



HAL
open science

Caractérisation des régulateurs transcriptionnels Rgg et étude du rôle de la protéine Rgg0182 de *Streptococcus thermophilus*

Romain Henry

► **To cite this version:**

Romain Henry. Caractérisation des régulateurs transcriptionnels Rgg et étude du rôle de la protéine Rgg0182 de *Streptococcus thermophilus*. Bactériologie. Université Henri Poincaré - Nancy 1, 2011. Français. NNT : 2011NAN10087 . tel-01746236

HAL Id: tel-01746236

<https://hal.univ-lorraine.fr/tel-01746236>

Submitted on 29 Mar 2018

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

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



AVERTISSEMENT

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

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

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

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

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

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

http://www.cfcopies.com/V2/leg/leg_droi.php

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

Ecole doctorale BioSE (Biologie - Santé - Environnement)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITE HENRI POINCARÉ

Mention : « Sciences de la Vie et de la Santé »

par **Romain Henry**

**Caractérisation des régulateurs transcriptionnels Rgg et étude
du rôle de la protéine Rgg₀₁₈₂ de *Streptococcus thermophilus***

Soutenance le 10 novembre 2011

Membres du Jury :

Rapporteurs :	M. Jean-François Cavin Mme Christine Delorme	Professeur, Université de Bourgogne Chargée de recherche (HDR), INRA, Jouy-en-Josas
Examineurs :	Mme Annie Dary-Mourot Mme Nathalie Leblond-Bourget	Professeur, Nancy-Université Professeur, Nancy-Université, co-directeur de thèse
Membres invités :	M. Bernard Decaris M. Pierre Leblond	Professeur, Nancy-Université, co-directeur de thèse Professeur, Nancy-Université

Ecole doctorale BioSE (Biologie - Santé - Environnement)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITE HENRI POINCARÉ

Mention : « Sciences de la Vie et de la Santé »

par **Romain Henry**

**Caractérisation des régulateurs transcriptionnels Rgg et étude
du rôle de la protéine Rgg₀₁₈₂ de *Streptococcus thermophilus***

Soutenance le 10 novembre 2011

Membres du Jury :

Rapporteurs :	M. Jean-François Cavin Mme Christine Delorme	Professeur, Université de Bourgogne Chargée de recherche (HDR), INRA, Jouy-en-Josas
Examineurs :	Mme Annie Dary-Mourot Mme Nathalie Leblond-Bourget	Professeur, Nancy-Université Professeur, Nancy-Université, co-directeur de thèse
Membres invités :	M. Bernard Decaris M. Pierre Leblond	Professeur, Nancy-Université, co-directeur de thèse Professeur, Nancy-Université

Remerciements

Je tiens tout d'abord à remercier Bernard Decaris et Pierre Leblond pour m'avoir accueilli au sein de leur laboratoire.

Je remercie Bernard Decaris pour m'avoir accueilli au sein de son équipe, pour sa confiance, pour ses conseils, remarques et discussions constructives. Vous avez eu un premier thésard... et je suis le dernier... Bonne retraite !

Je remercie également Nathalie Leblond pour m'avoir encouragé et soutenu, pour ses conseils et nos discussions scientifiques. Je t'ai fait relire plusieurs versions de ma thèse jusqu'à la dernière minute... et je t'en remercie. Je te remercie également pour ta sympathie et ta bonne humeur.

Je souhaite remercier Christine Delorme (Chargée de Recherche à l'INRA de Jouy-en-Josas) et Jean-François Cavin (Professeur à l'Université de Bourgogne) pour avoir accepté la lourde tâche de rapporter ce manuscrit. Je remercie vivement Annie Dary et Pierre Leblond pour avoir accepté de juger ce travail.

Ce travail est le fruit de collaborations qui, sans elles, n'aurait pas la même saveur et je souhaite vivement remercier l'ensemble des personnes qui ont participé à ce sujet :

- Emmanuelle Bruneau actuellement en thèse au LGM et qui continue ce sujet. Emmanuelle a réalisé les clonages pour la production de la protéine Rgg₀₁₈₂, sa purification et les expériences de retard de migration sur gel. Merci pour tes nombreux coups de mains sur la fin de ma thèse, pour ton acharnement à faire fonctionner les gels retards capricieux et pour nos discussions scientifiques toujours très animées. Je te souhaite de réussir et surtout de t'épanouir dans ce domaine un peu fou.

- Julien Hiblot qui a su, lors de son stage de master au LGM, redonner une bouffée d'oxygène à ce sujet. Ton acharnement nous a permis de prendre des orientations importantes pour le sujet.

- Stéphane Bertin (adjoint technique au LGM) pour avoir participé au sujet, notamment les stress, malgré le temps passé à la préparation des TP et le travail commun du labo. Tu as réussi à supporter le stress et à faire preuve de grand professionnalisme dans les moments de rush.

- Véronique Monnet, Rozenn Gardan et Betty Fleuchot du laboratoire Peptides et communications bactériennes. Rozenn et Betty ont réalisé les fusions transcriptionnelles des petites ORF bordant « notre rgg » et analyser la luminescence dans les souches sauvage et mutant. Merci également pour la relecture de l'article.

- Grégory Francius du LCPME à Villers-lès-Nancy. Grégory a réalisé les analyses par microscopie à force atomique et m'a également appris à utiliser cet appareil (mais pas les formules de physique très complexes ;-). Ton enthousiasme et ton implication pour les sujets qu'on te propose sont ta réussite ! Il te faudra un peu de bouteille pour être le spécialiste de la microscopie à force atomique chez les bactéries mais cela viendra, j'en suis sûr.

- Jérôme Duval et Jenny Merlin du LEM pour avoir réalisé les mobilités électrophorétiques.

Je n'oublie pas Brigitte, Ludo, K'ro ! Les trois premiers techniciens avec qui j'ai travaillé. Vous êtes restés peu de temps mais je garde d'excellents souvenirs de nos collaborations.

Je remercie également tous les stagiaires qui ont participé au sujet, Christophe, Nancy, Fabien, Alex, Justine, Maude, Stéphanie et Mathieu avec lesquels j'ai passé d'agréables moments. Je vois déjà la réussite des premiers et je suis fier d'y avoir contribué, même très peu !

Je remercie l'ensemble du personnel du LGM pour leur accueil, leur serviabilité et leur sympathie.

Une mention spéciale à Emmanuelle et Nico. Nico, mon collègue de rédaction d'article et de thèse, j'ai pris un peu d'avance sur toi au final ! Merci pour ton soutien et tes encouragements pendant ces mois de rédaction, pas toujours évidents ! Et Manu, tu as supporté mon humeur pendant la rédaction, tu m'as encouragé et, parfois, même boosté quand je n'avais plus de jus. Je te rendrai la pareille quand ce sera ton tour !

Je voudrais remercier plus particulièrement : tout d'abord, les « anciens » Miky, Mag, Nico, Virge, Guiguix, Céline et Alex (merci à vous 7 pour votre accueil au labo), mais aussi JY, JB, Sibel, Kevin, Axelle, Jérôme et Philippe, pour les soirées, les « murders », la wii, les bonnes bouffes , les jeux de cartes, de plateau,... et surtout pour cette amitié qui s'est créée et qui continue de se forger ; et ensuite les « petits » (ou ceux qui sont arrivés après moi ;-), Manu, Frisouille, Maxou, Jaja, Sarah, Aurore, Emilie, Ludo, PV et P'tit Manu pour les soirées, les restos et pour tous nos délires de toutes sortes. Pour ceux qui sont en thèse, je vous souhaite bon courage pour la suite, je vous assure qu'il y a une sortie au bout du tunnel !

Je souhaite remercier Bernard Vitoux, Laurence Hensch et Laurence Dendien du Centre d'Initiation à l'Enseignement Supérieur de Lorraine. Votre implication et votre soutien pour les thésards a toujours été une source d'énergie ! Le monitorat m'aura permis de développer mon goût de l'enseignement.

Je souhaite également remercier Mimi et Joëlle, les petites dames du RU qui m'ont fournis les calories nécessaires et le café du midi, pour tenir tous les après-midi mais aussi pour leur sympathie !

Je ne voudrai pas oublier les rencontres tels que Smith, Stout, K'ro, Zed et, plus récemment, Jaja et Maxou (et les rencontres faites par cet intermédiaire, Gaby, Kichou, Zed) pour avoir pu partager notre passion pour la musique, pour avoir joué ensemble et pour les nombreux, très nombreux bons moments passés avec vous. Ce fut une source d'énergie pendant ces années difficiles de thèse, même si cela n'aura duré qu'à peine deux ans...

Enfin, je voudrai remercier ma famille. Merci à mes grands-parents qui m'ont hébergé pendant toute la durée de mes études, ça fait déjà 11 ans ! Ces années m'auront permis de mieux vous connaître et de partager des moments uniques avec vous.

Merci à mes parents pour m'avoir encouragé et soutenu (financièrement aussi ;-). Votre réussite est un exemple pour moi ! Merci pour tout...

Merci à mes frères, Cyril et Sébastien, et leur famille respective pour les moments passés en votre compagnie. Ma soutenance sera pour vous l'occasion d'essayer de comprendre ce que je fais ;-). Vous avez réussi, formidablement, votre vie professionnelle ! Il ne reste qu'à moi de faire de même !

Enfin, et surtout, je remercie Emmanuelle, celle qui partage ma vie, qui m'encourage et me soutient dans tout ce que j'entreprends. Ta présence à mes côtés est une source d'énergie et aura été plus que bénéfique à la fin de ma thèse. Merci de ta compréhension dans les moments stressants. Restes telle que tu es !

Je dédicace ma thèse à Manu.

A la mémoire de mes grands-pères.

SOMMAIRE

SOMMAIRE	1
ABBREVIATIONS	5
PREAMBULE	9
I - INTRODUCTION	13
1. STREPTOCOCCUS THERMOPHILUS, UNE BACTERIE LACTIQUE LARGEMENT UTILISEE EN INDUSTRIE LAITIERE	13
1.1. <i>Les bactéries lactiques : utilisation en industrie agro-alimentaire</i>	13
1.1.1. Qu'est-ce qu'une bactérie lactique ?	13
1.1.2. Utilisation des bactéries lactiques dans la production de produits alimentaires	13
1.2. <i>Streptococcus thermophilus présente un intérêt économique important</i>	14
1.2.1. Les bactéries lactiques à l'origine des laits fermentés.....	14
1.2.2. Choix des bactéries lactiques en industrie laitière	15
1.2.3. <i>Streptococcus thermophilus</i> est l'une des bactéries les plus utilisées en industrie laitière	16
1.2.3.1. Les laits fermentés par <i>Streptococcus thermophilus</i>	16
1.2.3.2. Le genre <i>Streptococcus</i>	17
1.2.3.3. L'espèce <i>Streptococcus thermophilus</i>	19
1.3. <i>Les procédés industriels, des stress pour Streptococcus thermophilus</i>	22
1.3.1. Qu'est-ce qu'un stress ?	22
1.3.2. Les procédés industriels produisent du stress	22
2. LES RESEAUX DE REGULATION DE LA REPONSE AU STRESS PRESENTS CHEZ STREPTOCOCCUS THERMOPHILUS	25
2.1. <i>Des réseaux de régulation sont mis en place afin de répondre aux stress</i>	25
2.2. <i>Les facteurs sigma</i>	26
2.2.1. Qu'est-ce qu'un facteur sigma ?	26
2.2.2. Les facteurs sigma chez <i>Streptococcus thermophilus</i>	27
2.3. <i>Les régulateurs transcriptionnels</i>	29
2.3.1. La diversité des régulateurs transcriptionnels	29
2.3.2. Les systèmes à un composant ou régulateurs OCS	31
2.3.2.1. Qu'est-ce qu'un système à un composant ?.....	31
2.3.2.2. Les régulateurs OCS de <i>Streptococcus thermophilus</i>	32
2.3.2.2.1. Exemple du régulateur HrcA	32
2.3.2.2.2. Exemple du régulateur CtsR.....	33
2.3.3. Les systèmes à deux composants ou régulateurs TCS	34
2.3.3.1. Qu'est-ce qu'un système à deux composants ?.....	34
2.3.3.2. Les régulateurs TCS de <i>Streptococcus thermophilus</i>	35
3. LES PROTEINES DE TYPE RGG PARTICIPENT A LA REPONSE AUX STRESS	38
3.1. <i>Les protéines Rgg constituent une famille de régulateur transcriptionnel</i>	38
3.1.1. Les gènes <i>rgg</i> codent des régulateurs transcriptionnels	38
3.1.2. Certaines protéines Rgg sont des régulateurs cellulaires globaux	39
3.2. <i>Les protéines Rgg se lient spécifiquement à l'ADN</i>	40
3.2.1. Les protéines Rgg possèdent un domaine HTH.....	40
3.2.2. Les protéines Rgg se fixent dans la région promotrice de leur gènes cibles	41
3.3. <i>Implication des régulateurs Rgg dans la réponse aux stress</i>	42
3.4. <i>Implication du régulateur Rgg₁₃₅₈ de S. thermophilus LMD9 dans un mécanisme de</i>	

<i>quorum sensing</i>	43
3.4.1. Mise en évidence de petites ORF non annotées dans les banques de données	43
3.4.2. Mécanisme de quorum-sensing faisant intervenir les protéines Shp ₁₃₅₈ et Rgg ₁₃₅₈ ...	44
4. ETAT DE L'ART.....	46
II – RESULTATS.....	51
1. CARACTERISATION <i>IN SILICO</i> DES PROTEINES APPARTENANT A LA FAMILLE RGG.....	51
1.1. La majorité des protéines Rgg sont présentes au sein de la famille des <i>Streptococcaceae</i> et des genres <i>Lactobacillus</i> et <i>Listeria</i>	51
1.2. Plusieurs copies rgg peuvent être présentes par génome.....	54
1.3. Les protéines de la famille Rgg se caractérisent par un domaine HTH et un domaine médian.....	57
1.3.1. Deux domaines peuvent être identifiés au sein des protéines Rgg	57
1.3.2. Le domaine médian est spécifique des protéines Rgg	59
1.4. Parenté phylogénétique des copies Rgg	62
1.5. Caractérisation des copies Rgg présentes chez <i>S. thermophilus</i>	64
1.5.1. Onze copies Rgg distinctes sont présentes chez <i>S. thermophilus</i>	64
1.5.2. Polymorphisme et parenté des protéines Rgg de <i>S. thermophilus</i>	66
1.5.3. Etude de l'environnement génétiques des copies rgg de <i>S. thermophilus</i>	69
2. ROLE DES GENES RGG DE <i>S. THERMOPHILUS</i> LMG18311.....	71
2.1. Obtention d'une collection de mutants Δ rgg de <i>S. thermophilus</i> LMG18311	71
2.2. La délétion de gènes rgg n'a pas d'impact sur la croissance de <i>S. thermophilus</i> LMG18311 en condition de laboratoire.....	73
2.3. Impact des protéines Rgg dans l'organisation des cellules en chaînes	75
2.4. Tolérance des mutants à différents stress.....	79
3. LE GENE RGG₀₁₈₂ CODE UN REGULATEUR TRANSCRIPTIONNEL REQUIS DANS L'ADAPTATION AUX CHANGEMENTS DE TEMPERATURE DE <i>STREPTOCOCCUS THERMOPHILUS</i> LMG18311	82
3.1. Caractérisation du locus rgg ₀₁₈₂ de <i>S. thermophilus</i> LMG18311.....	82
3.1.1. Analyse de la séquence du gène rgg ₀₁₈₂	82
3.1.2. L'environnement génétique du gène rgg ₀₁₈₂ est similaire à celui du gène rgg ₁₃₅₈	84
3.2. Etude du taux de transcrit rgg ₀₁₈₂ de <i>S. thermophilus</i> LMG18311.....	85
3.3. Le gène rgg ₀₁₈₂ de <i>S. thermophilus</i> LMG18311 code un régulateur transcriptionnel	87
3.3.1. Implication du gène rgg ₀₁₈₂ de <i>S. thermophilus</i> LMG18311 dans la transcription des gènes adjacents.....	87
3.3.2. La protéine Rgg ₀₁₈₂ de <i>S. thermophilus</i> LMG18311 se fixe dans la région promotrice des gènes cibles.....	89
3.4. Implication du régulateur transcriptionnel Rgg ₀₁₈₂ de <i>S. thermophilus</i> LMG18311 dans la transcription de gènes codant des chaperons et protéases	90
3.5. Implication de la protéine Rgg ₀₁₈₂ de <i>S. thermophilus</i> LMG18311 dans la réponse au stress thermique.....	92
3.5.1. Implication de la protéine Rgg ₀₁₈₂ dans la réponse au stress chaud	92
3.5.2. Implication de la protéine Rgg ₀₁₈₂ dans la transcription des gènes <i>clpE</i> et <i>cspB</i> suite à un stress chaud.....	94
3.6. Effet du stress sur le taux de transcrit des chaperons et des protéases	95
4. IMPLICATION DU GENE RGG₀₁₈₂ DE <i>S. THERMOPHILUS</i> LMG18311 DANS L'HYDROPHOBICITE THERMO-INDUITE DES CELLULES	97
4.1. Implication de la protéine Rgg ₀₁₈₂ dans un phénomène d'adhésion cellulaire thermo-	

<i>induite</i>	97
4.2. <i>Caractérisation des forces hydrophobes sur les parois des microplaques</i>	101
4.3. <i>Analyse de l'hydrophobicité des cellules</i>	103
III – DISCUSSION ET PERSPECTIVES	107
1. COMMENT FONCTIONNENT LES PROTEINES RGG ?	107
1.1. <i>Comment les protéines Rgg contrôlent-elles l'expression de leurs cibles ?</i>	107
1.2. <i>Quelle est la forme active des protéines Rgg ?</i>	110
1.3. <i>Les protéines Rgg sont-elles toutes activées par un mécanisme de QS ?</i>	111
2. QUEL EST LE ROLE PHYSIOLOGIQUE DES PROTEINES RGG DE S. THERMOPHILUS ?	113
2.1. <i>Les protéines Rgg de S. thermophilus sont-elles impliquées dans la réponse aux stress ?</i>	113
2.1.1. <i>Les protéines Rgg ont-elles un effet antagoniste ou compensateur ?</i>	113
2.1.2. <i>La protéine Rgg₀₁₈₂ est-elle impliquée dans les réseaux de régulation de la réponse au stress thermique de S. thermophilus ?</i>	114
2.2. <i>La protéine Rgg₀₁₈₂ contrôle-t-elle la synthèse de protéines impliquées dans la division cellulaire ?</i>	117
2.3. <i>Est-ce que les protéines Rgg sont des régulateurs transcriptionnels globaux ?</i>	119
3. COMMENT EXPLIQUER LA REPARTITION DES PROTEINES RGG CHEZ LES BACTERIES ?	122
4. PERSPECTIVES DE TRAVAIL	126
4.1. <i>Perspectives à court et moyen terme</i>	126
4.2. <i>Perspectives à long terme</i>	126
BIBLIOGRAPHIE	131
ANNEXES	147
ARTICLE	147
ACTES DE CONGRES	183

ABBREVIATIONS

AA	: Acide aminé
°C	: Degrés Celsius
µg	: Microgramme (10 ⁻⁶ Gramme)
A	: Adénine
<i>A.</i>	: <i>Aerococcus</i>
ADN	: Acide DéshoxyriboNucléique
AFM	: Microscopie à force atomique
ARN	: Acide RiboNucléique
ATP	: Adénosine TriPhosphate
<i>B.</i>	: <i>Bacillus</i>
BLAST	: Basic Local Alignment Search Tool
C	: Cytosine
<i>C.</i>	: <i>Carnobacterium</i>
CDD	: Conserve Domain Database
CFM	: Chemical Force Microscopy
CIRCE	: <u>C</u> ontrolling <u>I</u> nverted <u>R</u> epeat of <u>C</u> haperone <u>E</u> xpression
CNRZ	: Centre National de Recherche Zoologique
DO	: Densité Optique
DR	: Répétition Directe
<i>E.</i>	: <i>Enterococcus</i>
Ery	: Erythromycine
G	: Guanine
<i>G.</i>	: <i>Granulicatella</i>
GABA	: Glutamate-γ-aminobutyrate
GTF	: Glucosyltransférase
HLH	: Helix-Loop-Helix
HMM	: Hidden Markov Model
HTH	: Hélice-Tour-Hélice
IFR	: Institut Fédératif de Recherche
INRA	: Institut National de Recherche Agronomique
IR	: Répétition Inversée
IS	: Séquence d'insertion
kb	: kilobase
kDa	: KiloDalton
kg	: kilogramme
<i>L.</i>	: <i>Lactococcus</i>
<i>Lb.</i>	: <i>Lactobacillus</i>
LCPME	: Laboratoire de Chimie Physique et Microbiologie pour l'Environnement
<i>Le.</i>	: <i>Leuconostoc</i>
LGM	: Laboratoire de Génétique et Microbiologie
<i>Li.</i>	: <i>Listeria</i>
LM17	: Medium 17 Lactose
MCD	: Milieu chimiquement défini
mg	: milligramme

ABBREVIATIONS

min	: minute
ML	: Maximum de vraisemblance
ml	: millilitre (10^{-3} Litre)
nN	: nano Newton
NJ	: Neighbour-Joining
<i>O.</i>	: <i>Oenococcus</i>
OCS	: One Component System
Osr	: Oxidative stress resistance
ORF	: Open Reading Frame
<i>Pa.</i>	: <i>Parascardovia</i>
pb	: paire de bases
PCR	: Polymerase Chain Reaction
<i>Pe.</i>	: <i>Pediococcus</i>
pH	: potentiel Hydrogène
PSI-BLAST	: <u>P</u> osition- <u>S</u> pecific <u>I</u> terated <u>B</u> asic <u>L</u> ocal <u>A</u> lignment <u>S</u> earch <u>T</u> ool
qPCR	: PCR quantitative en temps réel
QS	: quorum sensing
RBS	: Ribosome Binding Site
Rgg	: Regulatory gene glucosyltransferase
<i>S.</i>	: <i>Streptococcus</i>
<i>Sc.</i>	: <i>Scardovia</i>
Sod	: Superoxide dismutase
<i>St.</i>	: <i>Staphylococcus</i>
STK	: Sérine Thréonine Kinase
T	: Thymine
TCS	: Two Component signal transduction Systems
TH	: Transfert horizontal
TPR	: Tétratricopeptide
UFC	: Unité Formant Colonie
<i>W.</i>	: <i>Weissella</i>
XRE	: Xenobiototic Response Element

PREAMBULE

PREAMBULE

Les bactéries lactiques sont des microorganismes incluses dans un groupe hétérogène c'est-à-dire non regroupées entre-elles d'un point de vue phylogénétique mais fonctionnelle. Celles-ci sont utiles à l'homme pour la production de nombreux produits alimentaires fermentés, comme, par exemple, les yaourts et les fromages, issus de la transformation du lait.

L'intérêt des bactéries lactiques dans la production de produits laitiers remonte, suivant les traditions, à la plus haute antiquité lorsque l'homme est passé du statut de chasseur-cueilleur à celui d'éleveur, aux environs de 8 000 avant J.-C. Dès cette époque, l'homme sait contrôler le phénomène de caillage du lait (Taillez, 2001). Les laits fermentés constituent une source bon marché de protéines hautement digestibles et la présence d'une flore lactique vivante confère à ces produits une haute qualité nutritionnelle (Loones, 1994). En France, par exemple, la consommation annuelle de yaourts et de fromages par habitant est estimée à environ 21 et 24 kg, respectivement. Ainsi, les bactéries lactiques présentent de forts intérêts économiques, techniques et nutritionnels.

Streptococcus thermophilus est une bactérie lactique couramment employée dans l'industrie laitière. En effet, elle est utilisée en association avec *Lactococcus lactis* (*L. lactis*) pour la fabrication de fromages à pâte pressée cuite et avec *Lactobacillus bulgaricus* (*Lb. bulgaricus*) pour la fabrication de yaourts. Ces souches, indépendamment de leur pouvoir acidifiant, confèrent aux laits fermentés des caractères rhéologiques particuliers notamment de viscosité, de goût et permettent leur conservation et leur salubrité (Bouillanne et Desmazeaud, 1981 ; Giraffa et Bergère, 1987). Cependant, les procédés industriels conduisent à des modifications des conditions environnementales (modification de la concentration en oxygène ou en sel, de la température, de l'acidité) qui peuvent entraîner des stress pour les bactéries lactiques utilisées. Ainsi, l'un des défis les plus importants de l'industrie laitière est donc de choisir les genres, les espèces et les souches bactériennes les mieux adaptés aux fonctions qui leur sont imparties dans les diverses technologies de transformation du lait par fermentation (Schmidt *et al.*, 1994). Pour ce faire, les souches bactériennes utilisées doivent avoir la capacité à s'adapter aux stress afin de survivre et assurer une qualité sanitaire aux produits fermentés.

Pour s'adapter aux stress, les bactéries mettent en place des systèmes de défense. Les gènes impliqués dans ces systèmes sont nombreux et codent des protéines aux fonctions variées. En effet, ces gènes codent des régulateurs transcriptionnels qui contrôlent la synthèse de protéines spécifiques du stress, comme par exemple, des protéines de détoxification ou des protéines de réparation de l'ADN. L'identification de ces gènes permet d'explorer les aspects fondamentaux de l'adaptation des bactéries à leur environnement. Le séquençage complet de génomes de microorganismes, a

PREAMBULE

permis d'acquérir de nouvelles données, sur la présence potentielle des systèmes de défense contre les stress (Koonin *et al.*, 2000). Ces données montrent que de nombreux facteurs, tels que la taille du génome, la niche écologique, le transfert horizontal ou la perte de gènes, engendrent une diversité des systèmes de défense entre les microorganismes. Par ailleurs, l'écosystème dans lequel évolue une bactérie exerce également une pression de sélection sur ces gènes. Ainsi, la présence d'un système de défense au sein d'un genre bactérien n'est pas forcément transposable à un autre genre (Koonin *et al.*, 2000). En effet, la présence d'éléments non essentiels (plasmides, prophages, séquences d'insertions (IS),...) contribue à rendre le génome des bactéries lactiques plus flexibles et dynamiques et par conséquent à modifier les propriétés phénotypiques des souches, dans la mesure où ceux-ci peuvent être perdus ou acquis.

Les travaux réalisés au cours de cette thèse s'appuient sur l'exemple de la bactérie *S. thermophilus* et l'implication de régulateurs transcriptionnels appelés Rgg (Regulatory glucosyltransférase gene) dans l'adaptation de cette bactérie. En effet, l'équipe de Nathalie Leblond-Bourget, du laboratoire de Génétique et Microbiologie (LGM), a montré l'implication de gènes appelés *rgg* dans la réponse au stress oxydant de *S. thermophilus* (Thibessard *et al.*, 2004 ; Fernandez *et al.*, 2004a).

La thèse a fait l'objet d'une collaboration avec l'équipe « Peptides et communication bactérienne », dirigée par Véronique Monnet, de l'INRA de Jouy-en-Josas. Cette équipe est spécialisée dans l'étude et la caractérisation de petites ORF, appelé *shp*, qui sont impliqués dans des mécanismes de quorum-sensing faisant intervenir des protéines Rgg (Ibrahim *et al.*, 2007b ; Fleuchot *et al.*, 2011).

La thèse a également permis une collaboration avec Grégory Francius du Laboratoire de Chimie-Physique et Microbiologie pour l'Environnement (LCPME) de Villers-lès-Nancy. G. Francius est spécialisé dans l'utilisation de microscopie à force atomique permettant l'analyse de surfaces bactériennes.

INTRODUCTION

I - INTRODUCTION

1. *Streptococcus thermophilus*, une bactérie lactique largement utilisée en industrie laitière

1.1. Les bactéries lactiques : utilisation en industrie agro-alimentaire

1.1.1. Qu'est-ce qu'une bactérie lactique ?

Le terme « bactéries lactiques » (LAB, pour Lactic Acid Bacteria) sont incluses dans un groupe hétérogène et ne sont pas regroupées entre-elles d'un point de vue phylogénétique mais fonctionnelle (de Roissard, 1986). Toutes les LAB ont un métabolisme fermentaire saccharolytique qui, en utilisant des glucides, produit soit de l'acide lactique (bactéries homolactiques strictes), soit de l'acide lactique et de l'acide acétique (bactéries hétérolactiques facultatives) soit de l'acide lactique, de l'acide acétique ou de l'éthanol et du CO₂ (bactéries hétérolactiques strictes). Certaines espèces ou certaines souches peuvent en plus produire de l'acide formique ou de l'acide succinique. Aucune souche du groupe lactique n'est capable de produire des acides volatils ayant plus de deux atomes de carbone. Ce métabolisme carboné est assez rudimentaire et caractéristique des organismes ancestraux. En effet, ces bactéries seraient apparues il y a près de 3 milliards d'années (Taillez, 2001). D'autre part, les bactéries lactiques sont, pour la plupart, à coloration Gram-positives, généralement immobiles, asporulées et anaérobies aérotoles (Dellaglio *et al.*, 1994).

