionnement de <i>Streptococcus thermophilus</i> au sein des bactéries lactiques .....	25
<b>3. Analyse du génome : une clé pour comprendre l'adaptation de <i>Streptococcus thermophilus</i> à la niche écologique lait</b> .....	<b>25</b>
3. 1. Caractéristiques du génome de <i>Streptococcus thermophilus</i> .....	25
3. 2. Transfert génétique horizontal.....	27
3. 2. 1. Présence d'éléments génétiques contribuant aux évènements de transfert horizontal dans le génome de <i>Streptococcus thermophilus</i> .....	31
3. 3. Perte de gènes de virulence et acquisition de nouvelles fonctions permettant l'adaptation au lait.....	32
3. 3. 1. Evolution du génome de l'espèce <i>Streptococcus thermophilus</i> par perte de fonctions.....	32
3. 3. 1. 1. Perte de voies métaboliques.....	33
3. 3. 1. 2. Perte de gènes de virulence.....	35
3. 3. 2. Acquisition de nouvelles fonctions chez <i>Streptococcus thermophilus</i> .....	37
<b>4. Exigences nutritionnelles de <i>Streptococcus thermophilus</i></b> .....	<b>38</b>
4. 1. Exigence en vitamines .....	39
4. 2. Exigence en source azotée.....	39
4. 2. 1. Système protéolytique de <i>Streptococcus thermophilus</i> .....	40
4. 2. 1. 1. Protéases de paroi .....	40
4. 2. 1. 2. Les systèmes de transports des peptides.....	44
4. 2. 1. 2. 1. Les systèmes de transport de peptides bactériens .....	44
A. Les transporteurs de type ABC .....	45
a. Caractéristiques générales des systèmes de transport ABC : Organisation en domaines .....	46
b. La famille de transporteurs de type ABC des peptides chez les bactéries .....	48
b.1. Caractéristiques des protéines de liaison aux peptides.....	48
b.2. Caractéristiques des deux domaines transmembranaires .....	50
b.3. Caractéristiques des deux domaines de fixation de l'ATP.....	51
B. Mécanisme de transport .....	52
4. 2. 1. 2. 2. Les systèmes de transports des peptides chez <i>Streptococcus thermophilus</i> .....	54
A. Transporteur de peptides proton dépendant .....	54
B. Système de transport de peptides ATP dépendant .....	54

a. Organisation génétique du système de transport des oligopeptides .....	56
b. Spécificité de substrat des transporteurs de peptides .....	58
c. Mode d'affinité des protéines de liaison aux peptides.....	61
d. Différents rôles des transporteurs des peptides.....	62
❖ Rôle dans la nutrition azotée.....	62
❖ Rôle dans le renseignement sur les conditions environnementales .....	63
4. 2. 1. 3. Les peptidases intracellulaires de <i>Streptococcus thermophilus</i> .....	65
4. 2. 1. 4. Les transporteurs d'acides aminés libres .....	67
4. 3. Activité uréasique chez <i>Streptococcus thermophilus</i> .....	67
4. 3. 1. L'uréase .....	68
4. 3. 2. Organisation génétique des gènes codant l'uréase.....	71
4. 3. 3. Importance de l'uréase.....	73
4. 3. 4. Régulation génétique de l'opéron uréase .....	75
<b>II. Objectifs du travail .....</b>	<b>77</b>
<b>III. Matériels et Méthodes .....</b>	<b>81</b>
<b>1. Souches, milieux et conditions de culture .....</b>	<b>83</b>
1. 1. Souches bactériennes utilisées et conditions de conservation.....	83
1. 2. Milieux de culture.....	83
<b>2. Extraction d'ADN génomique et ARN totaux, réactions de PCR, conditions d'électrophorèse et séquençage .....</b>	<b>88</b>
2. 1. Extraction de l'ADN génomique de <i>Streptococcus thermophilus</i> .....	88
2. 2. Extraction des ARN totaux de <i>Streptococcus thermophilus</i> .....	88
2. 2. 1. Récupération des cellules bactériennes.....	88
2. 2. 2. Extraction et quantification de l'ARN total .....	89
2. 3. Design des amorces .....	90
2. 4. Conditions d'amplification par PCR.....	90
2. 4. 1. Réactions standard de PCR.....	90
2. 4. 2. Réaction de Long-Range PCR .....	95
2. 5. PCR en temps réel .....	96
2. 5. 1. Préparation des ADN complémentaires.....	96
2. 5. 2. Quantification des niveaux d'expression .....	96
2. 6. Electrophorèse d'ADN.....	98
2. 7. Séquençage.....	98
<b>3. MLST .....</b>	<b>98</b>
<b>4. Construction du mutant LMD-9<math>\Delta</math>ots .....</b>	<b>99</b>
4. 1. Préparation d'un plasmide recombinant pGhost9 $\Delta$ ots .....	102
4. 1. 1. Amplification des deux inserts.....	102
4. 1. 2. Digestion enzymatique des amplicons et du plasmide pGhost9 .....	102
4. 1. 3. Ligation des deux inserts avec le plasmide .....	103
4. 2. Transformation d' <i>Escherichia coli</i> EC101 et sélection des transformants.....	103
4. 2. 1. Préparation des cellules électro-compétentes d' <i>Escherichia coli</i> EC101.....	103
4. 2. 2. Transformation d' <i>Escherichia coli</i> EC101 par électroporation.....	104
4. 2. 3. Sélection des transformants .....	104

4. 3. Transformation de la souche <i>Streptococcus thermophilus</i> LMD-9 et sélection des transformants.....	105
4. 3. 1. Préparation des cellules compétentes de <i>Streptococcus thermophilus</i> LMD-9 et électroporation.....	105
4. 3. 2. Intégration du plasmide pGhost9 $\Delta$ ots dans le chromosome de la souche LMD-9.....	106
4. 3. 3. Excision du plasmide.....	109
<b>5. Caractérisation du rôle du mutant .....</b>	<b>111</b>
5. 1. Estimation de la croissance bactérienne.....	111
<b>IV. Résultats et Discussion .....</b>	<b>113</b>
<b>Chapitre I.....</b>	<b>115</b>
<b>Chapitre I : Publication .....</b>	<b>117</b>
<b>Variabilité d'hydrolyse des caséines <math>\beta</math>, <math>\alpha_{s1}</math> et <math>\alpha_{s2}</math> de 10 souches de <i>Streptococcus thermophilus</i> et leurs peptides bioactifs résultant.....</b>	<b>117</b>
<b>Chapitre II .....</b>	<b>131</b>
<b>1. Recherche de systèmes de transport de peptides au sein d'une collection de souches de <i>Streptococcus thermophilus</i> par PCR.....</b>	<b>133</b>
1. 1. Etablissement des relations phylogénétiques entre 22 souches de <i>Streptococcus thermophilus</i> .....	134
1. 2. Recherche du système Ami .....	141
1. 3. Recherche du transporteur DtpT.....	148
1. 4. Détection du cluster <i>ots</i> (Oligopeptide Transport <i>Streptococcus thermophilus</i> ) de la famille des transporteurs de peptides et de nickel de type ABC .....	151
1. 4. 1. Détection et organisation du cluster <i>ots</i> .....	151
1. 4. 2. Analyses des régions flanquant le cluster <i>ots</i> chez <i>Streptococcus thermophilus</i> LMD-9 et comparaison avec celle des souches CNRZ1066, LMG18311, ND03, JIM8232.....	152
1. 4. 3. Paramètres physicochimiques de chaque protéine du cluster <i>ots</i> .....	154
1. 4. 4. Recherche <i>in silico</i> d'une organisation génétique similaire au cluster <i>ots</i> .....	154
1. 4. 5. Recherche de motif caractéristique de protéine constituant les systèmes de transport de peptides et de nickel de type ABC au sein du cluster <i>ots</i> .....	156
1. 4. 5. 1. Recherche du motif consensus de lipomodification chez OtsA .....	156
1. 4. 5. 2. Motif signature caractéristique de la famille des protéines affines aux peptides et au nickel chez OtsA.....	156
1. 4. 5. 3. Prédiction de structures secondaires des perméases transmembranaires OtsB et OtsC.....	158
1. 4. 5. 4. Différents motifs de l'ATPase OtsD .....	159
1. 4. 5. 5. Conclusion.....	159
1. 4. 6. Analyse des régions adjacentes au cluster <i>ots</i> .....	161
1. 4. 7. Analyses des régions flanquant les clusters homologues au cluster <i>ots</i> chez <i>Bifidobacterium</i> .....	166
1. 4. 8. Recherche par PCR-niché de la présence des gènes du cluster <i>ots</i> au sein d'une collection de 19 souches de <i>Streptococcus thermophilus</i> .....	170
<b>2. Séquençage du cluster <i>ots</i> et analyse des séquences obtenues chez les souches PB180 et PB302.....</b>	<b>173</b>
<b>3. Organisation transcriptionnelle du cluster <i>ots</i> et expression au cours de la croissance des souches PB302, PB180 et LMD-9 en milieu M17 .....</b>	<b>173</b>
<b>4. Etude de l'expression du cluster <i>ots</i> par PCR-quantitative au cours de la croissance des souches PB302 et PB180 en milieu M17 et lait.....</b>	<b>177</b>
A. Croissance de <i>Streptococcus thermophilus</i> PB302 et PB180 en lait et en milieu M17 .....	177

B. Quantification de l'expression des gènes <i>amiF</i> , <i>otsA</i> et <i>dtpT</i> à différents stades de croissance des souches PB18O et PB302 cultivées en milieu M17 par RT-PCR en temps réel .....	180
C. Quantification de l'expression des gènes <i>amiF</i> , <i>otsA</i> et <i>dtpT</i> à différents stades de croissance des souches PB18O et PB302 cultivées en milieu lait par RT-PCR en temps réel.....	181
D. Conclusion .....	182
<b>5. Rôle du système Ots .....</b>	<b>182</b>
5. 1. Croissance des souches LMD-9 et LMD-9 $\Delta$ <i>ots</i> en présence de source azotée différentes.....	183
5. 1. 1. Croissance des souches LMD-9 et LMD-9 $\Delta$ <i>ots</i> en lait .....	184
5. 1. 2. Croissance des souches LMD-9 et LMD-9 $\Delta$ <i>ots</i> en milieu M17 .....	185
5. 1. 3. Croissance des souches LMD-9 et LMD-9 $\Delta$ <i>ots</i> en milieu YL.....	186
5. 1. 4 Conclusion.....	187
5. 2. Croissance des souches LMD-9 et LMD-9 $\Delta$ <i>ots</i> en milieu chimiquement défini supplémenté en peptides... ..	188
<b>6. Discussion .....</b>	<b>192</b>
6. 1. Transfert horizontal chez <i>Streptococcus thermophilus</i> .....	193
6. 1. 1. Variabilité de la région génomique contenant le cluster <i>ots</i> .....	193
6. 1. 2. Détection de transposon composite contenant le cluster <i>ots</i> .....	194
6. 1. 3. Origine du cluster <i>ots</i> .....	195
6. 1. 4. Mécanismes de transfert horizontal impliqués dans l'acquisition de cluster <i>ots</i> .....	197
6. 2. Rôle du système Ots : utilité chez <i>Streptococcus thermophilus</i> .....	198
6. 2. 1. Rôle du système <i>ots</i> chez <i>Streptococcus thermophilus</i> .....	198
6. 2. 2. Acquisition des gènes conférant des avantages pour l'adaptation de <i>Streptococcus thermophilus</i> en lait .....	199
6. 2. 3. Intégration du cluster <i>ots</i> dans le schéma de régulation de <i>Streptococcus thermophilus</i> .....	200
6. 3. Devenir du système Ots chez <i>Streptococcus thermophilus</i> .....	201
<b>V. Perspectives .....</b>	<b>204</b>
<b>VI. Références bibliographiques.....</b>	<b>208</b>

# Abréviations

A : Adénine  
ABC : ATP Binding Cassette  
ADN : Acide Désoxyribonucléique  
ADNc : ADN complémentaire  
ARNr 16S : Acide Ribonucléique ribosomal de coefficient de sédimentation 16 Sveldberg  
ARNr 23S : Acide Ribonucléique ribosomal de coefficient de sédimentation 23 Sveldberg  
ARNm : Acide Ribonucléique messenger  
ATP : Adénosine triphosphate  
BCA : Bacto-Casamino Acides  
BL : Bactérie Lactique  
BLAST : Basic Local Aligement Search Tool  
BT : Bacto-Tryptone  
C :Cytosine  
CDS : Coding Sequence  
CG% : Pourcentage en cytosine et guanine  
CNRZ : Centre National de recherche zootechnique  
ComR : Competence Regulator  
ComS : Competence Signal  
Ct: Cycle threshod  
Da : Dalton  
DdlA : D-alanine D-alanine ligase  
DnaE : ADN polymérase III  
dNTP : Désoxyribonucléotide triphosphate  
D.O.<sub>650 nm</sub> : Densité optique à une longueur d'onde de 650 nanomètre  
DTT : Dithiothréitol  
FBP : Fructose-1,6-bisphosphate  
EDTA : Ethylène diamine tétra acétique  
EPS : Exopolysaccharides  
G : Guanine  
GlcK : glucose kinase  
GRAS : Generally Recognized As Safe  
GTP : Guanosine triphosphate  
h : heure  
SHP : Short Hydrophobic Peptide  
ICE : Integrative Conjugative Element  
IlvC : ketol-acid reductoisomérase  
IS : Séquences d'insertion  
ITS : Internal Transcribed Sequence

kb : Kilo base  
kDa : kiloDalton  
LC-ESI-MS-MS : Liquid chromatography electrospray ionisation tandem mass spectrometry  
MCD : Milieu Chimiquement Défini  
MLST : Multi Locus Sequence Typing  
MM : Masse Moléculaire  
ms : Milliseconde  
m/v : Rapport masse/volume  
Mpb : Méga paire de base  
NCBI : National Center for Biological Sequence Analysis  
ORF : Phase ouverte de lecture (Open Reading Frame)  
pb : Paire de bases  
PCR : Polymerase Chain Reaction  
PepO : Endopeptidase  
pH : potentiel hydrogène  
*pI* : Point Isoélectrique  
PMSF : Phénylméthylsulfonylfluorure  
PyrE : Orotate phosphoribosyl transférase  
RBS : Ribosome Binding Site  
RT : Reverse Transcription  
SDS : Sodium dodécil sulfate  
ssp : Sous-espèce  
T : Thymine  
TAE : Tris acétate  
TE : Tris EDTA  
ThrS : Threonyl-tRNA synthétase  
Tkt : Transkétolase  
Tris :  $\alpha\alpha\alpha$ -Tris (hydroxyméthyl) méthylamine  
U : Uracile  
UV : Ultraviolet  
v/v : Rapport volume/volume  
YL : Yeast extract-Lactose

# Liste des publications et communications

## I- Publications

### A. Article sur le travail de thèse

Variability of hydrolysis of  $\beta$ -,  $\alpha_{s1}$ - and  $\alpha_{s2}$ -caseins by 10 strains of *Streptococcus thermophilus* and resulting bioactive peptides

Miclo, L., Roux, E., Genay, M., Brusseau-Lorson, E., Poirson, C., **Jameh, N.**, Perrin, C. and Dary A.

J Agric Food Chem. 2012 Jan 18;60(2):554-65.

### B. Article sur le travail de master

Can the mammopathogenic *Escherichia coli* P4 strain have a direct role on the caseinolysis of milk observed during bovine mastitis ?

Dufour D, **Jameh N**, Dary A, Le Roux Y.

J Dairy Sci. 2009 Apr;92(4):1398-403.

## II- Communications internationales avec actes

1. Variability of the proteolytic properties in *Streptococcus thermophilus*.

Galia, W., Perrin, C., **Jameh, N.**, Genay, M., Dary, A.

4th IDF Dairy Science and Technology Week, 20-24 Avril 2009, Rennes, France

2. Mise en évidence d'un système du transport de peptides et du Nickel potentiel chez *Streptococcus thermophilus*

**Jameh, N.**, Galia, W., Perrin, C., Dary, A.

Séminaire Ecole Doctorale RB2E, 28 janvier 2010, Nancy, France

3. Recherche du rôle de l'opéron *ots* chez *Streptococcus thermophilus*

**Jameh, N.**, Perrin, C., Carré, J., Galia, W., Roux, E., Dary, A.

XVIIe Colloque du Club des bactéries lactiques (CBL). 27-29 octobre 2010, Nancy



# Table des Figures

Figure 1 : Distances phylogénétiques entre les principaux genres constituant les bactéries lactiques basées sur les séquences des ARNr 16S (Stiles et Holzapfel, 1997).....	21
Figure 2 : Résultat de la comparaison des génomes de <i>Streptococcus thermophilus</i> et d'autres espèces de streptocoques pathogènes (Lefébure et Stanhope, 2007).....	24
Figure 3 : Représentation schématique de la structure des éléments transposables bactériens (Kung <i>et al.</i> , 2010).....	29
Figure 4 : Représentation schématique du transfert conjugatif d'un ICE transféré sous forme simple brin et s'intégrant grâce à une recombinase site-spécifique à tyrosine (d'après Burrus <i>et al.</i> , 2002b).....	30
Figure 5 : Îlot génomique contenant le gène <i>prtS</i> avec la région chromosomique adjacente chez <i>Streptococcus thermophilus</i> et <i>Streptococcus suis</i> (Delorme <i>et al.</i> , 2010).....	37
Figure 6 : Représentation schématique du système protéolytique chez <i>Streptococcus thermophilus</i> (d'après Hols <i>et al.</i> , 2005).....	41
Figure 7 : Représentation schématique des différents domaines composant les protéases de paroi, (Fernandez-Espla <i>et al.</i> , 2000 ; Savijoki <i>et al.</i> , 2006 ; Siezen, 1999).....	42
Figure 8 : Organisation structurale des domaines de divers transporteurs ABC (Biemans-Oldehinkel <i>et al.</i> , 2006).....	47
Figure 9 : Représentation schématique de la maturation des lipoprotéines chez les bactéries à Gram positif (Braun et Wu, 1994).....	49
Figure 10 : Structure d'une protéine OBP «Oligopeptides Binding Protein» liée à un peptide (Levdikov <i>et al.</i> , 2005).....	50
Figure 11 : Organisation générale des motifs conservés d'un domaine NBD (Nucleotide-Binding Domain) chez les transporteurs ABC.....	52
Figure 12 : Modèle de transport des peptides par le système Opp chez <i>Lactococcus lactis</i> (Doeven <i>et al.</i> , 2008 ; Lanfermeijer <i>et al.</i> , 1999).....	53
Figure 13 : Organisation des gènes codant les protéines des transporteurs d'oligopeptides chez différentes espèces bactériennes.....	57
Figure 14 : Représentation schématique du mécanisme de liaison du peptide à la protéine OppA de <i>Lactococcus lactis</i> (Detmers <i>et al.</i> , 2000).....	61
Figure 15 : Modèle d'activation de la compétence naturelle chez la souche <i>Streptococcus thermophilus</i> LMD-9 lors de la croissance en milieu chimiquement défini (Fontaine <i>et al.</i> , 2010).....	64
Figure 16 : Réactions de dégradation de l'urée à l'intérieur de la cellule bactérienne (Mobley et Hausinger, 1989).....	69
Figure 17 : Modèle d'activation de l'uréase (Mulrooney et Hausinger, 2003).....	70
Figure 18 : Organisation génétique de l'opéron uréasique chez (A) <i>Streptococcus salivarius</i> 57.I et (B) <i>Streptococcus thermophilus</i> LMD-9.....	72
Figure 19 : Représentation schématique de l'activité enzymatique impliquée dans la biosynthèse de l'aspartate et de la glutamine à partir des catabolites de l'urée chez <i>Streptococcus thermophilus</i> (d'après Arioli <i>et al.</i> , 2007, 2009).....	74
Figure 20 : Représentation schématique de la stratégie utilisée pour déléter les ORFs des gènes <i>otsA</i> et <i>otsB</i> de <i>Streptococcus thermophilus</i> LMD-9 à l'aide du plasmide pGhost9.....	101
Figure 21 : Représentation schématique des positions d'hybridation des amorces utilisées pour séquencer l'insert I-II porté par le plasmide pGhost9.....	105
Figure 22 : Représentation schématique des deux possibilités d'intégration du plasmide pGhost9 au sein du génome de la souche LMD-9.....	107

Figure 23 : Détection par PCR de l'intégration du plasmide pGhost9 $\Delta$ ots dans le chromosome de la souche <i>Streptococcus thermophilus</i> LMD-9 .....	108
Figure 24 : Recherche de la forme mutante $\Delta$ ots dans le chromosome de la souche <i>Streptococcus thermophilus</i> LMD-9.....	110
Figure 25 : Arbre phylogénétique de 22 souches de <i>Streptococcus thermophilus</i> basé sur les séquences concaténées de 8 «gènes de ménage » ( <i>glcK</i> , <i>ddlA</i> , <i>pepO</i> , <i>ilvC</i> , <i>thrS</i> , <i>tkl</i> , <i>pyrE</i> , <i>dnaE</i> ) et de l'ITS 16S-23S .....	140
Figure 26 : Représentation schématique de l'organisation génétique des gènes du système de transport de peptides Ami chez <i>Streptococcus thermophilus</i> CNRZ1066, LMG18311, ND03, LMD-9 et JIM8232 et positions des amorces utilisées pour les amplifier par PCR.....	143
Figure 27 : Détection par PCR des gènes <i>amiA2</i> (pistes 1 à 5), <i>amiA1</i> et <i>amiA3</i> (pistes 6 à 10) chez différentes souches de <i>Streptococcus thermophilus</i> .....	144
Figure 28 : Séquence nucléotidique du gène <i>ntpT</i> chez la souche <i>Streptococcus thermophilus</i> LMD-9 avec les positions putatives des promoteur et terminateur.....	149
Figure 29 : Structures tridimensionnelles de la perméase DtpT de <i>Streptococcus thermophilus</i> LMD-9 proposées par le logiciel de prédiction de structures transmembranaires TMpred .....	150
Figure 30 : Représentation schématique de la région contenant le cluster <i>ots</i> chez <i>Streptococcus thermophilus</i> LMD-9, CRZ1066, LMG18311, ND03 et JIM8232 .....	153
Figure 31 : Localisation du motif consensus de lipomodification (A) et du motif signature caractéristique de la famille des protéines affines aux peptides et au nickel (B) sur la séquence d'OtsA (signalée en rouge) et comparaison de ces motifs avec ceux des protéines affines aux oligopeptides AmiA1 et AmiA3 chez <i>Streptococcus thermophilus</i> LMD-9 .....	157
Figure 32 : Modèle topologique des perméases OtsB et OtsC de <i>Streptococcus thermophilus</i> LMD-9 déduit à l'aide du logiciel TMpred.....	158
Figure 33 : Séquence protéique de OtsD (double ATPase) de <i>Streptococcus thermophilus</i> LMD-9 déduite de la séquence nucléotidique du gène <i>otsD</i> .....	160
Figure 34 : Organisation de la région chromosomique contenant le cluster <i>ots</i> ou son homologue chez <i>Streptococcus thermophilus</i> LMD-9 (A) <i>Bifidobacterium bifidum</i> PRL2010 (B) <i>Bifidobacterium longum</i> DJO 10A (C) .....	162
Figure 35 : Comparaison de l'organisation génétique des gènes <i>udp</i> et <i>gly</i> chez <i>Streptococcus thermophilus</i> LMD-9 (A) et ND03 (B) .....	166
Figure 36 : Arbre phylogénétique des souches de <i>Streptococcus thermophilus</i> utilisées dans ce travail.....	172
Figure 37 : Recherche des transcrits du cluster <i>ots</i> chez <i>Streptococcus thermophilus</i> par RT-PCR.....	174
Figure 38 : Séquences nucléotidiques du cluster <i>ots</i> chez la souche <i>Streptococcus thermophilus</i> LMD-9 avec les régions promotrice et terminatrice.....	176
Figure 39 : Suivi de croissance en milieu M17 des souches <i>Streptococcus thermophilus</i> PB180 et PB302 par détermination de la densité cellulaire (mesure de DO à 650 nm) « A » et mesure du pH « B » .....	178
Figure 40 : Suivi de croissance en lait de <i>Streptococcus thermophilus</i> PB180 et PB302 par détermination de la densité cellulaire (mesure de DO à 480 nm) « A » (Thomas et Turner, 1977) et mesure du pH « B » ....	179
Figure 41 : Représentation schématique de l'organisation génétique des clusters <i>ots</i> et $\Delta$ ots .....	183
Figure 42 : Croissance des souches LMD-9 et LMD-9 $\Delta$ ots en lait .....	184
Figure 43 : Croissance des souches LMD-9 et LMD-9 $\Delta$ ots en milieu M17 .....	185
Figure 44 : Croissance des souches LMD-9 et LMD-9 $\Delta$ ots en milieu YL .....	186
Figure 45 : Croissance des souches LMD-9 et LMD-9 $\Delta$ ots en milieu chimiquement défini supplémenté en bacto-tryptone.....	189
Figure 46 : Croissance des souches LMD-9 et LMD-9 $\Delta$ ots en milieu chimiquement défini supplémenté en bacto-casamino acides.....	191
Figure 47 : Détection d'un éventuel motif de fixation de la protéine régulatrice CodY en amont des gènes <i>amiA1</i> , <i>amiA3</i> et <i>otsA</i> chez <i>Streptococcus thermophilus</i> LMD-9.....	202

# Table des Tableaux

Tableau 1 : Classement de 34 espèces de streptocoques établi sur la base des séquences des gènes codant les ARN 16S (Kawamura <i>et al.</i> , 1995).....	23
Tableau 2 : Caractéristiques des cinq génomes entièrement séquencés de <i>Streptococcus thermophilus</i> .....	26
Tableau 3 : Distribution des pseudogènes dans les différentes fonctions cellulaires chez <i>Streptococcus thermophilus</i> CNRZ1066 et LMG18311 (Bolotin <i>et al.</i> , 2004) .....	34
Tableau 4 : Identification des VGRs putatifs dans le génome de <i>Streptococcus thermophilus</i> CNRZ 1066 par comparaison avec ceux présents dans les génomes de <i>Streptococcus pyogenes</i> M1 et <i>Streptococcus pneumoniae</i> TIGR4 (Bolotin <i>et al.</i> , 2004).....	35
Tableau 5 : Pourcentages d'identité entre les protéines du système de transport d'oligopeptides Opt et Opp de <i>Lactococcus lactis</i> IL1403 d'après (Lamarque <i>et al.</i> , 2004).....	55
Tableau 6 : Pourcentages d'identité entre les protéines du système Ami de <i>Streptococcus thermophilus</i> ST18 et du système Opp de <i>Lactococcus lactis</i> IL1403 (d'après Garault <i>et al.</i> , 2002).....	59
Tableau 7 : Taille des peptides incorporés par des transporteurs de peptides bactériens de type ABC chez quelques espèces bactériennes .....	60
Tableau 8 : Tableau récapitulatif des peptidases caractérisées biochimiquement ou déduites de séquences codantes chez <i>Streptococcus thermophilus</i> .....	66
Tableau 9 : List des souches de <i>Streptococcus thermophilus</i> utilisées dans ce travail .....	84
Tableau 10 : Composition des milieux de culture utilisés pour la croissance de <i>Streptococcus thermophilus</i> ....	85
Tableau 11 : Milieux et solutions utilisés pour la préparation de cellules compétentes de la souche <i>Streptococcus thermophilus</i> LMD-9 .....	86
Tableau 12 : Composition des milieux de culture d' <i>Escherichia coli</i> .....	87
Tableau 13 : Séquence des amorces utilisées pour l'amplification de séquences d'ADN génomiques par PCR	91
Tableau 14 : Séquence des amorces utilisées pour l'amplification de séquences d'ADN par PCR .....	92
Tableau 15 : Amorces utilisées pour la réalisation d'amplification de séquences d'ADN génomiques par PCR pour faire MLST.....	93
Tableau 16 : Couples d'amorces utilisés dans les réactions de PCR.....	94
Tableau 17 : Couples d'amorces utilisés et taille des amplicons attendus pour l'amplification par PCR des séquences de « gènes de ménage » pour le typage par MLST .....	95
Tableau 18 : Génotype de la souche <i>Escherichia coli</i> EC101 utilisée pour la construction du mutant LMD-9 $\Delta$ ots.....	103
Tableau 19 : Répartition des différents allèles de 8 « gènes de ménage » chez des souches de <i>Streptococcus thermophilus</i> de la collection URAFFPA et des banques de donnés .....	135
Tableau 20 : Comparaison entre le nombre d'allèles trouvés pour chaque « gène de ménage » dans cette étude et celui obtenu à partir d'une collection de 27 souches par Delorme <i>et al.</i> en 2010.....	137
Tableau 21 : Pourcentages de divergence entre les allèles de chaque « gène de ménage » étudié .....	138
Tableau 22 : Tableau synthétique des pourcentages de divergences nucléotidiques maximales de chaque « gène de ménage » au sein de notre collection de souches de <i>Streptococcus thermophilus</i> par rapport à ceux obtenus par Delorme <i>et al.</i> (2010).....	139
Tableau 23 : Pourcentage d'identité entre les gènes <i>amiA</i> chez les cinq souches dont le génome a été séquencé .....	142
Tableau 24 : Classification des souches en fonction de leur pouvoir acidifiant en lait (d'après Galia <i>et al.</i> , 2009). Le type d'ITS, la présence et l'absence du gène <i>prtS</i> , l'activité caséinolytique ainsi que la présence du système <i>ami</i> sont mentionnés.....	146

Tableau 25 : Pourcentages d'identité entre les gènes <i>amiA</i> /protéines AmiA des souches ST14, CNRZ160, PB180 et PB302 et leurs homologues présents dans les génomes séquencés de <i>Streptococcus thermophilus</i> .....	147
Tableau 26 : Paramètres physicochimiques (masses moléculaires et points isoélectriques) calculés à partir de la composition des protéines OtsA, OtsB, OtsC et OtsD à l'aide du logiciel compute pI/Mw ( <a href="http://web.expasy.org/compute_pi/">http://web.expasy.org/compute_pi/</a> ) .....	154
Tableau 27 : Pourcentages d'identité entre les protéines Ots de <i>Streptococcus thermophilus</i> LMD-9 et leurs homologues présents au sein de cluster ayant une organisation génétique similaire à celle du cluster <i>ots</i> .....	155
Tableau 28 : Pourcentages d'identité entre les transposases des six IS localisées en amont et en aval du cluster <i>ots</i> chez <i>Streptococcus thermophilus</i> LMD-9.....	163
Tableau 29 : Pourcentages d'identité entre les séquences d'insertion adjacentes au cluster <i>ots</i> de <i>Streptococcus thermophilus</i> LMD-9 et leurs homologues présents dans la banque des données.....	165
Tableau 30 : Présence d'un cluster homologue au cluster <i>ots</i> au sein des génomes de <i>Bifidobacterium</i> disponibles dans les banques de données.....	168
Tableau 31 : Pourcentages d'identité entre les séquences protéiques déduites des gènes du cluster <i>opp</i> de <i>Bifidobacterium bifidum</i> PRL2010 et leurs homologues chez d'autres espèces du genre <i>Bifidobacterium</i> dont les génomes sont entièrement séquencés .....	169
Tableau 32 : Récapitulatif des différentes caractéristiques des 22 souches de <i>Streptococcus thermophilus</i> étudiées.....	171
Tableau 33 : Comparaison des niveaux <sup>a</sup> d'expression des gènes <i>amiF</i> , <i>dtpT</i> et <i>otsA</i> entre les souches PB302 et PB180 cultivées en milieu M17 .....	181
Tableau 34 : Comparaison des niveaux <sup>a</sup> d'expression des gènes <i>amiF</i> , <i>dtpT</i> et <i>otsA</i> entre les deux souches étudiées (souche PB302/souche PB180) cultivées en lait .....	181
Tableau 35 : Tableau récapitulatif des valeurs de pH et de D.O. <sub>650 nm</sub> obtenues après 10 h de croissance en milieu chimiquement défini supplémenté en bacto-tryptone .....	190
Tableau 36 : Tableau récapitulatif des valeurs de pH et de D.O. <sub>650 nm</sub> obtenues après 10 h de croissance en milieu chimiquement défini supplémenté en bacto-casamino acides .....	192

# **I. Synthèse Bibliographique**



# 1. Les bactéries lactiques

## 1. 1. Intérêt général des bactéries lactiques en industrie alimentaire

Le terme de bactéries lactiques (BL) désigne les bactéries capables de produire par fermentation de l'acide lactique à partir de glucides fermentescibles. Certaines d'entre-elles sont très utilisées en industrie agroalimentaire et interviennent dans de nombreuses fermentations alimentaires telles que les fabrications de fromage, yaourt, salamis, saucisses, jambons secs, choucroute, bières et pain (Adams, 1999).

Outre la modification de la saveur et de la texture de l'aliment grâce à la production d'acide lactique et de divers composés aromatiques tels que l'acétoïne, l'acétaldéhyde ou le diacétyle, les bactéries lactiques inhibent également le développement de la flore microbienne indésirable grâce à l'acidification du milieu et à la production de molécules bactéricides telles que les bactériocines.

## 1. 2. Les bactéries lactiques : un ensemble hétérogène

Les bactéries lactiques forment un groupe hétérogène. Il s'agit d'un ensemble de bactéries à Gram positif, généralement immobiles, asporulées, anaérobies mais aérotolérantes et qui se présentent sous forme de coques ou de bacilles. Généralement ces bactéries ne possèdent ni nitrate réductase, ni catalase (Desmazeud, 1983 ; Vandamme *et al.*, 1996). Cependant, *Lactobacillus sakei* et *Lactobacillus plantarum* possèdent une activité catalase hème-dépendante, qui peut être active dans les produits carnés (Hertel *et al.*, 1998 ; Mares *et al.*, 1994). En général, les BL ont des exigences nutritionnelles en vitamines, acides aminés, glucides et acides gras (Holzapfel *et al.*, 2001).

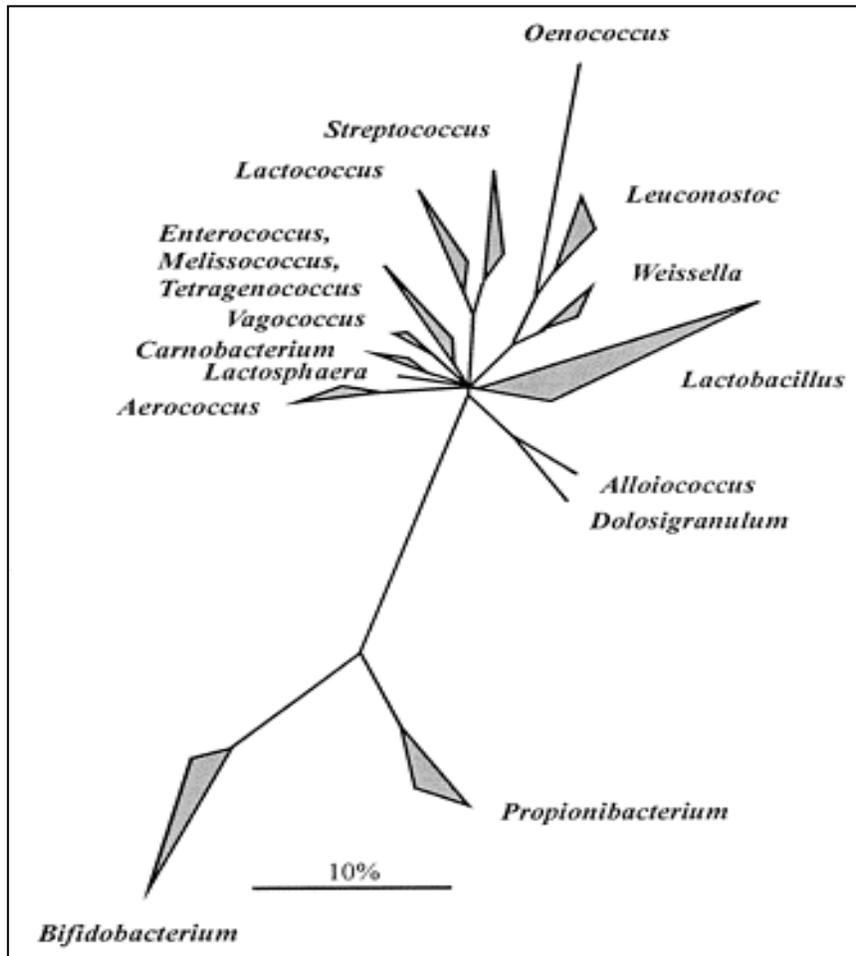
Selon les produits finaux de fermentations, les bactéries lactiques peuvent être classées en trois groupes : les bactéries lactiques homofermentaires (l'acide lactique représentant plus de 90% des produits de fermentation), les bactéries lactiques hétérofermentaires facultatives (production d'acide lactique ou d'acide lactique et d'acide acétique) et les bactéries lactiques hétérofermentaires strictes (production d'acide lactique, d'acide acétique ou d'éthanol et de CO<sub>2</sub>)

(Vandamme *et al.*, 1996).

### 1. 3. Classification phylogénétique des bactéries lactiques

La classification des bactéries lactiques proposée par (Stiles et Holzapfel, 1997) comporte 4 groupes phylogénétiques :

- I. Le premier est composé des genres : *Enterococcus*, *Melissococcus*, *Tetragenococcus*, *Vagococcus*, *Carnobacterium*, *Lactosphaera* et *Aerococcus*, genres isolés de différents habitats. Ainsi, *Enterococcus faecalis* et *Enterococcus faecium*, bactéries intestinales, dont certaines souches sont des pathogènes opportunistes de l'homme, ont été isolées d'eaux contaminées par des fèces et *Enterococcus seriolicida* a été isolée de cas d'infections de poissons (Aguirre et Collins, 1993 ; Kusuda *et al.*, 1991).
- II. Le deuxième groupe est constitué des genres *Streptococcus* et *Lactococcus* qui sont phylogénétiquement proches (Schleifer *et al.*, 1985). *Lactococcus lactis* est une bactérie très utilisée en fermentation fromagère. *S. thermophilus* est la seule espèce du genre *Streptococcus* à être utilisée comme ferment dans l'industrie laitière et à ne pas être pathogène. Elle est notamment utilisée pour la production de yaourts ou de fromage, en coculture avec d'autres BL.
- III. Le troisième groupe est formé des genres *Oenococcus*, *Leuconostoc* et *Weissella* (Collins *et al.*, 1993). Il comporte des bactéries importantes dans l'industrie agroalimentaire pour la fabrication du vin et de la choucroute.
- IV. Le quatrième groupe phylogénétique est constitué des genres *Lactobacillus* et *Pediococcus*. Si les trois genres *Bifidobacterium*, *Brevibacterium* et *Propionibacterium* sont aussi ajoutés au groupe des bactéries lactiques, de par la similarité de leurs propriétés physiologiques et biochimiques et de leur utilisation en industrie laitière, ces microorganismes sont phylogénétiquement éloignés des autres bactéries lactiques (Figure 1). Ce sont des bactéries à Gram positif dont le pourcentage en GC est supérieur à 50% (phylum des *Actinobacteria*), alors que les bactéries lactiques typiques ont une teneur en GC inférieure à 50% (phylum des *Firmicutes*) (Holzapfel *et al.*, 2001 ; Vandamme *et al.*, 1996).



**Figure 1 : Distances phylogénétiques entre les principaux genres constituant les bactéries lactiques basées sur les séquences des ARNr 16S (Stiles et Holzapfel, 1997)**

La barre indique une divergence de séquence de 10%. Les bactéries à bas GC (situées en haut de la figure) sont phylogénétiquement éloignées des bactéries à haut GC (situées en bas de la figure).

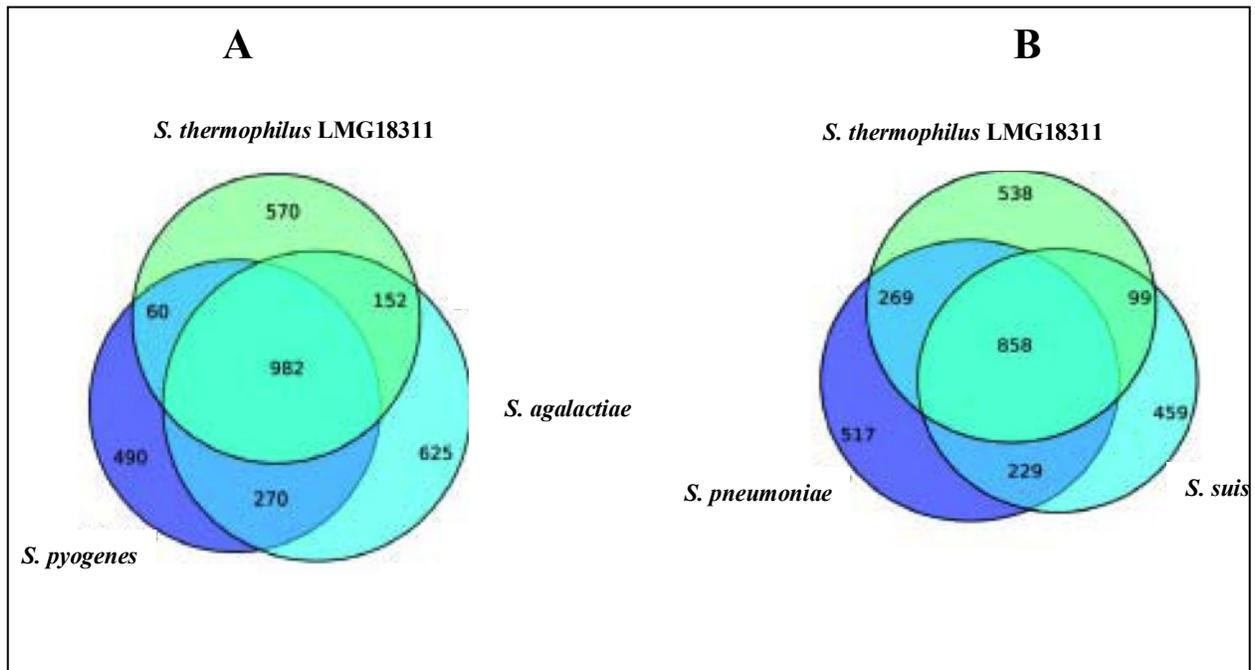
## **2. *Streptococcus thermophilus* : une espèce représentative de la diversité des bactéries lactiques**

### **2. 1. Positionnement de *Streptococcus thermophilus* au sein des streptocoques**

Le genre *Streptococcus* appartient au phylum des *Firmicutes* et regroupe des bactéries à Gram positif, dépourvues d'activité catalase, homofermentaires et à bas pourcentage en bases GC (Farrow et Collins, 1984). Les streptocoques sont divisés en 7 groupes : le groupe Mitis, Anginosus, Salivarius, Bovis, Mutans, Pyogène et un groupe renfermant des espèces non classées (*S. acidiminimus*, *S. suis* et *S. pleomorphus*). Le sixième groupe forme la division des streptocoques pyogènes tandis que 5 premiers groupes composent la division des streptocoques viridans. A l'exception de l'espèce *S. thermophilus*, la division viridans est constituée d'espèces bactériennes commensales voire pathogènes de la cavité orale, du tractus gastro-intestinal et des tractus génitaux des mammifères de l'homme, du singe, du rat, du porc ... Au sein de cette division, *S. thermophilus* est retrouvé dans le groupe salivarius qui comprend trois espèces : *S. salivarius*, *S. vestibularis* et *S. thermophilus* (Tableau 1) (Kawamura *et al.*, 1995). *S. thermophilus* est la seule de ces 3 espèces à être utilisée massivement en industrie agro-alimentaire et à être considérée comme GRAS (Generally Recognized As Safe). Les deux autres, *S. salivarius* et *S. vestibularis*, sont des bactéries commensales du tube digestif et/ou des pathogènes opportunistes. En effet, elles ont été récemment impliquées dans des cas d'infection : endocardite, méningite chez l'homme (Doyuk *et al.*, 2002 ; Idigoras *et al.*, 2001 ; Partridge, 2000). La comparaison des gènes codant les ARNr 16S et l'étude du polymorphisme du gène *sodA* (codant la superoxyde dismutase manganèse dépendante) chez ces trois espèces montrent qu'elles sont étroitement apparentées (Kawamura *et al.*, 1995 ; Poyart *et al.*, 1998). Outre la confirmation du lien de parenté entre *S. thermophilus* et des streptocoques pathogènes, la comparaison du génome de *S. thermophilus* LMG18311 avec ceux de quatre streptocoques pathogènes, *S. pneumoniae*, *S. suis*, *S. pyogenes* et *S. agalactiae*, a montré que plus de 54% des gènes de *S. thermophilus* sont orthologues de gènes d'autres streptocoques (Figure 2) (Hols *et al.*, 2005).

**Tableau 1 : Classement de 34 espèces de streptocoques établi sur la base des séquences des gènes codant les ARN 16S (Kawamura *et al.*, 1995)**

<b>Division des Streptocoques viridans</b>	<b>Groupe Mitis</b>
	<i>Streptococcus mitis</i>
	<i>Streptococcus gordonii</i>
	<i>Streptococcus pneumoniae</i>
	<i>Streptococcus oralis</i>
	<i>Streptococcus sanguis</i>
	<i>Streptococcus parasanguis</i>
	<b>Groupe Anginosus</b>
	<i>Streptococcus anginosus</i>
	<i>Streptococcus constellatus</i>
	<i>Streptococcus intermedius</i>
	<b>Groupe Salivarius</b>
	<i>Streptococcus salivarius</i>
	<i>Streptococcus thermophilus</i>
	<i>Streptococcus vestibularis</i>
	<b>Groupe Bovis</b>
	<i>Streptococcus bovis</i>
	<i>Streptococcus equinus</i>
	<i>Streptococcus alactolyticus</i>
	<b>Groupe Mutant</b>
<i>Streptococcus rnutans</i>	
<i>Streptococcus rattus</i>	
<i>Streptococcus cricetus</i>	
<i>Streptococcus downei</i>	
<i>Streptococcus sobrinus</i>	
<i>Streptococcus acaciae</i>	
<b>Division des Streptocoques pyogènes</b>	<b>Groupe Pyogène</b>
	<i>Streptococcus pyogenes</i>
	<i>Streptococcus agalactiae</i>
	<i>Streptococcus canis</i>
	<i>Streptococcus dysgalactiae</i>
	<i>Streptococcus equi</i>
	<i>Streptococcus iniae</i>
	<i>Streptococcus porcinus</i>
	<i>Streptococcus uberis</i>
	<i>Streptococcus parauberis</i>
	<i>Streptococcus hyointestinalis</i>
<b>Espèces non classé</b>	<i>Streptococcus acidominimus</i>
	<i>Streptococcus suis</i>
	<i>Streptococcus pleomorphus</i>



**Figure 2 : Résultat de la comparaison des génomes de *Streptococcus thermophilus* et d'autres espèces de streptocoques pathogènes (Lefébure et Stanhope, 2007)**

A. Répartition des gènes pour les trois génomes de *Streptococcus thermophilus* LMG1811, *Streptococcus agalactiae* et *Streptococcus pyogenes*

B. Répartition des gènes pour les trois génomes de *Streptococcus thermophilus* LMG1811, *Streptococcus pneumoniae* R6 et *Streptococcus suis*

## **2. 2. Positionnement de *Streptococcus thermophilus* au sein des bactéries lactiques**

*S. thermophilus* est utilisée en industrie laitière en tant que ferment dans la fabrication de produits laitiers tels que le yaourt et les fromages à pâte cuite (Emmental, Gruyère), à pâte pressée (Cheddar) ou à pâte filée (Mozzarella). Après *Lc. lactis* c'est la BL la plus utilisée dans ce domaine (Monnet, 2006).

Le rôle de *S. thermophilus* dans la fermentation du lait consiste en la conversion rapide du lactose en acide lactique, entraînant une diminution rapide du pH. Cette acidification engendre la précipitation des caséines provoquant la coagulation du lait. Elle est utilisée seule ou en association avec d'autres espèces bactérie lactiques, mais pour les yaourts, elle est toujours employée en co-culture avec *Lactobacillus delbrueckii* ssp. *bulgaricus* (*Lb. bulgaricus*) (Pette et Lolkema, 1950).

*S. thermophilus* est une bactérie homofermentaire, anaérobie aérotoleante, dont la température optimale de croissance se situe entre 37°C et 42°C, elle est exigeante d'un point de vue nutritionnel (Accolas *et al.*, 1980). Elle appartient au groupe des bactéries à Gram positif à bas pourcentage en base GC (39%) (Farrow et Collins, 1984 ; Roissart et Luquet, 1994).

## **3. Analyse du génome : une clé pour comprendre l'adaptation de *Streptococcus thermophilus* à la niche écologique lait**

### **3. 1. Caractéristiques du génome de *Streptococcus thermophilus***

A l'heure actuelle, les séquences de cinq génomes de *S. thermophilus* sont disponibles dans les banques des données EMBL/GenBank. Il s'agit du génome des souches CNRZ1066, LMG18311 et LMD-9 isolées de yaourt et du génomes des souches ND03 et JIM8232, isolées de lait (Bolotin *et al.*, 2004 ; Delorme *et al.*, 2011 ; Makarova *et al.*, 2006 ; Sun *et al.*, 2011). Pour chacune de ces souches le chromosome est circulaire, d'une taille comprise entre 1,8 et 1,9 Mpb,

et est caractérisé par un faible pourcentage en base GC (39%) (Tableau 2). Seule la souche LMD-9 héberge deux plasmides à faible nombre de copies.

La comparaison des génomes des trois souches CNRZ1066, LMG18311 et LMD-9 a révélé que 85% de leurs génomes sont bien conservés chez les trois souches (Lefébure et Stanhope, 2007). Les principales différences concernent les gènes impliqués dans la biosynthèse des exopolysaccharides, la synthèse de bactériocines, les résistances aux phages et les traces de prophages (Bolotin *et al.*, 2004 ; Deveau *et al.*, 2008 ; Horvath *et al.*, 2008). Néanmoins, les génomes des souches CNRZ1066 et LMG18311 semblent être très similaires puisque plus de 90% de leurs CDS sont communes. De plus, 80% de leurs gènes sont des orthologues de gènes d'autres streptococques (Bolotin *et al.*, 2004 ; Hols *et al.*, 2005).

L'analyse par hybridation de 47 génomes de *S. thermophilus* a révélé que le « core genome » de l'espèce *S. thermophilus* est constitué de 1271 gènes, soit 58% de leurs génomes (Rasmussen *et al.*, 2008). Une analyse *in silico* basée sur des biais compositionnels ont proposé qu'un ensemble de 197 gènes ont été acquis par transfert horizontal dans les trois génomes de *S. thermophilus* CNRZ1066, LMG18311 et LMD-9 (Liu *et al.*, 2009b).

En 2010, Delorme et ses collaborateurs ont estimé par « Multi Locus Sequence Typing » (MLST) la diversité génétique de 27 souches de *S. thermophilus*, 26 issues de différentes sources laitières et une isolée du sang humain (patient atteint d'une bactériémie). Cette analyse a révélé que la divergence de séquence nucléotidique est de 0,19% en moyenne (Delorme *et al.*, 2010), alors que la divergence au sein des espèces *S. salivarius*, *S. vestibularis* (aussi estimée par MLST) est de 6,6% et 3,6% respectivement. Ainsi, ces résultats suggèrent que l'espèce *S. thermophilus* serait une espèce clonale émergée récemment (Delorme *et al.*, 2007).

**Tableau 2 : Caractéristiques des cinq génomes entièrement séquencés de *Streptococcus thermophilus***

Souche	Origine	Taille de génome (pb)	CDS	CG%	Référence
CNRZ1066	Yaourt	1796226	1915	39,1	(Bolotin <i>et al.</i> , 2004)
LMG18311	Yaourt	1796846	1890	39,1	(Bolotin <i>et al.</i> , 2004)
LMD-9	Yaourt	1864178	1718	39,1	(Makarova <i>et al.</i> , 2006)
ND03	Lait	1831957	1919	39,1	(Sun <i>et al.</i> , 2011)
JIM8232	Lait	1929905	2145	38,9	(Delorme <i>et al.</i> , 2011)

CDS : pour Coding Sequence ; CG% : pourcentage en cytosine et guanine.

### 3. 2. Transfert génétique horizontal

Le transfert horizontal d'information génétique joue un rôle majeur dans l'évolution des génomes bactériens. Chez *S. thermophilus*, l'acquisition de nouvelles fonctions par transfert horizontal a permis l'adaptation de cette bactérie à sa seule niche écologique connue, le lait (Bolotin *et al.*, 2004). Une analyse *in silico* suggère également que les séquences acquises par transfert horizontal inter-espèces représenteraient environ 14% de la séquence des génomes séquencés de *S. thermophilus* (Hols *et al.*, 2005). Il s'agit souvent d'un échange d'information entre des espèces bactériennes proches phylogénétiquement. Néanmoins, des événements de transfert horizontal entre *S. thermophilus* et des espèces relativement éloignées phylogénétiquement, mais de proximité écologique étroite ont été observés (Delcour *et al.*, 2000 ; Guédon *et al.*, 1995). Le transfert génétique horizontal peut être résumé en deux étapes :

- Transfert d'une molécule d'ADN d'une cellule à une autre

Différents mécanismes peuvent être impliqués dans le transfert de l'ADN. La transformation correspond au processus d'incorporation d'un ADN extracellulaire à l'état libre à l'intérieur d'une cellule bactérienne compétente (Lorenz et Wackernagel, 1994) ; la transduction est décrite comme le transfert d'ADN entre les bactéries *via* l'intermédiaire des phages (Johnson et Chiu, 2007) ; la conjugaison permet le passage unidirectionnel de l'ADN d'une cellule donatrice vers une cellule réceptrice au travers d'un pont reliant les deux cellules (Thomas et Nielsen, 2005). Enfin, les nanotubes, permettent l'échange de constituants cytoplasmiques (des ARN, des protéines, des fragments d'ADN et des plasmides) entre deux cellules adjacentes par l'intermédiaire de structures tubulaires, présentes de façon non spécifique à la surface des cellules, et permettant de les associer physiquement (Dubey et Ben-Yehuda, 2011).

- Maintien de l'ADN acquis dans la descendance de la cellule réceptrice

La grande majorité de l'ADN acquis par transfert horizontal ne se maintient pas dans la population et est par conséquent perdue au fil des générations. En effet, l'ADN acquis doit être transmis à la descendance de la cellule réceptrice. Afin d'être transmis aux générations suivantes, l'ADN acquis doit être répliqué par la cellule réceptrice. Certains éléments tels que les plasmides sont capables de se répliquer de manière autonome car ils codent leur propre réplication. Cependant, la plupart des séquences acquises exige leur l'intégration dans le chromosome de la

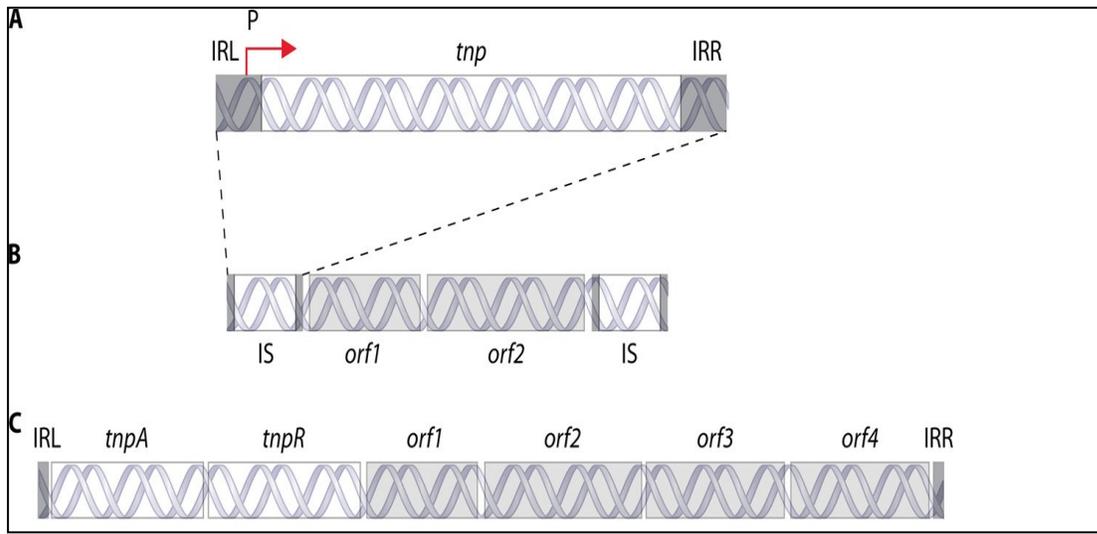
cellule hôte, permettant donc leur réplication en même temps que le chromosome.

Le transfert horizontal d'information génétique implique des éléments génétiques mobiles, comme les plasmides, les prophages, les éléments transposables et les éléments intégratifs et conjugatifs ICEs (pour Integratives and Conjugatives Elements). Ils se caractérisent par leur capacité spécifique à se disséminer et/ou à se maintenir dans les cellules réceptrice.

Les éléments transposables sont des entités génétiques qui assurent leur propre maintien par mouvement vers un ou plusieurs sites d'intégration (Bennett, 2004). Les principaux éléments transposables bactériens sont les séquences d'insertion et les transposons. Les séquences d'insertion (IS) sont des éléments mobiles capables de se transposer de façon autonome. Généralement, elles ne codent que des fonctions impliquées dans leur propre mobilité. Elles codent une transposase (Tnp) qui reconnaît les séquences inversées répétées (IR) terminales des IS et catalyse leur insertion avec une faible spécificité (Figure 3A) (Mahillon et Chandler, 1998 ; Mahillon *et al.*, 1999 ; Polard et Chandler, 1995). Les différentes IS ont été classées, selon la nature et la structure des IRs, la nature et la structure de la transposase, en 17 familles : IS1, IS3, IS4, IS5, IS6, IS21, IS30, IS66, IS91, IS110, IS200/605, IS256, IS630, IS982, IS1380, ISAs1, ISL3 (Mahillon et Chandler, 1998). Les transposons, en plus de fonction nécessaire à leur transposition, peuvent coder des fonctions capables de conférer un avantage sélectif à l'hôte. Bien que de nombreux types de ces éléments aient été identifiés, ceux-ci ont été traditionnellement répartis en deux grandes classes (Bennett, 2004) : les transposons composites qui sont constitués d'un ou plusieurs gènes encadrés par deux IS apparentées, les transposases codées par ces IS pourraient permettre la transposition de l'ensemble de cette structure (Roberts *et al.*, 2008) (Figure 3B) ; les transposons unitaires (ou non composites) qui se caractérisent par la présence d'une transposase (TnpA) et une recombinase (TnpR) (Figure 3C) (Kung *et al.*, 2010).

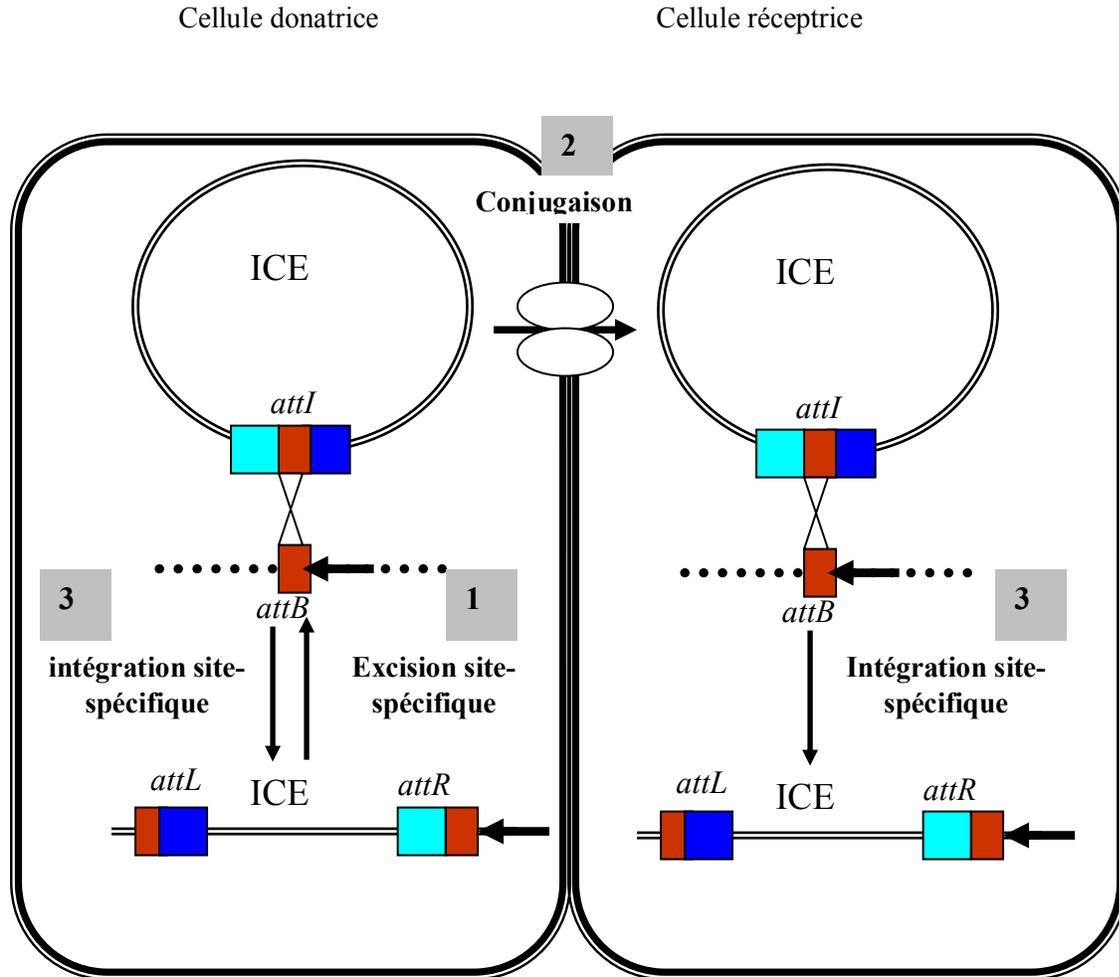
Les ICEs correspondent à des éléments capables de se transférer par conjugaison. Ils portent des gènes de recombinaison, qui leur permettent de s'intégrer dans un réplicon pour assurer leur stabilité. L'excision de la plupart des ICEs implique une recombinase site-spécifique à tyrosine qui catalyse la recombinaison entre les séquences d'attachement *attL* et *attR* bordant l'élément, et respectivement situées à gauche et à droite (Burrus *et al.*, 2002a, 2002b). Celle-ci génère une forme extrachromosomique circulaire de l'ICE portant un site d'attachement *attI* et un site d'attachement bactérien chromosomique *attB* libre (Figure 4). L'intégration se fait par recombinaison site-spécifique entre *attI* et *attB*. Bien que ces éléments s'excisent tous sous forme

circulaire et se répliquent de façon intercellulaire lors de la conjugaison, cette réplication ne semble pas être impliquée dans le maintien de l'élément dans la descendance de la cellule.



**Figure 3 : Représentation schématique de la structure des éléments transposables bactériens (Kung *et al.*, 2010).**

- A. Organisation d'une séquence d'insertion (IS) composée de répétitions inversées gauche (IRL), et droite (IRR), d'un gène codant une transposase (*tnp*). Le promoteur de la transposase peut être partiellement compris dans IRL.
- B. Organisation d'un transposon composite, composé de plusieurs ORFs (*orf1*, *orf2*) flanquées par deux IS très similaires
- C. Organisation d'un transposon unitaire, composé de plusieurs ORFs (*orf1*, *orf2*, *orf3*, *orf4*), d'un gène codant une transposase (*tnpA*), d'un gène codant la recombinase (*tnpR*). L'élément est flanqué de répétitions inversées IRL et IRR.



**Figure 4 : Représentation schématique du transfert conjugatif d'un ICE transféré sous forme simple brin et s'intégrant grâce à une recombinase site-spécifique à tyrosine (d'après Burrus *et al.*, 2002b)**

1. En présence de l'excisionase, l'intégrase catalyse la recombinaison site-spécifique entre les séquences identiques et de même orientation des sites *attL* et *attR* flanquant l'ICE. Ceci conduit à l'excision de l'ICE sous la forme d'une molécule circulaire porteuse d'un site *attI* et à la formation d'un site *attB* chromosomique.
  2. Le transfert de la forme circulaire de l'ICE se fait par conjugaison au travers d'un pore codé par l'élément (double ovale), et s'accompagne de sa réplication par cercle roulant.
  3. L'intégrase catalyse la recombinaison site-spécifique entre les séquences identiques du site *attI* de l'ICE et du site *attB* du chromosome conduisant à l'intégration de l'ICE et à la formation de sites *attL* et *attR*.
- Les rectangles bleus clairs et foncés représentent respectivement les sites *attR* et *attL* et leurs équivalents dans le site *attI*. Les rectangles rouges représentent les séquences identiques impliquées dans la recombinaison site-spécifique.

### 3. 2. 1. Présence d'éléments génétiques contribuant aux évènements de transfert horizontal dans le génome de *Streptococcus thermophilus*

L'analyse des génomes de souches de *S. thermophilus* révèle de nombreux événements de transfert horizontal qui ont eu lieu entre des différentes espèces des bactéries lactiques lors de co-cultures. Néanmoins, des échanges entre espèces phylogénétiquement proches mais de proximité écologique éloignées ont été observés. Plus de 50 IS ont été trouvées dans le génome des deux souches CNRZ1066 et LMG18311. Certaines d'entre-elles possèdent un pourcentage en GC différent du reste du génome et sont associées aux gènes liés à l'adaptation à la croissance en lait. Un exemple d'acquisition de gènes par transfert horizontal, est un îlot génomique de 17 kb, possédant trois séquences d'insertion de type IS1191. Grâce à cet îlot génomique, *S. thermophilus* a acquis les gènes de biosynthèse de la méthionine de *Lb. bulgaricus*. Cette voie de biosynthèse est indispensable à sa croissance en lait car la méthionine s'y trouve en très faible quantité (Bolotin *et al.*, 2004 ; Tamine et Deeth, 1980). Des séquences d'insertion (ISS1, IS981 et IS911) homologues aux séquences appartenant aux lactocoques ont été trouvées au sein du génome de *S. thermophilus* (Guédon *et al.*, 1995). Cela suggère la présence d'échanges d'informations génétiques par des évènements de transfert horizontal d'IS entre des différents genres de bactéries lactiques partageant la même niche écologique. Par contre, l'ensemble des souches testées de *Lc. lactis* possède l'IS981 à raison de 4 à 26 copies par souche et seules les souches de *S. thermophilus* isolées du fromage (et une seule souche utilisée dans la fabrication du yaourt) en possèdent mais avec un nombre de copies moins important (1 à 6 copies par souche). Les auteurs en ont donc conclu que leur présence chez *S. thermophilus* résulte de leur transfert de *Lc. lactis* vers *S. thermophilus*, ce qui est corrélé avec le fait que chez les souches de *S. thermophilus* issues du yaourt, la présence de cette IS est rare et les IS981 des deux espèces gardent des fortes identités entre-elles (Guédon *et al.*, 1995). Le transfert d'IS de *Lc. lactis* vers *S. thermophilus* a été également observé dans le cas de l'ISS1, trouvée chez les deux espèces (Bourgoin *et al.*, 1996). D'autres éléments mobiles ont été identifiés *S. thermophilus*, il s'agit des ICEs spécifiques de cette espèce. Les îlots génomiques ICESt1, ICESt3 ont été identifiés, ils sont capables de s'exciser et de s'intégrer par recombinaison site-spécifique dans l'extrémité 3' du gène *fda* codant la fructose-1,6 diphosphate aldolase (Burrus *et al.*, 2000 ; Pavlovic *et al.*, 2004).

Ces deux éléments sont capables de se transférer par conjugaison vers d'autres souches de *S. thermophilus*, d'autres bactéries lactiques telle que *Lc. lactis*, mais aussi vers des bactéries opportunistes telles que *S. pyogenes* et *En. faecalis* (Bellanger *et al.*, 2009). Par ailleurs, aucun plasmide conjugatif n'a été identifié chez *S. thermophilus*. Tous les plasmides identifiés chez cette espèce sont de taille inférieure à 8 kb, alors que la taille d'un module conjugatif fonctionnel est de 11 kb (Clewel D.B. *et al.*, 1995). Toutefois, des événements de transduction de plasmides de *S. thermophilus* vers *Lc. lactis* ont été détectés. Ce mécanisme de transfert pourrait constituer une alternative à la conjugaison pour le transfert des plasmides (Ammann *et al.*, 2008). Néanmoins, les mécanismes de transfert horizontal impliqués dans ces échanges d'information génétiques reste méconnus chez *S. thermophilus*. La transformation naturelle a été mise en évidence récemment chez certaines souches de *S. thermophilus* mais ce phénomène a été détecté en milieu chimiquement défini (Fontaine *et al.*, 2010 ; Gardan *et al.*, 2009). De ce fait, l'impact de ce mécanisme sur l'évolution du génome de *S. thermophilus* dans sa niche écologique reste encore méconnu.

### **3. 3. Perte de gènes de virulence et acquisition de nouvelles fonctions permettant l'adaptation au lait**

#### **3. 3. 1. Evolution du génome de l'espèce *Streptococcus thermophilus* par perte de fonctions**

La comparaison des génomes séquencés de *S. thermophilus* a révélé la présence de nombreux pseudogènes qui représentent environ 10% de la totalité des séquences codantes. Il s'agit de gènes inactivés par différentes mutations telles que des mutations non-sens, des décalages de cadre de lecture, des recodages ou encore des délétions occasionnées lors de réarrangements génomiques (Bolotin *et al.*, 2004). La proportion relativement élevée de pseudogènes chez *S. thermophilus* (celle des streptocoques pathogènes est estimée à 4,2%, 0,7% et 1,6% chez *S. pneumoniae*, *S. pyogenes* et *S. agalactiae*, respectivement) suggère que cette espèce a évolué principalement par perte de fonctions (Bolotin *et al.*, 2004 ; Hols *et al.*, 2005).

### 3. 3. 1. 1. Perte de voies métaboliques

Les génomes des souches LMG18311 et CNRZ1066 présentent des pseudogènes appartenant à différentes voies métaboliques. La proportion de pseudogènes n'est pas la même dans les différentes fonctions (Tableau 3). Trente pourcents des gènes liés au métabolisme énergétique sont tronqués alors que peu de pseudogènes sont retrouvés dans le métabolisme des acides aminés et des acides gras et qu'aucun des gènes liés au métabolisme des ARN ne le sont (Tableau 3).

Un tiers des gènes dédiés au transport et au métabolisme des glucides est des pseudogènes dont les gènes codant les transporteurs de type PTS (carbohydraPhosphoTransférase System) (Bolotin *et al.*, 2004). Des analyses réalisées chez 8 souches de *S. thermophilus* et 4 souches de *S. salivarius* révèlent que les 4 gènes *ptsG*, *fruA*, *bglP* et *treP*, qui coderaient respectivement des transporteurs de glucose, fructose,  $\beta$ -glucoside et de tréhalose, sont des pseudogènes chez *S. thermophilus*. Seule une souche de *S. thermophilus* présente une séquence fonctionnelle du gène *fruA*. A l'inverse, la séquence des gènes *ptsG*, *fruA* et *bglP* semble intacte chez les 4 souches de *S. salivarius* (pour cette espèce, le gène *treP* n'a pas été recherché). De plus, la majorité des souches de *S. thermophilus* est incapable de métaboliser le galactose, même si une partie de ces souches possède l'ensemble des gènes (*galKTE-galM*) nécessaires au métabolisme de ce glucide (Van den Bogaard *et al.*, 2004). Cette déficience proviendrait plutôt d'une faible activité du promoteur *galKTE* (Vaillancourt *et al.*, 2004 ; Vaughan *et al.*, 2001). Deux autres gènes, *butA* et *adhE*, impliqués dans le métabolisme carboné et codant l'acétoïne réductase et l'alcool acétaldéhyde déshydrogénase, sont inactivés chez *S. thermophilus* et ne semblent pas l'être chez *S. salivarius*. Inversement, un transporteur spécifique du lactose (le principal glucide du lait) est présent dans le génome de *S. thermophilus* et absent chez les autres streptocoques qui, par ailleurs, présentent une large capacité à utiliser différentes sources carbonées et qui sont ainsi capables de s'adapter à différentes niches écologiques (Bolotin *et al.*, 2004 ; Tettelin, 2001). Cette spécialisation a permis à *S. thermophilus* de s'adapter à la seule niche écologique connue pour cette bactérie : le lait.

**Tableau 3 : Distribution des pseudogènes dans les différentes fonctions cellulaires chez *Streptococcus thermophilus* CNRZ1066 et LMG18311 (Bolotin *et al.*, 2004)**

Catégories fonctionnelles	CNRZ1066			LMG18311		
	Total CDS	CDS tronquées <sup>1</sup>		Total CDS	CDS tronquées <sup>1</sup>	
		Nb <sup>2</sup>	%		Nb <sup>2</sup>	%
Biosynthèse des acides aminés	75	4	5,3	74	4	5,4
Biosynthèse des cofacteurs, groupes prosthétiques et transporteurs	45	8	17,8	45	6	13,3
Enveloppe cellulaire	93	18	19,4	93	16	17,2
Processus cellulaires	82	7	8,5	86	7	8,1
Métabolisme central intermédiaire	25	3	12,0	20	1	5,0
Métabolisme énergétique	115	34	29,6	119	39	32,8
Métabolisme des acides gras et phospholipides	33	2	6,1	34	2	5,9
Purines, pyrimidines, nucléotides et nucléosides	72	12	16,7	70	12	17,1
Fonctions de régulations	101	23	22,8	100	23	23,0
Réplication, Métabolisme d'ADN	93	15	16,1	93	13	14,0
Transcription, Métabolisme d'ARN	37	0	0,0	37	0	0,0
Traduction, Métabolisme des protéines	146	5	3,4	144	5	3,5
Protéines de transport et de liaisons	253	76	30,0	246	77	31,3
Autres fonctions (phage, tranposase..)	147	90	61,2	138	82	59,4
Potentielles fonctions	413	55	13,3	425	71	16,7
Inconnues <sup>3</sup>	185			166		
<b>Total</b>	<b>1,915</b>	<b>352</b>	<b>18,4</b>	<b>1,890</b>	<b>358</b>	<b>18,9</b>

<sup>1</sup>CDS : Séquences codantes

<sup>2</sup>Correspond au nombre de CDSs tronquées et retrouvées dans les génomes séquencés de *Streptococcus thermophilus*

<sup>3</sup>Fonctions inconnues exclues de cette analyse

### 3. 3. 1. 2. Perte de gènes de virulence

De nombreux gènes de virulence, Virulence Related Genes (VRGs), connus chez les streptocoques pathogènes sont absents du génome de *S. thermophilus* ou ne sont présents que sous forme de pseudogènes (Bolotin *et al.*, 2004). Au total, sur 50 VRGs de *S. pyogenes* SF370 et 118 VRGs de *S. pneumoniae* TIGR4, seuls 8 (dont 1 pseudogène) et 58 (dont 7 pseudogènes) sont respectivement retrouvés chez *S. thermophilus* (Tableau 4). Globalement, les déterminants principaux de la virulence (adhésine, protéines pneumococcales A et C de surface...) qui se trouvent à la surface des cellules bactériennes de streptocoques pathogènes sont soit absents, soit inactivés, chez *S. thermophilus*. Les déterminants de la résistance aux antibiotiques retrouvés chez les streptocoques pathogènes sont également absents du génome des souches de *S. thermophilus* (Hols *et al.*, 2005 ; Subbian *et al.*, 2004).

**Tableau 4 : Identification des VGRs putatifs dans le génome de *Streptococcus thermophilus* CNRZ 1066 par comparaison avec ceux présents dans les génomes de *Streptococcus pyogenes* M1 et *Streptococcus pneumoniae* TIGR4 (Bolotin *et al.*, 2004)**

	VRGs	VRGs présents chez <i>Streptococcus thermophilus</i>	VRGs inactivés <sup>1</sup> chez <i>Streptococcus thermophilus</i>	Pourcentage de perte des VRGs chez <i>Streptococcus thermophilus</i>
<b><i>Streptococcus pyogenes</i></b>				
-VGRs démontrés <sup>2</sup>	21	1	0	95,2
-VGRs putatifs <sup>3</sup>	29	7	1	75,9
<b>Total</b>	<b>50</b>	<b>8</b>	<b>1</b>	<b>84,0</b>
<b><i>Streptococcus pneumoniae</i></b>				
-VGRs démontrés <sup>2</sup>				
Adhérence	9	8	0	55,6
Défenses de l'hôte	23	20	2	78,3
Métabolisme	35	20	2	42,9
Transport/protéase	29	19	3	34,5
Autres	22	10	1	45,5
<b>Total</b>	<b>118</b>	<b>58</b>	<b>7</b>	<b>54,6</b>

1 Inactivation des gènes par un ou plusieurs décalages du cadre de lecture, par la présence d'un codon stop ou par la présence d'une extrémité 3' tronquée.

2 Démonstration expérimentale

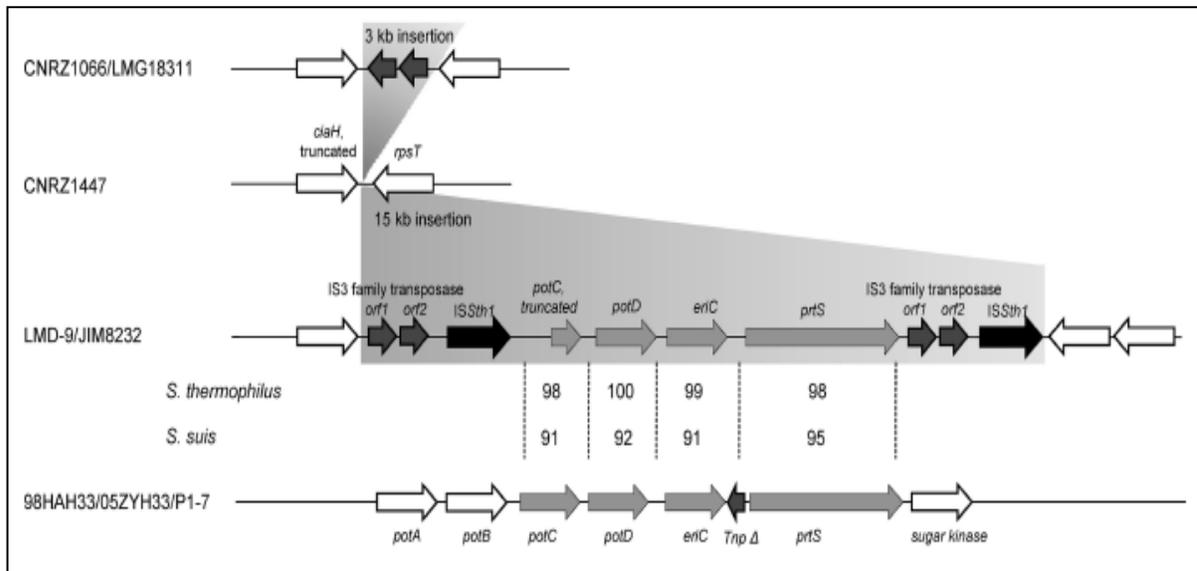
3 Gènes identifiés lors du séquençage des génomes

Cependant, pour certains auteurs (Bonifait *et al.*, 2010 ; Mora *et al.*, 2005a) certains facteurs conservés chez *S. thermophilus* peuvent être assimilés à des facteurs de virulence. Il s'agit notamment de l'uréase, du système de transport Ami, de la protéase PrtS, et de la capsule polysaccharidique. En effet, la présence d'activité uréasique est fréquemment liée à la pathogénicité de bactéries telles que *Helicobacter pylori*, *Clostridium perfringens*, *Klebsiella pneumoniae*, *Proteus mirabilis*, *Staphylococcus saprophyticus*, *Ureoplasma urealyticum* ou *Yersinia enterocolitica* (Burne et Chen, 2000 ; Mobley *et al.*, 1995). Par exemple, chez *H. pylori*, l'uréase est un déterminant crucial de la virulence : elle intervient à la fois dans le mécanisme de colonisation, la survie dans des conditions acides et dans la pathogénicité (Pflock *et al.*, 2006 ; Stingl *et al.*, 2002). Parmi 59 espèces du genre *Streptococcus*, l'opéron codant une activité uréasique fonctionnelle est retrouvé chez 5 espèces : *S. alactolyticus*, *S. hyointestinalis*, *S. salivarius*, *S. thermophilus* et *S. vestibularis*. Chez *S. thermophilus*, cet opéron présente de 97 à 100% d'identité avec celui de *S. salivarius* (Mora *et al.*, 2004). Chez *S. salivarius*, en plus de son rôle dans l'approvisionnement en source azotée, l'uréase permet à la bactérie de résister à des pH acides car la dégradation de l'urée en dioxyde de carbone et en ammoniacque engendre une remontée du pH extracellulaire qui permet à *S. salivarius* ainsi qu'à d'autres bactéries de la plaque dentaire de survivre en milieu acide (Chen *et al.*, 1996, 2000). Des gènes codant un système de transport d'oligopeptides (Ami) indispensable pour la croissance de *S. thermophilus* en lait ont été décrits chez *S. pneumoniae* comme étant impliqué dans la colonisation nasopharyngienne (Kerr *et al.*, 2004). De même, la protéase de paroi PrtS de *S. thermophilus*, qui a un rôle important pour la croissance de la bactérie en lait, présente 95,9% d'identité avec la protéase SspA (Bonifait *et al.*, 2010), protéase décrite comme un facteur de virulence chez *S. suis* (Hu *et al.*, 2010). Enfin, les gènes impliqués dans la synthèse des capsules polysaccharidiques sont importants pour la virulence chez les streptocoques pathogènes, une souche de *S. pneumoniae* dépourvue de capsule (*eps<sup>-</sup>/cps<sup>-</sup>*) n'est plus virulente. Chez *S. thermophilus*, ces gènes ne sont pas considérés comme déterminants pathogènes mais contribuent à la production de polysaccharides externes qui confèrent une texture recherchée pour le yaourt (Broadbent *et al.*, 2003 ; Tettelin, 2004).

### 3. 3. 2. Acquisition de nouvelles fonctions chez *Streptococcus thermophilus*

Plusieurs gènes qui, pourraient avoir été acquis par transfert horizontal, ont été identifiés dans le génome de *S. thermophilus*. Il s'agit de gènes codant des fonctions importantes, acquis d'une espèce commensale et/ou pathogène. Certains gènes interviennent dans le catabolisme des protéines et des glucides du lait, dans la synthèse de bactériocines, dans la synthèse d'exopolysaccharides et dans la tolérance à l'oxygène (Bourgoin *et al.*, 1999 ; Broadbent *et al.*, 2003 ; Fernandez-Espla *et al.*, 2000 ; Hols *et al.*, 2005 ; Mathot *et al.*, 2003 ; Rallu *et al.*, 2002 ; Siezen *et al.*, 1996 ; Vaughan *et al.*, 2001).

Un de ces exemples est la protéase PrtS de *S. thermophilus* qui a été acquise d'une espèce commensale et/ou pathogène proche à l'espèce *S. suis* (Delorme *et al.*, 2010). Le gène *prtS* fait partie d'un îlot génomique de 15 kb localisé entre le pseudogène *ciaH* et le gène *rpsT* (Figure 5).



**Figure 5 : Îlot génomique contenant le gène *prtS* avec la région chromosomique adjacente chez *Streptococcus thermophilus* et *Streptococcus suis* (Delorme *et al.*, 2010)**

Les pourcentages d'identité des séquences codantes de l'îlot entre différentes souches de *S. thermophilus* et *S. suis* sont respectivement indiqués en première ligne et deuxième ligne.

Cet îlot contient quatre gènes *potC* (tronqué), *potD*, *eriC* et *prtS* et il est entouré par deux éléments mobiles de type IS, le premier étant une transposase appartenant à la famille IS3 et le deuxième, un *ISSth1* appartenant à la famille ISL3. Ces quatre gènes présentent une forte identité avec l'ADN génomique de *S. suis* et présentent la même synténie chez les deux espèces. Ceci suggère que cet îlot a été acquis par transfert horizontal d'un streptocoque proche de *S. suis* (Delorme *et al.*, 2010). Ce cas donne l'exemple d'un échange de gènes entre espèces phylogénétiquement proches mais de proximité écologique éloignée puisque ces bactéries ne partagent pas la même niche écologique. En effet, *S. suis* est un pathogène retrouvé chez le porc qui peut être transmis à l'homme lors de contact avec des porcs malades (Lun *et al.*, 2007). De plus, cet îlot confère une propriété métabolique (activité protéolytique) importante pour l'adaptation de l'espèce *S. thermophilus* du lait. En effet, l'espèce était décrite précédemment comme étant peu protéolytique. L'émergence de cette protéase permet d'obtenir des souches ayant un intérêt technologique, puisque la capacité à croître en lait (propriétés recherchées en industrie) est liée à la propriété protéolytique (Galia *et al.*, 2009 ; Shahbal *et al.*, 1991). Bien que cet îlot ne possède aucun élément de transfert, *S. thermophilus* l'aurait accueilli grâce à sa capacité à développer une compétence naturelle (Delorme *et al.*, 2010).

#### **4. Exigences nutritionnelles de *Streptococcus thermophilus***

Les bactéries lactiques sont des bactéries très exigeantes d'un point de vue nutritionnel. Elles requièrent pour leur croissance une source carbonée fermentescible, une source d'acides aminés essentiels qu'elles ne peuvent pas synthétiser à partir d'une source azotée simple, des sources phosphatées et soufrées, mais aussi des facteurs de croissance comme les vitamines, des bases puriques et pyrimidiques et des acides gras. *S. thermophilus* est généralement moins exigeante que d'autres bactéries lactiques telles que *Lc. lactis* ou *Lactobacillus* (Cocaign - Bousquet *et al.*, 1995 ; Ledesma *et al.*, 1977 ; Morishita *et al.*, 1974, 1981).

Le lait comprend la majorité des facteurs de croissance nécessaires aux BL mais parfois en quantités limitantes pour leur croissance (Desmazeaud, 1990). Ainsi, le lait de vache est constitué de 90% d'eau, de 48 à 50 g.L<sup>-1</sup> de lactose, d'environ 35 g.L<sup>-1</sup> de matières grasses et de 34 g.L<sup>-1</sup> en moyenne de protéines, de minéraux et de vitamines (Alais, 1965). Le lactose est la source de carbone et d'énergie principale de ce milieu pour les bactéries. Les protéines représentent 95%

des matières azotées totales qui se décompose en 20% de protéines lactosériques ( $\beta$ -lactoglobuline,  $\alpha$ -lactalbumine, immunoglobulines, ...) et 80% de caséines (caséines  $\alpha_{s1}$ ,  $\alpha_{s2}$ ,  $\beta$  et  $\kappa$ ), les 5% restant étant constitués d'azote non protéique tels que des acides aminés libres, des petits peptides, de l'urée, des bases puriques et pyrimidiques et de la créatinine (Desmazeaud, 1990). Les caséines sont complexées avec le phosphate de calcium pour former des micelles stables dans le lait à l'état liquide mais qui sont facilement dégradées par des enzymes protéolytiques (Aimutis, 2004). Parmi les différents sources azotées du lait, les bactéries lactiques peuvent utiliser les acides aminés libres, les peptides et les caséines (après dégradation par leur(s) protéase(s) extracellulaires) et l'urée pour les souches possédant l'activité uréasique.

#### **4. 1. Exigence en vitamines**

*S. thermophilus* montre une exigence absolue en acide pantothénique (vitamine B<sub>5</sub>) et en riboflavine (vitamine B<sub>2</sub>) et à un moindre degré en thiamine (vitamine B<sub>1</sub>), en acide nicotinique (vitamine B<sub>3</sub>) et en biotine (vitamine B<sub>8</sub>). La pyridoxine (ou ses dérivés) (vitamine B<sub>6</sub>) stimule fortement leur croissance (Desmazeaud, 1983). Ces vitamines nécessaires à la croissance de *S. thermophilus* ne semblent pas être en quantité limitante dans le lait.

#### **4. 2. Exigence en source azotée**

Les BL sont des bactéries auxotrophes pour plusieurs acides aminés et leurs besoins varient suivant les souches et les espèces. Elles peuvent utiliser les acides aminés libres, ou les peptides présents dans le lait. Cependant les concentrations de ces derniers dans le lait sont nettement inférieures à celles requises pour atteindre une bonne densité cellulaire (Desmazeaud, 1990 ; Juillard et al., 1996 ; Letort et Juillard, 2001). Des études ont montré que les acides aminés libres et les peptides présents dans le lait ne sont respectivement responsables que de 2% et 8% de la croissance des lactococques (Juillard *et al.*, 1995).

*S. thermophilus* est capable de synthétiser plusieurs acides aminés et est donc de ce fait moins exigeante en acides aminés que les autres bactéries lactiques. Les exigences varient en fonction des souches (Desmazeaud, 1983 ; Garault *et al.*, 2000 ; Letort et Juillard, 2001). De nombreuses souches de *S. thermophilus* sont capables de croître avec un mélange de 3 à 6 acides

aminés (cystéine, méthionine, glutamine, leucine, isoleucine et valine) comme seule source d'acides aminés (Letort et Juillard, 2001).

Une croissance optimale de cette bactérie en milieu lait exige donc, soit une capacité de biosynthèse *de novo* d'acides aminés soit une capacité d'hydrolyse des caséines du lait en peptides. Ces peptides générés sont internalisés à l'intérieur de la cellule et ensuite dégradés en acides aminés (Fernandez-Espla *et al.*, 2000 ; Hols *et al.*, 2005 ; Savijoki *et al.*, 2006).

#### 4. 2. 1. Système protéolytique de *Streptococcus. thermophilus*

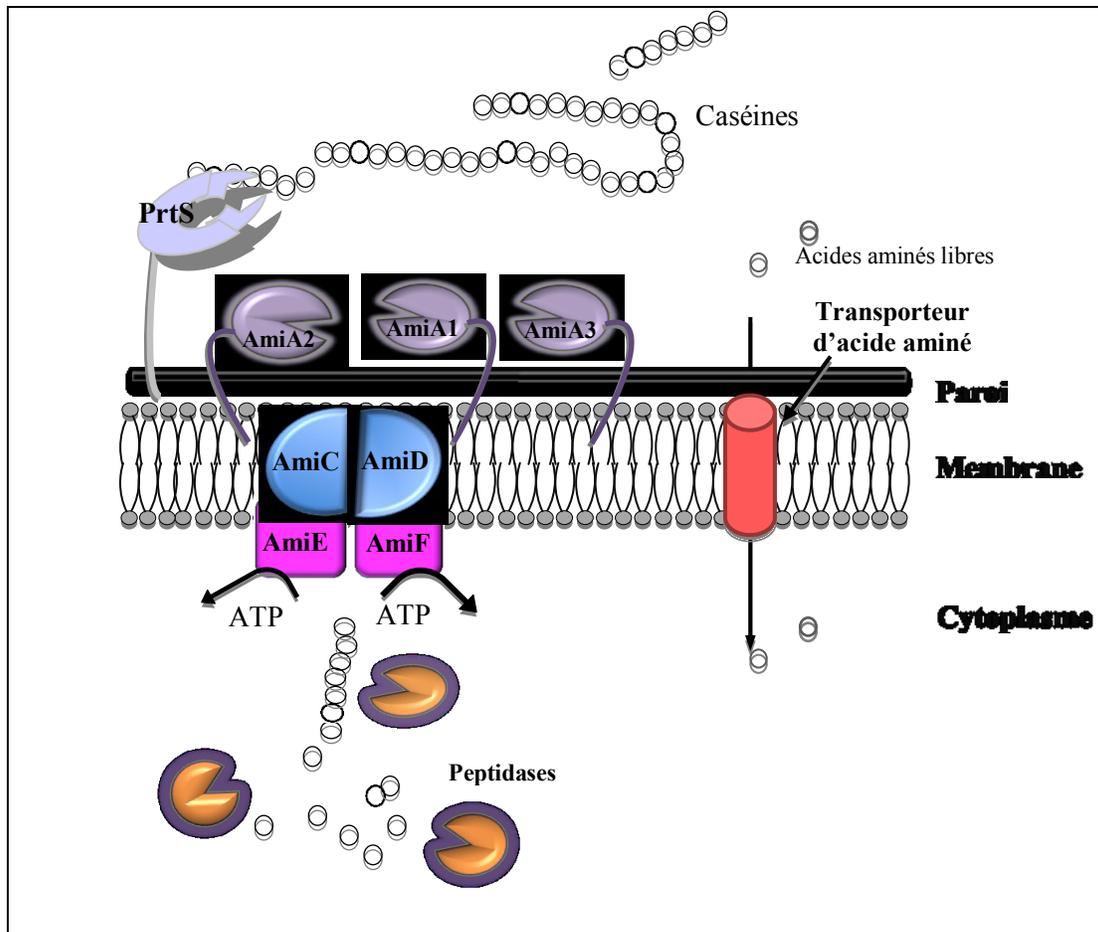
*S. thermophilus* possède un système protéolytique similaire à celui retrouvé chez les autres bactéries lactiques. Il est composé i) d'une protéase de paroi, appelée PrtS, capable d'hydrolyser les caséines du lait ii) de systèmes de transport de peptides et d'acides aminés iii) de peptidases intracellulaires indispensables à la dégradation des peptides internalisés en acides aminés (Hols *et al.*, 2005) (Figure 6).

##### 4. 2. 1. 1. Protéases de paroi

Cinq types de protéases de paroi capables d'hydrolyser les protéines du lait ont été identifiées chez les BL : PrtP chez *Lc. lactis* et *Lactobacillus paracasei*, PrtH chez *Lactobacillus helveticus*, PrtR chez *Lactobacillus rhamnosus*, PrtB chez *Lb. bulgaricus* et PrtS chez *S. thermophilus* (Figure 7) (Savijoki *et al.*, 2006 ; Siezen, 1999). Ces protéases sont synthétisées sous forme de préprotéine d'environ 2000 résidus et sont constituées de plusieurs domaines fonctionnels retrouvés chez les protéases de la famille de la subtilisine (Fernandez-Espla *et al.*, 2000 ; Siezen et Leunissen, 1997 ; Siezen, 1999).

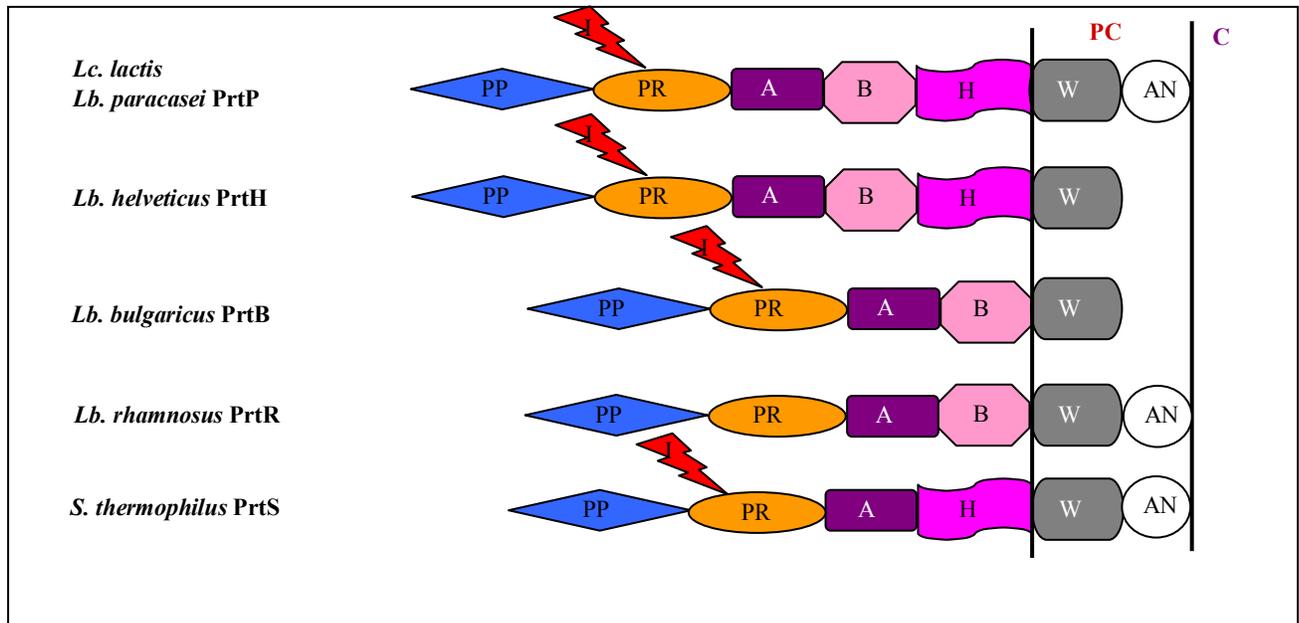
Chez *Lc. lactis*, une autre protéase membranaire, la protéase de ménage HtrA, a été mise en évidence. Cette enzyme hydrolyse des protéines anormales et/ou mal repliées, intervient dans la maturation de certaines protéines exportées par la cellule et est capable de couper le pro-peptide d'une pro-protéine hétérologue sécrétée. Il s'agit d'une protéase à sérine, dont le site catalytique potentiel est exposé à la surface et donc potentiellement capable de dégrader des protéines exportées (Poquet *et al.*, 2000). L'analyse du génome de *S. thermophilus* montre la présence du gène *htrA* codant une protéine homologue à celle de *Lc. lactis* (avec 48% d'identité

protéique) qui pourrait jouer le même rôle (Hols *et al.*, 2005). Cependant, cette protéase ne semble pas impliquée dans la caséinolyse observée lors de la croissance de *S. thermophilus* dans le lait car les souches CNRZ1066 et LMG18311 qui semblent ne posséder que HtrA comme protéase de surface se développent difficilement dans ce milieu (Galia *et al.*, communication personnelle).



**Figure 6 : Représentation schématique du système protéolytique chez *Streptococcus thermophilus* (d'après Hols *et al.*, 2005)**

- PrtS correspond à la protéase de paroi. Elle n'est pas présente chez toutes les souches
- AmiA1, A2 et A3 sont des protéines affines aux oligopeptides ; AmiC et D sont des perméases et Ami E et F sont des ATPases. L'ensemble des protéines Ami constitue un système de transport d'oligopeptides.



**Figure 7 : Représentation schématique des différents domaines composant les protéases de paroi, (Fernandez-Espla *et al.*, 2000 ; Savijoki *et al.*, 2006 ; Siezen, 1999)**

PC : paroi cellulaire ; C : cytoplasme ; PP : pré-prodomaine ; PR : domaine catalytique ; I : domaine d'insertion ; A : domaine globulaire ; B : domaine B ; H : domaine hélicoïdal ; W : domaine espaceur ; AN : domaine d'ancrage. Les domaines PP et AN sont clivés respectivement lors de l'excrétion-maturation et de l'ancrage de la protéase.

#### ❖ Caractéristiques générales de PrtS

La protéase PrtS comporte plusieurs domaines (Fernandez-Espla *et al.*, 2000 ; Galia *et al.*, 2009) :

- Le premier domaine (pré-prodomaine) est composé d'une séquence signal de 35 résidus et d'un propeptide formé de 109 résidus d'acides aminés. Lors de l'exportation et de la maturation de l'enzyme, ce domaine sera clivé.

- Le domaine PR, composé de 495 résidus d'acides aminés, constitue la partie N-terminale de la protéine mature et correspond au domaine catalytique. Le domaine catalytique renferme la triade catalytique d'une protéase à sérine comprenant un résidu d'acide aspartique, un résidu de sérine et un résidu d'histidine. Ce domaine renferme un sous-domaine appelé I pour insertion (~150 résidus) qui pourrait moduler la spécificité du substrat.

- Le domaine A est globulaire, il est composé de 438 résidus et sa fonction n'est pas clairement identifiée. Il est essentiel à l'activité protéolytique et serait impliqué dans la régulation de l'activité protéolytique et/ou la spécificité du substrat du domaine PR (Siezen, 1999).

La partie C-terminale de la protéine comporte, quant à elle, trois domaines :

- un domaine hélicoïdal (H) de 367 résidus d'acides aminés,
- un domaine hydrophile de 106 résidus acides aminés (W) qui serait un espaceur au niveau de la paroi cellulaire,
- et enfin, le domaine d'ancrage au peptidoglycane (AN), contenant 35 résidus d'acides aminés, renferme le motif LPXTG reconnu par la sortase A.

La protéase membranaire PrtS de *S. thermophilus* ne contient pas de domaine B. Ce domaine retrouvé au sein des protéases PrtP, PrtR, PrtH, PrtB pourrait être impliqué dans la stabilité de l'enzyme (Fernandez-Espla *et al.*, 2000).

La protéase PrtS de la souche CNRZ385 a été caractérisée. Elle possède une masse moléculaire de 153 kDa, est active sur une gamme de pH comprise entre 5,5 et 8,5 avec un pH optimum d'activité de 7,5. La température optimale d'activité déterminée avec le substrat synthétique Ac-Ala-Ala-Pro-Phe-pNA est de 37°C. La protéase PrtS est inhibée à 100% en présence de 1 mM de PMSF (phénylméthylsulfonylfluorure), à 83% en présence de 1 mM d'acide iodoacétique à 75% en présence de 0,1 mM de chymostatine. L'addition de CaCl<sub>2</sub> active fortement la protéase, de 2 à 10 fois, après addition de 2 à 10 mM de CaCl<sub>2</sub> respectivement (Fernandez-Espla *et al.*, 2000). Cette enzyme hydrolyse la caséine  $\beta$  préférentiellement (Delorme *et al.*, 2010 ; Miclo *et al.*, 2012 ; Shahbal *et al.*, 1993 ; Chang *et al.*, en préparation)

Trois allèles du gène *prtS* ont été retrouvés avec une divergence génétique maximale de 2,9% contre 1,3% chez les trois allèles trouvés chez *S. suis* (Delorme *et al.*, 2010). La principale différence au niveau des séquences codantes des différents allèles *prtS* de *S. thermophilus* trouvés au sein des banques de données, correspond à une duplication imparfaite de 32 résidus d'acides aminés, retrouvée dans la région du propeptide. Cette duplication imparfaite n'est pas présente chez les allèles *sspA* de *S. suis* et pourrait être importante pour la maturation de la protéase. Elle pourrait ainsi correspondre à une adaptation résultant d'une mutation apparue suite à l'acquisition du gène *prtS* par *S. thermophilus* (Delorme *et al.*, 2010). Même si la protéase PrtS chez *S. thermophilus* est similaire à la protéase de streptocoques pathogènes, elle n'a pas conservé la même fonction. En effet, elle est liée à la croissance dans le lait (Galia *et al.*, 2009 ; Shahbal *et al.*, 1991).

#### ❖ Répartition de PrtS au sein de *Streptococcus thermophilus*

*S. thermophilus* a longtemps été considérée comme étant une BL peu protéolytique. Cela provient des résultats obtenus lors de la caractérisation du système protéolytique d'une collection historique de souches d'origine différente. En effet, il avait été observé que seules 3 des 97 souches analysées présentaient une activité protéolytique de surface comparable à celle de *Lc. lactis* (Shahbal *et al.*, 1991). Par la suite, il a été observé que lorsqu'une souche de *S. thermophilus* possédant la protéase PrtS est co-cultivée avec *Lb. bulgaricus* en milieu lait, la croissance optimale de *S. thermophilus* dans ce milieu semble être plutôt dépendante des peptides libérés par l'activité de la protéase de *Lb. bulgaricus* (Courtin *et al.*, 2002).

Cependant, des études récentes ont rapporté l'émergence du gène *prtS* au sein de collections de souches de *S. thermophilus* (Delorme *et al.*, 2010 ; Galia *et al.*, 2009 ; Rasmussen *et al.*, 2008). Ainsi, Rasmussen *et al.* (2008) ont observé la présence du gène *prtS* dans 35 des 47 souches étudiées. De même, Galia *et al.* (2009), ont étudié le potentiel protéolytique de surface de 30 souches de *S. thermophilus* isolées de yaourts ou de fromages et montré que la moitié d'entre-elles possède une activité protéolytique dont le niveau varie entre les souches. Par contre, les auteurs ont montré que si seules 15 souches de leur collection présentent une activité protéolytique détectable, 23 possèdent le gène *prtS*. Le séquençage du gène *prtS* chez 4 des 8 souches de phénotype PrtS<sup>-</sup> et chez 1 souche PrtS<sup>+</sup> a révélé une parfaite identité de séquence entre les allèles, l'allèle retrouvé étant celui de la souche LMD-9. Les auteurs ont conclu que l'absence d'activité PrtS chez certaines souches n'est pas forcément due à une absence du gène mais peut résulter aussi d'une altération de sa régulation. De plus, une étude phylogénétique, basée sur la technique de « MLST », réalisée chez 27 souches de *S. thermophilus* a montré la présence du gène *prtS* chez 8 souches appartenant à des clusters phylogénétiques différents (Delorme *et al.*, 2010).

#### **4. 2. 1. 2. Les systèmes de transports des peptides**

##### 4. 2. 1. 2. 1. Les systèmes de transport de peptides bactériens

Chez les bactéries, la membrane cytoplasmique, de structure hydrophobe est imperméable à de nombreux composés. Les bactéries, possèdent donc des systèmes de transport qui font le lien

entre le milieu extracellulaire et le milieu intracellulaire. Deux grands groupes de transporteurs ont été décrits. Ils sont classés suivant la source d'énergie qu'ils utilisent :

- Le transport dépendant de la force protomotrice qui correspond au cas où le transport d'un soluté est lié au mouvement de protons à travers la membrane. Si le transport se fait dans le même sens que le mouvement de protons à travers la membrane, on parle de symport. S'il se fait dans le sens inverse, on parle alors d'antiport (Kakinuma, 1998). Une famille de transporteurs de peptides dépendant de la force protomotrice, la famille PTR (PTR pour **P**eptide **T**ransport) se distingue. Cette famille se trouve chez une grande variété d'êtres vivants, elle comprend plusieurs membres capables de prendre en charge une large gamme de substrats contenant de l'azote, y compris les di- et tri-peptides, les acides aminés et les ions nitrates (Chiang *et al.*, 2004 ; Frommer *et al.*, 1994 ; Stacey *et al.*, 2002 ; Steiner *et al.*, 1994 ; Tsay *et al.*, 1993). Ce sont des transporteurs formés d'une seule protéine constituée de 12 segments transmembranaires contenant au sein du 4<sup>ème</sup> ou du 5<sup>ème</sup> segment le motif signature FING (FYxxINxGSL) spécifique de transporteur des peptides, dont le rôle est mal connu (Hauser *et al.*, 2001 ; Steiner *et al.*, 1995).
- Les transporteurs ABC (ATP-Binding Cassette) connus sous le nom de «trafic ATPase», qui font intervenir directement l'énergie dégagée par l'hydrolyse d'ATP.

#### A. Les transporteurs de type ABC

Les protéines ABC sont des transporteurs actifs primaires : l'énergie est couplée au transport unidirectionnel de substrats à travers la membrane. Elle constitue vraisemblablement la famille de protéines la plus vaste, rassemble des systèmes d'import et d'export de solutés. Les systèmes d'import de type ABC n'ont été identifiés que chez les procaryotes tandis que les systèmes d'export ont été isolés chez les procaryotes, les eucaryotes et les archéobactéries (Higgins, 1992 ; Saurin *et al.*, 1999). Les transporteurs ABC sont impliqués dans le transport d'un grand nombre de substances biologiques : ions, glucides, acides aminés, peptides, phospholipides, stéroïdes, vitamines, antibiotiques, mais aussi de médicaments, lipopeptides et protéines de grande taille (>100000 Da) (Higgins, 1992). Ils se révèlent d'une importance toute particulière dans leur diversité de fonction. Ils sont impliqués dans des processus cellulaires divers comme la maintenance de l'homéostasie osmotique, l'importation des nutriments, la résistance aux xénobiotiques, l'immunité bactérienne, la division cellulaire, la régulation de

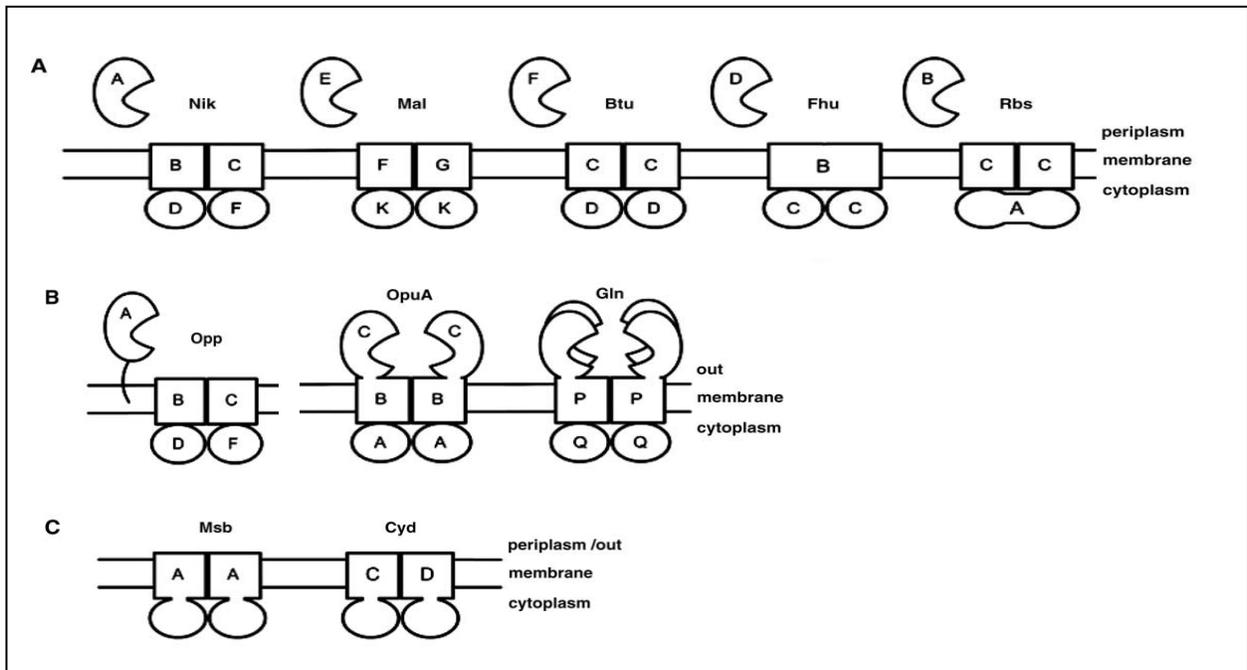
l'expression de certains gènes (la sporulation, la compétence et les facteurs de virulence), la présentation d'antigènes aux cytotoxiques (Biemans-Oldehinkel *et al.*, 2006).

#### a. Caractéristiques générales des systèmes de transport ABC : Organisation en domaines

Les transporteurs de type ABC possèdent généralement une structure commune composée d'au moins quatre domaines :

- Deux domaines hydrophobes transmembranaires (MSD : **M**embrane **S**panning **D**omain) formant un pore permettant au substrat de traverser la membrane. Ces domaines correspondent aux perméases.
- Deux domaines hydrophiles (NBD : **N**ucleotide **B**inding **D**omain), situés sur la face cytoplasmique de la membrane. Ces derniers sont responsables de la fixation d'ATP et du couplage de son hydrolyse au transport, ils correspondent aux ATPases,
- Un cinquième domaine, spécifique des systèmes d'import, est responsable de la capture du substrat (SBP : **S**ubstrat **B**inding **P**rotein) (Biemans-Oldehinkel *et al.*, 2006). Il s'agit d'une (ou des) protéine(s) affine(s) capable(s) de reconnaître le substrat assurant la spécificité et l'unidirectionnalité du transport (Higgins, 1992). Ces protéines se trouvent libres dans l'espace périplasmique chez les bactéries à Gram négatif, liées à la surface externe de la membrane cytoplasmique via des lipomodifications de leur extrémité N-terminale chez les bactéries à Gram positif (Gilson *et al.*, 1988 ; Tam et Saier, 1993).

Les systèmes de transport de type ABC présentent une grande variabilité dans l'organisation de leurs domaines. Ces domaines sont exprimés soit sous forme de chaînes polypeptidiques individuelles soit sous forme de protéines multidomaines (Figure 8) (Jones et George, 1999 ; Kerppola *et al.*, 1991). Par exemple, le transporteur de la glycine bêtaïne OpuA de *Lc. lactis* est une protéine qui comporte le domaine SBP et les domaines MSD (Van Der Heide et Poolman, 2000 ; Obis *et al.*, 1999). Par contre, le transporteur du ribose, Rbs, d'*Escherichia coli*, est formé de 3 protéines, une comportant les deux domaines NBD et les deux autres formant chacune un domaine MSD (Barroga *et al.*, 1996).



**Figure 8 : Organisation structurale des domaines de divers transporteurs ABC (Biemans-Oldenhinkel *et al.*, 2006)**

Les transporteurs ABC contiennent deux domaines « Nucleotide Binding Domain » (ovales) et deux domaines « Membrane Spanning Domain » (rectangles) et une (des) protéine(s) de liaison au substrat (chez les systèmes importateurs) (forme de pac-man). Les domaines NBD, TMD peuvent être des homo- ou des hétérodimères.

A) Exemples d'importateurs chez des bactéries à Gram négatif (*Escherichia coli*) : Nik : transporteur de nickel ; Mal : transporteur de maltose/dextrine ; Btu : transporteur de vitamine B12 ; Fhu : transporteur de sidérophore/hème/vitamine B12 ; Rbs : transporteur de ribose.

B) Exemples d'importateurs chez des bactéries à Gram positif (*Lactococcus lactis*) : Opp : transporteur d'oligopeptides ; OpuA : transporteur de glycine bêtaïne ; Gln : transporteur de glutamine/acide glutamique.

C) Exemples d'exportateurs bactériens (*Escherichia coli*) : Msb : flippase de lipides ; Cyd : transporteur de cystéine.

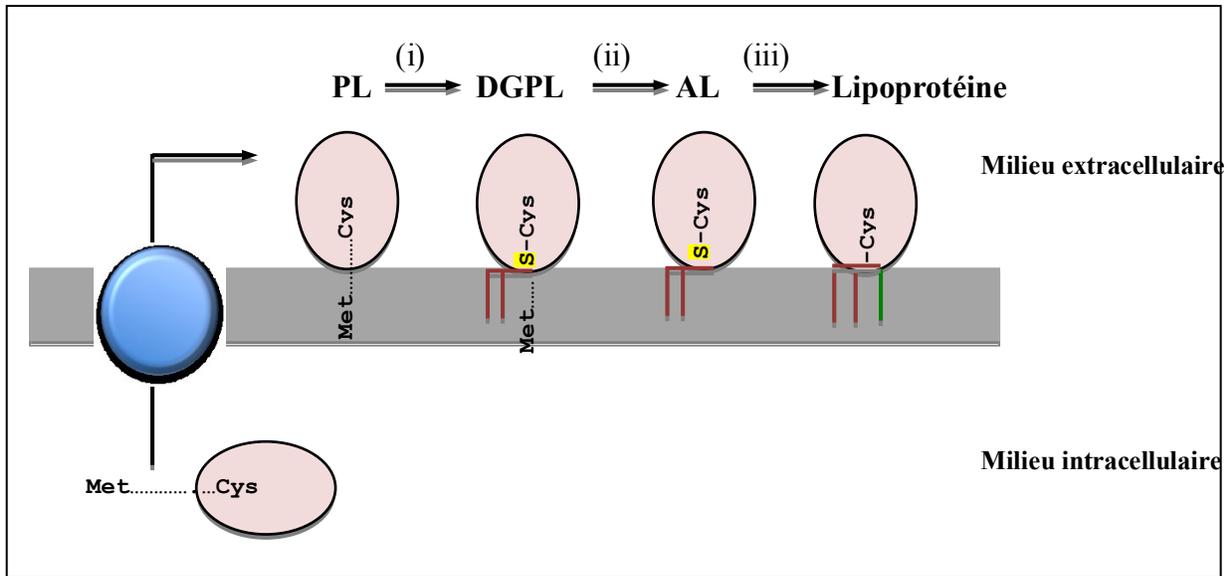
## b. La famille de transporteurs de type ABC des peptides chez les bactéries

### b.1. Caractéristiques des protéines de liaison aux peptides

L'alignement de 50 protéines de fixation de divers systèmes de transport bactérien (SBP) a permis de mettre en évidence l'existence de 8 groupes distincts (Tam et Saier, 1993). L'un d'entre-eux, le groupe OBP (Oligopeptide Binding Protein), renferme les protéines de fixation de peptides ayant comme séquence conservée (LIVM)AX<sub>2</sub>(WI)X<sub>1ou2</sub>(SN)(KE)DX<sub>4</sub>T(FY)X(LIV)RX<sub>3</sub>K. Cette séquence est également retrouvée chez les transporteurs de nickel. Il est donc impossible sur la base d'une comparaison de séquences de pouvoir différencier les transporteurs de nickel et de peptides.

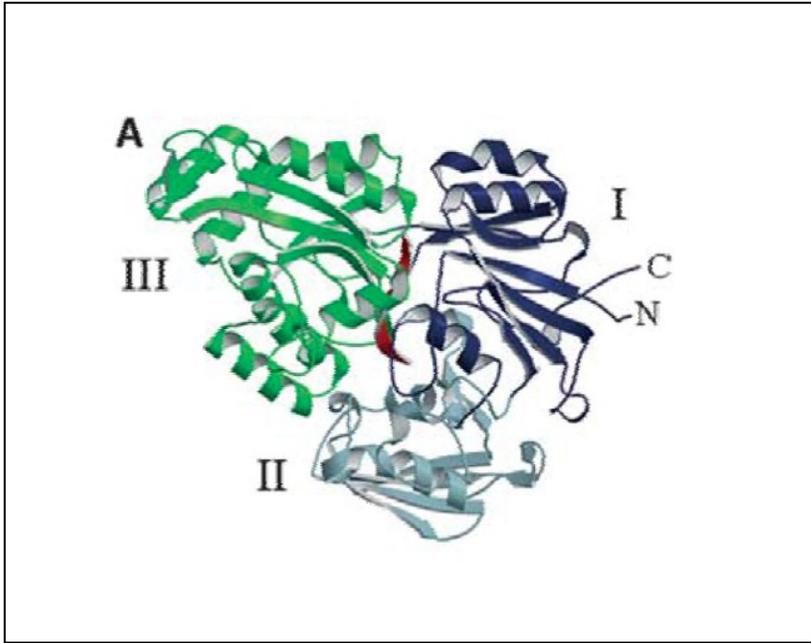
Chez les bactéries à Gram positif, ces protéines sont des lipoprotéines ancrées à la face externe de la membrane. Elles sont synthétisées sous forme de précurseur dans le cytoplasme, sont transloquées vers la face externe de la cellule par la voie générale de sécrétion Sec puis subissent trois modifications permettant leur ancrage à la membrane (Figure 9). Il y a tout d'abord formation d'une liaison thioéther entre la fonction thiol d'un résidu de cystéine situé du côté N-terminal du précurseur et un diacylglycérol formant la pro-lipoprotéine. Ensuite, le peptide signal de la pro-lipoprotéine est clivé par la « signal peptidase II » permettant ainsi la formation de l'apo-lipoprotéine et libérant la fonction amine du résidu de cystéine, résidu invariablement situé juste après la séquence signal, qui sera acylée (Figure 9) (Braun et Wu, 1994).

La structure tridimensionnelle de plusieurs OBP (OppA de *Salmonella typhimurium* et AppA de *Bacillus subtilis*) a révélé qu'elles sont constituées de deux lobes flexibles composées de deux domaines : domaines I et III reliés par deux segments permettant aux lobes de s'ouvrir et de se fermer pour pouvoir englober le peptide à transporter. De plus, un domaine spécifique de cette famille et de fonction inconnue a été trouvé (domaine II) (Figure 10) (Levdikov *et al.*, 2005 ; Tame *et al.*, 1994, 1995). Par rapport aux autres domaines du transporteur (NBD, MSD) où les motifs sont conservés, les séquences protéiques des OBP sont divergentes, il est donc très difficile de prédire la spécificité de substrat sur la base d'alignements de séquences.



**Figure 9 : Représentation schématique de la maturation des lipoprotéines chez les bactéries à Gram positif (Braun et Wu, 1994)**

- (i) Formation d'une liaison thioéther entre le résidu de cystéine (Cys) de l'extrémité N-terminale de la pro-lipoprotéine (PL) et un diacylglycérol (rouge), pour former un diacylglycéryl pro-lipoprotéine (DGPL)
- (ii) Clivage du peptide signal de la pro-lipoprotéine par la « signal peptidase II » pour former Apo-Lipoprotéine (AL)
- (iii) Acylation de la fonction amine du résidu de cystéine amino-terminal (acyl en vert)



**Figure 10 : Structure d'une protéine OBP «Oligopeptides Binding Protein» liée à un peptide (Levdikov *et al.*, 2005)**

Les domaines I et III forment les lobes flexibles (respectivement en bleu foncé et en vert), le domaine II n'a aucune fonction connue (bleu clair) et le peptide est coloré en rouge.

## b.2. Caractéristiques des deux domaines transmembranaires

Les protéines transmembranaires (ou MSD) sont des protéines hydrophobes formant un canal qui permet le passage du substrat à l'intérieur de la cellule. De plus, ces protéines font de plus le lien entre les protéines de fixation aux substrats et les ATPases, pour transmettre des signaux spécifiques. Fréquemment, chaque domaine est constitué de six segments hélice  $\alpha$  qui traversent la membrane (soit 12 segments par transporteur) avec plusieurs boucles extracellulaires et cytoplasmiques. Les extrémités amino et carboxy de chaque domaine sont localisées dans le cytoplasme. Toutefois, le nombre de segments peut varier entre 5 et 11 par domaine (Biemans-Oldehinkel *et al.*, 2006 ; Higgins, 1992). Concernant la séquence primaire, ces domaines sont plus variables que les domaines NBD (ou ATPase). Malgré ce faible niveau de conservation de séquence, un motif peptidique est conservé. Il s'agit du motif « EAA » (EAA-xxx-G-xxxxxxxxxx-I-x-LP, où x représente n'importe quel résidu d'acide aminé). Ce motif est localisé au niveau de la boucle cytoplasmique « L-loop » entre les segments transmembranaires 4 et 5. Il est présent

dans quelques systèmes d'importation mais pas dans les systèmes d'exportation (Mourez *et al.*, 1997 ; Saurin *et al.*, 1994). Jusqu'à présent son rôle n'est pas connu, il pourrait être impliqué dans la reconnaissance du substrat, ou dans l'interaction avec les domaines NBD (Dassa et Hofnung, 1985). Comme ce motif n'a pas été mis en évidence au sein de tous les systèmes d'exportation, il est peu probable qu'il joue un rôle majeur dans le fonctionnement du système.

### b.3. Caractéristiques des deux domaines de fixation de l'ATP

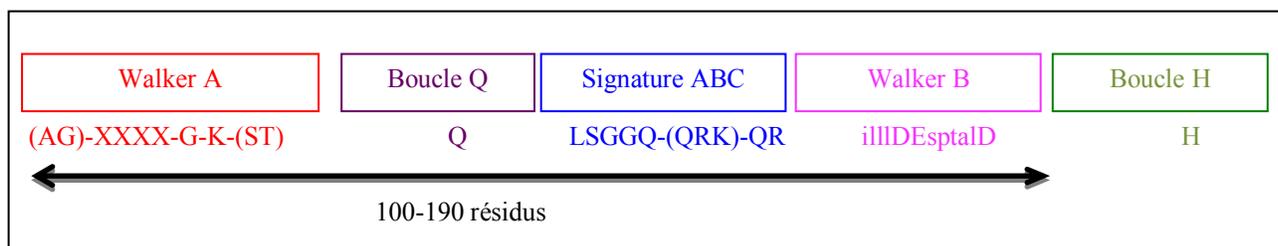
Les protéines ATPasiques (appelés NBD) sont des protéines hydrophiles situées sur la face cytoplasmique de la membrane, elles jouent un rôle primordial dans le transport puisqu'elles fournissent l'énergie nécessaire au transporteur par la fixation et l'hydrolyse de l'ATP. Généralement, chaque système de transport a besoin de deux domaines NBD pour sa fonctionnalité. Une des principales caractéristiques de tous les transporteurs ABC est une activité ATPase hautement dépendante des interactions coopératives entre les domaines NBD. Contrairement aux domaines MSD (formant le pore), les ATPases sont extrêmement conservées. Ainsi, chaque ATPase est constituée de deux sous-domaines : un sous-domaine catalytique RecA-like présent chez toutes les ATPases (Chen *et al.*, 2004), et un petit sous-domaine hélicoïdal spécifique de la famille des transporteurs ABC (Karpowich *et al.*, 2001). Ces deux sous-domaines sont reliés par une boucle flexible appelée « Q-loop ». Le sous-domaine RecA-like possède 3 motifs :

- Walker A appelé aussi boucle P « P-loop »,
- Walker B,
- et le motif « signature ABC ».

Ces 3 motifs participent à la fixation et à l'hydrolyse de l'ATP.

Le motif Walker A est toujours situé en amont du motif Walker B avec une distance entre ces deux motifs variant de 100 à 190 résidus (Saurin *et al.*, 1999 ; Walker *et al.*, 1982) (Figure 11). La séquence consensus proposée pour le motif Walker A est [AG]-xxxx-G-K-[ST]. Le motif Walker B est plus variable, néanmoins la séquence consensus proposée est illlDEpTsAlD (les acides aminés moins conservés sont en minuscules) (Saurin *et al.*, 1999). Ces deux motifs sont communs à toutes les protéines fixant l'ATP. Par contre, la « signature ABC » est spécifique des

transporteurs ABC et les distingue des autres ATPases, elle est composée jusqu'à 15 résidus d'acides aminés et commence généralement par la séquence LSGGQ. Une séquence consensus de la forme LSGGQ-[QRK]-QR a été proposée par Schneider et Hunke (1998), elle est située dans un domaine hélicoïdal en amont de Walker B. Ces trois domaines conservés sont essentiels à l'activité du transporteur, des mutations dans ces motifs conduisent à la perte d'activité de transport (Hung *et al.*, 1998 ; Schmees *et al.*, 1999). La boucle Q contenant un résidu de glutamine très conservé est localisée entre le motif Walker A et la « signature ABC » (Decottignies et Goffeau, 1997 ; Hyde *et al.*, 1990), cette boucle joue également un rôle dans l'hydrolyse de l'ATP. Enfin, la boucle H (contenant un résidu d'histidine) est une région conservée localisée à une trentaine de résidus après le motif Walker B, elle semble intervenir dans le couplage de l'hydrolyse de l'ATP et du transport (Schneider et Hunke, 1998).



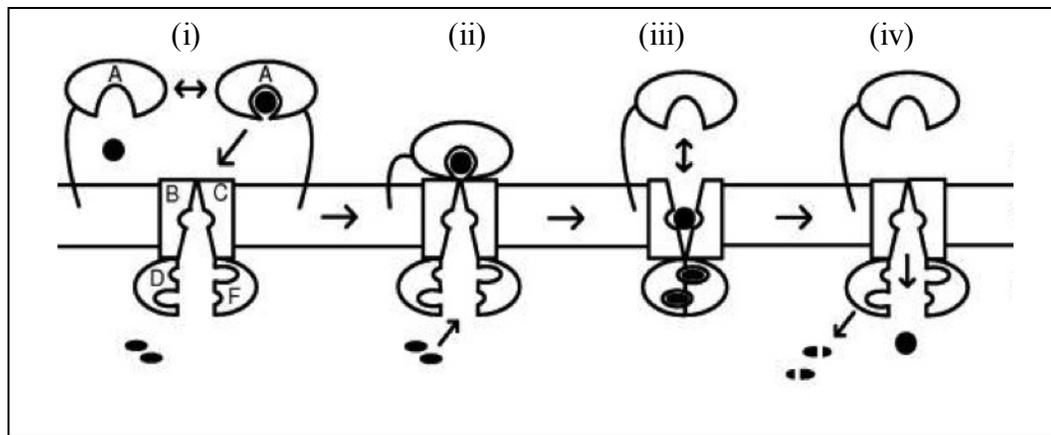
**Figure 11 : Organisation générale des motifs conservés d'un domaine NBD (Nucleotide-Binding Domain) chez les transporteurs ABC**

## B. Mécanisme de transport

Plusieurs modèles de fonctionnement de systèmes de transport ont été proposés, ils varient en fonction de l'organisation des domaines, du substrat, des bactéries. Le modèle de mécanisme de transport des oligopeptides du système Opp proposé chez *Lc. lactis* est présenté ci-après. Il se décompose en quatre étapes (Figure 12) :

- la première étape est une fixation réversible du substrat à la protéine affine OppA dans sa forme ouverte, ce qui entraîne un changement de conformation permettant la capture du peptide aboutissant à la forme OppA fermée ;

- au cours de la deuxième étape, la protéine OppA liée au substrat (forme fermée) se rapproche des perméases qui envoient un signal aux ATPases qui vont fixer l'ATP de manière coopérative ;
- les perméases subissent un changement conformationnel et le substrat est transféré de la protéine affine vers le canal transmembranaire ;
- enfin, l'hydrolyse de l'ATP provoque le changement conformationnel des perméases, aboutit à la libération du substrat dans le cytoplasme et au retour à l'état initial du système Opp (Doeven *et al.*, 2008 ; Lanfermeijer *et al.*, 1999).



**Figure 12 : Modèle de transport des peptides par le système Opp chez *Lactococcus lactis* (Doeven *et al.*, 2008 ; Lanfermeijer *et al.*, 1999)**

- fixation réversible du peptide à la protéine affine dans sa forme ouverte, puis changement de conformation de la protéine OppA permettant la capture du peptide ;
- rapprochement de la protéine OppA liée au peptide vers la perméase, et transmission d'un signal des protéines transmembranaires vers les ATPases, pour fixer l'ATP ;
- changement de conformation du complexe perméase de manière à permettre la translocation du substrat à travers la membrane ;
- transfert du peptide de façon irréversible de la protéine affine vers le canal transmembranaire jusqu'à la libération dans le cytoplasme

Les peptides issus des caséines hydrolysées par la protéase de paroi ou initialement présents dans le lait sont transportés à l'intérieur de la cellule par un ou plusieurs transporteurs (selon l'espèce et la souche). Chez les bactéries lactiques, les transporteurs de peptides appartiennent à deux grands groupes, ils diffèrent par le type d'énergie utilisée et par la spécificité de peptides pris en charge : les transporteurs de la famille PRT et les transporteurs ABC (Konings *et al.*, 1989 ; Kunji *et al.*, 1993 ; Tynkkynen *et al.*, 1993).

#### 4. 2. 1. 2. 2. Les systèmes de transports des peptides chez *Streptococcus thermophilus*

##### A. Transporteur de peptides proton dépendant

Un système de transport spécifique des di- tri-peptides a été caractérisé chez les lactocoques. Il s'agit du transporteur DtpT (**D**i-**t**ri-**p**eptide **T**ransporter). Initialement caractérisé chez *Lc. lactis* (Kunji *et al.*, 1993 ; Smid *et al.*, 1989), il appartient à la famille de transporteurs PTR. Il est constitué d'une seule protéine, codée par un gène chromosomique monocistronique, qui possède douze segments transmembranaires avec des extrémités N- et C- terminales extracellulaires (Guédon *et al.*, 2001a ; Hagting *et al.*, 1994). Il permet le transport d'une large variété de di- et tri-peptides, avec une meilleure affinité pour les dipeptides hydrophobes (Fang *et al.*, 2000). Chez *S. thermophilus*, aucun système de transport de di- tri-peptides n'a été caractérisé. Par contre, l'analyse des génomes des trois souches LMD-9, LMG18311 et CNRZ1066 révèle la présence d'un système de transport proton-dépendant homologue à DtpT de *Lc. lactis* (Bolotin *et al.*, 2004).

##### B. Système de transport de peptides ATP dépendant

Chez *Lc. lactis*, deux systèmes de transport de peptides de type ABC ont été identifiés : le premier est appelé Opp (**O**ligo-**p**eptide **p**erméase) et le deuxième Dpp (**D**i-**p**eptide **p**erméase) ou Opt. Le système Opp est constitué de cinq protéines : OppA, une protéine affine liant les oligopeptides, OppB et OppC, deux protéines transmembranaires, et enfin, OppD et OppF, deux protéines de fixation de l'ATP (Kunji *et al.*, 1993 ; Tynkkynen *et al.*, 1993). Le système Opt (Dpp) comprend 6 protéines : OptA et OptS, deux protéines de liaison aux peptides possédant 52,5% d'identité, OptB et OptC, deux protéines transmembranaires et OptD et OptF, deux ATPases (Lamarque *et al.*, 2004; Sanz *et al.*, 2001). Les pourcentages d'identité entre les protéines des systèmes Opt et Opp ne sont pas élevés (de 22 à 49,4%) (Tableau 5).

**Tableau 5 : Pourcentages d'identité entre les protéines du système de transport d'oligopeptides Opt et Opp de *Lactococcus lactis* IL1403 d'après (Lamarque *et al.*, 2004).**

Système Opt	Système Opp	Pourcentages d'identité
OptA	OppA	22,0%
OptS	(avec OppA)	24,5%
OptB	OppB	31,7%
OptC	OppC	30,8%
OptD	OppD	45,4%
OptF	OppF	49,4%

Ces deux transporteurs ne sont pas obligatoirement fonctionnels chez l'ensemble des souches. Ainsi, les deux systèmes Opp et Opt sont fonctionnels chez les deux souches SK11 et Wg2 (Lamarque *et al.*, 2004), mais chez la souche LMG1363, seul le système Opp semble responsable du transport d'oligopeptides, la protéine OptA (DppP) n'étant pas fonctionnelle. Chez la souche IL1403, le transport des oligopeptides semble dépendant du système Opt plutôt que du système Opp car la protéine OppA n'est pas fonctionnelle.

Un système homologue à Opt a été mis en évidence chez *S. pyogenes*. Il prend en charge le transport des dipeptides, et est constitué de cinq gènes organisés en opéron *dppABCDE*. L'expression de cet opéron varie en fonction des conditions de culture : un milieu riche en peptides réprime son expression (Podbielski et Leonard, 1998).

Chez *S. thermophilus*, un seul système de transport d'oligopeptides de type ABC a été caractérisé chez une souche industrielle, la souche ST18. Ce système a été nommé système Ami en raison de ces fortes similitudes avec le système de transport d'oligopeptides Ami de *S. pneumoniae*. Ce transporteur est constitué d'un translocon de quatre protéines : 2 protéines transmembranaires formant une perméase (AmiC et AmiD), de 2 protéines cytoplasmiques liant l'ATP (AmiE et AmiF), et contrairement aux systèmes de transport des oligopeptides chez *Lc. lactis*, le système Ami comporte trois protéines de liaison aux oligopeptides (AmiA1, AmiA2, AmiA3) (Figure 6). Les trois protéines AmiA sont associées à la membrane. Une séquence primaire (VLAACS) qui permettrait leur attachement à la membrane est retrouvée chez les protéines AmiA1, AmiA2 et AmiA3. Les trois protéines AmiA sont très apparentées, AmiA1 présente 97,6% d'identité avec AmiA2 et 87,1% avec AmiA3. Ce système est retrouvé chez de nombreux streptocoques. La

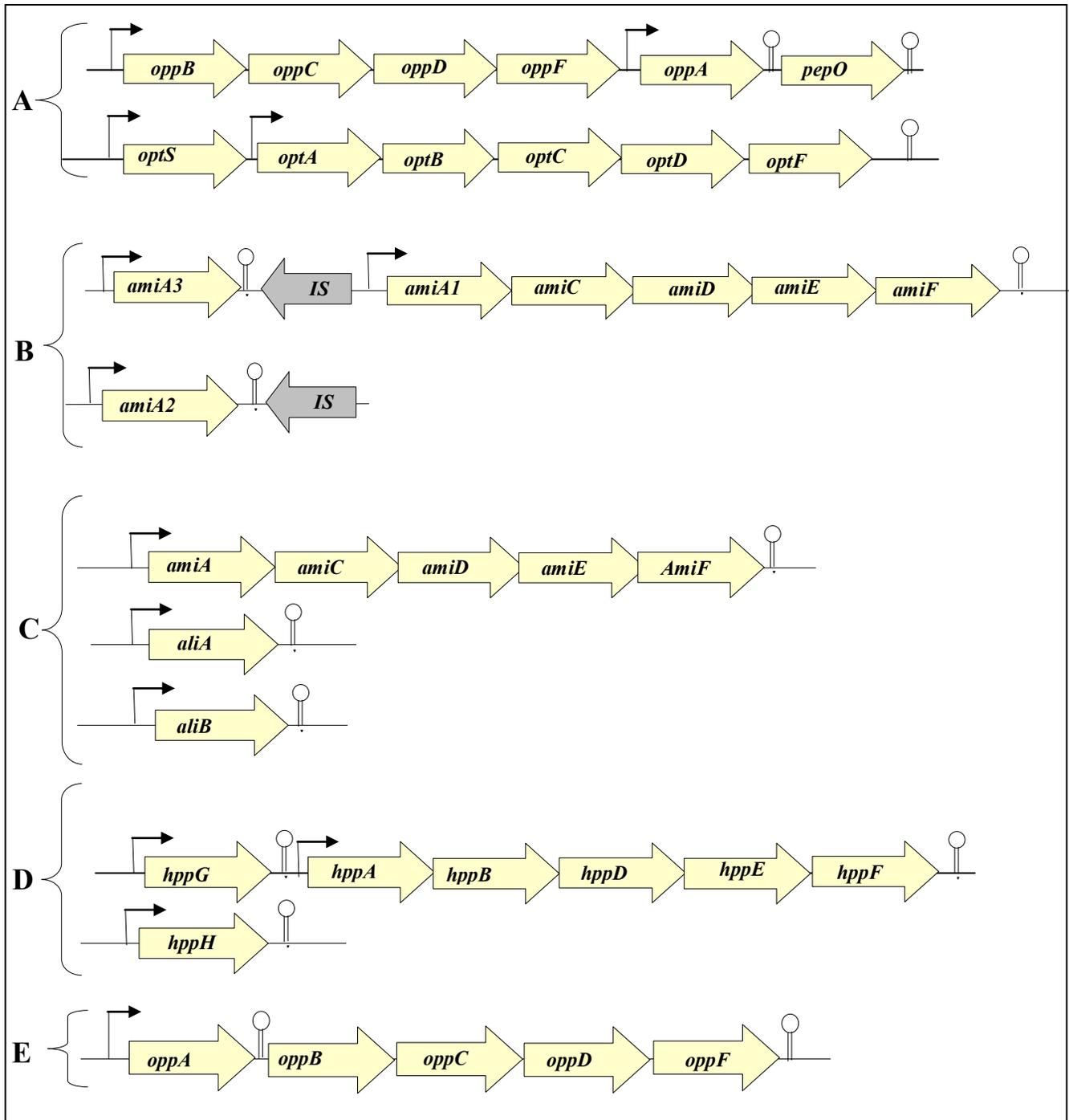
protéine AmiA1 de *S. thermophilus* présente 62%, 56% et 48% d'identité avec celle de *S. pneumoniae*, *S. gordonii* et *S. pyogenes* respectivement (Garault *et al.*, 2002). En revanche, aucun transporteur de di- et tri-peptides ATP-dépendant n'a été caractérisé chez *S. thermophilus*.

#### a. Organisation génétique du système de transport des oligopeptides

Les gènes codant les différentes protéines des systèmes de transport des oligopeptides sont généralement organisés en opéron et leur ordre est assez bien conservé, le gène codant la protéine OBP étant situé en amont. Cette organisation est retrouvée chez la plupart des bactéries à Gram positif telles que *B. subtilis*, *L. monocytogenes*, *S. pyogenes* ou *S. pneumoniae*. Cependant des variations dans l'ordre des gènes et le mode de transcription (mono/polycistronique) ont été observées chez certaines bactéries.

Chez *Lc. lactis* MG1363, les gènes *oppBCDFA* sont organisés en opéron avec le gène *pepO* qui code une oligopeptidase. L'opéron *oppBCDFApepO* comprend deux promoteurs qui sont respectivement situés en amont des gènes *oppB* et *oppA*, et deux terminateurs respectivement situés en aval des gènes *oppA* et *pepO* (Figure 13A). Cela implique donc que chez *Lc. lactis*, le gène *oppA* peut également être transcrit indépendamment des quatre autres gènes *oppBCDF* (Tynkkynen *et al.*, 1993). En ce qui concerne le système Opt, les deux gènes *optS* et *optA* peuvent être transcrits d'une part en opéron *optSABCD* mais aussi indépendamment, en raison de la présence d'un promoteur en amont des gènes *optS* et *optA*.

Chez *S. thermophilus* ST18, les quatre gènes codant le translocon (*amiC*, *amiD*, *amiE* et *amiF*) sont organisés en opéron avec le gène *amiA1*. Le gène *amiA1* se situe en amont de l'opéron *amiA1CDEF*, alors que les gènes *amiA2* et *amiA3* sont situés ailleurs sur le chromosome (Figure 13B). Cet opéron comprend un seul promoteur situé en amont du gène *amiA1*, et un seul terminateur situé en aval du gène *amiF*. De ce fait, le gène *amiA1* est transcrit avec les quatre gènes *amiC*, *amiD*, *amiE* et *amiF* tandis que les gènes *amiA2* et *amiA3* sont transcrits de façon indépendante de cet opéron (Garault *et al.*, 2002). La même organisation et le même mode de transcription ont été trouvés chez *S. pneumoniae* (*amiA*) et *S. gordonii* (*hppA*). Par contre, chez *S. pyogenes*, *oppA* peut être transcrit indépendamment grâce à la présence d'un terminateur faible en aval d'*oppA* (Figure 13CDE) (Podbielski *et al.*, 1996).



**Figure 13 : Organisation des gènes codant les protéines des transporteurs d'oligopeptides chez différentes espèces bactériennes**

A). *Lactococcus lactis* (systèmes Opp et Opt chez la souche LMG1363) ; B). *Streptococcus thermophilus* (système Ami chez la souche ST18) ; C). *Streptococcus pneumoniae* ; D). *Streptococcus gordonii* ; E). *Streptococcus pyogenes* (Alloing *et al.*, 1994 ; Garault *et al.*, 2002 ; Monnet *et al.*, 2003 ; Podbielski et Leonard, 1998).

↳ : Promoteur ; | : Termineur. ; IS : Séquence d'insertion.

Une situation originale a été constatée chez *S. thermophilus* ST18. En effet, la forte identité entre les trois protéines AmiA et la présence de séquence d'insertion IS en amont et/ou en aval de chaque gène *amiA*, suggère que les gènes *amiA2* et *amiA3* sont issus d'une double duplication récente du gène *amiA1*. Cette duplication semble avantageuse pour la bactérie et favoriserait la croissance de *S. thermophilus* en lait où le transport d'oligopeptides est essentiel pour une bonne croissance. La présence de plusieurs copies du gène *amiA* semble être une caractéristique de l'espèce *S. thermophilus*. Au moins une copie du gène *amiA* est présente chez toutes les souches étudiées (Garault *et al.*, 2002).

La présence de plusieurs protéines affines aux oligopeptides a déjà été décrite chez d'autres streptocoques. Par exemple, le système Ami de *S. pneumoniae* comporte trois protéines affines aux oligopeptides. Les gènes codant les protéines constituant le translocon et un des gènes codant une protéine affine sont organisés dans un opéron *amiACDEF*. Deux protéines affines aux oligopeptides AliA et AliB codée respectivement par les gènes *aliA* et *aliB*, tout comme la protéine affine AmiA sont fonctionnelles et utilisent le même translocon AmiCDEF (Figure 13C) (Alloing *et al.*, 1994). Chez *S. gordonii*, les quatre gènes *hppB*, *hppD*, *hppE*, *hppF* codant le translocon et le gène *hppA* codant une OBP sont organisés en opéron (*hppABDEF*). Les gènes *hppB* et *hppD* codent deux protéines formant une perméase et les gènes *hppE* et *hppF* codent deux ATPases. Deux autres gènes codant deux protéines affines sont codés par les gènes *hppG* et *hppH*. Le gène *hppG* se trouve en amont de l'opéron *hppABDEF* et le gène *hppH* est situé ailleurs sur le chromosome (Figure 13D) (Jenkinson *et al.*, 1996). Cependant l'identité entre les protéines affines chez ces deux espèces semble moins importante que chez *S. thermophilus* (chez *S. pneumoniae* 67% d'identité entre AmiA et AliA, 65% entre AmiA et AliB ; chez *S. gordonii* 68% d'identité entre HppA et HppG, 65% entre HppA et HppH) (Alloing *et al.*, 1994 ; Jenkinson *et al.*, 1996 ; Monnet *et al.*, 2003).

#### b. Spécificité de substrat des transporteurs de peptides

La spécificité de substrat des transporteurs de peptides dépend évidemment des OBP, puisque ce sont elles qui interagissent en premier avec le substrat, mais aussi, des perméases (Charbonnel *et al.*, 2003 ; Doeven *et al.*, 2005 ; Helinck *et al.*, 2003).

Chez *Lc. lactis*, OppA est capable de fixer des peptides d'une taille maximale de 35 résidus alors que la perméase du système (Opp) n'est capable de transporter que des peptides

contenant 4 à 18 résidus d'acides aminés (Detmers *et al.*, 1998, 2000 ; Doeven *et al.*, 2004) avec une préférence pour les peptides hydrophobes basiques dont la masse moléculaire est comprise entre 600 et 1100 Da (Juillard *et al.*, 1998). Concernant le système Opt, la protéine OptA fixe des peptides de 3 à 6 résidus, alors la protéine OptS semble avoir une spécificité de fixation plus restreinte, spécifique de certains di- et tri-peptides (Lamarque *et al.*, 2011).

Chez *S. thermophilus*, la spécificité des trois protéines AmiA1, AmiA2 et AmiA3 se chevauchent, bien qu'elles soient différentes. L'ensemble du système Ami est capable d'internaliser des peptides de 3 à 23 résidus d'acides aminés (Garault *et al.*, 2002) avec une préférence pour les peptides hydrophobes dont la masse moléculaire comprise entre 1000 et 3500 Da (Juille *et al.*, 2005). La spécificité de substrat du système Ami semble similaire à celle du système Opp de *Lc. lactis* et distincte de celles des autres streptocoques, et ce, malgré les différences qui existent entre le système Opp de *Lc. lactis* et le système Ami de *S. thermophilus* au niveau du nombre de leurs OBP, de l'organisation génétique ou du faible pourcentage d'identité de leurs séquences protéiques (de 24,2 à 43,4%) (Tableaux 6, 7) (Juille *et al.*, 2005).

Chez les autres streptocoques, la taille maximale des peptides transportés est inférieure à 10 résidus d'acides aminés. Par exemple, chez *S. gordonii*, le système Hpp transporte des peptides de 5 à 7 résidus (Jenkinson *et al.*, 1996) (Tableau 7).

**Tableau 6 : Pourcentages d'identité entre les protéines du système Ami de *Streptococcus thermophilus* ST18 et du système Opp de *Lactococcus lactis* IL1403 (d'après Garault *et al.*, 2002)**

<i>Streptococcus thermophilus</i>	<i>Lactococcus lactis</i>	Pourcentage d'identité
AmiA1	OppA	24,2%
AmiC	OppB	30,7%
AmiD	OppC	30,4%
AmiE	OppD	46,6%
AmiF	OppF	43,4%

**Tableau 7 : Taille des peptides incorporés par des transporteurs de peptides bactériens de type ABC chez quelques espèces bactériennes**

Organisme	Nom du système	Taille des peptides transportés (résidus)	Références
<i>Lactococcus lactis</i>	Opt	3-6 <sup>a</sup>	(Foucaud <i>et al.</i> , 1995 ; Lamarque <i>et al.</i> , 2004, 2011)
<i>Lactococcus lactis</i>	Opp	4-18	(Doeven <i>et al.</i> , 2004)
<i>Streptococcus thermophilus</i>	Ami	3-23	(Garault <i>et al.</i> , 2002)
<i>Streptococcus agalactiae</i>	Dpp	2	(Samen <i>et al.</i> , 2004)
<i>Streptococcus agalactiae</i>	Opp	2-6 <sup>b</sup>	(Samen <i>et al.</i> , 2004)
<i>Streptococcus pyogenes</i>	Dpp	2	(Podbielski et Leonard, 1998)
<i>Streptococcus pyogenes</i>	Opp	6 <sup>c</sup>	(Podbielski et Leonard, 1998)
<i>Streptococcus gordonii</i>	Hpp	6-7	(Jenkinson <i>et al.</i> , 1996)
<i>Streptococcus pneumoniae</i>	Ami	2-7 <sup>b</sup>	(Alloing <i>et al.</i> , 1994)
<i>Streptococcus uberis</i>	Opp	3-8	(Taylor <i>et al.</i> , 2003)
<i>Bacillus subtilis</i>	Opp	3-5 <sup>b</sup>	(Koide et Hoch, 1994 ; Picon et Van Wely, 2001)
<i>Bacillus subtilis</i>	App	4-9 <sup>b</sup>	(Koide et Hoch, 1994 ; Levnikov <i>et al.</i> , 2005)
<i>Escherichia coli</i>	Opp	2-5 <sup>b</sup>	(Alves et Payne, 1980 ; Guyer <i>et al.</i> , 1986)
<i>Escherichia coli</i>	Dpp	2-3	(Abouhamad et Manson, 1994 ; Dunten et Mowbray, 1995 ; Nickitenko <i>et al.</i> , 1995)
<i>Listeria monocytogenes</i>	Opp	3-8	(Borezee <i>et al.</i> , 2000 ; Verheul <i>et al.</i> , 1998)

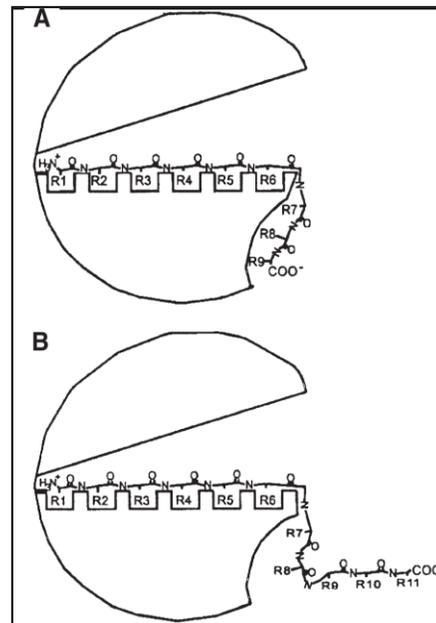
<sup>a</sup> La taille des peptides transportés change selon les souches, le système Opt possède une ou deux protéines affines dont la spécificité de transport est variée

<sup>b</sup> Indique qu'aucun ou très peu de peptides de taille supérieure à celle indiquée ont été testés dans l'étude.

<sup>c</sup> Des hexapeptides ont été uniquement testés.

### c. Mode d'affinité des protéines de liaison aux peptides

L'affinité de la protéine OppA à son substrat a été étudiée chez *Lc. lactis* dans des conditions proches de la physiologie des cellules par construction de protéoliposomes. Ainsi, OppA a été attachée à la surface de liposomes par son site de lipomodification afin de mimer son attachement à la surface des cellules. Une banque combinatoire de peptides a alors été utilisée pour déterminer sa spécificité (Detmers *et al.*, 2000). Il a ainsi été observé qu'OppA est capable de fixer une large gamme de peptides allant de 4 à 35 résidus avec une affinité accrue pour les nonapeptides. Les résidus dans les positions 4, 5 et 6 sont les plus sélectifs avec une affinité plus élevée pour les résidus hydrophobes, alors que la glycine, la proline et les résidus chargés négativement baissent l'affinité. L'utilisation de peptides fluorescents a montré que seuls les 6 premiers résidus de l'extrémité N-terminale sont enfouis dans la poche de la protéine de fixation alors que les résidus restants interagissent avec sa surface externe (Figure 14) (Lanfermeijer *et al.*, 2000).



**Figure 14 : Représentation schématique du mécanisme de liaison du peptide à la protéine OppA de *Lactococcus lactis* (Detmers *et al.*, 2000)**

A) le cas d'un nonapeptide ; B) le cas d'un décapeptide.