L'ensemble de ces caractères sont typiques et servent à définir le « cœur » du groupe lactique. Ainsi, le groupe des bactéries lactiques *sensu stricto* est formé de bactéries en forme de coque : *Streptococcus*, *Lactococcus*, *Vagococcus*, *Enterococcus*, *Pediococcus*, *Aerococcus*, *Tetragenococcus*, *Leuconostoc*, *Atopobium*, *Tetragenococcus* et *Oenococcus* ; et en forme de bacille : *Lactobacillus* et *Carnobacterium*. L'appellation « bactérie lactique » est, cependant, souvent étendue à d'autres bactéries qui leur sont apparentées : *Bifidobacterium*, *Micrococcus*, *Brevibacterium* et *Propionibacterium*. Ces 4 genres diffèrent des autres bactéries lactiques par leur capacité à tolérer l'oxygène : anaérobie stricte (*Bifidobacterium*, *Propionibacterium*) ou aérobie (*Micrococcus*, *Brevibacterium*).

1.1.2. Utilisation des bactéries lactiques dans la production de produits alimentaires

Les bactéries lactiques s'avèrent utiles à l'homme et contribuent à la production de nombreux produits alimentaires fermentés comme le fromage, le vin et la choucroute, dans nos régions tempérées, et des produits à base de céréales, de tubercules, de poisson et de viande dans les régions intertropicales (Taillez, 2001). Le tableau 1 présente quelques exemples d'utilisation des bactéries lactiques dans l'alimentation.

Tableau 1 : Quelques exemples de l'utilisation des bactéries lactiques dans l'industrie alimentaire.

Genre	Substrat	Exemples de produits
<i>Bifidobacterium</i>	Lait	Laits fermentés
<i>Lactobacillus</i>	Lait	Yaourts, laits fermentés, kéfirs, fromages
	Viande	Saucissons secs, jambons secs
	Végétaux	Choucroute, olives, « yaourts » au lait de soja
	Céréales	Pain au levain, bières
<i>Lactococcus</i>	Lait	Fromages, kéfirs
<i>Leuconostoc</i>	Végétaux	Choucroute, olives, vin
	Lait	Fromages, kéfirs
<i>Pediococcus</i>	Végétaux	Choucroute
	Viande	Saucisses semi-séchées
<i>Oenococcus</i>	Végétaux	Vin
<i>Streptococcus</i>	Lait	Yaourts, laits fermentés, fromages

Les études scientifiques montrent que derrière les traditions alimentaires, se trouvent des réalités physiologiques, biologiques et nutritionnelles qui légitiment ces coutumes. Ces qualités appartiennent à des souches de bactéries sélectionnées de manière empirique et transmises depuis des siècles. Actuellement, cette sélection se poursuit de façon plus scientifique, à la lumière des données acquises et avec l'aide des connaissances génétiques récentes. Néanmoins, les qualités liées à ces produits exigent le plus souvent le respect de la vitalité des bactéries employées, concept qui se retrouve dans les produits laitiers frais, aliments vivants, à temps de conservation limité (Antoine *et al.*, 1994).

1.2. *Streptococcus thermophilus* présente un intérêt économique important

1.2.1. Les bactéries lactiques à l'origine des laits fermentés

Les bactéries lactiques interviennent dans différentes technologies laitières (fromage frais et affinés, crème et beurre acides, laits fermentés) et leur rôle au niveau de ces fabrications est complexe et diversifié. En effet, elles peuvent intervenir sur : (i) l'acidification qui permet la coagulation du lait et limite les risques de développements microbiens indésirables, (ii) la formation de composants aromatiques qui assure aux produits laitiers les qualités organoleptiques recherchées, (iii) la protéolyse qui confère aux fromages affinés leurs caractères rhéologiques et organoleptiques,

(iv) la production de gaz carbonique qui permet l'ouverture de la pâte fromagère et, (v) la production d'agents épaississants qui améliore l'onctuosité des laits fermentés ou des fromages frais maigres et qui améliore la texture des fromages affinés maigres (Schmidt *et al.*, 1994).

Les laits fermentés se définissent comme étant le produit obtenu par fermentation du lait avec des bactéries lactiques ou des levures. Les laits fermentés sont principalement les yaourts, les fromages mais aussi la crème, le beurre ou encore le kéfir. Le décret n°88-1203 du 30 décembre 1988, relatif aux laits fermentés et au yaourt réserve la dénomination yaourt (ou yoghourt) au « lait fermenté obtenu, selon des usages loyaux et constants, par le développement des seules bactéries lactiques thermophiles spécifiques dites *S. thermophilus* et *Lb. bulgaricus* qui doivent êtreensemencées et se trouver vivantes dans le produit fini à raison d'au moins 10 millions de bactéries par gramme rapportées à la partie lactée ». La dénomination « fromage » est réservée, selon le décret n° 88-1206 du 30 décembre 1988, au « produit fermenté ou non, affiné ou non, obtenu à partir de matières d'origine exclusivement laitière (lait entier, lait partiellement ou totalement écrémé, crème, matière grasse, babeurre), utilisées seules ou en mélange, et coagulées en totalité ou en partie avant égouttage ou après élimination partielle de leur eau ».

1.2.2. Choix des bactéries lactiques en industrie laitière

Dans la plupart des procédés industriels de transformation du lait, les bactéries lactiques présentes dans le lait cru sont détruites lors de la pasteurisation. Elles doivent être réintroduites par des ferments composés de souches préalablement sélectionnées suivant des critères d'aptitudes technologiques et organoleptiques : acidification, production de composants de saveur et d'arôme, dégradation des protéines, résistance aux bactériophages, etc. L'un des défis les plus importants de l'industrie laitière est donc de choisir les genres, les espèces et les souches bactériennes les mieux adaptés aux fonctions qui leur sont imparties dans les diverses technologies de transformation du lait par fermentation (Schmidt *et al.*, 1994). Les premières précautions à prendre sont de choisir les associations de souches en recourant à des méthodes de sélection éprouvées, et d'appliquer des protocoles rigoureux de cultures des ferments. Les bactéries lactiques ont des propriétés (technologiques, sensorielles, antimicrobiennes ou probiotiques) qui sont spécifiques de l'espèce ou de la souche bactérienne. Ceci implique que la mise en évidence par des techniques d'identification et de différenciation soit précise. Cela est d'autant plus important que certaines bactéries lactiques ont un caractère pathogène commensaux ou opportunistes, comme la plupart des streptocoques (Curk *et al.*, 1994).

Concernant les propriétés fonctionnelles d'intérêt pour la fabrication des laits fermentés, deux

groupes se distinguent chez les bactéries lactiques. Le premier regroupe les bactéries dont les propriétés sont essentielles à l'élaboration même du produit, sans lesquelles ce dernier ne pourrait avoir la dénomination de fromages, de yaourt ou de beurre. Les propriétés de ce groupe sont liées à l'acidification, à la protéolyse et à la lipolyse et les bactéries lactiques qui assurent ce rôle sont appelées « starter ». Ce groupe contient les espèces *Lb. lactis*, *Lb. bulgaricus*, *Lb. helveticus*, *L. lactis*, *Leuconostoc sp.* et *S. thermophilus*. Le second groupe comprend les bactéries dont les propriétés permettent soit de donner de la valeur ajoutée ou de la typicité aux produits laitiers concernés soit d'influencer leurs interactions, positives et/ou négatives, avec les autres microorganismes. Ce groupe contient, par exemple, *Propionibacterium freudenreichii* ou *Brevibacterium limens* (Parente et Cogan, 2004).

Les bactéries lactiques les plus fréquemment rencontrées dans les produits laitiers sont les lactobacilles, les lactocoques et l'espèce *S. thermophilus*. Ces bactéries se différencient particulièrement par leur température optimale de croissance et leur propriété acidifiante (consommation des sucres, vitesse, quantité et nature de l'acide lactique produit). Ces différences phénotypiques sont évidemment la clé du choix de leur utilisation pour chacun des types de laits fermentés. Néanmoins, d'autres bactéries lactiques peuplent nos fromages et laits fermentés. Ces bactéries sont directement issues du lait à partir duquel ils sont fabriqués ou sélectionnées etensemencées volontairement pour leur propriétés technologiques comme par exemple des entérocoques, des leuconostocs ou encore *Penicillium roquefortii* (Jamet, 2009).

1.2.3. *Streptococcus thermophilus* est l'une des bactéries les plus utilisées en industrie laitière

1.2.3.1. Les laits fermentés par *Streptococcus thermophilus*

L'utilisation de *S. thermophilus* tend globalement à se généraliser dans un grand nombre de technologies fromagères. En effet, cette espèce présente le grand avantage de produire une activité acidifiante dans un large spectre de température. Sa température optimale de croissance est comprise entre 42 et 44°C, mais elle est capable de se développer, certes plus lentement, entre 19°C (température minimale) et 30°C. A l'opposé, beaucoup de souches de *S. thermophilus* résistent très bien à des températures supérieures à 50°C, avec un maximum autour de 55°C, lui permettant d'être utilisée de manière ubiquitaire dans les fromages à pâte pressée cuite, pour lesquels les températures de cuisson du caillé sont de cet ordre.

Cependant, bien qu'ayant une forte capacité d'acidification du lait, *S. thermophilus* ne permet

INTRODUCTION

quasiment jamais d'abaisser seule le pH du lait en dessous d'un pH de 4,8. En effet, *S. thermophilus* présente une sensibilité importante à pH acide (de l'ordre de 4,6/4,8), impliquant sa lyse rapide au cours de l'acidification du lait. Ainsi, l'utilisation de *S. thermophilus* est réalisée en co-culture dans la fabrication des laits fermentés (Jamet, 2009) : avec *Lb. bulgaricus* pour la fabrication des yaourts (Pette et Lolkema, 1950) et avec *L. lactis* pour la fabrication de certains fromages de type suisse, français et italien comme le gruyère, l'emmental, le grana, le provolone et le gorgonzola (Giraffa *et al.*, 2001).

Pour exemple, en France, environ 1,7 millions de tonnes de fromages et 2,2 millions de tonnes de yaourt sont produits chaque année. La filière du lait affiche un chiffre d'affaires de 25,6 milliards d'euros en 2009, ce qui la classe au deuxième rang des industries agro-alimentaires, juste après la filière de la viande. La consommation mondiale de cellules vivantes de *S. thermophilus* a été estimée à 1 000 milliard de milliard par an (source INRA).

1.2.3.2. Le genre *Streptococcus*

Le genre *Streptococcus* appartient au phylum des firmicutes qui regroupe des bactéries Gram positives possédant un génome dont le pourcentage en bases G et C est bas (compris entre 35,7 et 39,8 %) (Farrow et Collins, 1984 ; Tettelin *et al.*, 2002). Les chromosomes des bactéries du genre *Streptococcus* sont circulaires, de taille avoisinant les 2,2 Mb (Tettelin *et al.*, 2002).

Au sein du genre *Streptococcus*, une analyse comparative des ARNr 16S a permis de distinguer 6 groupes (le groupe pyogénique, le groupe anginosus, le groupe mitis, le groupe salivarius, le groupe bovis et le groupe mutans) englobant 34 espèces (Bentley *et al.*, 1991 ; Kawamura *et al.*, 1995) (figure 1).

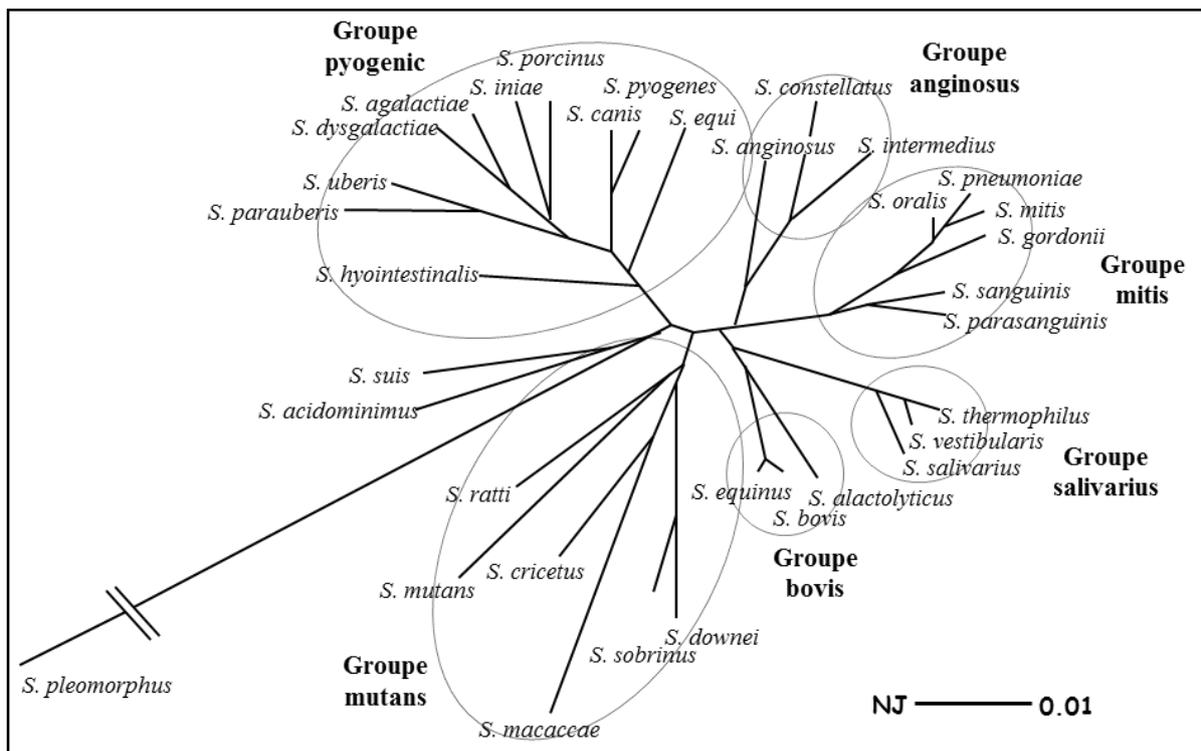


Figure 1 : Relations phylogénétiques établies à partir de l'ARN 16S issu de 34 espèces de *Streptococcus* (d'après Kawamura *et al.*, 1995). Les distances ont été calculées à l'aide la méthode du plus proche voisin (*Neighbour-Joining*, NJ).

Des études prenant en compte d'autres critères génétiques ont permis de mieux comprendre les relations phylogénétiques existant entre ces différentes espèces. Ainsi, une première étude analysant le polymorphisme du gène *sodA* (codant une superoxide dismutase) est en accord avec l'arbre proposé par Kawamura *et al.* (1995). D'autres caractères ont aussi été étudiés comme le polymorphisme du gène *rnpB* codant l'endoribonucléase P (Tapp *et al.*, 2003). Les relations phylogénétiques mises en évidence à l'aide de cette analyse sont en accord avec celles établies précédemment. Cependant, certaines différences ont pu être remarquées telles que l'absence de regroupement de *S. agalactiae* avec les espèces constituant le groupe pyogénique (Poyart *et al.*, 1998) ou le caractère non monophylétique du groupe mitis avec inclusion à l'intérieur du groupe anginosus. Le groupe anginosus (composé des espèces *S. anginosus*, *S. intermedius*, *S. constellatus*), le groupe bovis (composé des espèces *S. equinus*, *S. bovis* et *S. gallolyticus*) et le groupe salivarius (composé des espèces *S. salivarius*, *S. thermophilus* et *S. vestibularis*) constituent des clades spécifiques contenant un nombre restreint d'espèces, ce qui n'est pas le cas des autres groupes.

D'autres classifications non-phylogénétiques existent et se basent sur le pouvoir hémolytique de ces espèces mais aussi sur la présence et la nature des antigènes (établies par Lancefield). Ainsi, les streptocoques sont classés en trois groupes sur leur pouvoir hémolytique : les α -hémolytiques

INTRODUCTION

provoquant une hémolyse incomplète, les β -hémolytiques provoquant une hémolyse complète et les non-hémolytiques. De plus, la classification de Lancefield différencie 12 groupes antigéniques dont un regroupant les streptocoques non-hémolytiques (Lancefield, 1993).

Des analyses *in silico* comparant 26 génomes disponibles du genre *Streptococcus* permettent d'évaluer la taille du core-génome, c'est-à-dire l'ensemble des gènes présents chez tous les streptocoques, à 600 gènes. Le génome accessoire correspond à la partie variable du génome c'est à dire les gènes spécifiques de la souche, de l'espèce ou du genre. Le pan-génome est constitué de l'ensemble des gènes non redondants présents chez un organisme (souche, espèce ou genre) et correspond donc à la somme du core-génome et du génome accessoire. Chez les streptocoques, le pan-génome s'élèverait à environ 6 000 gènes. Le génome accessoire constitue ainsi les 90 % restant du pan-génome. Cela montre que les flux de gènes sont importants au sein de ce genre bactérien (Lefebure et Stanhope, 2007).

L'essentiel des espèces du genre *Streptococcus* regroupe des souches pouvant être pathogènes pour l'homme et/ou les animaux. *S. pyogenes* possède l'antigène A déterminé par Lancefield et est souvent qualifié comme streptocoque de groupe A (GAS). Cette bactérie est responsable de nombreuses maladies chez l'homme telles que les infections de la sphère rhinopharyngée ou de la peau et est impliquée dans des cas de chocs toxiques. L'espèce *S. agalactiae* ou streptocoque de groupe B (GBS), présente l'antigène du groupe B. Cette bactérie est responsable d'infections néonatales (méningites, septicémies). Jusqu'à récemment, les espèces du groupe salivarius étaient considérées comme commensales mais l'implication de *S. vestibularis* et *S. salivarius* dans des cas d'endocardites et de *S. salivarius* dans des cas de méningite a été observée (Cunliffe et Jacob, 1997 ; Partridge, 2000 ; Idigoras *et al.*, 2001). Parmi les espèces connues, seules les espèces *S. thermophilus* et *S. macedonicus* sont non pathogènes (Tsakalidou *et al.*, 1998).

1.2.3.3. L'espèce *Streptococcus thermophilus*

S. thermophilus a été décrite pour la première fois par Orla-Jensen en 1919. Le génome de *S. thermophilus* s'organise en une unité chromosomique unique (à l'exception de quelques souches disposant de petits plasmides) dont la taille est de 1,82 Mb (Roussel *et al.*, 1994). Par ailleurs, *S. thermophilus* pourrait posséder l'un des plus petits génomes des bactéries lactiques (Giliberti *et al.*, 2006).

Le génome de *S. thermophilus* présente 10% de pseudogènes suggérant que celui-ci évolue par perte de fonction (Delorme *et al.*, 2009). La forte proportion de pseudogènes suggère qu'un

INTRODUCTION

changement de mode de vie récent (passage en milieu lait) a entraîné l'inactivation de nombreux gènes (Bolotin *et al.*, 2001 ; Siezen *et al.*, 2005). En effet, l'adaptation à l'environnement du lait aurait résulté en une stabilisation de la structure du génome et contribué à la perte de fonctions non essentielles à la survie de *S. thermophilus* dans le lait (Bolotin *et al.*, 2004). L'inactivation des gènes par des mutations telles que les mutations non-sens, les décalages du cadre de lecture, les recodages ou encore les délétions lors de réarrangements génomiques, a conduit à l'apparition de ces pseudogènes dans le génome de *S. thermophilus* (Bolotin *et al.*, 2004). Ces pseudogènes devraient être à terme délévés ce qui contribuera à maintenir une forte densité de gènes fonctionnels (Leblond-Bourget et Guédon, 2009).

D'autre part, une analyse CGH (Comparative Genome Hybridization) portant sur 47 souches de *S. thermophilus* a permis d'estimer le core-génome de l'espèce *S. thermophilus* à 1 271 gènes (Rasmussen *et al.*, 2008), ce qui amène le génome variable à environ 700 gènes pour les souches entièrement séquencées. Cela révèle une grande quantité de gènes variables qui peuvent résulter de transferts horizontaux récents. En effet, bien que les mécanismes de transferts horizontaux restent peu caractérisés, le génome de *S. thermophilus* présente des séquences d'insertions, des traces d'événements d'intégration/excision de phages et des éléments conjugatifs et intégratifs pouvant être à l'origine de ces pertes de fonctions (Guédon *et al.*, 1995 ; Pavlovic *et al.*, 2004). Des analyses *in silico* basées sur des biais compositionnels (pourcentage en bases G et C, composition en dinucléotides) ont permis de proposer qu'un ensemble de 197 gènes ont été acquis par transfert horizontal dans le génome entièrement séquencé des souches LMG18311, CNRZ1066 et LMD9 de *S. thermophilus* (Liu *et al.*, 2009). Le transfert se serait effectué entre des bactéries lactiques alimentaires comme *L. lactis* et *L. bulgaricus* plutôt qu'entre des espèces bactériennes phylogénétiquement reliées à *S. thermophilus* (Figure 2).

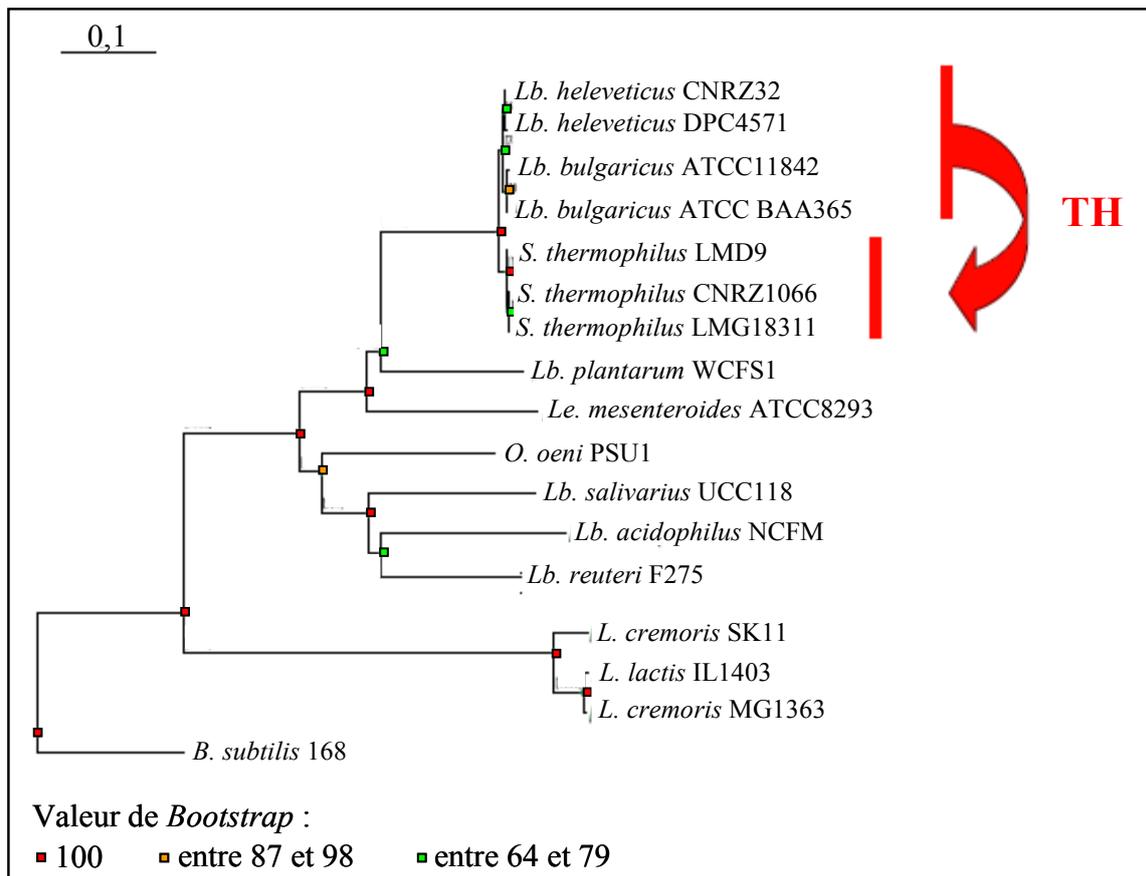


Figure 2 : Arbre phylogénétique du cluster de gènes *cbs-cblB (cglB)-cysE* présent dans les génomes de bactéries lactiques (d'après Liu *et al.*, 2009). L'arbre a été construit à partir de l'alignement de séquences en concatémère des gènes *cbs* avec *cblB* ou *cglB*. Bien que le gène *cysE* soit un pseudogène dans quelques génomes, celui-ci a été utilisé pour l'analyse phylogénétique. Les valeurs de *Bootstrap* ont été mesurées à partir de 1000 répliques. Les gènes *cbs*, *cblB* et *cglB* tronqués de *Lb. acidophilus* NCFM ont été inclus dans l'analyse. Des gènes de *B. subtilis* 168 ont été utilisés comme groupe externe. L'évènement de transfert horizontal (TH) est indiqué par la flèche en rouge.

Pendant plusieurs années, l'espèce *S. thermophilus* était classée comme une sous-espèce de *S. salivarius* (*S. salivarius* subsp. *thermophilus*). Cependant, des analyses récentes de phylogénie par MLST ont clairement démontré que *S. thermophilus* est une espèce à part entière. De plus, la faible variation des séquences du core-génome de cette espèce suggère qu'elle serait d'origine récente (Schleifer *et al.*, 1991 ; Delorme *et al.*, 2009). En effet, la spéciation de *S. thermophilus* aurait pu avoir lieu il y a environ 7000 ans, en parallèle avec le début de l'apparition de procédés de transformation du lait.

S. thermophilus est la bactérie lactique la plus utilisée en industrie laitière après *L. lactis* (Hols *et al.*, 2005). Celle-ci ne présentant aucun danger pour la santé des consommateurs, elle est reconnue comme saine pour la santé ou GRAS (Generally Recognized As Safe) (Bolotin *et al.*, 2004). La seule niche écologique de *S. thermophilus* connue est le lait où cette bactérie se développe de façon optimale à une température de croissance comprise entre 42°C et 44°C. Les génomes de cinq

souches, CNRZ1066 (Bolotin *et al.*, 2004), LMG18311 (Bolotin *et al.*, 2004), LMD-9 (Makarova *et al.*, 2006), ND06 (Sun *et al.*, 2011) et JIM8232 (Delorme *et al.*, 2011), sont actuellement entièrement séquencés et disponibles dans les bases de données.

1.3. Les procédés industriels, des stress pour *Streptococcus thermophilus*

1.3.1. Qu'est-ce qu'un stress ?

Hans Selye, fondateur et directeur de l'Institut de médecine et chirurgie expérimentale de l'Université de Montréal, a été l'inventeur de la « théorie du stress ». En effet, Hans Selye a introduit le terme « stress » au début du vingtième siècle dans le milieu médical. Dans son ensemble, le syndrome de stress (ou syndrome général d'adaptation) évolue selon trois stades successifs : (i) La « réaction d'alarme » pendant laquelle les forces de défense sont mobilisées ; (ii) Le « stade de résistance » qui reflète la complète adaptation à « l'agent stressant » ; (iii) Le « stade d'épuisement » qui suit inexorablement si l'agent stressant est assez puissant et agit assez longtemps, le pouvoir d'adaptation d'un être vivant étant toujours limité (Selye, 1975). Bien que s'appliquant en premier lieu au milieu médical, cette définition s'applique également aux bactéries.

Par ailleurs, des niveaux de sévérité distincts de stress peuvent être différenciés par l'étude de la croissance cellulaire. En effet, selon Storz et Hengge-Aronis (2000), le stress se définit plus simplement par une exposition à une situation environnementale qui conduirait à des dommages cellulaires en l'absence de réponse cellulaire. Ainsi, une exposition à un faible stress n'engendre aucune modification du taux de croissance ; une exposition à un stress d'intensité moyenne entraîne une réduction du taux de croissance ; enfin, une exposition à un stress extrême conduit à l'arrêt de la croissance, les cellules utilisant toutes leurs ressources pour tenter de survivre (Storz et Hengge-Aronis, 2000).

1.3.2. Les procédés industriels produisent du stress

Au cours des processus de production ou de conservation des laits fermentés, des changements des conditions de cultures peuvent survenir (modifications du pH, de la température, de la concentration en oxygène,...) et conduire à un stress (figure 3).

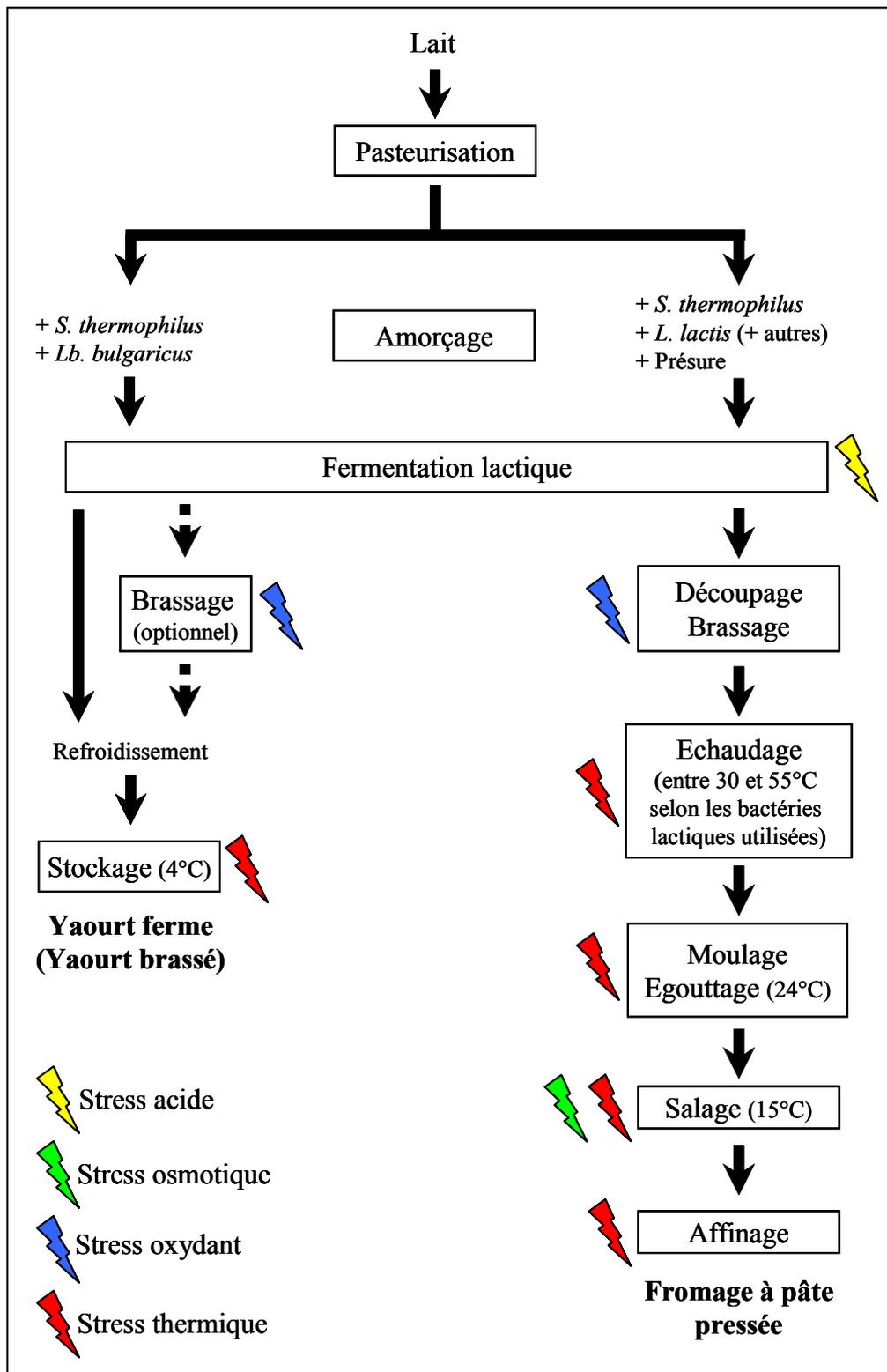


Figure 3 : Schématisation des différentes étapes de fabrication des yaourts et des fromages à pâte pressée.

Les stress générés par les processus de fabrication des yaourts et des fromages, peuvent affecter la survie de certaines souches et influencer l'équilibre des populations bactériennes. Le bouleversement de cet équilibre peut avoir des conséquences majeures, d'une part sur la qualité

INTRODUCTION

gustative du produit, et d'autre part sur le développement d'autres microorganismes, notamment pathogènes, comme les listérias. Par ailleurs, la capacité de *S. thermophilus* à faire face aux stress environnementaux qu'elle rencontre pendant la fabrication des yaourts et des fromages ou pendant sa conservation, en tant que ferment, est essentielle pour sa performance de starter (Zotta *et al.*, 2009).

La température, le pH et les concentrations en substrats et en produits sont celles qui agissent le plus sur la phase de croissance exponentielle des bactéries lactiques (Tayeb *et al.*, 1984 ; Bibal *et al.*, 1988 ; Beal *et al.*, 1989). Les caractéristiques de l'inoculum, l'osmolarité, les concentrations en oxygène et en substances inhibitrices du milieu sont aussi des facteurs importants dont les effets doivent être connus (Beal *et al.*, 1994).

Par exemple, lors de l'élaboration des produits alimentaires où elle intervient, *S. thermophilus* peut être soumise à des étapes de chauffage plus ou moins importantes. La température peut alors être nettement supérieure à celle de la croissance optimale comme lors de l'étape d'échaudage (52-55°C) des fromages à pâte pressée cuite. De plus, *S. thermophilus* peut être soumise à une soudaine baisse de température lors des étapes d'égouttage (environ 24°C), de salage (autour de 15°C) ou de stockage (généralement à 4°C) (Benachour *et al.*, 2009).

Un autre exemple est l'acidification de l'environnement. En effet, dans les produits fermentés, la fermentation permet l'amélioration de la conservation des aliments, en créant un environnement acide défavorisant la survie et la prolifération des contaminants potentiels. La croissance de *S. thermophilus* s'accompagne de la production d'une grande quantité d'acide lactique qui est excrété dans l'environnement. Elle est par conséquent très souvent soumise à des milieux acides (Benachour *et al.*, 2009).

2. Les réseaux de régulation de la réponse au stress présents chez *Streptococcus thermophilus*

2.1. Des réseaux de régulation sont mis en place afin de répondre aux stress

« Les bactéries ont réussi à coloniser toutes les niches écologiques concevables sur cette planète, y compris celles qui, d'un point de vue humain, paraissent les plus hostiles pour la vie. » (Nystrom et Neidhardt, 1994). Cette phrase montre à quel point l'adaptabilité des microorganismes est grande. En effet, dans la plupart des milieux naturels, les bactéries doivent faire face à plusieurs grandes catégories de problèmes : variation des conditions physico-chimiques, compétition pour les surfaces, exposition à des composés nocifs, épuisement de nutriments essentiels,... Depuis plusieurs décennies, des chercheurs concentrent leurs efforts sur l'analyse et la compréhension des phénomènes observés chez des bactéries, y compris les bactéries lactiques, placées dans des conditions de stress. L'identification de gènes et de protéines spécifiques de la réponse aux stress a montré que cette situation déclenche chez les bactéries une reprogrammation qui peut être qualifiée de « différenciation métabolique ». A l'instar des autres bactéries, les réponses adaptatives constituent un moyen usuel de protection contre les stress des bactéries lactiques (Benachour *et al.*, 2009).

Les premières études, sur la réponse d'un organisme vivant à une variation d'un paramètre environnemental, ont concerné la drosophile exposée à une élévation de température. Ce choc thermique a permis d'établir les principales caractéristiques d'une réponse à un stress, à savoir l'acquisition d'une résistance transitoire et le changement d'expression du programme génétique (Ritossa, 1962). Depuis, des études analogues se sont étendues à plusieurs organismes vivants, dont les bactéries, et à de nombreux stress tels que des variations de pH, de pressions osmotiques, de chocs oxydants, de carences en nutriments.

La coordination précise des réactions métaboliques nécessaires à l'adaptation rapide d'un microorganisme à de nouvelles conditions environnementales fait intervenir de nombreux réseaux de régulation. Les premières analyses de génomique comparative suggèrent que les systèmes de régulation varient d'une bactérie à une autre. Par ailleurs, le nombre de gènes intervenant dans les mécanismes de régulation semble être dépendant de la taille du génome du microorganisme analysé (Stover *et al.*, 2000). L'expression génique peut être régulée par diverses protéines comme les facteurs sigma (σ) ou des régulateurs transcriptionnels isolés ou inclus dans des systèmes à deux ou trois composants.

Ces modèles rendent compte de l'induction d'un ensemble de gènes, c'est-à-dire d'un régulon en particulier, qui est fonction de la nature de l'agent stressant. Différents réseaux de régulation

existent et certains agents en induisent plusieurs, ce qui illustre la complexité des mécanismes moléculaires qui gouvernent ces réseaux (Benachour *et al.*, 2009).

2.2. Les facteurs sigma

2.2.1. Qu'est-ce qu'un facteur sigma ?

Afin de faire varier l'expression des gènes, des facteurs protéiques peuvent interagir directement au niveau de l'ARN polymérase en modifiant sa capacité de fixation sur leur région promotrice. En effet, l'étape clé de la transcription des gènes chez les bactéries est celle de l'initiation qui résulte de l'interaction de l'ARN polymérase avec la région promotrice d'un gène et l'activation de l'ARN polymérase permettant l'ouverture de la double hélice de l'ADN.

L'ARN polymérase des bactéries est formée, sous sa forme apoenzyme, de 4 sous-unités ($\beta\beta'\alpha_2$) et est appelée l'ARN polymérase « core ». Les sous-unités α forment un dimère qui se lie aux sous-unités β et β' et à des protéines activatrices de la transcription (Zhang et Darst, 1998). La sous-unité α comporte deux domaines fonctionnels (un domaine C-terminal ou α -CTD et un domaine N-terminal ou α -NTD) séparés par une jonction flexible d'environ 14 nt (Jeon *et al.*, 1997).

Différents facteurs σ (ou sous-unité sigma) peuvent s'associer à l'apoenzyme et constituer l'holoenzyme ($\beta\beta'\alpha_2\sigma$), présentant, selon la nature du facteur σ , une plus ou moins grande spécificité pour la région promotrice des gènes. C'est la spécificité d'interaction enzyme-promoteur qui détermine la force promotrice c'est-à-dire, l'efficacité d'initiation de la transcription (Haldenwang, 1995). Le facteur σ reconnaît de manière spécifique un site cible composé généralement des boîtes dites « -10 » et « -35 » (ou SFBS pour Sigma Factor Binding Sites). Ces annotations correspondent à la position des boîtes par rapport au site du début de la synthèse de l'ARN messager (ARNm) (+1 de transcription) dans les régions promotrices des gènes. Cette reconnaissance spécifique permet une forte liaison entre l'holoenzyme à la séquence d'ADN. L'ARN polymérase recouvre une région de l'ADN qui s'étend de -50 à +20 nucléotides (par rapport au +1 de transcription) et l'ouverture de la double hélice d'ADN est réalisée au quatrième nucléotide composant le motif de la boîte -10. Celle-ci est formée par des appariements de faible stabilité (A et T) qui facilitent la séparation des deux brins d'ADN. Le facteur σ est relargué immédiatement après l'activation de la transcription d'un gène car il n'est pas nécessaire à la poursuite de la transcription et peut être ainsi recruté par une autre ARN polymérase « core ».

Deux grandes classes de facteurs σ peuvent être différenciées : (i) les facteurs σ végétatifs ou

généraux contrôlant essentiellement l'expression de gènes de ménage et assurant ainsi la transcription de base, (ii) les facteurs σ alternatifs contrôlant la transcription des gènes en réponse à des changements environnementaux.

Selon l'espèce bactérienne considérée, le nombre de facteurs σ varie. En effet, ce nombre est de 7 chez *E. coli* (Ishihama, 2000), 18 chez *B. subtilis* (Haldenwang, 1995), 13 chez *Mycobacterium tuberculosis* (Cole *et al.*, 1998) ou 65 chez *Streptomyces coelicolor* (Bentley *et al.*, 2002). Par contre, chez les streptocoques et les lactocoques, le nombre de facteurs σ est plus faible et variable selon l'espèce considérée. En effet, 2 facteurs σ chez *S. mutans*, 4 facteurs σ chez *S. pyogenes* et 3 facteurs σ chez *L. lactis* et chez *S. pneumoniae* ont pu être identifiés.

2.2.2. Les facteurs sigma chez *Streptococcus thermophilus*

L'analyse du génome des souches LMG18311, CNRZ1066, LMD9 et ND03 de *S. thermophilus* révèle la présence de 2 facteurs sigma : le facteur végétatif σ^{70} (encore appelé σ^D) et le facteur alternatif σ^X . Le génome de la souche JIM8232 ne possède que le facteur σ^{70} . Cependant, peu de données concernant le facteur σ^D sont disponibles chez les streptocoques ou les lactocoques et son rôle est pour l'instant inconnu.

Au contraire, le facteur σ^X a été décrit chez la souche LMD9 de *S. thermophilus* et est impliqué dans l'induction de la compétence. En effet, en milieu chimiquement défini, le peptide stimulant la compétence, appelé ComS, est produit, mûri, sécrété puis détecté par la sous-unité A3 du transporteur Ami quand il atteint une concentration seuil dans le milieu extracellulaire. La forme mûrie de ComS, renommé ComS*, est réimportée par le transporteur Ami et interagit avec le régulateur cytoplasmique ComR conduisant à l'activation de ce dernier. Sous forme active, ComR induit la transcription de gènes appelés *comS* et *comX* (figure 4) (Fontaine *et al.*, 2010).

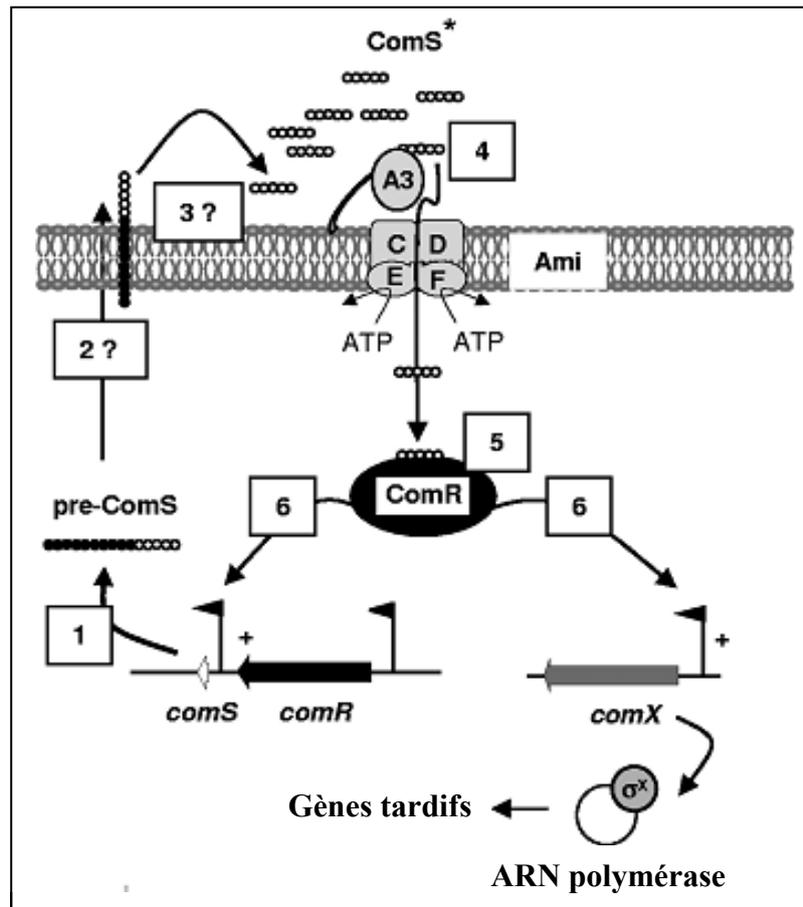


Figure 4 : Modèle d'activation de la compétence par un mécanisme de quorum-sensing (QS) impliquant la protéine ComX de *S. thermophilus* LMD9, en milieu chimiquement défini (d'après Fontaine *et al.*, 2010). (1) En condition de croissance en milieu chimiquement défini, la transcription du gène *comS* (*shp0316*) est déclenchée et le précurseur ComS est produit. (2) Le peptide pré-ComS est sécrété, probablement *via* le système général de sécrétion. (3) Le peptide pré-ComS subit une ou plusieurs étapes de maturation et la forme mature, ComS*, s'accumule dans le milieu extracellulaire. (4) Lorsque le peptide mature ComS* atteint une concentration seuil, celui-ci est reconnu par l'oligopeptide AmiA3 puis internalisé par le transporteur Ami (AmiCDEF). (5) La protéine ComS* intracellulaire interagit avec la protéine ComR ce qui active le régulateur transcriptionnel ComR (Ster_0316). (6) La protéine ComR se fixe sur la ComR-box localisé en amont des gènes *comS* et *comX*. Cela résulte en une boucle d'autoamplification du mécanisme et en l'activation des gènes de compétences tardifs. Les flèches pointées représentent les promoteurs. Les sites de fixation de la protéine ComR sont indiqués par le signe « + ». Les étapes inconnues sont indiquées par le signe « ? ».

Chez *S. pneumoniae*, des études ont montré que le gène *comX* code un facteur sigma, appelé σ^X (Lee et Morrison, 1999 ; Morrison et Lee, 2000). Le facteur σ^X va s'associer transitoirement avec l'ARN polymérase core et ensuite se fixer sur le promoteur des gènes cibles. Ces gènes sont appelés gènes tardifs car les produits de ces gènes interviennent en dernier dans la compétence. Ils sont requis dans la protection du double brin d'ADN ou dans l'intégration du nouveau matériel génétique par recombinaison homologue (Peterson *et al.*, 2004 ; Johnsborg et Havarstein, 2009).

Au sein du genre *Streptococcus*, des études ont montré que la protéine ComX pourrait être

INTRODUCTION

impliquée dans la réponse au stress thermique. En effet, la souche TIGR4 de *S. pneumoniae*, contient deux ORF *comX* appelées, *comX1* et *comX2*. Une analyse de transcriptomique comparant l'expression de gènes de la souche sauvage et d'un double mutant $\Delta comX1 \Delta comX2$ a montré que la délétion des gènes *comX* a un effet inhibiteur sur la transcription de gènes codant des protéines chaperonnes (*grpE*, *dnaK*, *dnaJ*, *groES* et *groEL*, *hrtA*), une protéase (*clpL*) et un répresseur du choc thermique (*hrcA*) (Peterson *et al.*, 2004). Les chaperons et les protéases sont requises pour le repliement, la dégradation et/ou la désagrégation des protéines dénaturées suite à un choc thermique mais également, pour la maturation des protéines (Gross *et al.*, 1996). Ainsi, ces données montrent que le facteur σ^X est impliqué dans la réponse au stress thermique, suggérant que le facteur σ^X de *S. thermophilus* pourrait également être impliqué dans la réponse à ce stress.

En conclusion, peu de facteurs σ , alternatifs ou végétatifs, sont présents au sein des génomes des souches de *S. thermophilus*. Ce faible nombre suggère potentiellement l'existence d'autres systèmes de régulation de l'expression génique. Par ailleurs, aucun homologue du facteur σ^B , intervenant dans la réponse générale aux stress de *B. subtilis*, n'a pu être identifié. Ainsi, la question de la régulation des gènes impliqués dans les mécanismes de défense en condition de stress reste ouverte.

2.3. Les régulateurs transcriptionnels

2.3.1. La diversité des régulateurs transcriptionnels

Les régulateurs transcriptionnels sont des protéines se liant à l'ADN et permettent soit d'activer soit de réprimer la transcription d'un ou plusieurs gène(s) cible(s) en réponse à un changement environnemental. Certains régulateurs peuvent réguler de nombreux gènes, codant des protéines aux fonctions très large alors que d'autres contrôlent un plus petit nombre de gènes, codants des protéines aux fonctions plus spécifiques. Certains régulateurs peuvent être activateurs pour certains gènes et répresseur pour d'autres. Les régulateurs ont leur site de fixation (opérateurs) dans la région promotrice du gène régulé. Lorsqu'un répresseur se fixe sur son opérateur, l'encombrement stérique empêche l'ARN polymérase de se fixer sur le promoteur, le gène ne peut donc plus être transcrit. Un activateur fixé sur son opérateur aura un effet inverse, son interaction avec l'ARN polymérase va augmenter l'affinité de celle-ci pour le promoteur du gène régulé, qui sera transcrit de manière plus efficace.

L'analyse des structures protéiques révèle que la plupart des régulateurs transcriptionnels se fixent sur des séquences d'ADN palindromique ou pseudopalindromique (opérateurs) par

l'intermédiaire de motifs particuliers.

Parmi ces motifs, les plus récurrents sont les motifs HTH (Helix-Turn-Helix), les motifs WHTH (Winged-HTH) et HLH (Helix-Loop-Helix), encore appelé motif en feuillet β (Pabo et Sauer, 1992 ; Phillips et Phillips, 1994). Les motifs HTH classiques contiennent deux hélices α qui forment un angle d'environ 120° , reliées par 4 résidus avec, généralement, une glycine en deuxième position. Une troisième hélice α stabilise le motif en un domaine compact (Gallegos *et al.*, 1997). Les motifs WHTH contiennent un motif HTH suivi par un ou deux coude(s) en épingle à cheveux (Brennan, 1993). Les motifs HLH sont formés de deux feuillets β et une hélice α (Pabo et Sauer, 1992).

Cependant, deux autres types de motifs existent. Le premier correspond aux motifs en « doigt de zinc » qui sont des petits domaines requérant la coordination d'un ou plusieurs atomes de zinc. Ceux-ci ont une structure constituée de deux feuillets β parallèle et d'une hélice α (Klug et Rhodes, 1987). Le second correspond aux motifs appelés « leucine zipper » qui sont constitués de répétitions de leucines (Landschulz *et al.*, 1988).

Les régulateurs transcriptionnels sont des systèmes moléculaires impliqués dans la transduction du signal et sont extrêmement diverses chez les bactéries. Ceux-ci comprennent de simples régulateurs transcriptionnels et des systèmes multi-composants entraînant des cascades de signalisation à des niveaux cruciaux du cycle cellulaire. Ces différents systèmes sont présentés dans la figure 5.

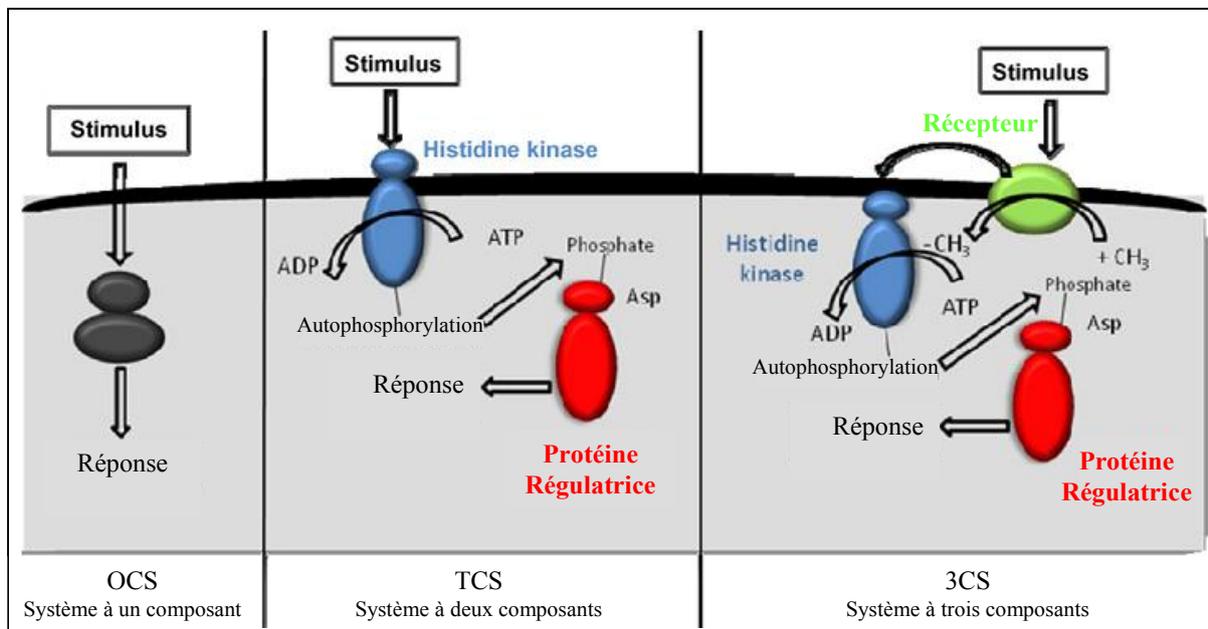


Figure 5 : Schématisation des différents chemins de transduction du signal développés par les bactéries (Marijuan *et al.*, 2010). Le stimulus externe est perçu par un récepteur/transducteur interne (OCS, à gauche), par une histidine kinase transmembranaire, connectée à une protéine régulatrice cytoplasmique (TCS, au milieu) ou par un récepteur indépendant associé à l'histidine kinase (3CS, à droite).

INTRODUCTION

Le premier niveau de complexité correspond aux simples régulateurs transcriptionnels ou OCS (One Component System). En effet, la plupart des protéines qui participent à l'adaptation cellulaire aux changements environnementaux pourrait être incluse dans cette catégorie (Galperin, 2005). Le niveau de complexité suivant implique des régulateurs à deux composants ou TCS (Two Components System) qui sont considérés comme le paradigme central de la transduction du signal chez les procaryotes. Un certain nombre de communications intercellulaires et interspèces se fait par ces systèmes. Récemment, le terme « système à trois composants » (3CS) est apparu dans la littérature (Li *et al.*, 2007 ; Ortiz de Orué Lucana et Groves, 2009). Ces systèmes impliquent un récepteur membranaire, non-kinase, qui active une protéine kinase intracellulaire.

2.3.2. Les systèmes à un composant ou régulateurs OCS

2.3.2.1. Qu'est-ce qu'un système à un composant ?

Les systèmes à un composant, OCS, sont constitués d'une seule protéine capable, à la fois, de percevoir un signal et de réguler la transcription de gènes cibles. Différentes familles d'OCS existent et sont définies par les domaines présents sur la protéine comme la nature du domaine de fixation à l'ADN (Aravind *et al.*, 2005). Le tableau 2 présente quelques exemples de familles de régulateurs transcriptionnels retrouvées sur les génomes de *S. thermophilus* (Ramos *et al.*, 2005).

Tableau 2 : Quelques exemples de familles de régulateurs transcriptionnels OCS présents chez *S. thermophilus*.

HTH : Helix Tour Helix ; CSD : Cold Shock Domain

Famille	Action	Fonctions des gènes régulés	Domaine de fixation	Position
LysR	Activateur/ Répresseur	Métabolisme du carbone et de l'azote	WHTH	N-ter
AraC	Activateur	Métabolisme du carbone, réponse aux stress et pathogénicité	HTH	C-ter
TetR	Répresseur	Biosynthèse d'antibiotiques, transporteur, stress osmotique,...	HTH	C-ter
Rgg	Activateur/ Répresseur	Virulence, métabolisme, réponse aux stress,...	HTH	N-ter
LacI	Répresseur	Utilisation des sources de carbone	WHTH	N-ter
MarR	Activateur/ Répresseur	Résistances aux antibiotiques, aux stress	HTH	Centrale
DeoR	Répresseur	Métabolisme des sucres	WHTH	N-ter
Csp	Activateur	Résistance aux basses températures	CSD	Variable
GntR	Répresseur	Métabolisme général, résistance aux stress	WHTH	N-ter

2.3.2.2. Les régulateurs OCS de *Streptococcus thermophilus*

Plus de 40 ORF codant potentiellement des régulateurs OCS sont présents sur les 5 génomes disponibles de *S. thermophilus*.

2.3.2.2.1. Exemple du régulateur HrcA

Le gène *hrcA* code un répresseur transcriptionnel de gènes de réponse au choc thermique (Schulz et Schumann, 1996). Ce régulateur transcriptionnel possède un motif WHTH en position N-terminale appartenant aux domaines WHTH de la famille GntR (tableau 2). Les protéines de la famille GntR sont caractérisées par deux domaines fonctionnels : (i) un domaine WHTH en position N-terminale nécessaire à leur fixation à l'ADN ; (ii) un domaine C-terminal impliqué dans la fixation effecteur/ligand. La fixation de la protéine à son ligand permet de réguler la fixation de la protéine à l'ADN (Hoskisson et Rigali, 2009).

Ainsi, la protéine HrcA se fixe, *via* son domaine WHTH, sur une séquence opératrice composée de séquences répétées inversées de 9 pb séparées d'un « spacer » de 9 pb (Yuan et Wong, 1995). Cette séquence opératrice, dont le consensus est TTAGCACTC-N9-GAGTGCTAA, est appelé élément CIRCE (pour Controlling Inverted Repeat of Chaperone Expression) (Zuber et Schumann, 1994). Le répresseur HrcA et l'élément CIRCE ont été décrits dans de nombreuses espèces bactériennes suggérant que ce mécanisme de régulation est très répandu chez les bactéries (Hecker *et al.*, 1996). En effet, ce régulateur a été décrit chez différents genres bactériens tels que *Bacillus*, *Streptomyces*, *Lactococcus*, *Bradyrhizobium* et serait restreint aux bactéries Gram positives.