#### d. Différents rôles des transporteurs des peptides

Chez les bactéries à Gram positif, le système de transport d'oligopeptides joue non seulement un rôle important dans la nutrition azotée, mais il intervient aussi dans les processus qui impliquent des peptides comme molécules de communication, tels que les processus de sporulation et compétence chez *B. subtilis* et *S. pneumoniae* (Cheng *et al.*, 1997 ; Lazizzera *et al.*, 1999 ; Stragier et Losick, 1996), de virulence chez plusieurs bactéries pathogènes (Borezee *et al.*, 2000 ; Podbielski *et al.*, 1996 ; Wang *et al.*, 2005), ou de conjugaison chez *En. faecalis* (Dunny et Leonard, 1997 ; Leonard *et al.*, 1996).

##### ❖ Rôle dans la nutrition azotée

L'impact de ces transporteurs dans la nutrition azotée a été évalué par inactivation des gènes codant ces transporteurs puis comparaison de la croissance de la souche sauvage avec celle de la souche mutante en lait ou en milieu chimiquement défini contenant les acides aminés sous forme d'oligopeptides. Ainsi, un mutant *opp<sup>-</sup>* chez *Lc. lactis* MG1363 est incapable de croître en lait ou en milieu contenant la source azotée sous forme d'oligopeptides (Juillard *et al.*, 1995 ; Picon *et al.*, 2000). Le système DtpT semble, quant à lui, jouer un rôle minoritaire dans la nutrition azotée en lait puisque son inactivation n'affecte pas la densité cellulaire du mutant en fin de culture (Juillard *et al.*, 1996 ; Kunji *et al.*, 1995).

Chez *S. thermophilus* ST18, il a été montré que le système Ami est essentiel pour la croissance en lait. Les trois protéines affines AmiA n'ont pas la même importance dans l'approvisionnement en peptides de la bactérie. L'inactivation de la protéine AmiA3 a l'effet le plus important sur la croissance en lait et MCD supplémenté en acides aminés sous forme de peptides, alors qu'AmiA3 est la protéine affine la moins apparentée aux protéines AmiA1 et AmiA2. Quant à la protéine AmiA1, elle semble avoir l'effet le plus faible sur la croissance en lait (Garault *et al.*, 2002).

Chez *S. pneumoniae* et *S. gordonii*, le système Ami permet la croissance dans un milieu complexe dont la source azotée est apportée sous forme de peptides (Alloing *et al.*, 1994 ; Jenkinson *et al.*, 1996). Chez *S. gordonii*, parmi les trois protéines affines aux oligopeptides (HppA, HppG et HppH), seule HppA semble avoir un rôle dans la croissance en milieu

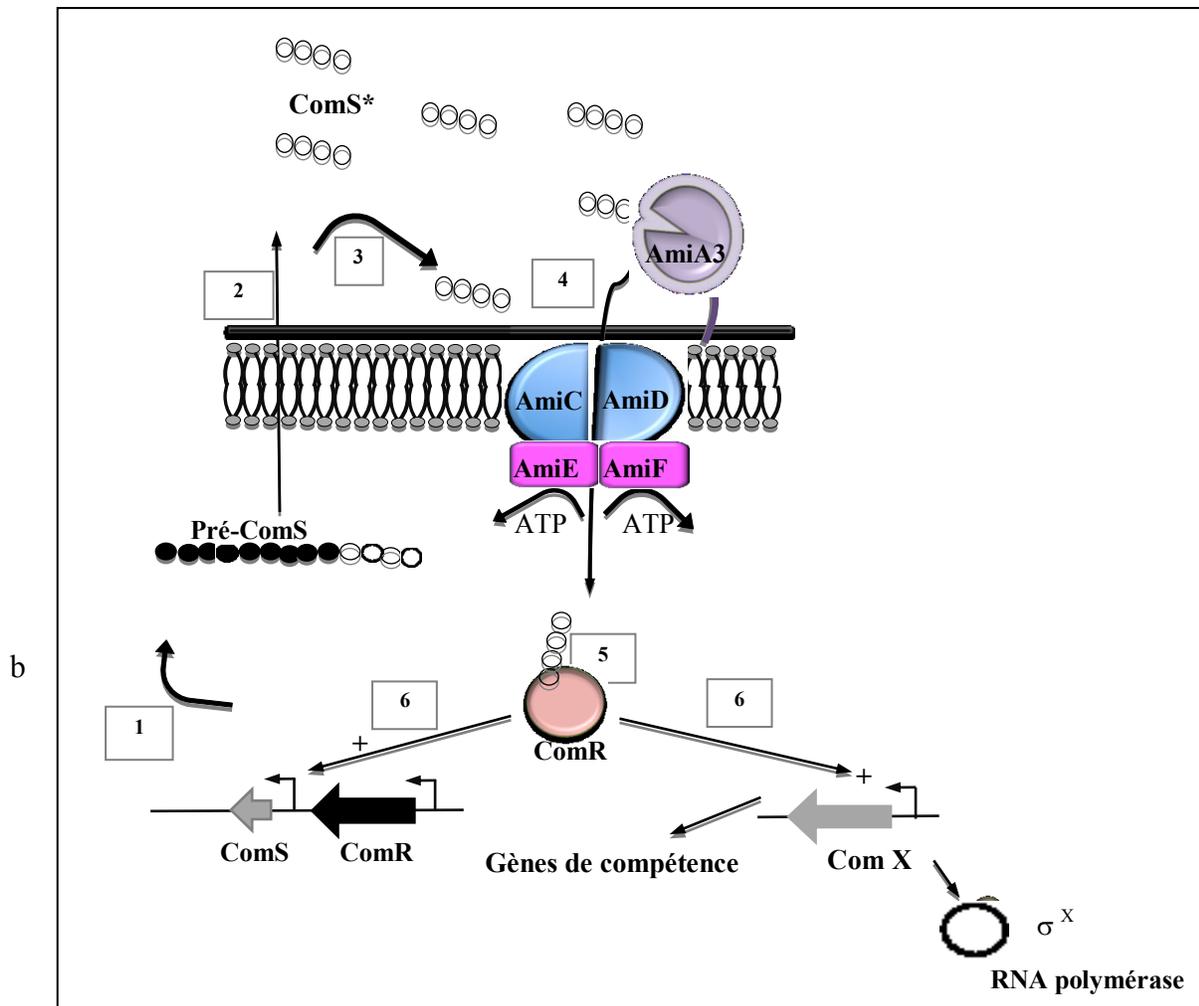
complexe, car la croissance d'un simple ou double mutant *hppG*<sup>-</sup> et/ou *hppH* ne semble pas affectée en milieu complexe (Jenkinson *et al.*, 1996).

❖ Rôle dans le renseignement sur les conditions environnementales

Les bactéries communiquent entre-elles grâce à de petites molécules. Cette communication, encore appelée *quorum sensing*, leur permet d'échanger des informations sur les conditions environnementales. Les molécules de signalisation s'accumulent jusqu'à une concentration seuil (*quorum*), qui constitue un signal à partir duquel la bactérie module l'expression de certains gènes cibles (*sensing*).

Les molécules de signalisation sont synthétisées dans le cytoplasme, sous forme de précurseurs, puis sécrétées. Dans certaines conditions et après maturation, elles sont ré-internalisées et permettent la modulation de l'expression de certains gènes par interaction avec certains régulateurs. In fine, cela aboutit à l'adaptation des bactéries aux changements de conditions environnementales (Lazazzera *et al.*, 1997 ; Slamti et Lereclus, 2002). Pour être internalisées, ces molécules empruntent un système de transport tel qu'un système de transport d'oligopeptides de type ABC.

Chez *S. thermophilus*, le système système Ami intervient de façon indirecte dans l'expression des gènes en jouant un rôle de « renseignement » sur les conditions environnementales. En effet, lorsque la souche LMD-9 est cultivée en milieu chimiquement défini, la compétence naturelle se déclenche par la production d'un petit peptide hydrophobe appelé ComS (pour **C**ompetence **S**ignal) de 24 résidus. Ce peptide signal est produit sous forme d'un précurseur, puis sécrété et mûri. A partir d'une certaine concentration dans le milieu extracellulaire, le peptide mature baptisé ComS\* est ré-internalisé essentiellement par AmiA3. Une fois à l'intérieur de la cellule, ComS\* interagit avec le régulateur ComR (**C**ompetence **R**egulator), qui appartient à la famille des régulateurs Rgg. ComR, une fois, activé, va à son tour, moduler la transcription de deux gènes : d'une part, le gène *comS*, et d'autre part le gène *comX* codant un facteur sigma alternatif qui régule les gènes tardifs de la compétence (Fontaine *et al.*, 2010 ; Gardan *et al.*, 2009) (Figure 15).



**Figure 15 : Modèle d'activation de la compétence naturelle chez la souche *Streptococcus thermophilus* LMD-9 lors de la croissance en milieu chimiquement défini (Fontaine *et al.*, 2010)**

1. Déclenchement de la transcription du gène *comS*, en milieu chimiquement défini, pour synthétiser le précurseur pré-ComS.
2. Sécrétion de pré-ComS, probablement *via* le système de sécrétion générale.
3. Maturation de pré-ComS et accumulation de sa forme mature (ComS\*) dans le milieu extracellulaire.
4. A partir d'une certaine concentration, le peptide ComS\* serait reconnu par la protéine AmiA3 et internalisé *via* le système Ami.
5. Interaction de ComS\* internalisé avec le régulateur transcriptionnel ComR.
6. Fixation de la protéine ComR à son opérateur situé en amont de la boîte -10 des gènes *comS* et *comX*, pour activer ces gènes respectivement. Les flèches noires indiquent les promoteurs ; + : activation des gènes cibles par ComR

Récemment chez *Lc. lactis* IL1403, un rôle dans la perception de l'environnement a été attribué au système Opt via la protéine affine OptS (Lamarque *et al.*, 2011). La protéine OptS, contrairement au DtpT permet, spécifiquement en lait, l'entrée de peptides de signalment qui peuvent agir sur le régulateur CodY et donc moduler l'activité du système protéolytique.

Chez les streptocoques pathogènes, en plus de leur rôle dans la nutrition azotée, les systèmes de transport des oligopeptides jouent également un rôle dans la régulation des gènes impliqués dans la virulence des bactéries. Chez *S. pneumoniae*, le système de transport Ami participe directement ou indirectement dans la régulation des gènes de la compétence ou de l'adhérence. Deux voies possibles de régulation de la compétence ont été proposées. La première mettrait en jeu les peptides internalisés à l'intérieur de la cellule par le système Ami qui une fois à l'intérieur de la cellule, seraient dégradés en acides aminés, dont certains joueraient un rôle important dans la régulation générale des gènes impliqués dans plusieurs fonctions comme le transport des acides aminés, la compétence et l'adhérence des cellules. La deuxième impliquerait probablement des molécules de signalment (des phéromones) qui pourraient être internalisées aussi par le système Ami (Claverys *et al.*, 2000, 2006).

#### **4. 2. 1. 3. Les peptidases intracellulaires de *Streptococcus thermophilus***

Les peptides internalisés par le système de transport sont ensuite dégradés par un pool de peptidases intracellulaires avec des spécificités partiellement chevauchantes (Kunji *et al.*, 1996). Un grand nombre de ces peptidases intracellulaires a été purifié et caractérisé biochimiquement chez différents genres de bactéries lactiques : *Lactococcus*, *Lactobacillus* et *Streptococcus*.

Chez *S. thermophilus* les caractéristiques biochimiques de quinze peptidases intracellulaires ont été étudiées (PepA, PepS, PepN, PepC, PepB, PepQ, PepV, PepT, PepM, PepXP, PepP, PepO, PepX, AP1 et EP), dont les deux dernières (AP1 : un amiopeptidase, EP : un oligopeptidase) spécifiques de *S. thermophilus* n'ont pas été retrouvées chez les lactococques (Chapot-Chartier *et al.*, 1994 ; Chavagnat *et al.*, 2000 ; Liu *et al.*, 2009a ; Rul et Monnet, 1997 ; Rul *et al.*, 1994 ; Tsakalidou *et al.*, 1997). *S. thermophilus* obtient donc grâce à sa diversité de peptidases intracellulaires une diversité d'acides aminés exogènes plus large que celle d'autres BL.

Les peptidases peuvent être classées en exopeptidases ou endopeptidases. Les exopeptidases hydrolysent les liaisons peptidiques à partir des extrémités amino (aminopeptidases) ou carboxy

(carboxypeptidase) terminales des chaînes polypeptidiques tandis que les endopeptidases hydrolysent les liaisons peptidiques internes à l'oligopeptide (Tableau 8).

**Tableau 8 : Tableau récapitulatif des peptidases caractérisées biochimiquement ou déduites de séquences codantes chez *Streptococcus thermophilus*.**

Peptidases ou activités peptidasiques Spécificité de coupure chez la(les) souche(s) étudiée(s) et caractéristiques biochimiques	Numéro d'accèsion du gène codant pour la peptidase chez les souches dont le génome est entièrement séquencé				
	LMD-9	LMG18311	CNRZ1066	ND03	JIM8232
<b>Endopeptidases</b>					
<b>PepO</b> : Métallopeptidase (endopeptidasique) Souche A : 70 kDa ; p <sub>H</sub> 6,5 ; T <sub>o</sub> 41°C (Chavagnat <i>et al.</i> , 2000)	YP_821179	YP_140299	YP_142213	ADQ63857	CCC20811
<b>PepF/PepB</b>	YP_819966	YP_140869	YP_138979	ADQ62516	CCC19292
<b>Oligoendopeptidase F</b>	YP_820673				CCC20219
<b>Amino-peptidases</b>					
<b>PepA</b> : Métallopeptidase (Glu-X et Asp-X) Souche CNRZ 302 : 45 kDa ; p <sub>H</sub> 8,5 ; T <sub>o</sub> 62°C (Rul <i>et al.</i> , 1995)	YP_821150	YP_140265	YP_142180	ADQ63821	CCC20773
<b>PepN</b> : Métallopeptidase à spécificité large : résidus ni acides ni Pro en position P2 Souche CNRZ302 : 97 kDa ; p <sub>H</sub> 7,0 ; T <sub>o</sub> 36°C (Rul <i>et al.</i> , 1994) Souche A : 95 kDa ; p <sub>H</sub> 7,0 ; T <sub>o</sub> 37°C (Chavagnat <i>et al.</i> , 1999)	YP_820418	YP_139482	YP_141400	ADQ63016	CCC19904
<b>PepC</b> : Peptidase à cystéine à spécificité large : résidus basiques, neutres hydrophobes et acides Souche CNRZ 302 : 50,4 kDa, p <sub>H</sub> 7,0, T <sub>o</sub> 50°C (Chapot-Chartier <i>et al.</i> , 1994)	YP_819783	YP_138769	YP_140658	ADQ62298	CCC19071
<b>PepS</b> : Métallopeptidase (Arg-X, Tyr-X, Phe-X et Trp-X). Souche CNRZ 302 : 45 kDa ; p <sub>H</sub> 8, ; T <sub>o</sub> 55°C (Fernandez-Espla et Rul, 1999)	YP_819646	YP_138607	YP_140495	ADQ62144	CCC18883
<b>PepM</b> (Liu <i>et al.</i> , 2009a)	YP_820865	YP_139975	YP_141902	ADQ63528	CCC20450
<b>Peptidases spécifiques des résidus de proline</b>					
<b>PepX</b> : Peptidase à sérine (Pro-X) Souche ACA-DC4 : 80 kDa ; p <sub>H</sub> 7,0 ; T <sub>o</sub> 50°C (Anastasiou <i>et al.</i> , 2002; Tsakalidou <i>et al.</i> , 1998)	YP_820974	YP_140089	YP_142016	ADQ63642	CCC20579
<b>PepP</b> : Activité peptidasique (X-Pro) Souche CNRZ302 (Rul et Monnet, 1997)	YP_821056	YP_140156	YP_142075	ADQ63712	CCC20657
<b>PepQ</b> : Activité peptidasique (X-Pro) Souche CNRZ302 (Rul et Monnet, 1997)	YP_820133	YP_139150	YP_141040	ADQ62690	CCC19530
<b>PepR</b> : Activité peptidasique (Pro-X) Souche CNRZ302 (Rul et Monnet, 1997)					
<b>Di-/tri-peptidase</b>					
<b>PepV</b>	YP_820475	YP_139581	YP_141490	ADQ63093	CCC19999
<b>PepT</b>	YP_820484	YP_139592	YP_141502	ADQ63103	CCC20010

#### 4. 2. 1. 4. Les transporteurs d'acides aminés libres

Plusieurs mécanismes de transport d'acides aminés ont été décrits chez les BL (Konings *et al.*, 1989).

- Des systèmes de transport actif dont l'énergie est fournie par l'ATP ou d'autres composés phosphorylés riches en énergie. Ces systèmes catalysent généralement des réactions unidirectionnelles, irréversibles et permettent l'accumulation intracellulaire des acides aminés (Driessen *et al.*, 1987). Chez *S. thermophilus*, ils sont utilisés pour transporter les acides aminés à chaîne ramifiée, la méthionine, la glutamine, et les acides aminés polaires (Hols *et al.*, 2005 ; Juilliard, 2007). Chez *S. thermophilus*, comme chez *S. pneumoniae*, le transport de la glutamine est assuré par plusieurs transporteurs de type ABC soulignant l'importance de la glutamine pour la croissance des streptocoques.

- Des transports couplés à la force protomotrice permettent le transport simultané et réversible d'acides aminés et de protons (ou d'autres cations comme le Na<sup>+</sup>). L'alanine, la glycine, la sérine et la thréonine peuvent utiliser ce type de transport (Ser/thr : Na) chez *S. thermophilus*.

- Enfin, le transport par système antiport implique des gradients de concentration des acides aminés transportés. Ainsi, l'entrée d'un acide aminé précurseur est couplée à la sortie d'un produit de ce précurseur comme c'est le cas du couple arginine/ornithine chez *S. thermophilus*.

#### 4. 3. Activité uréasique chez *Streptococcus thermophilus*

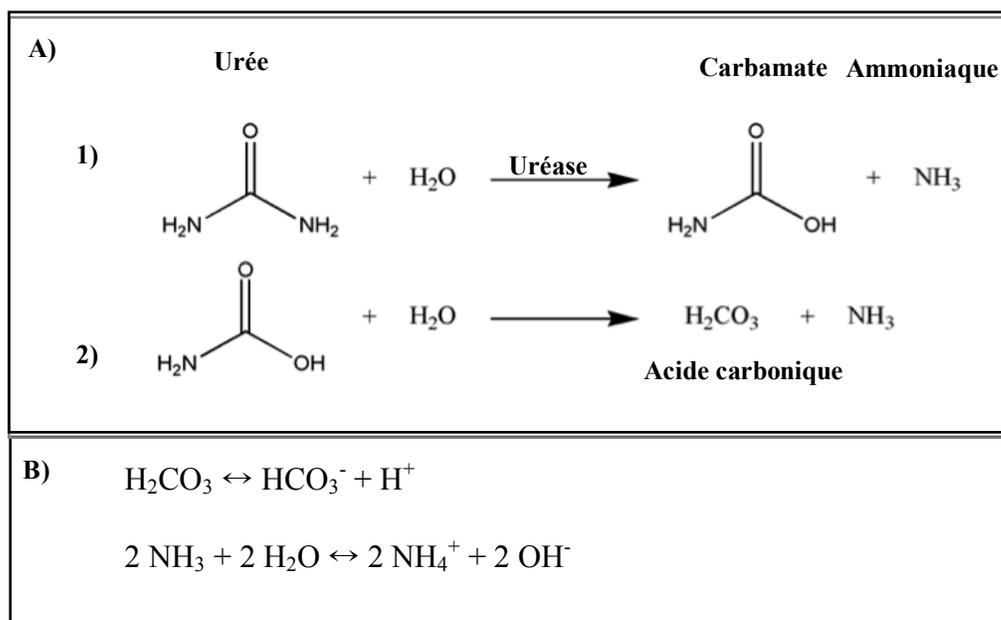
L'acide lactique obtenu lors de la fermentation du lactose par *S. thermophilus*, joue un rôle important dans la coagulation du lait et dans la sécurité sanitaire puisque le développement de microorganismes pathogènes dans le lait est empêché. Pour ces raisons, lors de la fabrication des produits laitiers, la vitesse d'acidification est un critère technologique très important pour les industriels. Elle dépend du métabolisme des souches, lui même influencé par divers facteurs tels que le système protéolytique, le métabolisme glucidique et l'activité uréasique (Dandoy *et al.*, 2011 ; Galia *et al.*, 2009 ; Monnet *et al.*, 2004 ; Mora *et al.*, 2004). Parmi les bactéries lactiques utilisées dans des produits laitiers fermentés, *S. thermophilus* est la seule bactérie à posséder une activité uréasique. L'uréase est une enzyme multimérique catalysant l'hydrolyse de l'urée en ammoniaque et en carbamate, celui-ci se décompose alors spontanément pour libérer une

deuxième molécule d'ammoniaque et une molécule d'acide carbonique (Juillard *et al.*, 1988 ; Mora *et al.*, 2004 ; Tinson *et al.*, 1982 ; Zourari et Desmazeaud, 1991 ; Zourari *et al.*, 1991). Au pH physiologique, l'acide carbonique ( $\text{H}_2\text{CO}_3$ ) se dissocie en hydrogénocarbonate ( $\text{HCO}_3^-$ ) et en proton ( $\text{H}^+$ ), les molécules d'ammoniaque sont en équilibre avec les ions ammonium (Mobley et Hausinger, 1989). (Figure 16). Au cours de la fermentation, en fonction de la quantité d'urée présente dans le lait ( $0,2 - 0,4 \text{ g.L}^{-1}$ ), ces réactions affectent le taux de diminution du pH (dû à l'acide lactique produit) voire engendrent une remontée temporaire du pH (Martin *et al.*, 1997 ; Mora *et al.*, 2004 ; Pernoud *et al.*, 2004 ; Spinnler et Corrieu, 1989 ; Tinson *et al.*, 1982). L'influence de l'activité uréasique sur la vitesse d'acidification a été constatée lors de fabrication du yaourt (Zourari *et al.*, 1991), et de différents types de fromages (Martin *et al.*, 1997 ; Sepulchre et Monnet, 2001). Dans les fabrications fromagères, les opérations technologiques (découpage du caillé, brassage ...) doivent avoir lieu à des valeurs de pH données, mais en pratique, ces opérations sont généralement réalisées à des temps déterminés. De ce fait, les variations d'activité acidifiante dues à l'hydrolyse de l'urée entraînent des défauts et des variabilités importantes dans les fromages (texture, taux d'humidité, affinage). (Martin *et al.*, 1997) ont ainsi observé que les variations des teneurs en urée provoquaient des modifications dans les cinétiques d'acidification et dans la texture des fromages de type reblochon, confirmant les résultats obtenus par Spinnler et Corrieu. (1989). Par ailleurs, comme le lait contient des quantités variables d'urée, la conséquence au cours de la fermentation est imprévisible si les souches de *S. thermophilus* possèdent une activité uréasique.

#### **4. 3. 1. L'uréase**

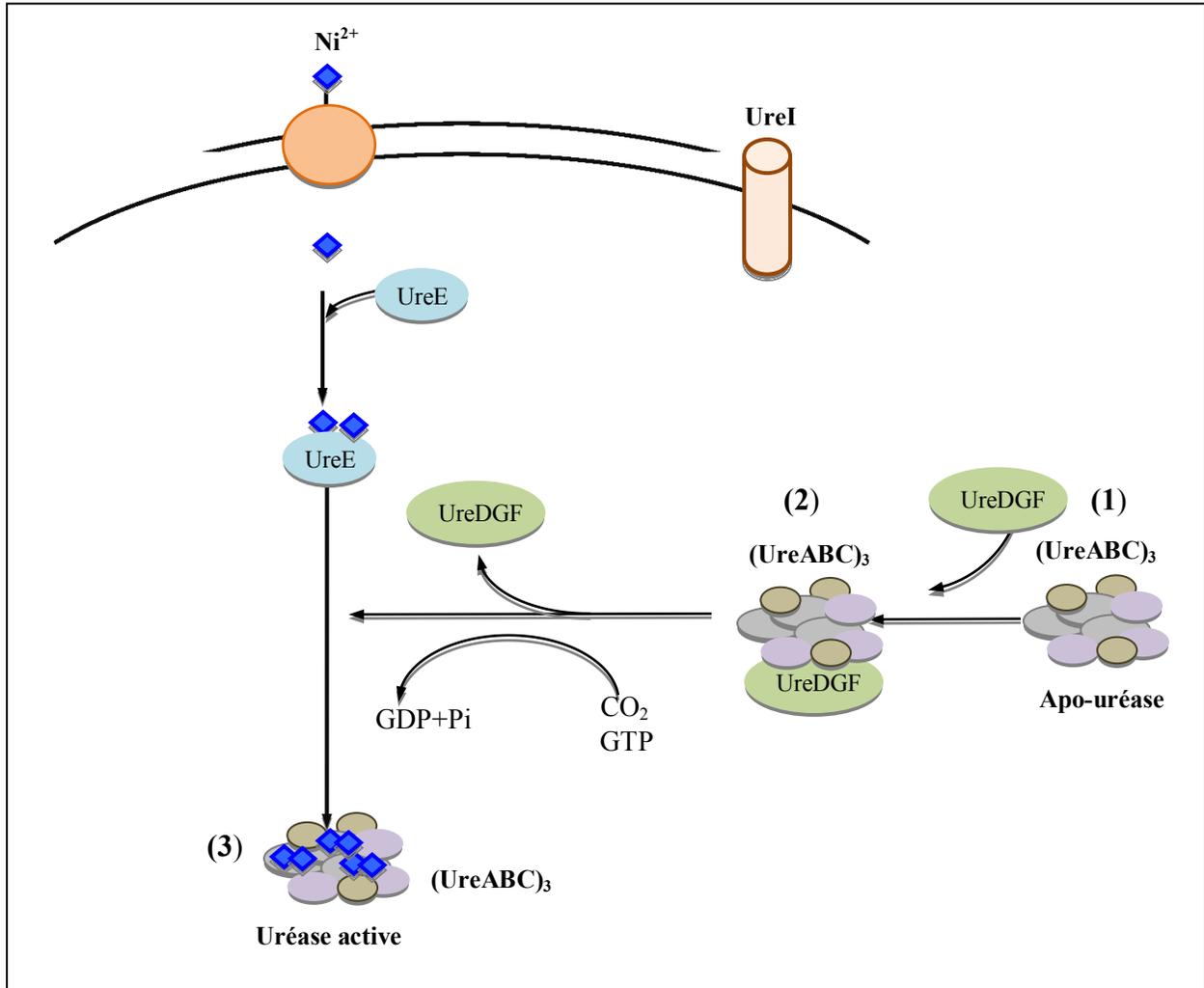
L'uréase est composée de trois sous-unités structurales différentes  $\alpha$ ,  $\beta$  et  $\gamma$  appelées respectivement UreC, UreB et UreA qui s'associent en trimère donnant une enzyme inactive (apo-uréase). La biogénèse de l'enzyme active se réalise grâce au concours de protéines auxiliaires UreD, UreF et UreG. Ces protéines sont nécessaires à l'incorporation de deux ions nickel au site actif de chaque trimère (qui se trouve dans la protéine UreC), qui se fait par une quatrième protéine accessoire, UreE. La synthèse du centre actif de l'uréase est complexe et nécessite la présence de nickel, de dioxyde de carbone et des protéines accessoires, ainsi que

l'hydrolyse du GTP (Benini *et al.*, 1999 ; Mobley *et al.*, 1995 ; Mulrooney et Hausinger, 2003). La figure 17 illustre les différentes étapes clés du processus d'activation cellulaire de l'uréase. L'activité de l'uréase est dépendante du nickel (Dixon *et al.*, 1975). Toutes les uréases identifiées jusqu'à présent contiennent dans leur site actif du nickel qui semble avoir pour rôle la polarisation de la fonction carbonyle de l'urée (Wages *et al.*, 1993). Le nickel qui est un élément nutritionnel indispensable pour la croissance des microorganismes se trouve à l'état de trace dans la nature (Thauer *et al.*, 1980). Par contre, au delà d'une certaine concentration, le nickel devient toxique et inhibe la croissance cellulaire (Wu *et al.*, 1994). De ce fait, le nickel doit pouvoir être transporté en quantité suffisante, mais non toxique, à l'intérieur de la cellule grâce à des transporteurs spécifiques de haute affinité.



**Figure 16 : Réactions de dégradation de l'urée à l'intérieur de la cellule bactérienne (Mobley et Hausinger, 1989)**

- A) Hydrolyse de l'urée en deux molécules d'ammoniaque et d'acide carbonique  
 1. Hydrolyse de l'urée par l'action de l'uréase en carbamate et ammoniaque  
 2. Décomposition spontanée du carbamate en acide carbonique et ammoniaque  
 B) Dissipation des métabolites de l'urée au pH physiologique



**Figure 17 : Modèle d'activation de l'uréease (Mulrooney et Hausinger, 2003)**

- 1) Formation d'une enzyme inactive (apo-uréease) à partir de l'assemblage de trois trimères de UreA, UreB et UreC.
- 2) Fixation de l'apoprotéine à un complexe préformé composé de trois protéines accessoires essentielles : UreD, UreF et UreG.
- 3) Activation de l'uréease par fixation de nickel dans le site catalytique de la sous-unité structurale UreC grâce à la protéine d'accessoire UreE en présence de CO<sub>2</sub> et GTP.

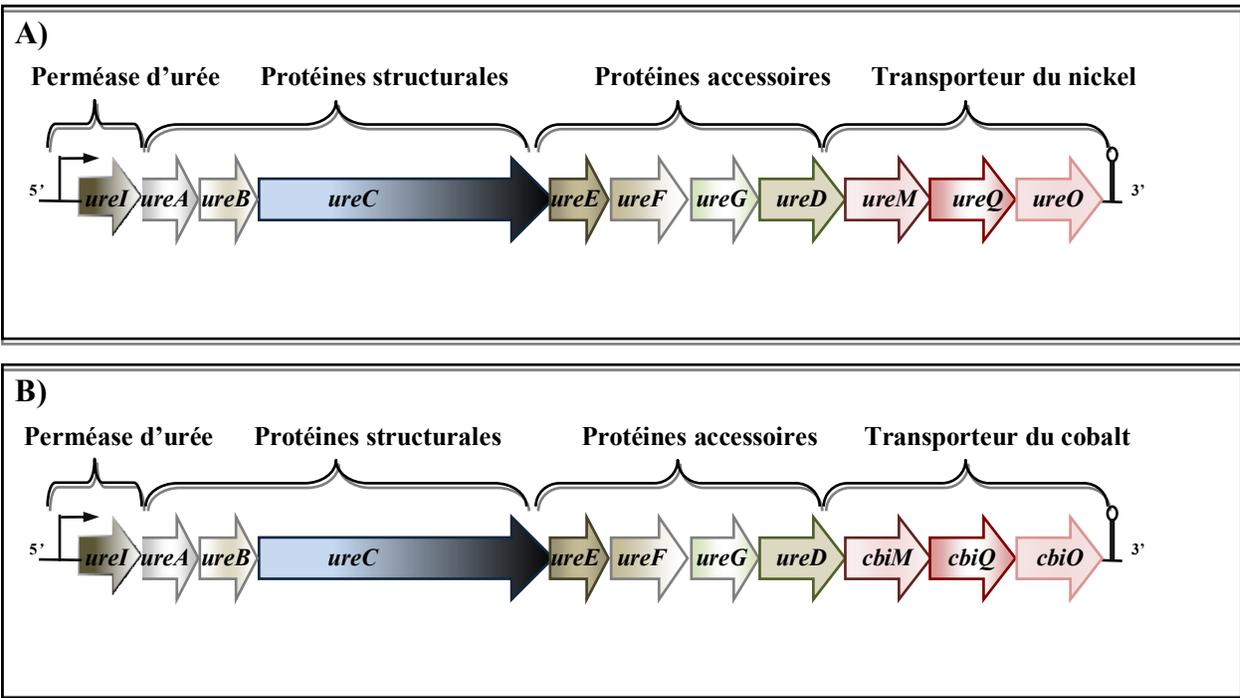
#### 4. 3. 2. Organisation génétique des gènes codant l'uréase

Chez la bactérie lactique *S. thermophilus*, l'opéron uréase présente une organisation génétique similaire à celle de la bactérie orale *S. salivarius* (Chen *et al.*, 1996 ; Chen et Burne, 2003 ; Mora *et al.*, 2004).

Chez *S. salivarius*, 11 ORFs organisées en opéron *ureIABCDEFGHIJMQO*, sont nécessaires pour que l'uréase soit active. Le gène *ureI* code un transporteur membranaire d'urée ; les gènes *ureA*, *ureB* et *ureC* codent les trois sous-unités  $\gamma$ ,  $\beta$  et  $\alpha$  qui constituent les protéines structurales ; les gènes *ureE*, *ureF*, *ureG* et *ureD* codent les protéines accessoires (auxiliaires) et les trois gènes *ureM*, *ureQ* et *ureO* codent trois protéines constituant un transporteur de nickel de type ABC. Les gènes *ureM* et *ureQ* codent deux protéines formant une perméase et le gène *ureO* code une ATPase (Chen *et al.*, 1996 ; Chen et Burne, 2003) (Figure 18A).

Selon Mora *et al.* (2004), chez *S. thermophilus*, l'opéron *ureIABCEFGD* est composé de 8 ORFs et est organisé comme celui de *S. salivarius* (Figure 18B). Les huit gènes de l'opéron *ure* de *S. thermophilus* montrent une très forte identité avec leurs homologues de *S. salivarius*, de 97 à 100% selon le gène.

L'analyse des génomes séquencés de *S. thermophilus* montre la présence de 3 gènes localisés dans l'opéron *ure* et annotés comme des transporteur du cobalt. Il s'agit des gènes *cbiM*, *cbiQ* et *cbiO*. Les protéines *cbiM*, *cbiQ* et *cbiO* présentent de 97 à 99% d'identité avec les protéines *UreM*, *UreQ* et *UreO* de *S. salivarius*. Le nickel, étant un cofacteur essentiel pour l'uréase, les transporteurs de cobalt ayant la capacité de transporter le nickel et enfin, la forte identité qui se trouve entre les protéines chez les deux espèces nous laissent penser que les protéines *CbiM*, *CbiQ* et *CbiO* seraient en fait capables de prendre en charge le nickel nécessaire pour le fonctionnement de l'uréase.



**Figure 18 : Organisation génétique de l'opéron uréasique chez (A) *Streptococcus salivarius* 57.I et (B) *Streptococcus thermophilus* LMD-9**

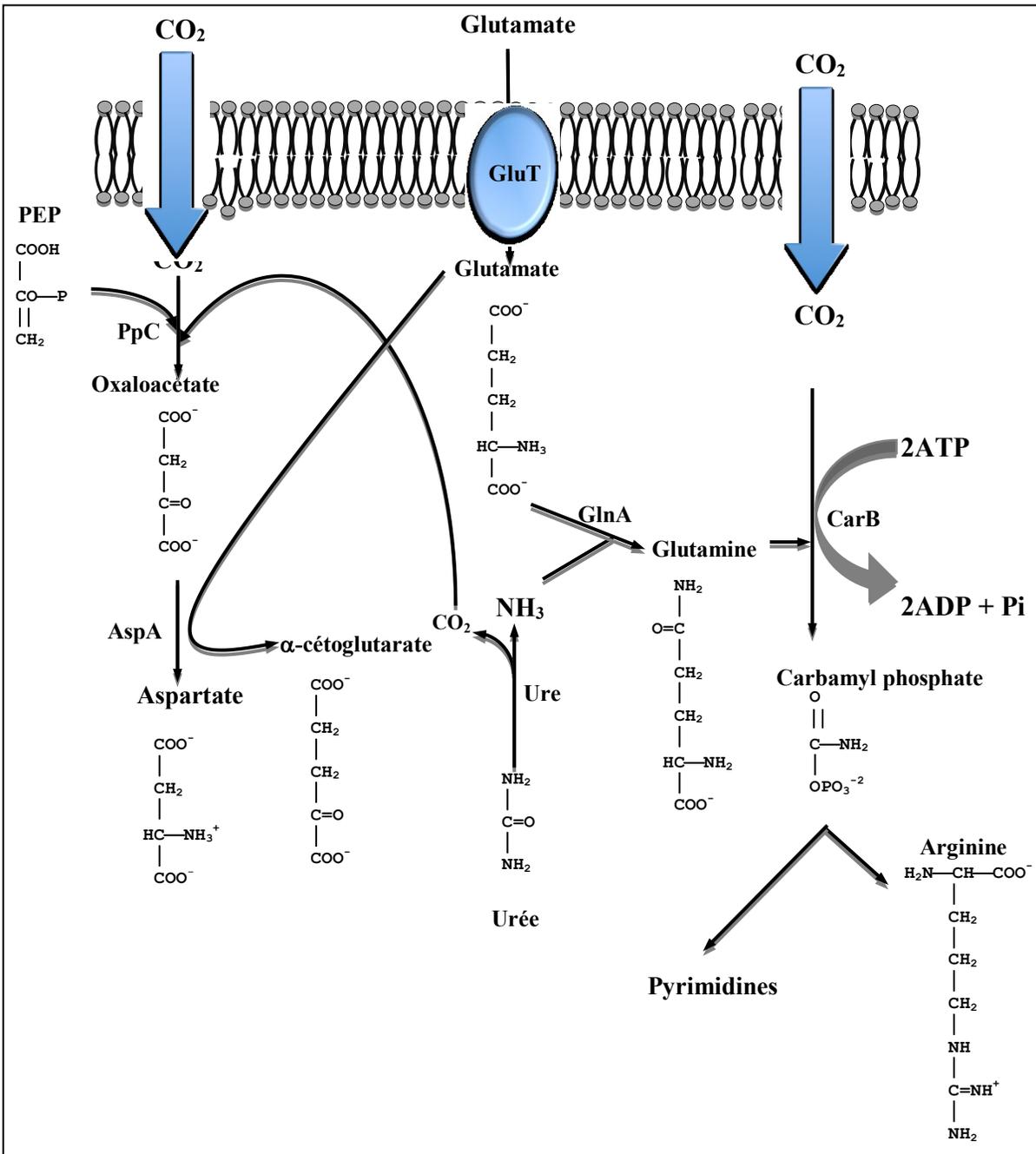
A. Chez *Streptococcus salivarius* l'opéron uréasique comporte 11 ORFs organisées en opéron *ureIABCEFGDMQO* : *ureI* code un transporteur membranaire de l'urée ; *ureA*, *ureB* et *ureC* codent les sous-unités  $\gamma$ ,  $\beta$  et  $\alpha$  qui constituent des protéines structurales ; *ureD*, *ureE*, *ureF* et *ureG* codent des protéines et accessoires ; *ureM*, *ureQ* et *ureO* codent les protéines formant un transporteur de nickel de type ABC (Chen *et al.*, 1996 ; Chen et Burne, 2003).

B. Chez *Streptococcus thermophilus*, 8 ORFs sont organisées en opéron *ureIABCEFGD* (Mora *et al.*, 2004). Les trois gènes situés en aval de gène *ureD* sont annotés comme codant des protéines constituant un transporteur de cobalt de type ABC. Les gène *cbiM* (STER\_0330) et *cbiQ* (STER\_0331) codent chacun une perméase et le gène *cbiO* code une ATPase (STER\_0332). Ils pourraient faire partie d'un même opéron. Les protéines CbiM, CbiQ et CbiO de *Streptococcus thermophilus* LMD-9 possèdent des pourcentages d'identité de 99, 97 et 97% avec les protéines UreM, UreQ et UreO de *Streptococcus salivarius* 57.I respectivement.

### 4. 3. 3. Importance de l'uréase

Chez *S. salivarius*, l'uréase a deux rôles principaux : elle permet à la bactérie de résister à des pH acides et d'utiliser l'urée comme source d'azote (Chen *et al.*, 2000). Malgré la forte identité entre l'uréase de *S. salivarius* et celle de *S. thermophilus*. L'uréase de *S. thermophilus* ne semble pas fournir un avantage sélectif contre l'acidité lors de la fermentation car l'activité uréasique se déclenche à des pH élevés (pH 5,5 et 6) (Mora *et al.*, 2005b ; Zotta *et al.*, 2008). En plus de l'effet du métabolisme de l'urée sur la cinétique d'acidification du milieu extracellulaire lors de la fermentation, Pernoud *et al.* (2004) ont montré que (selon les souches) le métabolisme de l'urée a aussi des conséquences sur la croissance et la production d'acide lactique. La consommation de l'urée augmente la production d'acide lactique, qui peut être due à la stimulation de la croissance par utilisation de métabolites de l'urée (dioxyde de carbone et ammoniacque) dans la synthèse cellulaire. De plus, les métabolites issus de la catalyse de l'urée peuvent être utilisés dans la biosynthèse des acides aminés et des acides nucléiques. L'ammoniacque est utilisée pour la synthèse de la glutamine *via* la glutamine synthétase GlnA (Figure 19), et le dioxyde de carbone est impliqué dans la biosynthèse des bases pyrimidiques et d'acides aminés tel que l'aspartate. (Arioli *et al.*, 2007, 2009 ; Monnet *et al.*, 2005).

Le dioxyde de carbone issu de l'hydrolyse de l'urée, est utilisé par la phosphoenolpyruvate carboxylase (PpC) qui joue un rôle fondamental dans la croissance de *S. thermophilus* en lait. PpC catalyse la réaction entre le phosphoenolpyruvate (issu de la voie de la glycolyse) et le dioxyde de carbone (issu de l'hydrolyse de l'urée ou importé de l'extérieur de la cellule) pour obtenir l'oxaloacétate, qui est ensuite converti en aspartate par l'aspartate aminotransférase AspA (Arioli *et al.*, 2007). Un mutant *ppC* présente une croissance très faible dans le lait et n'est plus capable de coaguler le lait.



**Figure 19 : Représentation schématique de l'activité enzymatique impliquée dans la biosynthèse de l'aspartate et de la glutamine à partir des catabolites de l'urée chez *Streptococcus thermophilus* (d'après Arioli *et al.* 2007, 2009)**

GluT : transporteur membranaire du glutamate ; PpC : phosphoénolpyruvate carboxylase ; AspA : aspartate aminotransférase ; GlnA : glutamine synthétase ; Ure : uréase ; CarB : carbamyl phosphate synthétase ; PEP : . phosphoénolpyruvate

#### 4. 3. 4. Régulation génétique de l'opéron uréase

L'expression de l'opéron uréase varie en fonction de l'espèce. Chez *S. thermophilus* et *S. salivarius* l'expression de l'uréase est régulée par le pH du milieu de culture. Chez *S. thermophilus*, elle est maximale à pH 6 et non réprimée à pH neutre (pH inférieur à 6 n'a pas été testé dans cette étude). Chez *S. Salivarius*, elle est maximale à pH 5,5, réprimé à pH neutre (pH 7) ou à des pH acides (< à pH 4,3). De plus, contrairement à *S. thermophilus*, chez *S. Salivarius*, l'opéron uréase semble être est régulé par le niveau de composés carbonés (Chen et Burne, 1996 ; Chen *et al.*, 1996). La différence observée entre les deux espèces peut être expliquée par la nature de leur niche écologique, pour *S. thermophilus* le lait où le lactose n'est pas en quantité limitante, la cavité buccale pour *S. salivarius*, où la disponibilité des glucides est strictement liée à l'apport alimentaire. Chez *S. thermophilus*, comme chez *S. salivarius*, l'expression de l'opéron *ure* n'est pas inductible par l'addition de nickel ou d'urée.



## **II. Objectifs du travail**



## Objectif du travail

L'équipe « Protéolyse et Biofonctionnalités des Protéines et Peptides » s'intéresse aux peptides bioactifs issus des protéines laitières. Ses recherches concernent (i) la production de peptides bioactifs à partir des protéines laitières par *S. thermophilus* et (ii) la compréhension de la biodisponibilité/mode d'action de peptides bioactifs modèles.

La production de peptides bioactifs à partir des caséines par *S. thermophilus* implique de bien connaître le système protéolytique de cette bactérie. L'objectif de ma thèse a été double :

- 1) évaluer la capacité du système protéolytique de *S. thermophilus* à produire des peptides bioactifs à partir des caséines bovines,
- 2) évaluer la variabilité des systèmes de transport de peptides.

Dans ce manuscrit les résultats obtenus sont organisés en 2 chapitres dont le premier a fait l'objet d'une publication parue en 2012. Les résultats du deuxième chapitre n'ayant pas encore été publiés, ils sont présentés *in extenso* et discutés.



## **III. Matériels et Méthodes**



# 1. Souches, milieux et conditions de culture

## 1. 1. Souches bactériennes utilisées et conditions de conservation

Seuls les matériels et méthodes employés pour l'obtention des résultats décrits dans le chapitre 2 seront décrits cette partie.

19 souches de *S. thermophilus* d'origines diverses ont été utilisées dans ce travail. Elle proviennent soit d'isolements réalisés au laboratoire à partir de produits laitiers d'origines différentes (yaourt ou fromages) (P. Bracquart, communication personnelle), soit de la collection CNRZ (Centre National de Recherches Zootechniques, INRA Jouy-en-Josas, France) soit de la collection ATCC (American Type Culture Collection, Manassas, VA, USA) (Tableau 9). Ces souches sont stockées au laboratoire à - 80°C en lait écrémé reconstitué à 10% (m/v). De plus, les trois souches LMG18311, ND03 et JIM8232 dont leur génome est entièrement séquencé sont mentionnées dans le tableau pour comparer leurs génotypes avec les autres souches de collection. Les souches d'*E. coli* utilisées dans ce travail sont conservées en milieu de culture de Luria Bertani (LB) glycérolé (20%, v/v).

## 1. 2. Milieux de culture

Les précultures de *S. thermophilus* ont été réalisées en lait écrémé reconstitué (10%) ou en milieu M17 supplémenté en lactose (20 g.L<sup>-1</sup>). Pour les suivis de croissance en milieu chimiquement défini, des précultures en milieu M17 supplémenté en lactose ont été réalisées de façon à récupérer les cellules en phase exponentielle (pH d'environ 5 à 5,5 et densité optique (D.O.) à 650 nm de 2 à 2,5). Il est à remarquer que dans les deux cas, les précultures ont été incubées à 42°C.

Cinq milieux de culture ont été utilisés pour la croissance de *S. thermophilus* (Tableau 10) : le lait écrémé reconstitué à 10% (m/v) ; le milieu de culture synthétique M17 (Terzaghi et Sandine, 1975) supplémenté en lactose à 20 g.L<sup>-1</sup> ; le milieu YL (Yeast extract-Lactose, Morice *et al.*, 1992) ; le milieu MCD (Milieu Chimiquement Définie) adapté pour *S. thermophilus* (Letort et Juillard, 2001) dans lequel les sources azotées ont été apportées sous forme de bacto-tryptone

(Fluka) ou bacto-casamino acids (Bacto™ Casamino Acids de la société Becton, Dickinson and Company) ; le milieu Belliker a été utilisé pour la préparation de cellules compétentes et leur transformation (Tableau 11).

Les cultures ont été placées à 42°C, sans agitation. Si nécessaire, de l'érythromycine 2 et 5 µg.mL<sup>-1</sup> a été ajoutée au milieu de culture M17.

Les précultures d'*E. coli* ont été incubées pendant une nuit en milieu LB à 37°C sous agitation avant chaque expérience. Deux milieux ont été utilisés pour réaliser les cultures d'*E. coli* (Tableau 12) : le milieu LB (Sambrook *et al.*, 1989) et le milieu SOC (Dower *et al.*, 1988). Les cultures sont incubées à 37°C, sous agitation, avec ou sans érythromycine 150 µg.mL<sup>-1</sup>.

**Tableau 9 : List des souches de *Streptococcus thermophilus* utilisées dans ce travail**

Souche	Origine	Pouvoir acidifiant*	Type d'ITS 16S-23S*	Gène <i>prtS</i> *	Activité caséinolytique *
CNRZ391	Yaourt	Fort	II	+	+
PB385	Yaourt		II	+	+
PB302	Yaourt		II	+	+
4F44	Fromage		I	+	+
LMD-9	Yaourt <sup>b</sup>		II	+	+
PB18	Fromage		II	+	+
ST87	Fromage		II	+	+
ST88	Fromage		II	+	+
ST85	Fromage		II	+	+
CNRZ385N	Yaourt	Moyen	I	+	+
HD17	Yaourt		I	+	-
CNRZ404	Yaourt		I	+	+/-
ST14	Yaourt		I	-	-
CNRZ160	Yaourt		I	-	-
CNRZ445	Fromage		I	+	+/-
ATE11PB6	Yaourt		I	+	-
PB180	Fromage		I	+	-
CNRZ1066	Yaourt <sup>b</sup>		Faible	I	-
CNRZ21	Yaourt	I		-	-
LMG18311 <sup>a</sup>	Yaourt <sup>b</sup>	Non étudié	I	-	Non étudiée
ND03 <sup>a</sup>	Lait <sup>b</sup>		I	-	Non étudiée
JIM8232 <sup>a</sup>	Lait <sup>b</sup>		I	+	Non étudiée

L'origine des souches, ainsi que le pouvoir acidifiant, leur type d'ITS 16S-23S, la présence des gènes *prtS* et l'activité caséinolytique déterminés par Galia *et al.* (2009) sont mentionnés

<sup>a</sup> Les souches LMG18311, ND03 et JIM8232 dont le génome est entièrement séquencé ont été ajoutées dans cette étude, mais leur pouvoir acidifiant et leur activité caséinolytique n'ont pas été déterminés

<sup>b</sup> Source de ces souches selon Makarova *et al.* (2006), Bolotin *et al.* (2004), Sun *et al.* (2010), Delorme *et al.* (2011).

**Tableau 10 : Composition des milieux de culture utilisés pour la croissance de *Streptococcus thermophilus***

Milieux et conditions de stérilisation	Composition		Utilisation
	Constituant	Concentration g.L <sup>-1</sup>	
Lait reconstitué à 10% (m/v) autoclave 110°C - 10 min	Lait en poudre	100	- Conservation des souches à -80°C - Préculture à 42°C - Suivi de pH et acidité Dornic
Milieu Chimiquement Défini sans acide aminé, (Letort et Juillard, 2001) microfiltration 22 µm Additioné de bacto-tryptone ou bacto-casamino acid	Lactose	5	- Détermination de la croissance par suivi du pH et D.O. <sub>650 nm</sub>
	Acetate de sodium	1	
	Citrate d'ammonium	0,6	
	KH <sub>2</sub> PO <sub>4</sub>	3	
	K <sub>2</sub> HPO <sub>4</sub>	2,5	
	CaCl <sub>2</sub>	0,01	
	MgCl <sub>2</sub>	0,16	
	Urée	0,24	
	Pyridoxamine-HCl	0,8.10 <sup>-3</sup>	
	Acide Nicotinic	0,1.10 <sup>-3</sup>	
	Riboflavine	0,05.10 <sup>-3</sup>	
	Panthoténate de calcium	0,1.10 <sup>-3</sup>	
Thiamine-HCl	0,005.10 <sup>-3</sup>		
Acide ascorbique	0,5		
M17 (Terzaghi et Sandine, 1975) autoclave 120°C - 20 min	Bacto-tryptone	2,5	- Extraction d'ADN, ARN - Détermination de la croissance par suivi du pH et D.O. <sub>650 nm</sub>
	Extrait de levure	2,5	
	Sulfate de magnésium	0,25	
	Acide ascorbique	0,5	
	Extrait de viande	3	
	Soyase	5	
	Lactose	20	
	Bacto-peptone	2,5	
Milieu YL* (Morice <i>et al.</i> , 1992) autoclave 120°C - 20 min	β-glycérophosphate de sodium	19	Détermination de la croissance par suivi du pH et D.O. <sub>650 nm</sub>
	Extrait de levure	8	
	Lactose	20	

**Tableau 11 : Milieux et solutions utilisés pour la préparation de cellules compétentes de la souche *Streptococcus thermophilus* LMD-9**

Milieu *	Composition		Utilisation
	Constituant	Concentration g.L <sup>-1</sup>	
Belliker (Fluka)	Elliker (D178)	48,5	Préparation des cellules compétentes et transformation
	Extrait de viande	10	
Belliker solide	Elliker (D178)	48,5	
	Extrait de viande	10	
	Agar	20	
Belliker avec thréonine (Sigma)	Elliker (D178)	48,5	
	Extrait de viande	10	
Belliker-Thr (20 mM)	Thréonine	2,38	
Belliker avec thréonine et Glycine (Roth)	Elliker (D178)	48,5	
	Extrait de viande	10	
	Thréonine	2,38	
Belliker-Thr (20 mM) -Gly (130 mM)	Glycine	10	

Solution	Composition		Utilisation
	Constituant	Concentration g.L <sup>-1</sup>	
Tampon d'électroporation à pH 6	HEPES	1,6675	Préparation des cellules compétentes
	MgCl <sub>2</sub>	0,2025	

\* Tous les milieux cités ci-dessus ont été stérilisés à l'autoclave à 110°C pendant 20 min

**Tableau 12 : Composition des milieux de culture d'*Escherichia coli***

Milieu et conditions de stérilisation	Composition		Utilisation
	Constituant	Concentration g.L <sup>-1</sup>	
SOC (Dower <i>et al.</i> , 1988) autoclave 120°C - 20 min	KCl	0,18	Régénération des cellules transformées après électroporation
	Bacto-tryptone	20	
	Extrait de levure	5	
	NaCl	0,5845	
	D-glucose	3,6	
	MgCl <sub>2</sub>	2,03	
	MgSO <sub>4</sub>	2,46	
LB (Sambrook <i>et al.</i> , 1989) autoclave 120°C - 20 min	Extrait de levure	5	Culture d' <i>Escherichia coli</i>
	Bacto-tryptone	10	
	NaCl	5	

Les milieux de culture gélosés ont été obtenus par l'ajout de 15 g.L<sup>-1</sup> d'agar

## **2. Extraction d'ADN génomique et ARN totaux, réactions de PCR, conditions d'électrophorèse et séquençage**

### **2. 1. Extraction de l'ADN génomique de *Streptococcus thermophilus***

A partir d'une culture de 24 h réalisée dans 50 mL de milieu M17, les cellules bactériennes ont été récupérées par centrifugation 10 min à 3900 x g à 4°C, lavées 2 fois avec 5 mL de tampon TE (Tris-HCl 10 mM, EDTA 8 mM pH 8) et remises en suspension dans 5 mL de TE. Afin d'hydrolyser la paroi des bactéries, 250 µL d'une solution de lysozyme (Roche) 5% (m/v) ont été additionnés. Après 1 h d'incubation à 37°C, 600 µL d'EDTA (0,5 M pH 8) et 50 µL d'une solution de protéinase K (Roche) 2% (m/v) ont été ajoutés. Après 15 min d'incubation à 37°C, 660 µL d'une solution de SDS 10% (m/v) ont été ajoutés et l'ensemble a été incubé à 50°C jusqu'à ce que les cellules soient lysées. Le lysat a ensuite été aliquoté à raison d'1 mL par tube et 350 µL d'acétate de potassium froid (5 M acide acétique, 3 M acétate de potassium) ont été ajoutés à chaque tube. Après 15 min d'incubation sur glace, les surnageants contenant les acides nucléiques ont été récupérés après centrifugation (15 min, 12000 x g à 4°C). Les ARN ont alors été éliminés par addition d'une solution de RNase A (Roche) 1% (m/v). Après 60 min d'incubation à 37°C, un volume d'isopropanol froid permet de précipiter l'ADN génomique. Le culot est alors lavé à l'éthanol 70% (v/v), séché pendant 5 min à 42°C puis dissout dans un volume d'eau ultra pure.

### **2. 2. Extraction des ARN totaux de *Streptococcus thermophilus***

#### **2. 2. 1. Récupération des cellules bactériennes**

Les ARN totaux ont été extraits à partir de cellules bactériennes cultivées en milieu M17 (45 mL) et récupérées par centrifugation pendant 15 min à 3900 x g et à 4°C. Les culots bactériens obtenus ont été directement congelés dans de l'azote liquide, puis conservés à -80°C.

## 2. 2. 2. Extraction et quantification de l'ARN total

Les ARN totaux ont été extraits à l'aide du réactif Trizol (Sigma) (Chomczynski et Sacchi, 1987 ; Chomczynski, 1993)

Le culot bactérien a été resuspendu dans 4 mL de Trizol et incubé pendant 5 min à température ambiante pour permettre la dissociation complète des complexes nucléoprotéiques. L'homogénat a ensuite été aliquoté. A 5 volumes d'homogénat a été ajouté 1 volume de chloroforme. Après homogénéisation, les protéines ont été séparées des acides nucléiques par centrifugation à 12000 x g et à 4°C pendant 15 min. La phase aqueuse a été récupérée et un volume d'isopropanol froid a été ajouté afin de précipiter l'ARN. Après une centrifugation à 12000 x g à 4°C pendant 20 min, le culot contenant l'ARN a été récupéré, lavé à l'éthanol 70% (v/v), séché puis dissout dans 90 µL d'eau ultra pure.

Afin d'éliminer l'ADN génomique présent dans les échantillons, les échantillons ont été traités une première fois à la DNase I. Pour cela, les échantillons ont été placés en présence de 1 µL de DNase I (10 U.µL<sup>-1</sup>, Roche) et de 10 µL de tampon concentré 10X (DNase I recombinant, RNase-free, Roche), ainsi que de 1 µL d'un inhibiteur de RNase (RNaseOUT™, Invitrogen) est incubés pendant 30 min à 37°C.

Après la première étape d'élimination de l'ADN génomique, l'ARN a été purifié. Pour cela, à deux volumes de solution d'ARN, un volume d'acétate de sodium (3 M, pH 6) et 3,5 volumes d'eau ultra pure ont été ajoutés. Cinq volumes de phénol/chloroforme ont ensuite été ajoutés et le mélange a été homogénéisé puis centrifugé pendant 5 min à 12000 x g et 4°C. La phase aqueuse a été récupérée et un volume d'isopropanol froid ajouté. Après une incubation de 10 min à température ambiante, une seconde centrifugation a été effectuée pendant 20 min à 12000 x g et 4°C. Le surnageant a été alors éliminé et le culot d'ARN lavé avec un volume d'éthanol à 70% (v/v), séché, et dissout dans de l'eau ultra pure. Les ARN totaux obtenus ont été quantifiés par mesure de l'absorbance à 260 nm ( $A_{260\text{ nm}}$ ), après dilution adéquate des ARN dans de l'eau ultra pure. La concentration d'ARN (µg.mL<sup>-1</sup>) a été obtenue grâce à la relation suivante :

$$[\text{ARN}] = A_{260\text{ nm}} \times 0,04 \times \text{facteur de dilution}$$

Enfin, 4 µg de l'ARN récupéré ont été traité avec la TURBO DNase (fourni dans le kit TURBO DNA-free™, Ambion), permettant d'assurer l'élimination complète d'ADN génomique. Pour cela, les ARN ont été placés 90 min à 37°C en présence de 3 µL de TURBO DNase et de 2 µL de

tampon concentré 10X. La réaction de 20  $\mu\text{L}$  a été interrompue par ajout de 4  $\mu\text{L}$  de réactif d'inactivation fourni dans le kit. Après 5 min d'incubation à température ambiante, le surnageant a été récupéré par centrifugation à 12000 x g et 4°C pendant 90 s.

### **2. 3. Design des amorces**

Les amorces utilisées dans ce travail ont été dessinées avec le logiciel Primer 3 input (<http://frodo.wi.mit.edu/primer3/>). Les caractéristiques des amorces utilisées (séquences, tailles et positions sur le génome) sont présentées dans les tableaux 13, 14 et 15.

### **2. 4. Conditions d'amplification par PCR**

#### **2. 4. 1. Réactions standard de PCR**

Des réactions de PCR classiques ont été réalisées pour amplifier des fragments d'ADN dont la taille est inférieure à 1 kb. Dans un volume final de 20  $\mu\text{L}$ , 80 ng d'ADN génomique ont été mélangés à 2  $\mu\text{L}$  de tampon PCR 10X (Invitrogen), 0,6  $\mu\text{L}$  de  $\text{MgCl}_2$  (50 mM, Invitrogen), 0,4  $\mu\text{L}$  d'une solution de dNTP (10 mM, BioRad), 1  $\mu\text{L}$  de chacun des oligonucléotides (25  $\mu\text{M}$ , synthétisés à façon par la société Eurogentec), 0,24  $\mu\text{L}$  de la Taq polymérase (5  $\text{U.mL}^{-1}$ , Invitrogen). La réaction se déroule en trois étapes. Une première étape de 4 min à 94°C permet la séparation des brins d'ADN. Ensuite, 30 cycles sont réalisés, chacun étant composé d'une phase de dénaturation de 45 s à 94°C, une phase d'hybridation des amorces de 30 s avec une température qui varie selon les amorces et une phase d'élongation de 45 s à 72°C (ce temps dépend de la taille du produit à amplifier). Les différents couples d'amorces utilisés, les tailles d'amplicon attendus, les températures d'hybridation ainsi que l'objectif de l'utilisation de chaque couple d'amorces sont indiqués dans les tableaux 16 et 17. Les réactions ont été effectuées dans un thermocycler My Cycler (BioRad).

**Tableau 13 : Séquence des amorces utilisées pour l'amplification de séquences d'ADN génomiques par PCR**

Amorces	Séquence du côté 5' au côté 3'	Position des amorces
amiC rev	GTCTGTCCTTGGCCTTGTGT	Séquence interne du gène <i>amiC</i>
amiC for	CGGCTTTGACCTACACGATT	
amiD rev	GTCTGTCCTTGGCCTTGTGT	Séquence interne du gène <i>amiD</i>
amiD for	CGGCTTTGACCTACACGATT	
amiE rev	CGTCCCTGGAATTGAGAAGA	Séquence interne du gène <i>amiE</i>
amiE for	ACCGTGGCAAAGACTTGACT	
amiF rev	TGAAACGGACCACTGACAAA	Séquence interne du gène <i>amiF</i>
amiF for	CGGCGAAGGAAGAAAGAAGT	
amiA3 rev	TCAAAGACTATGATGCCTTG	Séquence interne du gène <i>amiA1/amiA3</i>
amiA2N for	TGATGTCTAGGTTTCGAGTCC	Séquence interne du gène <i>ARNr</i>
amiA2LR rev	AATGCGTTGGTTCGTAAAGG	Séquence interne du gène <i>cfa</i>
dtpT for	TGCTTGGACATATCGTGCTC	Séquence interne du gène <i>dtpT</i>
dtpT rev	AATCGCCCAGAACAAAACAG	Séquence interne du gène <i>dtpT</i>
obp LM for	TGAGTTGACCGCTTCTGATG	Séquence interne du gène <i>otsA</i>
obp LM rev	ACGACCCCATGTTATGTGT	Séquence interne du gène <i>otsD</i>
obp ER for	CAAACATGCATGGAAACCTG	Séquence interne du gène <i>otsB</i>
obp ER rev	GACCCCAAGATAACACGAGA	Séquence interne du gène <i>otsC</i>
obp1 for	ACAGTCAAGTGGCGTCTGTAA	Région promotrice du gène <i>otsA</i>
obp2 rev	GGGAAGTTGAAAAGGAAAGGA	Séquence interne du gène <i>otsA</i>
obp3 for	ACTAAAATTATTATGGGAACAGC	Séquence interne du gène <i>otsA</i>
obp4 rev	AATGGACCTGCAAAGGCTTA	Séquence interne du gène <i>otsA</i>
obp5 for	TAACGTCTCTGGGAGCAAA	Séquence interne du gène <i>otsA</i>
obp6 rev	CTTTTGCTTTATCAACAGATGG	Séquence interne du gène <i>otsA</i>
obp7 for	AAGCAGTCGCTAATTTGGTTGA	Séquence interne du gène <i>otsA</i>
obp8 rev	CGATTTGCTTCCCTTCAAGTA	Séquence interne du gène <i>otsA</i>
obp9 for	AGTGGGCTCGAGTTAAGAA	Séquence interne du gène <i>otsA</i>
obp10 rev	ACCGATGACCATCGCTAAAAT	Séquence interne du gène <i>otsB</i>
obp11 for	TGACTGGCGATCCAATTTCC	Séquence interne du gène <i>otsB</i>
obp12 rev	AAAGTTGAAATCAAATTCGTTTCG	Séquence interne du gène <i>otsB</i>
obp13 for	CCATCAAATGGTCGTGTCTC	Séquence interne du gène <i>otsB</i>
obp14 rev	TAATCAAACCGGTCAGCAACAT	Séquence interne du gène <i>otsC</i>
obp15 for	CGAGATTTCAATGCCGTACA	Séquence interne du gène <i>otsB</i>
obp16 rev	TGAGGGACAAAACCACGGT	Séquence interne du gène <i>otsC</i>
obp17 for	ACTATTGCTCGCGATTGTCA	Séquence interne du gène <i>otsC</i>
obp18 rev	TTGATCAATCGTTAAACAAGCC	Séquence interne du gène <i>otsD</i>
obp19 for	ATGACTTAGCAGATCCTCG	Séquence interne du gène <i>otsC</i>
obp20 rev	GACAATCCGTTGCTTTTGA	Séquence interne du gène <i>otsD</i>
obp21 for	GGCCGTAAGTTGACAAATGC	Séquence interne du gène <i>otsD</i>
obp22 rev	AGCACCTTACCGTAGACTTTT	Séquence interne du gène <i>otsD</i>
obp23 for	TTTGAAATTTCGAAAGGAGAAG	Séquence interne du gène <i>otsD</i>
obp24 rev	TTCACCATGATATAAGACCGAA	Séquence interne du gène <i>otsD</i>
obp25 for	TGTCCAAAAGAAGTTTTGAAAT	Séquence interne du gène <i>otsD</i>
obp26 rev	AAAAACATGCTCACCTTAAAAAC	Région terminatrice du gène <i>otsD</i>
RT-obp for	GCAAATCGCTGTTTCAGGA	Séquence interne du gène <i>otsA</i>
RT-per rev	AAATCGCACCAAACAGTATCG	Séquence interne du gène <i>otsB</i>
RT-ATPase rev	CATCAAAATCGCCTTGGTCT	Séquence interne du gène <i>otsD</i>

**Tableau 14 : Séquence des amorces utilisées pour l'amplification de séquences d'ADN par PCR**

Amorces	Séquence du côté 5' au côté 3'	Position d'amorce
obp I for	<i>TAAGTTCTAGAATTC</i> <sup>a</sup> TTATGGGAACAGCATTATGTACT	+ 23 <sup>d</sup>
obp I rev	<i>ACTTAAGGGGATCC</i> <sup>b</sup> ATGGACCTGCAAAGGCTTA	+ 609 <sup>d</sup>
obp II for	<i>CGTATAGATGGATCC</i> <sup>b</sup> TTCAATGTCAATGCCCGTTA	+ 2026 <sup>d</sup>
obp II rev	<i>ATTTAATATCTCGAG</i> <sup>c</sup> TCGTCTCGGTTAGAATTGCAC	+ 2639 <sup>d</sup>
pGhost9 for	CCATTTTGAACGATGACCTC	plasmide pGhost9
pGhost9 rev	GACAGCTTCCAAGGAGCTAA	plasmide pGhost9
sigma for	CCAAATGGCTGTCATCCTCT	+ 826 <sup>d</sup>
sigma rev	GACCCAACACCTGAGCAAAT	+ 69 <sup>d</sup>
amiFquantfor	CTTCCACGATTGTCCTTAT	+ 694 <sup>d</sup>
amiFquantrev	TTCGTGCCCAAGTTTTGAATTT	+ 553 <sup>d</sup>
dtptquantfor	AACGTTCTGTTGCCGTAACC	+ 838 <sup>d</sup>
dtptquantrev	CGTTGTGATGAACCTTGTCG	+ 704 <sup>d</sup>
otsAquantfor	CATGAGTTGACCGCTTCTGA	+ 340 <sup>d</sup>
otsAquantrev	TTTGCTTCGACTGATTCCAT	+ 445 <sup>d</sup>

<sup>a, b, c</sup> Séquences soulignées indiquent respectivement les sites de restriction des enzymes *EcoRI*, *BamHI*, *XhoI*.

Les séquences en italique indiquent les nucléotides ajoutés aux amorces.

<sup>d</sup> Positions des amorces par rapport à l'ATG

**Tableau 15 : Amorces utilisées pour la réalisation d'amplification de séquences d'ADN génomiques par PCR pour faire MLST**

Amorce	Séquence du côté 5' au côté 3'	Position des amorces
glcK for	TGGGCAGAAACTCAAGA <b>A</b>	Séquence interne du gène <i>glcK</i> (glucose kinase)
glcK rev	AACACCACCACCGATAAC	
ddlA for	TCAAGTGT <b>A</b> GCTATGGAC <b>C</b>	Séquence interne du gène <i>ddlA</i> (D-alanine D-alanine ligase)
ddlA rev	GTAGATGGCTCCATCCTC	
pepO for	AACTCTACCACCC <b>A</b> TATGA	Séquence interne du gène <i>pepO</i> (endopeptidase)
pepO rev	GGTTT <b>A</b> TTCCACTTGCTCCAG	
ilvC for	<b>ATA</b> CAGGTCACGATGTTAT	Séquence interne du gène <i>ilvC</i> (ketol-acid reductoisomerase)
ilvC rev	GGTGCATATCCAGCTTCAGT	
thrS for	ATCACTGAAGATGGAAGC	Séquence interne du gène <i>thrS</i> (threonyl-tRNA synthétase)
thrS rev	CCAAGTTTACGGTGGTCA	
tkt for	CTGGCCACCCTGGTGTGG	Séquence interne du gène <i>tkt</i> (transketolase)
tkt rev	GGAGCACCGTGAACACCG	
pyrE for	CGCTTTACGGAGGAACAT	Séquence interne du gène <i>pyrE</i> (orotate phosphoribosyltransferase)
pyrE rev	GTCCGTCTGCAGTGATGT	
dnaE for	GGACTGGGAGCCTGGGAT	Séquence interne du gène <i>dnaE</i> (ADN polymérase III)
dnaE rev	ACTCCCTGCAGCAGACCC	

Les nucléotides surlignés en vert correspondent aux nucléotides ajoutés à l'amorce, en bleu aux nucléotides modifiés par rapport aux amorces utilisées par Delorme *et al.* (2007, 2010).

**Tableau 16 : Couples d'amorces utilisés dans les réactions de PCR**

Couple d'amorce	Taille d'amplicon (pb)	Température d'hybridation (°C)	Utilisation
amiC rev/amiC for	615	52	Présence/absence du gène <i>amiC</i>
amiD rev/amiD for	601	52	Présence/absence du gène <i>amiD</i>
amiE rev/amiE for	614	52	Présence/absence du gène <i>amiE</i>
amiF rev/amiF for	605	52	Présence/absence du gène <i>amiF</i>
amiC rev/ amiA3 rev	4700 ou 1100	52	Présence/absence des gènes <i>amiA1</i> et <i>amiA3</i>
amiA2N for/amiA2LR rev	5760 ou 1690	50	Présence/absence du gène <i>amiA2</i>
dtpT for/dtpT rev	608	52	Présence/absence du gène <i>dtpT</i>
obp LM for/obp LM rev	4129	56	Présence/absence du cluster <i>ots</i>
obp ER for/obp ER rev	536	52	
obp1 for/obp2 rev	667	58	Séquençage du cluster <i>ots</i>
obp3 for/obp4 rev	617	58	
obp5 for/obp6 rev	623	58	
obp7 for/obp8 rev	612	58	
obp9 for/obp10 rev	642	58	
obp11 for/obp12 rev	610	58	
obp13 for/obp14 rev	633	58	
obp15 for/obp16 rev	643	58	
obp17 for/obp18 rev	663	58	
obp19 for/obp20 rev	625	58	
obp21 for/obp22 rev	699	58	
obp23 for/obp24 rev	612	58	
obp25 for/obp26 rev	579	58	
obp8 rev/obp9 for	91	58	
obp12 rev/obp13 for	212	58	
obp16 rev/obp17 for	108	58	
obp22 rev/obp23 for	128	58	
RT-obp for/RT-per rev	250	56	
obp15 for/obp14 rev	216	58	
obp19 for/RT-ATPase rev	170	58	
sigma for/ sigma rev	130	58	Construction du mutant $\Delta$ <i>ots</i>
obpI for/obpI rev	633	58	
obpII for/obpII rev	640	58	
pGhost9 for/obp13 for	895	56	
obp4 rev/pGhost9 rev	710	56	
obp12 rev/obp LM for	745	56	

La taille des amplicons attendus, la température d'hybridation et leur objectif d'utilisation sont indiqués.

**Tableau 17 : Couples d’amorces utilisés et taille des amplicons attendus pour l’amplification par PCR des séquences de « gènes de ménage » pour le typage par MLST**

Couple d’amorce	Taille d’amplicon (pb)	Temperature d’hybridation (°C)	Utilisation
glcK for/glcK rev	540	50	MLST
ddlA for/ddlA rev	531	50	
pepO for/pepO rev	643	50	
ilvC for/ilvC rev	547	50	
thrS for/thrS rev	590	50	
tkfA for/tkfA rev	721	50	
pyrE for/pyrE rev	602	50	
dnaE for/dnaE rev	556	50	

#### 2.4.2. Réaction de Long-Range PCR

La Long-Range PCR a été utilisée pour amplifier des fragments d’ADN de taille supérieure à 1,5 kb. Le protocole expérimental de cette réaction consiste à mélanger 2 µL de tampon PCR 10X (Sigma), 0,4 µL de la solution de dNTP (10 mM, BioRad), 0,12 µL de chacun des oligonucléotides (100 µM, Eurogentec) et 1 unité de l’accu Taq (Sigma) à 80 ng d’ADN cible. De l’eau ultra pure est ajoutée pour compléter le volume final à 20 µL. L’appareil est programmé de la manière suivante : la première étape de dénaturation s’effectue pendant 5 min à 95°C, suivie de 32 cycles, chacun étant composé d’une phase de 45 s à 95°C, une phase de 30 s à la température d’hybridation (variable en fonction des amorces), une dernière phase à 68°C dont le temps varie en fonction de la taille du produit à amplifier. L’étape d’élongation finale se déroule à 68°C pendant 10 min. Les différents couples d’amorces utilisés, les tailles d’amplicon attendus, les températures d’hybridation ainsi que l’objectif de l’utilisation de chaque couple sont indiqués dans le tableau 16. Les réactions ont été effectuées dans un thermocycler My Cycler (BioRad).

## **2. 5. PCR en temps réel**

### **2. 5. 1. Préparation des ADN complémentaires**

Les ARN ont été convertis en ADN complémentaires en utilisant la « Reverse Transcriptase - PCR » (RT-PCR). Les ARN ont été dénaturés 5 min à 65°C en présence de 5 µL de Random primers (200 µg.mL<sup>-1</sup>), 2,5 µL de dNTP (10 mM) (BioRad) et 7,5 µL d'eau stérile. Après refroidissement rapide dans la glace, 2,5 µL d'un inhibiteur de RNase (RNaseOUT™, Invitrogen), 10 µL de tampon 5X (Invitrogen) et 5 µL de DTT 0,1 M (Invitrogen) ont été ajoutés et le mélange a été incubé 2 min à 37°C. Après ajout de 2,5 µL de Reverse Transcriptase M-MLV (Invitrogen), 10 min à 25°C puis 50 min à 37°C sont nécessaires pour réaliser la rétrotranscription. La rétrotranscription a été stoppée par incubation à 70°C pendant 15 min et les ADNc ont été conservés à - 20°C. Les ADNc ainsi obtenus ont ensuite été utilisés comme matrice dans la recherche de transcrit spécifique des gènes étudiés par PCR.

### **2. 5. 2. Quantification des niveaux d'expression**

La PCR en temps réel a été réalisée au moyen d'un thermocycleur couplé à un module de détection de type MJ Opticon Monitor™ (Biorad). Les réactions ont été effectuées sur des plaques 96 puits, chaque puits contenant 7,5 µL de MESA GREEN SYBR Assay No ROX (Eurogentec), 1,5 µL de chaque amorce (à 1 µM chacune), 25 ng d'ADNc et 2,5 µL d'eau ultra pure. Après avoir été recouverte, la plaque a été centrifugée 30 s à 5000 g. Le cycle de PCR utilisé comporte une étape de dénaturation primaire à 95°C de 5 min, suivie de 40 cycles d'amplification (95°C pendant 15 s et 58°C pendant 45 s).