La protéine HrcA est sous forme active lorsqu'elle est en interaction avec la protéine GroEL. En effet, en condition standard, la protéine chaperonne GroEL stabilise la protéine HrcA qui, à son tour, bloque la transcription de ses gènes cibles en se fixant sur l'élément CIRCE (Mogk et Schumann, 1997). En condition de stress, l'accumulation de protéines dénaturées nécessite la dégradation et/ou le repliement de ces protéines dont la protéine GroEL. La protéine HrcA n'est ainsi plus en interaction avec la protéine GroEL et devient inactive ce qui entraîne l'activation de la transcription de ses gènes cibles. La protéine GroEL agirait ainsi comme un thermosenseur de la cellule. Les gènes réprimés par le régulateur HrcA codent des chaperons et forment la classe I des gènes de réponse au choc thermique.

Martirani *et al.* (2001) ont mis en évidence la présence d'une protéine HrcA au sein de la souche ST11 de *S. thermophilus*. Les études réalisées ont permis de montrer que cette protéine est capable de se fixer sur un élément CIRCE. De plus, les auteurs ont montré que la protéine GroEL est requise pour la renaturation de la protéine HrcA (Martirani *et al.*, 2001). Une ORF annotée *hrcA* est présente sur le génome séquencé des 5 souches de *S. thermophilus*. La protéine putative HrcA de

ces 5 souches présente environ 30% d'identité avec la protéine HrcA de *B. subtilis* str.168. Des études ont montré l'implication du gène *hrcA* dans la réponse au stress thermique de *S. thermophilus* Sfi39 (Zotta *et al.*, 2009). En effet, l'inactivation du gène *hrcA* résulte en une augmentation significative de la présence de la protéine GroEL suggérant que la protéine HrcA agisse comme un répresseur de l'expression de la protéine GroEL.

2.3.2.2.2. Exemple du régulateur CtsR

A l'instar du gène *hrcA*, le gène *ctsR* code un répresseur de gènes de réponse au stress thermique appartenant à la famille MarR. La famille MarR se caractérise par la présence d'un motif WHTH et un domaine de dimérisation en position N-terminale (Derre *et al.*, 2000 ; Wilkinson et Grove, 2006). Le régulateur CtsR se fixe *via* son domaine WHTH sur une séquence, appelée CtsR-box, présente en amont de ses gènes cibles. La séquence consensus de la CtsR-box est A/GGTCAAANANA/GGTCAAA (Derre *et al.*, 1999 ; Derre *et al.*, 2000). Le régulateur CtsR et ses gènes cibles ne sont pas présents sur le génome des bactéries Gram positives à haut pourcentage en bases G et C suggérant que ce régulateur soit spécifique des firmicutes (Elsholz *et al.*, 2010).

Le régulateur CtsR agit comme un thermosenseur de la cellule. En effet, l'activité de la protéine CtsR dépend de la température intracellulaire. A 37°C, la protéine CtsR présente une forte affinité pour la CtsR-box, alors qu'à haute température, son affinité est faible. En effet, la protéine CtsR possède une boucle tétraglycine dans son domaine WHTH qui permet au domaine WHTH d'adopter une conformation lui permettant de se lier à l'ADN en fonction de la température (Derre *et al.*, 1999 ; Derre *et al.*, 2000). A 37°C, la conformation du domaine WHTH permet la fixation stable de la protéine CtsR à l'ADN et ses gènes cibles sont donc réprimés. Les gènes réprimés par le répresseur CtsR forment la classe III des gènes de réponse au stress thermique. Ceux-ci regroupent des gènes codant des protéases ATP-dépendants appelés Clp (Nair *et al.*, 2000).

Une ORF annotée *ctsR* est présente sur le génome disponible et entièrement séquencé des 5 souches de *S. thermophilus*. La protéine putative CtsR des souches de *S. thermophilus* présente 50% d'identité avec la protéine CtsR de *B. subtilis* str. 168. Chez un mutant Δ *ctsR* de *S. thermophilus* Sfi39, une augmentation significative du taux de la protéine ClpL est observée en phase stationnaire de croissance par rapport à la souche sauvage. Cela suggère que la protéine CtsR serait bien un répresseur chez *S. thermophilus* (Zotta *et al.*, 2009).

2.3.3. Les systèmes à deux composants ou régulateurs TCS

2.3.3.1. Qu'est-ce qu'un système à deux composants ?

Les systèmes à deux composants, notés TCS, permettent aux bactéries d'assurer des fonctions d'adaptation et de survie en percevant les changements environnementaux et en modulant la transcription d'un certain nombre de gènes en réponse à de nombreux stimuli. Les TCS se différencient des OCS puisqu'ils sont constitués de 2 protéines principales : une protéine senseur kinase membranaire et une protéine régulatrice cytoplasmique recevant le signal (Hoch, 1993a ; Hoch, 1993b ; Hoch, 2002) (figure 6).

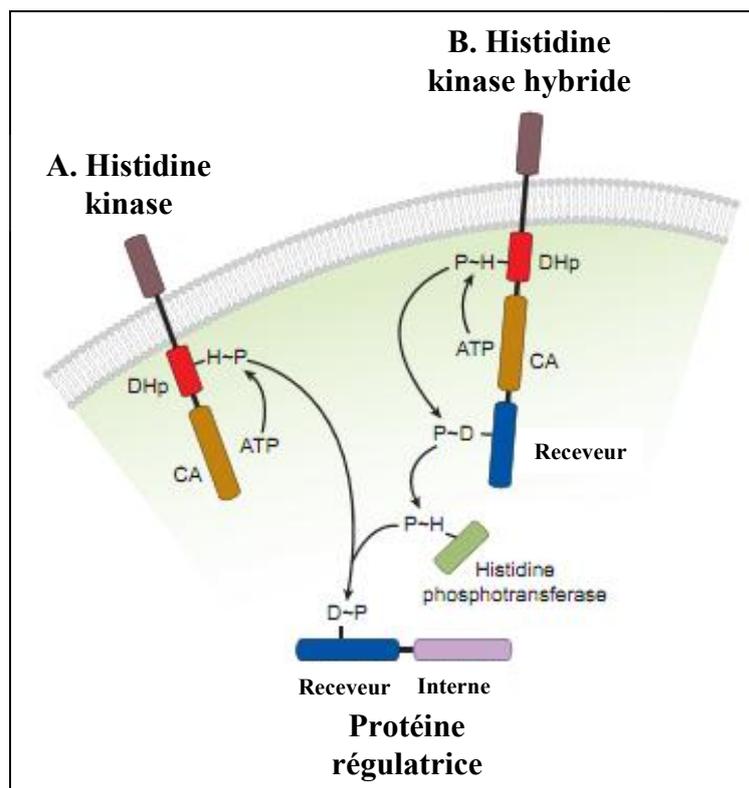


Figure 6 : Représentation schématique d'un système à deux composants et structure des domaines le constituant (d'après Laub et Goulian, 2007). A. L'histidine kinase s'autophosphoryle sur un résidu histidine puis transfère le groupement phosphate à la protéine régulatrice. Les domaines histidine kinase sont très répandus, typiques et ne présentent pas d'homologie substantielle entre eux. Le domaine catalytique et ATPase (CA, en brun) de l'histidine kinase est responsable de la fixation de l'ATP et catalyse l'autophosphorylation d'une histidine conservée à l'aide du domaine histidine phosphotransférase (DHp, en rouge). Le domaine DHp sert de donneur de phosphate à la protéine régulatrice. Certaines histidines kinases assurent une double fonction et peuvent déphosphoryler la protéine régulatrice. Un signal extérieur peut stimuler l'activité phosphatase de la kinase. Les protéines régulatrices contiennent deux domaines : le domaine receveur (en bleu) contient l'aspartate accepteur de phosphate et d'autres acides aminés stimulant la catalyse du phosphotransfère depuis l'histidine kinase ; le domaine interne (en violet), qui est activé par phosphorylation du domaine receveur, permet la fixation à l'ADN. Les TCS sont utilisés par les cellules afin de détecter et de répondre à des stimuli environnementaux en modifiant la transcription. B. Le phosphorelais est une variante du mécanisme des TCS. Le signal environnemental active l'histidine kinase hybride. Le groupe phosphate est transféré de manière intramoléculaire au domaine C-terminal receveur. Une histidine phosphotransférase (HPT, en vert) sert ensuite de navette pour le groupement phosphate entre l'histidine kinase hybride et le domaine interne de la protéine régulatrice.

INTRODUCTION

Les protéines senseur kinases présentent 2 domaines. Le domaine N-terminal permet de détecter le stimulus. La séquence et la structure de ce domaine dépendent du signal à percevoir et varient selon les protéines senseur kinases. Ce domaine est généralement localisé au niveau de la membrane cytoplasmique. Le domaine C-terminal constitue le domaine autokinase, pouvant être subdivisé en un sous-domaine phosphotransférase et un sous-domaine fixant et hydrolysant l'ATP. La protéine senseur kinase répond à des signaux environnementaux par autophosphorylation puis en transmet le signal à la protéine régulatrice. En effet, la reconnaissance du signal par la protéine senseur kinase induit l'hydrolyse d'une molécule d'ATP. Le phosphate ainsi libéré est transféré à un résidu, histidine en général, situé sur le domaine phosphotransférase puis transféré à la protéine régulatrice cytoplasmique (figure 6A). Un mécanisme de phosphorelais faisant intervenir une protéine intermédiaire pour le transfert de phosphate peut également intervenir (figure 6B).

La protéine régulatrice est en général un régulateur transcriptionnel. Cette protéine est composée de deux domaines : un domaine N-terminal régulateur et un domaine C-terminal de liaison à l'ADN (figure 6). Le groupement phosphate provenant de la protéine senseur kinase est transféré à un résidu aspartate du domaine régulateur. La phosphorylation de la protéine régulatrice entraîne un changement de conformation augmentant la capacité de fixation à l'ADN. Ainsi, la fonction de régulateur transcriptionnel est ainsi assurée en permettant d'activer ou de réprimer la transcription de ses gènes cibles (Boucher *et al.*, 1994 ; Rampersaud *et al.*, 1994).

Les sites de liaison de la protéine régulatrice à l'ADN sont situés en amont du promoteur des gènes cibles à réguler. La protéine régulatrice peut ainsi être en contact avec l'ARN polymérase et permettre de modifier la capacité de l'ARN polymérase à initier la transcription (Gralla et Collado-Vides, 1996).

Les TCS ont été décrits dans de nombreux mécanismes tels que le cycle cellulaire bactérien, la communication cellulaire ou la virulence en réponse aux changements de l'environnement (Dziejman et Mekalanos, 1995 ; Hoch, 2002).

2.3.3.2. Les régulateurs TCS de *Streptococcus thermophilus*

L'analyse des génomes entièrement séquencés de *S. thermophilus* révèle la présence d'au moins 6 ORF codant potentiellement des TCS.

L'un des régulateurs TCS bien décrit et présent chez *S. thermophilus* est le système CovS/CovR. Le système CovS/CovR (encore appelé CsrS/CsrR) est un régulateur à deux composants majeur des

INTRODUCTION

streptocoques des groupes A et B. En effet, environ 15% et 7% des gènes sont régulés par les protéines CovS et CovR chez *S. pyogenes* et chez *S. agalactiae*, respectivement (Federle et Scott, 2002 ; Jiang *et al.*, 2008). Le système CovS/CovR est encodé par l'opéron *COVRSX*, où le gène *covR* code la protéine régulatrice et le gène *covS* code la protéine senseur kinase. La fonction codée par le gène *covX* n'est cependant pas élucidée (Lee *et al.*, 2004). La protéine CovR réprime la transcription de l'opéron *COVRSX* en l'absence d'autres facteurs (Federle *et al.*, 1999). Cette autorégulation permet une réponse rapide à des stimuli environnementaux, en adaptant les quantités des protéines CovS/CovR nécessaire à la cellule.

Lorsque la protéine CovR est phosphorylée (CovR-P), celle-ci se multimérise (Miller *et al.*, 2001). La forme multimérique de CovR-P va induire une liaison sur l'ADN entre les boîtes de fixation -35 et -10 de l'ARN polymérase. Un site de fixation de CovR-P, ATTARA, a été proposé comme site consensus chez *S. pyogenes* (Federle et Scott, 2002). La liaison de CovR-P à l'ADN altère la topologie de ce dernier ce qui empêche l'accès de l'ARN polymérase au promoteur (Gusa et Scott, 2005). De ce fait, la protéine CovR agit principalement comme un répresseur de la transcription. De plus, la protéine CovR peut intervenir directement ou indirectement, *via* d'autres régulateurs, sur la transcription de ses gènes cibles (Graham *et al.*, 2002).

Chez les streptocoques pathogènes, le système CovS/CovR est requis dans la transcription de gènes de virulence. Néanmoins, des études ont montré que ce système est également requis dans la médiation de la réponse générale aux stress chez les streptocoques du groupe A qui n'encodent pas de facteur sigma alternatifs (Dalton et Scott, 2004). Ainsi, le système CovS/CovR est requis pour la croissance de *S. pyogenes* en condition de stress thermique, acide ou salin. De plus, la protéine CovR joue un rôle majeur dans la régulation des gènes de réponse au stress thermique chez *S. pyogenes*, *S. agalactiae* et chez *S. mutans*. Les gènes régulés par la protéine CovR codent notamment des Clp protéases qui appartiennent au core génome de *S. thermophilus* (Rasmussen *et al.*, 2008).

Les protéines annotées Rr01 des 5 souches de *S. thermophilus* dont le génome est disponible présentent 83% d'identité avec la protéine CovR de *S. pyogenes* MGAS8232 alors que les protéines annotées Hk01 ne présentent que 47% d'identité avec la protéine CovS. Cela suggère que les gènes codant ces protéines sont présents dans les génomes de *S. thermophilus*.

Chez *S. thermophilus*, seules des données concernant ce TCS chez la souche Sfi39 sont disponibles. L'inactivation du gène *rr01* de *S. thermophilus* Sfi39 entraîne une diminution significative, en phase exponentielle et stationnaire, de la tolérance au choc thermique par rapport à la souche sauvage. De plus, la préadaptation du mutant $\Delta rr01$ à un stress acide et la préadaptation à

INTRODUCTION

un stress thermique, en phase exponentielle, résulte en une légère augmentation et une diminution significative, respectivement, de la tolérance au stress thermique par rapport aux cellules non adaptées en phase exponentielle. Bien que le mécanisme de régulation de la protéine Rr01 chez *S. thermophilus* soit inconnu à ce jour (Zotta *et al.*, 2009), celle-ci pourrait être similaire à celle de *S. pyogenes*. Ainsi, la protéine Rr01 pourrait être une protéine régulatrice activée par phosphorylation *via* la protéine senseur Hk01.

3. Les protéines de type Rgg participent à la réponse aux stress

3.1. Les protéines Rgg constituent une famille de régulateur transcriptionnel

3.1.1. Les gènes *rgg* codent des régulateurs transcriptionnels

Les protéines Rgg (Regulator of glucosyltransferase gene) forment une famille de régulateurs transcriptionnels et leur présence serait restreinte aux bactéries Gram positives des genres *Streptococcus*, *Lactococcus*, *Lactobacillus*, et *Listeria* (Sulavik *et al.*, 1992 ; Lyon *et al.*, 1998 ; Sanders *et al.*, 1998 ; Qi *et al.*, 1999 ; Chaussee *et al.*, 2001 ; Rawlinson *et al.*, 2002 ; Fernandez *et al.*, 2004a ; Samen *et al.*, 2006 ; Ibrahim *et al.*, 2007b ; Zheng *et al.*, 2011). Les gènes membres de la famille *rgg* codent des régulateurs transcriptionnels qui peuvent être positifs ou négatifs (tableau 3).

Tableau 3 : Exemple de protéines Rgg décrites dans la littérature.

Nom de la protéine Rgg	Souche	Action	Référence
Rgg	<i>S. gordonii</i> Challis	Activateur	(Sulavik <i>et al.</i> , 1992)
GadR	<i>L. lactis</i> MG1363	Activateur	(Sanders <i>et al.</i> , 1998)
MutR	<i>S. mutans</i> UA140	Activateur	(Qi <i>et al.</i> , 1999)
RopB	<i>S. pyogenes</i> NZ131	Activateur Répresseur	(Lyon <i>et al.</i> , 1998 ; Chaussee <i>et al.</i> , 2002)
Rgg	<i>S. oralis</i> ATCC10557	Activateur	(Fujiwara <i>et al.</i> , 2000)
RovS	<i>S. agalactiae</i> NEM316	Activateur Répresseur	(Samen <i>et al.</i> , 2006)
Rgg	<i>S. suis</i> SS2	Activateur Répresseur	(Zheng <i>et al.</i> , 2011)
LasX	<i>Lb. sakei</i> L45	Activateur Répresseur	(Rawlinson <i>et al.</i> , 2002)

La première protéine de la famille Rgg a été identifiée chez *S. gordonii* Challis. Chez cet organisme, la protéine Rgg régule positivement le gène *gtfG*, situé à proximité immédiate et en aval du gène *rgg*. La protéine GtfG code une enzyme extracellulaire, la glucosyltransférase G essentielle au développement de la plaque dentaire chez l'homme (Sulavik *et al.*, 1992 ; Sulavik et Clewell, 1996). En absence de cette protéine, aucun transcrit du gène *gtfG* n'est observé.

INTRODUCTION

Chez *S. mutans* UA140, la protéine MutR, appartenant à la famille Rgg, régule positivement la transcription d'un opéron *mut_{AMTFEG}*, situé en amont du gène *mutR* et codant les protéines nécessaires à la biosynthèse d'un antibiotique appelé mutacine II (Qi *et al.*, 1999).

Chez *Lb. sakei* L45, la protéine LasX, appartenant à la famille Rgg, régule positivement l'opéron *las_{AMNTUVPJW}* intervenant dans la biosynthèse, la formation, l'export et l'immunité d'un antibiotique appelé lactocine S. La protéine LasX régule négativement l'opéron *las_{XY}* codant notamment la protéine LasX, situé en aval et en orientation inverse de l'opéron *las_{AMNTUVPJW}* (Rawlinson *et al.*, 2002).

Ces données montrent que les gènes *rgg* contrôlent, positivement ou négativement, la transcription de gènes qui leur sont adjacents (Sulavik *et al.*, 1992 ; Chaussee *et al.*, 1999 ; Samen *et al.*, 2006 ; Pulliainen *et al.*, 2008). De plus, ces gènes cibles sont impliqués dans de nombreux mécanismes tels que la virulence, le métabolisme, la régulation transcriptionnelle ou la résistance aux stress (Sulavik *et al.*, 1992 ; Sanders *et al.*, 1998 ; Fujiwara *et al.*, 2000 ; Chaussee *et al.*, 2001 ; Skaugen *et al.*, 2002 ; Zheng *et al.*, 2011).

3.1.2. Certaines protéines Rgg sont des régulateurs cellulaires globaux

Des analyses ont montré que certains régulateurs, membre de la famille Rgg, interviennent dans la régulation globale de l'expression génique (Chaussee *et al.*, 2002 ; Chaussee *et al.*, 2004 ; Samen *et al.*, 2006 ; Zheng *et al.*, 2011).

Par une approche de puces à ADN, comparant les niveaux de transcrits d'une souche présentant une délétion du gène *ropB* de *S. pyogenes* NZ131 et de la souche sauvage NZ131, Chaussee *et al.* (2002) ont montré que le gène *ropB* contrôlait l'expression de nombreux gènes dont plusieurs facteurs de virulence. De plus, le gène *ropB* agit sur la transcription de plusieurs régulateurs transcriptionnels qui eux même ont pour cibles de nombreux gènes. Par ailleurs, l'utilisation de la protéomique a révélé que la protéine RopB de *S. pyogenes* est un régulateur global qui module la synthèse ou la stabilité de protéines en fonction de la phase de croissance. Ainsi, la protéine RopB de *S. pyogenes* NZ131 est requise en phase stationnaire de croissance pour la régulation de la transcription de gènes de virulences impliqués dans la dissémination de la bactérie (Kreikemeyer *et al.*, 2003 ; Chaussee *et al.*, 2004).

Chez *S. suis* SS2, une expérience de puce à ADN réalisée dans un contexte sauvage et mutant Δ *rgg* a permis de montrer que la transcription d'environ 15% des gènes (soit 345) est sous le contrôle de la protéine Rgg. Parmi ces gènes, 195 sont sur-transcrits et 150 sont sous-transcrits chez le mutant Δ *rgg* par rapport à la souche sauvage, indiquant que la protéine Rgg contrôle un ensemble

de gènes codant des fonctions diverses intervenant dans les métabolismes cellulaires, la division cellulaire, la biogénèse de la paroi ou encore dans la régulation transcriptionnelle. Ces données montrent que la protéine Rgg de *S. suis* SS2 est un régulateur transcriptionnel global (Zheng *et al.*, 2011).

3.2. Les protéines Rgg se lient spécifiquement à l'ADN

3.2.1. Les protéines Rgg possèdent un domaine HTH

Les données de la littérature indiquent que le pourcentage d'identité existant entre les protéines de la famille Rgg est faible. Pour exemple, la protéine RovS de *S. agalactiae* NEM316 présente 25% d'identité avec la protéine Rgg de *S. gordonii* Challis et 21% avec la protéine RopB de *S. pyogenes* NZ131 (Samen *et al.*, 2006). Néanmoins, un alignement de la structure primaire de plusieurs protéines Rgg a permis de mettre en évidence la présence de 3 résidus très conservés (Loughman et Caparon, 2007) (figure 7).

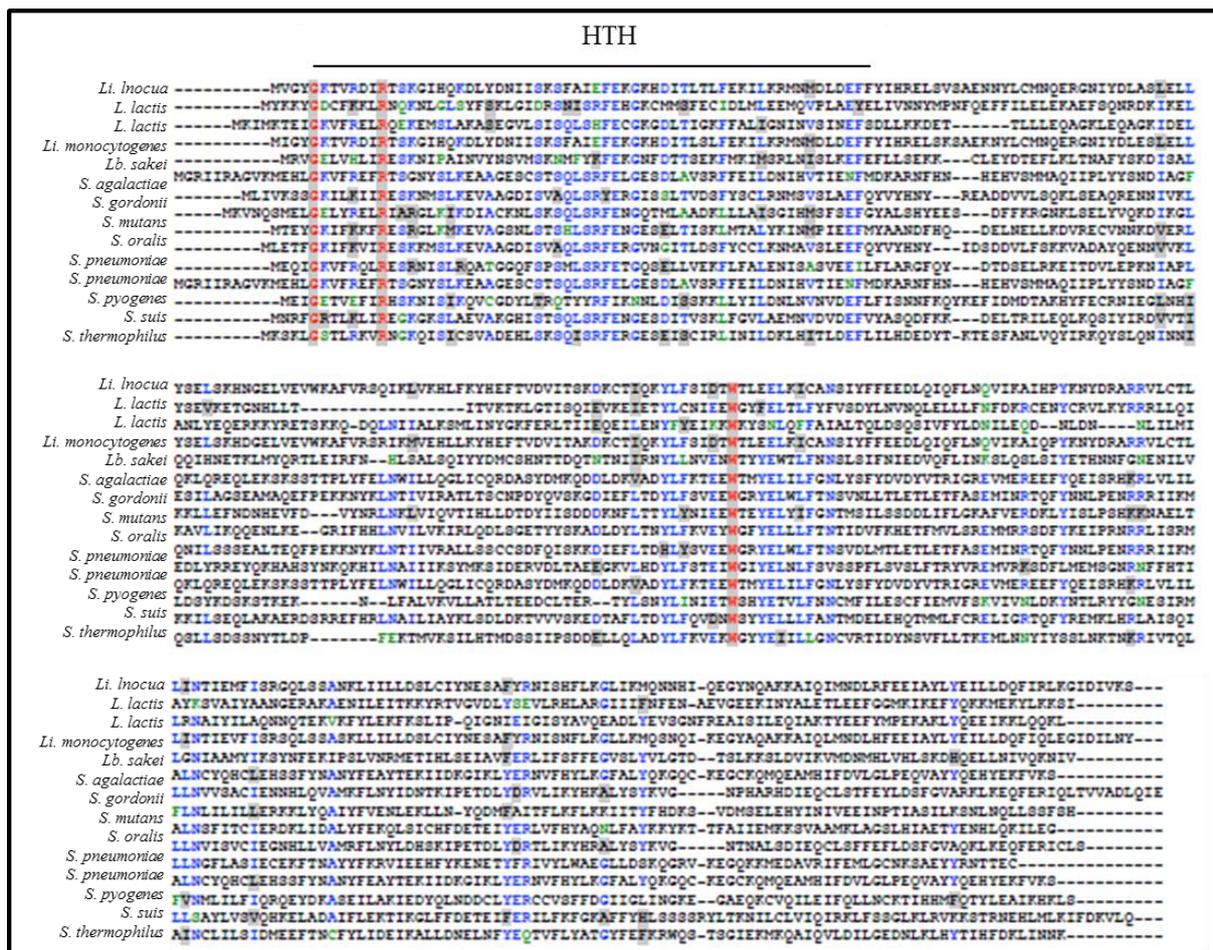


Figure 7 : Alignement multiples des séquences primaires des protéines de la famille Rgg révélant les résidus conservés (d'après Loughman et Caparon, 2007). Les séquences sont alignés par le méthode Clustal W (Thompson *et al.*, 1994). Les couleurs indiquent le degré de conservation des résidus. En rouge sont indiqués les résidus invariables.

INTRODUCTION

Les protéines Rgg possèdent un motif HTH appartenant à la famille XRE (pour Xenobiotic Response Element) en position N-terminale. Aucun autre domaine n'est répertorié dans les banques de données. Deux résidus, G₄ et R₁₁, localisés dans ce domaine HTH et un résidu W₁₄₂, localisé dans la région centrale sont parfaitement conservés dans la séquence primaire des protéines Rgg analysées.

Loughman et Caparon (2007) ont montré que les résidus R₁₁ et W₁₄₂ sont essentiels à la fonction de la protéine RopB de *S. pyogenes*. Le résidu R₁₁ est impliqué dans la liaison ADN-Rgg et le résidu W₁₄₂ pourrait être impliqué dans des liaisons Rgg-protéine (Loughman et Caparon, 2007). Néanmoins, le rôle d'un résidu glycine dans les domaines HTH de la famille XRE n'est pas clairement établi. L'analyse de la structure d'autres domaines HTH-XRE révèle la présence d'un résidu arginine à cette position impliqué dans la stabilité de la liaison ADN-protéine (Wigneshweraraj *et al.*, 2001).

3.2.2. Les protéines Rgg se fixent dans la région promotrice de leur gènes cibles

Certaines protéines Rgg se fixent, *via* leur domaine HTH, sur le promoteur de leurs gènes cibles au niveau de séquences répétées directes (DR) ou inversée (IR) (Skaugen *et al.*, 2002 ; Rawlinson *et al.*, 2005 ; Samen *et al.*, 2006 ; Anbalagan *et al.*, 2011 ; Fleuchot *et al.*, 2011).

Par exemple, le promoteur de l'opéron *las_{A-W}* de *Lb. sakei*, appelé P_{*lasA-W*}, est flanqué de deux heptamères (TTATCCC) en orientation directe (Skaugen *et al.*, 2002). Ces deux heptamères constitueraient les sites de liaison de la protéine LasX *via* son motif HTH. L'étude de la fixation de la protéine Rgg sur le promoteur du gène *lasA* a révélé l'existence d'une troisième répétition imparfaite (TTtCCC) fonctionnelle, localisée dans la région -40 à -70. Cette répétition imparfaite présente au sein du promoteur est requise pour l'activation du promoteur P_{*lasA-W*} alors que les deux autres répétitions, flanquant le promoteur, sont impliquées dans la modulation de l'interaction LasX-promoteur (Rawlinson *et al.*, 2005).

Chez *S. agalactiae* NEM316, des analyses de retard de migration sur gel ont permis de définir un site consensus de fixation de la protéine RovS : **A**W**A**A**W**V**H**T**D**A**W**-N6/7-**W**T**K**W**W**A**M**D**W**A**K** (les nucléotides conservés sont indiqués en gras) (Samen *et al.*, 2006).

Loughman et Caparon (2007) ont fait l'hypothèse que les séquences répétées localisées au sein du promoteur des gènes régulés par les protéines Rgg permettraient la formation d'une structure secondaire de type tige-boucle qui serait reconnue par la protéine Rgg. Celle-ci pourrait s'y fixer *via* son motif HTH. Ce mécanisme de fixation semble être une propriété commune à toutes les

protéines de la famille Rgg décrites jusqu'ici (Sanders *et al.*, 1998 ; Chaussee *et al.*, 1999 ; Qi *et al.*, 1999 ; Fujiwara *et al.*, 2000 ; Rawlinson *et al.*, 2002 ; Samen *et al.*, 2006). Cependant, la séquence DR ou IR serait spécifique de la protéine Rgg et de l'espèce (Rawlinson *et al.*, 2005 ; Samen *et al.*, 2006). Par ailleurs, la variété des fonctions codés par ces gènes suggèrent que les protéines Rgg interagissent avec d'autres protéines pour modifier la transcription de ses gènes cibles (Loughman et Caparon, 2007).

3.3. Implication des régulateurs Rgg dans la réponse aux stress

Plusieurs auteurs ont décrit l'implication de protéines Rgg dans la réponse à différents stress. Par exemple, chez *L. lactis*, la protéine Rgg (appelé GadR) intervient dans la résistance au stress acide. En effet, la résistance au stress acide fait intervenir l'antiport glutamate/ γ -aminobutyrate GadC et le glutamate décarboxylase GadB. L'homéostasie du pH est obtenue par importation du glutamate et sa décarboxylation, à l'aide de la protéine GadB, en γ -aminobutyrate qui sera ensuite exporté. Le taux de transcrite de l'opéron *gad_{BC}* est augmenté à faible pH et en présence de glutamate. Des expériences ont montré l'implication de la protéine GadR dans l'activation de la transcription de l'opéron *gad_{BC}*. La synthèse de des protéines GadBC, contrôlée par GadR, n'est induite qu'en phase exponentielle de croissance et en réponse à un pH acide (Sanders *et al.*, 1998).

Le niveau d'expression des oxydases Nox-1 et AhpC, protéines qui interviennent de façon coordonnée pour détoxifier les cellules de *Streptococcus* face à un stress oxydant, est aussi affecté chez le mutant $\Delta ropB$ de *S. pyogenes* NZ131. Nox-1 est une NADH oxydase qui catalyse la réduction de l' O_2 en H_2O_2 . L' H_2O_2 ainsi formé est réduit en H_2O par la peroxydase AhpC. Chez *S. pyogenes*, les protéines AhpC et Nox1 sont plus abondantes chez le mutant présentant une inactivation du gène *ropB* que chez la souche sauvage (Chaussee *et al.*, 2004). Par ailleurs, chez *S. pneumoniae* et chez *S. thermophilus*, des études ont montré l'implication de deux gènes *rgg* distincts dans la réponse au stress oxydant (Bortoni *et al.*, 2009) (Fernandez *et al.*, 2004a).

Chez *S. pyogenes* NZ131, l'implication du gène *ropB* dans la réponse aux stress thermique a été démontrée (Chaussee *et al.*, 2004 ; Pulliainen *et al.*, 2008). L'inactivation de ce gène affecte l'abondance des protéines associées aux stress thermique (les protéases ClpL, ClpP et ClpE. Ces protéines sont exprimées pendant les phases de croissance exponentielle et stationnaire chez le mutant mais pas chez la souche sauvage.

Enfin, d'autres exemples d'implication des protéines Rgg dans la réponse aux changements des conditions environnementales sont décrits chez d'autres espèces. En effet, *S. mutans* UA149 induit

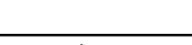
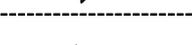
la synthèse de la mutacine quand les ressources nutritives sont épuisées, améliorant ainsi sa compétitivité vis-à-vis des autres espèces présentes dans son environnement. Ainsi, l'ensemble de ces exemples suggèrent que l'une des fonctions des protéines Rgg serait de réguler des gènes intervenant dans la réponse aux conditions environnementales et notamment en réponse à différents stress.

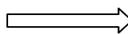
3.4. Implication du régulateur Rgg₁₃₅₈ de *S. thermophilus* LMD9 dans un mécanisme de quorum sensing

3.4.1. Mise en évidence de petites ORF non annotées dans les banques de données

Un certain nombre de petites ORF, non annotées dans les bases de données et présentes dans les génomes séquencés des souches LMG18311, CNRZ1066 et LMD9 de *S. thermophilus*, a été mis en évidence par des analyses *in silico* (Ibrahim *et al.*, 2007a). Plusieurs de ces ORF sont génétiquement associés à des gènes *rgg* de *S. thermophilus* (tableau 4).

Tableau 4 : Présentation des différents gènes *rgg* de *S. thermophilus* associés à des ORF *shp* et/ou *pep* (d'après Ibrahim *et al.* (2007a) et Fleuchot *et al.* (2011)).

Groupe Shp	Particularité de la séquence Shp	Gènes <i>rgg</i>	Environnement génétique
Shp de groupe I	Séquence atypique	<i>stu0182</i> <i>str0182</i>	←  →
	Un résidu aspartate conservé	<i>ster_1299</i>	 ←
		<i>stu1044</i> <i>str1044</i>	←  →
Shp de groupe II	Un résidu glutamate conservé	<i>ster_1358</i>	←  →
		<i>ster1922</i> <i>stu1950</i> <i>str1950</i>	←  →
		<i>ster_0934</i>	←  →  →
		<i>stu0917</i> <i>str0917</i>	←  →  →
		<i>ster_1530</i> <i>stu1572</i> <i>str1572</i>	←  →

← *shp*  *rgg*  inconnu  *pep*

ster : souche LMD9 ; *stu* : souche LMG18311 ; *str* : souche CNRZ1066

INTRODUCTION

Parmi les peptides codés par ces petites ORF, une nouvelle classe de peptides hydrophobes a été mise en évidence, les peptides Shp (pour Short hydrophobic peptide). Les gènes codant des peptides Shp sont associés de façon récurrente aux gènes *rgg*. En effet, au sein du génome des souches LMG18311 et CNRZ1066 de *S. thermophilus*, 5 gènes codant des peptides de type Shp ont été identifiés en amont de 5 gènes *rgg*. Dans le génome de la souche LMD9, 5 copies *rgg* sont présentes et toutes sont associées à une ORF *shp*. Toutes les ORF *shp* sont en orientation inverse à l'ORF *rgg* associée. Les peptides Shp ont été regroupés en deux classes en fonction de la présence d'un aspartate (groupe I) ou d'un glutamate (groupe II) (Ibrahim *et al.*, 2007a). Une étude de phylogénie récente montre que les ORF *shp* associées aux gènes *stu0182* et *str0182* sont apparentées au groupe I malgré leur séquence atypique sans résidu aspartate ou résidu glutamate (Fleuchot *et al.*, 2011).

D'autres petites ORF sont également présentes dans l'environnement génétique de certains gènes *rgg* associés à des ORF *shp*. En effet, 4 des gènes *rgg* associés à une ORF *shp* sont associés à une ou deux ORF *pep*. Néanmoins, les petites ORF *pep* codent des petits peptides (moins de 60 AA) aux fonctions inconnues. En effet, aucun homologue n'est retrouvé dans les banques de données (Ibrahim *et al.*, 2007a).

3.4.2. Mécanisme de quorum-sensing faisant intervenir les protéines Shp₁₃₅₈ et Rgg₁₃₅₈

Des auteurs ont mis en évidence que la présence d'un petit peptide, Pep_{1357c}, de *S. thermophilus* LMD9 nécessitait l'expression des gènes *shp₁₃₅₈* et *rgg₁₃₅₈* mais également un gène appelé *eep*, codant un transporteur de peptide. Ces données ont conduits les auteurs à émettre l'hypothèse d'un mécanisme de quorum-sensing (QS) (Ibrahim *et al.*, 2007b), confirmé par des études récentes (Fleuchot *et al.*, 2011).

Le QS est un mécanisme de communication cellulaire utilisé par les bactéries pour contrôler l'expression de gènes *via* la sécrétion de molécules appelée auto-inducteurs ou phéromones. Premièrement, la phéromone est exportée dans le milieu extracellulaire puis, lorsqu'elle atteint une concentration seuil, elle est détectée, soit par des protéines senseurs, soit par la réimportation de la phéromone dans le milieu intracellulaire *via* un transporteur appelé Ami, membre de la superfamille ubiquitaire des ABC transporteurs (Linton et Higgins, 2007). La détection de la phéromone conduit les cellules à moduler l'expression génique, de façon coordonnée, dans l'ensemble de la population bactérienne (Bassler, 2002 ; Waters et Bassler, 2005 ; Antunes *et al.*, 2010).

Le modèle de QS proposé par Fleuchot *et al.* (2011) est présenté dans la figure 8.

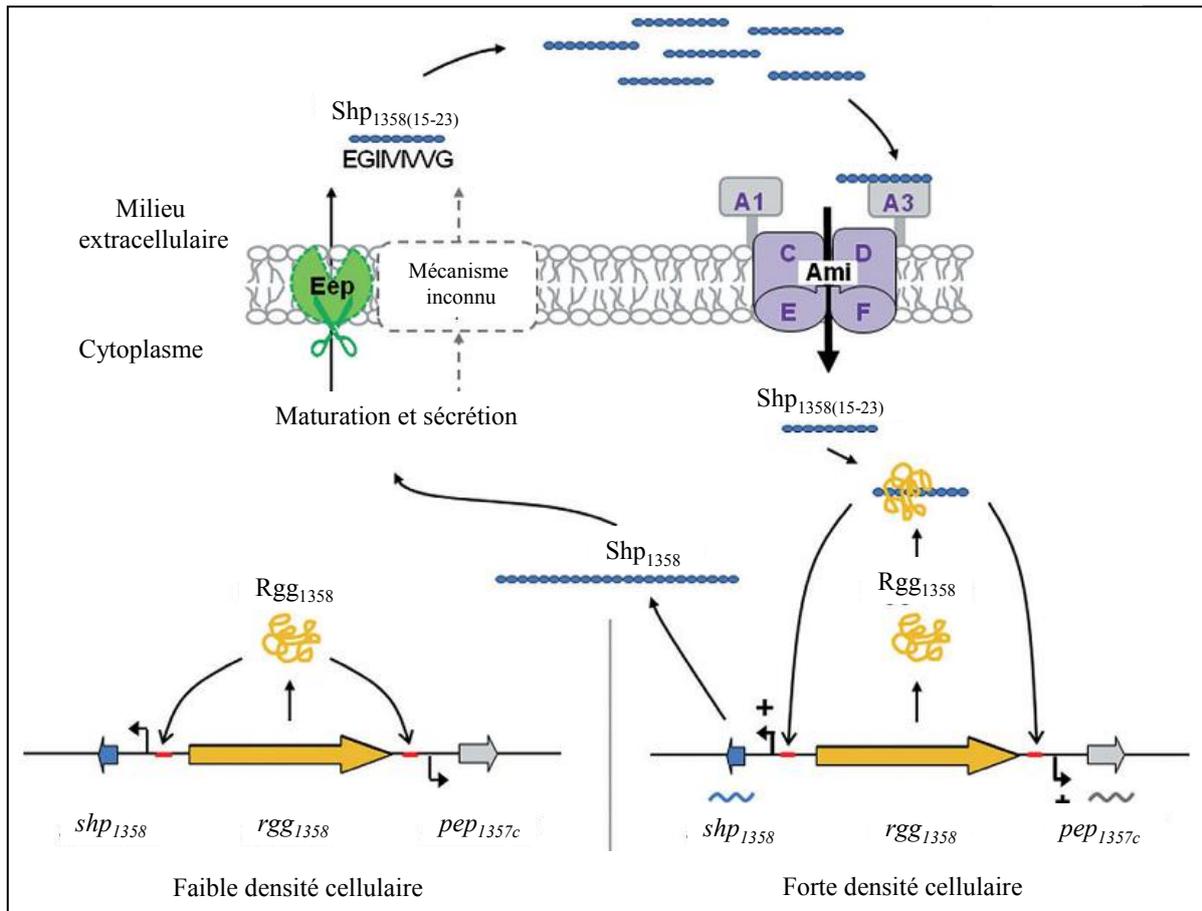


Figure 8 : Représentation schématique du mécanisme de quorum-sensing Rgg₁₃₅₈/Shp₁₃₅₈ de *S. thermophilus* LMD9 (d'après Fleuchot *et al.*, 2011). Le signal de QS est codé par le gène *shp*₁₃₅₈. La phéromone, un petit peptide hydrophobe appelé Shp₁₃₅₈₍₁₅₋₂₃₎ (EGIIVVVG), est produite par le clivage en C-terminal du peptide précurseur de 23 AA, un processus qui fait intervenir l'endopeptidase Eep localisée dans la membrane. A forte densité cellulaire, le peptide Shp₁₃₅₈₍₁₅₋₂₃₎ secrété est détecté par la lipoprotéine AmiA3 et réimporté dans la cellule par le transporteur AmiCDEF. Le peptide Shp₁₃₅₈₍₁₅₋₂₃₎ intracellulaire interagit ensuite avec la protéine Rgg₁₃₅₈. L'effet de l'interaction est inconnu mais la fixation du complexe Shp₁₃₅₈₍₁₅₋₂₃₎/Rgg₁₃₅₈ sur le promoteur de deux gènes cibles, *shp*₁₃₅₈ et *pep*_{1357c}, induit l'activation de leur transcription. A faible densité cellulaire, Fleuchot *et al.*, font l'hypothèse que Shp₁₃₅₈₍₁₅₋₂₃₎ est présent en quantité négligeable et que la protéine Rgg₁₃₅₈ se fixe seule sur la région promotrice des gènes *shp*₁₃₅₈ et *pep*_{1357c} dans un état inactif ne permettant pas leur transcription.

Le peptide Shp₁₃₅₈, du groupe II des Shp, est la phéromone du système de QS de *S. thermophilus* LMD9. Bien que la nature de son interaction avec la protéine Rgg₁₃₅₈ n'ait pas été déterminée, les auteurs font l'hypothèse qu'elle pourrait moduler l'activité de la protéine Rgg₁₃₅₈. Cette dernière induit, ensuite, la transcription des gènes *shp*₁₃₅₈ et *pep*_{1357c}. La fonction du peptide Pep_{1357c} reste cependant inconnue.

Ces données indiquent donc que des gènes *rgg* associés à des gènes *shp* et/ou *pep* pourraient être impliqués dans des mécanismes de QS.

4. Etat de l'art

Nathalie Leblond-Bourget et son équipe s'intéresse à l'adaptation de *S. thermophilus* aux changements de son environnement et notamment à la réponse au stress oxydant. Les premières études réalisées au laboratoire ont permis de montrer l'existence d'un système de défense inductible au H₂O₂ chez *S. thermophilus* CNRZ368 (Thibessard *et al.*, 2001). Afin d'identifier des gènes impliqués dans la réponse au stress oxydant de *S. thermophilus* CNRZ368, une banque de mutants a été obtenue par mutagenèse insertionnelle aléatoire à l'aide du plasmide pGh9 : ISS1 (Thibessard *et al.*, 2002). Des mutants ont ensuite été sélectionnés sur leur capacité de résistance/sensibilité en réponse au stress oxydant (Fernandez *et al.*, 2004a ; Thibessard *et al.*, 2004). Parmi eux, deux mutants appelé OsrA et OsrC (pour Oxidative stress resistance) étaient plus résistants que la souche sauvage à un stress oxydant et l'analyse des loci interrompus a montré qu'ils codaient potentiellement des protéines de la famille Rgg (Fernandez *et al.*, 2004a ; Fernandez, 2004b).

Au cours de sa thèse, Annabelle Fernandez a mis en évidence que les souches de *S. thermophilus* présentaient un grand nombre d'ORF *rgg*. En effet, l'analyse du génome des souches LMD9 et LMG18311 a permis d'identifier 9 copies différentes de gènes *rgg*. La recherche de chacune des copies a été effectuée au sein de 10 souches de *S. thermophilus* par Southern blot (Fernandez, 2004b) (tableau 5).

Tableau 5 : Polymorphisme des gènes *rgg* au sein de 10 souches de *S. thermophilus* (d'après Fernandez *et al.*, 2004b). Les souches de *S. thermophilus* analysées sont indiquées en gras.

+ : présence ; - : absence

	A054	ATCC 19258	CNRZ 302	CNRZ 308	CNRZ 368	CNRZ 445	CNRZ 702	CNRZ 1402	IP 6757	LMG 18311
<i>rggA</i>	+	+	-	+	+	+	+	-	+	+
<i>rggB</i>	+	+	+	+	+	+	+	+	+	+
<i>rggC</i>	+	+	+	+	+	+	+	+	+	+
<i>rggD</i>	-	-	-	-	+	-	+	-	+	+
<i>rggE</i>	-	-	-	-	-	-	-	-	+	+
<i>rggF</i>	+	-	+	+	+	+	+	+	+	+
<i>rggG</i>	+	-	-	-	+	+	+	+	+	+
<i>rggH</i>	-	-	+	-	-	-	+	+	+	-
<i>rggI</i>	-	-	-	-	-	-	+	+	+	-

Ainsi, cette étude montre qu'au sein des différentes souches analysées, entre 3 et 9 copies pourraient être présentes sur le génome de *S. thermophilus*. Alors que les streptocoques pathogènes présentes 2 ou 3 copies de gènes *rgg* (Sulavik *et al.*, 1992 ; Samen *et al.*, 2006), *S. thermophilus* peut présenter un nombre plus important de copies (Fernandez, 2004b ; Ibrahim *et al.*, 2007a). De

INTRODUCTION

plus, les régulateurs Rgg sont décrits comme des régulateurs majeurs de la virulence chez les streptocoques pathogènes (Chaussee *et al.*, 2004 ; Samen *et al.*, 2006 ; Ikebe *et al.*, 2010). Ces données posent la question du rôle des gènes *rgg* chez *S. thermophilus*. Sont-ils tous impliqués dans la régulation de la transcription de gènes de réponse aux stress ? Sont-ils impliqués dans des réseaux de régulation similaires ou différents ? Est-ce que toutes ces copies codent des régulateurs transcriptionnels globaux ou codent-ils des régulateurs transcriptionnels plus spécifiques ?

L'objectif de cette thèse est double. Le premier objectif a pour but de mieux caractériser les régulateurs de la famille Rgg. En effet, quelle est la répartition de ces gènes dans les génomes bactériens ? Quelle est la parenté phylogénétique entre ces gènes ? Qu'est-ce qui caractérise l'appartenance à cette famille ?

Le second objectif est de comprendre l'implication des gènes *rgg* dans l'adaptation de *S. thermophilus* à son environnement. En effet, les génomes de *S. thermophilus* se caractérisent par un grand nombre de gènes appartenant à la famille *rgg*. Quel est le rôle biologique de ces gènes ? Est-ce que toutes ces copies sont impliquées dans la réponse aux stress ?

RESULTATS

II – RESULTATS

1. Caractérisation *in silico* des protéines appartenant à la famille Rgg

Une partie de ces résultats fera l'objet d'une publication.

1.1. La majorité des protéines Rgg sont présentes au sein de la famille des *Streptococcaceae* et des genres *Lactobacillus* et *Listeria*

Les données de la littérature décrivent l'existence de plusieurs protéines Rgg au sein des génomes de bactéries des genres *Streptococcus*, *Lactococcus* et *Lactobacillus*. Afin de déterminer si la présence des gènes *rgg* est restreinte à ces genres bactériens, une recherche exhaustive de ces gènes a été effectuée par une analyse PSI-BLAST (Position-Specific Iterated Basic Local Alignment Search Tool). Cette recherche a été réalisée sur l'ensemble des séquences procaryotes disponibles, utilisant la séquence primaire en acide aminés des protéines Rgg décrites dans la littérature. Cette méthode permet de mettre en évidence des séquences significativement proches des séquences sélectionnées ($e\text{-value} < e^{-2}$). Plusieurs itérations sont réalisées afin de prendre en compte les nouvelles séquences mises en évidence par l'itération précédente, et ce, jusqu'à ce qu'aucune nouvelle séquence ne soit identifiée. Les résultats de cette analyse sont présentés dans le tableau 6.

Tableau 6 : Nombre de copies Rgg mise en évidence par l'analyse de PSI-BLAST au sein de différentes familles de la division des Firmicutes et des Actinobacteria. La recherche a été réalisée sur l'ensemble des séquences procaryotes disponibles dans les bases de données. Seuls les genres bactériens présentant au moins une souche dont le génome contient une (ou plusieurs) copies *rgg* sont indiqués dans ce tableau.

Division	Classe	Ordre	Famille	Genre	Nombre de copies
Firmicutes	Bacilli	Lactobacillales	<i>Aerococcaceae</i>	<i>Aerococcus</i>	1
			<i>Carnobacteriaceae</i>	<i>Carnobacterium</i>	1
				<i>Granulicatella</i>	1
			<i>Enterococcaceae</i>	<i>Enterococcus</i>	28
			<i>Lactobacillaceae</i>	<i>Lactobacillus</i>	109
				<i>Pediococcus</i>	1
			<i>Leuconostocaceae</i>	<i>Leuconostoc</i>	2
				<i>Oenococcus</i>	4
				<i>Weissella</i>	1
			<i>Streptococcaceae</i>	<i>Streptococcus</i>	441
		<i>Lactococcus</i>		26	
		Bacillales	<i>Bacillaceae</i>	<i>Bacillus</i>	1
<i>Listeriaceae</i>	<i>Listeria</i>		112		
<i>Staphylococcaceae</i>	<i>Staphylococcus</i>		2		
Actinobacteria	Actinobacteridale	Bifidobacteriales	<i>Bifidobacteriaceae</i>	<i>Parascardovia</i>	2
				<i>Scardovia</i>	1
				Somme	733

RESULTATS

Par l'analyse de PSI-BLAST, 733 copies Rgg putatives ont été mises en évidence, qu'elles soient identiques ou différentes. Lorsque des copies identiques sont retrouvées au sein de plusieurs génomes, elles sont comptabilisées autant de fois que de génomes les possèdent. A l'exception de 3 copies présentes chez des Bifidobactéries, toutes les copies sont retrouvées au sein de la division des Firmicutes et plus particulièrement au sein de la classe des Bacilli et du genre *Listeria*.

La majorité des copies *rgg*, soit plus de la moitié, est retrouvée au sein du genre *Streptococcus* indiquant une forte représentation au sein de ce genre bactérien. Un grand nombre de copies Rgg est également présent au sein des genres *Listeria* et *Lactobacillus* (environ 15% des copies Rgg pour chacun de ces deux genres). Les genres *Enterococcus* et *Lactococcus* possèdent plusieurs copies Rgg (environ 3% des copies) alors que seulement quelques copies (1 à 4 copie(s)) sont présentes au sein des autres genres bactériens présentés dans le tableau. Enfin, aucune copie Rgg n'est retrouvée dans d'autres genres bactériens.

Ainsi, la présence de copies Rgg pourraient être restreinte aux génomes des genres *Streptococcus*, *Listeria*, *Lactobacillus*, *Enterococcus* et *Lactococcus*. Cependant, l'analyse peut être biaisée par le nombre de génomes entièrement séquencés. En effet, le faible nombre de copies Rgg observé pour certains genres bactériens pourrait s'expliquer par un faible nombre de génomes disponibles. Ainsi, afin de caractériser plus finement la distribution des copies *rgg* chez les procaryotes, les génomes disponibles, au moment où le PSI-BLAST a été réalisé (16 mars 2011), ont également été dénombrés et sont indiqués dans le tableau 7.

Tableau 7 : Nombre de copies Rgg mises en évidence par analyse PSI-BLAST et répartition de ces copies au sein de genres bactériens pour lesquels des génomes ont été séquencés. Les genres bactériens ne possédant aucune copie *rgg* au sein de leur génome ne sont pas mentionnés. En noir sont indiqués les genres bactériens dont la probabilité de trouver des copies *rgg* est forte; en gris, ceux qui en possèdent peu et, en rouge, ceux dont peu de génomes sont disponibles. Les copies identiques sont comptabilisées autant de fois que de génomes les possèdent.

Genre bactérien	Nombre de génomes séquencés	Nombre de génomes séquencés présentant une (des) copie(s) <i>rgg</i>	Nombre de copies <i>rgg</i> identifiées
<i>Streptococcus</i>	65	57	194
<i>Listeria</i>	10	7	27
<i>Lactococcus</i>	5	5	26
<i>Lactobacillus</i>	30	17	38
<i>Staphylococcus</i>	31	2	2
<i>Bacillus</i>	39	1	1
<i>Leuconostoc</i>	4	1	2
<i>Oenococcus</i>	1	1	1
<i>Pediococcus</i>	1	1	1
Somme		92	292

Parmi l'ensemble des génomes séquencés au moment de l'analyse, seuls 92 d'entre eux possédaient au moins une copie *rgg*, ce qui implique que cette famille de gènes n'est pas ubiquitaire dans les génomes bactériens.

Les protéines *rgg* sont très répandues au sein des génomes appartenant aux genres *Streptococcus*, *Lactococcus*, *Listeria*, et *Lactobacillus*. En effet, 100% des génomes séquencés de *Lactococcus* possèdent au moins une copie *rgg*, contre 88% pour *Streptococcus*, 70% pour *Listeria* et 57% pour *Lactobacillus*. Au contraire, les copies *rgg* sont peu répandues dans les génomes des genres *Bacillus*, et *Staphylococcus* (3%, et 7%, respectivement).

La faible proportion de génomes disponibles appartenant aux genres *Leuconostoc*, *Oenococcus*, *Pediococcus* et *Enterococcus* ne permet pas de conclure quant à la faible ou forte répartition des copies *rgg* au sein de ces genres bactériens. C'est le cas également des genres *Aerococcus*, *Carnobacterium*, *Granulicatella*, *Weissella*, *Parascardovia* et *Scardovia* dont aucun génome n'a, pour l'instant, été séquencé.

Le tableau 7 indique que les genres *Streptococcus*, *Lactococcus*, *Listeria*, et *Lactobacillus* sont particulièrement riches en copies *Rgg*. Ainsi, cela semblait intéressant de déterminer si ces copies étaient présentes au sein de toutes les espèces de ces genres ou restreintes à certaines d'entre-elles. La figure 9 répertorie le nombre de copies *rgg* retrouvées au sein des différentes espèces des 4 genres bactériens qui en possèdent le plus.

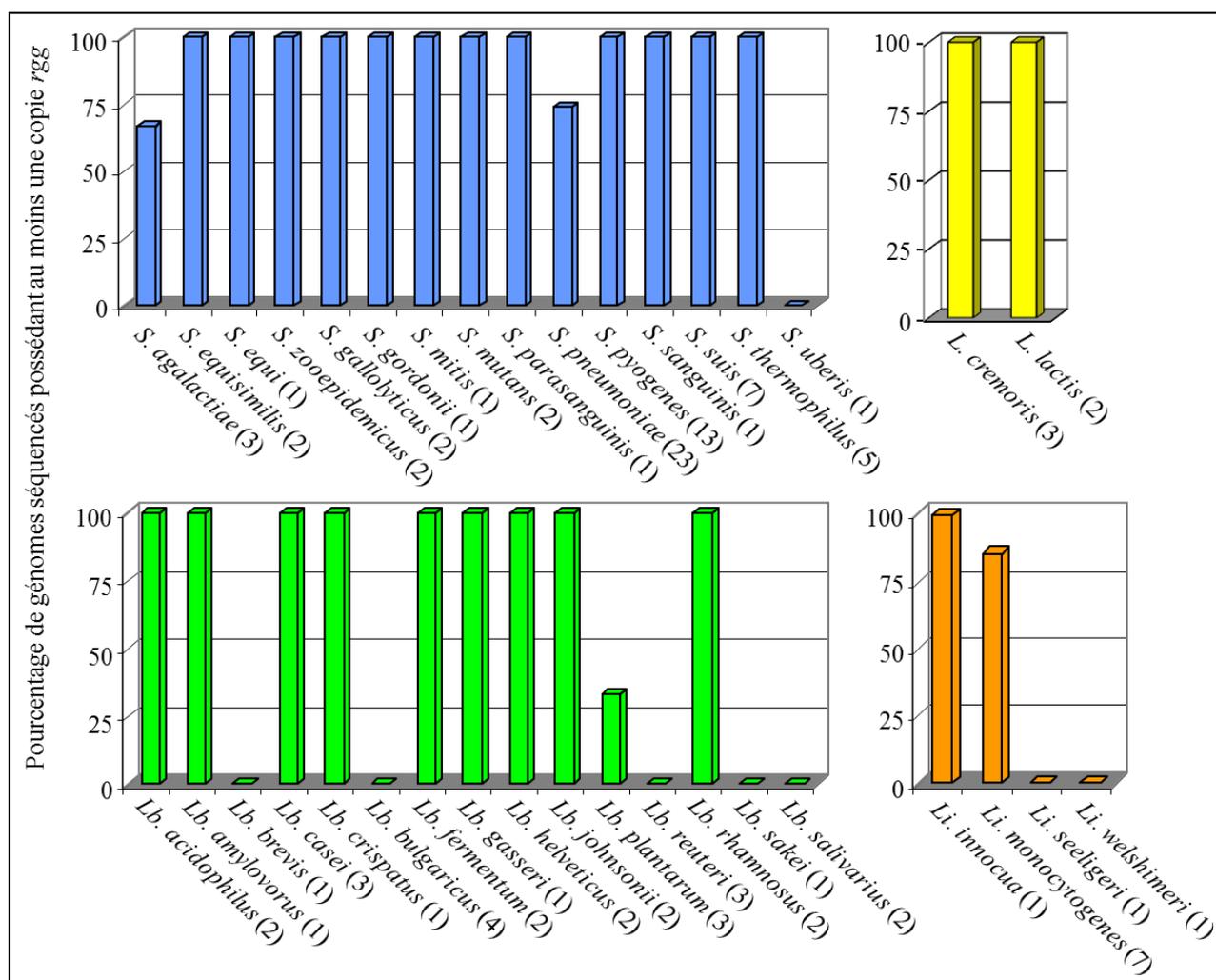


Figure 9 : Histogramme représentant le pourcentage de génomes séquencés possédant une ou plusieurs copie(s) *rgg* au sein de chaque espèce considérée. Le genre *Lactobacillus* est représenté en vert, le genre *Listeria* en orange, le genre *Lactococcus* en jaune et le genre *Streptococcus* en bleu. Le nombre de génomes séquencés par espèces est indiqué entre parenthèse. Seules les espèces dont le génome d’au moins une souche a été séquencé sont représentées.