Le principe de la PCR quantitative en temps réel est basé sur la détection et la quantification d'un indicateur fluorescent (dans cette étude, il s'agit du Sybr-Green) de la production d'amplicons durant chaque cycle d'amplification. L'augmentation du signal fluorescent est directement proportionnelle à la quantité d'amplicons générés durant la réaction de PCR. Ainsi, grâce à un système de détection de la fluorescence, la courbe d'amplification de l'ADNc au cours du temps peut être tracée. Le seuil de détection de la fluorescence n'est atteint qu'après un certain nombre

de cycles d'amplification. Le premier cycle de démarrage de la phase d'amplification correspond au cycle seuil (Threshold Cycle) définie par la valeur Ct. Ce nombre-seuil est en fonction de la quantité initiale de séquences d'ADNc cibles présentes dans la réaction de PCR. En fait, plus la valeur Ct est élevée, plus la quantité initiale d'ADNc présents dans la réaction de PCR est faible. Pour valider le Ct de chaque échantillon, la réaction est réalisée en triplicat et la moyenne des trois valeurs de Ct est utilisée.

Dans cette étude, pour normaliser la quantité d'ADNc présents dans chaque échantillon, le niveau de transcription des gènes cibles a été comparé à celui d'un gène de référence : le gène codant le facteur sigma  $\sigma_{70}/\sigma_{32}$ , qui contrôle les fonctions de ménage essentielles et assure ainsi la transcription de base chez les bactéries. Ce gène est donc supposé être exprimé à toutes les phases de croissance à des niveaux similaires entre les deux souches étudiées en lait et en milieu M17.

La technologie de la PCR quantitative passe par un traitement mathématique des données. La quantification des niveaux d'expression a été effectuée en utilisant la méthode  $\Delta\Delta Ct$  (Livak et Schmittgen, 2001), méthode utilisée de façon habituelle pour déterminer des niveaux d'expression en RT-PCRq. Pour chaque condition analysée (X et X0), la différence de Ct obtenue entre le gène d'intérêt et le gène de référence donne les valeurs  $\Delta Ct_X$  et  $\Delta Ct_{X0}$ . Le  $\Delta\Delta Ct$  est obtenu par la différence  $\Delta Ct_X - \Delta Ct_{X0}$ . Ainsi, trois types de calcul ont été réalisés :

$$\Delta\Delta Ct = (Ct_{\text{cible}} - Ct_{\text{référence}})_{\text{souche X}} - (Ct_{\text{cible}} - Ct_{\text{référence}})_{\text{souche X0}}$$

$$\Delta\Delta Ct = (Ct_{\text{cible}} - Ct_{\text{référence}})_{\text{Phase X}} - (Ct_{\text{cible}} - Ct_{\text{référence}})_{\text{Phase X0}}$$

Ces relations, ont permis de comparer le niveau de transcription de chaque gène entre les deux souches étudiées (stade de croissance et milieu de culture identique), entre chaque phase de croissance (pour une même souche et un même milieu).

Pour obtenir la quantification, le calcul est basé sur le fait que la quantité d'ADN doit doubler à chaque nouveau cycle de PCR, soit après N cycles une quantité d'ADN  $2^N$  fois dans le cas où l'efficacité de PCR est de 100%.

Le niveau d'expression est ainsi donné par la relation :

$$\text{Niveau expression} = 2^{-\Delta\Delta Ct}$$

## 2. 6. Electrophorèse d'ADN

Un volume d'échantillon à analyser a été additionné à 0,2 volume de tampon de dépôt (composé de 0,25% (m/v) de bleu de bromophénol, de 70% (m/v) de saccharose et de 0,1 M d'EDTA à pH 8) et soumis à une migration électrophorétique horizontale dans du tampon TAE (Tris acétate 40 mM, EDTA 1 mM). Le pourcentage en agarose des gels a été choisi en fonction de la taille des fragments d'ADN à séparer. Les fragments d'ADN ont été ensuite visualisés, après révélation des gels dans une solution de bromure d'éthidium à 0,5  $\mu\text{g}\cdot\text{mL}^{-1}$ , sous UV à 254 nm grâce à l'utilisation d'un appareil GelDoc (BioRad) et du logiciel d'analyse d'images QuantityOne. Les marqueurs de taille utilisés sont les EZ-Load 100 pb et 1 kb de la société BioRad.

## 2. 7. Séquençage

Le séquençage des fragments d'ADN génomiques amplifiés par PCR a été effectué par la société GENOME EXPRESS. Après évaluation de la spécificité d'amplification par électrophorèse sur gel d'agarose, les mélanges réactionnels de PCR ont été directement envoyés à la société GENOME EXPRESS qui a réalisé une élimination de l'ADN simple brin par action d'une exonucléase spécifique et un séquençage selon la méthode de Sanger (Sanger *et al.*, 1977). Les séquences nucléotidiques obtenues et les séquences protéiques déduites ont été analysées par les programmes « BlastN » et « BlastP » afin de rechercher les pourcentages d'identité avec différentes séquences présentes dans les banques de données, *via* le site NCBI (<http://www.ncbi.nih.gov/>). Les alignements multiples ont été réalisés à l'aide du programme « ClustalW ».

## 3. MLST

Dans cette étude les 8 « gènes de ménage » : *glcK*, *ddlA*, *pepO*, *ilvC*, *thrS*, *tkl*, *pyrE* et *dnaE* utilisés par Delorme *et al.* (2010) pour estimer la diversité génétique des souches de *S. thermophilus* ont été utilisés. De 531 à 721 pb de chaque gène ont été amplifiés par PCR chez 17 souches de notre collection en utilisant les mêmes oligonucléotides que ceux utilisés par

Delorme *et al.* (2007, 2010). Cependant, notre étude étant limitée à l'espèce *thermophilus*, un nucléotide a été rajouté dans les amorces « *glcK for* » et « *ddlA for* » du côté 3', un nucléotide a été modifié dans les amorces « *ddlA for* », « *pepO for* », « *pepO rev* » et trois nucléotides dans l'amorce « *ilvC for* » pour correspondre exactement aux séquences des gènes de *S. thermophilus* référencés dans les banques des données (Tableaux 15 et 17). Les amplicons ont ensuite été séquencés. Pour chaque gène, les séquences obtenues ont été alignées à l'aide du logiciel « ClustalW » et les numéros d'allèles attribués par Delorme *et al.* (2010) assignés à chaque séquence correspondante. Dans le cas de la découverte d'un nouvel allèle, un nouveau numéro a été attribué. En outre, les séquences des «gènes de ménage» des cinq souches de *S. thermophilus* dont le génome est entièrement séquencé (LMD-9, LMG18311, CNRZ1066, ND03 et JIM8232) ont été extraites des banques de données et parfois corrigées (gènes *glcK* des souches LMD-9 et CNRZ1066) selon Delorme (communication personnelle). Pour construire l'arbre phylogénétique, les séquences de 8 «gènes de ménage» et la séquence intergénique ITS 16S-23S ont été concaténées. Afin d'enraciner l'arbre phylogénétique, les séquences de 8 «gènes de ménage» (identiques à ceux utilisés pour *S. thermophilus*) et la séquence intergénique ITS 16S-23S de trois souches de *S. salivarius* (JIM8777, CCHSS3, 57.I) ont été concaténées. Ensuite, la méthode de « Neighbor joining » a été appliquée avec le logiciel MEGA\_5.0. Les valeurs de « bootstrap » ont été calculées avec 1000 répliques. Le modèle p-distance a été utilisé pour estimer la distance de séquences nucléotidiques.

#### **4. Construction du mutant LMD-9 $\Delta$ ots**

Un mutant LMD-9 $\Delta$ ots a été construit par double recombinaison homologue afin d'effectuer la délétion d'un fragment des gènes *otsA* et *otsB* sans insertion de gène de résistance à un antibiotique. La Figure 20 résume les différentes étapes utilisées pour la construction de ce mutant.

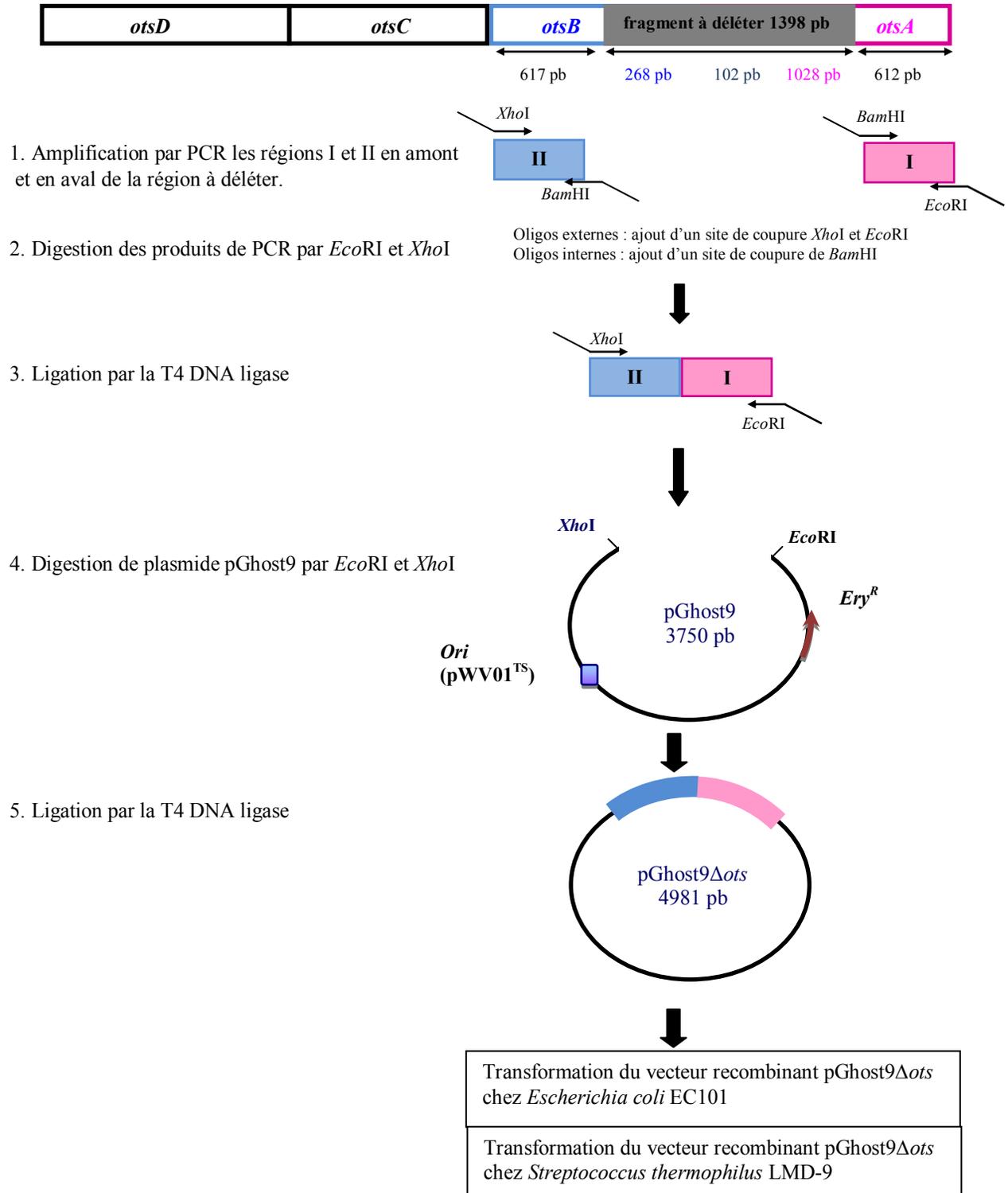
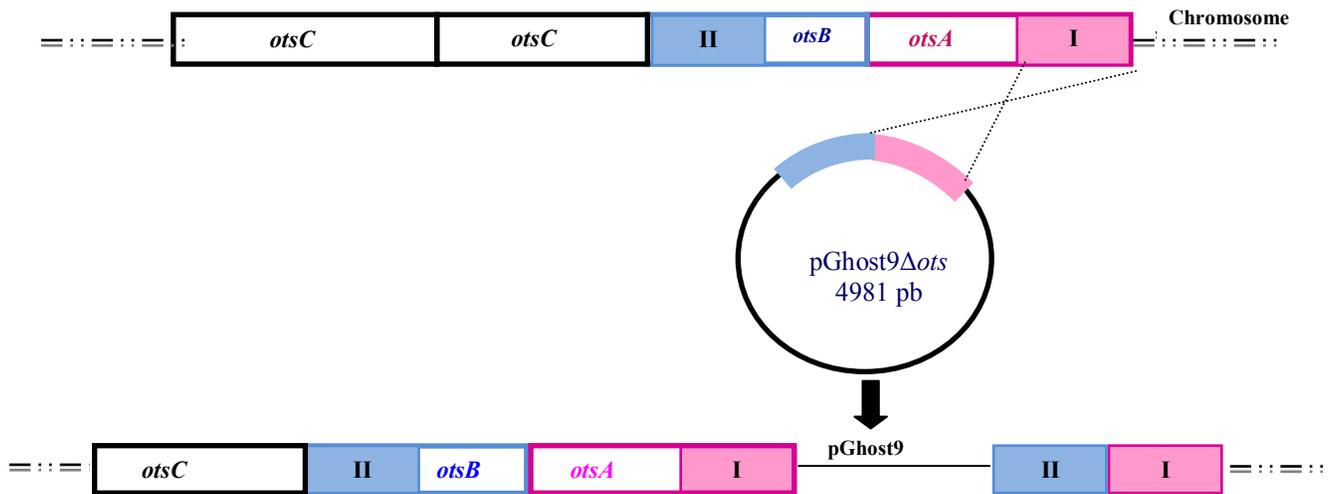
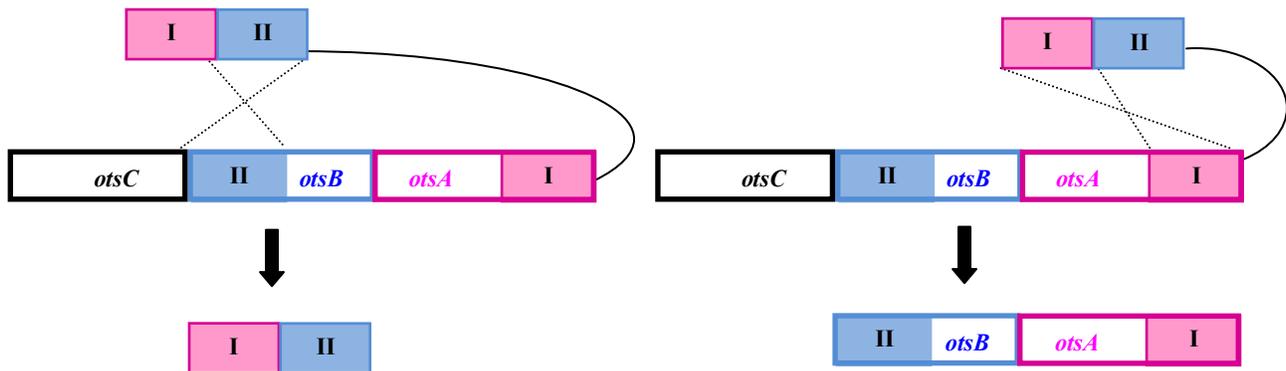


Figure 20 : 1/2

### A. Premier événement de recombinaison homologue



### B. Deuxième événement de recombinaison homologue



Remplacement des deux gènes *otsA* et *otsB* sauvages par deux gènes mutés tronqués

Retour au génotype sauvage

Figure 20 : 2/2

**Figure 20 : Représentation schématique de la stratégie utilisée pour déléter les ORFs des gènes *otsA* et *otsB* de *Streptococcus thermophilus* LMD-9 à l'aide du plasmide pGhost9**

Cette construction est basée sur la stratégie décrite par Biswas *et al.* (1993) et Maguin *et al.* (1996).

*Ery<sup>R</sup>*: gène conférant une résistance à l'érythromycine ; *ori* (pWV01<sup>TS</sup>) : origine de réplication active chez *Streptococcus thermophilus*

## 4. 1. Préparation d'un plasmide recombinant pGhost9 $\Delta$ ots

### 4. 1. 1. Amplification des deux inserts

Les régions I (633 pb) et II (640 pb) qui se trouvent respectivement en amont et en aval de la partie des gènes à déléter (le fragment à déléter est composé de 1398 pb) ont été amplifiées par PCR (Figure 20). Pour cela, deux couples d'amorces ont été dessinés à partir de la séquence de la souche LMD-9. Des sites de restriction *Bam*HI, *Xho*I, *Eco*RI et *Bam*HI ont été ajoutés respectivement à l'extrémité 5' de chaque oligonucléotide : « obpII for », « obpII rev », « obpI for » et « obpI rev » (Tableau 14 et 16). Dans un premier temps, les deux inserts I et II ont été amplifiés par PCR, selon le programme suivant : 3 min à 94°C puis 30 cycles de 45 s à 94°C, 30 s à 58°C et 45 s à 72°C, suivis d'une élongation de 10 min à 72°C. Enfin, les amplicons obtenus ont été purifiés par un kit de purification de la société Roche en suivant les recommandations du fournisseur.

### 4. 1. 2. Digestion enzymatique des amplicons et du plasmide pGhost9

Les produits de PCR purifiés ont été digérés avec les enzymes de restriction : *Bam*HI et *Eco*RI pour l'insert I et *Bam*HI et *Xho*I pour l'insert II avec 2 U. $\mu$ g<sup>-1</sup> d'ADN amplifié pendant 2 h à 37°C. Les enzymes de restriction ont été fournies par la société Roche.

Le plasmide pGhost9 (3752 pb) a été utilisé dans cette étude comme un vecteur de clonage. Il s'agit d'un plasmide thermosensible, possédant une origine de répllication thermosensible et une cassette de résistance à l'érythromycine. Il est capable de se répliquer à 37°C chez *E. coli* EC101 (Tableau 18) et à 30°C chez *S. thermophilus* mais il est incapable de se répliquer à 42°C chez ce dernier hôte (Biswas *et al.*, 1993 ; Maguin *et al.*, 1992). Parallèlement, le plasmide pGhost9 a été digéré par les enzymes *Eco*RI et *Xho*I pendant 2 h à 37°C avec 5 U. $\mu$ g<sup>-1</sup> d'ADN plasmidique, puis le plasmide a été déphosphorylé pendant 1 h à 37°C par l'enzyme CIAP « Alkaline Phosphatase, Calf Intestinal » (New England Biolabs) à raison de 5 U. $\mu$ g<sup>-1</sup> d'ADN plasmidique. Enfin, les inserts et le plasmide ont été purifiés à l'aide du kit Roche.

### 4. 1. 3. Ligation des deux inserts avec le plasmide

Les deux produits de PCR et le plasmide pGhost9 (préalablement digérés et purifiés) ont été ligaturés par incubation à 16°C pendant une nuit avec la T4 DNA ligase (New England Biolabs). La ligation a été faite dans un volume réactionnel de 20 µL final par 4 unités de ligase avec un rapport plasmide/insert de 1/3.

**Tableau 18 : Génotype de la souche *Escherichia coli* EC101 utilisée pour la construction du mutant LMD-9Δots**

Souche	Génotype	Référence
<i>Escherichia coli</i> EC101	<i>supE hsd-5 thi Δ(lac-proAB) F<sup>+</sup> (traD6 proAB<sup>+</sup> lacI<sup>f</sup> lacZΔM15) repA<sup>+</sup></i>	(Leenhouts, 1995)

## 4. 2. Transformation d'*Escherichia coli* EC101 et sélection des transformants

### 4. 2. 1. Préparation des cellules électro-compétentes d'*Escherichia coli* EC101

La souche *E. coli* EC101 a été utilisée pour le clonage (Tableau 18). Les cellules compétentes ont été préparées selon le protocole suivant (toutes les manipulations ont été effectuées stérilement et sur glace) :

- Les bactéries ont étéensemencées à 1% avec une préculture saturée, dans 100 mL de milieu LB et incubées à 37°C sous agitation jusqu'à l'obtention d'une D.O.<sub>600 nm</sub> comprise entre 0,5 et 0,7.
- Les cellules ont été récupérées par centrifugation pendant 10 min, 4000 x g à 4°C.
- Les culots cellulaires ont ensuite été lavés deux fois avec 100 mL d'eau ultra pure, puis avec 20 mL de glycérol à 10% (v/v).
- Le culot bactérien a finalement été repris avec 500 µL de glycérol 10% (v/v) et réparti en fractions aliquotes de 40 µL.
- Les cellules compétentes ont été, soit directement utilisées pour la transformation, soit conservées à - 80°C.

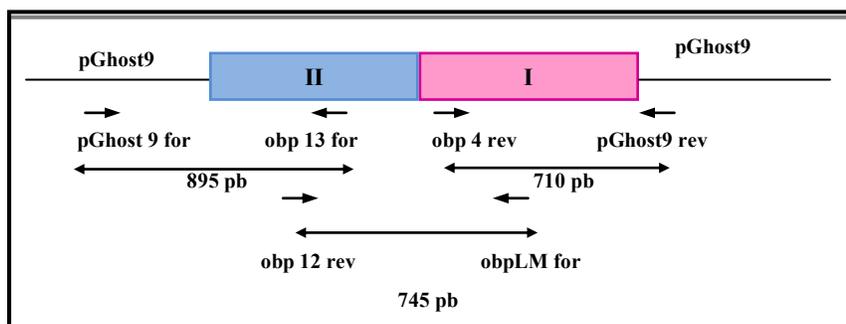
#### 4. 2. 2. Transformation d'*Escherichia coli* EC101 par électroporation

Un aliquot de cellules compétentes a été déposée dans une cuve d'électroporation stérile préalablement refroidie. 0,5 µL de plasmide recombinant pGhost9 $\Delta$ ots ont été ajoutés (équivalent à 250 ng). Le mélange a été incubé 1 min sur glace. L'appareil MicroPulser (BioRad) applique une tension de 3 kV, soit un temps de passage d'environ 5 ms (programme Ec3 de l'appareil), ce qui provoque l'introduction du plasmide dans les cellules. Pour régénérer les cellules, 1 mL de milieu SOC a été immédiatement ajouté, puis les cellules ont été incubées 1 h à 37°C sous agitation. Finalement, les transformants ont été étalés sur milieu LB gélosé contenant 150 µg.mL<sup>-1</sup> d'érythromycine.

#### 4. 2. 3. Sélection des transformants

Plusieurs clones résistants à l'érythromycine ont été sélectionnés après incubation d'une nuit à 37°C. Leurs plasmides ont été extraits à l'aide du kit « GenElute™ Plasmid Miniprep » (Sigma). Afin de contrôler la présence de l'insert I-II, les plasmides ont ensuite été digérés par les enzymes de restriction *EcoRI* et *XhoI*. Les produits de digestion ont ensuite été séparés par électrophorèse sur gel d'agarose.

Afin de vérifier la séquence de l'insert I-II, les plasmides positifs sur gel d'agarose ont été séquencés par la société GENOME EXPRESS avec les couples d'amorces suivants : « pGhost9 rev/obp4 rev », « obpLM for/obp12 rev » et « obp13 for/pGhost9 for ». Les séquences des amorces, les tailles d'amplicons et les températures d'hybridation sont indiquées dans les tableaux 13, 14 et 16. Les positions d'hybridation des amorces sur le plasmide sont schématisées dans la Figure 21.



**Figure 21 : Représentation schématique des positions d’hybridation des amorces utilisées pour séquencer l’insert I-II porté par le plasmide pGhost9**

### **4. 3. Transformation de la souche *Streptococcus thermophilus* LMD-9 et sélection des transformants**

#### **4. 3. 1. Préparation des cellules compétentes de *Streptococcus thermophilus* LMD-9 et électroporation**

La souche *S. thermophilus* LMD-9 a été rendue compétente selon le protocole suivant : Du lait reconstitué a été ensemencé avec la souche LMD-9 à 1% (v/v) et incubé à 42°C pendant 8 h. Cette préculture a ensuite été utilisée pour ensemencer à 1% (v/v) du milieu Belliker liquide (Tableau 11). Pour récupérer des cellules en fin de phase exponentielle de croissance, des dilutions décimales ont été effectuées dans le même milieu. Enfin, les différentes cultures ont été incubées une nuit à 42°C. Cette deuxième préculture a été utilisée pour ensemencer 100 mL de milieu Belliker-Thr (20 mM) à raison de 1% (v/v). La culture a été incubée à 42°C jusqu’à l’obtention d’une D.O.<sub>600 nm</sub> supérieure ou égale à 0,3. Les cellules ont été récoltées par centrifugation pendant 15 min, 3000 x g à 20°C, puis le culot a été resuspendu dans 200 mL de milieu Belliker-Thr (20 mM)-Gly (130 mM) et incubé 1 h à 42°C. Les cellules ont ensuite été lavées deux fois avec 20 mL de tampon d’électroporation (confère composition dans le tableau

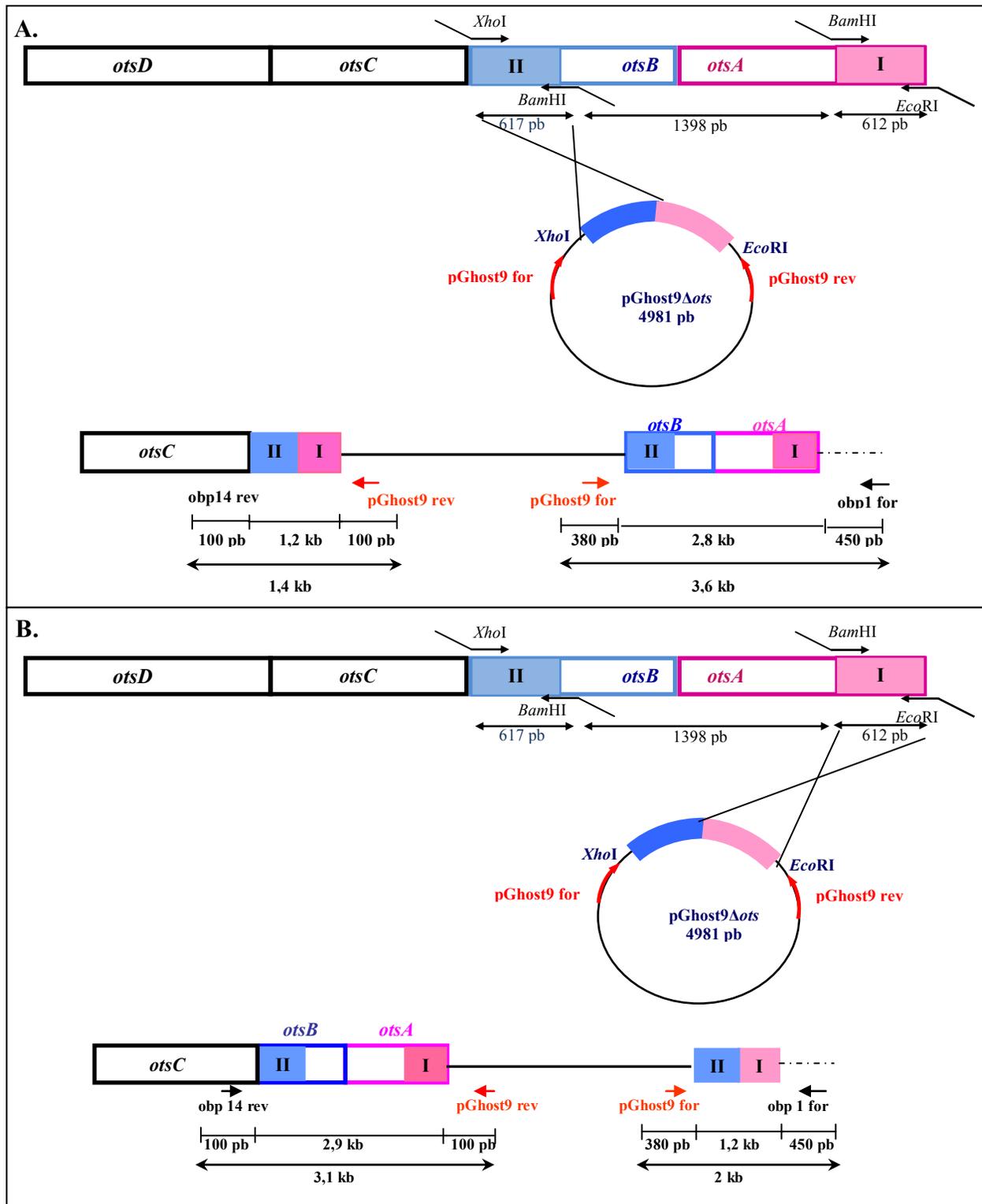
11). Le culot a alors été repris dans 2 mL de tampon d'électroporation et réparti en aliquots de 80  $\mu\text{L}$ .

Une quantité de 250 ng d'ADN plasmidique a été ajoutée à 80  $\mu\text{L}$  de cellules compétentes et le mélange a été incubé 5 min sur glace. Le plasmide pGhost9 $\Delta$ ots a été ensuite introduit dans la souche LMD-9 par électroporation (à 2,5 kV, 200  $\Omega$ , 25  $\mu\text{F}$  avec un temps de passage de environ 4 ms). Après électroporation, les cellules ont été régénérées par l'ajout de 1 mL de milieu Belliker. Après un repos de 5 min sur glace, les cellules transformées en milieu liquide ont été incubées 3 h à 30°C avant étalement sur milieu Belliker gélosé contenant de l'érythromycine à 5  $\mu\text{g}\cdot\text{mL}^{-1}$ . Les milieux gélosés ont alors été incubés en conditions anaérobies à 30°C pendant 48 h pour permettre la réplication du plasmide. Finalement, des transformants ont été isolés, régénérés puis stockés dans du lait contenant de l'érythromycine à 5  $\mu\text{g}\cdot\text{mL}^{-1}$  à - 80°C.

#### **4. 3. 2. Intégration du plasmide pGhost9 $\Delta$ ots dans le chromosome de la souche LMD-9**

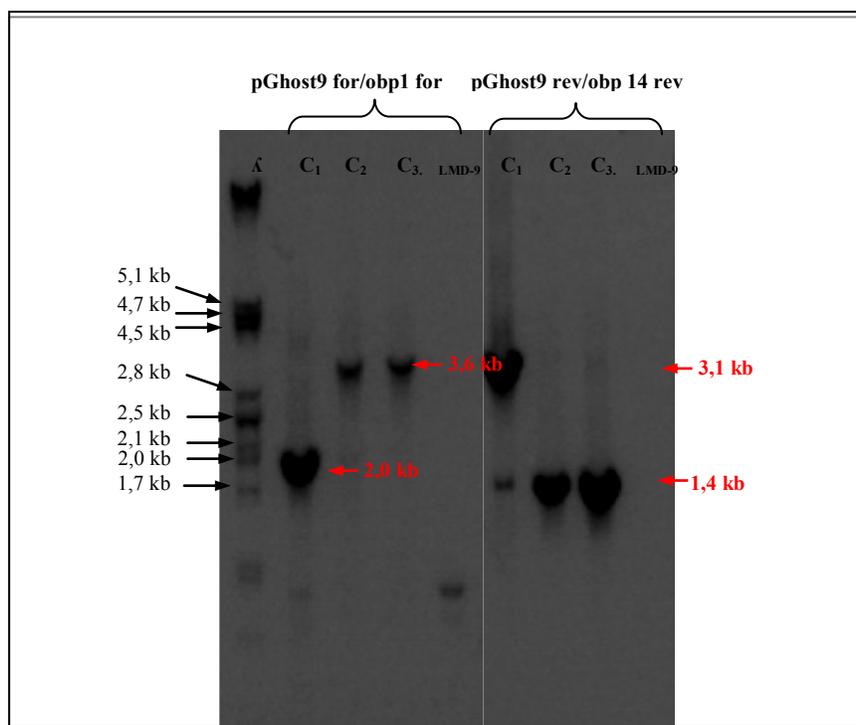
Le premier événement de recombinaison homologue permet l'intégration du plasmide pGhost9 $\Delta$ ots dans le chromosome. En effet, les régions I ou II portées par ce plasmide sont identiques au chromosome bactérien (Figure 20). Les colonies portant le plasmide intégré dans le chromosome ont été sélectionnées par culture à 42°C sur milieu M17 gélosé contenant de l'érythromycine à 2  $\mu\text{g}\cdot\text{mL}^{-1}$ .

L'intégration a alors été vérifiée : quelques colonies ont été cultivées en lait additionné de 2  $\mu\text{g}\cdot\text{mL}^{-1}$  d'érythromycine à 42°C pendant une nuit. Ces précultures ont été utilisées pour ensemencher des milieux M17 liquides à 1% (v/v) en présence de 2  $\mu\text{g}\cdot\text{mL}^{-1}$  d'érythromycine puis incubées à 42°C pendant 24 h. L'ADN génomique a été extrait. L'intégration du plasmide pGhost9 $\Delta$ ots dans le chromosome a alors été vérifiée par PCR en utilisant les couples d'amorces « pGhost9 for/obp1 for » et « opb14 rev/pGhost9 rev », mentionnés dans les tableaux 13, 14 et 16. Les positions des amorces et les tailles des amplicons attendus sont illustrées dans la Figure 22. La migration électrophorétique des produits d'amplification a révélé plusieurs colonies répondant positivement à l'intégration génomique du plasmide pGhost9 $\Delta$ ots (Figure 23).



**Figure 22 : Représentation schématique des deux possibilités d'intégration du plasmide pGhost9 au sein du génome de la souche LMD-9**

Les positions d'hybridation des amorces utilisées pour détecter l'intégration par PCR et les tailles de produits d'amplification attendus sont mentionnées.



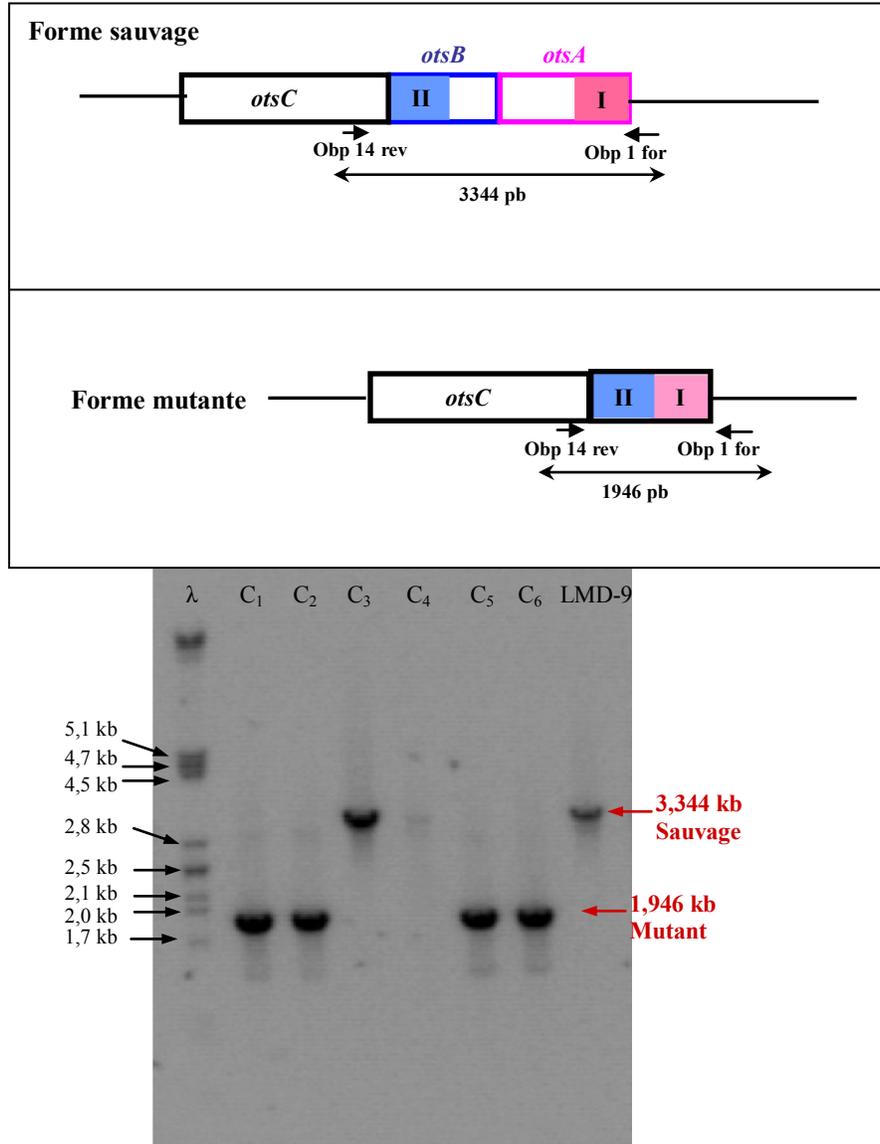
**Figure 23 : Détection par PCR de l'intégration du plasmide pGhost9 $\Delta$ ots dans le chromosome de la souche *Streptococcus thermophilus* LMD-9**

L'amplification a été réalisée par PCR à l'aide des couples d'amorces « pGhost9 for/obp1 for », « pGhost9 rev/obp14 rev » à partir d'ADN génomique.

Le puits  $\lambda$  : correspond à l'ADN du phage  $\lambda$  digéré par l'enzyme de restriction *Hind*III utilisé comme marqueur de taille ; puits C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub> correspondent à trois colonies résistantes à l'érythromycine présentées ; LMD-9 : représente la souche LMD-9 utilisée comme témoin négatif.

### 4. 3. 3. Excision du plasmide

Un deuxième évènement de recombinaison homologe permet l'excision du plasmide en laissant sur le chromosome soit la forme sauvage, soit la forme mutée (Figure 20). Pour provoquer le deuxième évènement de recombinaison homologe, plusieurs colonies contenant le plasmide pGhost9 $\Delta$ *ots* intégré dans le chromosome ont été placées en milieu M17 liquide sans érythromycine à 30°C pendant 12 h. Cette culture a servi à ensemercer à 1% (v/v) du milieu M17 sans érythromycine qui a été incubé à 42°C pendant 12 h. La culture ainsi obtenue a servi à son tour pour ensemercer (à 1%) du milieu M17 sans érythromycine à 30°C pendant 12 h. Ces étapes ont été répétées 3 fois. Les colonies comportant le plasmide inséré dans le chromosome sont résistantes à l'érythromycine, alors que les colonies ayant perdu le plasmide sont redevenues sensibles à l'antibiotique. L'excision du plasmide a été vérifiée par repiquage sur un milieu M17 gélosé en présence et en absence de 2  $\mu\text{g}\cdot\text{mL}^{-1}$  d'érythromycine. La présence de la forme mutée a été recherchée par PCR chez les colonies sensibles à l'érythromycine, en utilisant le couple d'amorce « obp1 for/opb14 rev », mentionné dans les tableaux 13 et 16 . La migration électrophorétique des produits d'amplification a révélé la présence de la forme mutée pour certaines colonies (Figure 24). Enfin, pour vérifier la délétion, le cluster  $\Delta$ *ots* a été entièrement séquencé. L'analyse de la séquence nucléotidique de  $\Delta$ *ots* a montré qu'un fragment de taille attendue de 1398 pb a été délété. Il comprend 1028 pb du côté 3' du gène *otsA*, 268 pb du côté 5' du gène *otsB* et 102 pb entre les deux gènes.



**Figure 24 : Recherche de la forme mutante  $\Delta$ *ots* dans le chromosome de la souche *Streptococcus thermophilus* LMD-9**

A. Position du couple d'amorces utilisé pour détecter la présence de la délétion souhaité ou la forme sauvage

B. Migration électrophorétique des produits d'amplification obtenus par PCR

Les puits C<sub>1</sub>, C<sub>2</sub>, C<sub>5</sub>, C<sub>6</sub> correspondent à quatre colonies possédant la délétion attendue ; C<sub>3</sub>, C<sub>4</sub> : deux colonies possédant la forme sauvage ;  $\lambda$  : l'ADN du phage  $\lambda$  par l'enzyme de restriction *Hind*III utilisés comme un marqueur de taille pour estimer la taille des produits

## 5. Caractérisation du rôle du mutant

### 5. 1. Estimation de la croissance bactérienne

En milieux M17, YL et MCD, la croissance des souches LMD-9 et LMD-9 $\Delta$ *ots* ont été suivies par mesure du pH et de la densité optique à 650 nm. En lait, la croissance a été estimée par mesure du pH et du degré d'acidité Dornic.

L'acidité Dornic a été directement mesurée sur 3 tubes avant incubation (temps zéro). Les autres tubes ont été incubés à 42°C et l'acidité Dornic déterminée en triplicat à 2 h et 6 h de culture. Les calculs du degré d'acidité Dornic ont alors été effectués selon la méthode suivante :

Une mole de soude versée neutralise une mole d'acide lactique ; la soude Dornic a un titre de 1/9 M ; l'acide lactique a une masse molaire de 90 g.mol<sup>-1</sup>. Ainsi, 1 mL de NaOH versé correspond à  $1/9 \cdot 10^{-3}$  mole d'OH<sup>-</sup> soit à  $1/9 \cdot 10^{-3}$  mole d'acide lactique à l'équivalence ou  $(1/9) \cdot 10^{-3} \times 90$  g d'acide lactique apportés par 10 mL de lait, soit 1 g d'acide lactique par litre de lait. Sachant qu'un degré Dornic est égal à 0,1 g d'acide lactique par litre de lait, 1 mL de soude versé correspond à 10 degrés Dornic.



## **IV. Résultats et Discussion**



# Chapitre I



## Chapitre I : Publication

### **Variabilité d'hydrolyse des caséines $\beta$ , $\alpha_{s1}$ et $\alpha_{s2}$ de 10 souches de *Streptococcus thermophilus* et leurs peptides bioactifs résultant.**

#### **Résumé :**

Les protéines du lait renferment dans leur séquence de nombreux peptides potentiellement bioactifs, qui peuvent être libérés soit par l'action de protéases digestives soit par le système protéolytique des bactéries lactiques lors de la fabrication des produits alimentaires. La capacité de *S. thermophilus* à générer des peptides à partir des caséines du lait bovin, en particulier des peptides bioactifs, a été étudiée. Dix souches exprimant différemment la protéase, PrtS, ancrée à la paroi cellulaire, ont été incubées en présence d'une des trois caséines  $\alpha_{s1}$ ,  $\alpha_{s2}$  ou  $\beta$ . Les analyses des surnageant par LC-ESI-MS/MS ont montré que la caséine  $\beta$  a été préférentiellement hydrolysée, suivie par la caséine  $\alpha_{s2}$  et la caséine  $\alpha_{s1}$ . Le nombre et les types de peptides libérés ont été dépendant de la souche utilisée. L'hydrolyse semble être liée à l'accessibilité des différentes régions des caséines à la protéase. L'analyse des liaisons peptidiques hydrolysées dans la région 1-23 de la caséine  $\alpha_{s1}$  suggère que la protéase PrtS est responsable de la production de peptides. Enfin, parmi les peptides générés ; 13 peptides générés à partir de la caséine  $\beta$ , 5 peptides à partir de la caséine  $\alpha_{s2}$  et 2 peptides à partir de la caséine  $\alpha_{s1}$  ont été référencés dans la littérature comme des peptides bioactifs, dont 15 étant des peptides inhibiteurs de l'enzyme de conversion de l'angiotensine.



## Variability of Hydrolysis of $\beta$ -, $\alpha_{s1}$ -, and $\alpha_{s2}$ -Caseins by 10 Strains of *Streptococcus thermophilus* and Resulting Bioactive Peptides

Laurent Miclo,<sup>||</sup> Émeline Roux,<sup>||</sup> Magali Genay, Émilie Brusseau, Chantal Poirson, Nawara Jameh, Clarisse Perrin, and Annie Dary\*

Unité de Recherche "Animal et Fonctionnalités des Produits Animaux" (UR AFPA), Équipe "Protéolyse et Biofonctionnalités des Protéines et des Peptides" (PB2P), Nancy-Université, 54506 Vandœuvre-lès-Nancy, France

**ABSTRACT:** Milk proteins contain numerous potential bioactive peptides, which may be released by digestive proteases or by the proteolytic system of lactic acid bacteria during food processing. The capacity of *Streptococcus thermophilus* to generate peptides, especially bioactive peptides, from bovine caseins was investigated. Strains expressing various levels of the cell envelope proteinase, PrtS, were incubated with  $\alpha_{s1}$ -,  $\alpha_{s2}$ -, or  $\beta$ -casein. Analysis of the supernatants by LC-ESI-MS/MS showed that the  $\beta$ -casein was preferentially hydrolyzed, followed by  $\alpha_{s2}$ -casein and then  $\alpha_{s1}$ -casein. Numbers and types of peptides released were strain-dependent. Hydrolysis appeared to be linked with the accessibility of different casein regions by protease. Analysis of bonds hydrolyzed in the region 1–23 of  $\alpha_{s1}$ -casein suggests that PrtS is at least in part responsible for the peptide production. Finally, among the generated peptides, 13 peptides from  $\beta$ -casein, 5 from  $\alpha_{s2}$ -casein, and 2 from  $\alpha_{s1}$ -casein have been reported as bioactive, 15 of them being angiotensin-converting enzyme inhibitors.

**KEYWORDS:** *Streptococcus thermophilus*, bioactive peptides, cell envelope proteinase, casein hydrolysis

### ■ INTRODUCTION

In recent years, a link between the prevalence of certain diseases commonly encountered in Western countries, such as cardiovascular diseases, obesity, diabetes, cancer, and dietary factors, has been emphasized. The engineering and manufacturing of functional ingredients and foods is emerging to counter these pathologies. Among them, fermented dairy products have received much attention because they contain several compounds, such as bioactive peptides, obtained by milk protein hydrolysis.<sup>1,2</sup> Bioactive peptides are encrypted within many dietary protein sequences and, when released by various proteases, exercise physiological functions in the human body with a wide range of potential applications (heart, bone, dental, and digestive health; weight management; immunomodulation; mood swings, memory, and stress control).<sup>3</sup> Bioactive peptides described in the literature have been released from dietary proteins either by enzymes of the digestive tract such as trypsin or by the proteolytic system of microorganisms during food manufacture.<sup>3,4</sup> Numerous studies deal with the capacity of lactococci or lactobacilli to generate bioactive peptides from milk proteins during fermentation.<sup>3</sup> These studies were carried out on pure cultures of species such as *Lactobacillus delbrueckii* subsp. *bulgaricus*, *Lactococcus lactis* subsp. *cremoris*, *Lb. acidophilus*, *Lb. casei*, *Lb. jensenii*, *Lb. reuteri*, *Lb. rhamnosus*, *Lc. lactis* ssp. *lactis*, *Lc. raffinolactis*, *Leuconostoc mesenteroides* ssp. *cremoris*, and numerous strains of *Lb. helveticus* or cocultures with industrial yogurt starters (*Lb. delbrueckii* ssp. *bulgaricus* and *Streptococcus thermophilus*).<sup>5–9</sup>

Although it belongs to the *Streptococcus* genus, *S. thermophilus* is classified as "Generally Recognized As Safe".<sup>10</sup> It is the second lactic acid bacterium (LAB), after *Lc. lactis*, in terms of industrial use.<sup>11</sup> *S. thermophilus*, having strong acidifying capacities, plays an important role in the initial stages of dairy product manufacture. To our knowledge, very

few studies have proved the ability of this bacterium to produce bioactive peptides from milk proteins, perhaps because its proteolytic capacity was considered to be very weak. In particular, 16 peptides were characterized in a study on milk proteolysis by pure or mixed cultures of *Lb. delbrueckii* subsp. *bulgaricus* and *S. thermophilus*.<sup>12</sup> Two of these peptides correspond to potential angiotensin-converting enzyme inhibitors.

The proteolytic system of *S. thermophilus* consists of a facultative cell envelope proteinase (CEP) named PrtS, an ABC-transporter of oligopeptides and di/tripeptide transporter, and a pool of intracellular peptidases.<sup>13</sup> It has long been believed that the high cellular density of *S. thermophilus* in milk depends upon its cocultivation with other bacteria such as *Lb. bulgaricus* during the manufacture of fermented dairy products;<sup>14</sup> with this assumption in mind, the screening of 97 strains of the collection of the Institut National de la Recherche Agronomique (INRA, France) showed that only 3 strains were PrtS<sup>+</sup> and probably able to grow at high density in milk because they were able to strongly acidify it during their growth.<sup>15</sup> However, recent publications report the presence of the CEP PrtS in numerous strains.<sup>16–18</sup> In a previous work, we clustered into 3 groups 30 strains of our laboratory collection on the basis of their ability to acidify milk during their growth: high (H)-, low (L)-, and medium (M)-acidifying strains.<sup>16</sup> We showed that all of the H-strains possess the *prtS* gene (*prtS*<sup>+</sup> genotype) and expressed it (PrtS<sup>+</sup> phenotype), whereas the L-strains were *prtS*<sup>-</sup>. For the M-strains, the situation was more complex. Indeed, 62% had a *prtS*<sup>+</sup> genotype and displayed a

Received: June 1, 2011

Revised: November 20, 2011

Accepted: November 21, 2011

Published: November 21, 2011

PrtS<sup>-</sup> or PrtS<sup>±</sup> phenotype, this last phenotype corresponding to a very weak protease activity.

The aim of this work was to determine (i) whether *S. thermophilus* was able to generate bioactive peptides from purified bovine  $\alpha_{s1}$ - and  $\alpha_{s2}$ - and  $\beta$ -casein (CN) as protein source through the activity of its proteolytic surface system; (ii) whether such a capacity could vary among the *S. thermophilus* species; (iii) whether the CEP PrtS could account for such a production, at least in part. Hence, the proteolytic capacities of six H-strains, four M-strains, and one L-strain (strain CNRZ1066, which is prtS<sup>-</sup>) have been investigated.

## MATERIALS AND METHODS

**Bacterial Strains, Media, and Growth Conditions.** *S. thermophilus* strains used in this work are listed in Table 1. They

**Table 1.** *S. thermophilus* Strains Used in This Work<sup>a</sup>

<i>S. thermophilus</i> strain	origin	strain type	prtS	PrtS
ATE19PB8	yogurt	H	+	+
Y4	yogurt	H	+	+
4F44	cheese	H	+	+
LMD-9*	yogurt	H	+	+
PB302	yogurt	H	+	+
PB385	yogurt	H	+	+
CNRZ404	yogurt	M	+	±
CNRZ445	cheese	M	+	±
ATCC19258	pasteurized milk	M	+	±
HAD8 $\alpha$	yogurt	M	+	-
CNRZ1066*	yogurt	L	-	-

<sup>a</sup>Key: \*, genomes entirely sequenced; strain type H (high acidifying strain), L (low acidifying strain), and M (medium acidifying strain); genotype prtS (+, presence of the gene; -, absence). Phenotype PrtS is only qualitative (+, strong proteolytic activity; ±, weak proteolytic activity; -, no proteolytic activity) and was determined by zymography.<sup>16</sup>

were isolated in our laboratory either from yogurt or from cheese (P. Bracquart, personal communication) or came from the Centre National de Recherches Zootechniques (CNRZ) (INRA, Jouy-en-Josas, France) or American Type Culture Collection (ATCC) (Manassas, VA) collections. The genome of the L-strain CNRZ1066 and that of the H-strain LMD-9 are sequenced.<sup>19,20</sup> All of the strains used in this work were prtS<sup>+</sup> (except strain CNRZ1066) but express PrtS at different levels. Strain CNRZ1066 was used to evaluate CN proteolysis in the absence of PrtS. Strains were stored at -80 °C in reconstituted skim milk (10%, w/v). *S. thermophilus* strains were inoculated at 1% in skim milk and incubated overnight at 42 °C before each experiment. Precultures were then inoculated at 1% in M17 medium<sup>21</sup> supplemented with lactose (20 g/L) and incubated at 42 °C until OD<sub>650 nm</sub> reached 1.

**Measurement of Intracellular Peptidase Activity.** An intracellular peptidase activity, potentially released into the medium after cell lysis, was sought using the chromogenic substrate Lys-pNA. The substrate was prepared at a concentration of 2 mM, in 50 mM sodium phosphate buffer, pH 7.5. Each strain used in this work was grown in M17 medium supplemented with lactose (20 g/L) until OD<sub>650 nm</sub> reached 1, and cells were harvested by centrifugation at 4000g for 5 min at 20 °C. The cell pellet was washed two times with 50 mM sodium phosphate buffer, pH 7.5, prewarmed at 42 °C, and resuspended at an OD<sub>650 nm</sub> of 1 in the same buffer. The cells were then incubated at 42 °C, and an aliquot was collected immediately and then after 4 and 22 h. Cells were eliminated by centrifugation at 12000g during 5 min at 4 °C, and the supernatant was filtered through a 0.45  $\mu$ m filter. One hundred microliters of the supernatant was then mixed with 1 mL of Lys-pNA solution and incubated for 2 h at 37 °C,

and the absorbance was determined spectrophotometrically at 410 nm (Uvikon, Kontron, Switzerland). To ensure that intracellular peptidases of *S. thermophilus* were able to hydrolyze Lys-pNA substrate, intracellular proteins were extracted from the cell pellets, as previously described,<sup>16</sup> and incubated in the same conditions as above.

**Casein Preparation.** Bovine sodium caseinate was prepared and fractionated to obtain pure  $\beta$ -CN variant A2, according to the method of Girardet et al.,<sup>22</sup> and  $\alpha_{s2}$ -CN variant A and  $\alpha_{s1}$ -CN variant B, as previously described except that EDTA was not added to the CN fractions after their elution.<sup>23</sup>

**Casein Hydrolysis.** Cells were grown in M17 medium supplemented with 20 g/L lactose until OD<sub>650 nm</sub> reached 1 and harvested by centrifugation at 4000g for 5 min at 20 °C. The cell pellet was washed in 50 mM sodium phosphate buffer, pH 7.5, prewarmed at 42 °C and was resuspended in 1.5 mL of the same buffer. Five hundred microliters of the cell suspension was added in 15 mL of sodium phosphate buffer at 42 °C containing 1 mg/mL  $\alpha_{s1}$ ,  $\alpha_{s2}$ , or  $\beta$ -CN to obtain an OD<sub>650 nm</sub> of 1. The samples were incubated at 42 °C, and an aliquot was collected immediately and after 2 and 4 h of incubation. Caffeine was added (3.26 mg/L final concentration) as internal standard for HPLC. Samples were centrifuged (12000g, 5 min, 4 °C), filtered through 0.22  $\mu$ m filters (Millipore, Molsheim, France), and kept at -20 °C.

**Peptide Identification by LC-ESI-MS/MS.** Nano-LC experiments were performed using an online liquid chromatography tandem mass spectrometry (MS/MS) set up using a nano-LC system (LC-Packings, Dionex, Voisins le Bretonneux, France) fitted to a QSTAR XL (MDS SCIEX, Ontario, Canada) equipped with a nanoelectrospray ion source (ESI) (Proxeon Biosystems A/S, Odense, Denmark). Samples were first concentrated on a PepMap 100 C<sub>18</sub> column (5  $\times$  0.3 mm, 5  $\mu$ m particule size, Dionex, Amsterdam, The Netherlands). Peptides were separated on a PepMap C<sub>18</sub> column (150  $\times$  0.075 mm, 3  $\mu$ m particule size, Dionex). Solvent A contained 2% (v/v) acetonitrile, 0.08% (v/v) formic acid, and 0.01% (v/v) TFA in deionized water and solvent B, 95% (v/v) acetonitrile, 0.08% (v/v) formic acid, and 0.01% (v/v) TFA in deionized water. A linear gradient from 10 to 50% of solvent B in 65 min was applied for the elution at a flow rate of 0.2  $\mu$ L/min. Eluted peptides were directly electrosprayed into the mass spectrometer operated in positive mode. A full continuous MS scan was carried out followed by three data-dependent MS/MS scans. Spectra were collected in the range of  $m/z$  450–1500 for MS and in the range of  $m/z$  60–2000 for MS/MS spectra. The three most intense ions from the MS scan were selected individually for collision-induced dissociation (1+ to 4+ charged ions were considered for the MS/MS analysis). The mass spectrometer was operated in data-dependent mode automatically switching between MS and MS/MS acquisition using Analyst QS 1.1 software. The instrument was calibrated by multipoint calibration using fragment ions that resulted from the collision-induced decomposition of the  $\beta$ -CN(f193–209). The peptides present in the samples were identified from MS and MS/MS data by using MASCOT v. 2.2 software against a homemade database dealing with major milk proteins and corresponding to a portion of the Swissprot database. No specific enzyme cleavage was used, and the peptide mass tolerance was set to 0.3 Da for MS and to 0.2 Da for MS/MS. For a database containing 25 protein sequences ( $N$ ) and a significance level of 0.04 ( $\alpha$ ), the cutoff score is equal to  $-10 \times \log(\alpha/N) = 28$ . Matches with a Mascot score superior or equal to this cutoff score were considered to be statistically significant.

## RESULTS AND DISCUSSION

The capacity of *S. thermophilus* strains to generate peptides from CNs was evaluated by a biochemical approach (Table 1). The strains were grown to midexponential phase; cell pellets were recovered and further incubated with purified  $\alpha_{s1}$ -CN,  $\alpha_{s2}$ -CN, or  $\beta$ -CN in a 50 mM sodium phosphate buffer, pH 7.5. After 2 or 4 h of incubation at 42 °C, the peptides contained in the supernatants were separated by HPLC and characterized by

Table 2. Peptides Generated from Bovine  $\beta$ -Casein by the Proteolytic System of *S. thermophilus* H-Strains<sup>a</sup>

Sequence	Fragment	STRAINS					
		PB302	PB385	ATE 19PB8	4F44	Y4	LMD-9
SLVYFPFGPIPN	57-68						4
VPPFLQPEVM	84-93						4
<b>GYSKYKEAM*</b>	94-102	2 4	2 4	4	4	4	2 4
<b>GYSKYKEAMAPK*</b>	94-105	2 4	2 4	2 4	4	2 4	2 4
VSKVKEAMAPK*	95-105	2 4	2 4	2 4		2	2 4
SKVKEAMAPK*	96-105	2 4	2 4	2		2 4	2
<b>KVKEAMAPK*</b>	97-105	2 4	2 4	2 4	2 4	2 4	2 4
KVKEAMAPKHK*	97-107		2		2		
VKEAMAPK	98-105				4		
VKEAMAPKHK	98-107					4	
EAMAPK*	100-105		4			2	2
<b>EAMAPKHK*</b>	100-107	2 4	2 4	2 4	2 4	2 4	2
EAMAPKHKKE	100-108					4	
EAMAPKHKEMPFPKYPVEPFTESQ	100-123				4		
<b>HKEMPFPK*</b>	106-113	2 4	2 4	2 4	2 4	2 4	2 4
<b>HKEMPFPKYPVEPF*</b>	106-119	2 4	2 4	2 4	2 4	2 4	2 4
HKEMPFPKYPVEPFTE	106-121		4				4
<b>HKEMPFPKYPVEPFTESQ*</b>	106-123	2 4	2 4	2 4	4	2 4	2
KEMPFPKYPVEPF*	107-119	4	4	2 4		2 4	2 4
<b>EMPFPKYPVEPF*</b>	108-119	2 4	4	2 4	4	2 4	4
EMPFPKYPVEPFTESQ*	108-123		2		4		2
YPVEPF	114-119						4
TESQSLTL*	120-127		4	4		4	2
SLTLTDVENL*	124-133						2 4
SLTLTDVENLHLPLPL*	124-139			2			
SLTLTDVENLHLPLPLQ*	124-141	2				2	2
TLTDVENL*	126-133			4		2	2 4
TLTDVENLHLPLPL	126-139						4
TLTDVENLHLPLPLQ*	126-141	2 4	4	2		2	2 4
TDVENL	128-133						4
TDVENLHLPLPL*	128-139		4	2 4		2	4
TDVENLHLPLPLQ	128-141	4					
DVENLHLPLPL*	129-139					2	4
NLHLPLPL	132-139						4
LHLPLPLQ*	133-141	4	4		4	2 4	2 4
HLPLPL*	134-139			4		2	2 4
<b>HLPLPLQ*</b>	134-141	2 4	4	2 4	4	2	2 4
<b>HLPLPLQSWM*</b>	134-144	2 4	2 4	2 4	4	2	2 4
LPLPL	135-139						4
LPLPLQ*	135-141	4		4		2 4	2 4
LPLPLQSW	135-143						4
LPLPLQSWM*	135-144			4		2	2 4
LQSWMHQPH	140-148	4					
SWMHQPHQLPPT*	142-154		4	4			2 4
MHQPHQLPPT	144-154		4				
HQPHQLPPT*	145-154	4	4	4		2 4	2 4
VMFPQSVL*	155-163	4	2 4	2			2 4
MFPPQSVL*	156-163			4		2	
FPPQSVL*	157-163					4	2 4
FPPQSVLSL*	157-165			2 4	4	4	2 4
<b>SVLSLS*</b>	161-166	2 4	2 4	2 4	4	2 4	2 4
<b>SVLSLSQ*</b>	161-167	2 4	2 4	2 4	2 4	2 4	2 4
SVLSLSQS*	161-168			2	4		4
VLLSQ	162-167				4		
LSQSKVLPVPQ*	165-175	2					
<b>SQSKVLPVPQ*</b>	166-175	2 4	2 4	2 4	2 4	2 4	2 4
<b>SQSKVLPVPQK*</b>	166-176	2 4	2 4	2 4	2 4	2 4	2 4
SQSKVLPVPQKA*	166-177				2 4	4	
SQSKVLPVPQKAVPPQ*	166-182				2 4		
QSKVLPVPQK*	167-176	2 4	2 4	2 4		2 4	2 4
<b>SKVLPVPQ*</b>	168-175	2 4	2 4	2 4	2 4	2 4	2 4
<b>SKVLPVPQK*</b>	168-176	2 4	2 4	2 4	2 4	2 4	2 4
SKVLPVPQKA*	168-177		2		2		
SKVLPVPQKAVPPQ*	168-182				2		
<b>KVLPVPQ*</b>	169-175	4	4	2 4	4	2 4	2 4
KVLPVPQK*	169-176			4	4	4	2 4
<b>VLPVPQ*</b>	170-175	2 4	2 4	2 4	2 4	2 4	2 4
VLPVPQKAVPPQ*	170-182	2	2		4		
LPVPQ	171-175			4		4	4
LPVPQK	171-176					4	
LPVPQKAVPPQ	171-182					4	
<b>KAVPPQ*</b>	176-182	2 4	2 4	2 4	2 4	2 4	2 4
<b>KAVPPQR*</b>	176-183	2 4	2 4	2	4	2	2 4
<b>KAVPPQRDMPQ*</b>	176-188	2 4	2 4	2	2	2	2

Table 2. continued

Sequence	Fragment	STRAINS					
		PB302	PB385	ATE 19PB8	4F44	Y4	LMD-9
<b>KAVYPQRDMPIQA*</b>	<b>176-189</b>	2 4	2 4	2 4	4	2 4	2
KAVYPQRDMPIQAF	176-190		4				
<b>AVYPYQ*</b>	<b>177-182</b>	2 4	2 4	2 4	2 4	2 4	2 4
AVYPYQRDMPIQ*	177-188	2 4	4	2 4		2 4	2 4
AVYPYQRDMPIQA	177-189					4	
VPYPQ*	178-182	4		2 4	4	2 4	2 4
VPYPQRDMPI	178-187					4	4
VPYPQRDMPIQ	178-188					4	
<b>RDMPIQA*</b>	<b>183-189</b>	2 4	2 4	2	2 4	2	2 4
<b>RDMPIQAF*</b>	<b>183-190</b>	2 4	2 4	2 4	2 4	2 4	2 4
<b>RDMPIQAFLL*</b>	<b>183-191</b>	2 4	2 4	2 4	2 4	2 4	2 4
<b>RDMPIQAFLL*</b>	<b>183-192</b>	4	2 4	2 4	2 4	4	4
RDMPIQAFLLY*	183-193	2	2 4		2 4		
<b>DMPIQAF*</b>	<b>184-190</b>	4	2 4	4	2	2 4	2 4
<b>DMPIQAFLL*</b>	<b>184-191</b>	2	2 4	2 4	2	2 4	2 4
<b>DMPIQAFLL*</b>	<b>184-192</b>	4	2 4	2	4	2 4	2 4
MPIQAF	185-190					4	4
MPIQAFLL	185-191					4	4
PIQAFLL	186-192					4	4
IQAFLL	187-192					4	4
AFLLYQEPVL*	189-198		2 4				
<b>FLLYQEPVL*</b>	<b>190-198</b>	2 4	2 4	2	2	2	2
<b>LLYQEPVL*</b>	<b>191-198</b>	2 4	2 4	2 4	2 4	2 4	4
<b>LLYQEPVLGPVR*</b>	<b>191-202</b>	2 4	2 4	2	2 4	2	2
<b>LLYQEPVLGPVRGPFPIV*</b>	<b>191-209</b>	2 4	2 4	2	2	2	2
<b>LYQEPVL*</b>	<b>192-198</b>	4	2 4	2 4	2 4	2 4	2 4
<b>LYQEPVLGPVR*</b>	<b>192-202</b>	2 4	2 4	2	2 4	2	2
<b>LYQEPVLGPVRGPFPIV*</b>	<b>192-209</b>	2 4	2 4	2 4	2 4	2 4	2 4
YQEPV	193-197						4
<b>YQEPVL*</b>	<b>193-198</b>	2 4	2 4	2 4	2 4	2 4	2 4
<b>YQEPVLGPV*</b>	<b>193-201</b>	2 4	2	2 4	2 4	2 4	2 4
<b>YQEPVLGPVR*</b>	<b>193-202</b>	2 4	2 4	2 4	2 4	2 4	2 4
YQEPVLGPVRGPFV*	193-206					2 4	
<b>YQEPVLGPVRGPFPIV*</b>	<b>193-209</b>	2 4	2 4	2 4	2 4	2 4	2 4
<b>QEPVLGPVRGPFPIV*</b>	<b>194-209</b>	2 4	2 4	2	2 4	2 4	2 4
LGPVRGPFPIV*	198-209		2 4				
GPVRGPFV*	199-206	4					4
<b>GPVRGPFPIV*</b>	<b>199-209</b>	2 4	2 4	2 4	2 4	2 4	2 4
VRGPFPIV*	201-209					2	2
<b>GPPIV*</b>	<b>203-209</b>	2 4	2 4	4	4	4	2
PPPIV	204-209			4			
FPIV	205-209			4			
Number of peptides present after 2 h of incubation		48	52	53	38	61	65
Number of peptides present after 4 h of incubation and absent after 2 h		15	14	15	23	20	23
Number of peptides present after 4 h of incubation		58	61	54	53	61	73
Number of peptides present after 2 h of incubation and absent after 4 h		5	5	14	8	20	15
Total number of peptides formed by the strain (2 h and 4 h)		63	66	68	61	81	88
Number of specific peptides produced (2 h and 4 h)		5	6	4	8	17	20
Peptides average size (residues) after 2 h		10.6	10.3	10.3	10.4	10.1	10.0
± standard deviation (residues)		3.5	3.3	3.4	3.5	3.5	3.4
Peptides average size (residues) after 4 h		10.2	10.3	9.3	10.1	9.5	9.2
± standard deviation (residues)		3.4	3.2	3.1	3.9	3.1	3.0

\*Peptides were detected after 2 or 4 h of hydrolysis in 50 mM phosphate buffer, pH 7.5. Empty box refers to a lack of detection. The asterisk indicates the peptides that are generated after 2 h of incubation. A peptide is considered to be "specific" when it is produced by a maximum of two different strains and is shaded light gray. The peptides produced by all of the H-strains are shaded dark gray.

MS/MS. The L-strain CNRZ1066, which is *prtS*<sup>-</sup>, displayed, as expected, no proteolytic activity on the three CNs tested after 2 or 4 h of incubation. As no peptidase activity, from a cell lysis, was detected in the medium in which the cells were incubated, the hydrolysis of caseins observed here resulted only from a surface proteolytic activity.

**Identification of Peptides Generated from  $\beta$ -,  $\alpha_2$ -, and  $\alpha_1$ -Casein Hydrolysis.** Identification of Peptides Generated from  $\beta$ -CN. Not only was  $\beta$ -CN hydrolyzed more quickly by the proteolytic system of the *S. thermophilus* strains tested, but its hydrolysis degree was also higher. After 2 h of incubation, no hydrolysis was detected with the M-strains (data not shown), whereas 23.1% of the peptide bonds (48 of 208 bonds) of the  $\beta$ -CN were cleaved by all H-strains taken

together, resulting in a production of 84 different peptides (Table 2). The mean number of peptides produced by each strain was  $52.8 \pm 9.6$ . The number of generated peptides varied from one strain to another because only 38 peptides were liberated by strain 4F44, whereas 65 peptides were generated by strain LMD-9 (Table 2).

After 4 h of incubation, 31.3% of the peptide bonds of the  $\beta$ -CN were hydrolyzed by the 6 H-strains, leading to a total of 109 different peptides. The average number of peptides produced by each H-strain was  $60.0 \pm 7.2$  (Table 2). The proteolytic profiles differed between the strains analyzed. For example, only 53 peptides were detected for strain 4F44, whereas 73 peptides were generated by strain LMD-9 after 4 h of hydrolysis. Strains Y4 and LMD-9 were the most proteolytic

Table 3. Peptides Generated from Bovine  $\beta$ -Casein by the Proteolytic System of *S. thermophilus* M-Strains<sup>a</sup>

Sequence	Fragment	STRAINS			
		CNRZ 404	CNRZ 445	ATCC 19258	HAD8 $\alpha$
RELEELNVPGEIVE	1 - 14				X
RELEELNVPGEIVSL	1 - 16				X
KIEKFQSEEQQTTEDELQDKIHFP	29 - 52		X		
KIEKFQSEEQQTTEDELQDKIHFAQTQ	29 - 56		X		
KVKEAMAPK*	97 - 105		X		
EMPPFKYPVEPF*	108 - 119				X
MHQPHOPLPPTVMFPPOSVLSLSQSKVLPVPQK	144 - 176	X	X		
SLSQSKVLPVPQK	164 - 176	X	X		
SQSKVLPVPQ*	166 - 175	X		X	X
SQSKVLPVPQK*	166 - 176	X	X	X	
SQSKVLPVPQKAVPYQORDMPIQA	166 - 189		X		
SQSKVLPVPQKAVPYQORDMPIQAF	166 - 190		X		
SKVLPVPQK*	168 - 176		X		
AVPYQORDMPIQA	177 - 189				X
AVPYQORDMPIQAFLLYQEPVLGVRGPFPIIV	177 - 209				X
VPYQORDMPIQAF	178 - 190	X			
VPYQORDMPIQAFL	178 - 191	X	X		X
VPYQORDMPIQAFLLY	178 - 193		X		
RDMPIQAF*	183 - 190			X	X
RDMPIQAFL*	183 - 191				X
DMPIQAFL*	184 - 191				X
FLLYQEPVLGVRGPFPIIV	190 - 209			X	
LLYQEPVL*	191 - 198				X
LLYQEPVLGVRGPFPIIV*	191 - 209		X		
LYQEPVL*	192 - 198				X
YQEPVL*	193 - 198			X	X
YQEPVLGVR*	193 - 202	X	X		
YQEPVLGVRGPF	193 - 206	X			X
YQEPVLGVRGPFPIIV*	193 - 209	X	X	X	X
QEPVLGVRGPFPIIV*	194 - 209	X	X	X	X
Total number of peptides generated		10	15	7	16
Peptides average size (residues)		15.1	17.9	12.6	12.8
$\pm$ standard deviation (residues)		6.4	7.2	4.8	6.2
Average size of peptides only generated by M-strains (residues)		18.3	22.1	20.0	19.3
$\pm$ standard deviation (residues)		8.5	6.7	-	8.0

<sup>a</sup>Peptides were detected only after 4 h of hydrolysis in 50 mM phosphate buffer, pH 7.5. "X" indicates that the peptide is produced by the corresponding strain; empty box refers to a lack of detection. Peptides shaded gray are produced only by M-strains. An asterisk next to the sequence indicates a peptide also generated by all of the H-strains studied. "S" in a sequence indicates a residue of phosphoserine.

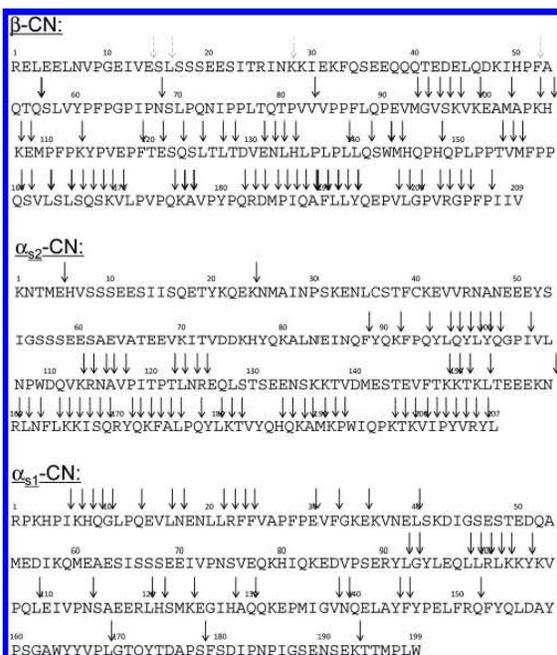
toward  $\beta$ -CN because almost all new peptides produced by strains PB302 and PB385 between 2 and 4 h of hydrolysis were already released after 2 h of hydrolysis by strains LMD-9 and Y4. Only 44 of the peptides identified were formed by all of the H-strains studied. Globally, during the hydrolysis of 4 h, 116 different peptides could be produced by the 6 H-strains tested, and only 7 peptides found after 2 h of hydrolysis were not recovered after 4 h ( $\beta$ -CN(f97–107),  $\beta$ -CN(f124–139),  $\beta$ -CN(f124–141),  $\beta$ -CN(f165–175),  $\beta$ -CN(f168–177),  $\beta$ -CN(f168–182), and  $\beta$ -CN(f201–209); Table 2).

The number of specific peptides (a specific peptide was defined as a peptide produced by a maximum of 2 different strains) varied from one strain to another: strain ATE19PB8 produced only 4 specific peptides, whereas strains Y4 and LMD-9 generated 17 and 20, respectively. Such differences could not be attributed only to the import of peptides which could differ between the strains or to an additional hydrolysis, because few of the peptides generated after 2 h of  $\beta$ -CN hydrolysis disappeared after 4 h (Table 2). Therefore, most of the peptides generated from  $\beta$ -CN after 2 h of hydrolysis by the H-strains corresponded to final peptides. The large proportion of proline residues (1–4) in their sequence could have increased their resistance toward proteolysis.<sup>24</sup>

Only 30 peptides were generated by the M-strains after 4 h of incubation, suggesting that their proteolytic capacities were lower than those of the H-strains (Table 3). Thirteen of the 30 peptides were specifically formed by the 4 M-strains studied in

this work (Table 3). The mean size of these peptides was about 21 residues, that is, significantly longer than for H-strains (about 10 residues). Fifteen of 17 peptides generated from  $\beta$ -CN by both M- and H-strains were systematically produced by all of the H-strains studied. Therefore, these peptides were generated from accessible zones of caseins by the proteolytic system of *S. thermophilus* even by the less efficient system of the M-strains. Galia et al.<sup>16</sup> showed that the *prtS* allele of H-strains LMD-9, PB302, and PB385 and of M-strain HAD8 $\alpha$  are the same and that M-strains expressed PrtS at lower levels than H-strains in M17 medium. Therefore, a lower proteolytic enzyme expression could slow the kinetics of hydrolysis, and large peptides observed in the case of M-strains might correspond to intermediate peptides.

Cleavage site distribution analysis of the whole  $\beta$ -CN sequence revealed a varied distribution of cleaved bonds (Figure 1). Thus, no peptide bond located between residues 1 and 92 was cleaved after 2 h of incubation. Even after 4 h of hydrolysis, the N-terminal part of the  $\beta$ -CN was very poorly hydrolyzed because only 2 peptide bonds (14–15 and 16–17) were cleaved by M-strain HAD8 $\alpha$ , 3 (28–29, 52–53, and 56–57) by M-strain CNRZ445, and 3 (56–57, 68–69, and 83–84) by H-strain LMD-9 (Figure 1). Most of the peptides generated by all of the H-strains came from the 160–209 region of the  $\beta$ -CN. On the basis of hydrophobicity estimated according to the method of Kyle and Doolittle,<sup>25</sup> it appeared in a first approximation that the hydrophilic amino-terminal extremity



**Figure 1.** Peptide bonds of bovine  $\beta$ -,  $\alpha_{22}$ -, and  $\alpha_{51}$ -CN hydrolyzed after 4 h of incubation by the proteolytic system of *S. thermophilus* strains. The arrows represent peptide bonds that are hydrolyzed, gray dotted arrows indicate bonds cleaved only by the M-strains tested in this work, and bold arrows indicate bonds cleaved by both M- and H-strains.

of the protein (residues 1–50) was more resistant to hydrolysis than the hydrophobic carboxy-terminal extremity (153–175 and 187–209 residues).