Les genres *Streptococcus* et *Lactococcus* se distinguent des deux autres genres. En effet, à l’exception de *S. uberis*, des copies *Rgg* sont présentes dans toutes les espèces appartenant aux genres *Streptococcus* et *Lactococcus*. Au sein du genre *Lactobacillus*, des copies *rgg* sont présentes dans 10 espèces sur 15. Enfin, des copies *rgg* sont présentes dans la moitié des espèces appartenant au genre *Listeria*.

1.2. Plusieurs copies *rgg* peuvent être présentes par génome

Des données de la littérature indiquent que plusieurs copies *rgg* différentes peuvent être présentes au sein d’un seul et même génome. Ainsi, le génome de la souche CNRZ1066 de *S. thermophilus* possède 7 copies *rgg* (Ibrahim *et al.*, 2007b) alors que la souche NEM316 de

S. agalactiae en possède 3 (Samen *et al.*, 2006). Afin de déterminer si la présence en multicopie des gènes *rgg* est une caractéristique partagée (ou non), le nombre de copies *rgg* par génome a été recensé pour l'ensemble des 92 génomes bactériens séquencés possédant au moins une copie *rgg* (figure 10).

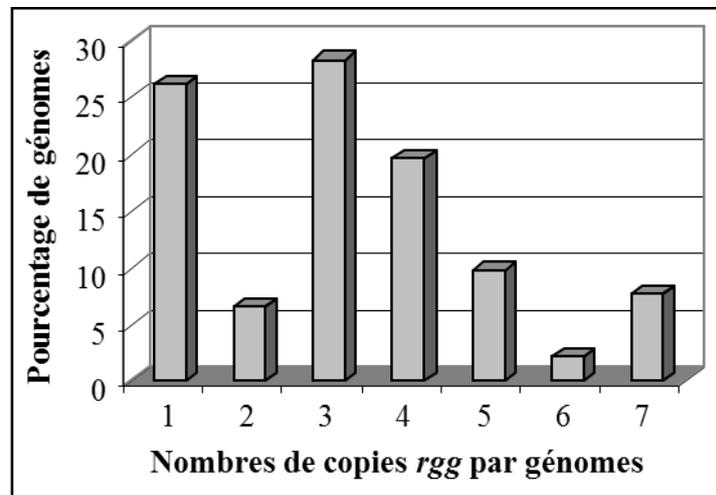


Figure 10 : Histogramme représentant le pourcentage du nombre de copies *rgg* par génome séquencé.

Les résultats de cette analyse indiquent que dans la majorité des cas (75%), les génomes contiennent plusieurs copies *rgg*. Le nombre de copies *rgg* au sein des génomes séquencés est très variable puisqu'il varie de 1 à 7 copies en fonction des souches. Afin d'analyser plus en détail cette variation, le nombre de copies *rgg* par génome a été étudié au sein des différents genres bactériens (figure 11).

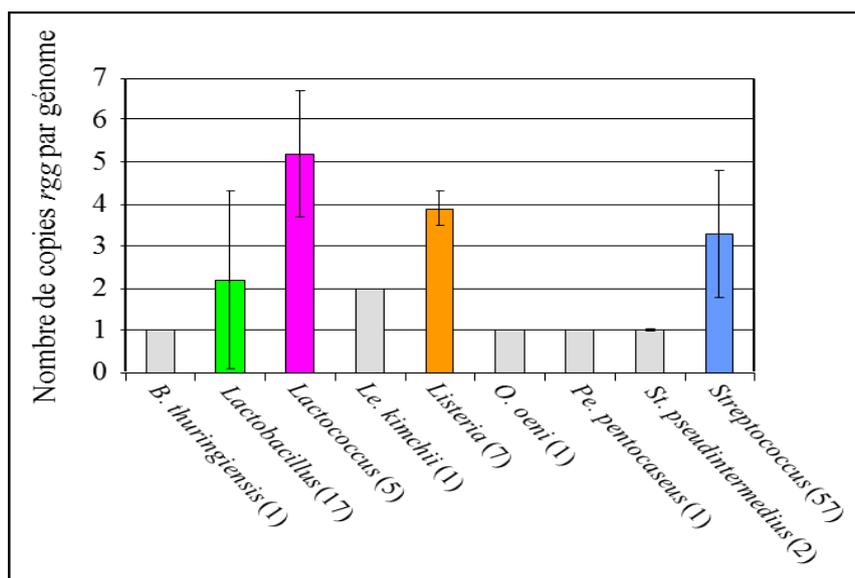


Figure 11 : Moyenne du nombre de copies *rgg* présentes dans les génomes qui en possèdent. Les résultats présentés correspondent à la moyenne \pm la déviation standard (quand elle a pu être

calculée). Lorsqu'un genre contient seulement une espèce présentant une ou plusieurs copies *rgg*, le nom de l'espèce est indiqué et l'histogramme est coloré en gris. Le nombre de génomes séquencés au sein de chacun des genres considérés est indiqué entre parenthèses.

A l'heure actuelle, les génomes des espèces bactériennes où peu de copies *rgg* sont présentes possèdent, pour la plupart, une seule copie *rgg*. La seule exception concerne le génome de l'espèce *Le. kimchii* qui possède deux copies *rgg*. Les génomes des genres *Lactobacillus*, *Lactococcus*, *Listeria* et *Streptococcus* sont ceux qui possèdent le plus de copies *rgg*. La figure 12 détaille le nombre de copies *rgg* portées par les génomes des différentes espèces de ces 4 genres.

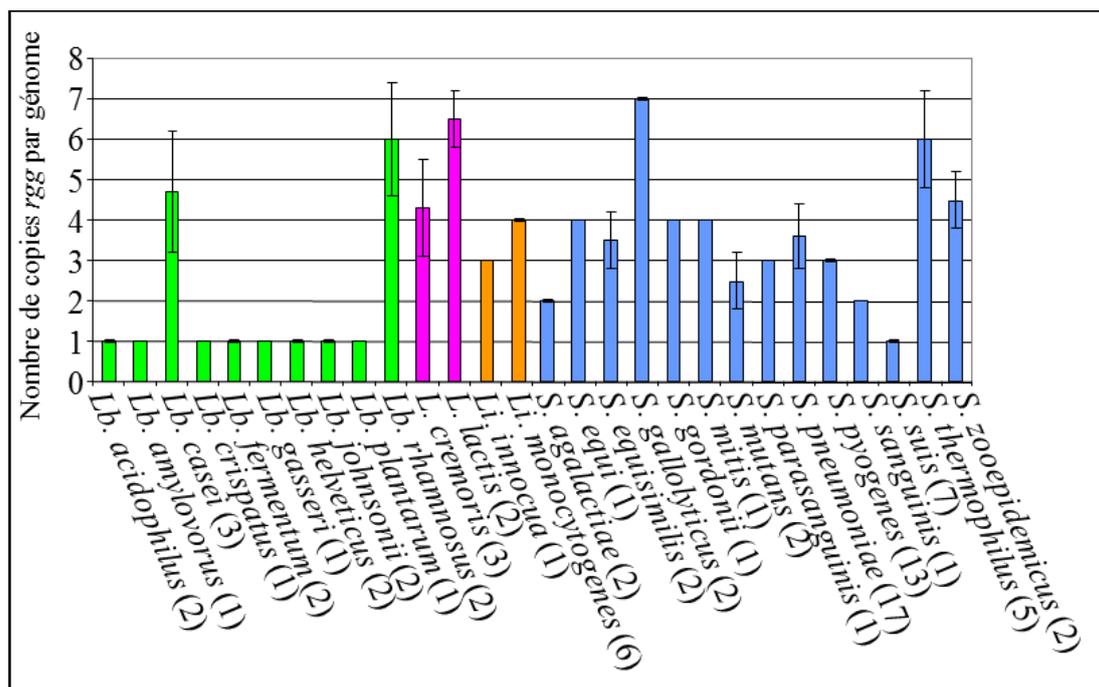


Figure 12 : Moyenne du nombre de copies *rgg* par génome des différentes espèces des genres *Lactobacillus*, *Lactococcus*, *Listeria* et *Streptococcus*. Les résultats présentés correspondent à la moyenne \pm la déviation standard (quand elle a pu être calculée). Le nombre de génomes séquencés au sein de chacun des espèces considérées est indiqué entre parenthèses.

Le genre *Lactobacillus* se différencie des 3 autres genres. En effet, au sein du genre *Lactobacillus*, seuls les génomes des espèces *Lb. casei* et *Lb. rhamnosus* possèdent plusieurs copies *rgg*, les autres n'en possédant qu'une. Au contraire, à l'exception des génomes de l'espèce *S. suis*, tous les génomes des espèces appartenant aux genres *Lactococcus*, *Listeria* et *Streptococcus* possèdent plusieurs copies *rgg*. De plus, les résultats montrent, qu'au sein des espèces du genre *Streptococcus*, le nombre de copies *rgg* par génome est très variable et est compris entre 1 et 7 copies *rgg*. Les génomes des espèces *Lb. rhamnosus*, *L. lactis*, *S. gallolyticus* et *S. thermophilus* sont ceux qui possèdent le plus grand nombre de copies *rgg* (entre 5 et 7 copies selon les souches). Par ailleurs, un test *t*-student a permis de valider cette observation, au seuil de risque 1%.

1.3. Les protéines de la famille Rgg se caractérisent par un domaine HTH et un domaine médian

1.3.1. Deux domaines peuvent être identifiés au sein des protéines Rgg

Une analyse réalisée par Louhman et Caparon (2007) a permis de mettre en évidence la conservation de 3 acides aminés au sein des séquences primaires de 57 protéines Rgg : les résidus G₄ et R₁₁ impliqués dans la fixation protéine-ADN et le résidu W₁₄₂ probablement impliqué dans des liaisons protéines-protéines. L'alignement réalisé par Loughman et Caparon (2007) indique par ailleurs un fort polymorphisme entre les protéines Rgg. Afin de compléter ces données, un alignement des séquences primaires des copies Rgg a été réalisé à l'aide du logiciel ClustalW. Pour cela, un seul exemplaire de chaque copie Rgg a été retenu parmi les 733, soit 499 séquences. La figure 13 présente un alignement de 10 séquences, choisies aléatoirement, pour illustrer le polymorphisme des protéines de la famille Rgg.

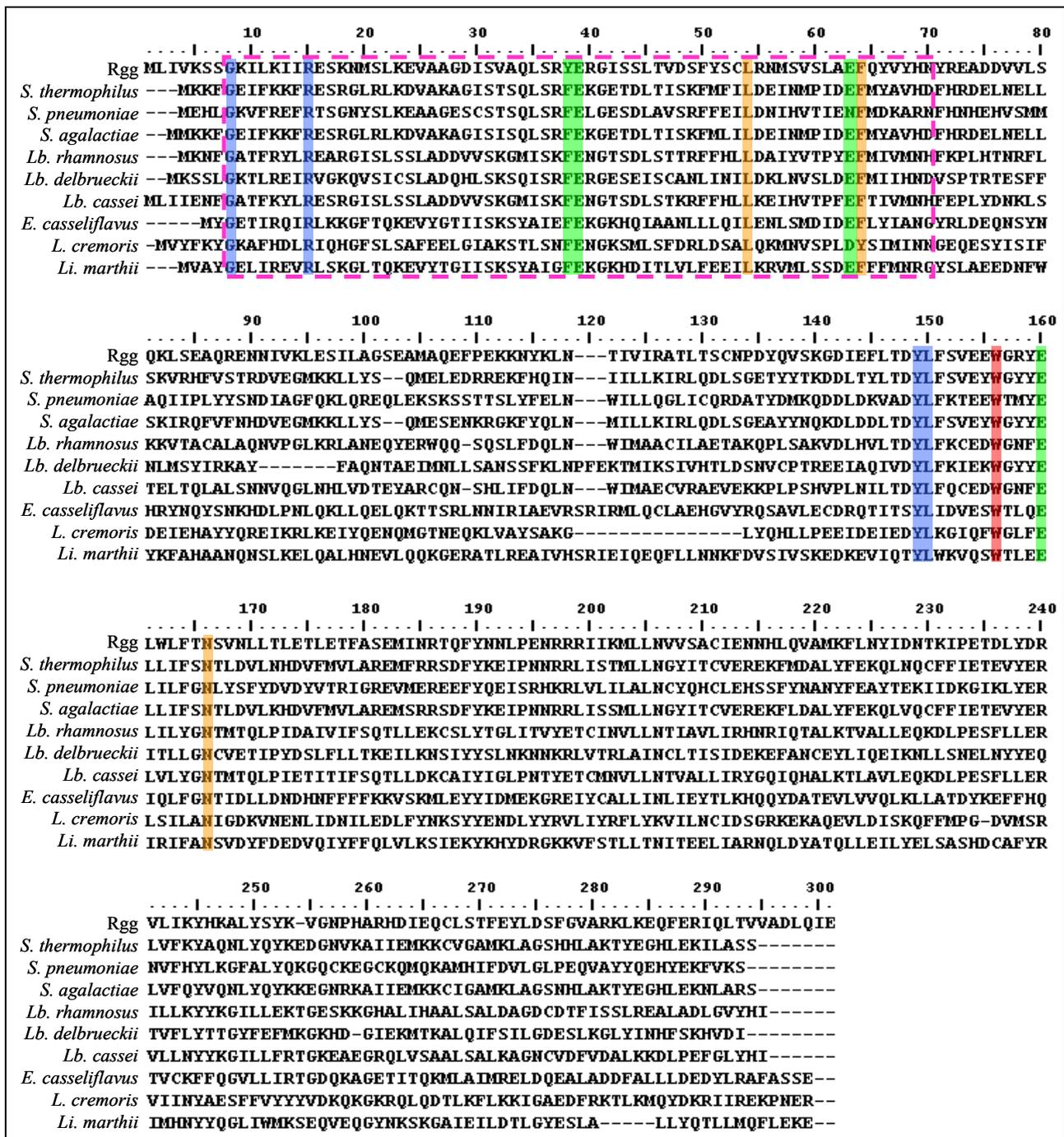


Figure 13 : Alignement de la séquence primaire des copies Rgg, réalisé à l'aide du logiciel ClustalW. Cet alignement a été réalisé à partir de 9 copies Rgg et illustre celui obtenu avec les 499 séquences primaires Rgg. La séquence primaire de la protéine Rgg de *S. gordonii* Challis a été utilisé comme référence (Sulavik *et al.*, 1992). Sur fond rouge apparaissent les résidus conservés dans 95% des 499 séquences primaires ; sur fond bleu dans 90% des séquences ; sur fond vert dans 85% des séquences et sur fond orange dans 80% des séquences. Les acides aminés constituant le domaine HTH-XRE, tel qu'annoté dans les bases de données, sont entourés par une ligne pointillée rose.

Cet alignement confirme le fort polymorphisme existant entre les séquences primaire des copies Rgg et montre qu'aucun résidu n'est conservé dans la totalité des séquences étudiées. Seuls 12 résidus sont conservés dans plus de 80% des 499 copies Rgg étudiées. Le résidu le plus conservé est le tryptophane en position 156 (W₁₅₆) qui est conservé dans 95% des séquences étudiées. La glycine

(G₈), l'arginine (R₁₅), la tyrosine (Y₁₄₉) et la leucine (L₁₅₀) sont conservées dans 90% des copies. La phénylalanine (F₃₈) et les trois glutamates (E₃₉, E₆₃, E₁₆₀) sont conservés dans 85% des séquences. Enfin, la leucine (L₅₄), la phénylalanine (F₆₄) et l'asparagine (N₁₆₆) sont conservées dans 80% des cas.

Parmi les résidus conservés, les résidus G₈, R₁₅, Y₃₈, E₃₉, L₅₄, E₆₃ et F₆₄ appartiennent au domaine HTH, décrit dans la littérature comme un domaine de fixation des protéines à l'ADN (Sulavik et Clewell, 1996 ; Neely *et al.*, 2003 ; Samen *et al.*, 2006 ; Loughman et Caparon, 2007). Ces résidus pourraient ainsi définir le « core » de ce domaine au sein des protéines Rgg. Celui-ci est situé de la position 8 à 70.

Les autres résidus conservés (Y₁₄₉, L₁₅₀, W₁₅₆, E₁₆₀ et N₁₆₆) sont tous localisés dans la région médiane. La conservation de ces résidus suggère qu'ils sont importants pour la fonctionnalité de la protéine Rgg et donc suggère qu'un domaine fonctionnel, non reconnu par les bases de données, pourrait être localisé dans cette région. Ainsi, ce domaine, bordé par les résidus Y₁₄₉ et N₁₆₆, a été appelé « domaine médian ». Les données de la littérature suggèrent que les protéines Rgg puissent agir sous la forme de dimères (Samen *et al.*, 2006 ; Loughman et Caparon, 2007 ; Fleuchot *et al.*, 2011). Ainsi, l'hypothèse privilégiée est que ce domaine puisse être impliqué dans des interactions protéines-protéines nécessaire à l'oligomérisation des protéines Rgg.

Par ailleurs, l'alignement montre qu'à l'exception de quelques courtes délétions/insertions, seules des mutations par substitutions sont observées entre les différentes séquences. Celles-ci ne sont pas localisées à un endroit précis mais réparties tout au long de la séquence.

1.3.2. Le domaine médian est spécifique des protéines Rgg

L'analyse des structures des régulateurs transcriptionnels révèle que la plupart des facteurs de transcriptions procaryotes sont des homodimères qui se fixent sur des séquences d'ADN, impliquant deux domaines particuliers. Les deux domaines impliqués dans ces fonctions permettent, la plupart du temps, à définir les familles de régulateurs transcriptionnels. En effet, tous les régulateurs transcriptionnels se fixent à l'ADN *via* un domaine, spécifique. Trois motifs récurrents de fixation à l'ADN ont été décrit : les motifs HTH, les motifs WHTH et les motifs HLH (Pabo et Sauer, 1992 ; Phillips et Phillips, 1994). Cette caractéristique pourrait être commune aux protéines de la famille Rgg qui se caractérise par un domaine HTH localisé en position N-terminale. Par ailleurs, des données de la littérature indiquent que certaines des protéines Rgg sont capables de se fixer l'ADN (Neely *et al.*, 2003 ; Samen *et al.*, 2006 ; Anbalagan *et al.*, 2011 ; Fleuchot *et al.*, 2011).

RESULTATS

Afin de déterminer si les résidus conservés dans les différentes copies Rgg, au niveau du domaine HTH et du domaine médian, sont caractéristiques des protéines Rgg ou s'ils sont plus largement conservés, les séquences primaires des 10 copies Rgg (celles présentées dans la figure 13) ont été comparées avec les séquences de plusieurs protéines caractéristiques appartenant à des familles distinctes de régulateurs transcriptionnels.

Sept protéines ont été choisies pour la diversité de leurs motifs de fixation à l'ADN. De plus, afin de ne pas biaiser l'alignement, les protéines ont été choisies de sorte que leur longueur en AA soit du même ordre de grandeur que celle des protéines Rgg. La protéine HrcA (360 AA), appartenant à la famille GntR, se caractérise par un domaine WHTH et un domaine de fixation effecteur/ligand (Hoskisson et Rigali, 2009). La protéine LysR (304 AA) se caractérise par un domaine WHTH et un domaine de fixation d'un co-inducteur (Maddocks et Oyston, 2008). La protéine GalR (331 AA), de la famille LacI, se caractérise par un domaine WHTH, un domaine d'oligomérisation et un site de fixation d'un inducteur (de Boer et Glickman, 1998). La protéine DeoR (254 AA) possède un motif WHTH et un domaine d'oligomérisation (Garces *et al.*, 2008). La protéine ComR (299 AA) possède un motif HTH-XRE (pour Xenobiotic Response Element). La protéine AraC (202 AA) se caractérise par la présence de deux motifs HTH et d'un domaine de dimérisation (Gallegos *et al.*, 1997). Enfin, la protéine Arc (338 AA) se caractérise par un motif HLH et un domaine de dimérisation. Le résultat de l'alignement obtenu est présenté sur la figure 14.

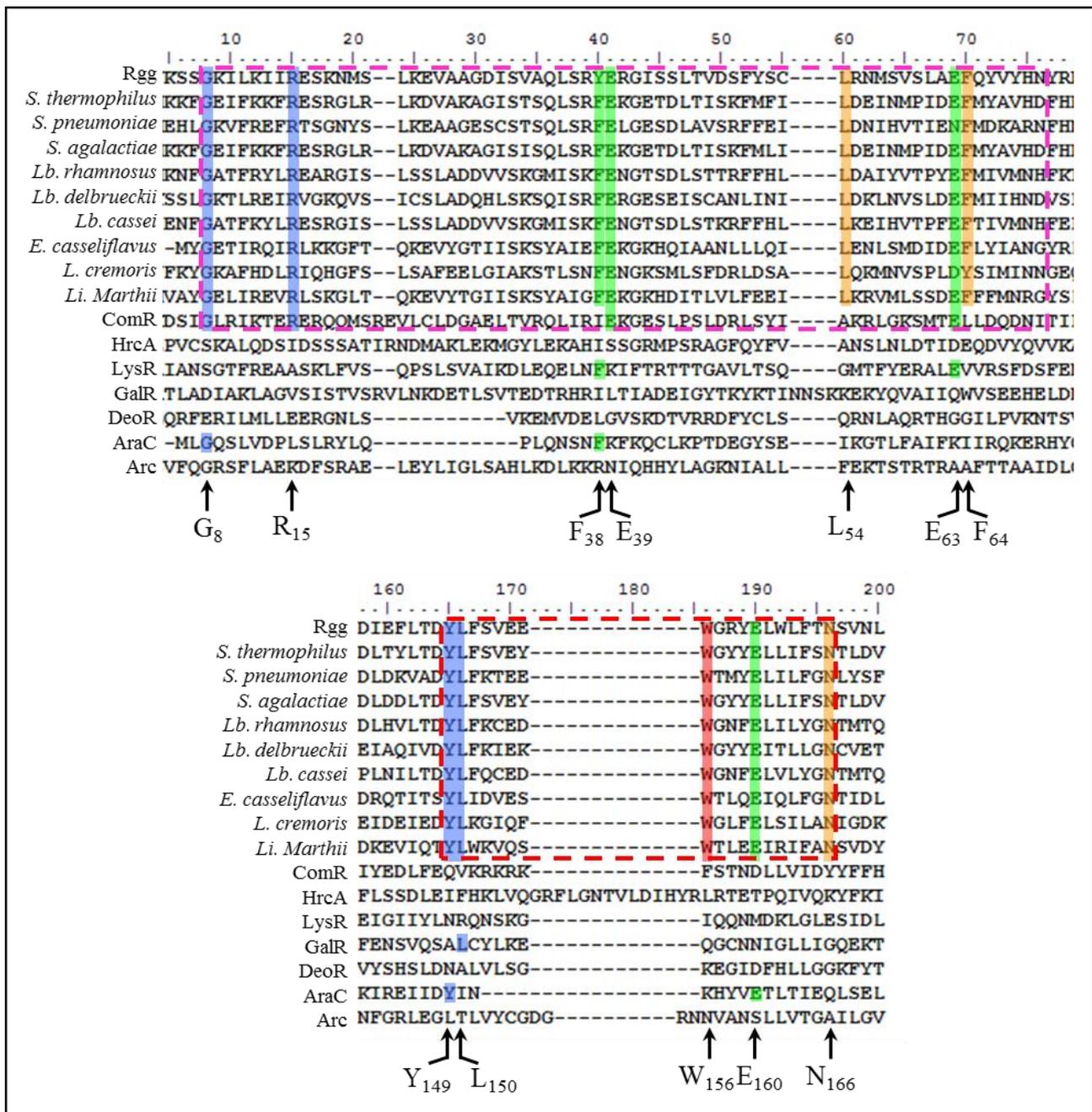


Figure 14 : Alignement de la séquence primaire de 10 copies Rgg, comparé avec la séquence primaire de 7 protéines appartenant à des familles de régulateurs transcriptionnels différents, réalisé à l'aide du logiciel ClustalW. Les 10 séquences primaires Rgg correspondent à celles utilisées dans la figure 13. Sur fond rouge apparaissent les résidus conservés dans 95% des 499 séquences primaires Rgg ; sur fond bleu dans 90% des séquences ; sur fond vert dans 85% des séquences et sur fond orange dans 80% des séquences. Les acides aminés constituant le domaine HTH des protéines Rgg sont entourés par une ligne pointillée rose et ceux constituant le domaine médian sont entourés par une ligne pointillée rouge. Les AA conservés au sein des protéines Rgg sont indiqués par des flèches.

A l'exception de la protéine ComR, aucune des protéines analysées ne possèdent de motifs conservés avec les protéines Rgg. Néanmoins, deux résidus sont conservés au sein du domaine HTH des protéines LysR (F₃₈ et E₆₃) et AraC (G₈ et F₃₈), par rapport aux protéines Rgg. Au contraire, 4 résidus sur les 6 résidus du domaine HTH des protéines Rgg (G₈, R₁₅, E₃₉ et E₆₃) sont conservés dans la séquence primaire de la protéine ComR. Ainsi, le domaine HTH de ces deux

types de protéines pourrait être similaire. Cela est cohérent avec des analyses récentes qui indiquent que ces protéines sont proches phylogénétiquement (Fleuchot *et al.*, 2011).

Au niveau du domaine médian, seul deux résidus pour la protéine AraC (Y₁₄₉ et E₁₆₀) et un pour la protéine GalR (L₁₅₀) sont conservés avec les protéines Rgg. Cela indique que le domaine médian serait absent des protéines analysées. Ainsi, le domaine médian pourrait être spécifique de la famille Rgg. Par ailleurs, un BLAST réalisé à partir de la séquence primaire du domaine médian n'a permis de mettre en évidence que des protéines appartenant à la famille Rgg, renforçant l'hypothèse que ce domaine serait spécifique des protéines de la famille Rgg.

1.4. Parenté phylogénétique des copies Rgg

Les analyses de Loughman et Caparon (2007) et l'alignement des 499 copies Rgg ont montré que les protéines Rgg sont très polymorphes. De ce fait, l'analyse de la parenté phylogénétique entre les différentes copies Rgg a été analysée. Pour cela, un arbre phylogénétique a été réalisé à l'aide du logiciel Seaview en utilisant la méthode du plus proche voisin (ou *Neighbour-Joining*). Pour ce faire, seules les copies différentes ont été prises en compte. De plus, afin d'optimiser l'alignement des séquences et ne pas introduire de biais, les copies présentant de grandes délétions n'ont pas été prises en compte dans l'analyse. Au total, 457 copies Rgg de taille comprises entre 275 et 339 AA ont été utilisées. La différence de taille s'explique par des insertions/délétions principalement positionnées au niveau de la région C-terminale, le logiciel ne prenant pas en compte ces régions afin de ne pas fausser l'analyse.

La figure 15 présente l'arbre phylogénétique enraciné à l'aide de la séquence primaire de 3 protéines ComR de streptocoques. Ce choix est justifié par une analyse récente démontrant que les protéines ComR forment un groupe proche mais distinct des protéines de la famille Rgg (Fleuchot *et al.*, 2011).