**Identification of Peptides Generated from  $\alpha_{22}$ -CN.** The H- and M-strains did not generate any hydrolysis products from  $\alpha_{22}$ -CN after 2 h, and only the H-strains produced peptides after 4 h (Table 4). Therefore,  $\alpha_{22}$ -CN appeared to be hydrolyzed at a slower rate than  $\beta$ -CN by *S. thermophilus*. Sixty-two peptide bonds of 206 were hydrolyzed (30.0%) after 4 h of hydrolysis of this CN by the H-strains, leading to the production of 89 different peptides. The mean number of peptides produced by each strain was  $35.8 \pm 12.4$ . The mean number of residues was  $9.5 \pm 3.8$ . As observed for  $\beta$ -CN, a great variability between the six H-strains tested was observed. Strains PB302 and ATE19PB8 generated only 25 peptides from  $\alpha_{22}$ -CN, whereas strains PB385 and 4F44 produced 45 and 56 peptides, respectively. These two strains were able to produce a great number of specific peptides compared to the other four. Only six of the peptides recovered in the hydrolysis medium were common for the six H-strains studied, and 58.4% of the peptides released appeared to be specific.

As observed for  $\beta$ -CN, the cleavage sites were not uniformly distributed along the  $\alpha_{22}$ -CN sequence (Figure 1).  $\alpha_{22}$ -CN is composed of three regions: a negatively charged region (3–18), a hydrophobic region (78–109), and a positively charged region (126–207). As shown by Tauzin et al.<sup>23</sup> using a tryptic digestion, the association of  $\alpha_{22}$ -CN in a tetrameric form in phosphate buffer (pH 8.1) leads to a protection of some regions toward hydrolysis. Thereby, the carboxy extremity of  $\alpha_{22}$ -CN (153–207) was more susceptible to trypsin hydrolysis.

On the contrary, the amino extremity (1–91) and the central region (114–150) were almost not accessible to trypsin. A similar phenomenon was observed with the proteolytic system of *S. thermophilus* H-strains (Figure 1). Indeed, the region 159–207 was extensively hydrolyzed, whereas only the two bonds 5–6 and 24–25 were cleaved in the amino-terminal part of the protein (residues 1–88) by strains 4F44 and LMD-9 (Figure 1; Table 4). Therefore, the most hydrolyzed region of  $\alpha_{22}$ -CN was hydrophilic, and not hydrophobic as previously observed with  $\beta$ -CN. This suggested that the hydrophobicity was not the main parameter to determine CN susceptibility to hydrolysis by the proteolytic system of *S. thermophilus*. Indeed, accessibility to substrate appeared to be a crucial parameter because the CN regions that are not involved in oligomerization were hydrolyzed by the proteolytic system. CN accessibility is also the determinant parameter for the action of the proteolytic system of *Lb. helveticus*.<sup>26</sup> Besides the peptides generated by all H-strains, the specific peptides were also released from accessible zones of caseins, which might appear surprising because the proteolytic system of all bacteria should be able to easily hydrolyze these regions. PrtS is the sole caseinolytic CEP that has been identified in *S. thermophilus*, and the polymorphism concerning this CEP mainly consists of its presence or absence. If PrtS is responsible for the hydrolysis observed, this implies that when the affinity of the CEP for a bond located inside an accessible region is low, the cleavage of this bond requires a higher quantity of the CEP and, therefore, a higher expression of this protease.

**Identification of Peptides Generated from  $\alpha_{51}$ -CN.** As observed with  $\alpha_{22}$ -CN, the H- and M-strains did not hydrolyze  $\alpha_{51}$ -CN after 2 h, and only H-strains were able to generate peptides after 4 h of hydrolysis (Table 5).  $\alpha_{51}$ -CN was the least hydrolyzed of the three CNs studied because only 39 of the 198 bonds (i.e., 19.7%) were cleaved by the 6 *S. thermophilus* H-strains taken together, leading to the formation of only 40 peptides. The mean number of peptides produced by each strain was  $20.0 \pm 5.0$ , and the mean number of residues was  $9.6 \pm 4.6$ . Only 20.0% of the peptides were found to be produced by all six H-strains studied, with strain LMD-9 being the most proteolytic strain, as observed on  $\beta$ -CN. This strain generated 29 peptides, whereas strains PB385 and Y4 liberated only 16 peptides. Many of the peptides produced from  $\alpha_{51}$ -CN by the six H-strains were specific, this type of peptide representing 55% of the peptides produced. Half of the specific peptides was produced only by strain LMD-9. Four peptides ( $\alpha_{51}$ -CN(f1–8),  $\alpha_{51}$ -CN(f1–9),  $\alpha_{51}$ -CN(f1–13), and  $\alpha_{51}$ -CN(f1–17)) have been characterized after milk hydrolysis by a pure culture of *S. thermophilus* CNRZ385.<sup>12</sup> These peptides, except  $\alpha_{51}$ -CN(f1–17), were also generated by all of the H-strains studied in this work and therefore belonged to the group of nonspecific peptides.

Kumosinski et al.<sup>27</sup> described the structure of  $\alpha_{51}$ -CN as follows: a hydrophilic region from residue 1–12, a hydrophobic region from residue 13 to 40 approximately, a hydrophilic region until residue 100, and finally a highly hydrophobic region from residue 100 to the carboxy-terminal part. The  $\alpha_{51}$ -CN monomers could associate by hydrophobic interactions or via the charged regions. Hence, some peptide bonds could become inaccessible to proteolysis in the multimers formed. Analysis of the cleavages observed in our experiments revealed that the regions 90–110 and 1–40 were the most susceptible to hydrolysis as about 40% of the bonds in these regions were cleaved (Figure 1). On the contrary, the region 41–91

Table 4. Peptides Generated from Bovine  $\alpha_{2}$ -Casein by the Proteolytic System of *S. thermophilus* H-Strains<sup>a</sup>

Sequence	Fragment	STRAINS					
		PB302	PB385	ATE19PB8	4F44	Y4	LMD-9
KNTMEHVSSEESHSQETYSQEK	1-24				X		X
HVSSSEESHSQETYSQEK	6-24				X		
YQKFPQYLQ	89-97	X			X		
YLQVLY	95-100			X	X		
QVLYQGPIVLPWDQVQRN	97-115					X	
YLYQGPIVLPWDQVQK	98-113				X	X	X
YLYQGPIVLPWDQVQR	98-114		X	X	X		
YLYQGPIVLPWDQVQRN	98-115				X		
LYQGPIVLPWDQVQK	99-113		X				
LYQGPIVLPWDQVQR	99-114		X		X		
LYQGPIVLPWDQVQRN	99-115				X	X	X
YQGPIVLPWDQVQK	100-113	X	X	X			X
YQGPIVLPWDQVQR	100-114	X	X	X			X
YQGPIVLPWDQVQRN	100-115		X			X	
<b>QGPIVLPWDQVQK</b>	<b>101-113</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>
<b>QGPIVLPWDQVQR</b>	<b>101-114</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>
<b>QGPIVLPWDQVQRN</b>	<b>101-115</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>
GPVLPWDQVQR	102-114			X			X
VLPWDQVQK	105-113						X
VLPWDQVQR	105-114	X	X	X	X		X
VLPWDQVQRN	105-115					X	
NAVPIPT	115-122	X				X	
NAVPIPTL	115-123					X	
AVPIPT	116-122			X			
<b>AVPIPTL</b>	<b>116-123</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>
AVPIPTLN	116-124				X	X	
AVPIPTLNR	116-125				X	X	X
VPITPL	117-123			X			X
PIPTLNR	118-125				X	X	X
KTKLTKK	150-159		X				
TKLTKK	151-159	X	X	X			X
KTKLTKKRNRL	151-161				X		
KTKLTKKRNRLNFL	151-164				X		
KLTKKRNRL	152-161				X		
TKKRNRL	154-161				X		
TKKRNRLNFL	154-164				X		
RLNFK	160-165				X		
RLNFKK	160-166	X	X		X		
LNFK	161-165			X	X		
NFKKISQ	162-169				X		
FLKKISQRY	163-171				X		
KISQRYQKF	166-174		X				
ISQRYQKF	167-174		X				
SQRYQKF	168-174				X	X	
QRYQKF	169-174		X				
QKFALPQ	172-178					X	
QKFALPQYLK	172-181					X	X
KFALPQ	173-178				X		X
KFALPQYL	173-180				X		
KFALPQYLK	173-181		X		X		
KFALPQYLKT	173-182				X	X	X
FALPQ	174-178	X	X	X	X		
FALPQYLK	174-181		X				
ALPQYLK	175-181		X		X	X	X
FALPQYLKT	175-182		X		X	X	
LPQYLK	176-181			X	X	X	X
PQYLK	177-181		X		X	X	X
POYLKT	177-182				X		
YLKTVYQH	179-186	X	X		X		
YLKTVYQH	179-187		X				
KTVYQHOKA	181-189		X				
TVYQHOKA	182-189	X	X	X		X	
KAMKPWIQPK	188-197				X	X	X
AMKPWIQPK	189-197	X	X		X		
AMKPWIQPKT	189-198	X	X		X		
AMKPWIQPKTK	189-199	X	X	X	X		
MKPWIQPK	190-197	X	X		X		
MKPWIQPKT	190-198		X		X		
MKPWIQPKTK	190-199	X	X		X		
KPWQPK	191-197	X	X	X		X	X
<b>KPWQPKT</b>	<b>191-198</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>
<b>KPWQPKTK</b>	<b>191-199</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>
PWIQPK	192-197		X		X		X
PWIQPKT	192-198						X

Table 4. continued

Sequence	Fragment	STRAINS					
		PB302	PB385	ATE19PB8	4F44	Y4	LMD-9
PWIQPKTK	192-199				X		
WIQPK	193-197		X				
WIQPKTK	193-199	X		X		X	
TKVIP	198-202	X			X	X	
TKVIPY	198-203		X			X	
TKVIPYV	198-204		X				
TKVIPYVR	198-205	X	X	X	X	X	
TKVIPYVRY	198-206		X		X		
TKVIPYVRYL	198-207		X	X	X		
KVIPYVRYL	199-207		X		X		
VIPYVR	200-205					X	
VIPYVRYL	200-207					X	
IPYVR	201-205			X			X
IPYVRYL	201-207					X	
PYVRYL	202-207		X				
Total number of peptides generated		25	45	25	56	34	29
Number of specific peptides produced		2	17	6	25	12	8
Peptides average size (residues)		9.4	9.4	9.5	9.8	9.5	10.3
± standard deviation (residues)		2.8	3.2	3.6	4.0	3.6	4.3

<sup>a</sup>Peptides were detected after 4 h of hydrolysis in 50 mM phosphate buffer, pH 7.5. "X" indicates that the peptide is produced by the corresponding strain; empty box refers to a lack of detection. A peptide is considered to be "specific" when it is produced by a maximum of two different strains and is shaded light gray. The peptides produced by all of the strains are shaded dark gray.

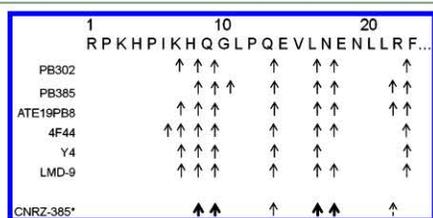
Table 5. Peptides Generated from Bovine  $\alpha_{s1}$ -Casein by the Proteolytic System of *S. thermophilus* H-Strains<sup>a</sup>

Sequence	Fragment	STRAINS					
		PB302	PB385	ATE19PB8	4F44	Y4	LMD-9
<b>RPKHPIKH</b>	<b>1-8</b>	X	X	X	X	X	X
<b>RPKHPIKHQ</b>	<b>1-9</b>	X	X	X	X	X	X
RPKHPIKHQG	1-10		X				
<b>RPKHPIKHQGLPQ</b>	<b>1-13</b>	X	X	X	X	X	X
RPKHPIKHQGLPQEVLENLLRF	1-23		X		X	X	X
KHQGLPQEV	7-16				X		
HQGLPQ	8-13	X		X	X	X	X
QGLPQEV	9-16	X		X	X		
QGLPQEVN	9-17	X	X	X	X		X
GLPQEVLENLL	10-21		X	X			
EVLNENLLR	14-22	X	X	X		X	X
<b>NENLR</b>	<b>17-22</b>	X	X	X	X	X	X
<b>NENLLRF</b>	<b>17-23</b>	X	X	X	X	X	X
<b>NENLLRFF</b>	<b>17-24</b>	X	X	X	X	X	X
ENLLR	18-22						X
FFVAPPEVF	23-32						X
FVAPPE	24-30						X
FVAPPEVF	24-32		X	X	X	X	X
FVAPPEVFGKE	24-35			X			
VAPPEVF	25-32			X			
GKEVNL	33-40			X			X
GYLEQL	93-98	X			X		
<b>GYLEQLL</b>	<b>93-99</b>	X	X	X	X	X	X
GYLEQLLR	93-100	X			X		
<b>GYLEQLRL</b>	<b>93-101</b>	X	X	X	X	X	X
GYLEQLRLK	93-102	X	X	X	X	X	
YLEQLL	94-99			X			
YLEQLRL	94-101			X	X	X	X
YLEQLRLK	94-102				X	X	
KVPQLEIVPN	105-114						X
EIVPNSAERLH	110-121						X
SAEERLHSMK	115-124	X		X		X	X
HSMKEGIIH	121-128						X
SMKEGIIHAQ	122-130	X	X	X	X		X
AQQKEPMIGVN	129-139						X
NQELAY	139-144						X
QELAY	140-144						X
YPELFRQ	146-152						X
GTQYTDAPFSFDIPNPIGSENSEKTTMLPW	170-199						X
FSDIPNPIGSENSEK	179-193						X
Total nb of peptides formed		17	16	22	20	16	29
Nb of specific peptides produced		1	2	5	3	1	13
Peptides average size (residues)		8.4	9.9	8.7	9.1	9.4	9.8
± standard deviation (residues)		1.8	3.9	1.9	3.7	4.0	5.2

<sup>a</sup>Peptides were detected after 4 h of hydrolysis in 50 mM phosphate buffer, pH 7.5. "X" indicates that the peptide is produced by the corresponding strain; empty box refers to a lack of detection. A peptide is considered to be "specific" when it is produced by a maximum of two different strains and is shaded light gray. The peptides produced by all of the strains are shaded dark gray.

appeared to be very resistant to hydrolysis and the region 152–199 was cleaved only four times by the sole strain LMD-9. As both hydrophobic and hydrophilic regions could be hydrolyzed or, on the contrary, protected from hydrolysis, the accessibility of a region seems to be the most accurate hypothesis to explain the susceptibility to hydrolysis.

**Implication of Protease PrtS in the Formation of the Peptides.** The CEP of lactococci have been divided into two main groups on the basis of their substrate specificity. The P<sub>I</sub>-type proteases hydrolyze predominantly  $\beta$ -CN, and the P<sub>III</sub>-type proteases also cleave  $\alpha$ - and  $\kappa$ -CNs.<sup>28,29</sup> The differences between the two groups are not so clear, and mixed-type specificities have been highlighted. Besides, on the basis of their action on the  $\alpha_{s1}$ -CN(f1–23) fragment, Exterkate et al.<sup>30</sup> proposed to divide the proteases of lactococci into seven groups and identified six potential cleavage sites on the  $\alpha_{s1}$ -CN(f1–23) fragment. The same cleavage sites have been observed using a purified PrtS from *S. thermophilus* CNRZ385:<sup>13</sup> four were major sites of hydrolysis, one was a medium site, and one was a minor site (Figure 2).



**Figure 2.** Hydrolysis of the region 1–23 of  $\alpha_{s1}$ -CN by the proteolytic system of *S. thermophilus* PB302, PB385, ATE19PB8, 4F44, Y4, and LMD-9 in 4 h in 50 mM phosphate buffer, pH 7.5. The cleavage sites are represented by arrows. CNRZ-385\* refers to the specificity of PrtS of the strain CNRZ385, determined on the peptide 1–23 of  $\alpha_{s1}$ -CN. Its six major, medium, and minor cleavage sites are indicated by a large, a medium, and a dotted arrow, respectively.<sup>13</sup>

In our experiments, *S. thermophilus* strains were able to degrade  $\beta$ -CN, but also  $\alpha$ -CNs at a slower rate, suggesting that their protease(s) displayed intermediate P<sub>I</sub>- and P<sub>III</sub>-types. This observation is consistent with that obtained for the sole CEP of *S. thermophilus* described today, that is, PrtS.<sup>13</sup> The analysis of the cleavage sites in the 1–23 region of the  $\alpha_{s1}$ -CN (Figure 2) revealed that they corresponded to those that were observed on the peptide  $\alpha_{s1}$ -CN(f1–23) after hydrolysis by PrtS purified from *S. thermophilus* CNRZ385.<sup>13</sup> Four of the six bonds of  $\alpha_{s1}$ -CN(f1–23), which have been cleaved by PrtS of the strain CNRZ385, were cut by all of the H-strains studied in this work; they corresponded to three major sites, H<sup>8</sup>–Q<sup>9</sup>, Q<sup>9</sup>–G<sup>10</sup>, L<sup>16</sup>–N<sup>17</sup>, and to the medium site Q<sup>13</sup>–E<sup>14</sup> (Figure 2). The major cleavage site N<sup>17</sup>–E<sup>18</sup> was cut by all of the H-strains, except strain Y4, and the minor site L<sup>21</sup>–R<sup>22</sup> was hydrolyzed by two of the six H-strains. This bond was hydrolyzed at a very slow rate in the study of Fernandez-Espla et al.,<sup>13</sup> and this tendency was also found in our study. The bond R<sup>22</sup>–F<sup>23</sup> of the  $\alpha_{s1}$ -CN was cut by all six H-strains, whereas PrtS, purified from strain CNRZ385, did not cut this bond in peptide  $\alpha_{s1}$ -CN(f1–23). The pure PrtS was unable to generate the cleavage between the residues R22 and F23 in peptide  $\alpha_{s1}$ -CN(f1–23), because this bond is the last one and could have been hydrolyzed only if PrtS exhibited an exopeptidase-like activity. Contrarily, this

bond being internal in  $\alpha_{s1}$ -CN, it was not surprising that it had been cut by an endopeptidase activity.

Three new bonds were also hydrolyzed in our study: I<sup>6</sup>–K<sup>7</sup>, K<sup>7</sup>–H<sup>8</sup>, and G<sup>10</sup>–L<sup>11</sup>. The cleavage of the I<sup>6</sup>–K<sup>7</sup> bond of  $\alpha_{s1}$ -CN (hydrolyzed only by strain 4F44) was already reported for *Lb. helveticus* CNRZ32,<sup>31</sup> CNRZ303, and CP790.<sup>28</sup> The K<sup>7</sup>–H<sup>8</sup> bond of  $\alpha_{s1}$ -CN was hydrolyzed by five of the six H-strains considered. This is the first time that this bond is reported as being cleaved by a CEP of a LAB. Nevertheless, the importance of this cleavage site could not be evaluated because the generated peptides were identified but not quantified. The G<sup>10</sup>–L<sup>11</sup> bond was hydrolyzed only by strain PB385.

The correspondence between the major and medium cleavage sites observed with PrtS of strain CNRZ 385 on peptide  $\alpha_{s1}$ -CN(f1–23), and those observed with the proteolytic system of the six H-strains studied here on the region 1–23 of  $\alpha_{s1}$ -CN (the hydrolysis of the major site N<sup>17</sup>–E<sup>18</sup> was lacking only for strain Y4), led to the conclusion that PrtS of these strains was probably responsible for the hydrolysis of CNs reported here. Knowing that strains PB302, PB385, and LMD-9 and M-strain HAD8 $\alpha$  possess the same *prtS* allele, how can the fact that strain LMD-9 generated 88 peptides, of which 20 were specific, from  $\beta$ -CN be explained, whereas strains PB302 and PB385 generated 63 and 66 peptides, of which 5 and 6 were specific, respectively? One hypothesis is to consider that all of these strains did not display the same level of expression of PrtS, leading to different enzyme/substrate ratios between the strains. Such an assumption is based on the results of Galia et al.,<sup>16</sup> who observed that the M-strains possessing the *prtS* gene expressed the protease PrtS at a lower level than the H-strains. Thus, if the level of expression of PrtS could vary between the M- and H-strains, it may be possible that it could be also differ within the H-strain group.

***S. thermophilus* Is Able To Release Bioactive Peptides from Caseins.** Hydrolysis of  $\beta$ -,  $\alpha_{s2}$ -, and  $\alpha_{s1}$ -CN by the proteolytic system of *S. thermophilus* H-strains led to approximately 250 peptides. As the bioactive peptides can exert numerous different effects on the body<sup>3,32</sup> and no rapid method of screening of the different bioactivities exists, the sequence of the peptides obtained in this work was compared to that of the bioactive peptides published in the literature or listed in the “database of biologically active peptide sequences”.<sup>33</sup> Among the peptides generated by the tested *S. thermophilus* strains, 13 peptides from  $\beta$ -CN, 5 from  $\alpha_{s2}$ -CN, and 2 from  $\alpha_{s1}$ -CN were listed as bioactive. Although many bioactive peptides from  $\alpha_{s1}$ -CN have been described in the literature, only two bioactive peptides were identified in this study. It must be noted that this work focused on the presence of bioactive peptides and not on their quantification. Thus, some of them can be present in low concentration, but the overall biological activity of a hydrolysate may also result from additive effects.

**Angiotensin I-Converting Enzyme (ACE) Inhibitors.** Fifteen peptides were ACE inhibitors (Table 6). Nine of these peptides were identified in  $\beta$ -CN hydrolysates, five were identified in  $\alpha_{s2}$ -CN hydrolysates, and one was identified in  $\alpha_{s1}$ -CN hydrolysates (Table 6). Finally, all of the bioactive peptides resulting from  $\alpha_{s2}$ -CN hydrolysis were ACE inhibitors, one being also an antioxidant. It must be noted that the peptides generated from  $\alpha_{s2}$ -CN were often specific peptides, generated by one or a maximum of three strains of *S. thermophilus* (Table 4).

**Other Bioactive Peptides.** Several other potential bioactive peptides were detected in the hydrolysis media of  $\beta$ -CN (Table

**Table 6. Potential Bioactive Peptides Generated by the Proteolytic System of Various *S. thermophilus* Strains from Bovine  $\beta$ -,  $\alpha_2$ -, and  $\alpha_1$ -Caseins after 2 or 4 h of Hydrolysis<sup>a</sup>**

source	sequence	strains	activity	ref	
$\beta$ -CN	98–105	VKEAMAPK	4F44	antioxidant	34
	168–175	SKVLPVPQ	all	ACE inhibitor	35
	169–175	KVLPVPQ	all	ACE inhibitor	36
	169–176	KVLPVPQK	ATE19PB8, 4F44, Y4, LMD-9	LOX inhibitor	37
	183–190	RDMPIQAF	all	ACE inhibitor	35
	191–209	LLYQEPVLPVVRGPFPIIV	all	ACE inhibitor and immunomodulating	35, 38
	192–209	LYQEPVLPVVRGPFPIIV	all	mitogene	39
	193–198	YQEPVL	all	ACE inhibitor	40
	193–202	YQEPVLPVVR	all	ACE inhibitor	35, 41
	193–209	YQEPVLPVVRGPFPIIV	all	immunomodulating and antibacterial	2, 42
	199–209	GPVVRGPFPIIV	all	ACE inhibitor	43
	201–209	VRGPFPIIV	Y4, LMD-9	ACE inhibitor	44
	203–209	GPFPIIV	all	ACE inhibitor	43
	$\alpha_2$ -CN	174–181	FALPQYLK	PB385	ACE inhibitor
189–197		AMKPWQPK	PB385, PB302, 4F44	ACE inhibitor	36
190–197		MKPWQPK	PB385, PB302, 4F44	ACE inhibitor	36
198–202		TKVIP	PB302, Y4, 4F44	ACE inhibitor	36
202–207		PVVRYL	PB385	ACE inhibitor and antioxidant	46
$\alpha_1$ -CN	1–9	RPKHPIKHQ	all	ACE inhibitor	47
	1–23	RPKHPIKHQGLPQEVLNENLLRF	LMD-9, PB385, 4F44, Y4	immunomodulating and antibacterial	48

<sup>a</sup>“all” means that the peptide was produced by the six H-strains tested in this study. ACE, angiotensin-converting enzyme (ACE transforms angiotensin I to angiotensin II and also hydrolyzes bradykinin into inactive peptides); LOX, lipoxygenase (lipoxygenase is a dioxygenase that catalyzes oxygenation of certain polyunsaturated fatty acids).

6). Thus, the peptide  $\beta$ -CN(f98–105) displays an antioxidant activity<sup>34</sup> and the peptide  $\beta$ -CN(f169–176) inhibits the lipoxygenase activity.<sup>37</sup> Some of the highlighted bioactive peptides may act on the immune system because the peptide  $\beta$ -CN(f193–209) was described as both an immunomodulator and an antibacterial<sup>2,42</sup> and the peptide  $\beta$ -CN(f191–209), already described as an ACE inhibitor, was also characterized as an immunomodulating peptide.<sup>38</sup> The peptide  $\alpha_1$ -CN(f1–23), which has been first described by Lahov and Regelson<sup>48</sup> as a antimicrobial peptide called “isracidin”, is also considered to be an immunomodulating peptide.<sup>2</sup> Finally, the peptide  $\beta$ -CN(f192–209) was labeled as potentially mitogenic.<sup>39</sup>

The potential bioactive peptides cited above were generated by all of the H-strains tested, except isracidin,  $\beta$ -CN(f98–105), and  $\beta$ -CN(f169–176). Isracidin was produced by strains PB385, 4F44, Y4, and LMD-9,  $\beta$ -CN(f98–105) by strain 4F44, and  $\beta$ -CN(f169–176) by strains ATE19PB8, 4F44, Y4, and LMD-9 (Table 6). The peptide  $\beta$ -CN(f191–209) present at 2 h of hydrolysis was further degraded by the four strains ATE19PB8, 4F44, Y4, and LMD-9.

Thus, all bioactive peptides, except one, highlighted in this work appeared to be resistant to hydrolysis by the proteolytic system of *S. thermophilus*, confirming the results of Somkuti and Paul,<sup>49</sup> who showed that isracidin remained intact after 4 h of incubation with *S. thermophilus* strains but was degraded by exposure to *Lb. delbrueckii* ssp. *bulgaricus* strains.

Numerous studies have been conducted on bioactive peptides generated from fermentations with *Lactococcus* or *Lactobacillus* strains. Studies on the release of bioactive peptides (or their precursors) from CNs by *S. thermophilus* in coculture with *Lactobacillus* (with the possible addition of enzymes) were also carried out.<sup>7,50,51</sup> To our knowledge, very few studies dealt with the genesis of peptides from the hydrolysis of CNs with *S.*

*thermophilus* alone.<sup>12</sup> Among the bioactive peptides released by *S. thermophilus*, some of them are also generated from CNs by other LAB. Thus, the peptides  $\beta$ -CN(f191–209),  $\beta$ -CN(f192–209), and  $\beta$ -CN(f199–209) are produced by *Lc. lactis*, the peptide  $\alpha_1$ -CN(f1–9) is produced by *Lb. helveticus*, and the peptides  $\beta$ -CN(f168–175),  $\beta$ -CN(f169–175), and  $\beta$ -CN(f183–190) are produced by both bacteria.<sup>52–55</sup> We considered only the bioactive peptides with a proven activity, especially those listed in the “database of biologically active peptide sequences”,<sup>33</sup> but the real number of bioactive peptides may be higher. Moreover, the peptides with  $m/z < 400$  were not identified by the LC-ESI-MS/MS analysis, which explains that the reported peptides have a sequence of at least five residues. The identified peptides could also be the precursors of bioactive peptides that could be released in vivo by gastrointestinal enzymes. Thus, Parrot et al.<sup>56</sup> suggested that digestive enzymes diminished the ACE inhibition activity of the peptides detected in Emmental cheese, whereas the digestion of peptides of higher molecular weight would generate new ACE inhibitory peptides. New techniques in vitro would be required to rapidly evaluate the biological activity of a hydrolysate, but finally, only tests on animals would give a global knowledge of the bioactive effect taking into account the presence of antagonist(s) or agonist(s).

The aim of this work was to determine (i) whether the proteolytic system of *S. thermophilus* is able to generate bioactive peptides from some bovine CNs; (ii) whether such a capacity could vary among the *S. thermophilus* strains; and (iii) whether the CEP PrtS could account for such a production, at least in part. The results obtained showed that *S. thermophilus* is able to generate numerous peptides from CNs and that some of them are strain-specific. Among the peptides, several have bioactive potential, leading to the conclusion that *S.*

*thermophilus*, which is widely used in dairy product manufacture, could be used for the release of bioactive peptides from CNs during processing. Finally, we showed that the protease PrtS could account at least in part for the observed production of the peptides from CNs and, so, for the bioactive peptides detected.

This work poses several questions. The first concerns the origin of the specific peptides that are produced by few strains. Is PrtS responsible for this variability and how? The second question concerns the stability of the bioactive peptides during the storage of the products and their bioavailability for the consumer. Finally, the third concerns the real number of bioactive peptides that are generated by the *S. thermophilus* strains and the possibility of driving their liberation by modifying the proteolytic system of *S. thermophilus*.

#### Abbreviations Used

ACE, angiotensin I-converting enzyme; CN, casein; H, high acidifying; L, low acidifying; LAB, lactic acid bacteria; M, medium acidifying.

#### AUTHOR INFORMATION

##### Corresponding Author

\*Postal address: UR AFPA, Équipe PB2P, Faculté des Sciences et Technologies, Université Henri Poincaré – Nancy 1, B.P. 70239, 54506 Vandœuvre-lès-Nancy Cedex, France. Phone: +33 (0)3 83 68 42 67. Fax: +33 (0)3 83 68 42 74. E-mail: Annie.Dary@scbiol.uhp-nancy.fr.

##### Author Contributions

<sup>||</sup>These authors contributed equally to this work.

##### Funding

This work was supported by Mission scientifique Syndifrais – CNIEL.

#### ACKNOWLEDGMENTS

We thank Oun-Ki Chang for his valuable assistance and Maira Junjua for her critical reading and corrections of our manuscript.

#### REFERENCES

- (1) Smacchi, E.; Gobetti, M. Bioactive peptides in dairy products: synthesis and interaction with proteolytic enzymes. *Food Microbiol.* **2000**, *17*, 129–141.
- (2) Silva, S. V.; Málcata, F. X. Caseins as source of bioactive peptides. *Int. Dairy J.* **2005**, *15*, 1–15.
- (3) Korhonen, H. Milk-derived bioactive peptides: from science to applications. *J. Funct. Food* **2009**, *1*, 177–187.
- (4) Phelan, M.; Aherne, A.; FitzGerald, R. J.; O'Brien, N. M. Casein-derived bioactive peptides: biological effects, industrial uses, safety aspects and regulatory status. *Int. Dairy J.* **2009**, *19*, 643–654.
- (5) Kuwabara, Y.; Nagai, S.; Yoshimitsu, N.; Nakagawa, L.; Watanabe, Y.; Tamai, Y. Antihypertensive effect of the milk fermented by culturing various lactic acid bacteria and a yeast. *J. Ferment. Bioeng.* **1995**, *80*, 294–295.
- (6) Gobetti, M.; Stepaniak, L.; De Angelis, M.; Corsetti, A.; Di Cagno, R. Latent bioactive peptides in milk proteins: proteolytic activation and significance in dairy processing. *Crit. Rev. Food Sci. Nutr.* **2002**, *42*, 223–239.
- (7) Donkor, O. N.; Henriksson, A.; Singh, T. K.; Vasiljevic, T.; Shah, N. P. ACE-inhibitory activity of probiotic yoghurt. *Int. Dairy J.* **2007**, *17*, 1321–1331.
- (8) Pihlanto, A.; Virtanen, T.; Korhonen, H. Angiotensin I converting enzyme (ACE) inhibitory activity and antihypertensive effect of fermented milk. *Int. Dairy J.* **2010**, *20*, 3–10.

(9) Sadat-Mekmene, L.; Genay, M.; Atlan, D.; Lortal, S.; Gagnaire, V. Original features of cell-envelope proteinases of *Lactobacillus helveticus*. A review. *Int. J. Food Microbiol.* **2011**, *146*, 1–13.

(10) Delorme, C. Safety assessment of dairy microorganisms *Streptococcus thermophilus*. *Int. J. Food Microbiol.* **2008**, *126*, 274–277.

(11) Hansen, E. B. Microorganisms. In *Handbook of Food and Beverage Fermentation Technology*; Hui, Y. H., Meunier-Goddik, L., Hansen, A. S., Josephsen, J., Nip, W.-K., Stanfield, P. S., Toldrá, F., Eds.; Dekker: New York, 2004; pp 10–25.

(12) Courtin, P.; Rul, F. Interactions between microorganisms in a simple ecosystem: yogurt bacteria as a study model. *Lait* **2004**, *84*, 125–134.

(13) Fernandez-Espla, M. D.; Garault, P.; Monnet, V.; Rul, F. *Streptococcus thermophilus* cell wall-anchored proteinase: release, purification, and biochemical and genetic characterization. *Appl. Environ. Microbiol.* **2000**, *66*, 4772–4778.

(14) Courtin, P.; Monnet, V.; Rul, F. Cell-wall proteinases PrtS and PrtB have a different role in *Streptococcus thermophilus*/*Lactobacillus bulgaricus* mixed cultures in milk. *Microbiology* **2002**, *148*, 3413–3421.

(15) Shahbal, S.; Hemme, D.; Desmazeaud, M. High cell wall-associated proteinase activity of some *Streptococcus thermophilus* strains (H-strains) correlated with a high acidification rate in milk. *Lait* **1991**, *71*, 351–357.

(16) Galia, W.; Perrin, C.; Genay, M.; Dary, A. Variability and molecular typing of *Streptococcus thermophilus* strains displaying different proteolytic and acidifying properties. *Int. Dairy J.* **2009**, *19*, 89–95.

(17) Rasmussen, T.; Danielsen, M.; Valina, O.; Garrigues, C.; Johansen, E.; Pedersen, M. *Streptococcus thermophilus* core genome: comparative genome hybridization study of 47 strains. *Appl. Environ. Microbiol.* **2008**, *74*, 4703–4710.

(18) Delorme, C.; Bartholini, C.; Bolotine, A.; Ehrlich, S. D.; Renault, P. Emergence of a cell wall protease in the *Streptococcus thermophilus* population. *Appl. Environ. Microbiol.* **2010**, *76*, 451–460.

(19) Bolotin, A.; Quinquis, B.; Renault, P.; Sorokin, A.; Dusko Ehrlich, S.; Kulakauskas, S.; Lapidus, A.; Goltsman, E.; Mazur, M.; Pusch, G. D.; Fonstein, M.; Overbeek, R.; Kyprides, N.; Purnelle, B.; Prozzi, D.; Ngui, K.; Masuy, D.; Hancy, F.; Burteau, S.; Boutry, M.; Delcour, J.; Goffeau, A.; Hols, P. Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nat. Biotechnol.* **2004**, *22*, 1554–1558.

(20) Makarova, K.; Slesarev, A.; Wolf, Y.; Sorokin, A.; Mirkin, B.; Koonin, E.; Pavlov, A.; Pavlova, N.; Karamychev, V.; Polouchine, N.; Shakhova, V.; Grigoriev, L.; Lou, Y.; Rohksar, D.; Lucas, S.; Huang, K.; Goodstein, D. M.; Hawkins, T.; Plengvidhya, V.; Welker, D.; Hughes, J.; Goh, Y.; Benson, A.; Baldwin, K.; Lee, J. H.; Diaz-Muniz, L.; Dosti, B.; Smeianov, V.; Wechter, W.; Barabote, R.; Lorca, G.; Altermann, E.; Barrangou, R.; Ganesan, B.; Xie, Y.; Rawsthorne, H.; Tamir, D.; Parker, C.; Breidt, F.; Broadbent, J.; Hutkins, R.; O'sullivan, D.; Steele, J.; Unlu, G.; Saier, M.; Klaenhammer, T.; Richardson, P.; Kozyavkin, S.; Weimer, B.; Mills, D. Comparative genomics of the lactic acid bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 15611–15616.

(21) Terzaghi, B. E.; Sandine, W. E. Improved medium for lactic Streptococci and their bacteriophages. *J. Appl. Microbiol.* **1975**, *29*, 807–813.

(22) Girardet, J.-M.; Humbert, G.; Creusot, N.; Chardot, V.; Campagna, S.; Courthaudon, J.-L.; Gaillard, J.-L. Dilational rheology of mixed  $\beta$ -casein/Tween 20 and  $\beta$ -casein (f114–169)/Tween 20 films at oil-water interface. *J. Colloid Interface Sci.* **2001**, *243*, 515–522.

(23) Tauzin, J.; Micol, L.; Roth, S.; Mollé, D.; Gaillard, J.-L. Tryptic hydrolysis of bovine  $\alpha_2$ -casein: identification and release kinetics of peptides. *Int. Dairy J.* **2003**, *13*, 15–27.

(24) Kaspari, A.; Diefenthal, T.; Grosche, G.; Schierhorn, A.; Demuth, H.-U. Substrates containing phosphorylated residues adjacent to proline decrease the cleavage by proline-specific peptidases. *Biochim. Biophys. Acta* **1996**, *1293*, 147–153.

(25) Kyte, J.; Doolittle, R. F. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **1982**, *157*, 105–132.

- (26) Sadat-Mekmene, L.; Jardin, J.; Corre, C.; Mollé, D.; Richoux, R.; Delage, M. M.; Lortal, S.; Gagnaire, V. Simultaneous presence of PrtH and PrtFH2 proteinases in *Lactobacillus helveticus* strains improves breakdown of the pure  $\alpha_1$ -casein. *Appl. Environ. Microbiol.* **2011**, *77*, 179–186.
- (27) Kumosinski, T. F.; Brown, E. M.; Farrell, H. M. Jr. Three-dimensional molecular modeling of bovine caseins:  $\alpha_1$ -casein. *J. Dairy Sci.* **1991**, *74*, 2889–2895.
- (28) Kunji, E. R. S.; Mierau, L.; Hagting, A.; Poolman, B.; Konings, W. N. The proteolytic systems of lactic acid bacteria. *Antonie van Leeuwenhoek* **1996**, *70*, 187–221.
- (29) Tsakalidou, E.; Anastasiou, R.; Vandenberghe, I.; Van Beeumen, J.; Kalantzopoulos, G. Cell-wall-bound proteinase of *Lactobacillus delbrueckii* subsp. *lactis* ACA-DC 178: characterization and specificity for  $\beta$ -casein. *Appl. Environ. Microbiol.* **1999**, *65*, 2035–2040.
- (30) Exterkate, F. A.; Alting, A. C.; Bruinenberg, P. G. Diversity of cell envelope proteinase specificity among strains of *Lactococcus lactis* and its relationship to charge characteristics of the substrate-binding region. *Appl. Environ. Microbiol.* **1993**, *59*, 3640–3647.
- (31) Pederson, J. A.; Mileski, G. J.; Weimer, B. C.; Steele, J. L. Genetic characterization of a cell envelope-associated proteinase from *Lactobacillus helveticus* CNRZ32. *J. Bacteriol.* **1999**, *181*, 4592–4597.
- (32) Hartmann, R.; Meisel, H. Food-derived peptides with biological activity: from research to food applications. *Curr. Opin. Biotechnol.* **2007**, *18*, 163–169.
- (33) Dziuba, J.; Minkiewicz, P.; Nalecz, D.; Iwaniak, A. Database of biologically active peptide sequences. *Nahrung* **1999**, *43*, 190–195.
- (34) Korhonen, H.; Pihlanto, A. Technological options for the production of health-promoting proteins derived from milk and colostrum. *Curr. Pharm. Des.* **2007**, *13*, 829–843.
- (35) Yamamoto, N.; Akino, A.; Takano, T. Antihypertensive effect of peptides derived from casein by an extracellular proteinase from *Lactobacillus helveticus* CP790. *J. Dairy Sci.* **1994**, *77*, 917–922.
- (36) Maeno, M.; Yamamoto, N.; Takano, T. Identification of an antihypertensive peptide from casein hydrolysate produced by a proteinase from *Lactobacillus helveticus* CP790. *J. Dairy Sci.* **1996**, *79*, 1316–1321.
- (37) Rival, S.; Boeriu, C. G.; Wichers, H. J. Caseins and casein hydrolysates. 2. Antioxidative properties and relevance to lipoxygenase inhibition. *J. Agric. Food Chem.* **2001**, *49*, 295–302.
- (38) Hayes, M.; Stanton, C.; Fitzgerald, G. F.; Ross, R. P. Putting microbes to work: dairy fermentation, cell factories and bioactive peptides. Part II: Bioactive peptide functions. *Biotechnol. J.* **2007**, *2*, 435–449.
- (39) Coste, M.; Rochet, V.; Léonil, J.; Mollé, D.; Bouhallab, S.; Tomé, D. Identification of C-terminal peptides from  $\beta$ -casein that enhance proliferation of rat lymphocytes. *Immunol. Lett.* **1992**, *33*, 41–46.
- (40) Pihlanto-Leppälä, A.; Rokkaa, T.; Korhonen, H. Angiotensin I converting enzyme inhibitory peptides derived from bovine milk proteins. *Int. Dairy J.* **1998**, *8*, 325–331.
- (41) Rokka, T.; Syväoja, E.-L.; Tuominen, J.; Korhonen, H. Release of bioactive peptides by enzymatic proteolysis of *Lactobacillus* GG fermented UHT milk. *Milchwissenschaft* **1997**, *52*, 675–678.
- (42) Sandré, C.; Gleizes, A.; Forestier, F.; Gorges-Kergot, R.; Chilmonec, S.; Léonil, J.; Moreau, M.-C.; Labarre, C. A peptide derived from bovine  $\beta$ -casein modulates functional properties of bone marrow-derived macrophages from germfree and human flora-associated mice. *J. Nutr.* **2001**, *131*, 2936–2942.
- (43) Hayes, M.; Stanton, C.; Slattey, H.; O'Sullivan, O.; Hill, C.; Fitzgerald, G. F.; Ross, R. P. Casein fermentate of *Lactobacillus animalis* DPC6134 contains a range of novel propeptide angiotensin-converting enzyme inhibitors. *Appl. Environ. Microbiol.* **2007**, *73*, 4658–4667.
- (44) Miguel, M.; Alexandre, M. A.; Ramos, M.; López-Fandiño, R. Effect of simulated gastrointestinal digestion on the antihypertensive properties of ACE-inhibitory peptides derived from ovalbumin. *J. Agric. Food Chem.* **2006**, *54*, 726–731.
- (45) Tauzin, J.; Miclo, L.; Gaillard, J.-L. Angiotensin-I-converting enzyme inhibitory peptides from tryptic hydrolysate of bovine  $\alpha_2$ -casein. *FEBS Lett.* **2002**, *531*, 369–374.
- (46) López-Expósito, I.; Quirós, A.; Amigo, L.; Recio, I. Casein hydrolysates as a source of antimicrobial, antioxidant and antihypertensive peptides. *Dairy Sci. Technol.* **2007**, *87*, 241–249.
- (47) Saito, T.; Nakamura, T.; Kitazawa, H.; Kawai, Y.; Itoh, T. Isolation and structural analysis of antihypertensive peptides that exist naturally in Gouda cheese. *J. Dairy Sci.* **2000**, *83*, 1434–1440.
- (48) Lahov, E.; Regelson, W. Antibacterial and immunostimulating casein-derived substances from milk: casecidin, isracidin peptides. *Food Chem. Toxicol.* **1996**, *34*, 131–145.
- (49) Somkuti, G. A.; Paul, M. Enzymatic fragmentation of the antimicrobial peptides casecidin and isracidin by *Streptococcus thermophilus* and *Lactobacillus delbrueckii* ssp. *bulgaricus*. *Appl. Microbiol. Biotechnol.* **2010**, *87*, 235–242.
- (50) Schieber, A.; Brückner, H. Characterization of oligo- and polypeptides isolated from yoghurt. *Eur. Food Res. Technol.* **2000**, *210*, 310–313.
- (51) Tsai, J. S.; Chen, T. J.; Pan, B. S.; Gong, S. D.; Chung, M. Y. Antihypertensive effect of bioactive peptides produced by protease-facilitated lactic acid fermentation of milk. *Food Chem.* **2008**, *106*, 552–558.
- (52) Juillard, V.; Laan, H.; Kunji, E. R. S.; Jeronimus-Stratingh, C. M.; Bruins, A. P.; Konings, W. N. The extracellular P(1)-type proteinase of *Lactococcus lactis* hydrolyzes  $\beta$ -casein into more than one hundred different oligopeptides. *J. Bacteriol.* **1995**, *177*, 3472–3478.
- (53) Kunji, E. R. S.; Fang, G.; Jeronimus-Stratingh, C. M.; Bruins, A. P.; Poolman, B.; Konings, W. N. Reconstruction of the proteolytic pathway for use of  $\beta$ -casein by *Lactococcus lactis*. *Mol. Microbiol.* **1998**, *27*, 1107–1118.
- (54) Otte, J.; Lenhard, T.; Flambard, B.; Sørensen, K. I. Influence of fermentation temperature and autolysis on ACE-inhibitory activity and peptide profiles of milk fermented by selected strains of *Lactobacillus helveticus* and *Lactococcus lactis*. *Int. Dairy J.* **2011**, *21*, 229–238.
- (55) Robert, M.-C.; Razaname, A.; Mutter, M.; Juillerat, M. A. Identification of angiotensin-I-converting enzyme inhibitory peptides derived from sodium caseinate hydrolysates produced by *Lactobacillus helveticus* NCC 2765. *J. Agric. Food Chem.* **2004**, *52*, 6923–6931.
- (56) Parrot, S.; Degraeve, P.; Curia, C.; Martial-Gros, A. *In vitro* study on digestion of peptides in Emmental cheese: analytical evaluation and influence on angiotensin I converting enzyme inhibitory peptides. *Nahrung* **2003**, *47*, 87–94.

## **Chapitre II**



# 1. Recherche de systèmes de transport de peptides au sein d'une collection de souches de *Streptococcus thermophilus* par PCR

Comme cela a été décrit précédemment, nous avons montré que le système protéolytique de surface de *S. thermophilus* est capable de générer des peptides bioactifs à partir des caséines bovines  $\beta$ ,  $\alpha_{s2}$  et  $\alpha_{s1}$ . Néanmoins, l'utilisation de cette bactérie pour la production de tels peptides dans l'aliment requiert qu'elle en internalise le moins possible. Dans les conditions utilisées précédemment, peu ou pas de peptides générés ont été internalisés mais cette situation ne peut se produire lors de la fabrication d'un lait fermenté. Aussi, nous nous sommes intéressés aux systèmes de transport de peptides. Le système Ami ayant été très étudié chez *S. thermophilus*, nous avons étudié la variabilité génétique de ce système *ami*, du transporteur *dpT* et d'étudier la variabilité et le rôle du cluster *ots* découvert lors du séquençage du génome de la souche LMD-9. Vingt deux souches ont été choisies pour cette étude ; 19 appartiennent à la collection du laboratoire, ont été isolées de yaourts ou de fromages et possèdent des caractéristiques phénotypiques et génétiques différentes (Galia *et al.*, 2009) ; les 3 autres souches (LMG18311, ND03 et JIM8232) correspondent à des souches dont le génome a été séquencé et est disponible dans les banques de données. Sur ces dernières, les caractéristiques phénotypiques n'ont pas été déterminées (Tableau 9). Au niveau phénotypique, les 19 souches appartenant à la collection du laboratoire présentent les caractéristiques suivantes : 10 souches ont une forte activité caséinolytique de surface (PrtS<sup>+</sup>), 2 une légère activité protéolytique (PrtS<sup>+/-</sup>) et 7 souches en sont dépourvues (PrtS<sup>-</sup>). En termes de capacités acidifiantes, ces souches ont été classées en trois groupes : fort, moyen et faible pouvoir acidifiant. Il est à noter que toutes les souches présentant un fort pouvoir acidifiant ont également une forte activité protéolytique (PrtS<sup>+</sup>). La variabilité génotypique des souches a été évaluée par l'analyse de la séquence intergénique ITS 16S-23S. Toutes les souches ayant une forte activité protéolytique (PrtS<sup>+</sup>) et un fort pouvoir acidifiant ont une région intergénique 16S-23S de type II.

## **1. 1. Etablissement des relations phylogénétiques entre 22 souches de *Streptococcus thermophilus***

Dans un premier temps, nous avons cherché à établir les proximités phylogénétiques existant entre les 22 souches étudiées. Le séquençage de l'ITS 16S-23S des souches de la collection du laboratoire a permis de distinguer 2 groupes différents, mais à l'intérieur de ces groupes, la proximité phylogénétique entre les souches n'a pu être évaluée. En effet, si l'analyse de la séquence intergénique ITS 16S-23S permet de distinguer les différentes espèces et sous-espèces (Daffonchio *et al.*, 2003 ; Mora *et al.*, 2003), elle ne permet pas toujours de distinguer les souches. A l'inverse, le typage moléculaire par MLST peut permettre de distinguer entre souches possédant le même ITS 16S-23S, et donc d'évaluer la diversité génétique à l'intérieur d'un même groupe et de proposer une classification phylogénétique des souches, le cas échéant. Le typage par MLST consiste à établir la séquence de portions de « gènes de ménage », indispensables au métabolisme cellulaire, présents dans le « core génome » et répartis sur le chromosome.

Huit « gènes de ménage » (*glcK*, *ddlA*, *pepO*, *ilvC*, *thrS*, *tkl*, *pyrE*, *dnaE*), utilisés par Delorme *et al.* (2010) lors d'une étude utilisant la MLST, ont été choisis pour notre analyse. Dans un premier temps, des portions de 531 à 721 pb des huit « gènes de ménage » ont été amplifiées par PCR chez 17 souches de notre collection à l'aide des couples d'amorces présentés dans les tableaux 15 et 17 et séquencés. Les séquences de ces gènes chez les 5 souches LMD-9, CNRZ1066, LMG18311, JIM8232 et ND03 (génomés entièrement caractérisés), ont été obtenues à partir des banques de données (<http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi>). Les résultats de l'étude MLST, obtenus pour l'ensemble des gènes des 22 souches (17 souches provenant de la collection et 5 souches dont le génome est entièrement séquencé) sont présentés dans le tableau 19. Pour chaque gène, et à chaque séquence unique, un numéro d'allèle a été assigné. Ce numéro correspond à ceux attribués par Delorme *et al.* (2010) pour les séquences identiques, et dans le cas d'une séquence différente, un nouveau numéro d'allèle a été attribué. Pour le gène *glcK*, codant la glucose kinase, 4 allèles sur les 7 listés par Delorme *et al.* (2010) ont été retrouvés. Il s'agit des allèles *glcK*-1, -2, -3 et -5 avec une prédominance de l'allèle 1 au sein de notre collection de souches et celles des banques de données puisqu'il est présent dans 15 souches des 22 référencées ici (Tableau 19). Au contraire, l'allèle *glcK*-3 n'est présent que chez la souche *S. thermophilus* ST14.

Pour le gène *ddlA*, codant la D-alanine D-alanine ligase, deux allèles répertoriés par Delorme *et al.* (2010) ont été détectés au sein de notre collection. L'allèle *ddlA-2* est le plus répandu puisque détecté dans 18 souches sur 22. D'autre part, nous avons détecté un nouvel allèle, nommé *ddlA-4*, uniquement retrouvé chez la souche 4F44 (Tableau 19). Pour le gène *pepO*, codant une endopeptidase, 3 allèles (*pepO-1*, -2 et -3) sur les 6 publiés par Delorme *et al.* (2010) ont été détectés. Les allèles 1 et 2 sont majoritaires, l'allèle 3 n'étant présent que chez 2 souches de notre collection (Tableau 19).

**Tableau 19 : Répartition des différents allèles de 8 « gènes de ménage » chez des souches de *Streptococcus thermophilus* de la collection URAFPA et des banques de donnés**

souches/gènes	Numéros d'allèles								Type d'ITS <sup>b</sup>	Gène <i>prtS</i> <sup>b</sup>
	<i>glcK</i>	<i>ddlA</i>	<i>pepO</i>	<i>ilvC</i>	<i>thrS</i>	<i>tkl</i>	<i>pyrE</i>	<i>dnaE</i>		
CNRZ21	5	2	1	1	1	1	3	1	I	-
CNRZ1066	5	2	1	1	1	1	3	1	I	-
CNRZ445	5	2	1	1	1	1	3	1	I	+
CNRZ404	5	2	1	1	1	1	3	1	I	+
HD17	1	2	1	1	1	1	3	1	I	+
ST14	3	2	1	1	1	8	13	1	I	-
ATE11PB6	2	2	1	1	1	3	1	1	I	+
CNRZ160	2	2	1	1	1	3	1	1	I	-
PB180	1	2	3	2	1	3	6	1	I	+
CNRZ385N	1	2	3	2	2	3	8	2	I	+
4F44	1	4 <sup>a</sup>	1	1	2	12 <sup>a</sup>	6	1	I	+
LMD-9	1	1	2	1	2	7	5	1	II	+
PB385	1	2	2	1	2	7	5	1	II	+
ST88	1	2	2	1	2	7	5	1	II	+
ST85	1	2	2	1	2	7	5	1	II	+
PB18	1	2	2	1	2	7	5	1	II	+
ST87	1	2	2	1	2	7	5	1	II	+
PB302	1	2	2	1	2	7	5	1	II	+
CNRZ391	1	2	2	1	2	7	5	1	II	+
ND03	1	1	1	1	1	3	6	1	I	-
LMG18311	1	2	1	1	2	1	3	1	I	-
JIM8232	1	1	1	1	2	7	12	1	I	+

Les numéros d'allèles proposés par Delorme *et al.* (2010) ont été attribués aux allèles de notre collection. Les noms surlignés en gris correspondent aux souches dont le génome a été entièrement séquencé.

<sup>a</sup> Nouvel allèle (cette étude)

<sup>b</sup> Les types d'ITS et la présence/absence du gène *prtS* sont issus de l'étude de Galia *et al.* (2009).

Pour le gène *ilvC* codant la céto-acide réductoisomérase, les deux allèles (*ilvC*-1 et -2) détecté par Delorme *et al.* (2010) ont été retrouvés, l'allèle 2 n'étant présent que chez 2 souches de notre collection, toutes les autres présentant l'allèle *ilvC*-1 (Tableau 19). Pour le gène *thrS*, codant la thréonyl-tRNA synthétase, 2 allèles (*thrS*-1 et -2) sur les 5 signalés par Delorme *et al.* (2010), ont été détectés. Ils sont distribués de façon équivalente au sein de notre étude (Tableau 19). Pour le gène *tkl*, codant la transcétolase, 4 (*tkl*-1, -3, -7 et -8) sur 11 allèles mentionnés par Delorme *et al.* (2010) dans leur étude ont été retrouvés. De plus, un nouvel allèle, nommé *tkl*-12, a été mis en évidence chez la souche 4F44. Seule la souche ST14 présente l'allèle *tkl*-8 (Tableau 19). Pour le gène *pyrE*, codant l'orotate phosphoribosyltransférase, 7 (*pyrE*-1, -3, -5, -6, -8, -12 et -13) des 13 allèles listé par Delorme *et al.* (2010) ont été détectés. L'allèle *pyrE*-5 est l'allèle majoritaire. Les allèles 8, 12 et 13 sont respectivement présents chez les souches CNRZ385N, JIM8232 et ST14 (Tableau 19). Pour le gène *dnaE*, codant l'ADN polymérase III, 2 allèles (*dnaE*-1 et -2) sur les 5 listés par Delorme *et al.* (2010) ont été détecté dans notre étude. Seule la souche CNRZ385N possède l'allèle *dnaE*-2 (Tableau 19).

Ainsi, les résultats de l'analyse des 8 « gènes de ménage » chez les 22 souches de *S. thermophilus* considérées dans cette étude a permis de montrer la présence de 2 à 7 allèles par gène (Tableau 20), contre 2 à 13 pour les 27 souches de *S. thermophilus* analysées par Delorme *et al.* (2010). Même en tenant compte d'un nombre plus restreint de souches analysées dans notre étude, la diversité génétique observée chez les souches du laboratoire est significativement plus faible que celle mise en évidence par Delorme *et al.* (2010).

Les pourcentages de divergence entre les différents allèles de chaque gène de ménage ont été calculés à l'aide du logiciel « clustalW » (% divergence = 100 - % identité), ils sont tous inférieurs à 1%. Le pourcentage de divergence le plus bas (0,2%) est obtenu pour les trois gènes *ilvC*, *thrS* et *dnaE* pour lesquels deux allèles ont été détectés dans cette étude. Les gènes *tkl* et *pyrE* ayant respectivement 5 et 7 allèles différents, conservent un pourcentage de divergence faible (0,8%) (Tableau 21). Hormi les gènes *ddlA*, *pepO*, *ilvC*, pour lesquels les pourcentages de divergence sont identiques à ceux par Delorme *et al.* (2010), pour les autres gènes, les pourcentages de divergence de notre étude sont inférieurs (en raison d'un nombre inférieur d'allèles) (Tableau 22).

Le pourcentage de diversité des souches de notre collection a été déterminé à l'aide du logiciel MEGA\_5.0 et estimé à 0,16%, ce qui est proche de celui (0,19%) obtenu par Delorme *et al.*

(2010). Ces résultats sont en accord avec le faible polymorphisme observé chez l'espèce *S. thermophilus*, espèce clonale émergée récemment avec le début de fabrications de produits laitiers (depuis environ 7000 ans) (Bolotin *et al.*, 2004 ; Delorme *et al.*, 2010 ; Fox, 1993).

**Tableau 20 : Comparaison entre le nombre d'allèles trouvés pour chaque « gène de ménage » dans cette étude et celui obtenu à partir d'une collection de 27 souches par Delorme *et al.* en 2010**

Gène	Nombre d'allèles trouvés chez <i>S. thermophilus</i>	
	dans cette étude	dans la collection de Delorme <i>et al.</i> (2010)
<i>glcK</i>	4	7
<i>ddlA</i>	2+1 <sup>a</sup>	3
<i>pepO</i>	3	6
<i>ilvC</i>	2	2
<i>thrS</i>	2	5
<i>ikt</i>	4+1 <sup>a</sup>	11
<i>pyrE</i>	7	13
<i>dnaE</i>	2	5

<sup>a</sup> Nouvel allèle (cette étude)

**Tableau 21 : Pourcentages de divergence entre les allèles de chaque « gène de ménage » étudié**

<b>gène <i>glcK</i></b>	<b>Allèle 1</b>	<b>Allèle 2</b>	<b>Allèle 3</b>	<b>Allèle 5</b>
Allèle 1	0			
Allèle 2	0,2	0		
Allèle 3	0,4	0,2	0	
Allèle 5	0,2	0,4	0,6	0

<b>gène <i>dlla</i></b>	<b>Allèle 1</b>	<b>Allèle 2</b>	<b>Allèle 4</b>
Allèle 1	0		
Allèle 2	0,2	0	
Allèle 4	0,2	0,4	0

<b>gène <i>pepO</i></b>	<b>Allèle 1</b>	<b>Allèle 2</b>	<b>Allèle 3</b>
Allèle 1	0		
Allèle 2	0,2	0	
Allèle 3	0,2	0,4	0

<b>gène <i>ilvC</i></b>	<b>Allèle 1</b>	<b>Allèle 2</b>
Allèle 1	0	
Allèle 2	0,2	0

<b>gène <i>thrS</i></b>	<b>Allèle 1</b>	<b>Allèle 2</b>
Allèle 1	0	
Allèle 2	0,2	0

<b>gène <i>tkt</i></b>	<b>Allèle 1</b>	<b>Allèle 3</b>	<b>Allèle 7</b>	<b>Allèle 8</b>	<b>Allèle 12</b>
Allèle 1	0				
Allèle 3	0,2	0			
Allèle 7	0,6	0,4	0		
Allèle 8	0,4	0,6	0,6	0	
Allèle 12	0,4	0,2	0,6	0,8	0

<b>gène <i>pyrE</i></b>	<b>Allèle 1</b>	<b>Allèle 3</b>	<b>Allèle 5</b>	<b>Allèle 6</b>	<b>Allèle 8</b>	<b>Allèle 12</b>	<b>Allèle 13</b>
Allèle 1	0						
Allèle 3	0,4	0					
Allèle 5	0,4	0,4	0				
Allèle 6	0,2	0,2	0,2	0			
Allèle 8	0,4	0,4	0,4	0,2	0		
Allèle 12	0,6	0,6	0,2	0,4	0,6	0	
Allèle 13	0,6	0,6	0,6	0,4	0,6	0,8	0

<b>gène <i>dnaE</i></b>	<b>Allèle 1</b>	<b>Allèle 2</b>
Allèle 1	0	
Allèle 2	0,2	0

Les pourcentages de divergence ont été déterminés à l'aide du logiciel « ClustalW »

**Tableau 22 : Tableau synthétique des pourcentages de divergences nucléotidiques maximales de chaque « gène de ménage » au sein de notre collection de souches de *Streptococcus thermophilus* par rapport à ceux obtenus par Delorme *et al.* (2010)**

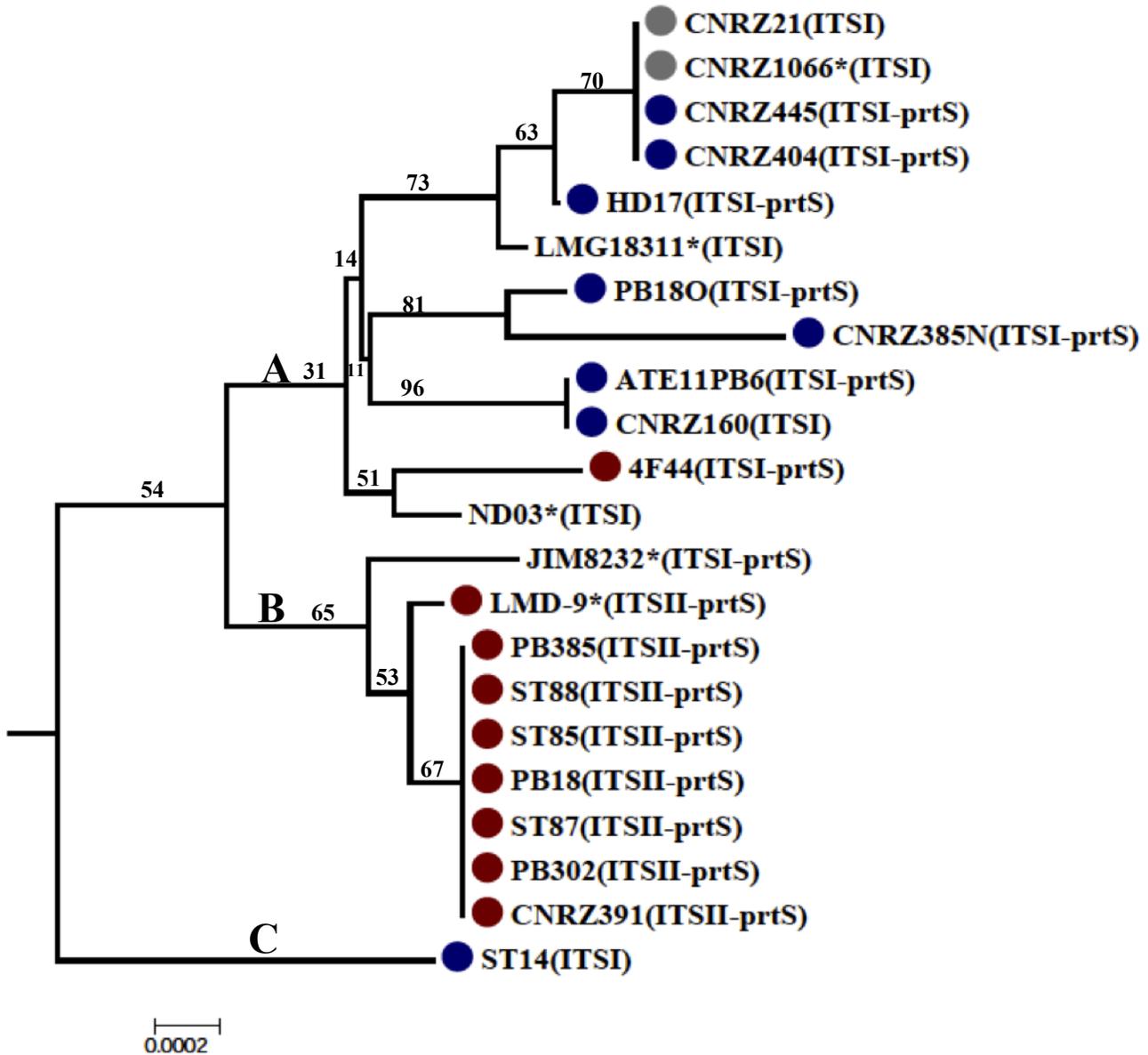
Gène	Pourcentage de divergence maximale obtenue	
	dans cette étude	par Delorme <i>et al.</i> 2010
<i>glcK</i>	0,6	1,7
<i>ddlA</i>	0,4	0,4
<i>pepO</i>	0,4	0,4
<i>ilvC</i>	0,2	0,2
<i>thrS</i>	0,2	0,6
<i>tkt</i>	0,8	1,0
<i>pyrE</i>	0,8	1,2
<i>dnaE</i>	0,2	0,6

Les séquences des 8 « gènes de ménage » des souches de notre collection et de celles des banques de données ont été utilisées pour construire un arbre phylogénétique. Pour cela, les séquences des huit loci et d'ITS 16S-23S ont été concaténées pour chaque souche puis la méthode « Neighbour Joining » a été appliquée à l'aide du logiciel MEGA\_5.0. Les souches étudiées peuvent être principalement réparties en deux-groupes (Figure 25), permettant une séparation nette entre les souches à fort pouvoir acidifiant et les souches à moyen/faible pouvoir acidifiant.

Le premier sous-groupe (branche B) est composé de 9 souches dont huit souches de type ITSII, à forts pouvoirs acidifiant et protéolytique, sept souches (PB302, CNRZ391, ST87, PB18, ST85, ST88 et PB385) ayant les mêmes allèles pour les 8 « gènes de ménage » étudiés, et la souche LMD-9 diffère des souches précédentes par un seul allèle, il s'agit de l'allèle *ddlA*. La souche JIM8232 de type ITSI possède le gène *prtS*, mais ses pouvoirs acidifiant et protéolytique n'ont pas été déterminés. En effet, posséder le gène ne signifie pas l'exprimer comme cela a été montré par Galia *et al.* (2009).

Le deuxième sous-groupe (branches A et C) est composé de souches de type ITSI à moyen ou faible pouvoir acidifiant sauf la souche 4F44 qui est à fort pouvoir acidifiant. Certaines souches ne possèdent pas le gène *prtS*, d'autres le possèdent, mais l'expriment faiblement (souches PrtS<sup>+/-</sup> selon l'étude de Galia *et al.*, 2009). Le pouvoir acidifiant des souches LMG18311 et ND03 n'a pas été déterminé. Cependant, compte tenu des faibles valeurs de « bootstrap » obtenues (inférieures à 90%), l'arbre phylogénétique obtenu reste indicatif. Par

conséquences, l'ajout d'autres « gènes de ménage » dans l'analyse pourrait aider à mieux discriminer les souches de notre collection.



**Figure 25 : Arbre phylogénétique de 22 souches de *Streptococcus thermophilus* basé sur les séquences concaténées de 8 «gènes de ménage» (*glcK*, *ddlA*, *pepO*, *ilvC*, *thrS*, *tkl*, *pyrE*, *dnaE*) et de l'ITS 16S-23S**

L'arbre a été construit en utilisant la méthode « Neighbour Joining » à l'aide du logiciel MEGA\_5.0. L'échelle indique le nombre de substitutions de nucléotides par site. Les lettres de A à C représentent les différents sous-groupes, les étoiles correspondent aux souches dont les génomes sont complètement séquencés. L'absence de « prtS » entre parenthèse renseigne sur le génotype *prtS* de la souche. Les valeurs de « bootstrap » (1000 répétitions) sont indiquées sur les branches. Le pouvoir acidifiant (d'après Galia *et al.*, 2009) en lait de chaque souche de la collection du laboratoire est mentionné par une pastille colorée :

- fort pouvoir acidifiant
- moyen pouvoir acidifiant
- faible pouvoir acidifiant

## 1. 2. Recherche du système Ami

Sachant que le système Ami est présent chez les 5 souches dont le génome est entièrement séquencé et que la présence des gènes *amiA2* et *amiA3* codant chacun une « Oligopeptide Binding Protein » (OBP) varie en fonction des souches de *S. thermophilus*, nous avons recherché par PCR la présence du système Ami au sein des 17 souches étudiées. Pour cela, des amorces spécifiques ont été conçues (Tableaux 13, 16) à partir des séquences des gènes *ami* présents dans les génomes séquencés de *S. thermophilus*. Afin de mettre en évidence la présence ou l'absence des gènes *amiCDEF* organisés en opéron (opéron *amiA1CDEF*) des PCR ont été effectuées avec un couple d'amorces approprié à chaque gène (Tableaux 13, 16). Les résultats montrent que pour chacune des souches testées, une bande correspondant à la taille du produit attendue (environ 600 pb) a été obtenue pour chaque gène chez toutes les souches, ce qui montre que les quatre gènes *amiC*, *amiD*, *amiE*, *amiF* sont présents chez les 17 souches testées.

Pour déterminer la présence ou l'absence des trois paralogues *amiA1*, *amiA2* et *amiA3*, des PCR ont été réalisées avec des couples d'amorces bien spécifiques (Tableaux 13, 16). La conception de ces amorces a été rendue difficile en raison du fort pourcentage d'identité existant d'une part entre les 3 gènes *amiA1*, *amiA2*, *amiA3* (entre 84 et 99% d'identité, Tableau 23), et d'autre part, entre les 2 copies d'*IS1193* situées en aval des gènes *amiA2* et *amiA3* (Figure 26). L'amplification des gènes *amiA1* et *amiA3* a été réalisée avec le même couple d'amorces et le gène *amiA2* avec un deuxième couple d'amorces (Figure 26).

La recherche des gènes *amiA1* et *amiA3*, a été réalisée en utilisant le couple d'amorces « *amiA3* for/*amiC* rev ». Compte tenu du pourcentage d'identité existant entre les gènes *amiA1* et *amiA3*, l'amorce « *amiA3* for » s'hybride à la fois sur les gènes *amiA1* et *amiA3* tandis que l'amorce « *amiC* rev » s'hybride au niveau du gène *amiC* (Figure 26). Ainsi sur gel d'agarose, en fonction des souches, le nombre d'amplicons obtenu diffère :

- en présence du gène *amiA1* et en l'absence du gène *amiA3*, une seule bande correspondant à un amplicon de 1,1 kb est détectée ;
- en cas de présence des deux gènes *amiA1* et *amiA3*, deux bandes correspondant à deux amplicons de 1,1 kb et de 4,7 kb sont visualisées (Figure 27).

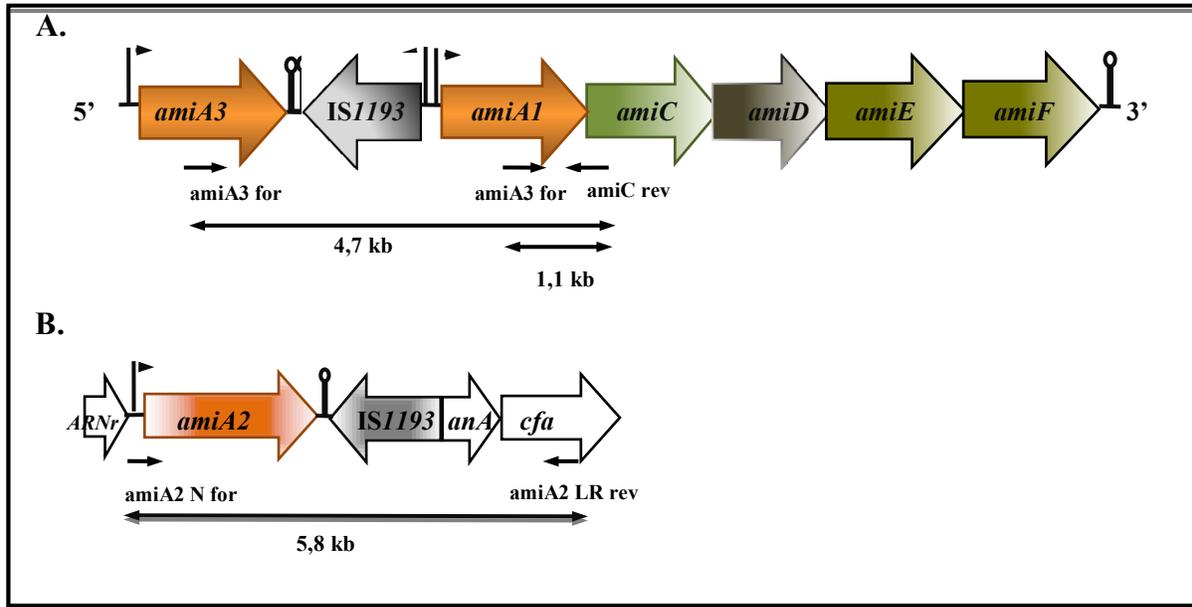
La présence du gène *amiA2* a été recherchée en utilisant le couple d'amorce « *amiA2N* for/*amiA2LR* rev ». L'amorce « *amiA2N* for » s'hybride sur un gène codant un ARN

ribosomique qui se trouve en 5' du gène *amiA2* et l'oligonucléotide « *amiA2LR rev* » s'hybride sur le gène *cfa* codant la « cyclopropane-fatty-acyl-phospholipid synthase », situé en 3' du gène *amiA2*. En présence du gène *amiA2* un amplicon de 5,8 kb est détecté sur gel d'agarose, alors qu'en son absence la taille de l'amplicon est de 1,7 kb (Figure 27).