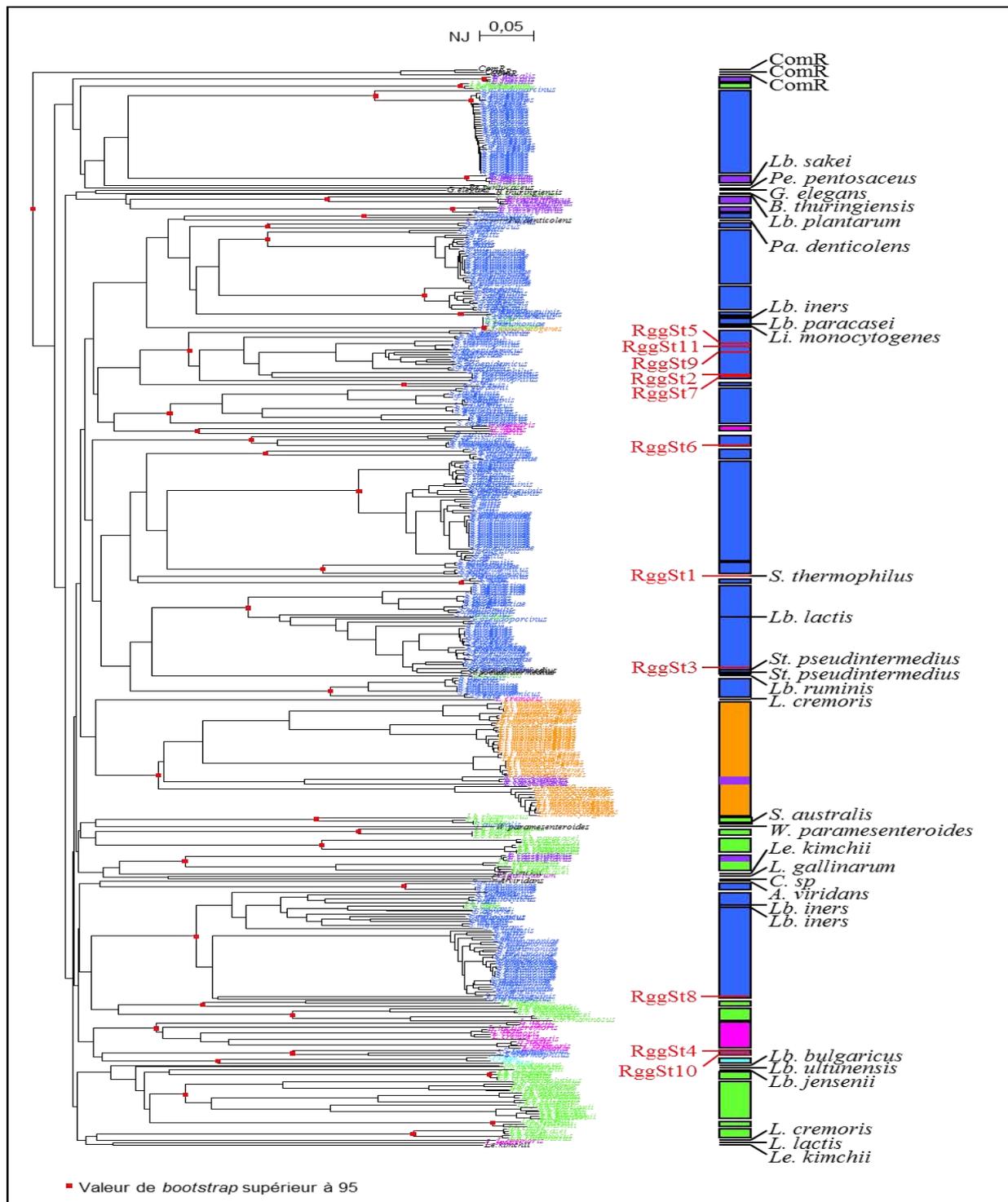


Figure 15 : Arbre phylogénétique des copies Rgg putatives réalisé à l'aide du logiciel Seaview par la méthode du *Neighbour-Joining* (NJ). L'arbre obtenu est présenté à gauche. Afin de permettre l'interprétation de l'arbre, une représentation, à droite, a été réalisée avec un regroupement monophylétique des copies Rgg (valeur de *Bootstrap* supérieure à 95). Les valeurs de *Bootstrap* ont été obtenues à partir de 1000 répliques. L'arbre a été enraciné à l'aide de la séquence primaire de 3 copies ComR de streptocoques. Les copies Rgg présentes chez les streptocoques sont présentées en bleu foncé, en vert chez les lactobacilles, en rose chez les lactococques, en violet chez les entérocoques, en orange chez les listeria et en bleu clair chez les oenocoques. Les copies Rgg isolées apparaissent en noir et celles de *S. thermophilus* apparaissent en rouge.

Les protéines Rgg ne forment pas un seul groupe monophylétique mais peuvent être classées en au moins 39 groupes distincts définis par des valeurs de Bootstrap supérieures à 95, bien que ne tenant pas compte de la distance d'évolution. Hormis les copies Rgg des genres *Listeria* et *Oenococcus* qui sont toutes regroupées au sein d'un même groupe pour chacun de ces deux genres, la topologie de l'arbre ne montre pas de lien avec la taxonomie puisque les protéines issues des différents genres ne sont pas regroupées entre elles. En effet, au moins 19 groupes distincts chez les streptocoques, 11 chez les lactobacilles, 6 chez les entérocoques et 2 chez les lactocoques peuvent être observés.

Par ailleurs, deux des groupes d'entérocoques sont inclus dans d'autres groupes : l'un est inclus dans le groupe des listérias et l'autre dans un groupe de lactobacilles. De plus, 9 copies Rgg sont incluses dans des groupes différents du genre pour lesquels elles sont issues. Ceci suggère que des copies Rgg aient pu être transférées d'un genre à l'autre. Enfin, 20 copies Rgg sont isolées et n'ont pu être associées à un groupe phylogénétique puisque les valeurs de *Bootstrap* les concernant sont trop faibles.

L'ensemble de cette analyse montre que les copies Rgg ne forment pas un seul groupe monophylétique mais, au contraire, de nombreux groupes, qui pour la plupart sont issues du même genre bactérien. Ainsi, des duplications auraient pu avoir lieu au sein de ces groupes. De plus, cet arbre montre que des copies issues des génomes de genres bactériens différents sont proches phylogénétiquement ce qui suggère que celles-ci aient pu être transférées d'un genre à l'autre.

1.5. Caractérisation des copies Rgg présentes chez *S. thermophilus*

1.5.1. Onze copies Rgg distinctes sont présentes chez *S. thermophilus*

Lorsque l'analyse PSI-Blast a été réalisée, les génomes complets de 5 souches appartenant à l'espèce *S. thermophilus* étaient disponibles. Leur analyse a démontré, qu'au total, 31 copies *rgg* sont présentes et la totalité de ces copies ne sont présentes que dans les génomes de cette espèce (tableau 8).

Tableau 8 : Copies *rgg* identifiées au sein des 5 génomes entièrement séquencés de *S. thermophilus*.

La présence de chaque copie est indiquée par le signe « + » et son absence par le signe « - ». Les copies localisées sur une même ligne sont homologues.

Nom de la protéine	Nom de la souche				
	LMG18311	CNRZ1066	LMD9	ND03	JIM8232
RggSt1	+	+	-	-	-
RggSt2	+	+	+	+	+
RggSt3	+	+	-	-	-
RggSt4	+	+	-	-	-
RggSt5	+	+	+	+	+
RggSt6	+	+	+	-	+
RggSt7	+	+	-	+	+
RggSt8	-	-	+	-	+
RggSt9	-	-	+	+	+
RggSt10	-	-	-	+	-
RggSt11	-	-	-	-	+
Total	7	7	5	5	7

Ce tableau révèle que tous les génomes de l'espèce *S. thermophilus* actuellement disponibles portent plusieurs copies *rgg* par génome allant de 5 à 7. La comparaison deux à deux de ces protéines montre que plusieurs de ces copies sont homologues (99 ou 100% d'identité). Ainsi, parmi ces 31 exemplaires, 11 copies Rgg différentes sont identifiées (renommées de RggSt1 (pour Rgg de *S. thermophilus*) à RggSt11).

Hormis les génomes des souches LMG18311 et CNRZ1066 qui se caractérisent par un contenu identique en copies *rgg*, tous les autres génomes de l'espèce *S. thermophilus* présentent un contenu en copies Rgg différent en nombre et/ou en nature des copies. Deux copies Rgg (RggSt2 et RggSt5) sont communes aux cinq souches suggérant qu'elles aient pu être héritées d'un ancêtre commun. Leur maintien suggère par ailleurs qu'elles puissent conférer aux souches qui les possèdent un avantage adaptatif. Au contraire, 3 autres copies (RggSt10, RggSt11 et RggSt12) ne sont présentes que dans une seule souche suggérant une acquisition récente.

Par ailleurs, le contenu en gènes *rgg* des différentes souches est conforme à leur parenté phylogénétique comme le montre la figure 16.

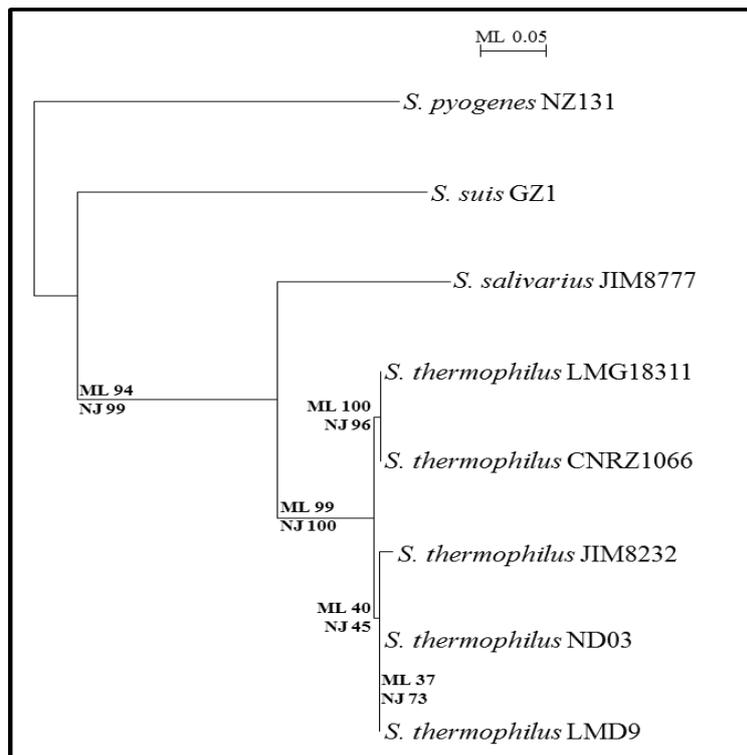


Figure 16 : Arbre phylogénétique des 5 souches de *S. thermophilus* dont le génome a été entièrement séquencé, réalisé à l'aide du logiciel Seaview. L'arbre a été construit à partir de l'alignement de la séquence primaire de l'ARN 16S par la méthode du *Neighbour-Joining* (NJ) et la méthode du maximum de vraisemblance (ML). L'arbre a été enraciné par la séquence de l'ARN 16S des souches NZ131 de *S. pyogenes*, GZ1 de *S. suis* et JIM8777 de *S. salivarius*. Les valeurs de *Bootstrap* ont été mesurées à partir de 100 répliques pour les deux méthodes.

Les souches LMG18311 et CNRZ1066 sont très proches phylogénétiquement et possèdent le même contenu en gènes *rgg*. La phylogénie des souches LMD9, ND03 et JIM8232 est cependant moins claire. L'hétérogénéité des copies *Rgg* au sein des souches de *S. thermophilus* est en faveur d'une évolution par perte ou par gain.

1.5.2. Polymorphisme et parenté des protéines *Rgg* de *S. thermophilus*

Les copies *Rgg* des génomes des 5 souches de *S. thermophilus* ont été comparées et leur pourcentage d'identité calculé à l'aide du logiciel BLASTTwoSeq (tableau 10).

Tableau 9 : Pourcentage d'identité entre les différentes copies Rgg de *S. thermophilus*. En gras sont indiqués les pourcentages d'identité supérieurs à 50.

	RggSt2	RggSt3	RggSt4	RggSt5	RggSt6	RggSt7	RggSt8	RggSt9	RggSt10	RggSt11
RggSt1	28%	30%	25%	35%	33%	31%	28%	36%	23%	31%
RggSt2		25%	26%	62%	26%	64%	31%	63%	24%	44%
RggSt3			23%	27%	29%	24%	28%	26%	22%	25%
RggSt4				28%	30%	23%	25%	28%	56%	26%
RggSt5					28%	61%	29%	81%	26%	47%
RggSt6						30%	28%	28%	27%	27%
RggSt7							31%	62%	26%	46%
RggSt8								31%	24%	27%
RggSt9									26%	47%
RggSt10										25%

A l'exception de quelques-unes d'entre elles, l'ensemble de ces données indique que les protéines Rgg de *S. thermophilus*, prises 2 à 2, sont caractérisées, dans la majorité des cas, par de faibles pourcentages d'identité de séquence et ne sont pas plus proches entre-elles qu'avec les autres protéines Rgg. Leur polymorphisme suggère qu'elles pourraient avoir des fonctions ou un mode d'action différent bien que les 11 copies possèdent les motifs HTH et médian.

Seules, les protéines RggSt2, RggSt5, RggSt7 et RggSt9 présentent un fort pourcentage d'identité entre-elles (entre 61 et 81%) et sont classées dans le même groupe monophylétique (figure 15, page 68). Ces données suggèrent que ces différentes copies pourraient résulter de la duplication d'un gène *rgg* ancestral au sein des génomes de *S. thermophilus*, avec une évolution divergente des différentes copies sœurs. La même hypothèse peut être émise pour expliquer la parenté des copies RggSt4 et RggSt10 qui appartiennent toutes deux au même groupe monophylétique. Afin de savoir si certaines régions sont plus conservées, ou non, entre ces séquences, un alignement a été réalisé à partir des séquences primaires des protéines possédant un fort pourcentage d'identité entre-elles. Les copies RggSt2, RggSt5, RggSt7, RggSt9 ont été regroupées entre-elles et nommées « groupe RggSt I ». La protéine RggSt11 étant incluse dans le même groupe monophylétique que ces 4 protéines, elle a été incluse dans ce groupe. De la même façon, les copies RggSt4 et RggSt10 ont été regroupées et forment le « groupe RggSt II ». Le résultat de l'alignement des séquences du groupe I est présenté sur la figure 17 et celui du groupe RggSt II sur la figure 18.

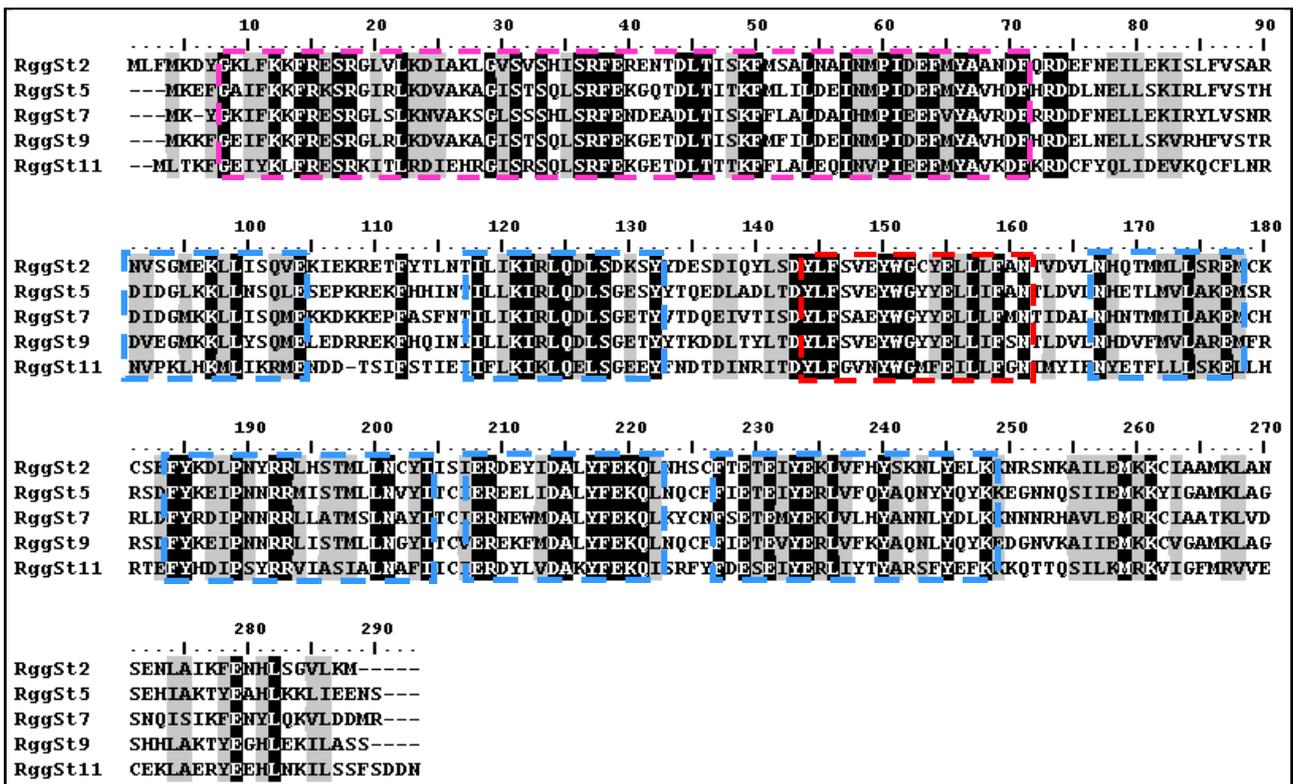


Figure 17 : Alignement de la séquence primaire des copies du groupe RggSt I (RggSt2, RggSt5, RggSt7, RggSt9 et RggSt11) de *S. thermophilus*, réalisé à l'aide du logiciel ClustalW. Les acides aminés sur fond noir sont conservés dans toutes les séquences et ceux sur fond gris sont similaires. Les acides aminés constituant le domaine HTH des protéines Rgg sont entourés par une ligne pointillée rose et ceux constituant le domaine médian sont entourés par une ligne pointillée rouge. Les régions où peu de polymorphisme est observé sont entourées par une ligne pointillée bleue.

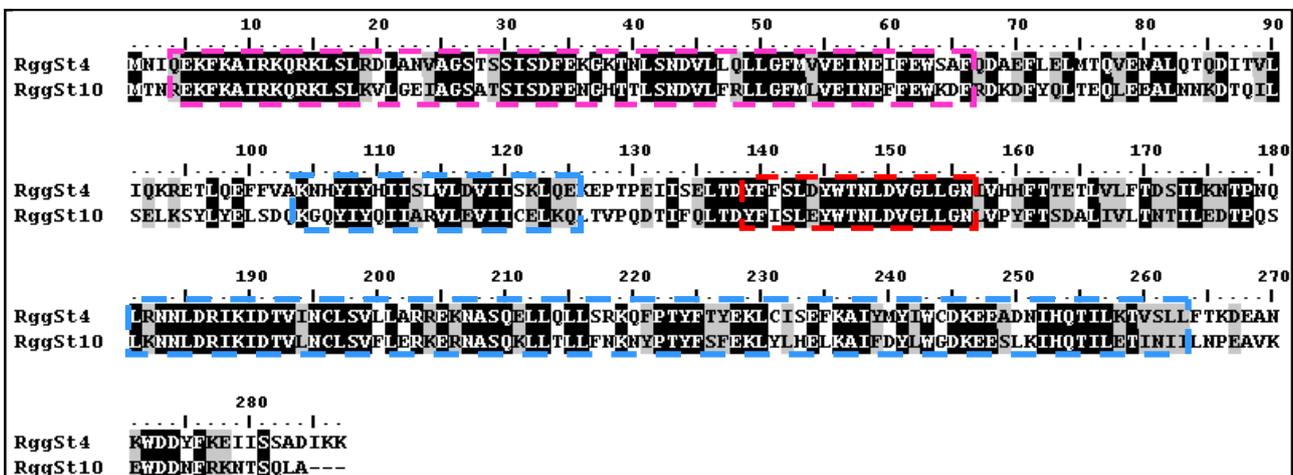


Figure 18 : Alignement de la séquence primaire des copies du groupe RggSt II (RggSt4 et RggSt10) de *S. thermophilus*, réalisé à l'aide du logiciel ClustalW. Les acides aminés sur fond noir sont conservés dans toutes les séquences et ceux sur fond gris sont similaires. Les acides aminés constituant le domaine HTH des protéines Rgg sont entourés par une ligne pointillée rose et ceux constituant le domaine médian sont entourés par une ligne pointillée rouge. Les régions où peu de polymorphisme est observé sont entourées par une ligne pointillée bleue.

Pour les deux groupes RggSt étudiés, l'alignement de séquences montre que le polymorphisme de séquence au sein de ces groupes ne semble pas aléatoire. Certaines régions, notamment les

domaines HTH et médian, sont mieux conservées alors que d'autres sont plus polymorphes. La conservation de ces régions au sein du groupe RggStI et II suggèrent que celles-ci seraient importantes pour le fonctionnement de ces protéines.

1.5.3. Etude de l'environnement génétiques des copies *rgg* de *S. thermophilus*

Sachant que le produit des gènes *rgg* régule, pour la plupart, des gènes adjacents, l'environnement génétique des ORF *rgg* de *S. thermophilus* a été analysé à l'aide du logiciel Artemis. Les résultats de cette analyse sont présentés dans le tableau 10.

Tableau 10 : Présentation des environnements génétiques des différentes ORF *rgg* présentes sur les génomes séquencés de *S. thermophilus*. Les noms d'ORF *rgg* correspondent aux annotations des banques de données. NG : Non Groupé; *stk* : sérine/thréonine kinase; *prs* : protéine à radical SAM; *IS* : Séquence d'Insertion. Les régions indiquées entre parenthèses présentent un pourcentage en bases G et C plus faible (entre 25 et 31%) que la moyenne de l'ensemble du génome (39% ± 4) sur lequel elles se trouvent.

Groupe RggSt	Protéine Rgg	Nom de l'ORF <i>rgg</i>	Environnement génétique	
Groupe RggSt I	RggSt2	<i>str0917</i>	(← shp → rgg → pep → prs → stk → transporteur → IS → autre ou inconnu →)	
		<i>stu0917</i> <i>ster0934</i> <i>std0883</i> <i>sth1116</i>	(← shp → rgg → pep → prs → stk → transporteur → IS → autre ou inconnu →)	
		RggSt5	<i>stu1572</i> <i>str1572</i> <i>ster1530</i> <i>std1506</i> <i>sth1809</i>	(← shp → rgg → transporteur →)
		RggSt7	<i>stu1950</i> <i>str1950</i> <i>std1893</i> <i>sth2252</i>	(← shp → rgg → prs → transporteur →)
	RggSt9	<i>ster1358</i> <i>std1322</i> <i>sth1611</i>	(← transporteur ← prs ← shp → IS →)	
	RggSt11	<i>sth1792</i>	(← shp → rgg → pep → prs → stk → transporteur → IS → autre ou inconnu →)	
Groupe RggSt II	RggSt4	<i>stu1511</i> <i>str1511</i>	(← shp → rgg → prs → transporteur →)	
	RggSt10	<i>std0397</i>	(← transporteur ← prs ← shp → IS →)	
NG	RggSt1	<i>stu0182</i> <i>str0182</i>	(← shp → rgg → pep → stk → transporteur →)	
	RggSt3	<i>stu1044</i> <i>str1044</i>	(← shp → rgg → IS →)	
	RggSt6	<i>ster1919</i>	(← shp → rgg → stk → transporteur →)	
		<i>stu1947</i> <i>str1947</i> <i>sth2247</i>	IS → (← shp → rgg → stk → transporteur →)	
RggSt8	<i>ster1299</i> <i>sth1541</i>	(← shp → rgg → prs → stk → transporteur → IS → autre ou inconnu →)		

← shp → rgg → pep → prs → stk → transporteur → IS → autre ou inconnu →
stu : souche LMG18311 ; *str* : souche CNRZ1066 ; *ster* : souche LMD9 ; *std* : souche ND03 ; *sth* (*sth8232*) : souche JIM8232

RESULTATS

Les numéros des ORF, annotés dans les banques de données, indiquent leur position dans le génome par rapport à l'origine de réplication et permettent de positionner les différentes ORF *rgg* les unes par rapport aux autres. L'analyse du positionnement de ces ORF montre, tout d'abord, qu'elles sont réparties tout le long du génome et non pas toutes regroupées dans une seule région, suggérant qu'elles sont insérées aléatoirement dans le génome. Par ailleurs, l'étude des gènes situés dans l'environnement proche des gènes *rgg* indique qu'ils sont toujours localisés à proximité de gène codant des transporteurs putatifs et souvent localisés à proximité d'ORF codant potentiellement des sérine/thréonine kinases et des petits peptides. Cela suggère que le produit de ces ORF pourrait jouer un rôle commun avec celui des ORF *rgg* de *S. thermophilus*. Par ailleurs, 9 des 11 gènes *rgg* de *S. thermophilus* sont flanqués d'un gène codant un peptide hydrophobe de type Shp.

Les peptides Shp ont été regroupés en deux classes en fonction de la conservation dans leur séquence un d'AA conservé de type aspartate (groupe I) ou glutamate (groupe II) (Ibrahim *et al.*, 2007a). Les résultats présentés dans le tableau 10 indiquent que toutes les ORF codant des Shp appartenant au groupe II sont localisées en amont des ORF *rgg* classées dans le groupe *rggSt* I, alors que les ORF codant des Shp appartenant au groupe I sont présentes uniquement en amont de gènes *rgg* appartenant au groupe *rggSt* II ou de copies *rgg* isolées. Ces résultats sont concordants avec l'analyse phylogénétique récente réalisée par Fleuchot *et al.* (2011) indiquant que les gènes *rgg* associés aux gènes *shp* du groupe I sont distincts de ceux associés aux gènes *shp* du groupe II.

Par ailleurs, le pourcentage en bases G et C des ORF *rgg* étant plus faible que celui du génome et la présence de séquences IS dans l'environnement génétique de 4 des gènes *rgg* sont en cohérence avec leur probable acquisition par transfert horizontal.

2. Rôle des gènes *rgg* de *S. thermophilus* LMG18311

2.1. Obtention d'une collection de mutants Δ *rgg* de *S. thermophilus* LMG18311

Lorsque ce travail a débuté, seuls les génomes des souches LMG18311 et CNRZ1066 étaient disponibles. Ils possèdent tous deux le même contenu en ORF *rgg*. La présence de ces gènes suggérait qu'ils puissent conférer un avantage adaptatif aux bactéries qui les portent. Cette hypothèse était confortée par des données de la littérature démontrant l'implication du gène *rgg*₁₉₅₀ dans la réponse au stress oxydant de la souche CNRZ368 (Fernandez *et al.*, 2004a) et, d'autre part, l'implication de gènes *rgg* dans la réponse aux stress chez *S. pyogenes*, *S. agalactiae* et *S. pneumoniae* (Chaussee *et al.*, 2004 ; Samen *et al.*, 2006 ; Pulliainen *et al.*, 2008).

Ainsi, l'objectif de ce travail était d'étudier l'implication des ORF *rgg* de *S. thermophilus* LMG18311 dans l'adaptation de *S. thermophilus* aux modifications de son environnement afin de déterminer, parmi ces gènes, lesquels sont importants pour la réponse de *S. thermophilus* aux stress, et de connaître la nature du changement environnemental auxquels ils répondent. Pour ce faire, la stratégie a consisté, dans un premier temps, à construire une série de mutants présentant chacun la délétion d'une ORF *rgg* puis, de comparer la tolérance à divers stress environnementaux des mutants et de la souche sauvage (non délétée).

La stratégie de délétion a été réalisée selon le protocole décrit par Layec *et al.* (2007) et les différentes étapes sont présentées dans la figure 19.

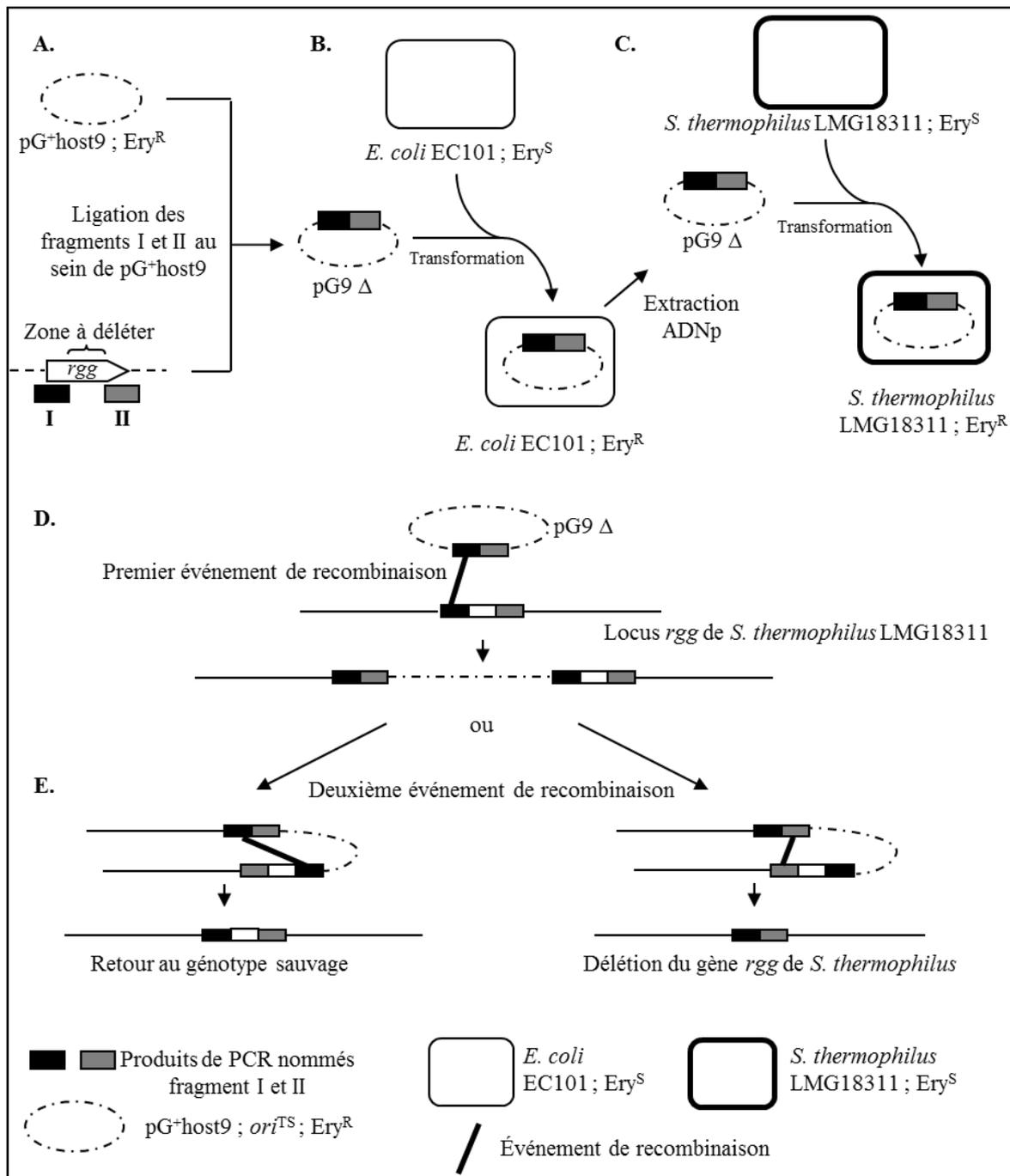


Figure 19 : Schéma de la stratégie utilisée pour déléter les ORF *rgg* de *S. thermophilus* LMG18311.