**Tableau 23 : Pourcentage d'identité entre les gènes *amiA* chez les cinq souches dont le génome a été séquencé**

	LMG18311		CNRZ1066		ND03			LMD-9		JIM8232	
LMG18311	<i>amiA2</i>	<i>amiA3</i>	<i>amiA2</i>	<i>amiA3</i>	<i>amiA1</i>	<i>amiA2</i>	<i>amiA3</i>	<i>amiA1</i>	<i>amiA3</i>	<i>amiA1</i>	<i>amiA3</i>
<i>amiA2</i>	100,0										
<i>amiA3</i>	84,14	100,0									
<b>CNRZ1066</b>											
<i>amiA2</i>	99,9	84,1	100,0								
<i>amiA3</i>	84,9	99,0	84,8	100,0							
<b>ND03</b>											
<i>amiA1</i>	98,9	84,5	98,8	85,2	100,0						
<i>amiA2</i>	99,9	84,2	99,9	84,9	98,9	100,0					
<i>amiA3</i>	83,5	97,4	83,5	97,3	84,0	83,6	100,0				
<b>LMD-9</b>											
<i>amiA1</i>	99,7	84,1	99,6	84,80	99,0	99,7	83,5	100,0			
<i>amiA3</i>	84,0	99,1	84,0	98,6	84,5	84,1	98,0	84,04	100,0		
<b>JIM8232</b>											
<i>amiA1</i>	98,2	85,0	98,2	85,6	99,3	98,3	84,7	98,3	85,2	100,0	
<i>amiA3</i>	83,1	97,6	83,1	97,6	84,4	84,1	99,5	83,1	98,1	85,1	100,0

<sup>a</sup>: *amiA1* est un pseudogène chez les souches LMG18311 et CNRZ1066 (Hols *et al.*, 2005)



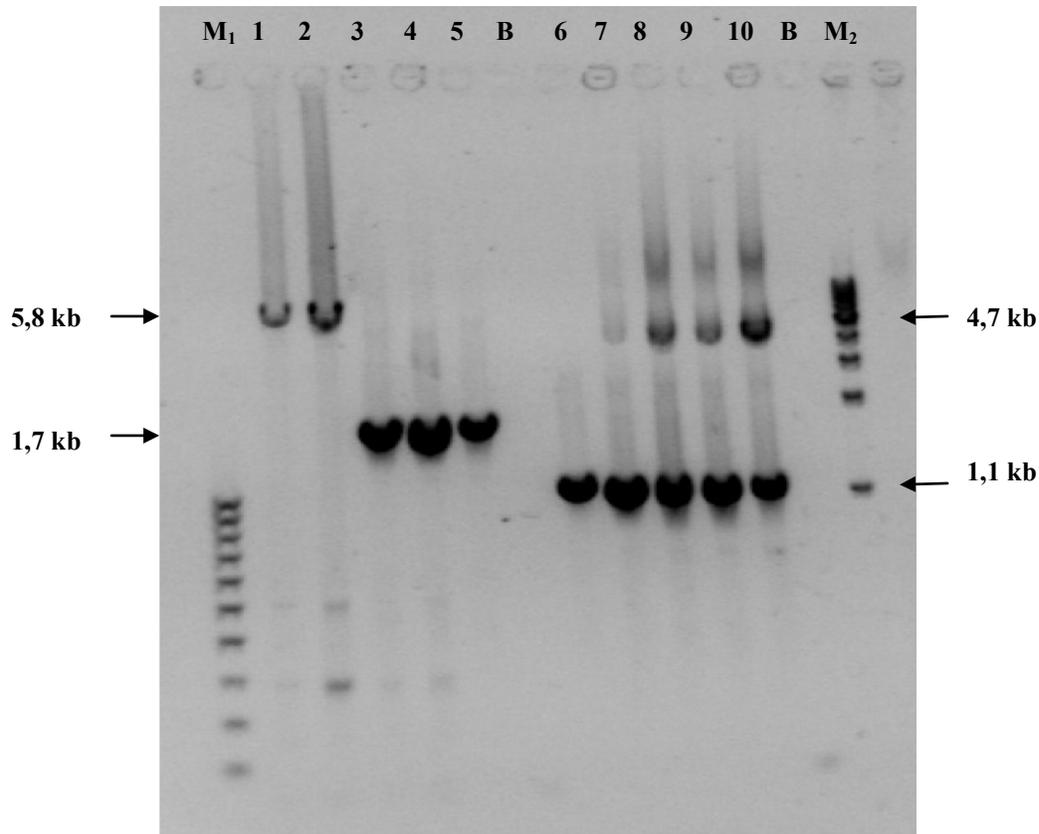
**Figure 26 : Représentation schématique de l'organisation génétique des gènes du système de transport de peptides Ami chez *Streptococcus thermophilus* CNRZ1066, LMG18311, ND03, LMD-9 et JIM8232 et positions des amorces utilisées pour les amplifier par PCR**

A. Opéron *amiA1CDEF* et gène *amiA3* situé en amont de cet opéron,

B. Gène *amiA2* (absent chez les souches LMD-9 et JIM8232).

Les positions d'hybridation de deux couples d'amorces « *amiC* rev/*amiA3* for » et « *amiA2*N for/*amiA2*LR rev » utilisés pour détecter la présence des gènes *amiA1-amiA3* et *amiA2*, respectivement sont indiquées, ainsi que les tailles attendues de chaque amplicon.

*ARNr* : gène codant un ARN ribosomique ; *IS1193* : séquence d'insertion de la famille *ISL3* ; *anaA* : gène codant une aminoacylase/N-acyl-L-amino acid amidohydrolase ; *cfa* : gène codant une cyclopropane-fatty-acyl-phospholipid synthase.



**Figure 27 : Détection par PCR des gènes *amiA2* (pistes 1 à 5), *amiA1* et *amiA3* (pistes 6 à 10) chez différentes souches de *Streptococcus thermophilus***

Les amorces « AmiA2N for/AmiA2LR rev » ont été utilisées pour la détection du gène *amiA2* tandis que les amorces « AmiA3 for/AmiC rev » ont été utilisées pour la détection des gènes *amiA1* et *amiA3*.

Pistes 1 à 5 : la présence d'un amplicon de 5,8 kb signifie que le gène *amiA2* est présent alors que la présence d'un amplicon de 1,7 kb signifie qu'il est absent. Pistes 6 à 10 : l'amplicon de 4,7 kb est attendu si le gène *amiA3* est présent tandis qu'un amplicon de 1,1 kb est attendu dans le cas du gène *amiA1*.

Les puits 1 et 6 correspondent à la souche ST14 ; les puits 2 et 7 à la souche CNRZ160 ; les puits 3 et 8 à la souche 4F44 ; les puits 4 et 9 à la souche ST88 ; les puits 5 et 10 à la souche LMD-9 (témoin).

M<sub>1</sub> : marqueur de taille de 100 pb à 1 kb , M<sub>2</sub> : marqueur de taille de 1 à 10 kb, B : blanc.

L'opéron *amiA1CDEF* est présent chez toutes les souches étudiées (Tableau 24). Trois souches possèdent le gène *amiA2* mais pas le gène *amiA3* ; 11 souches possèdent uniquement le gène *amiA3* ; et 8 souches possèdent les deux (*amiA2* et *amiA3*). La variabilité du nombre de gènes *amiA* a déjà été montrée au sein de l'espèce *S. thermophilus* par Garault *et al.* (2002). Cependant, la PCR permet de détecter la présence ou l'absence des gènes mais elle ne permet pas de savoir s'ils sont fonctionnels. En effet, le gène *amiA1* est un pseudogène chez les souches CNRZ1066 et LMG18311 (Hols *et al.*, 2005). Ainsi, afin de déterminer si ces gènes codent des protéines fonctionnelles, le séquençage des gènes *amiA* chez les souches PB302, ST14, CNRZ160 et PB180 a été réalisé. La souche PB302 a été choisie pour représenter les souches de type ITSII de fort pouvoir acidifiant et les souches ST14, CNRZ160 et PB180 ont été choisies pour représenter les souches de type ITSI de moyen pouvoir acidifiant. En outre, ces souches appartiennent à trois branches différentes dans l'arbre phylogénétique établi précédemment (Figure 25). Les deux souches CNRZ160 et PB180 sont positionnées sur la branche A, la souche PB302 est positionnée sur la branche B et la souche ST14 est positionnée sur la branche C. Les quatre souches possèdent le gène *amiA1*, les souches PB302 et PB180 ne possèdent pas le gène *amiA2* et la souche ST14 ne possède pas le gène *amiA3* (Tableau 24). L'analyse des séquences obtenues a révélé que les gènes *amiA* possèdent des forts pourcentages d'identité allant de 97 à 100% avec leurs homologues présents dans les génomes séquencés de *S. thermophilus*. Au niveau protéique, les pourcentages sont compris entre 96 et 100% chez les souches dont les séquences des génomes établies. Enfin, d'après les séquences géniques obtenues, les protéines semblent fonctionnelles (Tableau 25).

**Tableau 24 : Classification des souches en fonction de leur pouvoir acidifiant en lait (d'après Galia *et al.*, 2009). Le type d'ITS, la présence et l'absence du gène *prtS*, l'activité caséinolytique ainsi que la présence du système *ami* sont mentionnés**

Souche	Origine	Pouvoir acidifiant	Type d'ITS 16S-23S	Gène <i>prtS</i>	Activité caséinolytique	Gènes :		
						<i>amiA1</i> <i>CDEF</i>	<i>amiA2</i>	<i>amiA3</i>
CNRZ391	Yaourt	Fort	II	+	+	+	-	+
PB385	Yaourt		II	+	+	+	-	+
PB302	Yaourt		II	+	+	+	-	+
4F44	Fromage		I	+	+	+	-	+
LMD-9 <sup>a</sup>	Yaourt		II	+	+	+	-	+
PB18	Fromage		II	+	+	+	-	+
ST87	Fromage		II	+	+	+	-	+
ST88	Fromage		II	+	+	+	-	+
ST85	Fromage		II	+	+	+	-	+
CNRZ385N	Yaourt	Moyen	I	+	+	+	+	+
HD17	Yaourt		I	+	-	+	+	-
CNRZ404	Yaourt		I	+	+/-	+	+	+
ST14	Yaourt		I	-	-	+	+	-
CNRZ160	Yaourt		I	-	-	+	+	+
CNRZ445	Fromage		I	+	+/-	+	+	+
ATE11PB6	Yaourt		I	+	-	+	+	-
PB18O	Fromage		I	+	-	+	-	+
CNRZ1066 <sup>a</sup>	Yaourt	Faible	I	-	-	+	+	+
CNRZ21	Yaourt		I	-	-	+	+	+
LMG18311 <sup>a</sup>	Yaourt	Non étudié	I	-	Non étudiée	+	+	+
ND03 <sup>a</sup>	Lait		I	-	Non étudiée	+	+	+
JIM8232 <sup>a</sup>	Lait		I	+	Non étudiée	+	-	+

<sup>a</sup> Chez les souches dont les génomes sont entièrement séquencés LMD-9, CNRZ1066, LMG18311, ND03 et JIM8232 les résultats de présence/absence des gènes ont été obtenus à partir des banques de donnés.

**Tableau 25 : Pourcentages d'identité entre les gènes *amiA*/protéines AmiA des souches ST14, CNRZ160, PB180 et PB302 et leurs homologues présents dans les génomes séquencés de *Streptococcus thermophilus***

<b>AmiA1</b>						
<b>(Nucleotide : Protéine)</b>	<b>ND03</b>	<b>LMD-9</b>	<b>JIM8232</b>	<b>CNRZ160</b>	<b>ST14</b>	<b>PB180</b>
<b>CNRZ160</b>	(98,9 : 98,2)	(99,6 : 99,1)	(98,3 : 97,6)	(100 : 100)		
<b>ST14</b>	(98,9 : 98,2)	(99,6 : 99,1)	(98,3 : 97,6)	(99,9 : 99,7)	(100 : 100)	
<b>PB180</b>	(99,4 : 99,3)	(99,1 : 98,6)	(98,8 : 98,6)	(99,0 : 98,3)	(99,0 : 98,3)	(100 : 100)
<b>PB302</b>	(99,4 : 99,2)	(99,1 : 98,6)	(98,8 : 98,6)	(99,0 : 98,3)	(99,0 : 98,3)	(100 : 100)

<b>AmiA2</b>					
<b>(Nucleotide : Protéine)</b>	<b>ND03</b>	<b>LMG18311</b>	<b>CNRZ1066</b>	<b>CNRZ160</b>	<b>ST14</b>
<b>CNRZ160</b>	(99,9 : 99,7)	(99,9 : 99,8)	(100 : 100)	(100 : 100)	
<b>ST14</b>	(99,9 : 99,7)	(99,9 : 99,8)	(100 : 100)	(100 : 100)	(100 : 100)

<b>AmiA3</b>							
<b>(Nucleotide : Protéine)</b>	<b>ND03</b>	<b>LMG18311</b>	<b>CNRZ1066</b>	<b>LMD-9</b>	<b>JIM8232</b>	<b>CNRZ160</b>	<b>PB180</b>
<b>CNRZ160</b>	(97,5 : 96,5)	(99,4 : 99,1)	(99,4 : 98,9)	(99,0 : 98,2)	(97,7 : 96,9)	(100 : 100)	
<b>PB180</b>	(98,0 : 96,9)	(99,0 : 98,5)	(98,6 : 97,4)	(100 : 100)	(98,0 : 97,1)	(99,0 : 98,2)	(100 : 100)
<b>PB302</b>	(98,0 : 96,9)	(99,0 : 98,5)	(98,6 : 97,4)	(100 : 100)	(98,0 : 97,1)	(99,0 : 98,2°)	(100 : 100)

### 1. 3. Recherche du transporteur DtpT

Jusqu'à maintenant aucun système de transport de di- tri- peptides n'a été caractérisé chez *S. thermophilus*. L'étude de Garault *et al.* (2002) a montré qu'un triple mutant *amiA1/A2/A3* de la souche *S. thermophilus* ST18 était capable de croître dans un milieu contenant des di- tri-peptides comme unique source azotée. Cela a permis aux auteurs de suggérer la présence d'au moins un autre système de transport de di- tri-peptides fonctionnel (Garault *et al.*, 2002). De plus, l'étude de Herve-Jimenez *et al.* (2008) a montré que le gène *dtpT* de la souche LMG18311 était surexprimé en fin de phase exponentielle en milieu lait.

Les promoteur et terminateur potentiels ont été détectés respectivement en amont et en aval du gène *dtpT* chez la souche LMD-9 (Figure 28). Ce gène code un transporteur de 483 résidus d'acide aminé. Deux modèles topologiques sont proposés par le logiciel de prédiction de structure TMPRED (<http://www.ch.embnet.org/>). Le premier, propose une perméase de 12 domaines hélicoïdaux transmembranaires composés de 19 à 27 résidus d'acide aminé, avec six boucles du côté extracellulaire et cinq du côté cytoplasmique. Dans ce modèle, les extrémités N- et C-terminales seraient cytoplasmiques (Figure 29A). Le deuxième, propose une perméase composée de 13 domaines hélicoïdaux transmembranaires composés de 19 à 27 résidus d'acide aminés, avec six boucles du côté extracellulaire et six du côté cytoplasmique, et dans ce modèle, l'extrémité N-terminale serait extracellulaire tandis que l'extrémité C-terminale serait cytoplasmique (Figure 29B). L'analyse des séquences protéiques déduites à partir des allèles *dtpT* des 5 génomes connus (quatre allèles différents) révèle que la séquence protéique est bien conservée avec des pourcentages d'identité de 98 à 100% entre les souches. Le gène *dtpT* est bien conservé entre les souches de *S. thermophilus* et possède donc toutes les caractéristiques pour être fonctionnel.

La présence du gène *dtpT*, au sein de toutes les souches de notre collection a été recherchée par PCR en amplifiant un fragment interne de 610 pb avec le couple d'amorce « *dtpT* for/*dtpT* rev » (Tableaux 13, 16). Ce gène a été détecté chez l'ensemble des souches.

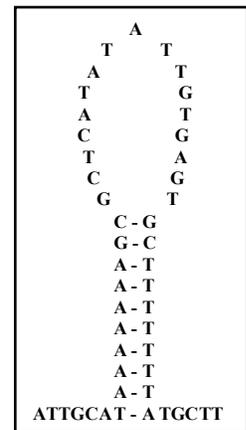
```

1 TAACCAAGGC TTTTGTGACA TCCGCATAAC GGCAAGAAGA ATAGCAAAAA
51 AGTTTCCAAT GATAAAGCTC AAAACTGAGA TTCCTAAGGT ATATATAAGG
101 CAGACCTGAT AGGACATAAG GAATGCTTTT TAGAATCAGT GAAAAATCAA
151 TAATCATAAG TCTAACTTTC TAAACTTTTC TTTTGCCATA CTTTTGGACA
201 TCCCAACGGG TACCTCATT CTTTAACTTA AATTCGTTCA CCCTCTATTG
251 TATAGATAAA TCCACGGTCT GACCACATA TCGCCCTATT TTACGATATC
301 TTCACGTCAA CCTATCGGTT AAGCACGGCT AATATGATAT TATTTTGATA
351 CTAAATTGAG TAAATAATTT GTATTTTACA CCATTTTAGC ATAAAATAGA
401 AAGATATTTT TAATAAATAT GTTTGTAA CG TTGTTAATTC AAATAA CG
-35 boîte -10 boîte

RBS Codon Start dpT
451 TATCAAAAAT CAACCGGAGG ATTTTATGGA AGACAAAGGA AAGACATTTT
501 TCGGGCAACC TTTAGGACTG TTAACACTAT TCATGACTGA AATGTGGGAG
551 CGTTTCTCAT ACTATGGTAT GCGCGCCATC TTGCTTTACT ATATGTGGTT
601 CCTCATCAGT ACTGGTGATC TACACATTAC ACGAGCAACA GCCGCATCAA
651 TCATGGCAAT TTACGCTTCA ATGGTTTACC TTTCTGGTAC GATCGGTGGT
701 TTCGTAGCCG ATCGTATCAT CGGAGCACGT CCAGCTGTAT TCTGGGGTGG
751 TGTTCTCATC ATGCTTGGAC ATATCGTGCT CACTTTGCCT TTTGGCGCAA
801 GCGCATTGTT TGGGTCAATC ATCTTAATTA TTATCGGAAC TGGTTTCTTG
851 AAGCCAAACG TTTCAACATT GGTTGGGACA TTGTATGACG AACATGATCG
901 TCGTCGTGAC GCTGGTTTCT CAATCTTGT GTTTGGTATT AACTTGGGTG
951 CGTTCATTGC ACCACTTATT GTTGGTGCTG CACAAGAAGC TGCAGGTTAC
1001 CACGTGGCCT TCTCATTGGC AGCAATCGGT ATGTTTATTG GTTTGCTGGT
1051 TTATTACTTC GGTGGTAAGA AAACATTGGA TCCGCGTTAC TTGCGCCCAA
1101 CTGATCCACT AGCACCAGAA GAAGTTAAGC CATTACTGGT TAAAGTAAGC
1151 TTAGCTGTTG CTGGCTTTAT TGCAATCATC GTTGTGATGA ACCTTGTCGG
1201 TTGGAATTCA TTGCCTGCCT ACATCAATCT CTTGACGATT GTTGCGATCG
1251 CCATTCCAGT CTTCTATTTT GTTTGGATGA TTGCTCAGT CAAGGTTACG
1301 GCAACAGAAC GTTTGCGTGT TGTATCATA ATTCCACTCT TCATCGCCGC
1351 TGTTTTGTTT TGGGCGATTG AAGAACAAGC CTCAGTTGTG TTGGCAACAT
1401 TTGCTGCTGA ACGTGTTGAT TCATCTTGGT TCCCTGTCTC ATGGTTCCAA
1451 TCATTGAACC CACTCTTCAT CATGCTTTAC ACACCATTCT TTGCTTGGCT
1501 ATGGACTGCC TGGGAAGAAA ATCAACCATC ATCACCACA AAGTTTGCCG
1551 TTGGTTTGAT GTTTGCCGGC TTAAGCTTCT TGCTCATGGC CATTCCCTGGT
1601 GCTTTGTATG GCACAAGTGG TAAAGTATCA CCACTATGGT TAGTTGGGTC
1651 ATGGGCCCTT GTTATCGTCG GTGAAATGTT AATTTACCT GTCGGCCTAT
1701 CCGTCACAAC CAAGTTGGCA CAAAAGCCT TCAACTCTCA AATGATGTCA
1751 ATGTGGTTCT TGTCATCAGC GGTAGGTTCT GCTTTGAACG CACAACGGT
1801 GACATTGTAT AACGCCAAT CTGATGTCGC TTAATTCTCT TACTTCGGAT
1851 TAGGTTCTGT TGTACTTGGT ATTGTTCTCG TCTTCCATC AAAGCGTATT
1901 CAAGGTTTGA TGCAAGGTGT TGAATAATT GCATAAAAA AGCGCTCATA

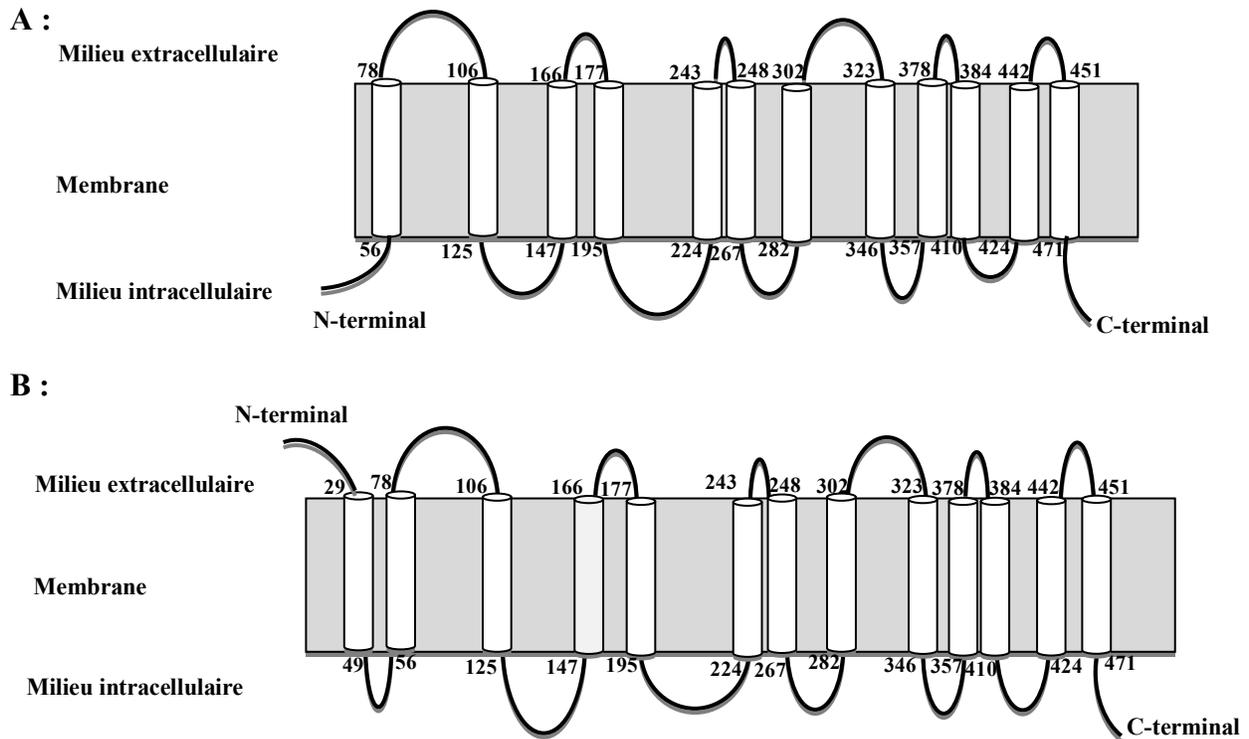
Terminateur putatif dpT
1951 TATTGTGAGT GCTTTTTTTA TGCTTAAATTA TTTCACATCG TCGGGCTTTG
2001 TCAGCAACTT TTTCAGCAGC CTTGTCAACA TTTCATTAA CTTCTTTAGC
2051 CTTATCAGCA GCTTTATCAG AGAAATCTTC AGCTTCAACT TGACCCGCGC

```



**Figure 28 : Séquence nucléotidique du gène *dpT* chez la souche *Streptococcus thermophilus* LMD-9 avec les positions putatives des promoteur et terminateur**

La séquence codante du gène *dpT* est présentée en bleu. Les codons « start » et « stop » sont surlignés en gris. Le promoteur a été recherché en se basant sur la séquence consensus du promoteur définie pour les *Firmicutes* (TTGACA-17nt-TATAAT) (Helmann, 1995 ; Morrison et Jaurin, 1990). Les séquences putatives promotrice et terminatrice sont encadrées, le site RBS (pour Ribosomal Binding Site) est indiqué en caractères gras (Rocha *et al.*, 1999). Le zoom représente le terminateur replié.



**Figure 29 : Structures tridimensionnelles de la perméase DtpT de *Streptococcus thermophilus* LMD-9 proposées par le logiciel de prédiction de structures transmembranaires TMpred**

A, B : les deux structures proposées par le logiciel TMpred ([http://www.ch.embnet.org/software/TMPRED\\_form.html](http://www.ch.embnet.org/software/TMPRED_form.html)) diffèrent par le nombre de domaines transmembranaires et la localisation des extrémités N- et C- terminales de la perméase.

## 1. 4. Détection du cluster *ots* (Oligopeptide Transport Streptococcus *thermophilus*) de la famille des transporteurs de peptides et de nickel de type ABC

### 1. 4. 1. Détection et organisation du cluster *ots*

Une recherche *in silico* dans les 5 génomes séquencés a été menée afin de déterminer s'il existe des systèmes de transport de peptides potentiels. Un cluster constitué de 4 gènes (STER\_0142, STER\_0143, STER\_0144, STER\_0145) et précédé d'un pseudogène (STER\_0141) a été retrouvé chez la souche LMD-9. Il est annoté comme codant potentiellement un système de transport de peptides et de nickel de type ABC, sachant que sur la base de la structure primaire il est impossible de différencier les 2. Ce cluster a été cité comme provenant d'un événement de transfert horizontal dans l'étude de Liu *et al.* (2009b) qui avait comme objectif la détection d'évènements potentiels de transferts de gènes entre *S. thermophilus* et *Lb. bulgaricus*. Le pourcentage en GC est de 24, 36, 39, 41 et 40% pour les gènes STER\_0141, STER\_0142, STER\_0143, STER\_0144, STER\_0145. L'origine de ce cluster n'a pas été déterminée par Liu *et al.* (2009b).

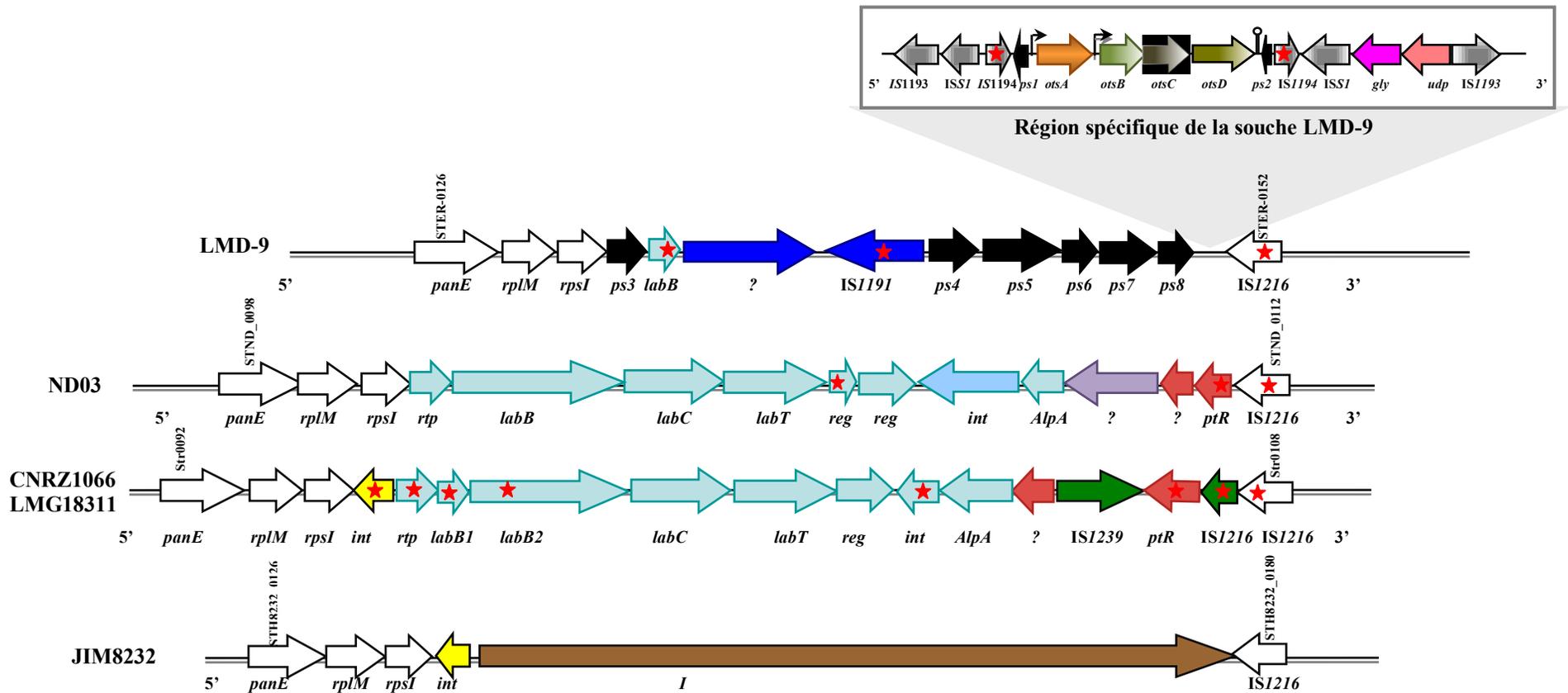
Pour la suite de notre étude ce cluster a été nommé *ots* pour : Oligopeptide Transport Streptococcus *thermophilus*. Il est composé de 4 gènes :

- *otsA* codant une protéine affine aux peptides ou au nickel de 551 résidus d'acide aminé ;
- *otsB* et *otsC* codant deux protéines membranaires de 352 et 317 résidus d'acide aminé respectivement, qui forment le pore du système. Ces deux perméases partagent 15% d'identité ;
- *otsD* codant une ATPase double de 565 résidus d'acide aminés, qui fournit l'énergie au système.

#### **1. 4. 2. Analyses des régions flanquant le cluster *ots* chez *Streptococcus thermophilus* LMD-9 et comparaison avec celle des souches CNRZ1066, LMG18311, ND03, JIM8232**

Afin de mieux connaître la région d'intégration du cluster *ots*, les régions flanquant ce cluster chez la souche LMD-9 ont été analysées et comparées aux régions homologues retrouvées chez les 4 autres souches dont les génomes sont entièrement séquencés (CNRZ1066, LMG18311, ND03 et JIM8232), et qui ne possèdent pas ce cluster (Figure 30).

Cette analyse montre que le cluster *ots* est situé dans une région chromosomique très variable. Des îlots génomiques différents, acquis par transfert horizontal, ont été retrouvés chez les différentes souches. Chez la souche JIM8232, un îlot génomique contenant des gènes codant des fonctions impliquées dans la biosynthèse de pigments jaunes a été décrit, il s'agit de gènes atypiques de l'espèce *S. thermophilus*, probablement acquis par transfert horizontal à partir de l'espèce *S. agalactiae* (Delorme *et al.*, 2010). Tout comme les deux souches CNRZ1066 et LMG18311, la souche ND03 en possède un cluster *labBCT*, mentionné comme étant acquis par transfert horizontal (Hols *et al.*, 2005). Ce cluster est composé des gènes *labB* et *C* qui coderaient des protéines impliquées dans la biosynthèse de lantibiotique. Et du gène *labT* qui coderait une protéine responsable de l'efflux de lantibiotique.



**Figure 30 : Représentation schématique de la région contenant le cluster *ots* chez *Streptococcus thermophilus* LMD-9, CRZ1066, LMG18311, ND03 et JIM8232**

Les flèches noires correspondent à des pseudogènes, les blanches à des gènes présents chez les cinq souches, les flèches bleu clair à des gènes présents chez les souches CNRZ1066/LMG18311, ND03 et LMD-9, les rouges à des gènes présents chez les souches LMG18311/CNRZ1066 et ND03, les jaunes à des gènes présents chez les souches CNRZ1066/LMG1 et JIM8232, les vertes à des gènes présents chez les souches CNRZ1066/LMG18311, les brunes à des gènes spécifiques de la souche JIM8232, les violettes à des gènes spécifiques de la souche ND03, le bleu foncé à des gènes spécifiques de la souche LMD-9. Les points d'interrogation correspondent à des gènes codant des protéines de fonction inconnue, les étoiles correspondent à des gènes tronqués, *ps1*, *2*, *3*, *4*, *5*, *6*, *7*, *8* : des pseudogènes. *panE* : cétopantoate réductase ; *rplM* : protéine ribosomale 50 S (L13) ; *rpsI* : protéine ribosomale 30S (S9) ; *rtp* : régulateur transcriptionnel putative ; *labB* : « lantibiotic biosynthesis protein » ; *labC* : « lantibiotic biosynthesis protein » ; *labT* : « lantibiotic efflux protein » ; *int* : une intégrase ; *AlpA* : un régulateur transcriptionnel de la famille AlpA ; *ptR* : « putative transfert protein ». La région I correspond à des gènes présents uniquement chez la souche JIM8232 et comporte des gènes codant des fonctions impliquées dans la pigmentation jaune mentionnés par Delorme *et al.* (2011)

### 1. 4. 3. Paramètres physicochimiques de chaque protéine du cluster *ots*

Les paramètres physicochimiques de chacune des protéines du cluster *ots*, point isoélectrique (pI) et masse moléculaire (MM), ont été estimés à l'aide du logiciel pI/mr d'EXPASY et sont présentés dans le tableau 26.

**Tableau 26 : Paramètres physicochimiques (masses moléculaires et points isoélectriques) calculés à partir de la composition des protéines OtsA, OtsB, OtsC et OtsD à l'aide du logiciel compute pI/Mw ([http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/))**

	OtsA	OtsB	OtsC	OtsD
<b>Masse moléculaire (kDa)</b>	60,0	38,7	34,1	62,0
<b>Point isoélectrique</b>	5,05	9,97	9,97	5,72

### 1. 4. 4. Recherche *in silico* d'une organisation génétique similaire au cluster *ots*

A l'aide du logiciel « BlastP », un cluster homologue au cluster *ots* a été recherché dans les banques de données. Aucun homologue n'a été retrouvé chez d'autres espèces de streptocoques, ni chez d'autres bactéries lactiques utilisées en industrie agro-alimentaire. Par contre, au sein de différents genres tels que *Bifidobacterium*, *Arthrobacter*, *Mycobacterium*, *Brevibacterium*, *Clavibacter* et *Corynebacterium* ce cluster a été retrouvé. Dans tous les cas, le cluster garde la même synténie avec des pourcentages d'identité élevés compris entre 47 et 66% (Tableau 27). Les bactéries comportant ce cluster appartiennent toutes au phylum des *Actinobacteria*, bactéries à Gram positif dont le pourcentage en GC est supérieur à 50%, qui colonisent un nombre important de niches écologiques, telles que le sol (*Arthrobacter*), les plantes (*Clavibacter*), les mammifères (*Mycobacterium*), le tractus digestif humain ou les produits laitiers fermentés (*Brevibacterium* et *Bifidobacterium*). Ces clusters portent également des annotations non vérifiées et la fonction de ces protéines n'a donc pas été étudiée chez ces différentes bactéries.

**Tableau 27 : Pourcentages d'identité entre les protéines Ots de *Streptococcus thermophilus* LMD-9 et leurs homologues présents au sein de cluster ayant une organisation génétique similaire à celle du cluster *ots***

Les annotations des banques de données sont mentionnées entre guillemets. Les pourcentages présentés ont été obtenus à l'aide du logiciel « Blast P »

<b>souches</b>	<b>OtsA</b>	<b>OtsB</b>	<b>OtsC</b>	<b>OtsD</b>
<i>S. thermophilus</i> LMD-9	« ABC-type oligopeptide transport system periplasmic component »	« ABC-type dipeptide/oligopeptide/nickel transport system, permease component »	« ABC-type dipeptide/oligopeptide/nickel transport system, permease component »	« ABC-type dipeptide/oligopeptide/nickel transport system, ATPase component »
<i>Bifidobacterium bifidum</i> PRL2010	52% « Peptides ABC transporter solute binding protein »	63% « Oligopeptide transport system permease protein »	66% « Oligopeptide transport ATP-binding protein »	54% « Oligopeptide transport ATP-binding protein »
<i>Bifidobacterium longum</i> ssp. <i>infantis</i> ATCC 15697	52% « Extracellular solute-binding protein, family 5 »	67% « Binding-protein-dependent transport systems inner membrane component »	64% « Binding-protein-dependent transport systems inner membrane component »	55% « ABC transporter related »
<i>Bifidobacterium longum</i> ssp. <i>infantis</i> ATCC 55813	51% « Peptide ABC superfamily ATP binding cassette transporter, solute-binding protein »	66% « Dipeptide/oligopeptide/nickel (Ni <sup>2+</sup> ) ABC superfamily ATP binding cassette transporter, membrane protein »	64% « Dipeptide/oligopeptide/nickel ABC superfamily ATP binding cassette transporters, permease protein »	55% « ABC superfamily ATP binding cassette transporter, ABC protein »
<i>Bifidobacterium longum</i> DJO10A	51% « ABC-type dipeptide transport system periplasmic component »	66% « ABC-type dipeptide/oligopeptide/nickel transport systems permease component »	64% « ABC-type dipeptide/oligopeptide/nickel transport systems, permease components »	55% « ATPase components of various ABC-type transport systems, contain duplicated ATPase »
<i>Bifidobacterium longum</i> NCC2705	52% « Solute-binding protein of ABC transporter for peptides »	66% « Peptide ABC transporter permease »	65% « Peptide ABC transporter permease »	55% « ATP binding protein of ABC transporter »
<i>Bifidobacterium adolescentis</i> ATCC 15703	51% « Solute-binding protein of ABC transporter for peptides »	65% « Peptide ABC transporter permease »	64% « Peptide ABC transporter permease »	55% « ATP binding protein of ABC transporter »
<i>Bifidobacterium angulatum</i> DSM 20098	52% « Hypothetical protein BIFANG_01657 »	66% « Hypothetical protein BIFANG_01655 »	62% « Hypothetical protein BIFANG_01654 »	56% « Hypothetical protein BIFANG_01653 »
<i>Bifidobacterium pseudocatenulatum</i> DSM 20438	53% « Hypothetical protein BIFPSEUDO_03007 »	65% Hypothetical protein BIFPSEUDO_03006 »	65% « Hypothetical protein BIFPSEUDO_03005 »	54% « Hypothetical protein BIFPSEUDO_03004 »
<i>Brevibacterium linens</i> BL2	52% « ABC-type dipeptide transport system, periplasmic component »	59% « ABC-type dipeptide/oligopeptide/nickel transport systems, permease components »	59% « ABC-type dipeptide/oligopeptide/nickel transport systems, permease components »	47% « Putative ABC transporter ATP-binding protein »
<i>Arthrobacter</i> sp. FB24	57% « Extracellular solute-binding protein »	61% « Binding-protein-dependent transport systems inner membrane component »	61% « Binding-protein-dependent transport systems inner membrane component »	58% « ABC transporter-like protein »
<i>Arthrobacter chlorophenolicus</i> A6	55% « Extracellular solute-binding protein family 5 »	59% « Binding protein-dependent transport systems inner membrane component »	61% « Binding-protein-dependent transport systems inner membrane component »	59% « ABC transporter related »
<i>Mycobacterium abscessus</i> ATCC 19977	47% « Putative oligopeptide ABC transporter, solute-binding protein »	51% « Oligopeptide ABC transporter, permease protein »	56% « Oligopeptide ABC transporter, permease protein »	51% « Putative oligopeptide ABC transporter, ATP-binding protein »
<i>Marine actinobacterium</i> PHSC20C1	51% « ABC transporter, solute-binding protein »	59% « ABC transporter, permease protein »	66% « ABC transporter, permease protein »	58% « ABC transporter, ATP-binding protein »
<i>Clavibacter michiganensis</i> ssp. <i>Sepedonicus</i>	53% « Putative substrate-binding transport protein »	59% « Putative integral membrane transport protein »	61% « Putative integral membrane transport protein »	57% « Putative ABC transporter ATP-binding subunit »
<i>Clavibacter michiganensis</i> ssp. <i>michiganensis</i> NCPPB 382	53% « Putative oligopeptide ABC transporter, substrate-binding protein »	59% « Putative oligopeptide ABC transporter, permease component »	59% « Putative oligopeptide ABC transporter, permease component »	50% « Putative oligopeptide ABC transporter, ATP-binding protein »
<i>Corynebacterium ammoniagenes</i> DSM 20306	47% « Putative ABC-type di/oligopeptide transport system »	57% « Dipeptide ABC transporter, permease protein »	62% « Dipeptide ABC transporter, permease protein »	55% « ABC transporter, ATP-binding protein »

### **1. 4. 5. Recherche de motif caractéristique de protéine constituant les systèmes de transport de peptides et de nickel de type ABC au sein du cluster *ots***

Toutes les protéines constituant les systèmes de transports ABC possèdent plusieurs motifs caractéristiques. Ces motifs ont été recherchés au sein des protéines Ots.

#### **1. 4. 5. 1. Recherche du motif consensus de lipomodification chez OtsA**

Comme cela a été évoqué précédemment, chez les bactéries à Gram positif, les protéines affines aux peptides sont des lipoprotéines attachées à la membrane. Un motif consensus de lipomodification a été recherché et identifié dans le domaine N-terminal d'OtsA. Ce motif est très similaire à celui des OBP Ami A1 et A3 (Figure 31A).

#### **1. 4. 5. 2. Motif signature caractéristique de la famille des protéines affines aux peptides et au nickel chez OtsA**

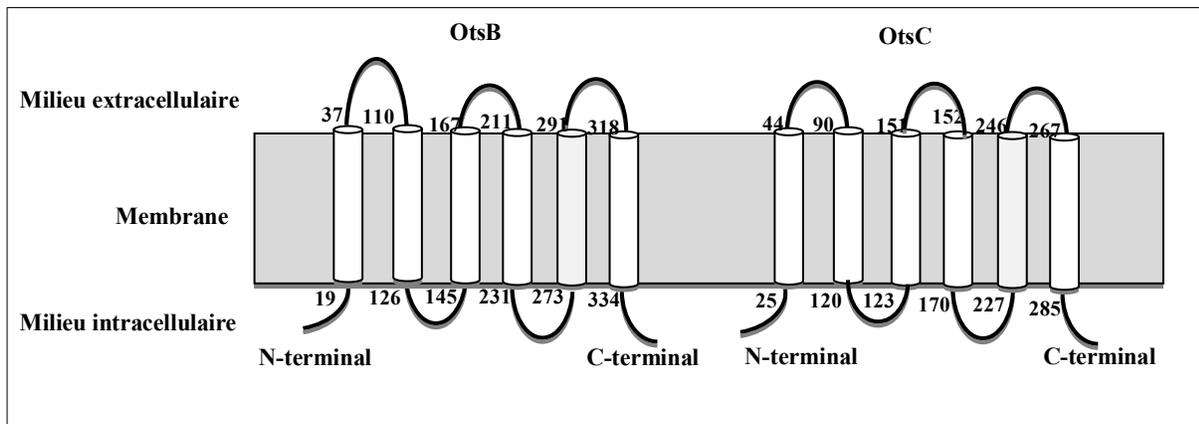
La protéine OtsA, tout comme les protéines AmiA1 et AmiA3, possède également un motif signature spécifique (MAXXIXXNKDXXXXTYXVRXXXXK) de la famille des protéines affines aux peptides et au nickel proche du motif signature décrit par Tam et Saier en 1993 (Figure 31B). Néanmoins trois différences sont à mentionner par rapport à la séquence consensus: un résidu de sérine semble être intercalé en position 88 et deux résidus conservés au sein des séquences consensus sont substitués : un résidu de glycine remplace un résidu de tryptophane ou d'isoleucine et un résidu de valine remplace un résidu de phénylalanine ou de tyrosine (Figure 31B).



### 1. 4. 5. 3. Prédiction de structures secondaires des perméases transmembranaires OtsB et OtsC

Les protéines OtsC et OtsB étant annotées comme étant des perméases, la prédiction de leur structure a été réalisée à l'aide du logiciel TMPRED (<http://www.ch.embnet.org/>). Chaque perméase présenterait 6 domaines hélicoïdaux transmembranaires composés de 17 à 31 résidus d'acide aminé (Figure 32). Les domaines N- et C-terminaux et deux boucles cytoplasmiques seraient orientés vers le milieu intracellulaire tandis que trois boucles seraient orientées vers le milieu extracellulaire.

Le motif « EAA » signature spécifique des systèmes d'importation de substrat, connu pour être très peu conservé (Mourez *et al.*, 1997), n'a pas été retrouvé mais cela ne compromet en rien la fonctionnalité du système.



**Figure 32 : Modèle topologique des perméases OtsB et OtsC de *Streptococcus thermophilus* LMD-9 déduit à l'aide du logiciel TMpred**

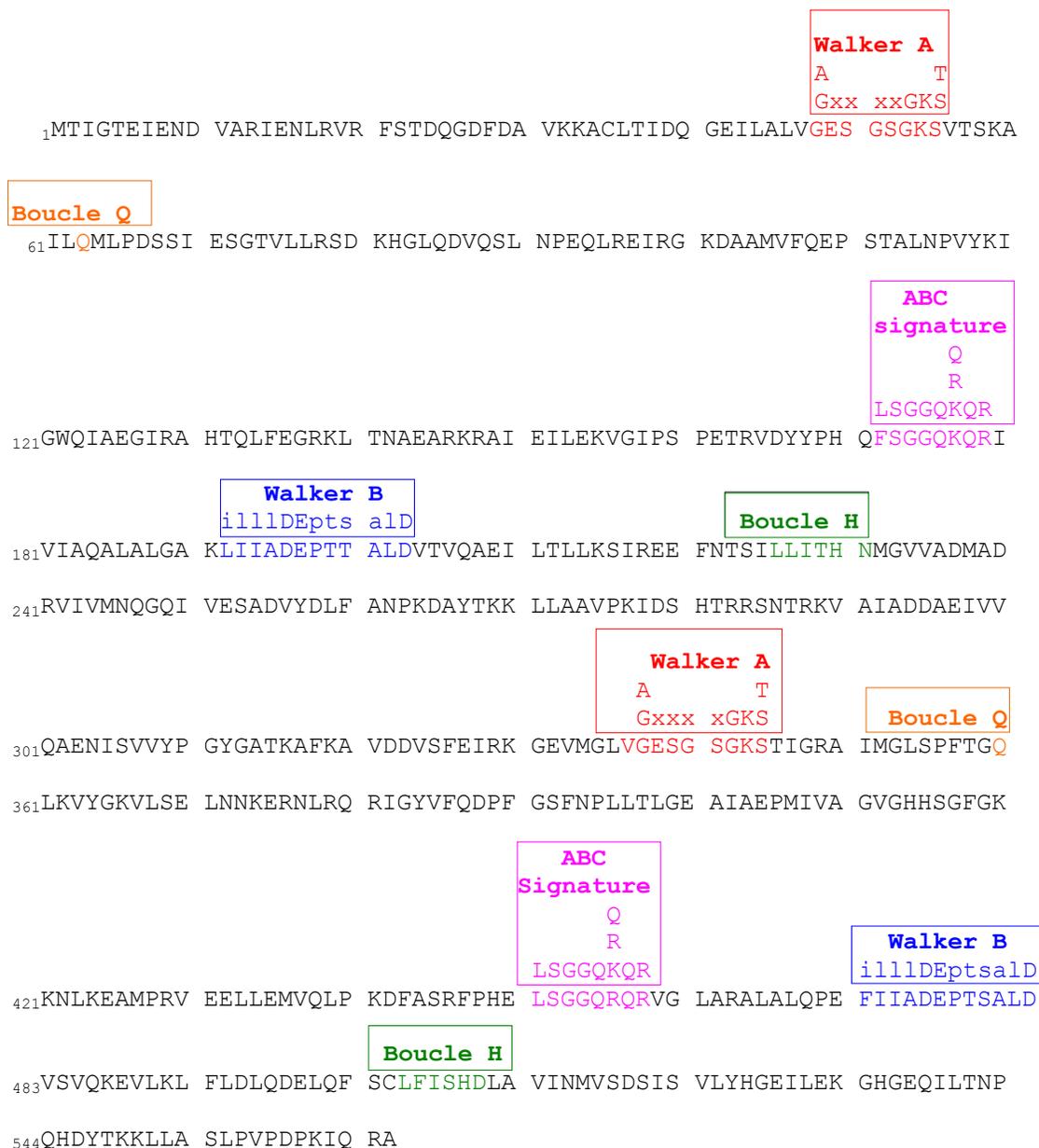
([http://www.ch.embnet.org/software/TMPRED\\_form.html](http://www.ch.embnet.org/software/TMPRED_form.html))

#### **1. 4. 5. 4. Différents motifs de l'ATPase OtsD**

OtsD est une ATPase qui fournirait de l'énergie au système. Contrairement au système Ami où chaque perméase est couplée à une ATPase contenant chacune un domaine NBD, dans le système Ots, une seule séquence protéique ayant une double activité ATPase a été retrouvée. En Figure 33, les 5 motifs peptidiques conservés, Walker A, Walker B, « signature ABC », boucle Q et boucle H, importants pour l'activité ATPase sont présents en double exemplaire.

#### **1. 4. 5. 5. Conclusion**

En conclusion, le système Ots aurait tous les domaines caractéristiques d'un transporteur de la famille des transporteurs de peptides et de nickel de type ABC. Il est composé d'une protéine affine au substrat, OtsA, ayant les motifs de lipomodification et de fixation de peptides et de nickel caractéristiques ; de deux perméases ayant 6 domaines caractéristiques par protéine et permettant le passage du substrat, et enfin, d'une ATPase possédant deux domaines (chaque domaine ayant les 5 motifs spécifiques à l'activité ATPasique) capables de fournir l'ATP nécessaire à la translocation du substrat. D'après sa séquence nucléotidique, ce cluster semble fonctionnel. Il n'est pas présent chez toutes les souches de *S. thermophilus* et ne fait donc pas partie du « core génome » de cette bactérie lactique. Il pourrait provenir d'un événement de transfert horizontal (Liu *et al.*, 2009b).



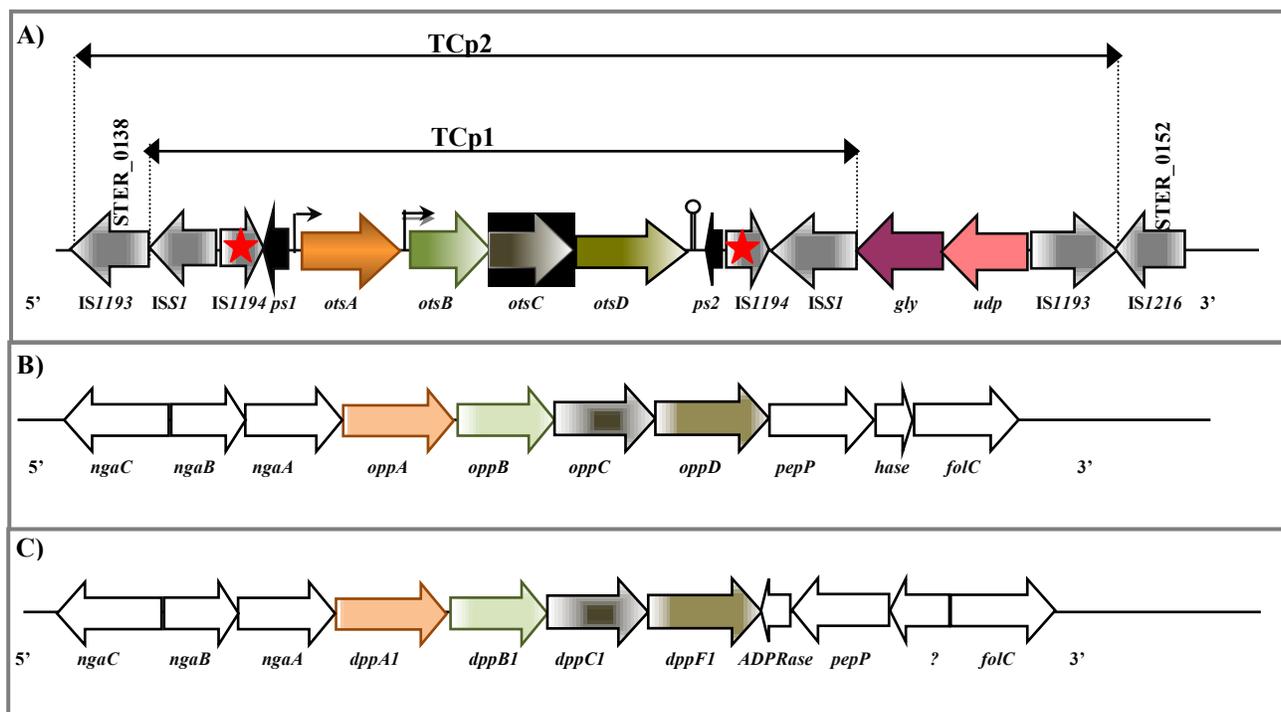
**Figure 33 : Séquence protéique de OtsD (double ATPase) de *Streptococcus thermophilus* LMD-9 déduite de la séquence nucléotidique du gène *otsD***

Les différents motifs signatures spécifiques de la famille des ATPases (Walker A, boucle Q, ABC signature, Walker B puis boucle H organisés en double et indispensables à la fixation et à l'hydrolyse de l'ATP) sont indiqués.

#### 1. 4. 6. Analyse des régions adjacentes au cluster *ots*

La présence du cluster *ots* chez seulement un des 5 génomes séquencés de *S. thermophilus* suggère qu'il n'était pas présent chez l'ancêtre de *S. thermophilus* et qu'il a été acquis récemment. L'absence d'un cluster homologue chez tous les streptocoques dont les séquences génomiques sont disponibles et plus précisément chez *S. salivarius* et *S. vestibularis*, deux espèces phylogénétiquement proches de *S. thermophilus*, renforce cette hypothèse. *S. thermophilus* aurait ainsi acquis le cluster *ots* après son individualisation au sein du groupe *salivarius*. Enfin, la présence d'un cluster homologue à *ots* chez d'autres genres bactériens, tous issus du phylum des *Actinobacteria*, permet d'émettre l'hypothèse que le cluster *ots* provient initialement d'une actinobactérie. Afin d'explorer davantage cette hypothèse, l'analyse des régions flanquant le cluster *ots* chez la souche LMD-9 a été menée (Figure 34A).

L'analyse des séquences situées en amont et en aval du cluster *ots* a révélé la présence de copies d'IS apparemment fonctionnelles. Tout d'abord, deux copies apparemment fonctionnelles d'IS1 de type  $\alpha$  sont retrouvées. La première est localisée en amont du gène *otsA*, au locus STER\_0139, et la deuxième, en aval du gène *otsD*, au locus STER\_0148 (Figure 34A). Les transposases de ces deux copies possédant 98% d'identité (Tableau 28), ces deux copies d'IS1 pourraient constituer et délimiter un transposon composite que nous avons appelé TCp1 (pour Transposon Composite putatif 1). TCp1 semble être entièrement inclus dans une deuxième structure délimitée par deux copies apparemment fonctionnelles d'IS1193 (Figure 34A), dont les transposases possèdent 96% d'identités (Tableau 28). L'une est située au locus STER\_0138 et l'autre au locus STER\_0151. Compte tenu de la ressemblance entre leurs transposases, ces deux copies pourraient délimiter, elles aussi, un transposon composite que nous avons désigné TCp2 (pour Transposon Composite putatif 2). L'élément IS1193 peut être considéré comme typique de *S. thermophilus* (G. Guédon, communication personnelle). L'interrogation des banques de données en utilisant les séquences des deux copies d'IS1193, qui bornent TCp2 montre l'existence de plusieurs copies d'IS1193 au sein des génomes séquencés de *S. thermophilus*.



**Figure 34 : Organisation de la région chromosomique contenant le cluster *ots* ou son homologue chez *Streptococcus thermophilus* LMD-9 (A) *Bifidobacterium bifidum* PRL2010 (B) *Bifidobacterium longum* DJO 10A (C)**

**A.**

*otsA* : gène codant une protéine affine aux peptides et au nickel ; *otsB* et *otsC* : gènes codant deux protéines transmembranaires ; *otsD* : gène codant une ATPase ; *gly* : gène codant une glycosyltransférase, probablement impliquée dans la biogenèse de la paroi cellulaire ; *udp* : gène codant une UDP-glucose 6-deshydrogénase prédite. Les deux flèches noires correspondent à des pseudogènes, les étoiles au milieu des flèches correspondent à des gènes tronqués. TCp1 et TCp2, pour Transposon Composite putatif 1 et 2 respectivement, représentent deux transposons composites potentiels.

**B.**

*nagC* : gène codant un régulateur transcriptionnel ; *nagB* : gène codant une Glucosamine-6-phosphate isomérase ; *nagA* : gène codant une N-acétylglucosamine-6-phosphate desacétylase ; *oppA* : gène codant une protéine affine aux peptides et au nickel ; *oppB* et *oppD* : gènes codant deux protéines transmembranaires ; *oppC* : gène codant une ATPase ; *pepP* : gène codant une aminopeptidase ; *phase* : gène codant une phosphohydrolase ; *folC* : gène codant une folylpolyglutamate synthétase.

**C.**

*nagC* : gène codant un régulateur transcriptionnel ; *nagB* : gène codant une Glucosamine-6-phosphate isomérase ; *nagA* : gène codant une N-acétylglucosamine-6-phosphate desacétylase ; *dppA1* : gène codant une protéine affine aux peptides et au nickel ; *dppB1* et *dppC1* : gènes codant deux protéines transmembranaires ; *dppF1* : gène codant une ATPase ; *ADPRase* : gène codant une ADP-ribose pyrophosphatase ; *pepP* : gène codant une aminopeptidase ; *folC* : gène codant une folylpolyglutamate synthétase.

**Tableau 28 : Pourcentages d'identité entre les transposases des six IS localisées en amont et en aval du cluster *ots* chez *Streptococcus thermophilus* LMD-9**

	STER_0138 IS1193	STER_0139 ISS1	STER_0147 IS1194	STER_0148 ISS1	STER_0151 IS1193	STER_0152 IS1216
STER_0138 IS1193	100%					
STER_0139 ISS1		100%				
STER_0147 IS1194			100%			
STER_0148 ISS1		98%		100%		
STER_0151 IS1193	96%				100%	
STER_0152 IS1216		76%		75%		100%

Cellules en vert correspondent aux pourcentages inférieurs à 15%

Les pourcentages d'identité ont été calculés avec le logiciel « clustalW »

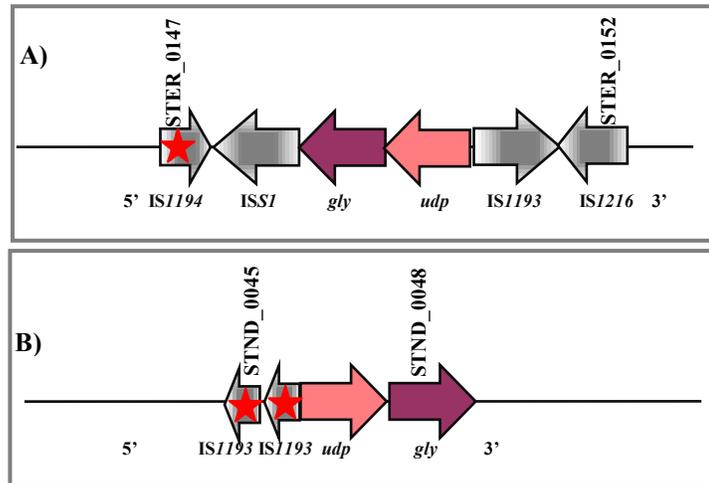
Les copies présentes au locus STER\_0138 et STER\_0151 présentent plus de 95% d'identités avec des copies d'IS1193 présentes dans d'autres souches de *S. thermophilus* et plus de 90% d'identités avec des copies présentes chez *S. salivarius* et *S. vestiularis* (Tableau 29). Les éléments ISSI de type  $\alpha$ ,  $\beta$ ,  $\gamma$  sont typiques de *Lc. lactis*, bien que le type  $\gamma$  soit aussi retrouvé au sein du genre *Enterococcus* (Bourgoin *et al.*, 1996). La présence chez *S. thermophilus* de quelques copies d'ISSI  $\alpha$  et  $\beta$  présentant de forts pourcentages d'identité avec celles retrouvées chez *Lc. lactis*, suggère qu'elles auraient été acquises par transfert horizontal à partir de *Lc. lactis*. Ainsi, l'analyse de la région contenant le cluster *ots* chez *S. thermophilus* LMD-9 a permis la détection de deux éléments, l'un (TCp1) contenant la totalité du cluster *ots* et borné par des copies d'ISSI probablement acquises par transfert horizontal et un plus grand, (TCp2), contenant TCp1 borné par deux copies d'IS, IS1193, endogènes.

Outre les gènes *otsA*, *B*, *C* et *D*, l'élément TCp1 contient 2 copies tronquées d'IS1194, localisées au loci STER\_0140 et STER\_0147, et 2 pseudogènes *ps1* et 2 (loci STER\_0141 et 0146). L'élément TCp2 comporte entre la borne droite de TCp1 et la deuxième copie d'IS1193, constituant la borne droite de TCp2, les gènes *gly* et *udp* qui ne sont présents que dans les génomes séquencés des souches LMD-9 et ND03. Chez la souche ND03, ces gènes, bien que présents mais localisés à un locus différent et en orientation inverse (Figure 35). En amont, 2 copies tronquées d'IS1193 sont également présentes. La présence de ces gènes à des localisations et des orientations différentes dans les deux souches suggèrent qu'il y a eu réarrangements génomiques pouvant impliquer IS1194. Enfin, en aval de l'élément TCp2, une copie d'IS1216 (locus STER\_0152) non fonctionnelle, en raison d'une mutation frameshift dans le gène codant la transposase, est retrouvée. Cet élément appartient à la famille d'ISSI de type  $\gamma$  et est présent chez les bactéries du genre *Enterococcus*. Des copies de cet élément sont retrouvées chez toutes les souches de *S. thermophilus* dont les génomes ont été séquencés (Tableau 29). Des copies possédant des forts pourcentages d'identité ont aussi été trouvées chez plusieurs espèces telles que *En. faecium* et *Lc. lactis*.

**Tableau 29 : Pourcentages d'identité entre les séquences d'insertion adjacentes au cluster *ots* de *Streptococcus thermophilus* LMD-9 et leurs homologues présents dans la banque des données**

Pourcentage d'identité	Organisme
<b>IS1193 de locus STER_0138</b>	
<b>(possédant 96% d'identité avec celle de locus STER_0151)</b>	
99%	<i>Streptococcus thermophilus</i> JIM8232
96%	<i>Streptococcus thermophilus</i> ND03
95%	<i>Streptococcus thermophilus</i> LMG18311
95%	<i>Streptococcus thermophilus</i> CNRZ1066
92%	<i>Streptococcus salivarius</i> 57.1
92%	<i>Streptococcus salivarius</i> CCHSS3
91%	<i>Streptococcus vestibularis</i> JIM8777
91%	<i>Streptococcus vestibularis</i> f0396
<b>ISS1 de type <math>\alpha</math>de locus STER_0139</b>	
<b>(possédant 99% d'identité avec celle de locus STER_0148)</b>	
99%	<i>Lactococcus lactis</i> ssp. <i>cremoris</i> SK11
99%	<i>Lactococcus lactis</i> ssp. <i>cremoris</i> A76
99%	<i>Lactococcus lactis</i> ssp. <i>cremoris</i> MG1363
99%	<i>Lactococcus lactis</i> ssp. <i>lactis</i> CV56
99%	<i>Lactococcus lactis</i> ssp. <i>lactis</i> IL594
99%	<i>Enterococcus faecium</i>
99%	<i>Enterococcus faecalis</i>
<b>IS1194 de locus STER_0147</b>	
95%	<i>Lactococcus lactis</i> ssp. <i>cremoris</i> A76
95%	<i>Leuconostoc lactis</i> KCTC 3528
<b>IS1193 de locus STER_0151</b>	
99%	<i>Streptococcus thermophilus</i> ND03
99%	<i>Streptococcus thermophilus</i> JIM8232
97%	<i>Streptococcus thermophilus</i> CNRZ1066
97%	<i>Streptococcus thermophilus</i> LMG18311
94%	<i>Streptococcus vestibularis</i> F0396
92%	<i>Streptococcus salivarius</i> 57.1
92%	<i>Streptococcus salivarius</i> CCHSS3
92%	<i>Streptococcus salivarius</i> JIM8780
<b>IS1216 de locus STER_0152</b>	
100%	<i>Streptococcus thermophilus</i> ND03
99%	<i>Streptococcus thermophilus</i> CNRZ1066
99%	<i>Streptococcus thermophilus</i> JIM8232
99%	<i>Streptococcus thermophilus</i> LMG18311
99%	<i>Enterococcus faecium</i>
98%	<i>Enterococcus faecalis</i> EF-01
98%	<i>Enterococcus faecium</i> TX0133a04
98%	<i>Lactococcus lactis</i> ssp. <i>cremoris</i> SK1
98%	<i>Lactococcus lactis</i> ssp. <i>lactis</i> ssp. <i>diacetylactis</i>
98%	<i>Streptococcus oralis</i>
98%	<i>Streptococcus suis</i> BM407

<sup>a</sup> Seule la séquence possédant le plus fort pourcentage d'identité a été mentionnée.  
Les pourcentages d'identité sont ceux indiqués par le logiciel « BlastN »



**Figure 35 : Comparaison de l'organisation génétique des gènes *udp* et *gly* chez *Streptococcus thermophilus* LMD-9 (A) et ND03 (B)**

Les étoiles au milieu des flèches indiquent les gènes tronqués

#### 1. 4. 7. Analyses des régions flanquant les clusters homologues au cluster *ots* chez *Bifidobacterium*

Les régions flanquant les clusters homologues au cluster *ots* ont été analysées chez deux espèces du genre *Bifidobacterium* dont les génomes sont entièrement séquencés. Elle révèle que ces clusters ne possèdent pas les mêmes régions voisines (Figure 34).

Par exemple, chez *Bifidobacterium bifidum* PRL2010, dont le cluster *opp* (composé d'*oppA*, *oppB*, *oppC*, *oppD*) est homologue au cluster *ots* (Tableau 27), aucune transposase n'a été trouvée en amont et en aval du cluster (Figure 34B). En amont du gène *oppA*, dans l'ordre, sont présents un gène codant le régulateur transcriptionnel (*nagC*), un gène codant la protéine glucosamine-6-phosphate isomérase (*nagB*) et un gène codant la protéine N-acétylglucosamine-6-phosphate déacétylase (*nagA*). En aval du gène *oppD* sont retrouvés un gène codant

l'aminopeptidase P (*pepP*), un gène codant la phosphohydrolase (*phase*) et un gène codant la folypolyglutamate synthétase (*folC*).

La souche *Bifidobacterium longum* DJO 10A possède un cluster *dppI* de quatre gènes (*dppA1*, *dppB1*, *dppC1*, *dppF1*) homologue au cluster *ots* de *S. thermophilus* LMD-9 (Tableau 27). Les régions flanquant ce cluster sont dépourvues de transposases et similaires à la souche *Bi. bifidum* PRL2010 avec une différence en aval du gène *dppF1* (Figure 34C) où sont retrouvés le gène *ADPRase*, codant l'ADP-ribose pyrophosphatase, le gène *pepP*, codant la phosphohydrolase, un gène codant une protéine non identifiée et le gène *folC* qui code une folypolyglutamate synthétase.

Les séquences génomiques de 22 souches appartenant à 6 espèces du genre *Bifidobacterium* étant disponibles, une recherche systématique de cluster homologue au cluster *ots* de *S. thermophilus* LMD-9 a été menée.

La présence d'un cluster homologue a été détecté chez toutes les bactéries du genre *Bifidobacterium* sauf pour deux souches de l'espèce *Bi. breve*, les souches ACS-071-V-Sch8b et UCC2003 (Tableau 30). La comparaison entre les séquences protéiques déduites des clusters homologues pour ces 20 souches (5 espèces différentes) et celles du cluster *opp* de *Bi. bifidum* PRL2010 a révélé de fortes similitudes (78 à 100% d'identité, Tableau 31). De plus, chez le genre *Bifidobacterium*, les régions flanquant ce cluster sont homologues mais différentes de celles de *S. thermophilus* LMD-9 (Figure 34). Ce cluster serait donc bien conservé au sein du genre *Bifidobacterium* et pourrait faire partie du « core génome ». Ainsi, ces résultats suggèrent que le cluster *ots* de *S. thermophilus* pourrait provenir de *Bifidobacterium* dont certaines souches sont utilisées en industrie laitière. Néanmoins, ces résultats ne permettent pas de déterminer si le transfert s'est fait directement de *Bifidobacterium* à *S. thermophilus*, ou s'il a impliqué un ou plusieurs hôtes intermédiaires, tel que *Lc. lactis*, sachant tout de même, que le cluster *ots* n'est pas présent dans les génomes séquencés de cette bactérie.

**Tableau 30 : Présence d'un cluster homologue au cluster *ots* au sein des génomes de *Bifidobacterium* disponibles dans les banques de données**

En jaune, absence du cluster homologue au cluster *ots*.

<i>Bifidobacterium</i>	Présence du cluster homologue au cluster <i>ots</i> chez les souches de <i>Bifidobacterium</i>			
	<i>otsA</i>	<i>otsB</i>	<i>otsC</i>	<i>otsD</i>
<i>Bifidobacterium adolescentis</i> ATCC 15703	+	+	+	+
<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> AD011	+	+	+	+
<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> BB-12	+	+	+	+
<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> BLC1	+	+	+	+
<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> B1-04	+	+	+	+
<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> CNCM I-2494	+	+	+	+
<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> DSM 10140	+	+	+	+
<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> V9	+	+	+	+
<i>Bifidobacterium bifidum</i> PRL2010	+	+	+	+
<i>Bifidobacterium bifidum</i> S17	+	+	+	+
<i>Bifidobacterium breve</i> ACS-071-V-Sch8b				
<i>Bifidobacterium breve</i> UCC2003				
<i>Bifidobacterium dentium</i> Bd1	+	+	+	+
<i>Bifidobacterium longum</i> DJO10A	+	+	+	+
<i>Bifidobacterium longum</i> NCC2705	+	+	+	+
<i>Bifidobacterium longum</i> ssp. <i>infantis</i> 157F	+	+	+	+
<i>Bifidobacterium longum</i> ssp. <i>infantis</i> ATCC 15697	+	+	+	+
<i>Bifidobacterium longum</i> ssp. <i>longum</i> BBM68	+	+	+	+
<i>Bifidobacterium longum</i> ssp. <i>longum</i> F8	+	+	+	+
<i>Bifidobacterium longum</i> ssp. <i>longum</i> JCM 1217	+	+	+	+
<i>Bifidobacterium longum</i> ssp. <i>longum</i> JDM301	+	+	+	+
<i>Bifidobacterium longum</i> ssp. <i>longum</i> KACC 91563	+	+	+	+

**Tableau 31 : Pourcentages d'identité entre les séquences protéiques déduites des gènes du cluster *opp* de *Bifidobacterium bifidum* PRL2010 et leurs homologues chez d'autres espèces du genre *Bifidobacterium* dont les génomes sont entièrement séquencés**

<i>Bifidobacterium</i>	Pourcentages d'identités protéiques avec les séquences protéiques de la souche <i>Bi. bifidum</i> PRL2010			
	OppA	OppB	OppC	OppD
<i>Bifidobacterium bifidum</i> PRL2010	100	100	100	100
<i>Bifidobacterium bifidum</i> S17	99	99	100	99
<i>Bifidobacterium longum</i> DJO10A	87	90	82	89
<i>Bifidobacterium longum</i> NCC2705	87	88	87	89
<i>Bifidobacterium longum</i> ssp. <i>longum</i> BBMN68	87	88	82	89
<i>Bifidobacterium longum</i> ssp. <i>longum</i> F8	87	88	82	89
<i>Bifidobacterium longum</i> ssp. <i>longum</i> JCM 1217	87	88	82	89
<i>Bifidobacterium longum</i> ssp. <i>longum</i> JDM301	87	88	82	90
<i>Bifidobacterium longum</i> ssp. <i>longum</i> KACC 91563	87	87	82	89
<i>Bifidobacterium longum</i> ssp. <i>infantis</i> 157F	86	88	82	89
<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> AD011	85	84	78	79
<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> BB-12	85	84	78	79
<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> BLC1	85	84	78	79
<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> BI-04	85	84	78	79
<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> CNCM I-2494	85	84	78	79
<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> DSM 10140	85	84	78	79
<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> V9	85	84	78	79
<i>Bifidobacterium adolescentis</i> ATCC 15703	82	88	81	90
<i>Bifidobacterium dentium</i> Bd1	82	89	80	90
<i>Bifidobacterium longum</i> ssp. <i>infantis</i> ATCC 15697	82	89	84	89

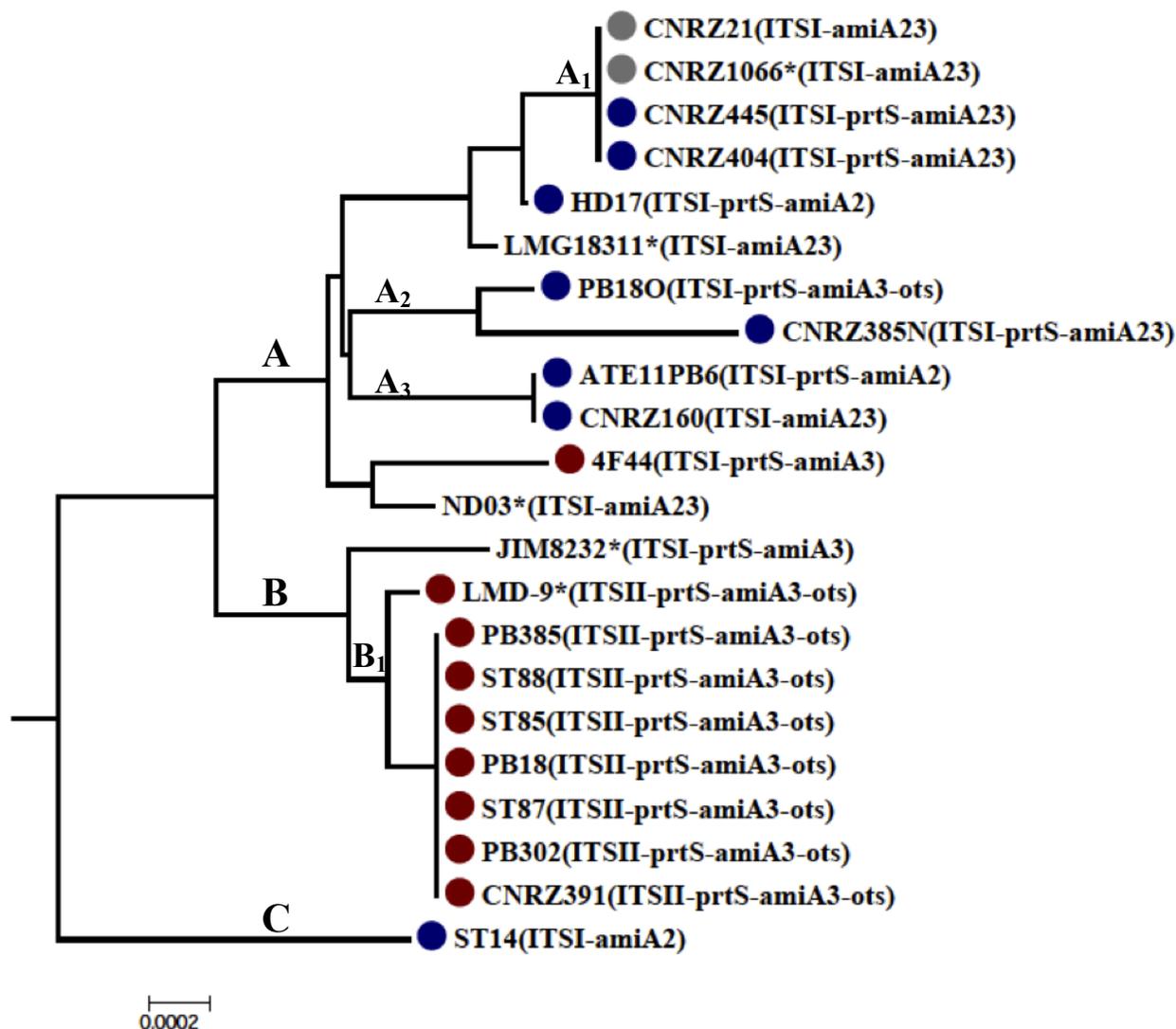
#### 1. 4. 8. Recherche par PCR-nichée de la présence des gènes du cluster *ots* au sein d'une collection de 19 souches de *Streptococcus thermophilus*

La présence du cluster *ots* a été recherchée au sein de notre collection de 19 souches. Les souches CNRZ1066 et LMD-9, dont les génomes ont été séquencés, ont servi de témoins positifs. Une approche de PCR nichée a été employée pour vérifier la présence des quatre gènes et leur organisation. Deux couples d'oligonucléotides (« obpLM for/obpLM rev » et « obpER for/obpER rev »), conçus à partir de la séquence du cluster de la souche *S. thermophilus* LMD-9 ont été utilisés (Tableaux 13, 16). Les oligonucléotides « obpLM for », et « obpLM rev » s'hybrident respectivement sur *otsA* et *otsD* alors que les oligonucléotides « obpER for » et « obpER rev » sur *otsB* et *otsC*. Des amplicons d'une taille attendue de 4129 pb (premier couple d'amorces) et 536 pb (deuxième couple d'amorces) ont été obtenus chez 9 souches. La présence ou l'absence du cluster *ots* chez l'ensemble des souches de notre collection est récapitulée dans le tableau 32. Il est intéressant de noter que, excepté la souche 4F44, toutes les souches à fort pouvoir acidifiant possèdent ce cluster, alors que toutes les souches à moyen ou faible pouvoir acidifiant, excepté la souche PB180, ne le possèdent pas. Sur la base de l'arbre phylogénétique établi précédemment, nous pouvons constater que toutes les souches appartenant à la branche B possèdent, à l'exception de la souche JIM8232, un ITS de type II, l'opéron *amiA1CDEF*, le transporteur *ntpT*, le gène *amiA3* et le cluster *ots*. Hormis la souche PB180 de la sous-branche A<sub>2</sub> de la branche A qui possède le cluster *ots*, toutes les autres souches des branches A et C possèdent un ITS de type I, l'opéron *amiA1CDEF*, le transporteur *ntpT*, les gènes *amiA2* et *amiA3* pour la plupart mais pas le cluster *ots* (Figure 36). Ainsi, la distribution du cluster *ots* au sein de notre collection de souches de *S. thermophilus* confirme l'hypothèse de l'acquisition récente de ce cluster.

**Tableau 32 : Récapitulatif des différentes caractéristiques des 22 souches de *Streptococcus thermophilus* étudiées**

Nom	Origine	Pouvoir acidifiant	Type d'ITS	<i>prtS</i>	Activité caséinolytique	Présence/absence des gènes :			
						<i>amiA1 CDEF</i>	<i>amiA2</i>	<i>amiA3</i>	<i>otsABCD</i>
CNRZ391	Yaourt	Fort	II	+	+	+	-	+	+
PB385	Yaourt		II	+	+	+	-	+	+
PB302	Yaourt		II	+	+	+	-	+	+
4F44	Fromage		I	+	+	+	-	+	-
LMD-9 <sup>a</sup>	Yaourt		II	+	+	+	-	+	+
PB18	Fromage		II	+	+	+	-	+	+
ST87	Fromage		II	+	+	+	-	+	+
ST88	Fromage		II	+	+	+	-	+	+
ST85	Fromage		II	+	+	+	-	+	+
CNRZ385N	Yaourt	Moyen	I	+	+	+	+	+	-
HD17	Yaourt		I	+	-	+	+	-	-
CNRZ404	Yaourt		I	+	+/-	+	+	+	-
ST14	Yaourt		I	-	-	+	+	-	-
CNRZ160	Yaourt		I	-	-	+	+	+	-
CNRZ445	Fromage		I	+	+/-	+	+	+	-
ATE11PB6	Yaourt		I	+	-	+	+	-	-
PB18O	Fromage		I	+	-	+	-	+	+
CNRZ1066	Yaourt	Faible	I	-	-	+	+	+	-
CNRZ21	Yaourt		I	-	-	+	+	+	-
LMG18311 <sup>a</sup>	Yaourt	Non étudié	I	-	Non étudiée	+	+	+	-
ND03 <sup>a</sup>	Lait		I	-	Non étudiée	+	+	+	-
JIM8232 <sup>a</sup>	Lait		I	+	Non étudiée	+	-	+	-

<sup>a</sup> Chez les souches LMD-9, CNRZ1066, LMG18311, ND03 et JIM8232 dont les génomes sont entièrement séquencés, les résultats de présence/absence des gènes ont été obtenus à partir des banques de données.



**Figure 36 : Arbre phylogénétique des souches de *Streptococcus thermophilus* utilisées dans ce travail**

Cet arbre est basé sur les séquences concaténées de 8 « gènes de ménage » et d'ITS. La présence/absence des gènes *amiA2* ou *amiA3*, du cluster *ots*, du gène *prtS* et du type d'ITS est signalée entre parenthèses tandis que le pouvoir acidifiant est mentionné par une pastille colorée

L'échelle indique le nombre de substitutions de nucléotides par site. Les étoiles correspondent aux souches dont les génomes sont complètement séquencés.

Le génotype (type ITS, présence des gènes *prtS*, *amiA2,3*, cluster *ots*) est mentionné entre parenthèses. Pour ce qui concerne le cluster *ots* et les gènes *prtS*, *amiA2* et *amiA3*, ils sont notés uniquement s'il sont présents, sachant que toutes les souches possèdent le gène *dtpT* et l'opéron *amiA1CDEF*.

Le pouvoir acidifiant selon Galia *et al.* (2009) de chaque souche de la collection du laboratoire est mentionné par une pastille colorée :

- fort pouvoir acidifiant
- moyen pouvoir acidifiant
- faible pouvoir acidifiant

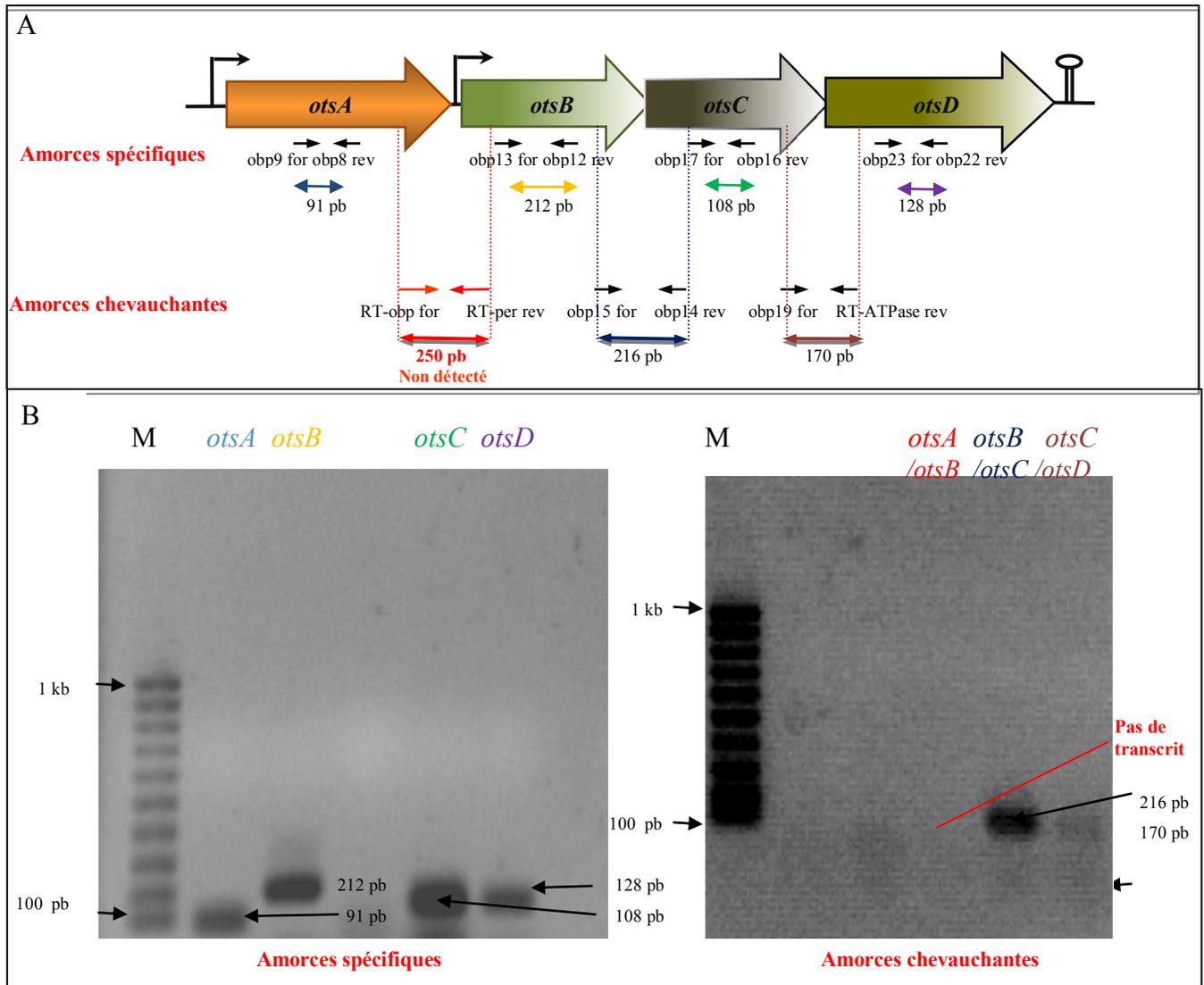
## **2. Séquençage du cluster *ots* et analyse des séquences obtenues chez les souches PB18O et PB302**

Sachant que la détection par PCR du cluster *ots* ne permet de déterminer ni sa variabilité de séquence ni sa fonctionnalité potentielle, il a été entièrement amplifié par PCR en utilisant plusieurs couples d'amorces (Tableaux 13, 16), puis séquencé chez les souches PB302 et PB18O. La souche PB302 a été choisie pour représenter les 8 souches de type ITSII à fort pouvoir acidifiant positionnées sur la même sous-branche ( $B_1$ ) de l'arbre phylogénétique présenté dans la Figure 36. La souche PB18O a, quant à elle, été choisie car c'est la seule souche à moyen pouvoir acidifiant qui possède le cluster (Figure 36). Il est à souligner que sur les 8 « gènes de ménage » partiellement séquencés chez ces deux souches, 5 présentent des allèles différents. La séquence du cluster *ots* établie chez ces deux souches est identique à celle de la souche LMD-9.

## **3. Organisation transcriptionnelle du cluster *ots* et expression au cours de la croissance des souches PB302, PB18O et LMD-9 en milieu M17**

Afin de déterminer si le cluster *ots* est un opéron et d'étudier son expression, des expériences de RT-PCR ont été menées chez les souches PB18O, PB302 et LMD-9 cultivées en milieu M17. Les ARN totaux ont été extraits en début, milieu, fin de phase exponentielle et, enfin, après 2 h de phase stationnaire. Après rétro-transcription des ARN extraits, les ADNc spécifiques de chaque gène ont été recherchés par PCR en utilisant des amorces spécifiques (« obp8 rev/obp9 for », « obp12 rev/obp13 for », « obp16 rev/obp17 rev », « obp 22 rev/obp 23 for ») (Tableaux 13, 16, Figure 37). Pour valider cette étude, le gène codant le facteur sigma ( $\sigma_{70}/\sigma_{32}$ ) a été utilisé comme témoin, ce gène étant supposé être exprimé à toutes les phases de croissance puisque, codant un facteur sigma qui contrôle des fonctions essentielles du métabolisme et assure ainsi la transcription de base chez les bactéries. Quel que soit le temps de croissance testé, des ADNc de chacun des cinq gènes (*otsA*, *otsB*, *otsC* et *otsD* et facteur *sigma*

$\sigma 70/\sigma 32$ ) ont été détectés chez les trois souches. La transcription des quatre gènes du cluster *ots* semble donc être constitutive chez les trois souches dans les conditions de culture testées.



**Figure 37 : Recherche des transcrits du cluster *ots* chez *Streptococcus thermophilus* par RT-PCR**

A. Position des amorces utilisées pour les réactions de RT-PCR sur le cluster *ots*.

B. Migration électrophorétique des produits d'amplification obtenus après RT-PCR chez *S. thermophilus*.

Amorces utilisées pour amplifier les transcrits *otsA* : « obp9 for/obp8 rev » ; *otsB* : « obp13 for/obp12 rev » ; *otsC* : « obp17 for/obp16 rev » ; *otsD* : « obp23 for/obp22 rev » ; *otsA/otsB* : « RT-obp for/RT-per rev » ; *otsB/otsC* : « obp15 for/obp14 rev » et *otsC/otsD* : « obp19 for/RT-ATPase rev ». Pour chaque transcrit, la taille des amplicons attendue est mentionnée sur le schéma. M. Marqueur de taille de 100 pb à 1kb.

Afin d'étudier l'organisation transcriptionnelle du système *ots*, des amorces chevauchant deux gènes différents ont été utilisées, (« RT-obp for/RT-per rev », « obp15 for/obp14 rev », « obp19 for/RT-ATPase » rev qui chevauchent les gènes *otsA-otsB*, *otsB-otsC* et *otsC-otsD* respectivement) (Figure 37A). Des amplicons de tailles attendues ont été obtenus, avec les amorces chevauchant *otsB-otsC* (216 pb) et *otsC-otsD* (170 pb). Par contre, aucun amplicon n'a été obtenu avec les amorces chevauchant *otsA-otsB*. Ces résultats suggèrent donc une organisation polycistronique des gènes *otsB*, *otsC* et *otsD*. Par contre, le gène *otsA* pourrait être transcrit indépendamment (Figure 37B). L'analyse de la séquence nucléotidique du cluster *ots* a permis de détecter, en amont du gène *otsA*, un promoteur potentiel constitué d'une boîte -10 et d'une boîte -35. En amont du gène *otsB*, seule une boîte -10 a été retrouvée (Figure 38). Les ORFs *otsA*, *otsB*, *otsC* et *otsD* sont précédées d'une séquence RBS putative. Un seul terminateur de transcription potentiel constitué d'une séquence pouvant adopter une structure tige-boucle a été retrouvé en aval du codon stop de l'ORF *otsD*. L'ensemble de ces données suggère que le cluster *ots* est fonctionnel, et qu'au moins les gènes *otsB*, *C* et *D* constituent un opéron.



## **4. Etude de l'expression du cluster *ots* par PCR-quantitative au cours de la croissance des souches PB302 et PB180 en milieu M17 et lait**

Lors de ses travaux de thèse, Wessam Galia a montré que si les souches PB302 et PB180 possèdent les mêmes allèles de différents gènes appartenant au système protéolytique (*prtS*, *amiA1CDEF*, *amiA3*, *codY*, *pepX*, *pepC*, *pepN*, *ntpT*, *livJHMGF*, *ilvC*, *ilvDBN*, *bcaT*, *ackA*, *ldh* et *relA*), elles les expriment très différemment lorsqu'elles sont cultivées en milieu lait. Dans le cadre de ce travail, nous avons montré que ces deux souches possèdent les mêmes allèles pour les gènes *ntpT* et du cluster *ots*. Aussi, nous avons quantifié l'expression des gènes *otsA*, *ntpT* et *amiF* chez ces deux souches. Ce dernier permet de quantifier l'expression de l'opéron *amiA1CDEF*.

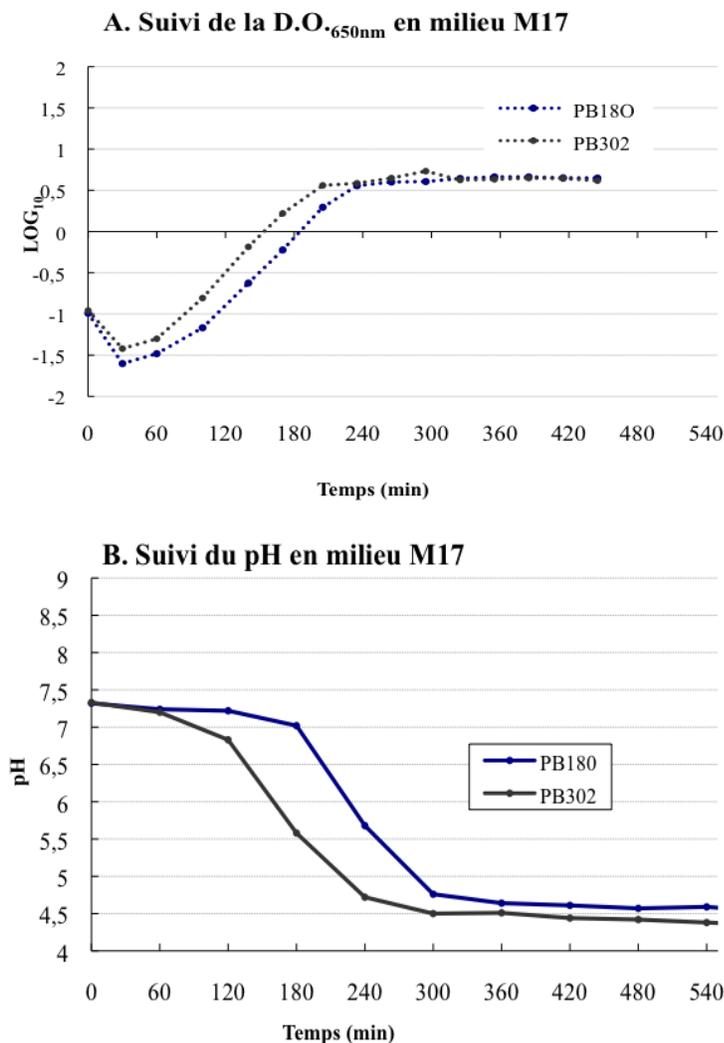
### **A. Croissance de *Streptococcus thermophilus* PB302 et PB180 en lait et en milieu M17**

La croissance des souches PB302 et PB180 a été suivie en milieu M17 (Figure 39) et en lait (Figure 40) par mesure de pH et détermination de la densité cellulaire. Ces deux milieux contiennent la même source de carbone, le lactose mais diffèrent au niveau de leur source azotée. En effet, le lait est pauvre en peptides et acides aminés libres et riche en protéines, alors que le milieu M17 est riche en peptides et acides aminés (hydrolysate de caséine, de viande, de soja, extrait de levure) (Terzaghi et Sandine, 1975).

En milieu M17, les deux souches possèdent une croissance similaire (Figure 39). Le temps nécessaire pour passer d'un pH extracellulaire de 6 (début de phase exponentielle) à un pH extracellulaire de 5 (fin de phase exponentielle/début de phase stationnaire) est le même pour les deux souches (proche de 50 min). La D.O.<sub>650 nm</sub> maximale atteinte en fin de phase exponentielle est de 3,6 pour la souche PB302 et 3,4 pour la souche PB180. Cependant quelques petites différences ont été observées (Figure 39). La phase de latence semble plus courte chez la souche

PB302 et le temps de génération de cette souche est légèrement inférieur (20 min contre 23 min pour la souche PB180).

En lait, les différences de croissance entre les souches semblent plus marquées (Figure 40). La  $D.O_{650\text{ nm}}$  maximale atteinte à la fin de la phase exponentielle est de 3,3 pour la souche PB180 contre 5,8 pour la souche PB302. Enfin, le temps nécessaire pour passer d'un pH extracellulaire de 6 (début de phase exponentielle) à un pH extracellulaire de 5 (fin de phase exponentielle/début de phase stationnaire) est de 86 minutes pour la souche PB302 contre 122 pour la souche PB180.



**Figure 39 : Suivi de croissance en milieu M17 des souches *Streptococcus thermophilus* PB180 et PB302 par détermination de la densité cellulaire (mesure de DO à 650 nm) « A » et mesure du pH « B »**

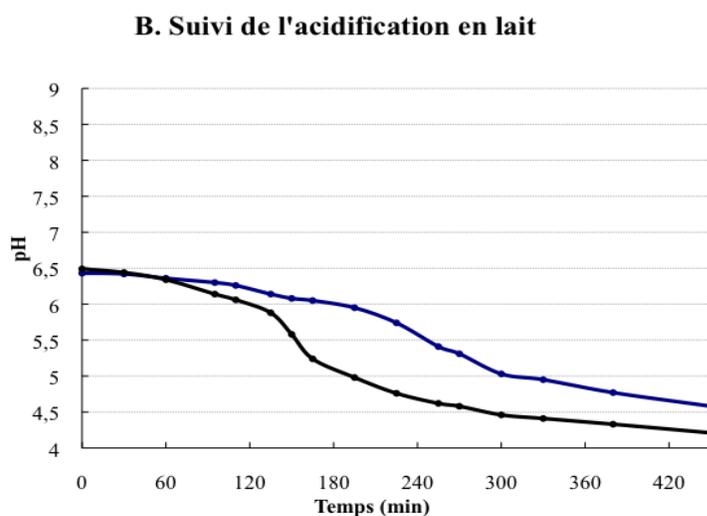
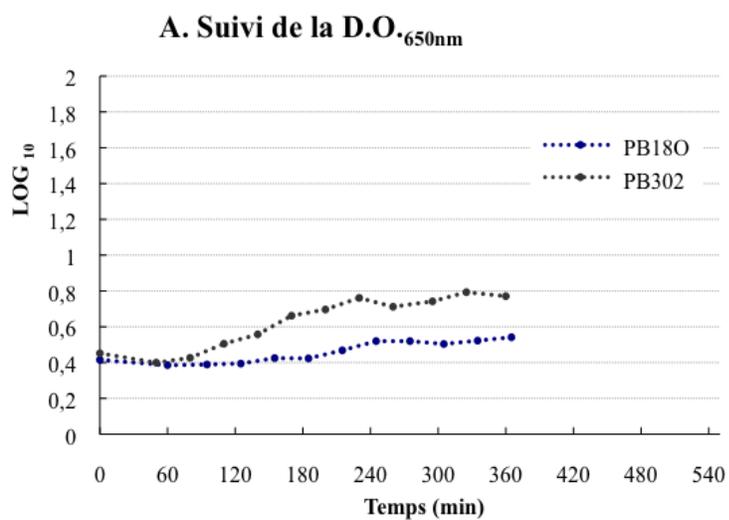


Figure 40 : Suivi de croissance en lait de *Streptococcus thermophilus* PB180 et PB302 par détermination de la densité cellulaire (mesure de DO à 480 nm) « A » (Thomas et Turner, 1977) et mesure du pH « B »

## **B. Quantification de l'expression des gènes *amiF*, *otsA* et *dtpT* à différents stades de croissance des souches PB18O et PB302 cultivées en milieu M17 par RT-PCR en temps réel**

L'expression des gènes *amiF* (opéron *amiA1CDEF*), *otsA* et *dtpT* a été analysée au cours de la croissance des souches PB18O et PB302 cultivées en milieu M17 par RT-PCR quantitative. Les ARN totaux ont été extraits en début (D.O.<sub>650 nm</sub> de 0,7 et pH de 6,5 et de 6,8 pour les souches PB302 et PB18O), milieu (D.O.<sub>650 nm</sub> de 2 et pH de 6 et de 6,2 pour les souches PB302 et PB18O), et fin de phase exponentielle/début de phase stationnaire (D.O.<sub>650 nm</sub> de 4 et pH de 4,7 et 5,1 pour les souches PB302 et PB18O). Après rétro-transcription des ARN extraits, les ADNc spécifiques des gènes *amiF*, *otsA* et *dtpT* ont été quantifiés par PCRq en utilisant les couples d'amorces spécifiques « *amiF*quant for/*amiF*quant rev », « *otsA*quant for/*otsA*quant rev » et « *dtpT*quant for/*dtpT*quant rev » respectivement (Tableau 14).

Dans le but de valider le Ct (Threshold Cycle) de chaque échantillon, la réaction a été réalisée en triplicat et l'ensemble du travail de PCRq a été réalisé deux fois. Une moyenne et un écart type ont été déterminés. Seul un ratio d'expression supérieur à 2,0 ou inférieur à 0,5 et observé dans les deux expériences indépendantes sera considéré dans la suite de nos analyses. Pour normaliser la quantité d'ADNc présente dans chaque échantillon, le niveau de transcription des gènes a été comparé à celui du gène codant le facteur sigma  $\sigma_{70}/\sigma_{32}$ , utilisé comme référence. Si les trois gènes présentent un niveau d'expression similaires chez les deux souches en début et milieu de phase exponentielle, les gènes *amiF* et *otsA* semblent s'exprimer différemment en début de phase stationnaire. En effet, le gène *amiF* semble moins exprimé chez la souche PB302 alors que c'est le cas inverse pour le gène *otsA* (Tableau 33).

**Tableau 33 : Comparaison des niveaux<sup>a</sup> d'expression des gènes *amiF*, *ntpT* et *otsA* entre les souches PB302 et PB180 cultivées en milieu M17**

Phase de croissance	Expression de gène <i>amiF</i> (PB302/PB180)	Expression de gène <i>ntpT</i> (PB302/PB180)	Expression de gène <i>otsA</i> (PB302/PB180)
Début de phase exponentielle (D.O. <sub>650 nm</sub> = 0,7)	0,78 ; 0,83 <b>0,80 ± 0,04</b>	4,53 ; 0,99 <b>2,76 ± 2,50</b>	2,2 ; 1,0 <b>1,6 ± 0,85</b>
Milieu de phase exponentielle (D.O. <sub>650 nm</sub> = 2)	0,40 ; 0,64 <b>0,52 ± 0,17</b>	1,71 ; 1,33 <b>1,52 ± 0,27</b>	1,0 ; 1,1 <b>1,05 ± 0,07</b>
Début de phase stationnaire (D.O. <sub>650 nm</sub> = 4)	0,21 ; 0,18 <b>0,19 ± 0,02</b>	4,38 ; 1,58 <b>2,98 ± 1,98</b>	6,9 ; 4,1 <b>5,5 ± 1,98</b>

<sup>a</sup> Il s'agit de taux moyens déterminés après 2 répétitions indépendantes. L'écart type de chacun de ces taux est également indiqué.

### C. Quantification de l'expression des gènes *amiF*, *otsA* et *ntpT* à différents stades de croissance des souches PB180 et PB302 cultivées en milieu lait par RT-PCR en temps réel

Dans le cas de cette étude, les ARN ont été extraits en fin de phase exponentielle (pH extracellulaire de 5,3) puisque des différences d'expression entre les deux souches avaient été observées pour cette phase lors de la croissance des souches en milieu M17. Les ARN ont été retrotranscrits et un ADNc spécifique de chacun des gènes étudiés recherché. L'expérience a été réalisée en triplicat et reproduite une fois. Comme cela est présenté dans le tableau 34, aucune différence significative d'expression de chacun des gènes testés n'a été détectée dans ces conditions.

**Tableau 34 : Comparaison des niveaux<sup>a</sup> d'expression des gènes *amiF*, *ntpT* et *otsA* entre les deux souches étudiées (souche PB302/souche PB180) cultivées en lait**

Phase de croissance	Expression de gène <i>amiF</i> (PB302/PB180)	Expression de gène <i>ntpT</i> (PB302/PB180)	Expression de gène <i>otsA</i> (PB302/PB180)
Milieu de phase exponentielle (pH = 5,3)	0,79 ; 1,42 <b>1,11 ± 0,44</b>	2,76 ; 0,88 <b>1,82 ± 1,33</b>	1,9 ; 1,1 <b>1,5 ± 0,57</b>

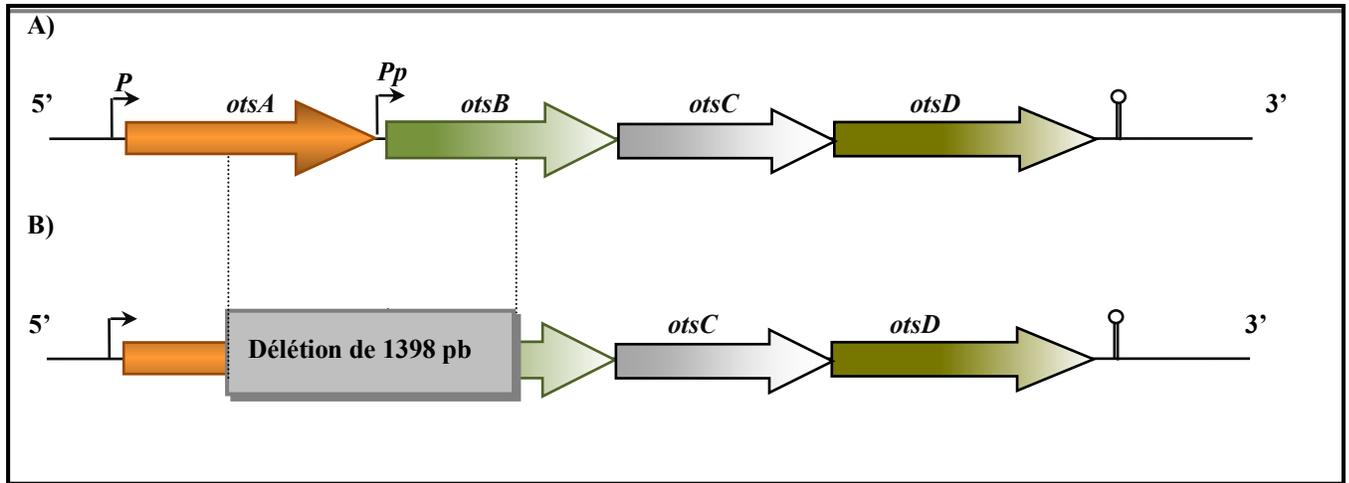
<sup>a</sup> Il s'agit de taux moyens déterminés après 2 répétitions indépendantes. L'écart type de chacun de ces taux est également indiqué.

## D. Conclusion

L'ensemble des résultats obtenus suggère ainsi que les différences de croissance observées entre les souches PB18O et PB302 n'impliquent pas de fortes variations d'expression des gènes de transport de peptides testés. Le cluster *ots* apparaît fonctionnel dans les deux souches d'après sa séquence, trois des 4 gènes sont organisés en opéron, et le cluster est transcrit tout au long de la croissance en milieu M17. Enfin, le gène *otsA* est exprimé à des taux similaires chez deux souches le possédant et appartenant à des groupes phylogénétiques différents.

## 5. Rôle du système Ots

Comme cela a été évoqué précédemment, le système Ots est annoté en tant que transporteur de peptides et de nickel, systèmes qui ne peuvent être distingués sur la base de leur structure primaire (Tam et Saier, 1993). Aussi, un mutant  $\Delta ots$  (LMD-9 $\Delta ots$ ) a été construit chez la souche LMD-9 afin de déterminer la fonction du cluster. Chez ce mutant, un fragment de 1398 pb correspondant à l'extrémité 3' du gène *otsA* et à l'extrémité 5' du gène *otsB* (soit 1028 pb du gène *otsA*, 102 pb de la région intergénique et 268 pb du gène *otsB* (Figure 20) a été délété. De fait, cette construction aboutit à l'inactivation des gènes *otsA* et *B*. Dans la mesure où le promoteur du gène *otsA* n'est pas délété et il peut y avoir transcription du gène mutant et synthèse, après traduction, d'une protéine OtsA tronquée. Pour ce qui concerne les gènes *otsC* et *D*, la délétion ne les affectant pas, ils continuent à coder des protéines fonctionnelles. Cependant, le promoteur putatif que nous avons identifié en amont du gène *otsB*, et à partir duquel l'opéron *ostBCD* pourrait être transcrit, étant délété chez LMD-9 $\Delta ots$ , il est peu probable que les protéines OtsC et D soient synthétisées (Figure 41).



**Figure 41 : Représentation schématique de l'organisation génétique des clusters *ots* et  $\Delta$ *ots***

(A) cluster *ots* de la souche sauvage LMD-9 ; (B) cluster  $\Delta$ *ots* de la souche LMD-9 $\Delta$ *ots*.

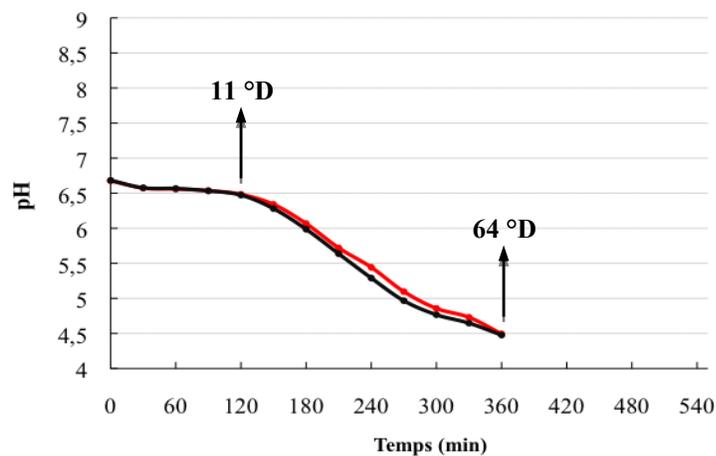
*P* : promoteur du gène *otsA* ; et *Pp* : promoteur putative du gène *otsB*.

## 5. 1. Croissance des souches LMD-9 et LMD-9 $\Delta$ *ots* en présence de source azotée différentes

La croissance des souches LMD-9 et LMD-9 $\Delta$ *ots* a été comparée dans trois milieux de croissance qui diffèrent notamment au niveau de leur source azotée. Il s'agit des milieux lait, M17 et YL. Le lait contient peu d'acides aminés libres et d'oligopeptides et est riche en protéines (caséines notamment). Pour une croissance à haute densité cellulaire, *S. thermophilus* doit donc être capable d'hydrolyser ces protéines et d'incorporer les peptides générés. Les deux souches testées possédant la protéase PrtS, l'étape de production des peptides à partir des caséines est donc réalisée seule l'étape d'internalisation des peptides peut différer. Le milieu M17 est fréquemment utilisé pour la croissance de *S. thermophilus*, il est riche en acides aminés et en peptides (hydrolysate de caséine, de viande, de soja, et d'extrait de levure) et est adapté pour la croissance des bactéries lactiques car tamponné (ajout de  $\beta$ -glycérophosphate) (Terzaghi et Sandine, 1975). Enfin, le milieu YL est un milieu non tamponné et composé d'une source azotée en quantité limitée, présente dans l'extrait de levure.

### 5. 1. 1. Croissance des souches LMD-9 et LMD-9 $\Delta$ ots en lait

La croissance des souches LMD-9 et LMD-9 $\Delta$ ots en milieu lait a été estimée par suivi du pH et détermination de l'acidité Dornic du milieu extracellulaire. Aucune différence significative n'a été observée entre les deux souches testées (Figure 42). L'absence d'un système Ots fonctionnel ne semble donc pas représenter un facteur limitant pour la croissance de la souche mutante LMD-9 $\Delta$ ots en lait. Par conséquent, il semble que l'acquisition de ce système par transfert horizontal ne confère pas un avantage pour la croissance de la bactérie en milieu lait.



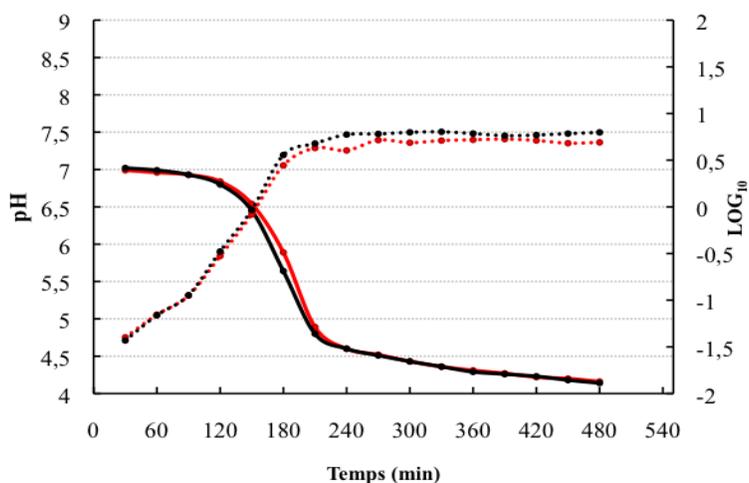
**Figure 42 : Croissance des souches LMD-9 et LMD-9 $\Delta$ ots en lait**

La croissance en lait a été estimée par mesures du pH du milieu extracellulaire et de l'acidité Dornic. Les valeurs d'acidité Dornic après 2 et 6 h de croissance sont mentionnés en degré Dornic (°D).

En noir : souche sauvage LMD-9 ; en rouge : souche mutante LMD-9 $\Delta$ ots. Les suivis de croissance en lait ont été reproduits 5 fois

### 5. 1. 2. Croissance des souches LMD-9 et LMD-9 $\Delta$ ots en milieu M17

En milieu M17, la croissance des deux souches a été évaluée par mesure du pH extracellulaire et détermination de la densité cellulaire (densité optique à 650 nm (D.O.<sub>650 nm</sub>)) (Figure 43). Les courbes d'acidification obtenues sont similaires chez les deux souches. Dans les deux cas, la baisse de pH est significative après 2 h de croissance. Pour ce qui concerne la densité cellulaire, il a été constaté dans 5 expériences indépendantes, que la souche sauvage atteignait des valeurs de D.O.<sub>650 nm</sub> finales légèrement supérieures à celles de la souche mutante. En effet, après 8 h de croissance (phase stationnaire), la souche LMD-9 atteint une valeur de D.O.<sub>650 nm</sub> proche de 6 alors que la souche LMD-9 $\Delta$ ots atteint une valeur proche de 5. Deux paramètres importants interviennent dans l'entrée en phase stationnaire de *S. thermophilus* : l'acidité du milieu et sa capacité à incorporer des nutriments. La présence d'une densité cellulaire légèrement plus faible en phase stationnaire chez la souche mutante pourrait donc provenir soit d'une sensibilité plus importante de cette souche à l'acidité du milieu ou de sa plus faible capacité de cette souche à incorporer des nutriments tels que des peptides.



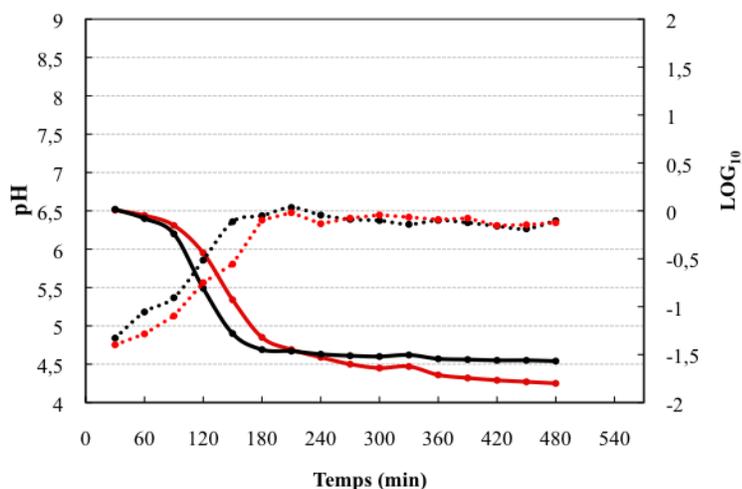
**Figure 43 : Croissance des souches LMD-9 et LMD-9 $\Delta$ ots en milieu M17**

La croissance en milieu M17 a été évaluée par mesure du pH du milieu extracellulaire et détermination de la D.O.<sub>650 nm</sub>.

En noir : souche sauvage LMD-9 ; en rouge : souche mutante LMD-9 $\Delta$ ots ; les traits continus représentent les suivis de pH et les traits pointillés de D.O.<sub>650 nm</sub>. Les suivis de croissance en milieu M17 ont été reproduits 5 fois.

### 5. 1. 3. Croissance des souches LMD-9 et LMD-9 $\Delta$ ots en milieu YL

Comme en milieu M17, la croissance des deux souches a été suivie par détermination du pH extracellulaire et de la densité cellulaire. Dans ce milieu et dans 9 expériences indépendantes, il a été constaté que la croissance des souches LMD-9 et LMD-9 $\Delta$ ots différait dès le démarrage, avec un retard de croissance du mutant d'environ 30 minutes (Figure 44). En phase exponentielle de croissance, la souche sauvage atteint un pH extracellulaire de 5,0 au bout de 140 min alors que la souche mutante l'atteint après 170 min. Les deux souches entrent en phase stationnaire à pH 4,9 après 2,5 h de croissance pour la souche sauvage et 3 h pour la souche mutante. Si le pH extracellulaire de la souche sauvage se stabilise à des valeurs proches de 4,6, celui de la souche mutante continue à diminuer pour atteindre une valeur de 4,2 après 8 h de croissance. Pour ce qui concerne les densités cellulaires, les deux souches présentent des courbes similaires et atteignent une densité finale située autour de  $D.O_{.650\text{ nm}} = 0,7$ . Les retards de croissance de la souche mutante détectés lors du suivi du pH extracellulaire sont retrouvés lors du suivi de la  $D.O_{.650\text{ nm}}$ . Par contre, les densités cellulaires en fin de croissance apparaissent similaires chez les deux souches.



**Figure 44 : Croissance des souches LMD-9 et LMD-9 $\Delta$ ots en milieu YL**

La croissance a été évaluée par mesure du pH du milieu extracellulaire et de la  $D.O_{.650\text{ nm}}$ .  
En noir : souche sauvage LMD-9 ; en rouge : souche mutante LMD-9 $\Delta$ ots ; les traits continus représentent les suivis de pH et les traits pointillés de  $D.O_{.650\text{ nm}}$ . Les suivis de croissance en milieu YL ont été reproduits 9 fois.

#### 5. 1. 4 Conclusion

Si l'étude de la croissance des souches LMD-9 et LMD-9 $\Delta$ *ots* n'a pas de révélée de différence en milieu lait, des différences tant au niveau de la densité cellulaire que du pH extracellulaire ont été observées lorsque les souches étaient cultivées en milieu M17 et YL.

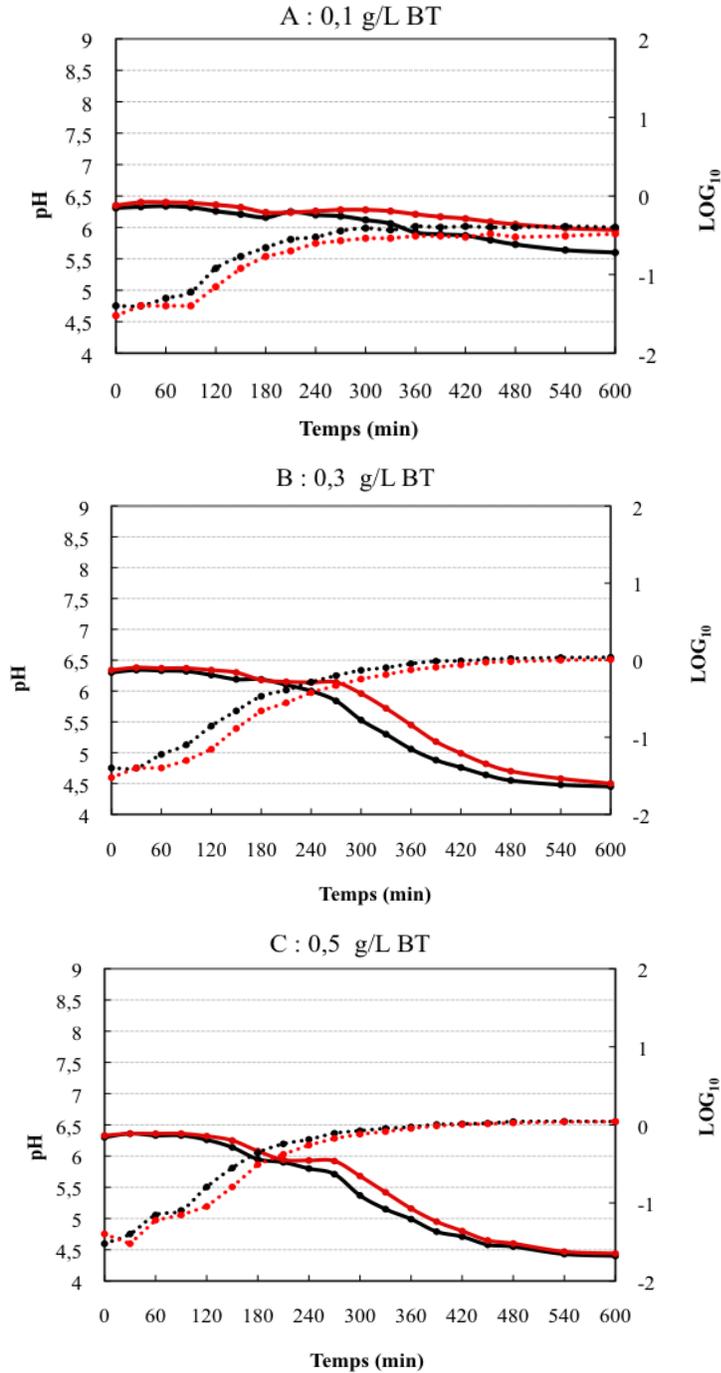
Il a été observé en milieu M17, que la densité cellulaire atteinte en fin de phase exponentielle par la souche sauvage LMD-9 est légèrement supérieure à celle atteinte par le mutant LMD-9 $\Delta$ *ots*. En milieu YL, cette différence n'est pas retrouvée en fin de phase exponentielle mais lors de la phase exponentielle de croissance des deux souches, un léger retard de croissance de la souche mutante a été observé (Figure 44). Ces deux résultats pourraient indiquer que la souche mutante présente ce retard en raison de sa déficience pour l'un de ses systèmes de transport de peptides, et donc de là, que le cluster *ots*, intervient dans ce transport. Concernant le pH extracellulaire, une différence a été observée en milieu YL uniquement. Dans ce milieu, si le pH extracellulaire reste stable (autour de 4,5) lorsque la souche sauvage est en phase stationnaire, il continue à baisser dans le cas de la souche mutante LMD-9 $\Delta$ *ots*, comme si celle-ci continuait à produire de l'acide lactique ou n'était pas capable de produire de l'ammoniaque qui tamponne le milieu. Cette observation apparaît donc contradictoire avec l'observation précédente. En effet, la différence observée entre les souches au niveau de la densité cellulaire nous a amenée à proposer que l'incorporation de peptides était à l'origine des différences observées entre les deux souches, alors que le suivi du pH extracellulaire, nous amène à proposer que la production d'ammoniaque pourrait varier entre les deux souches. Or la production d'ammoniaque résulte de l'activité de l'uréase, enzyme qui nécessite la présence de nickel pour son fonctionnement. En d'autres termes, si ces suivis de croissance ont permis de détecter des différences entre la souche sauvage et mutante, ils ne nous permettent pas de conclure quant au rôle du cluster *ots*.

## 5. 2. Croissance des souches LMD-9 et LMD-9 $\Delta$ ots en milieu chimiquement défini supplémenté en peptides

Afin de savoir si la différence de densité cellulaire observée en phase stationnaire entre les deux souches dans certaines conditions de culture pouvait résulter, au moins en partie, d'une différence de capacité de transport de peptides, nous avons utilisé un milieu de culture chimiquement défini (Letort et Juillard, 2001) dont les acides aminés libres ont été remplacés par des peptides en concentration variable (0,1, 0,3 ou 0,5 g.L<sup>-1</sup>). Deux sources de peptides ont été testées. La première, le bacto-tryptone (BT), correspond à un hydrolysate pancréatique de caséine contenant majoritairement des oligopeptides (80% d'oligopeptides et 20% de di- tri-peptides et d'acides aminés d'après le constructeur (Fluka)). La deuxième, les bacto-casamino acids (BCA), est issue d'une hydrolyse acide de caséine et contient majoritairement des acides aminés libres et des di et tripeptides (20% d'oligopeptides et 80% de di- tri-peptides et d'acides aminés d'après le constructeur (Becton, Dickinson and Company)).

En présence de bacto-tryptone, un retard de croissance est observé chez la souche mutante LMD-9 $\Delta$ ots (Figure 45 A, B et C). Cela est décelable par suivi du pH extracellulaire et de la densité cellulaire. A des concentrations de 0,3 et 0,5 g.L<sup>-1</sup> de BT, le pH extracellulaire et la densité cellulaire qui sont atteints par les deux souches en fin de phase stationnaire sont similaires (Figure 45 B et C et Tableau 35).

Par contre, lorsque la concentration en bacto-tryptone est plus faible (0,1 g.L<sup>-1</sup>), cette différence demeure. Ainsi, ces expériences suggèrent que si le transport d'oligopeptides est actif chez les deux souches, il apparaît plus efficace chez le sauvage.



**Figure 45 : Croissance des souches LMD-9 et LMD-9Δots en milieu chimiquement défini supplémenté en bacto-tryptone**

Trois concentrations de bacto-tryptone ont été testées (0,1 ; 0,3 ; 0,5 g.L<sup>-1</sup>).

La croissance a été évaluée par mesures du pH du milieu extracellulaire et de la D.O.<sub>650 nm</sub>.

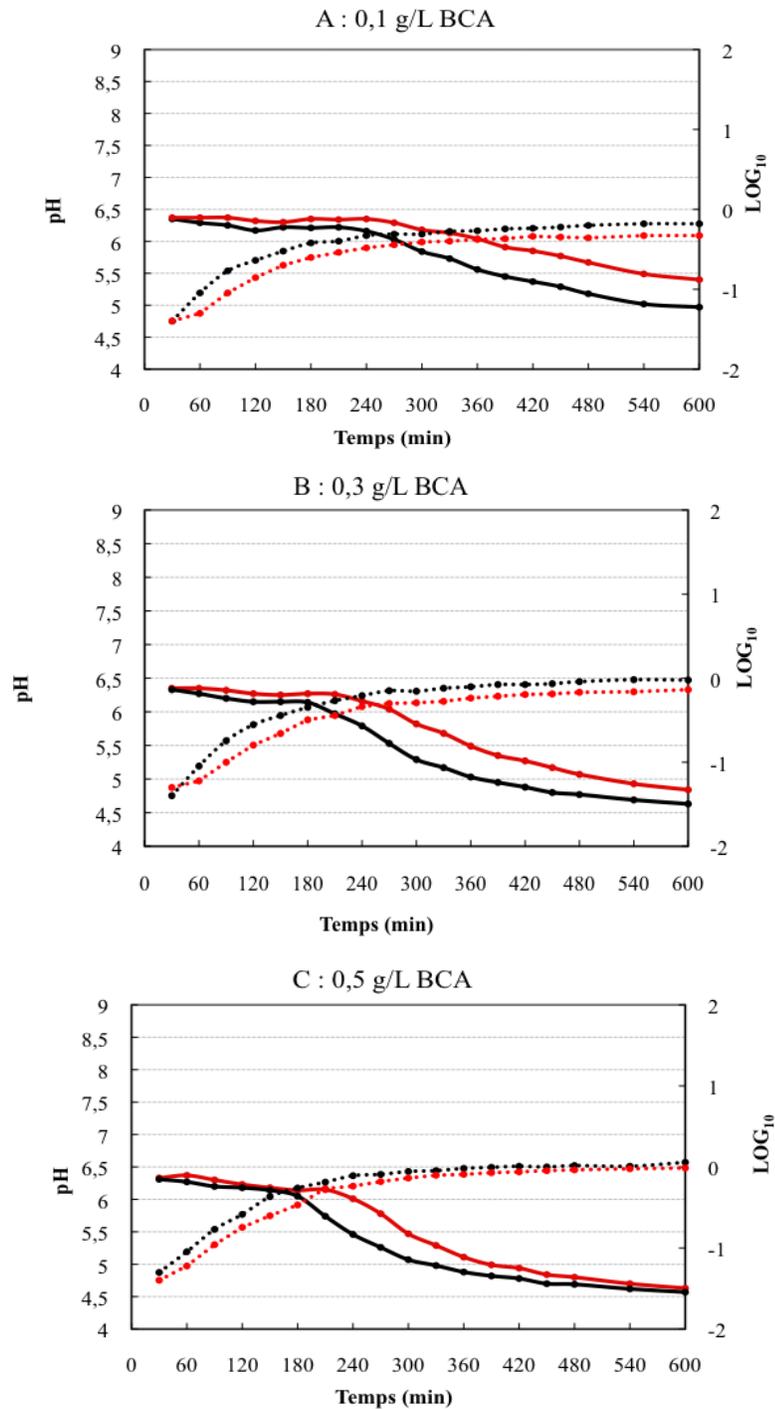
En noir : la souche sauvage LMD-9, en rouge : la souche mutante LMD-9Δots, les traits continus : pH et les traits pointillés : D.O.<sub>650 nm</sub>. Les suivis de croissance dans ces conditions ont été reproduits 3 fois.

**Tableau 35 : Tableau récapitulatif des valeurs de pH et de D.O. <sub>650 nm</sub> obtenues après 10 h de croissance en milieu chimiquement défini supplémenté en bacto-tryptone**

	pH		D.O. <sub>650 nm</sub>	
	après 10 h de culture		après 10 h de culture	
	LMD-9	LMD-9 $\Delta$ <i>ots</i>	LMD-9	LMD-9 $\Delta$ <i>ots</i>
0,1 g.L <sup>-1</sup> BT	5,60	6,00	0,40	0,30
0,3 g.L <sup>-1</sup> BT	4,40	4,40	1,00	1,00
0,5 g.L <sup>-1</sup> BT	4,40	4,40	1,00	1,00

En présence de bacto-casamino acids, le retard de croissance observé entre les souches sauvage et mutante est plus marqué que précédemment (Figure 46). En outre, en présence de 0,1 et 0,3 g.L<sup>-1</sup> de bacto-casamino acids, la souche mutante LMD-9 $\Delta$ *ots* ne récupère pas son retard (Figure 46 A et B et Tableau 36).

Ainsi, les différences de croissance observées entre les souches sauvage et mutante pourraient résulter d'une capacité moindre de la souche LMD-9 $\Delta$ *ots* à transporter les di- tri-peptides présent dans les bacto-casamino acids. Dans le cadre de cette hypothèse la différence observée en bacto-tryptone est moins marquée car la source azotée étant davantage constituée de peptides de plus grande taille qui peuvent être internalisés par le système Ami, qui est identique chez les deux souches. Il est à noter cependant qu'on ne peut formellement exclure l'hypothèse selon laquelle le système Ots serait impliqué dans le transport d'oligopeptides et que l'effet de son inactivation ne serait visible que lorsque la concentration en oligopeptide est faible.



**Figure 46 : Croissance des souches LMD-9 et LMD-9 $\Delta$ ots en milieu chimiquement défini supplémenté en bacto-casamino acides**

Trois concentrations de bacto-casamino acides ont été testées (0,1 ; 0,3 ; 0,5 g.L<sup>-1</sup>). La croissance a été évaluée par mesures du pH du milieu extracellulaire et de la D.O.<sub>650 nm</sub>. En noir : la souche sauvage LMD-9, en rouge : la souche mutante LMD-9 $\Delta$ ots, les traits continus : pH et les traits pointillés : D.O.<sub>650 nm</sub>. Les suivis de croissance dans ces conditions ont été reproduits 3 fois.

**Tableau 36 : Tableau récapitulatif des valeurs de pH et de D.O.  $_{650\text{ nm}}$  obtenues après 10 h de croissance en milieu chimiquement défini supplémenté en bacto-casamino acides**

	pH		D.O. $_{650\text{ nm}}$	
	après 10 h de culture		après 10 h de culture	
	LMD-9	LMD-9 $\Delta$ <i>ots</i>	LMD-9	LMD-9 $\Delta$ <i>ots</i>
0,1 g.L <sup>-1</sup> BCA	5,00	5,40	0,70	0,50
0,3 g.L <sup>-1</sup> BCA	4,60	4,80	1,00	0,70
0,5 g.L <sup>-1</sup> BCA	4,60	4,60	1,10	1,00

## 6. Discussion

Un des objectifs de l'équipe PB2P de l'URAFPA est de sélectionner des souches de *S. thermophilus* capables de produire des peptides bioactifs directement dans les produits laitiers. Cela implique d'étudier deux aspects complémentaires : la capacité de la bactérie à produire des peptides d'intérêt à partir des protéines du lait et sa capacité à les internaliser. C'est ce deuxième aspect qui a été abordé dans ce chapitre où nous nous sommes plus particulièrement intéressés à la variabilité des systèmes de transport de peptides. C'est ainsi que nous avons étudié au sein d'une collection de souches, dont nous avons déterminé la proximité phylogénétique, la variabilité du système Ami, DtpT et rechercher d'autre(s) système(s) susceptible(s) d'intervenir dans le transport de peptides chez *S. thermophilus*.

Pour ce qui concerne le système Ami, nous avons montré que l'opéron *amiAICDEF* est présent chez toutes les souches étudiées de la collection et une variabilité en termes de présence/absence des gènes *amiA2* et *amiA3* a été retrouvée. La variabilité observée dans le cadre de cette étude apparaît similaire à ce qui a été observé par Garault *et al.* (2002). Trois souches possèdent uniquement le gène *amiA2*, 11 souches possèdent le gène *amiA3* et 8 souches possèdent les deux. D'après l'arbre phylogénétique établi, il apparaît que toutes les souches qui présentent de fortes capacités acidifiantes, donc une bonne croissance en lait, et dont la majorité appartient à la branche B de l'arbre phylogénétique établi (Figure 36), sont dépourvues du gène *amiA2*. Il apparaît donc que la perte de ce gène n'altère pas les capacités de *S. thermophilus* à croître,

comme cela avait été observé précédemment (Garault *et al.*, 2002). Quant au transporteur DtpT qui pourrait intervenir dans le transport de di- et tri-peptides, le gène *dtpT* a été détecté chez toutes les souches de *S. thermophilus* étudiées.

Les analyses *in silico* réalisées à partir de la séquence du génome de *S. thermophilus* LMD-9 nous ont permis de détecter un nouveau cluster, spécifique de la souche LMD-9, et annoté comme codant un système de transport de peptides et de nickel de type ABC, cluster que nous avons appelé *ots*. Plusieurs résultats montrent que ce cluster a été acquis par transfert horizontal. Tout d'abord, il est absent des autres streptocoques dont les génomes sont disponibles dans les banques de données et notamment chez les deux espèces proches de *S. thermophilus*, à savoir *S. salivarius* et *S. vestibularis*. Il n'est pas présent chez toutes les souches de *S. thermophilus* que nous avons étudiées, il est situé dans une région où des événements de transfert horizontal ont été détectés et enfin, les protéines constituant ce système de transport partagent plus de 50% d'identité avec des protéines homologues présentes chez des actinobactéries du genre *Bifidobacterium* et *Brevibacterium*. Ainsi, nous avons conclu que ce système provenait d'actinobactéries et qu'il avait été acquis récemment par *S. thermophilus*.

L'analyse transcriptionnel du cluster *ots* suggère qu'il est exprimé de manière constitutive, au cours de la croissance en milieu M17, chez les souches PB18O, PB302 et LMD-9. D'autre part, un transcrite polycistronique des gènes *otsB*, *otsC* et *otsD* a été détecté. Le gène *otsA* serait, quant à lui, transcrite indépendamment. La comparaison de la croissance de la souche mutante LMD-9 $\Delta$ *ots* avec celle de la souche sauvage LMD-9 a permis de conclure que le cluster *ots* participe à l'internalisation des peptides. D'autre part, un rôle dans l'internalisation du nickel a pu être exclu. Cependant le système *ots* ne semble pas conférer un avantage pour la croissance de *S. thermophilus* en lait, car l'inactivation de ce système chez la souche LMD-9 ne paraît pas être un facteur limitant pour sa croissance.

## **6. 1. Transfert horizontal chez *Streptococcus thermophilus***

### **6. 1. 1. Variabilité de la région génomique contenant le cluster *ots***

Le cluster *ots* est absent chez 4 souches (CNRZ1066, LMG18311, ND03 et JIM8232) dont le génome a été séquencé et il est situé dans une région dont les gènes, acquis par transfert

horizontal, différent d'une souche à l'autre (Figure 30). Ainsi, chez la souche JIM8232, un îlot génomique portant des gènes conférant une pigmentation jaune et présentant des homologies avec ceux de *S. agalactiae* est présent dans cette région (Delorme *et al.*, 2011). Cet événement, est un exemple d'échange génétique entre espèces phylogénétiquement proches mais de proximité écologique éloignée, bien que *S. agalactiae*, puisse être retrouvée dans du lait cru en raison de son implication dans des mammites subcliniques. Chez les souches LMG18311, CNRZ1066 et ND03, des gènes impliqués dans la synthèse de l'antibiotique et acquis eux aussi par transfert horizontal sont présents dans cette région (Hols *et al.*, 2005).

### 6. 1. 2. Détection de transposon composite contenant le cluster *ots*

Plusieurs IS sont présents dans la région contenant le cluster *ots* que ce soit chez les souches CNRZ1066, LMG18311, ND03 et JIM8232 ou chez la souche LMD-9. Cependant, chez cette dernière, le nombre d'IS détectés est supérieur. La présence de ces éléments constitue un indice d'événement(s) de transfert horizontal (Zaneveld *et al.*, 2008).

Lors de nos analyses, nous avons détecté deux transposons composites potentiels, dénommés TCp1 et TCp2, qui délimitent le cluster *ots* (Figure 34A). De tels éléments sont connus dans la littérature pour véhiculer les séquences encadrées grâce aux deux IS qui les bornent (Roberts *et al.*, 2008). TCp1 contient la totalité du cluster *ots* et est borné par des copies d'ISS1 de type  $\alpha$  qui sont typiques de *Lc. lactis*. TCp2, quant à lui, contient TCp1 et est borné par deux copies d'IS1193, élément typique de *S. thermophilus*. Le fait que TCp1 soit flanqué de 2 éléments typiques de *Lc. lactis* suggère que le cluster *ots* a été acquis *via* cet élément. Outre la présence des 2 copies d'IS1193 qui le flanquent, TCp2 comporte les gènes *gly* et *udp* qui sont présents dans le génome de *S. thermophilus* ND03, mais à une localisation différente. Ainsi, les caractéristiques de cet élément suggèrent qu'il a été l'objet de 2 événements : un événement d'intégration de TCp1 et un événement de transposition dans un autre environnement génomique, l'ordre de ces deux événements étant impossible à déterminer. Les éléments transposables sont connus pour provoquer de nombreux remaniements génomiques tels que des délétions, inversions ou des insertions (Bennett, 2004). Environ 75% des IS trouvées chez *S. thermophilus* ont été associées à des événements de réarrangements génomiques (Hols *et al.*, 2005).

### 6. 1. 3. Origine du cluster *ots*

Des analyses *in silico* ont montré que toutes les bactéries possédant un cluster homologue au cluster *ots* appartiennent au phylum des *Actinobacteria* telles que des bactéries du genre *Bifidobacterium* et *Brevibacterium*. Dans tous les cas, il s'agit de bactéries éloignées d'un point de vue phylogénétique de *S. thermophilus* mais pouvant partager la même niche écologique, le lait. Ainsi, ce cluster est retrouvé chez des espèces utilisées dans la fabrication des produits laitiers probiotiques, telles que *Bi. bifidium*, *Bi. longum*, *Bi. infantis*, ou dans l'affinage de fromages telle que *Brevibacterium linens*. Leurs génomes se caractérisent par un haut pourcentage en GC (supérieur à 50%) par rapport aux bactéries lactiques qui appartiennent au phylum des *Firmicutes* et dont le pourcentage en GC est inférieur à 50% (Holzapfel *et al.*, 2001). Très peu d'études rapportent des échanges d'information génétique entre des bactéries du genre *Streptococcus* et du genre *Bifidobacterium* ou *Brevibacterium*. La recherche de biais de composition en nucléotides du génome de *S. thermophilus* a conduit Liu *et al.* (2009b) à détecter chez la souche CNRZ1066 des gènes impliqués dans la biosynthèse d'EPS et pouvant provenir des bactéries du genre *Bifidobacterium*.

Généralement, les séquences acquises récemment par transfert horizontal possèdent des caractéristiques différentes de celles de l'hôte. Ces caractéristiques sont utilisées pour leur recherche systématique au sein des génomes (Garcia-Vallve *et al.*, 2003). Une d'entre-elles est le pourcentage en bases GC. En effet, celui-ci est relativement homogène dans un génome et peu variable au sein d'une même espèce bactérienne, alors qu'il varie de façon importante entre espèces bactériennes (Muto et Osawa, 1987). Une corrélation entre ce caractère et la niche écologique occupée par la bactérie a été constatée. Les génomes des bactéries dépendantes d'un hôte (endosymbiotes ou pathogènes obligatoires) se caractérisent par une forte teneur en bases AT. Le pourcentage en GC est, en moyenne, de 38% pour les génomes des bactéries dépendant d'un hôte et de 49% pour les génomes de celles vivant de façon autonome. Pour expliquer une telle différence, l'hypothèse avancée implique le coût énergétique qui est plus élevé pour la synthèse des nucléotides GTP/CTP que pour la synthèse des ATP/UTP. Le biais en bases AT serait donc sélectionné car plus avantageux pour la cellule et les mutations de (G ou C) en (A ou T) avantageuses. Ainsi, dans des conditions de ressources nutritives limitées, la richesse en AT deviendrait un avantage sélectif. La même hypothèse peut être appliquée sur les séquences

acquises par transfert horizontal, y compris les éléments génétiques mobiles comme les plasmides, transposons, IS et les phages qui présentent une composition en bases GC inférieure de 2 à 4% au pourcentage du génome hôte. Les séquences exogènes étant en compétition avec le reste du génome, leur intégration et donc leur maintien dans le génome seraient favorisés par un abaissement de leur coût énergétique (Rocha et Danchin, 2002).

Si d'après les séquences protéiques il semble que le cluster *ots* provient d'une actinobactérie, le pourcentage en GC de ce cluster est proche de celui du génome de *S. thermophilus* LMD-9 et n'a donc pas pu être utilisé pour repérer ce cluster sur la séquence nucléotidique. En effet, le taux en GC des gènes du cluster *ots* est de 36, 39, 41 et 40% pour les gènes *otsA*, *otsB*, *otsC*, *otsD* contre 39,1% en moyenne pour celui du génome de la souche LMD-9.

Plusieurs hypothèses peuvent être formulées quant à la manière dont le cluster *ots* a été acquis par *S. thermophilus*. La première, la plus simple, implique une acquisition directe de ce cluster à partir d'une bactérie du phylum des *Actinobacteria* telle que *Bifidobacterium* ou *Brevibacterium*, qui peuvent partager le même habitat écologique que *S. thermophilus*. Dans ce cas là, il est attendu que le cluster *ots* ait été acquis depuis un temps suffisamment long pour permettre l'homogénéisation du pourcentage en GC de sa séquence.

Cette hypothèse implique, alors, que les deux copies d'ISSI qui délimitent l'élément TCp1 ont elles-mêmes été acquises à partir d'une actinobactérie. Des copies d'IS appartenant à 8 familles (IS3, IS21, IS110, IS150, IS256, IS607/IS200 et ISL3) ont aussi été détectées au sein des génomes de bifidobactéries. Par ailleurs, l'analyse de la région adjacente au cluster *ots* n'a pas permis de détecter des éléments mobiles spécifiques des bifidobactéries et aucune ISSI de type  $\alpha$  n'a été détectée dans les génomes séquencés de bifidobactéries. Le fait que les deux copies d'ISSI qui délimitent l'élément TCp1 soient caractéristiques de *Lc. lactis* suggère que le transfert du cluster *ots* d'une actinobactérie à *S. thermophilus* a impliqué un hôte intermédiaire tel que la bactérie lactique *Lc. Lactis* ou son ancêtre. Cette bactérie aurait pu ainsi acquérir le cluster *ots* dans un premier temps, et par la suite, la structure TCp1, *i.e.* le cluster *ots* flanqué par les deux copies d'ISSI, aurait pu être transférée à *S. thermophilus*. En outre, il faut souligner que le pourcentage en GC de *Lc. lactis* est proche de celui de *S. thermophilus*. Cependant, l'absence dans les banques des données d'une souche de *Lc. lactis* possédant un cluster homologue au cluster *ots* nous amène à dire que, soit la séquence génomique d'une telle souche n'est pas encore disponible dans les banques de données, soit cette structure a été perdue chez *Lc. lactis* au cours de

l'évolution. En effet, *Lc. lactis* ayant déjà deux systèmes de transport de peptides de type ABC (Opp et Opt), elle pourrait avoir perdu ce cluster car il n'apporterait pas d'avantage sélectif. Enfin, le cluster *ots* pourrait avoir été acquis par transfert horizontal à partir d'une bactérie dont le génome n'ait pas encore été étudié.

#### 6. 1. 4. Mécanismes de transfert horizontal impliqués dans l'acquisition de cluster *ots*

S'il semble clair que le cluster *ots* a été acquis par transfert horizontal et que les éléments transposables ont pu intervenir dans l'intégration du cluster dans le chromosome de *S. thermophilus* LMD-9, le vecteur ayant permis ce passage d'information génétique d'une bactérie à une autre reste totalement inconnu. De manière générale, plusieurs types de mécanismes permettent le transfert horizontal de gènes. Il s'agit de la conjugaison, de la transformation, de la transduction ou du nanotube.

Chez les bifidobactéries, des éléments pouvant participer aux mécanismes de transfert horizontal ont été trouvés au sein des 22 génomes disponibles dans les banques de données. Des plasmides portant des origines de transfert et des gènes codant des protéines impliquées dans la mobilisation ont été détectés. Cependant, à l'exception d'un plasmide présent chez *Bi. Bifidum* NCFB 1454 qui porte des gènes codant des protéines impliquées dans la production d'une bifidocine, tous ces plasmides sont cryptiques. Aucun ne comporte des gènes homologues à ceux du cluster *ots*.

Les ICEs sont des éléments génétiques codant les fonctions nécessaires à leur excision sous forme circulaire libre, à leur transfert par conjugaison d'une cellule donneuse à une cellule receveuse puis à leur intégration dans le génome de cette dernière (Burrus *et al.*, 2002a). L'analyse de la région contenant le cluster *ots* chez la souche LMD-9 a révélé la présence de pseudogènes présents chez d'autres ICEs de *S. thermophilus* (G. Guédon, communication personnelle). Par exemple, le pseudogène *ps3* (Figure 30) est très similaire à un gène présent sur ICE*St1* de *S. thermophilus*. De même, les gènes *ps6* et *ps7* sont très similaires de gènes présents dans des structures dérivant d'ICEs et intégrées au locus *rpsI* de *S. thermophilus* CNRZ1066, *i.e.*, locus où est situé le cluster *ots* dans la souche LMD-9. En outre, des gènes appartenant à des ICEs ont été détectés au niveau de ce locus chez les souches LMG18311, ND03 et JIM8232. Ainsi, cela pourrait signifier qu'un élément de type ICE ait été impliqué dans l'acquisition du cluster *ots* par la souche LMD-9. Acquisition, qui se serait donc faite par conjugaison.

Cependant, et sachant que des souches de *S. thermophilus* ont été qualifiées de « naturellement transformables » en milieu chimiquement défini (Gardan *et al.*, 2009), l'acquisition par transformation naturelle ne peut être écartée. Nos résultats montrent que toutes les souches ayant le cluster *ots* possèdent aussi les gènes *amiA1* et *amiA3* (Figure 36). Or, les travaux de Gardan *et al.* (2009) suggèrent que le système Ami, chez la souche LMD-9, joue un rôle clé dans la compétence naturelle. En effet, les deux protéines affines aux oligopeptides AmiA1 et AmiA3 contrôlent indirectement la compétence *via* le contrôle de l'expression du gène *comX*, lui-même impliqué dans le contrôle des gènes de compétence naturelle. Cela suggère donc que les souches *ots*<sup>+</sup>, comme la souche LMD-9, pourraient posséder les autres gènes nécessaires pour la compétence naturelle. Ces données nous permettent donc d'émettre l'hypothèse que le cluster *ots* pourrait avoir été acquis par transformation naturelle.

## **6. 2. Rôle du système Ots : utilité chez *Streptococcus thermophilus***

### **6. 2. 1. Rôle du système *ots* chez *Streptococcus thermophilus***

L'inactivation du cluster *ots* affecte clairement la croissance du mutant en milieu MCD. Dans ce milieu, la source azotée a majoritairement été apportée soit sous forme d'oligopeptides soit sous forme de di- et tri-peptides. Dans les conditions testées, la souche mutante présente un retard de croissance par rapport à la souche sauvage. Néanmoins, ce retard est plus marqué en présence de di- et tri-peptides qu'en présence d'oligopeptides. Cela pourrait résulter d'une capacité moindre de la souche LMD-9 $\Delta$ *ots* à transporter des di- tri-peptides. En effet, en présence d'oligopeptides, où la différence de croissance est moins marquée, le système Ami serait capable d'assurer une bonne croissance *via* l'internalisation des oligopeptides chez les deux souches. Quand le milieu contient majoritairement des di- et tri-peptides, la capacité du système Ami est limitée. Le rôle du système Ots serait alors plus important et son effet sur la croissance plus prononcé, et ce, malgré la présence du transporteur DtpT. En plus de son rôle dans la croissance, le système Ots pourrait exercer une autre fonction *via* le transport de peptides ayant un rôle dans le *quorum sensing*, et par conséquent, intervenir dans la régulation de différentes réponses telles que la compétence naturelle ou la conjugaison.

## 6. 2. 2. Acquisition des gènes conférant des avantages pour l'adaptation de *Streptococcus thermophilus* en lait

*S. thermophilus* est une espèce clonale ayant récemment émergé du groupe *salivarius*. L'émergence de cette espèce s'est accompagnée d'une évolution de son génome permettant d'exploiter au mieux sa nouvelle niche écologique, le lait. C'est ainsi que de nombreux gènes acquis par transfert horizontal et permettant l'adaptation de *S. thermophilus* au lait ont été détectés. Ces fonctions permettent, notamment, une meilleure utilisation des ressources nutritives disponibles dans le lait, dont la source azotée. Par exemple, la protéase PrtS, dont le gène a été acquis par transfert horizontal, permet la dégradation des caséines en peptides qui sont consommés par la bactérie au cours de sa croissance (Delorme *et al.*, 2010 ; Fernandez *et al.*, 2000). D'ailleurs, lorsque cette protéase est absente ou inactive, la croissance de *S. thermophilus* à haute densité cellulaire dans le lait est altérée. Par ailleurs, des gènes portés par un îlot acquis par transfert horizontal et intervenant dans la biosynthèse de méthionine pourraient jouer un rôle important dans la croissance bactérienne, d'autant que la méthionine est présente en faible quantité dans le lait.

Chez les streptocoques, aucun système de transport de peptides n'a à notre connaissance été décrit comme issu d'un événement de transfert horizontal. Ce n'est pas le cas chez d'autres bactéries lactiques. Ainsi, chez *Lc. lactis* SK11, l'opéron codant le système Opp a été retrouvé sur le plasmide pSK11L (Yu *et al.*, 1996). La présence d'IS982 et d'ISS1 respectivement localisée en amont et en aval de cet opéron pourrait être responsable de son instabilité et les auteurs ont proposé qu'un événement de transposition aboutisse chez certaines souches à l'intégration de cet opéron dans le chromosome ou à sa perte. Néanmoins, la présence de ce système confère un avantage sélectif à la bactérie qui le possède puisqu'il a été observé que la perte du plasmide pSK11L altère la croissance en lait de la souche mutante (Yu *et al.*, 1996).

L'étude du système de transport de peptides Opt chez *Lc. lactis* SK11 et Wg2 a permis aux auteurs de conclure que ce système pourrait en fait provenir d'un streptocoque. En effet, les gènes *opt* possèdent la même organisation que ceux présents chez les streptocoques. Ce système possède deux protéines affines aux oligopeptides, codées par les gènes *optA* et *optS*, or chez les streptocoques il y a souvent multiplicité des protéines OBP. L'arbre phylogénétique construit sur la base des séquences de plusieurs protéines OBP présentes chez plusieurs bactéries a montré que

les protéines OptA et OptS sont plus proches des protéines OppA présentes chez *S. uberis* et *S. mutans* que de celles présentes chez *Lc. lactis* et *Lb. bulgaricus* (Lamraque *et al.*, 2004).

Si l'étude de l'expression du cluster *ots* a permis de mettre en évidence une structure opéronique et une expression constitutive, son inactivation ne semble pas influencer de manière notable la croissance du mutant en milieu lait. Il est également à noter que la souche 4F44 dépourvue du cluster *ots* montre un fort pouvoir acidifiant en lait, reflet de sa croissance, comparable à celui de la souche LMD-9. Ainsi, soit effectivement le cluster *ots* n'apporte pas de réel avantage sélectif pour la croissance de la bactérie en lait car normalement celui-ci contient peu de di- et tri-peptides libres, soit l'inactivation du système *ots* est compensée par les autres systèmes de transport tels que les systèmes Ami et DtpT, soit son rôle majeur est l'internalisation de peptides de signalisation impliqués dans la perception de l'environnement.

### **6. 2. 3. Intégration du cluster *ots* dans le schéma de régulation de *Streptococcus thermophilus***

Généralement, les séquences acquises par transfert horizontal ont tendance à converger vers les propriétés compositionnelles du génome de l'organisme hôte. Ce phénomène est connu sous le nom de processus d'amélioration. En effet, la séquence exogène subit un processus d'accumulation de substitutions nucléotidiques permettant d'homogénéiser ses caractères avec celles de la souche hôte. Il est important de signaler que ces processus rendent plus difficile l'identification de la séquence exogène lors de l'analyse de la séquence du génome hôte. Néanmoins, cette approche a été employée pour estimer la date d'acquisition des gènes acquis par transfert horizontal (Lawrence et Ochman, 1997). Le contenu en GC du cluster *ots*, proche de celui de la bactérie réceptrice, est probablement le résultat des processus d'amélioration qui ont permis d'homogénéiser ces caractères avec ceux de *S. thermophilus*. Ceci dans le cas où le cluster *ots* aurait été acquis à partir d'une bactérie phylogénétiquement éloignée de *S. thermophilus*.

Outre, l'avantage sélectif rapide que doit conférer la séquence acquise pour son maintien dans le génome de la receveuse, elle doit aussi s'intégrer dans le schéma de régulation de la cellule. Si le rôle du cluster *ots* est d'intervenir dans l'approvisionnement azoté de la cellule, alors il pourrait appartenir au même régulon que d'autres gènes du système protéolytique de *S. thermophilus*. Chez *Lc. lactis*, il est connu que l'expression de certains gènes du système protéolytique sont

régulés par CodY. CodY est un régulateur transcriptionnel pléiotrope présent dans plusieurs espèces bactériennes à Gram positif dont les génomes à bas pourcentage en GC (Guédon *et al.*, 2001a, 2001b; den Hengst *et al.*, 2005a, 2005b). Il permet à la bactérie d'adapter son métabolisme à la quantité et aux types de nutriments présents dans son environnement (Sonenshein, 2005). La protéine CodY est capable de se fixer à l'ADN grâce à son motif hélice-tour-hélice (HTH), et de réguler les gènes cibles d'une manière directe. Le consensus établi chez *Lc. lactis* concernant sa séquence de fixation à l'ADN est AATTTTCWGAAAATT (Guédon *et al.*, 2005a ; den Hengst *et al.*, 2005b). La régulation du système protéolytique chez *S. thermophilus* a été très peu étudiée. L'analyse de la séquence génomique des souches de *S. thermophilus* disponible dans les banques de données indique la présence d'un homologue du gène *codY*. Une interaction entre la protéine CodY et les promoteurs des gènes *prtS*, *amiA1*, *pepA*, *pepS*, *pepN*, *pepC*, *pepB*, *pepQ*, *pepV*, *pepT*, *pepM*, *pepXP*, *pepP* et *pepO* a été démontrée *in vitro* (Liu *et al.*, 2009a). L'analyse de la séquence de la région promotrice du gène *otsA* chez la souche LMD-9 a révélé, tout comme pour les gènes *amiA1* et *amiA3*, la présence d'un éventuel motif de fixation de CodY en amont de ce gène (Figure 47). Cela suggère que le cluster *ots* pourrait appartenir au régulon CodY.

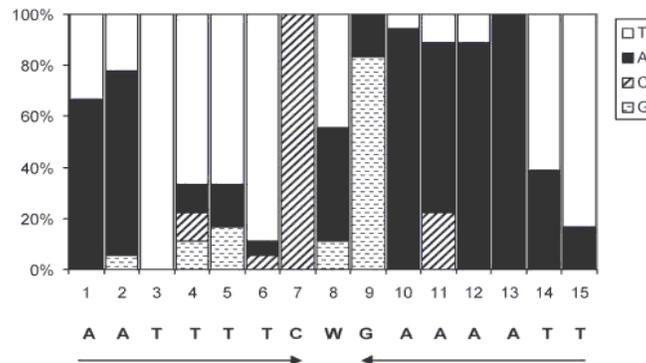
### 6. 3. Devenir du système Ots chez *Streptococcus thermophilus*

Pour que l'ADN acquis par transfert horizontal puisse se maintenir dans une cellule réceptrice, il doit conférer un avantage à la bactérie. Dans le cas où les gènes ne confèrent aucune utilité à la cellule bactérienne receveuse, ils vont accumuler des mutations ponctuelles conduisant à leur inactivation et donc à la formation de pseudogènes. Ainsi, les analyses du génome des souches LMG18311 et CNRZ1066 ont permis de détecter 226 gènes (équivalent 14% des gènes totaux) acquis potentiellement par transfert horizontal. Seuls 47 de ces 226, soit 21%, semblent être tronqués et/ou présumés inactifs (Hols *et al.*, 2005). Cela semble résulter de la contre-sélection de ces gènes car ils ne confèreraient pas d'avantage sélectif à la bactérie.

Les séquences du cluster *ots* codent des protéines fonctionnelles. Ce système est bien exprimé chez les souches étudiées PB302, PB180 et LMD-9 en milieu M17 et en lait. Enfin, la comparaison de la croissance de la souche sauvage LMD-9 avec celle de la souche mutante LMD-9 $\Delta$ *ots* en milieu riche en peptides a permis de lui attribuer un rôle dans l'internalisation de

di- et tri-peptides. L'ensemble de ces résultats suggère que le cluster *ots* serait utile pour la bactérie, bien qu'en lait nous n'ayons pas observé d'effet de son inactivation sur la croissance de la souche mutante. Ainsi, si le cluster *ots* n'a pas d'autre rôle que celui mis en évidence dans ce travail et qu'il n'apporte qu'un très faible avantage à *S. thermophilus*, il pourrait être inactivé par la suite.

A.



B.

Gène	Boîte CodY <sup>a</sup>	Distance/ATG <sup>b</sup>
<i>amiA1</i>	ATTATTCTGAAAAAT	-106
<i>amiA3</i>	AATTTTCTGATAAAA	-63
<i>otsA</i>	GGTTACAGATGATA	-75

**Figure 47 : Détection d'un éventuel motif de fixation de la protéine régulatrice CodY en amont des gènes *amiA1*, *amiA3* et *otsA* chez *Streptococcus thermophilus* LMD-9**

A. Séquence consensus permettant la fixation de la protéine régulatrice CodY mise en évidence par den Hengst *et al.* (2005b) chez *Lc. lactis*

B. Séquence et localisation d'un éventuel motif de fixation de la protéine CodY en amont des gènes *amiA1*, *amiA3* et *otsA* chez *S. thermophilus* LMD-9

Les substitutions sont mentionnées en rouge

<sup>a</sup> Les résidus qui diffèrent de la séquence consensus (AATTTTCWGAAAAT) sont en rouge (W = A ou T).

<sup>b</sup> Position relative en paires de bases par rapport à l'ATG du gène correspondant



## **V. Perspectives**



## Perspectives

Comme cela a été évoqué précédemment, un des objectifs de l'équipe PB2P est d'utiliser *S. thermophilus* pour la production de peptides bioactifs lors de la fabrication de produits laitiers tels que des laits fermentés, ce qui implique de connaître les capacités de production de peptides de cette bactérie à partir des caséines et son potentiel d'internalisation des peptides générés.

Les résultats obtenus au cours de cette étude ont montré que le système protéolytique de *S. thermophilus* était capable de générer des peptides, dont des peptides identifiés dans la littérature comme étant bioactif, à partir des caséines  $\alpha_{s1}$ ,  $\alpha_{s2}$  et  $\beta$  bovines. Afin de s'affranchir de l'intervention des systèmes de transport de peptides d'une part, et de la structure complexe (micelle de caséine) de la caséine d'autre part, notre démarche a consisté à placer les différentes souches de *S. thermophilus* en conditions de non croissance (incubation en tampon phosphate) en présence des caséines  $\alpha_{s1}$ ,  $\alpha_{s2}$  et  $\beta$  purifiées. Dans ces conditions, l'essentiel des peptides générés proviennent de l'activité des protéases/peptidases de surface. La variabilité de peptides générés par les souches a été interprétée comme le résultat d'une expression variable de la protéase PrtS. Il est vraisemblable que si cette expérience était répétée en incubant les souches de *S. thermophilus* en lait, la variabilité du nombre et du type de peptide généré différencierait de celle qui a été observée, et ce, pour deux raisons principales : la complexité de la caséine (substrat du système protéolytique) et l'intervention des différents systèmes de transport de peptides. Ainsi, même si les premiers résultats obtenus ont eu le mérite de montrer que *S. thermophilus* est capable de produire à partir des caséines de nombreux peptides dont des peptides bioactifs, il est maintenant indispensable de complexifier notre étude. Dans un premier temps, et afin de déterminer ce que le système protéolytique de *S. thermophilus* est capable de générer à partir des micelles de caséines, il sera nécessaire d'incuber les différentes souches protéolytiques que nous avons précédemment testées en présence de caséines micellaires, mais dans des conditions de non croissance des bactéries. Dans un deuxième temps, il sera nécessaire de déterminer le type de peptides présents dans le surnageant de culture de différents mutants de système de transport de peptides cultivés en lait. La comparaison des peptides présents avec ceux identifiés dans les conditions précédentes permettra de mieux appréhender le rôle des systèmes de transport de peptides.

Les résultats obtenus dans ce travail nous ont permis d'émettre l'hypothèse que le cluster *ots* pourrait participer à la nutrition azotée de *S. thermophilus* via l'internalisation des peptides et plus particulièrement ceux de petite taille. Plusieurs questions doivent maintenant être abordées : la spécificité de transport de ce système, la régulation de son expression et notamment le rôle du régulateur pléiotrope CodY dans cette régulation et le rôle du cluster *ots* dans des réponses bactériennes impliquant le *quorum sensing*. Répondre aux deux premières questions implique la fabrication préalable de différents mutants. Ainsi, la détermination de la spécificité de transport du cluster *ots* nécessite la fabrication de mutant du système Ami et du transporteur DtpT. L'étude de la régulation de l'expression du cluster nécessite, quant à elle, la construction d'un mutant *codY*. L'utilisation de la compétence naturelle pour la construction de mutants a déjà permis à l'équipe d'obtenir plusieurs mutants, dont 1 mutant *codY*. Ainsi, l'obtention de mutants du système Ami et du transporteur DtpT ne devrait pas *a priori* poser de problème particulier. Pour ce qui concerne l'étude de la spécificité, il s'agira dès lors d'étudier le type de peptides que les différents mutants sont capables d'internaliser comme cela a été réalisé pour le système Ami (Garault *et al.*, 2002). Pour ce qui concerne la régulation de l'expression du cluster *ots*, la comparaison de l'expression du cluster dans un contexte  $\Delta codY$  et sauvage devrait permettre de déterminer si le cluster *ots* appartient au régulon CodY. Enfin, l'étude de réponses dépendantes du *quorum sensing* telle que la compétence devront être étudiées chez le mutant LMD-9 $\Delta ots$ .

## **VI. Références bibliographiques**



- Abouhamad, W.N., Manson, M.D., 1994. The dipeptide permease of *Escherichia coli* closely resembles other bacterial transport systems and shows growth-phase-dependent expression. *Mol. Microbiol.* 14(5), 1077-1092.
- Accolas, J.P., Hemme, D., Desmazeaud, M.J., Vassal, L., Bouillanne, C., Veaux, M., 1980. Les levains lactiques thermophiles : Propriétés et comportement en technologie laitière. Une revue. *Le Lait.* 60(598), 487-524.
- Adams, M.R., 1999. Safety of industrial lactic acid bacteria. *J. Biotechnol.* 68(2-3), 171-178.
- Aguirre, M., Collins, M.D., 1993. Lactic acid bacteria and human clinical infection. *Journal of Applied Microbiology.* 75(2), 95-107.
- Aimutis, W.R., 2004. Bioactive Properties of Milk Proteins with Particular Focus on Anticariogenesis. *J. Nutr.* 134(4), 989S-995S.
- Alais, C., 1965. *Science du lait: principes des techniques laitières.* Editions Sep.
- Alloing, G., de Philip, P., Claverys, J.P., 1994. Three highly homologous membrane-bound lipoproteins participate in oligopeptide transport by the Ami system of the gram-positive *Streptococcus pneumoniae*. *J. Mol. Biol.* 241(1), 44-58.
- Alves, R.A., Payne, J.W., 1980. The number and nature of the peptide-transport systems of *Escherichia coli*: characterization of specific transport mutants. *Biochem. Soc. Trans.* 8(6), 704-705.
- Ammann, A., Neve, H., Geis, A., Heller, K.J., 2008. Plasmid Transfer Via Transduction from *Streptococcus Thermophilus* to *Lactococcus Lactis*. *J. Bacteriol.* 190(8), 3083-3087.
- Anastasiou, R., Papadelli, M., Georgalaki, M.D., Kalantzopoulos, G., Tsakalidou, E., 2002. Cloning and sequencing of the gene encoding X-prolyl-dipeptidyl aminopeptidase (PepX) from *Streptococcus thermophilus* strain ACA-DC 4. *J. Appl. Microbiol.* 93(1), 52-59.
- Arioli, S., Monnet, C., Guglielmetti, S., Parini, C., De Noni, I., Hogenboom, J., Halamí, P., Mora, D., 2007. Aspartate Biosynthesis is Essential for the Growth of *Streptococcus thermophilus* in Milk and its Availability Modulates the Level of Urease Activity. *Appl. Environ. Microbiol.* AEM.00533-07.
- Arioli, S., Roncada, P., Salzano, A.M., Deriu, F., Corona, S., Guglielmetti, S., Bonizzi, L., Scaloni, A., Mora, D., 2009. The relevance of carbon dioxide metabolism in *Streptococcus thermophilus*. *Microbiology.* 155(6), 1953 -1965.
- Barroga, C.F., Zhang, H., Wajih, N., Bouyer, J.H., Hermodson, M.A., 1996. The proteins encoded by the rbs operon of *Escherichia coli*: I. Overproduction, purification, characterization, and functional analysis of RbsA. *Protein Sci.* 5(6), 1093-1099.
- Bellanger, X., Roberts, A.P., Morel, C., Choulet, F., Pavlovic, G., Mullany, P., Decaris, B., Guédon, G., 2009. Conjugative Transfer of the Integrative Conjugative Elements ICES<sub>t1</sub> and ICES<sub>t3</sub> from *Streptococcus Thermophilus*. *J. Bacteriol.* 191(8), 2764-2775.
- Benini, S., Rypniewski, W.R., Wilson, K.S., Miletti, S., Ciurli, S., Mangani, S., 1999. A new proposal for urease mechanism based on the crystal structures of the native and inhibited enzyme from *Bacillus pasteurii*: why urea hydrolysis costs two nickels. *Structure.* 7(2), 205-216.
- Bennett, P.M., 2004. Genome plasticity: insertion sequence elements, transposons and integrons, and DNA rearrangement. *Methods Mol. Biol.* 266, 71-113.
- Biemans-Oldehinkel, E., Doeven, M.K., Poolman, B., 2006. ABC transporter architecture and regulatory roles of accessory domains. *FEBS Letters.* 580(4), 1023-1035.
- Biswas, I., Gruss, A., Ehrlich, S.D., Maguin, E., 1993. High-Efficiency Gene Inactivation and Replacement System for Gram-Positive Bacteria. *J. Bacteriol.* 175(11), 3628-3635.
- Van den Bogaard, P.T., Hols, P., Kuipers, O.P., Kleerebezem, M., de Vos, W.M., 2004. Sugar

- utilisation and conservation of the gal-lac gene cluster in *Streptococcus thermophilus*. *System.* 27, 10–17.
- Bolotin, A., Quinquis, B., Renault, P., Sorokin, A., Ehrlich, S.D., Kulakauskas, S., Lapidus, A., Goltsman, E., Mazur, M., Pusch, G.D., Fonstein, M., Overbeek, R., Kyprides, N., Purnelle, B., Prozzi, D., Ngui, K., Masuy, D., Hancy, F., Burteau, S., Boutry, M., Delcour, J., Goffeau, A., Hols, P., 2004. Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nat. Biotechnol.* 22(12), 1554-1558.
- Bonifait, L., de la Cruz Dominguez-Punaro, M., Vaillancourt, K., Bart, C., Slater, J., Frenette, M., Gottschalk, M., Grenier, D., 2010. The cell envelope subtilisin-like proteinase is a virulence determinant for *Streptococcus suis*. *BMC Microbiol.* 10(1), 42.
- Borezee, E., Pellegrini, E., Berche, P., 2000. OppA of *Listeria Monocytogenes*, an Oligopeptide-Binding Protein Required for Bacterial Growth at Low Temperature and Involved in Intracellular Survival. *Infect. Immun.* 68(12), 7069-7077.
- Bourgoin, F., Guédon, G., Pébay, M., Roussel, Y., Panis, C., Decaris, B., 1996. Characterization of a mosaic ISS1 element and evidence for the recent horizontal transfer of two different types of ISS1 between *Streptococcus thermophilus* and *Lactococcus lactis*. *Gene.* 178(1–2), 15-23.
- Bourgoin, F., Pluvinet, A., Gintz, B., Decaris, B., Guédon, G., 1999. Are horizontal transfers involved in the evolution of the *Streptococcus thermophilus* exopolysaccharide synthesis loci? *Gene.* 233(1-2), 151-161.
- Braun, V., Wu, H.C., 1994. Chapter 14 Lipoproteins, structure, function, biosynthesis and model for protein export, *Bacterial Cell Wall*. Elsevier, p. 319-341.
- Broadbent, J.R., McMahon, D.J., Welker, D.L., Oberg, C.J., Moineau, S., 2003. Biochemistry, Genetics, and Applications of Exopolysaccharide Production in *Streptococcus thermophilus*: A Review. *J. Dairy Sci.* 86(2), 407-423.
- Burne, R.A., Chen, Y.-Y.M., 2000. Bacterial ureases in infectious diseases. *Microbes and Infection.* 2(5), 533-542.
- Burrus, V., Pavlovic, G., Decaris, B., Guédon, G., 2002a. The ICES<sub>t1</sub> element of *Streptococcus thermophilus* belongs to a large family of integrative and conjugative elements that exchange modules and change their specificity of integration. *Plasmid.* 48(2), 77-97.
- Burrus, V., Pavlovic, G., Decaris, B., Guédon, G., 2002b. Conjugative transposons: the tip of the iceberg. *Mol. Microbiol.* 46(3), 601-610.
- Burrus, V., Roussel, Y., Decaris, B., Guedon, G., 2000. Characterization of a Novel Integrative Element, ICES<sub>t1</sub>, in the Lactic Acid Bacterium *Streptococcus thermophilus*. *Appl. Environ. Microbiol.* 66(4), 1749-1753.
- Chapot-Chartier, M.P., Rul, F., Nardi, M., Gripon, J.C., 1994. Gene cloning and characterization of PepC, a cysteine aminopeptidase from *Streptococcus thermophilus*, with sequence similarity to the eucaryotic bleomycin hydrolase. *Eur. J. Biochem.* 224(2), 497-506.
- Charbonnel, P., Lamarque, M., Piard, J.-C., Gilbert, C., Juillard, V., Atlan, D., 2003. Diversity of Oligopeptide Transport Specificity in *Lactococcus Lactis* Species A TOOL TO UNRAVEL THE ROLE OF OppA IN UPTAKE SPECIFICITY. *J. Biol. Chem.* 278(17), 14832-14840.
- Chavagnat, F., Casey, M.G., Meyer, J., 1999. Purification, Characterization, Gene Cloning, Sequencing, and Overexpression of Aminopeptidase N from *Streptococcus thermophilus* A. *Appl. Environ. Microbiol.* 65(7), 3001-3007.
- Chavagnat, F., Meyer, J., Casey, M.G., 2000. Purification, characterisation, cloning and

- sequencing of the gene encoding oligopeptidase PepO from *Streptococcus thermophilus* A. FEMS microbiology letters. 191(1), 79–85.
- Chen, M., Abele, R., Tampé, R., 2004. Functional non-equivalence of ATP-binding cassette signature motifs in the transporter associated with antigen processing (TAP). *J. Biol. Chem.* 279(44), 46073-46081.
- Chen, Y.Y., Burne, R.A., 1996. Analysis of *Streptococcus salivarius* urease expression using continuous chemostat culture. *FEMS Microbiol. Lett.* 135(2-3), 223-229.
- Chen, Y.Y., Clancy, K.A., Burne, R.A., 1996. *Streptococcus salivarius* urease: genetic and biochemical characterization and expression in a dental plaque streptococcus. *Infect. Immun.* 64(2), 585-592.
- Chen, Y.Y., Weaver, C.A., Burne, R.A., 2000. Dual functions of *Streptococcus salivarius* urease. *J. Bacteriol.* 182(16), 4667-4669.
- Chen, Y.-Y.M., Burne, R.A., 2003. Identification and characterization of the nickel uptake system for urease biogenesis in *Streptococcus salivarius* 57.I. *J. Bacteriol.* 185(23), 6773-6779.
- Cheng, Q., Campbell, E.A., Naughton, A.M., Johnson, S., Masure, H.R., 1997. The com locus controls genetic transformation in *Streptococcus pneumoniae*. *Mol. Microbiol.* 23(4), 683-692.
- Chiang, C.-S., Stacey, G., Tsay, Y.-F., 2004. Mechanisms and Functional Properties of Two Peptide Transporters, AtPTR2 and fPTR2. *Journal of Biological Chemistry.* 279(29), 30150 -30157.
- Chomczynski, P., 1993. A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *BioTechniques.* 15(3), 532-534, 536-537.
- Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162(1), 156-159.
- Claverys, J.-P., Prudhomme, M., Martin, B., 2006. Induction of competence regulons as a general response to stress in gram-positive bacteria. *Annu. Rev. Microbiol.* 60, 451-475.
- Claverys, J.P., Prudhomme, M., Mortier-Barrière, I., Martin, B., 2000. Adaptation to the environment: *Streptococcus pneumoniae*, a paradigm for recombination-mediated genetic plasticity? *Mol. Microbiol.* 35(2), 251-259.
- Clewell D.B., Flannagan S.E., Jaworski D.D., 1995. Unconstrained bacterial promiscuity: the Tn916-Tn1545 family of conjugative transposons. *Trends in Microbiology.* 3(6), 229-236.
- Cocaign - Bousquet, M., Garrigues, C., Novak, L., Lindley, N.D., Loublere, P., Cocaign - Bousquet, M., Garrigues, C., Novak, L., Lindley, N.D., Loublere, P., 1995. Rational development of a simple synthetic medium for the sustained growth of *Lactococcus lactis*, Rational development of a simple synthetic medium for the sustained growth of *Lactococcus lactis*. *Journal of Applied Microbiology, Journal of Applied Microbiology.* 79, 79(1, 1), 108, 108-116, 116.
- Collins, M.D., Samelis, J., Metaxopoulos, J., Wallbanks, S., 1993. Taxonomic studies on some leuconostoc-like organisms from fermented sausages: description of a new genus *Weissella* for the *Leuconostoc paramesenteroides* group of species. *J. Appl. Bacteriol.* 75(6), 595-603.
- Courtin, P., Monnet, V., Rul, F., 2002. Cell-wall proteinases PrtS and PrtB have a different role in *Streptococcus thermophilus*/*Lactobacillus bulgaricus* mixed cultures in milk. *Microbiology (Reading, Engl.)*. 148(Pt 11), 3413-3421.

- Daffonchio, D., Cherif, A., Brusetti, L., Rizzi, A., Mora, D., Boudabous, A., Borin, S., 2003. Nature of Polymorphisms in 16S-23S rRNA Gene Intergenic Transcribed Spacer Fingerprinting of *Bacillus* and Related Genera. *Appl. Environ. Microbiol.* 69(9), 5128-5137.
- Dandoy, D., Fremaux, C., Henry de Frahan, M., Horvath, P., Boyaval, P., Hols, P., Fontaine, L., 2011. The fast milk acidifying phenotype of *Streptococcus thermophilus* can be acquired by natural transformation of the genomic island encoding the cell-envelope proteinase PrtS. *Microb. Cell Fact.* 10 Suppl 1, S21.
- Dassa, E., Hofnung, M., 1985. Sequence of gene *malG* in *E. coli* K12: homologies between integral membrane components from binding protein-dependent transport systems. *EMBO J.* 4(9), 2287-2293.
- Decottignies, A., Goffeau, A., 1997. Complete inventory of the yeast ABC proteins. *Nat Genet.* 15(2), 137-145.
- Delcour, J., Ferain, T., Hols, P., 2000. Advances in the genetics of thermophilic lactic acid bacteria. *Curr. Opin. Biotechnol.* 11(5), 497-504.
- Delorme, C., Bartholini, C., Bolotine, A., Ehrlich, S.D., Renault, P., 2010. Emergence of a cell wall protease in the *Streptococcus thermophilus* population. *Appl. Environ. Microbiol.* 76(2), 451-460.
- Delorme, C., Bartholini, C., Luraschi, M., Pons, N., Loux, V., Almeida, M., Guédon, E., Gibrat, J.-F., Renault, P., 2011. Complete Genome Sequence of the Pigmented *Streptococcus thermophilus* Strain JIM8232. *Journal of Bacteriology.* 193(19), 5581 -5582.
- Delorme, C., Poyart, C., Ehrlich, S.D., Renault, P., 2007. Extent of Horizontal Gene Transfer in Evolution of *Streptococci* of the *Salivarius* Group. *J. Bacteriol.* 189(4), 1330-1341.
- Van Der Heide, T., Poolman, B., 2000. Osmoregulated ABC-transport system of *Lactococcus lactis* senses water stress via changes in the physical state of the membrane. *Proceedings of the National Academy of Sciences of the United States of America.* 97(13), 7102-7106.
- Desmazeaud, M., 1983. L'état des connaissances en matière de nutrition des bactéries lactiques Current state of research on lactic acid bacteria nutrition. *Le Lait.* 63(629-630), 50.
- Desmazeaud, M.J., 1990. Le lait milieu de culture. *M.A.N. Microbiologie, aliments, nutrition.* 8(4), 313-325.
- Detmers, F.J., Kunji, E.R., Lanfermeijer, F.C., Poolman, B., Konings, W.N., 1998. Kinetics and specificity of peptide uptake by the oligopeptide transport system of *Lactococcus lactis*. *Biochemistry.* 37(47), 16671-16679.
- Detmers, F.J., Lanfermeijer, F.C., Abele, R., Jack, R.W., Tampe, R., Konings, W.N., Poolman, B., 2000. Combinatorial peptide libraries reveal the ligand-binding mechanism of the oligopeptide receptor OppA of *Lactococcus lactis*. *Proc. Natl. Acad. Sci. U.S.A.* 97(23), 12487-12492.
- Deveau, H., Barrangou, R., Garneau, J.E., Labonte, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., Moineau, S., 2008. Phage Response to CRISPR-Encoded Resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190(4), 1390-1400.
- Dixon, N.E., Gazzola, T.C., blakeley, R.L., Zermer, B., 1975. Letter: Jack bean urease (EC 3.5.1.5). A metalloenzyme. A simple biological role for nickel? *J. Am. Chem. Soc.* 97(14), 4131-4133.
- Doeven, M.K., Abele, R., Tampé, R., Poolman, B., 2004. The Binding Specificity of OppA Determines the Selectivity of the Oligopeptide ATP-binding Cassette Transporter. *Journal of Biological Chemistry.* 279(31), 32301 -32307.

- Doeven, M.K., van den Bogaart, G., Krasnikov, V., Poolman, B., 2008. Probing receptor-translocator interactions in the oligopeptide ABC transporter by fluorescence correlation spectroscopy. *Biophys. J.* 94(10), 3956-3965.
- Doeven, M.K., Kok, J., Poolman, B., Doeven, M.K., Kok, J., Poolman, B., 2005. Specificity and selectivity determinants of peptide transport in *Lactococcus lactis* and other microorganisms, Specificity and selectivity determinants of peptide transport in *Lactococcus lactis* and other microorganisms. *Molecular Microbiology, Molecular Microbiology.* 57, 57(3, 3), 640, 640-649, 649.
- Dower, W.J., Miller, J.F., Ragsdale, C.W., 1988. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.* 16(13), 6127-6145.
- Doyuk, E., Ormerod, O.J., Bowler, I.C.J.W., 2002. Native valve endocarditis due to *Streptococcus vestibularis* and *Streptococcus oralis*. *J. Infect.* 45(1), 39-41.
- Driessen, A.J.M., Hellingwerf, K.J., Konings, W.N., 1987. Mechanism of energy coupling to entry and exit of neutral and branched chain amino acids in membrane vesicles of *Streptococcus cremoris*. *J. Biol. Chem.* 12438-12443.
- Dubey, G.P., Ben-Yehuda, S., 2011. Intercellular Nanotubes Mediate Bacterial Communication. *Cell.* 144(4), 590-600.
- Dunny, G.M., Leonard, B.A., 1997. Cell-cell communication in gram-positive bacteria. *Annu. Rev. Microbiol.* 51, 527-564.
- Dunten, P., Mowbray, S.L., 1995. Crystal structure of the dipeptide binding protein from *Escherichia coli* involved in active transport and chemotaxis. *Protein Sci.* 4(11), 2327-2334.
- Fang, G., Konings, W.N., Poolman, B., 2000. Kinetics and Substrate Specificity of Membrane-Reconstituted Peptide Transporter DtpT of *Lactococcus lactis*. *J. Bacteriol.* 182(9), 2530-2535.
- Farrow, J.A.E., Collins, M.D., 1984. DNA Base Composition, DNA-DNA Homology and Long-chain Fatty Acid Studies on *Streptococcus thermophilus* and *Streptococcus salivarius*. *Journal of General Microbiology.* 130(2), 357 -362.
- Fernandez-Esplá, M.D., Garault, P., Monnet, V., Rul, F., 2000. *Streptococcus thermophilus* Cell Wall-Anchored Proteinase: Release, Purification, and Biochemical and Genetic Characterization. *Appl. Environ. Microbiol.* 66(11), 4772-4778.
- Fernandez-Esplá, M.D., Rul, F., 1999. PepS from *Streptococcus thermophilus*. A new member of the aminopeptidase T family of thermophilic bacteria. *Eur. J. Biochem.* 263(2), 502-510.
- Fontaine, L., Boutry, C., de Frahan, M.H., Delplace, B., Fremaux, C., Horvath, P., Boyaval, P., Hols, P., 2010. A Novel Pheromone Quorum-Sensing System Controls the Development of Natural Competence in *Streptococcus thermophilus* and *Streptococcus salivarius*. *J. Bacteriol.* 192(5), 1444-1454.
- Foucaud, C., Kunji, E., Hagting, A., Richard, J., Konings, W., Desmazeaud, M., Poolman, B., 1995. Specificity of peptide transport systems in *Lactococcus lactis*: evidence for a third system which transports hydrophobic di- and tripeptides. *J. Bacteriol.* 177(16), 4652-4657.
- Fox, P.F., 1993. *Cheese: Chemistry, Physics and Microbiology.* Nature Publishing Group.
- Frommer, W.B., Hummel, S., Rentsch, D., 1994. Cloning of an *Arabidopsis* histidine transporting protein related to nitrate and peptide transporters. *FEBS Letters.* 347(2-3), 185-189.
- Galia, W., Perrin, C., Genay, M., Dary, A., 2009. Variability and molecular typing of *Streptococcus thermophilus* strains displaying different proteolytic and acidifying

- properties. *International Dairy Journal*. 19(2), 89-95.
- Garault, P., Le Bars, D., Besset, C., Monnet, V., 2002. Three Oligopeptide-binding Proteins Are Involved in the Oligopeptide Transport of *Streptococcus thermophilus*. *Journal of Biological Chemistry*. 277(1), 32-39.
- Garault, P., Letort, C., Juillard, V., Monnet, V., 2000. Branched-chain amino acid biosynthesis is essential for optimal growth of *Streptococcus thermophilus* in milk. *Appl. Environ. Microbiol.* 66(12), 5128-5133.
- Garcia-Vallve, S., Guzman, E., Montero, M.A., Romeu, A., 2003. HGT-DB: A Database of Putative Horizontally Transferred Genes in Prokaryotic Complete Genomes. *Nucl. Acids Res.* 31(1), 187-189.
- Gardan, R., Besset, C., Guillot, A., Gitton, C., Monnet, V., 2009. The oligopeptide transport system is essential for the development of natural competence in *Streptococcus thermophilus* strain LMD-9. *J. Bacteriol.* 191(14), 4647-4655.
- Gilson, E., Alloing, G., Schmidt, T., Claverys, J.P., Dudler, R., Hofnung, M., 1988. Evidence for high affinity binding-protein dependent transport systems in gram-positive bacteria and in *Mycoplasma*. *EMBO J.* 7(12), 3971-3974.
- Guédon, E., Renault, P., Ehrlich, S.D., Delorme, C., 2001a. Transcriptional Pattern of Genes Coding for the Proteolytic System of *Lactococcus lactis* and Evidence for Coordinated Regulation of Key Enzymes by Peptide Supply. *J. Bacteriol.* 183(12), 3614-3622.
- Guédon, E., Serror, P., Ehrlich, S.D., Renault, P., Delorme, C., 2001b. Pleiotropic transcriptional repressor CodY senses the intracellular pool of branched-chain amino acids in *Lactococcus lactis*. *Mol. Microbiol.* 40(5), 1227-1239.
- Guédon, G., Bourgoin, F., Pébay, M., Roussel, Y., Colmin, C., Simonet, J.M., Decaris, B., 1995. Characterization and distribution of two insertion sequences, *IS1191* and *iso-IS981*, in *Streptococcus thermophilus*: does intergeneric transfer of insertion sequences occur in lactic acid bacteria co-cultures? *Molecular Microbiology*. 16(1), 69-78.
- Guyer, C.A., Morgan, D.G., Staros, J.V., 1986. Binding specificity of the periplasmic oligopeptide-binding protein from *Escherichia coli*. *Journal of Bacteriology*. 168(2), 775 - 779.
- Hagting, A., Kunji, E.R., Leenhouts, K.J., Poolman, B., Konings, W.N., 1994. The di- and tripeptide transport protein of *Lactococcus lactis*. A new type of bacterial peptide transporter. *Journal of Biological Chemistry*. 269(15), 11391-11399.
- Hauser, M., Narita, V., Donhardt, A.M., Naider, F., Becker, J.M., 2001. Multiplicity and regulation of genes encoding peptide transporters in *Saccharomyces cerevisiae*. *Mol. Membr. Biol.* 18(1), 105-112.
- Helinck, S., Charbonnel, P., Foucaud-Scheunemann, C., Piard, J.-C., Juillard, V., 2003. Charged casein-derived oligopeptides competitively inhibit the transport of a reporter oligopeptide by *Lactococcus lactis*. *J. Appl. Microbiol.* 94(5), 900-907.
- Helmann, J.D., 1995. Compilation and analysis of *Bacillus subtilis* sigma A-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. *Nucleic Acids Res.* 23(13), 2351-2360.
- den Hengst, C.D., Curley, P., Larsen, R., Buist, G., Nauta, A., van Sinderen, D., Kuipers, O.P., Kok, J., 2005a. Probing Direct Interactions between CodY and the *oppD* Promoter of *Lactococcus lactis*. *J. Bacteriol.* 187(2), 512-521.
- den Hengst, C.D., van Hijum, S.A.F.T., Geurts, J.M.W., Nauta, A., Kok, J., Kuipers, O.P., 2005b. The *Lactococcus lactis* CodY Regulon: IDENTIFICATION OF A CONSERVED cis-REGULATORY ELEMENT. *Journal of Biological Chemistry*. 280(40), 34332-34342.

- Hertel, C., Schmidt, G., Fischer, M., Oellers, K., Hammes, W.P., 1998. Oxygen-Dependent Regulation of the Expression of the Catalase Gene *katA* of *Lactobacillus sakei* LTH677. *Appl. Environ. Microbiol.* 64(4), 1359-1365.
- Herve-Jimenez, L., Guillouard, I., Guedon, E., Gautier, C., Boudebbouze, S., Hols, P., Monnet, V., Rul, F., Maguin, E., 2008. Physiology of *Streptococcus thermophilus* during the late stage of milk fermentation with special regard to sulfur amino-acid metabolism. *Proteomics.* 8(20), 4273-4286.
- Higgins, C.F., 1992. ABC Transporters: From Microorganisms to Man. *Annu. Rev. Cell. Biol.* 8(1), 67-113.
- Hols, P., Hancy, F., Fontaine, L., Grossiord, B., Prozzi, D., Leblond-Bourget, N., Decaris, B., Bolotin, A., Delorme, C., Dusko Ehrlich, S., Guédon, E., Monnet, V., Renault, P., Kleerebezem, M., 2005. New insights in the molecular biology and physiology of *Streptococcus thermophilus* revealed by comparative genomics. *FEMS Microbiol. Rev.* 29(3), 435-463.
- Holzappel, W.H., Haberer, P., Geisen, R., Björkroth, J., Schillinger, U., 2001. Taxonomy and important features of probiotic microorganisms in food and nutrition. *Am. J. Clin. Nutr.* 73(2 Suppl), 365S-373S.
- Horvath, P., Romero, D.A., Coute-Monvoisin, A.-C., Richards, M., Deveau, H., Moineau, S., Boyaval, P., Fremaux, C., Barrangou, R., 2008. Diversity, Activity, and Evolution of CRISPR Loci in *Streptococcus thermophilus*. *J. Bacteriol.* 190(4), 1401-1412.
- Hu, Q., Liu, P., Yu, Z., Zhao, G., Li, J., Teng, L., Zhou, M., Bei, W., Chen, H., Jin, M., 2010. Identification of a cell wall-associated subtilisin-like serine protease involved in the pathogenesis of *Streptococcus suis* serotype 2. *Microb. Pathog.* 48(3-4), 103-109.
- Hung, L.W., Wang, I.X., Nikaido, K., Liu, P.Q., Ames, G.F., Kim, S.H., 1998. Crystal structure of the ATP-binding subunit of an ABC transporter. *Nature.* 396(6712), 703-707.
- Hyde, S.C., Emsley, P., Hartshorn, M.J., Mimmack, M.M., Gileadi, U., Pearce, S.R., Gallagher, M.P., Gill, D.R., Hubbard, R.E., Higgins, C.F., 1990. Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport. *Nature.* 346(6282), 362-365.
- Idigoras, P., Valiente, A., Iglesias, L., Trieu-Cout, P., Poyart, C., 2001. Meningitis due to *Streptococcus salivarius*. *J. Clin. Microbiol.* 39(8), 3017.
- Jenkinson, H., Baker, R., Tannock, G., 1996. A binding-lipoprotein-dependent oligopeptide transport system in *Streptococcus gordonii* essential for uptake of hexa- and heptapeptides. *J. Bacteriol.* 178(1), 68-77.
- Johnson, J.E., Chiu, W., 2007. DNA packaging and delivery machines in tailed bacteriophages. *Curr. Opin. Struct. Biol.* 17(2), 237-243.
- Jones, P.M., George, A.M., 1999. Subunit interactions in ABC transporters: towards a functional architecture. *FEMS Microbiology Letters.* 179(2), 187-202.
- Juillard, V., Le Bars, D., Kunji, E., Konings, W., Gripon, J., Richard, J., 1995. Oligopeptides are the main source of nitrogen for *Lactococcus lactis* during growth in milk. *Appl. Environ. Microbiol.* 61(8), 3024-3030.
- Juillard, V., Desmazeaud, M.J., Spinnler, H.E., 1988. Mise en évidence d'une activité uréasique chez *Streptococcus thermophilus*. *Canadian Journal of Microbiology.* 34, 818-822.
- Juillard, V., Foucaud, C., Desmazeaud, M., Richard, J., 1996. Utilization of nitrogen sources during growth of *Lactococcus lactis* in milk. Utilisation des sources d'azote du lait par *Lactococcus lactis*. *Le Lait.* 76(1-2), 12.
- Juillard, V., Guillot, A., Le Bars, D., Gripon, J.-C., 1998. Specificity of Milk Peptide Utilization

- by *Lactococcus lactis*. *Appl. Environ. Microbiol.* 64(4), 1230-1236.
- Juille, O., Bars, D.L., Juillard, V., 2005. The specificity of oligopeptide transport by *Streptococcus thermophilus* resembles that of *Lactococcus lactis* and not that of pathogenic streptococci. *Microbiology.* 151(6), 1987-1994.
- Juillard, V., 2007. Transport des acides aminés et des peptides, *Livre de bactéries lactiques.*
- Kakinuma, Y., 1998. Inorganic Cation Transport and Energy Transduction in *Enterococcus hirae* and Other Streptococci. *Microbiol. Mol. Biol. Rev.* 62(4), 1021-1045.
- Karpowich, N., Martsinkevich, O., Millen, L., Yuan, Y.R., Dai, P.L., MacVey, K., Thomas, P.J., Hunt, J.F., 2001. Crystal structures of the MJ1267 ATP binding cassette reveal an induced-fit effect at the ATPase active site of an ABC transporter. *Structure.* 9(7), 571-586.
- Kawamura, Y., Hou, X.G., Sultana, F., Miura, H., Ezaki, T., 1995. Determination of 16S rRNA sequences of *Streptococcus mitis* and *Streptococcus gordonii* and phylogenetic relationships among members of the genus *Streptococcus*. *Int. J. Syst. Bacteriol.* 45(2), 406-408.
- Kerppola, R.E., Shyamala, V.K., Klebba, P., Ames, G.F., 1991. The membrane-bound proteins of periplasmic permeases form a complex. Identification of the histidine permease HisQMP complex. *J. Biol. Chem.* 266(15), 9857-9865.
- Kerr, A.R., Adrian, P.V., Estevão, S., de Groot, R., Alloing, G., Claverys, J.-P., Mitchell, T.J., Hermans, P.W.M., 2004. The Ami-AliA/AliB permease of *Streptococcus pneumoniae* is involved in nasopharyngeal colonization but not in invasive disease. *Infect. Immun.* 72(7), 3902-3906.
- Koide, A., Hoch, J., 1994. Identification of a second oligopeptide transport system in *Bacillus subtilis* and determination of its role in sporulation. *Mol. Mic.* 13, 417-426.
- Konings, W.N., Poolman, B., Driessen, A.J.M., Maloney, P.C., 1989. Bioenergetics and Solute Transport in Lactococci. *Critical Reviews in Microbiology.* 16(6), 419-476.
- Kung, V.L., Ozer, E.A., Hauser, A.R., 2010. The Accessory Genome of *Pseudomonas Aeruginosa*. *Microbiol. Mol. Biol. Rev.* 74(4), 621-641.
- Kunji, E.R., Mierau, I., Hagting, A., Poolman, B., Konings, W.N., 1996. The proteolytic systems of lactic acid bacteria. *Antonie Van Leeuwenhoek.* 70(2-4), 187-221.
- Kunji, E.R., Smid, E.J., Plapp, R., Poolman, B., Konings, W.N., 1993. Di-tripeptides and oligopeptides are taken up via distinct transport mechanisms in *Lactococcus lactis*. *J. Bacteriol.* 175(7), 2052-2059.
- Kunji, E.R.S., Hagting, A., De Vries, C.J., Juillard, V., Haandrikman, A.J., Poolman, B., Konings, W.N., 1995. Transport of -Casein-derived Peptides by the Oligopeptide Transport System Is a Crucial Step in the Proteolytic Pathway of *Lactococcus lactis*. *Journal of Biological Chemistry.* 270(4), 1569-1574.
- Kusuda, R., Kawai, K., Salati, F., Banner, C.R., Fryer, J.L., 1991. *Enterococcus seriolicida* sp. nov., a Fish Pathogen. *International Journal of Systematic Bacteriology.* 41(3), 406-409.
- Lamarque, M., Aubel, D., Piard, J.-C., Gilbert, C., Juillard, V., Atlan, D., 2011. The peptide transport system Opt is involved in both nutrition and environmental sensing during growth of *Lactococcus lactis* in milk. *Microbiology.* 157(6), 1612-1619.
- Lamarque, M., Charbonnel, P., Aubel, D., Piard, J.-C., Atlan, D., Juillard, V., 2004. A Multifunction ABC Transporter (Opt) Contributes to Diversity of Peptide Uptake Specificity within the Genus *Lactococcus*. *J. Bacteriol.* 186(19), 6492-6500.
- Lanfermeijer, F.C., Detmers, F.J., Konings, W.N., Poolman, B., 2000. On the binding mechanism of the peptide receptor of the oligopeptide transport system of *Lactococcus lactis*. *EMBO*

- J. 19(14), 3649-3656.
- Lanfermeijer, F.C., Picon, A., Konings, W.N., Poolman, B., 1999. Kinetics and consequences of binding of nona- and dodecapeptides to the oligopeptide binding protein (OppA) of *Lactococcus lactis*. *Biochemistry*. 38(44), 14440-14450.
- Lawrence, J.G., Ochman, H., 1997. Amelioration of bacterial genomes: rates of change and exchange. *J. Mol. Evol.* 44(4), 383-397.
- Lazazzera, B.A., Kurtser, I.G., McQuade, R.S., Grossman, A.D., 1999. An Autoregulatory Circuit Affecting Peptide Signaling in *Bacillus subtilis*. *J. Bacteriol.* 181(17), 5193-5200.
- Lazazzera, B.A., Solomon, J.M., Grossman, A.D., 1997. An Exported Peptide Functions Intracellularly to Contribute to Cell Density Signaling in *B. subtilis*. *Cell*. 89(6), 917-925.
- Ledesma, O.V., Holgado, D.R., P, A., Oliver, G., Giori, G.S.D., Raibaud, P., Galpin, J.V., Ledesma, O.V., De Ruiz Holgado, A.P., Oliver, G., Giori, G.S.D., Raibaud, P., Galpin, J.V., 1977. A Synthetic Medium for Comparative Nutritional Studies of Lactobacilli, A Synthetic Medium for Comparative Nutritional Studies of Lactobacilli. *Journal of Applied Microbiology, Journal of Applied Microbiology*. 42, 42(1, 1), 123, 123-133, 133.
- Leenhouts, K., 1995. Integration strategies and vectors. *Dev. Biol. Stand.* 85, 523-530.
- Lefébure, T., Stanhope, M.J., 2007. Evolution of the core and pan-genome of *Streptococcus*: positive selection, recombination, and genome composition. *Genome Biol.* 8(5), R71-R71.
- Leonard, B.A., Podbielski, A., Hedberg, P.J., Dunny, G.M., 1996. *Enterococcus faecalis* pheromone binding protein, PrgZ, recruits a chromosomal oligopeptide permease system to import sex pheromone cCF10 for induction of conjugation. *Proceedings of the National Academy of Sciences*. 93(1), 260 -264.
- Letort, C., Juillard, V., 2001. Development of a minimal chemically-defined medium for the exponential growth of *Streptococcus thermophilus*. *J. Appl. Microbiol.* 91(6), 1023-1029.
- Levdikov, V.M., Blagova, E.V., Brannigan, J.A., Wright, L., Vagin, A.A., Wilkinson, A.J., 2005. The structure of the oligopeptide-binding protein, AppA, from *Bacillus subtilis* in complex with a nonapeptide. *J. Mol. Biol.* 345(4), 879-892.
- Liu, F., Du, L., Du, P., Huo, G., 2009a. Possible promoter regions within the proteolytic system in *Streptococcus thermophilus* and their interaction with the CodY homolog. *FEMS Microbiol. Lett.* 297(2), 164-172.
- Liu, M., Siezen, R.J., Nauta, A., 2009b. In silico prediction of horizontal gene transfer events in *Lactobacillus bulgaricus* and *Streptococcus thermophilus* reveals proto-cooperation in yogurt manufacturing. *Appl. Environ. Microbiol.* 75(12), 4120-4129.
- Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*. 25(4), 402-408.
- Lorenz, M.G., Wackernagel, W., 1994. Bacterial gene transfer by natural genetic transformation in the environment. *Microbiol. Rev.* 58(3), 563-602.
- Lun, Z.-R., Wang, Q.-P., Chen, X.-G., Li, A.-X., Zhu, X.-Q., 2007. *Streptococcus suis*: an emerging zoonotic pathogen. *The Lancet Infectious Diseases*. 7(3), 201-209.
- Maguin, E., Duwat, P., Hege, T., Ehrlich, D., Gruss, A., 1992. New thermosensitive plasmid for gram-positive bacteria. *J Bacteriol.* 174(17), 5633-5638.
- Maguin, E., Prévost, H., Ehrlich, S.D., Gruss, A., 1996. Efficient Insertional Mutagenesis in Lactococci and Other Gram-Positive Bacteria. *J. Bacteriol.* 178(3), 931-935.
- Mahillon, J., Chandler, M., 1998. Insertion sequences. *Microbiol. Mol. Biol. Rev.* 62(3), 725-774.
- Mahillon, J., Léonard, C., Chandler, M., 1999. IS elements as constituents of bacterial genomes.

- Research in Microbiology. 150(9–10), 675-687.
- Makarova, K., Slesarev, A., Wolf, Y., Sorokin, A., Mirkin, B., Koonin, E., Pavlov, A., Pavlova, N., Karamychev, V., Polouchine, N., Shakhova, V., Grigoriev, I., Lou, Y., Rohksar, D., Lucas, S., Huang, K., Goodstein, D.M., Hawkins, T., Plengvidhya, V., Welker, D., Hughes, J., Goh, Y., Benson, A., Baldwin, K., Lee, J.-H., Díaz-Muñiz, I., Dosti, B., Smeianov, V., Wechter, W., Barabote, R., Lorca, G., Altermann, E., Barrangou, R., Ganesan, B., Xie, Y., Rawsthorne, H., Tamir, D., Parker, C., Breidt, F., Broadbent, J., Hutkins, R., O'Sullivan, D., Steele, J., Unlu, G., Saier, M., Klaenhammer, T., Richardson, P., Kozyavkin, S., Weimer, B., Mills, D., 2006. Comparative genomics of the lactic acid bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 103(42), 15611-15616.
- Mares, A., Neyts, K., Debevere, J., 1994. Influence of pH, salt and nitrite on the heme-dependent catalase activity of lactic acid bacteria. *International journal of food microbiology.* 24(1-2), 191–198.
- Martin, B., Coulon, J.B., Chamba, J.F., Bugaud, C., 1997. Effect of milk urea content on characteristics of matured Reblochon cheeses. *Le Lait.* 77(4), 505-514.
- Mathot, A.G., Beliard, E., Thuault, D., 2003. *Streptococcus thermophilus* 580 Produces a Bacteriocin Potentially Suitable for Inhibition of *Clostridium tyrobutyricum* in Hard Cheese. *Journal of Dairy Science.* 86(10), 3068-3074.
- Miclo, L., Roux, E., Genay, M., Brusseau, E., Poirson, C., Jameh, N., Perrin, C., Dary, A., 2012. Variability of hydrolysis of  $\beta$ -,  $\alpha$ 1-, and  $\alpha$ 2-caseins by 10 strains of *Streptococcus thermophilus* and resulting bioactive peptides. *J. Agric. Food Chem.* 60(2), 554-565.
- Mobley, H.L., Hausinger, R.P., 1989. Microbial ureases: significance, regulation, and molecular characterization. *Microbiol. Rev.* 53(1), 85-108.
- Mobley, H.L.T., Island, M.D., Hausinger, R.P., 1995. Molecular biology of microbial ureases. *Microbiological Reviews.* 59(3), 451-480.
- Monnet, C., Mora, D., Corrieu, G., 2005. Glutamine Synthesis Is Essential for Growth of *Streptococcus thermophilus* in Milk and Is Linked to Urea Catabolism. *Appl. Environ. Microbiol.* 71(6), 3376-3378.
- Monnet, C., Nardi, M., Hols, P., Gulea, M., Corrieu, G., Monnet, V., 2003. Regulation of branched-chain amino acid biosynthesis by  $\alpha$ -acetolactate decarboxylase in *Streptococcus thermophilus*. *Letters in Applied Microbiology.* 36(6), 399-405.
- Monnet, C., Pernoud, S., Sepulchre, A., Fremaux, C., Corrieu, G., 2004. Selection and properties of *Streptococcus thermophilus* mutants deficient in urease. *J. Dairy Sci.* 87(6), 1634-1640.
- Monnet, V., 2006. *Streptococcus thermophilus*, un streptocoque atypique et sympathique. *Bull Soc Fr Microbiol.* 23-28.
- Mora, D., Fortina, M.G., Parini, C., Ricci, G., Gatti, M., Giraffa, G., Manachini, P.L., 2002. Genetic diversity and technological properties of *Streptococcus thermophilus* strains isolated from dairy products. *Journal of Applied Microbiology.* 93(2), 278-287.
- Mora, D., Maguin, E., Masiero, M., Parini, C., Ricci, G., Manachini, P.L., Daffonchio, D., 2004. Characterization of urease genes cluster of *Streptococcus thermophilus*. *Journal of Applied Microbiology.* 96(1), 209-219.
- Mora, D., Monnet, C., Daffonchio, D., 2005a. Balancing the Loss and Acquisition of Pathogenic Traits in Food-Associated Bacteria. *Microbiology.* 151(12), 3814-3816.
- Mora, D., Monnet, C., Parini, C., Guglielmetti, S., Mariani, A., Pintus, P., Molinari, F., Daffonchio, D., Manachini, P.L., 2005b. Urease biogenesis in *Streptococcus thermophilus*. *Research in Microbiology.* 156(9), 897-903.
- Mora, D., Ricci, G., Guglielmetti, S., Daffonchio, D., Fortina, M.G., 2003. 16S-23S rRNA

- intergenic spacer region sequence variation in *Streptococcus thermophilus* and related dairy streptococci and development of a multiplex ITS-SSCP analysis for their identification. *Microbiology (Reading, Engl.)*. 149(Pt 3), 807-813.
- Morice, M., Bracquart, P., Linden, G., 1992. Colonial Variation and Freeze-Thaw Resistance of *Streptococcus thermophilus*. *Journal of Dairy Science*. 75(5), 1197-1203.
- Morishita, T., Deguchi, Y., Yajima, M., Sakurai, T., Yura, T., 1981. Multiple nutritional requirements of lactobacilli: genetic lesions affecting amino acid biosynthetic pathways. *J. Bacteriol.* 148(1), 64-71.
- Morishita, T., Fukada, T., Shiota, M., Yura, T., 1974. Genetic Basis of Nutritional Requirements in *Lactobacillus Casei*. *J. Bacteriol.* 120(3), 1078-1084.
- Morrison, D.A., Jaurin, B., 1990. *Streptococcus pneumoniae* possesses canonical *Escherichia coli* ( $\sigma$  70) promoters. *Mol. Microbiol.* 4(7), 1143-1152.
- Mourez, M., Hofnung, M., Dassa, E., 1997. Subunit interactions in ABC transporters: a conserved sequence in hydrophobic membrane proteins of periplasmic permeases defines an important site of interaction with the ATPase subunits. *EMBO J.* 16(11), 3066-3077.
- Mulrooney, S.B., Hausinger, R.P., 2003. Nickel uptake and utilization by microorganisms. *FEMS Microbiol. Rev.* 27(2-3), 239-261.
- Muto, A., Osawa, S., 1987. The Guanine and Cytosine Content of Genomic DNA and Bacterial Evolution. *PNAS.* 84(1), 166-169.
- Nickitenko, A.V., Trakhanov, S., Quijcho, F.A., 1995. 2 A resolution structure of DppA, a periplasmic dipeptide transport/chemosensory receptor. *Biochemistry.* 34(51), 16585-16595.
- Obis, D., Guillot, A., Gripon, J.C., Renault, P., Bolotin, A., Mistou, M.Y., 1999. Genetic and biochemical characterization of a high-affinity betaine uptake system (BusA) in *Lactococcus lactis* reveals a new functional organization within bacterial ABC transporters. *J. Bacteriol.* 181(20), 6238-6246.
- Partridge, S.M., 2000. Prosthetic valve endocarditis due to *Streptococcus vestibularis*. *J. Infect.* 41(3), 284-285.
- Pavlovic, G., Burrus, V., Gintz, B., Decaris, B., Guedon, G., 2004. Evolution of genomic islands by deletion and tandem accretion by site-specific recombination: ICES<sub>T1</sub>-related elements from *Streptococcus thermophilus*. *Microbiology.* 150(4), 759-774.
- Pette, J.W., Lolkema, H., 1950. Yoghurt. 1. Symbiosis and antibiosis in mixed cultures of *Lactobacillus bulgaricus* and *Streptococcus thermophilus*. *Nederlandsch Melk- en Zuiveltijdschrift.* 4, 197-208.
- Pflock, M., Finsterer, N., Joseph, B., Mollenkopf, H., Meyer, T.F., Beier, D., 2006. Characterization of the ArsRS regulon of *Helicobacter pylori*, involved in acid adaptation. *J. Bacteriol.* 188(10), 3449-3462.
- Picon, A., Kunji, E.R.S., Lanfermeijer, F.C., Konings, W.N., Poolman, B., 2000. Specificity Mutants of the Binding Protein of the Oligopeptide Transport System of *Lactococcus lactis*. *J. Bacteriol.* 182(6), 1600-1608.
- Picon, A., Van Wely, K.H., 2001. Peptide binding to the bacillus subtilis oligopeptide-binding proteins OppA and AppA.
- Podbielski, A., Leonard, B.A., 1998. The group A streptococcal dipeptide permease (Dpp) is involved in the uptake of essential amino acids and affects the expression of cysteine protease. *Mol Microbiol.* 28(6), 1323-34.
- Podbielski, A., Pohl, B., Woischnik, M., Körner, C., Schmidt, K.H., Rozdzinski, E., Leonard, B.A., 1996. Molecular characterization of group A streptococcal (GAS) oligopeptide

- permease (opp) and its effect on cysteine protease production. *Mol. Microbiol.* 21(5), 1087-1099.
- Polard, P., Chandler, M., 1995. Bacterial transposases and retroviral integrases. *Mol. Microbiol.* 15(1), 13-23.
- Poquet, I., Saint, V., Sez nec, E., Simoes, N., Bolotin, A., Gruss, A., Poquet, I., Saint, V., Sez nec, E., Simoes, N., Bolotin, A., Gruss, A., 2000. HtrA is the unique surface housekeeping protease in *Lactococcus lactis* and is required for natural protein processing, HtrA is the unique surface housekeeping protease in *Lactococcus lactis* and is required for natural protein processing. *Molecular Microbiology*, *Molecular Microbiology*. 35, 35(5, 5), 1042, 1042-1051, 1051.
- Poyart, C., Quesne, G., Coulon, S., Berche, P., Trieu-Cuot, P., 1998. Identification of Streptococci to Species Level by Sequencing the Gene Encoding the Manganese-Dependent Superoxide Dismutase. *J. Clin. Microbiol.* 36(1), 41-47.
- Rallu, F., Taillez, P., Ehrlich, S.D., Renault, P., 2002. Common scheme of evolution between eps clusters of the food bacteria *Streptococcus thermophilus* and cps clusters of the pathogenic streptococci. *Proceedings of the 6th American Society of Microbiology Conference on Streptococcal Genetics*. 112.
- Rasmussen, T.B., Danielsen, M., Valina, O., Garrigues, C., Johansen, E., Pedersen, M.B., 2008. *Streptococcus thermophilus* Core Genome: Comparative Genome Hybridization Study of 47 Strains. *Appl. Environ. Microbiol.* 74(15), 4703-4710.
- Roberts, A.P., Chandler, M., Courvalin, P., Guédon, G., Mullany, P., Pembroke, T., Rood, J.I., Smith, C.J., Summers, A.O., Tsuda, M., Berg, D.E., 2008. Revised nomenclature for transposable genetic elements. *Plasmid*. 60(3), 167-173.
- Rocha, E.P., Danchin, A., Viari, A., 1999. Translation in *Bacillus subtilis*: roles and trends of initiation and termination, insights from a genome analysis. *Nucleic Acids Res.* 27(17), 3567-3576.
- Rocha, E.P.C., Danchin, A., 2002. Base composition bias might result from competition for metabolic resources. *Trends Genet.* 18(6), 291-294.
- Roissart, H., Luquet, F., 1994. *Bactéries lactiques: aspects fondamentaux et technologiques*. France, 605.
- Rul, F., Gripon, J.-C., Monnet, V., 1995. St-PepA, a *Streptococcus Thermophilus* Aminopeptidase with High Specificity for Acidic Residues. *Microbiology*. 141(9), 2281-2287.
- Rul, F., Monnet, V., Gripon, J.C., 1994. Purification and characterization of a general aminopeptidase (St-PepN) from *Streptococcus salivarius* ssp. *thermophilus* CNRZ 302. *J. Dairy Sci.* 77(10), 2880-2889.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory.
- Samen, U., Gottschalk, B., Eikmanns, B.J., Reinscheid, D.J., 2004. Relevance of Peptide Uptake Systems to the Physiology and Virulence of *Streptococcus agalactiae*. *Journal of Bacteriology*. 186(5), 1398 -1408.
- Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*. 74(12), 5463 -5467.
- Sanz, Y., Lanfermeijer, F.C., Renault, P., Bolotin, A., Konings, W.N., Poolman, B., 2001. Genetic and functional characterization of dpp genes encoding a dipeptide transport system in *Lactococcus lactis*. *Arch. Microbiol.* 175(5), 334-343.

- Saurin, W., Hofnung, M., Dassa, E., 1999. Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. *J. Mol. Evol.* 48(1), 22-41.
- Saurin, W., Köster, W., Dassa, E., 1994. Bacterial binding protein-dependent permeases: characterization of distinctive signatures for functionally related integral cytoplasmic membrane proteins. *Mol. Microbiol.* 12(6), 993-1004.
- Savijoki, K., Ingmer, H., Varmanen, P., 2006. Proteolytic systems of lactic acid bacteria. *Applied Microbiology and Biotechnology.* 71(4), 394-406.
- Schleifer, K.H., Kraus, J., Dvorak, C., Kilpper-Bälz, R., Collins, M.D., Fischer, W., 1985. Transfer of *Streptococcus lactis* and Related Streptococci to the Genus *Lactococcus* gen. nov. *Systematic and Applied Microbiology.* 6(2), 183-195.
- Schmees, G., Stein, A., Hunke, S., Landmesser, H., Schneider, E., 1999. Functional consequences of mutations in the conserved « signature sequence » of the ATP-binding-cassette protein MalK. *Eur. J. Biochem.* 266(2), 420-430.
- Schneider, E., Hunke, S., 1998. ATP-binding-cassette (ABC) transport systems: functional and structural aspects of the ATP-hydrolyzing subunits/domains. *FEMS Microbiol. Rev.* 22(1), 1-20.
- Shahbal, S., Hemme, D., Desmazeaud, M., 1991. High cell wall-associated proteinase activity of some *Streptococcus thermophilus* strains (H-strains) correlated with a high acidification rate in milk La forte activité protéinase de paroi de certaines souches de *Streptococcus thermophilus* (souches H) explique leur forte vitesse d'acidification du lait. *Le Lait.* 71(3), 7 pages.
- Shahbal, S., Hemme, D., Renault, P., 1993. Characterization of a Cell Envelope-Associated Proteinase Activity from *Streptococcus thermophilus* H-Strains. *Appl. Environ. Microbiol.* 59(1), 177-182.
- Siezen, R.J., 1999. Multi-domain, cell-envelope proteinases of lactic acid bacteria. *Antonie Van Leeuwenhoek.* 76(1-4), 139-155.
- Siezen, R.J., Kuipers, O.P., Vos, W.M., 1996. Comparison of lantibiotic gene clusters and encoded proteins. *Antonie van Leeuwenhoek.* 69(2), 171-184.
- Siezen, R.J., Leunissen, J.A.M., 1997. Subtilases: The superfamily of subtilisin-like serine proteases. *Protein Science.* 6(3), 501-523.
- Slanti, L., Lereclus, D., 2002. A cell-cell signaling peptide activates the PlcR virulence regulon in bacteria of the *Bacillus cereus* group. *EMBO J.* 21(17), 4550-4559.
- Smid, E.J., Driessen, A.J., Konings, W.N., 1989. Mechanism and energetics of dipeptide transport in membrane vesicles of *Lactococcus lactis*. *J. Bacteriol.* 171(1), 292-298.
- Sonenshein, A.L., 2005. CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria. *Current Opinion in Microbiology.* 8(2), 203-207.
- Stacey, M.G., Koh, S., Becker, J., Stacey, G., 2002. AtOPT3, a Member of the Oligopeptide Transporter Family, Is Essential for Embryo Development in Arabidopsis. *The Plant Cell Online.* 14(11), 2799 -2811.
- Steiner, H.Y., Naider, F., Becker, J.M., 1995. The PTR family: a new group of peptide transporters. *Mol. Microbiol.* 16(5), 825-834.
- Steiner, H.Y., Song, W., Zhang, L., Naider, F., Becker, J.M., Stacey, G., 1994. An Arabidopsis Peptide Transporter Is a Member of a New Class of Membrane Transport Proteins. *The Plant Cell Online.* 6(9), 1289 -1299.
- Stiles, M.E., Holzappel, W.H., 1997. Lactic acid bacteria of foods and their current taxonomy. *Int. J. Food Microbiol.* 36(1), 1-29.

- Stingl, K., Uhlemann, E.-M., Schmid, R., Altendorf, K., Bakker, E.P., 2002. Energetics of *Helicobacter Pylori* and Its Implications for the Mechanism of Urease-Dependent Acid Tolerance at pH 1. *J. Bacteriol.* 184(11), 3053-3060.
- Stragier, P., Losick, R., 1996. Molecular genetics of sporulation in *Bacillus subtilis*. *Annu. Rev. Genet.* 30, 297-241.
- Subbian, E., Yabuta, Y., Shinde, U., 2004. Positive selection dictates the choice between kinetic and thermodynamic protein folding and stability in subtilases. *Biochemistry.* 43(45), 14348-14360.
- Sun, Z., Chen, X., Wang, J., Zhao, W., Shao, Y., Wu, L., Zhou, Z., Sun, T., Wang, L., Meng, H., Zhang, H., Chen, W., 2011. Complete genome sequence of *Streptococcus thermophilus* strain ND03. *J. Bacteriol.* 193(3), 793-794.
- Sutcliffe, I.C., Russell, R.R., 1995. Lipoproteins of gram-positive bacteria. *J Bacteriol.* 177(5), 1123-1128.
- Tam, R., Saier, M.H., 1993. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol. Mol. Biol. Rev.* 57(2), 320-346.
- Tame, J.R., Dodson, E.J., Murshudov, G., Higgins, C.F., Wilkinson, A.J., 1995. The crystal structures of the oligopeptide-binding protein OppA complexed with tripeptide and tetrapeptide ligands. *Structure.* 3(12), 1395-1406.
- Tame, Murshudov, G., Dodson, E., Neil, T., Dodson, G., Higgins, C., Wilkinson, A., 1994. The structural basis of sequence-independent peptide binding by OppA protein. *Science.* 264(5165), 1578 -1581.
- Tamine, A.Y., Deeth, H.C., 1980. Yogurt: technology and biochemistry. *J. Food Protection.* 43, 939-977.
- Taylor, D.L., Ward, P.N., Rapier, C.D., Leigh, J.A., Bowler, L.D., 2003. Identification of a Differentially Expressed Oligopeptide Binding Protein (OppA2) in *Streptococcus uberis* by Representational Difference Analysis of cDNA. *Journal of Bacteriology.* 185(17), 5210 -5219.
- Terzaghi, B.E., Sandine, W.E., 1975. Improved medium for lactic streptococci and their bacteriophages. *Appl Microbiol.* 29(6), 807-813.
- Tettelin, H., 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science.* 293, 498-450.
- Tettelin, H., 2004. Streptococcal genomes provide food for thought. *Nature biotechnology.* 22(12), 1523–1524.
- Thauer, R.K., Diekert, G., Schönheit, P., 1980. Biological role of nickel. *Trends in Biochemical Sciences.* 5(11), 304-306.
- Thomas, C.M., Nielsen, K.M., 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* 3(9), 711-721.
- Thomas, T., Turner, K.W., 1977. Preparation of skim milk to allow harvesting of starter cells from milk cultures. *N Z J Dairy Sci Technol.* 15–21.
- Tinson, W., Broome, M.C., Hillier, A.J., Jago, G.R., 1982. Metabolism of *Streptococcus thermophilus*. 2. Production of carbon dioxide and ammonia from urea. *Australian Journal of Dairy Technology.*
- Tsakalidou, E., Anastasiou, R., Papadimitriou, K., Manolopoulou, E., Kalantzopoulos, G., 1998. Purification and characterisation of an intracellular X-prolyl-dipeptidyl aminopeptidase from *Streptococcus thermophilus* ACA-DC 4. *Journal of Biotechnology.* 59(3), 203-211.
- Tsay, Y.-F., Schroeder, J.I., Feldmann, K.A., Crawford, N.M., 1993. The herbicide sensitivity

- gene CHL1 of arabisopsis encodes a nitrate-inducible nitrate transporter. *Cell*. 72(5), 705-713.
- Tynkkynen, S., Buist, G., Kunji, E., Kok, J., Poolman, B., Venema, G., Haandrikman, A., 1993. Genetic and biochemical characterization of the oligopeptide transport system of *Lactococcus lactis*. *J. Bacteriol.* 175(23), 7523-7532.
- Vaillancourt, K., LeMay, J.-D., Lamoureux, M., Frenette, M., Moineau, S., Vadeboncoeur, C., 2004. Characterization of a Galactokinase-Positive Recombinant Strain of *Streptococcus thermophilus*. *Appl. Environ. Microbiol.* 70(8), 4596-4603.
- Vandamme, P., Pot, B., Gillis, M., de Vos, P., Kersters, K., Swings, J., 1996. Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol. Rev.* 60(2), 407-438.
- Vaughan, E.E., van den Bogaard, P.T.C., Catzeddu, P., Kuipers, O.P., de Vos, W.M., 2001. Activation of Silent gal Genes in the lac-gal Regulon of *Streptococcus thermophilus*. *J. Bacteriol.* 183(4), 1184-1194.
- Verheul, A., Rombouts, F.M., Abee, T., 1998. Utilization of Oligopeptides by *Listeria monocytogenes* Scott A. *Appl. Environ. Microbiol.* 64(3), 1059-1065.
- Wages, H.E., Taft, K.L., Lippard, S.J., 1993. Tris(acetato)bis(N,N,N',N'-tetramethylethylenediamine)(urea)dinickel(1+) triflate and [Ni(OAc)(urea)2(tmen)](OTf), model complexes for the enzyme urease. *Inorg. Chem.* 32(23), 4985-4987.
- Walker, J.E., Saraste, M., Runswick, M.J., Gay, N.J., 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1(8), 945-951.
- Wang, C.-H., Lin, C.-Y., Luo, Y.-H., Tsai, P.-J., Lin, Y.-S., Lin, M.T., Chuang, W.-J., Liu, C.-C., Wu, J.-J., 2005. Effects of Oligopeptide Permease in Group A Streptococcal Infection. *Infect. Immun.* 73(5), 2881-2890.
- Wu, L.F., Navarro, C., de Pina, K., Quénard, M., Mandrand, M.A., 1994. Antagonistic effect of nickel on the fermentative growth of *Escherichia coli* K-12 and comparison of nickel and cobalt toxicity on the aerobic and anaerobic growth. *Environ Health Perspect.* 102(Suppl 3), 297-300.
- Yu, W., Gillies, K., Kondo, J.K., Broadbent, J.R., McKay, L.L., 1996. Loss of Plasmid-Mediated Oligopeptide Transport System in *Lactococci*: Another Reason for Slow Milk Coagulation. *Plasmid.* 35(3), 145-155.
- Zaneveld, J.R., Nemergut, D.R., Knight, R., 2008. Are All Horizontal Gene Transfers Created Equal? Prospects for Mechanism-Based Studies of HGT Patterns. *Microbiology.* 154(1), 1-15.
- Zotta, T., Ricciardi, A., Rossano, R., Parente, E., 2008. Urease production by *Streptococcus thermophilus*. *Food Microbiology.* 25(1), 113-119.
- Zourari, A., Desmazeaud, M.J., 1991. Characterization of lactic acid bacteria isolated from Greek yogurts. II. Strains of *Lactobacillus delbrueckii* subsp *bulgaricus* and mixed cultures with *Streptococcus salivarius* subsp *thermophilus*. Caractérisation de bactéries lactiques thermophiles isolées de yaourts artisanaux grecs. II. Souches de *Lactobacillus delbrueckii* subsp *bulgaricus* et cultures mixtes avec *Streptococcus salivarius* subsp *thermophilus*. *Le Lait.* 71(4), 20.
- Zourari, A., Roger, S., Chabanet, C., Desmazeaud, M.J., 1991. Characterization of lactic acid bacteria isolated from Greek yogurts. I. *Streptococcus salivarius* subsp *thermophilus* strains. Caractérisation de bactéries lactiques thermophiles isolées de yaourts artisanaux grecs. I. Souches de *Streptococcus salivarius* subsp *thermophilus*. *Le Lait.* 71(4), 17.

**AUTORISATION DE SOUTENANCE  
DU DOCTORAT DE L'UNIVERSITE DE LORRAINE**

**o0o**

VU LES RAPPORTS ETABLIS PAR :

**Monsieur Vincent JULLIARD, Directeur de Recherches, INRA/MICALIS**

**Jouy-en-Josas,**

**Monsieur Jean-François CAVIN, Professeur, AgroSup, Université de Dijon.**

Le Président de l'Université de Lorraine, autorise :

**Madame JAMEH Nawara**

à soutenir devant un jury de l'UNIVERSITE DE LORRAINE, une thèse intitulée :

**"Système protéolytique de surface de *Streptococcus thermophilus* :**

**- variabilité des capacités d'hydrolyse des caséines,**

**- caractérisation d'un nouveau système de transport de peptides"**

en vue de l'obtention du titre de :

**DOCTEUR DE L'UNIVERSITE DE LORRAINE**

Intitulé du doctorat : **"Sciences Agronomiques"**

Fait à Vandoeuvre, le **4 juin 2012**

Le Président de l'Université de Lorraine,

Pierre MUTZENHARDT



## **Système protéolytique de surface de *Streptococcus thermophilus* :**

- variabilité des capacités d'hydrolyse des caséines**
- caractérisation d'un nouveau système de transport de peptides**

### **Résumé :**

*S. thermophilus* est une bactérie largement employée dans la fabrication des produits laitiers. La capacité des souches de *S. thermophilus* à générer des peptides bioactifs à partir des caséines bovines a été étudiée. Dix souches exprimant différemment la protéase de surface, PrtS, ont été incubées en présence de la caséine  $\alpha_{s1}$ ,  $\alpha_{s2}$  ou  $\beta$ . Le nombre et le type de peptides libérés dépendent de la souche utilisée. Des peptides connus comme des peptides bioactifs ont été détectés : 13 peptides ont été générés à partir de la caséine  $\beta$ , 5 peptides à partir de la caséine  $\alpha_{s2}$  et 2 peptides à partir de la caséine  $\alpha_{s1}$ . L'utilisation de cette bactérie pour la production de tels peptides dans l'aliment requiert qu'elle en internalise le moins possible. Nous nous sommes intéressés aux systèmes de transport de peptides présents au sein de l'espèce *S. thermophilus*. Une collection de 22 souches de *S. thermophilus* a été choisie pour étudier la variabilité des systèmes de transport des peptides présents au sein de l'espèce. Toute d'abord, la proximité phylogénétique entre les souches a été évaluée par MLST, puis la variabilité génétique du système de transport des oligopeptides Ami et du transporteur de di- et tripeptides DtpT a été étudiée au sein de cette collection. Un cluster, composé de 4 gènes, annoté en tant que transporteur ABC de peptides et de nickel a été détecté au sein du génome de la souche LMD-9, et appelé Ots. Il est présent chez 9 sur 22 souches de *S. thermophilus* et est transcrit tout au long de la croissance en milieu M17. La caractérisation du système Ots suggère qu'il est impliqué dans l'internalisation de peptides de petites tailles.

**Mots clés :** *Streptococcus thermophilus*, système de transport des peptides, cluster *ots*, transfert horizontal, MLST.

## **Proteolytic system of the surface of *Streptococcus thermophilus*:**

- Variability in the capacity of casein hydrolysis**
- Characterization of a novel peptide transport system**

### **Summary:**

*S. thermophilus* is a widely used bacterium in the manufacture of dairy products. The capacity of *S. thermophilus* to generate bioactive peptides from bovine caseins was studied. Ten strains expressing different levels of the cell envelope protease, PrtS, were incubated with  $\alpha_{s1}$ -,  $\alpha_{s2}$ - or  $\beta$ -casein. Number and type of peptides released were strain-dependent. Peptides known as bioactive peptides were detected: 13 peptides were generated from  $\beta$ -casein, 5 peptides from  $\alpha_{s2}$ -casein and 2 peptides from  $\alpha_{s1}$ -casein. The use of this bacterium for the production of such peptides in the food products requires the least internalization of these peptides by this bacterium. We were interested in knowing the peptide transport system present in the species *S. thermophilus*. A collection of 22 strains of *S. thermophilus* was chosen to study the genetic variability of peptide transport systems present within the species. First of all, we evaluated the phylogenetic proximity between selected strains by MLST, and then the genetic variability of the transport system of oligopeptides Ami and di- and tripeptides transporter, DtpT were studied in this collection. A cluster consisting of four genes, annotated as ABC transporter of peptides and nickel was detected in the genome of strain LMD-9, and called Ots. It is present in 9 of 22 strains of *S. thermophilus* and is transcribed throughout the growth in M17 medium. The Ots system seems to be involved in the internalization of smaller sized peptides.

**Keywords:** *Streptococcus thermophilus*, peptide transport system, cluster *ots*, horizontal transfer, MLST.