A. Deux régions d'ADN flanquant la zone à déléter sont amplifiées puis liguées entre elles et à l'ADN du plasmide pG⁺host9.

B. Le mélange de ligation est utilisé pour transformer la souche *E. coli* EC101. Les clones porteurs du plasmide recombinant sont ainsi résistants à l'érythromycine [Ery^R].

C. *S. thermophilus* LMG18311 est ensuite transformé par le vecteur recombinant. Les cellules transformées sont cultivées en présence d'érythromycine, à 30°C (température à laquelle le vecteur pG⁺host9 est capable d'autoréplication).

D. L'intégration du vecteur recombinant au sein du génome de *S. thermophilus* LMG18311 est ensuite sélectionnée. Le premier événement de recombinaison implique, sur le schéma, les fragments I, mais peut avoir lieu entre les fragments II. Les cellules sont cultivées à 42°C (température à laquelle le vecteur pG⁺host9 n'est plus capable d'autoréplication) en présence d'érythromycine.

E. La perte du plasmide conduisant soit au génotype sauvage soit à la délétion du gène cible est enfin sélectionnée. Un clone de *S. thermophilus* ayant intégré le vecteur recombinant est soumis à des cycles de cultures successifs avec changements de température (16 h à 30°C et 6 h à 42°C).

Cette stratégie a permis l'obtention de 9 mutants présentant la délétion d'une ou plusieurs ORF *rgg* chez *S. thermophilus* LMG18311 (tableau 11).

Tableau 11 : Liste des mutants Δrgg de la souche LMG18311 de *S. thermophilus*.

Nom	Genotype
Δrgg_{0182}	Souche LMG18311 portant une délétion de 699 pb dans le gène <i>rgg</i> ₀₁₈₂
Δrgg_{0917}	Souche LMG18311 portant une délétion de 807 pb dans le gène <i>rgg</i> ₀₉₁₇
Δrgg_{1511}	Souche LMG18311 portant une délétion de 777 pb dans le gène <i>rgg</i> ₁₅₁₁
Δrgg_{1947}	Souche LMG18311 portant une délétion de 825 pb dans le gène <i>rgg</i> ₁₉₄₇
Δrgg_{1950}	Souche LMG18311 portant une délétion de 723 pb dans le gène <i>rgg</i> ₁₉₅₀
$\Delta rgg_{0182}\text{-}\Delta rgg_{0917}$	Souche Δrgg_{0182} portant une délétion de 807 pb dans le gène <i>rgg</i> ₀₉₁₇
$\Delta rgg_{0182}\text{-}\Delta rgg_{1511}$	Souche Δrgg_{0182} portant une délétion de 777 pb dans le gène <i>rgg</i> ₁₅₁₁
$\Delta rgg_{0182}\text{-}\Delta rgg_{1947}$	Souche Δrgg_{0182} portant une délétion de 825 pb dans le gène <i>rgg</i> ₁₉₄₇
$\Delta rgg_{0182}\text{-}\Delta rgg_{1950}$	Souche Δrgg_{0182} portant une délétion de 723 pb dans le gène <i>rgg</i> ₁₉₅₀

L'obtention de 5 simples mutants Δrgg (*rgg*₀₁₈₂, *rgg*₀₉₁₇, *rgg*₁₅₁₁, *rgg*₁₉₄₇ et *rgg*₁₉₅₀) et de 4 doubles mutants $\Delta rgg_{0182}\text{-}\Delta rgg$ indique que les gènes *rgg* étudiés ne sont pas essentiels à la survie de *S. thermophilus* dans les conditions de laboratoire.

2.2. La délétion de gènes *rgg* n'a pas d'impact sur la croissance de *S. thermophilus* LMG18311 en condition de laboratoire

Afin d'étudier l'impact de la délétion d'un ou plusieurs gènes *rgg* de *S. thermophilus* LMG18311, la croissance des différents mutants a été étudiée en conditions de laboratoire. Pour ce faire, une culture de chacune des souches a étéensemencée en milieu LM17 et incubée à 42°C et la densité optique (DO) à 600nm a été mesurée au cours du temps (figure 20).

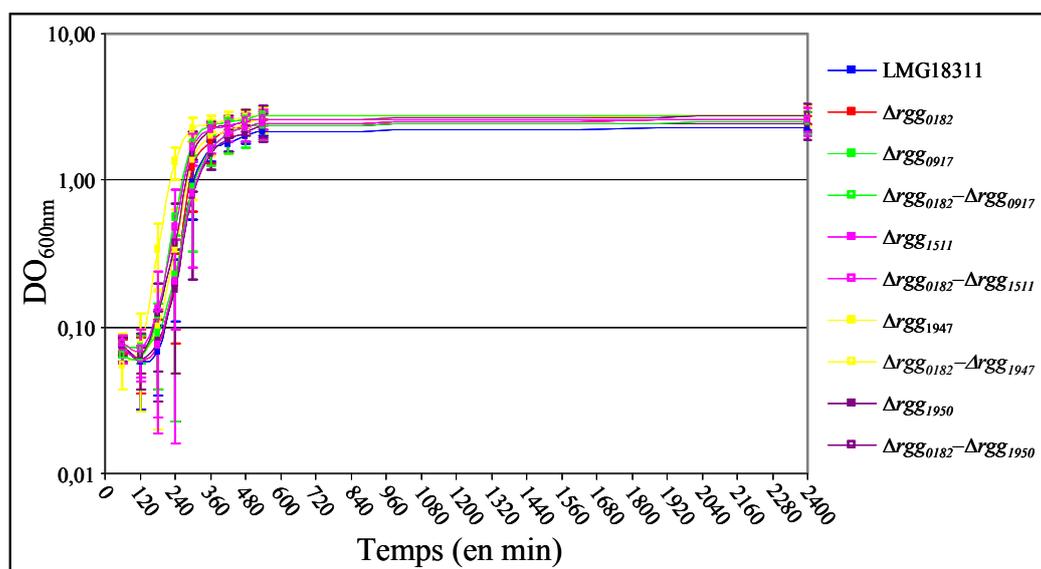


Figure 20 : Croissance des souches LMG18311 et Δ rgg de *S. thermophilus* LMG18311 réalisée à 42°C en milieu LM17. Les résultats correspondent à la moyenne \pm la déviation standard de 3 expériences indépendantes réalisées en triplicat.

Les courbes de croissance obtenues pour l'ensemble des souches étudiées sont similaires. En effet, les cellules entrent en phase exponentielle de croissance au bout de 2 heures de cultures et atteignent la phase stationnaire de croissance en 6 heures. Les paramètres de la croissance ont, ainsi, été mesurée et sont présentés dans le tableau 12.

Tableau 12 : Paramètres de croissance des souches LMG18311 et Δ rgg de *S. thermophilus* LMG18311 mesurées lors de cultures réalisées en milieu LM17 à 42°C.

Souche	DO finale	Temps de génération (en min)
LMG18311	2,3 \pm 0,1	31 \pm 4
Δ rgg ₀₁₈₂	2,6 \pm 0,2	30 \pm 1
Δ rgg ₀₉₁₇	2,5 \pm 0,4	28 \pm 5
Δ rgg ₁₅₁₁	2,6 \pm 0,6	28 \pm 5
Δ rgg ₁₉₄₇	2,8 \pm 0,4	30 \pm 4
Δ rgg ₁₉₅₀	2,4 \pm 0,5	27 \pm 3
Δ rgg ₀₁₈₂ - Δ rgg ₀₉₁₇	2,7 \pm 0,6	30 \pm 6
Δ rgg ₀₁₈₂ - Δ rgg ₁₅₁₁	2,6 \pm 0,5	32 \pm 2
Δ rgg ₀₁₈₂ - Δ rgg ₁₉₄₇	2,8 \pm 0,4	28 \pm 4
Δ rgg ₀₁₈₂ - Δ rgg ₁₉₅₀	2,8 \pm 0,6	33 \pm 3

Ces résultats montrent que la DO finale est d'environ 2,5 et le temps de génération d'environ 30 minutes pour chacune des souches étudiées. Cela suggère que, dans les conditions testées, les gènes

rgg étudiés n'influent pas sur la croissance de *S. thermophilus*.

2.3. Impact des protéines Rgg dans l'organisation des cellules en chaînes

Afin d'étudier les caractéristiques morphologiques des cellules des mutants Δrgg de *S. thermophilus*, des cultures ont étéensemencées en milieu complet LM17 à 42°C pendant une nuit et la morphologie des cellules en phase stationnaire de croissance a, ensuite, été étudiée au microscope photonique les cultures obtenues en phase stationnaire ont, ensuite, été analysées au microscope photonique. De plus, un comptage du nombre de cellules par chaînes a été réalisé. Pour cela, au moins 1000 cellules ont été comptées pour chacune des souches, en triplicat, à partir de cellules obtenues en phase stationnaire de croissance en milieu LM17. Les résultats de cette analyse sont présentés sur la figure 21 pour les simples mutants et sur la figure 22 pour les doubles mutants.

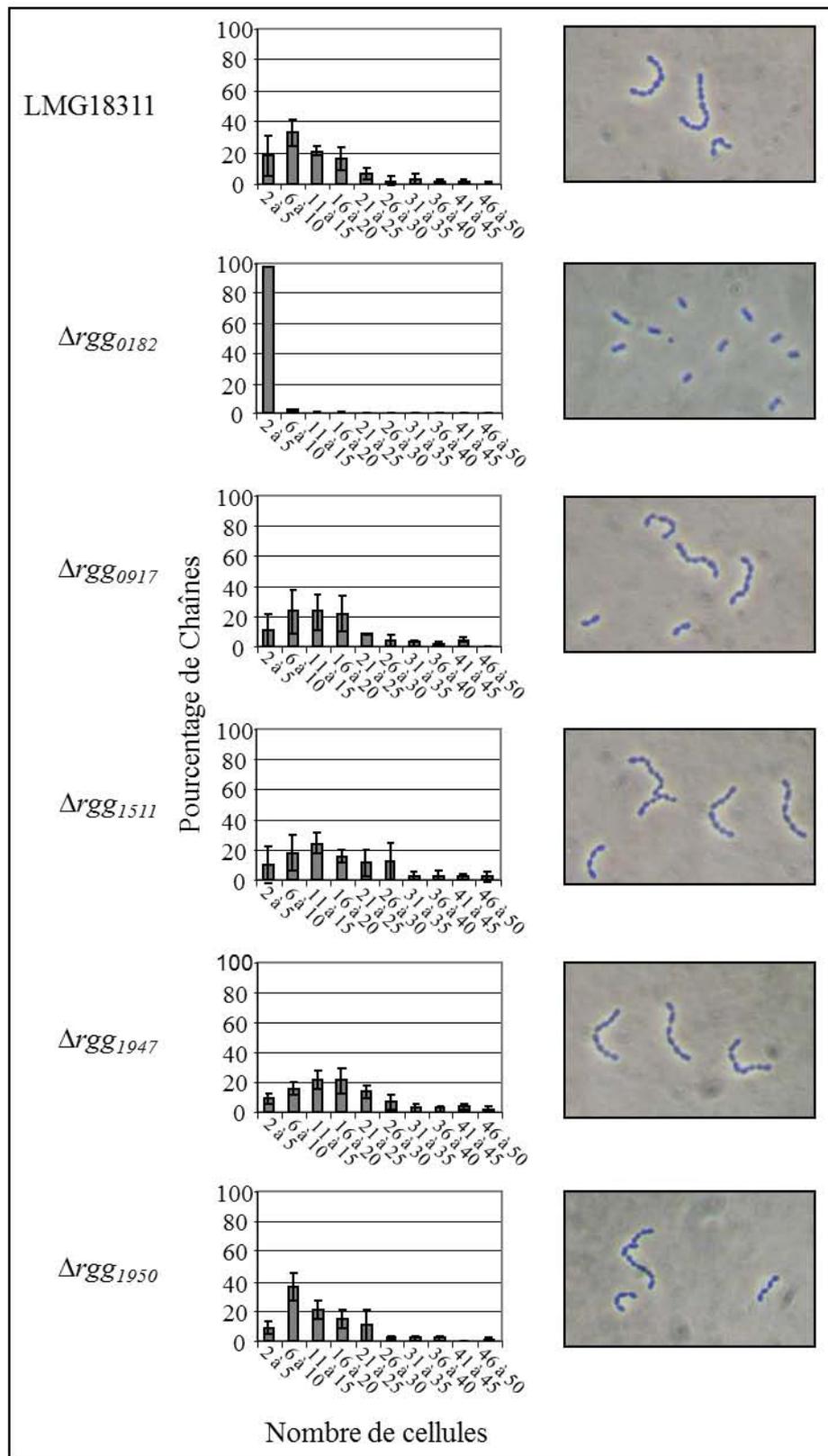


Figure 21 : Répartition du nombre de cellules par chaînes (à gauche) et photographies des cellules observées au microscope photonique (à droite) des simples mutants Δrgg et de la souche LMG18311 de *S. thermophilus*. Grossissement x1000.

RESULTATS

S. thermophilus se développe sous forme de chaînes de cellules comme le montre la photo de la souche LMG18311 de *S. thermophilus* de la figure 21. L'ensemble des mutants, à l'exception de celui portant la délétion du gène *rgg0182*, présente des chaînes de cellules similaires à celles de la souche sauvage. En effet, les chaînes de la souche sauvage sont formées, majoritairement, de 2 à 20 cellules (80% des chaînes), dont seulement 18% des chaînes comportent moins de 5 cellules. Les mutants Δrgg_{0917} , Δrgg_{1511} , Δrgg_{1947} et Δrgg_{1950} ont une répartition du nombre de cellules par chaînes similaire à celle de la souche sauvage. Au contraire, le mutant Δrgg_{0182} se développe sous forme de chaînes d'environ 2 à 5 cellules dans 97% des cas et seulement 3% des chaînes sont composées de 6 à 20 cellules. Ces résultats suggèrent que le gène *rgg0182* est impliqué dans l'organisation des cellules de *S. thermophilus* en chaînes.

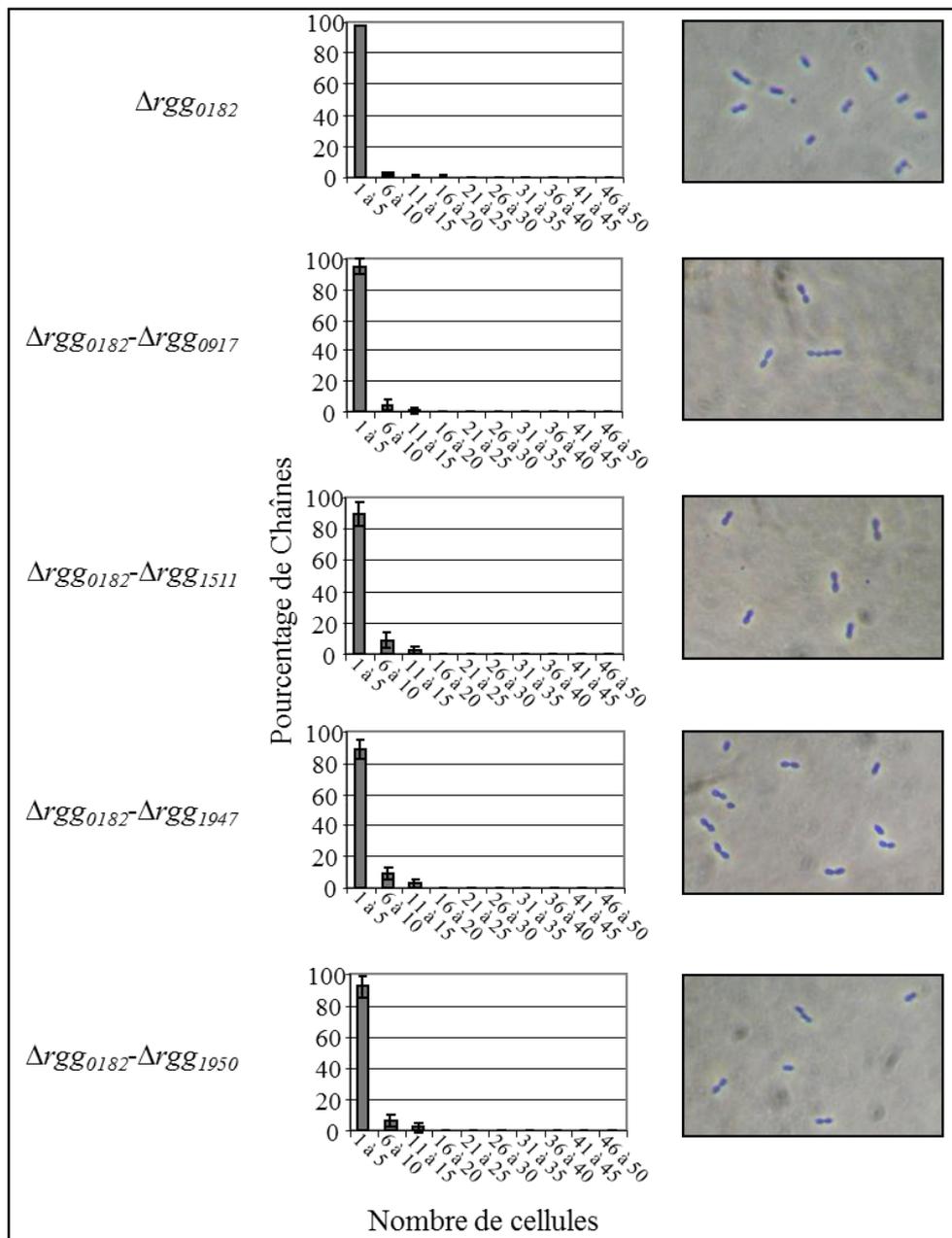


Figure 22 : Répartition du nombre de cellules par chaîne (à gauche) et photographie des cellules observées au microscope photonique (à droite) des mutants et $\Delta rgg_{0182}\text{-}\Delta rgg$ de la souche LMG18311 de *S. thermophilus*. Grossissement x1000.

Les doubles mutants portant la délétion du gène *rgg0182* (souches Δrgg_{0182} , $\Delta rgg_{0182}\text{-}\Delta rgg_{0917}$, $\Delta rgg_{0182}\text{-}\Delta rgg_{1511}$, $\Delta rgg_{0182}\text{-}\Delta rgg_{1947}$ et $\Delta rgg_{0182}\text{-}\Delta rgg_{1950}$) présentent des chaînes de cellules plus courtes que celles de la souche sauvage. Toutefois, leur phénotype est similaire à celui du simple mutant Δrgg_{0182} de *S. thermophilus* (figure 22) indiquant l'absence d'effet cumulatif des délétions. Cela suggère que seule la mutation du gène *rgg0182* est nécessaire et suffisante pour expliquer ce phénotype. Ainsi, l'hypothèse privilégiée est que seul le gène *rgg0182* serait impliqué dans le mécanisme de la séparation cellulaire.

Des analyses similaires, réalisées en phase exponentielle de croissance, ont donné des résultats identiques pour les 10 souches étudiées. Cela indique que la répartition du nombre de cellules par chaînes ne serait pas dépendante de la phase de croissance.

2.4. Tolérance des mutants à différents stress

Les données de la littérature indiquent que certaines protéines Rgg sont impliquées dans la réponse à différents stress. Pour exemple la protéine Rgg₁₉₅₀ de *S. thermophilus* CNRZ368 est impliquée dans la défense contre le stress oxydant (Thibessard 2004). Ainsi, afin de savoir si les différents gènes *rgg* de *S. thermophilus* LMG18311 participent à la réponse adaptative, la stratégie a consisté en la comparaison de la survie des différents mutants à celle de la souche sauvage, exposée à un stress.

De cette manière, la réponse au stress acide a été analysée. Afin de déterminer la valeur du pH à laquelle le stress doit être réalisé, un suivi de la production d'acide des souches sauvages et mutants Δrgg cultivées en milieu LM17 à 42°C et a été suivie (figure 23).

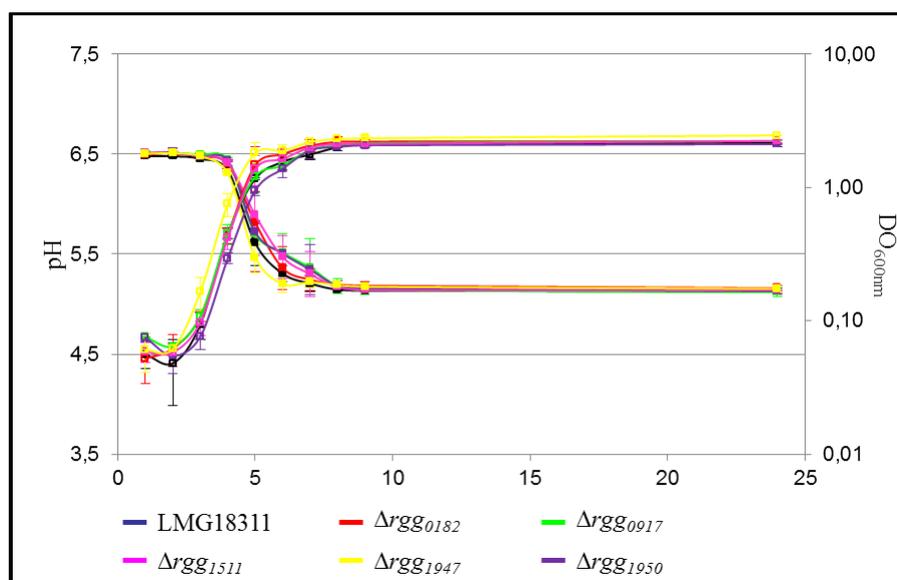


Figure 23 : Suivi de l'acidité du milieu de culture et de la croissance des souches LMG18311 et Δrgg de *S. thermophilus*, à 42°C en milieu LM17. Les carrés plein correspondent aux valeurs de pH et les carrés vides correspondent aux valeurs de DO. Les résultats correspondent à la moyenne \pm la déviation standard de 3 expériences indépendantes réalisées en triplicat.

Aucune différence n'est observée entre la souche sauvage et les simples mutants Δrgg . En effet, pour les 6 souches étudiées, le pH est de 6,5 lors de l'ensemencement des cultures puis cette valeur diminue au cours de la phase exponentielle pour atteindre une valeur de 5,1 lors de l'entrée en phase stationnaire de croissance. Ces données suggèrent qu'aucun des gènes *rgg* étudié n'est impliqué

dans la production d'acide lactique chez *S. thermophilus*, dans les conditions testées.

La tolérance au stress acide a ensuite été évaluée sur des cellules des souches sauvage et Δrgg en phase stationnaire de croissance, cultivées pendant 10h en milieu LM17 avant et après 30 minutes d'incubation à pH acide (pH = 4,5 ; 4 ; 3,5 et 3, c'est-à-dire à des pH inférieurs à celui atteint par les souches en phase stationnaire de croissance). A pH = 4,5, quasiment 100% de survie a été observée alors qu'à pH = 3, quasiment aucune survie n'est observée, pour toutes les souches étudiées.

Les taux de survie obtenus à pH = 4 et 3,5 sont comparables pour les deux pH étudiés. Les résultats obtenus à pH = 3,5 sont présentés dans la figure 24.

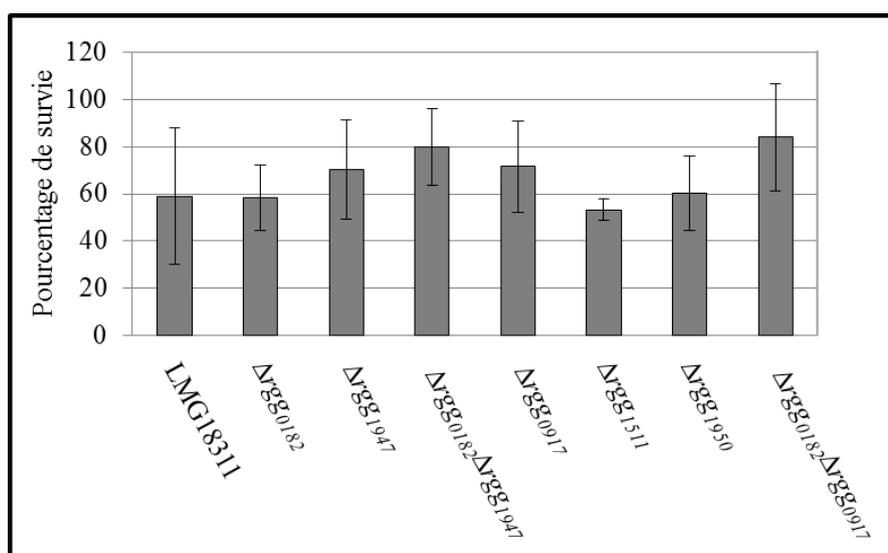


Figure 24 : Taux de survie des souches LMG18311 et des mutants Δrgg de *S. thermophilus* après un stress acide (pH = 3,5, pendant 30 minutes). La culture de *S. thermophilus* a été réalisée en milieu LM17 à 42°C puis exposée à un stress acide. Le pourcentage de survie a été calculé par la formule $N/N_0 \times 100$ où N_0 est le nombre d'UFC obtenu dans la condition contrôle et N est le nombre d'UFC après exposition au stress. Les résultats correspondent à la moyenne \pm la déviation standard de 3 expériences indépendantes réalisées en triplicat.

Aucune différence n'est observée entre la souche sauvage et les mutants Δrgg de *S. thermophilus*. En effet, le taux de survie obtenu est d'environ 60% quelle que soit la souche considérée. Les mêmes expériences ont été réalisées en phase exponentielle de croissance et des résultats similaires ont été obtenus. Ainsi, les gènes *rgg* de *S. thermophilus* ne seraient pas impliqués dans la réponse au stress acide.

Par ailleurs, la tolérance au stress oxydant a été analysée dans les mêmes conditions de cultures. Différents temps d'incubation (15 et 30 minutes) en présence d'agent oxydant (ménadione et tellurite) à différentes concentrations (1, 5 et 10 M), en phase exponentielle et en phase stationnaire de croissance ont été analysés. Cependant, aucune des conditions testées n'a permis de révéler de

différences phénotypiques particulières entre les différentes souches testées.

Ainsi, la stratégie qui consistait à étudier l'ensemble des mutants Δrgg n'a pas permis d'étudier l'implication de tous ces gènes *rgg* conduisant à se focaliser sur un seul gène *rgg*. Du fait de l'implication du gène *rgg₀₁₈₂* dans l'organisation des cellules en chaînes, celui-ci a été choisi pour la suite des travaux.

3. Le gène *rgg0182* code un régulateur transcriptionnel requis dans l'adaptation aux changements de température de *Streptococcus thermophilus* LMG18311

Cette partie a fait l'objet d'une publication intitulée « *The rgg0182 gene encodes a transcriptional regulator required for the full Streptococcus thermophilus LMG18311 thermal adaptation* » dans BMC Microbiology.

3.1. Caractérisation du locus *rgg0182* de *S. thermophilus* LMG18311

3.1.1. Analyse de la séquence du gène *rgg0182*

Le séquençage du gène *rgg0182* a été réalisé et montre que la séquence obtenue diffère de celle présente dans les banques de données (annotée *stu0182*) obtenue lors du séquençage complet du génome de la souche LMG18311 de *S. thermophilus* (Bolotin *et al.*, 2004). La figure 25 illustre ces différences.

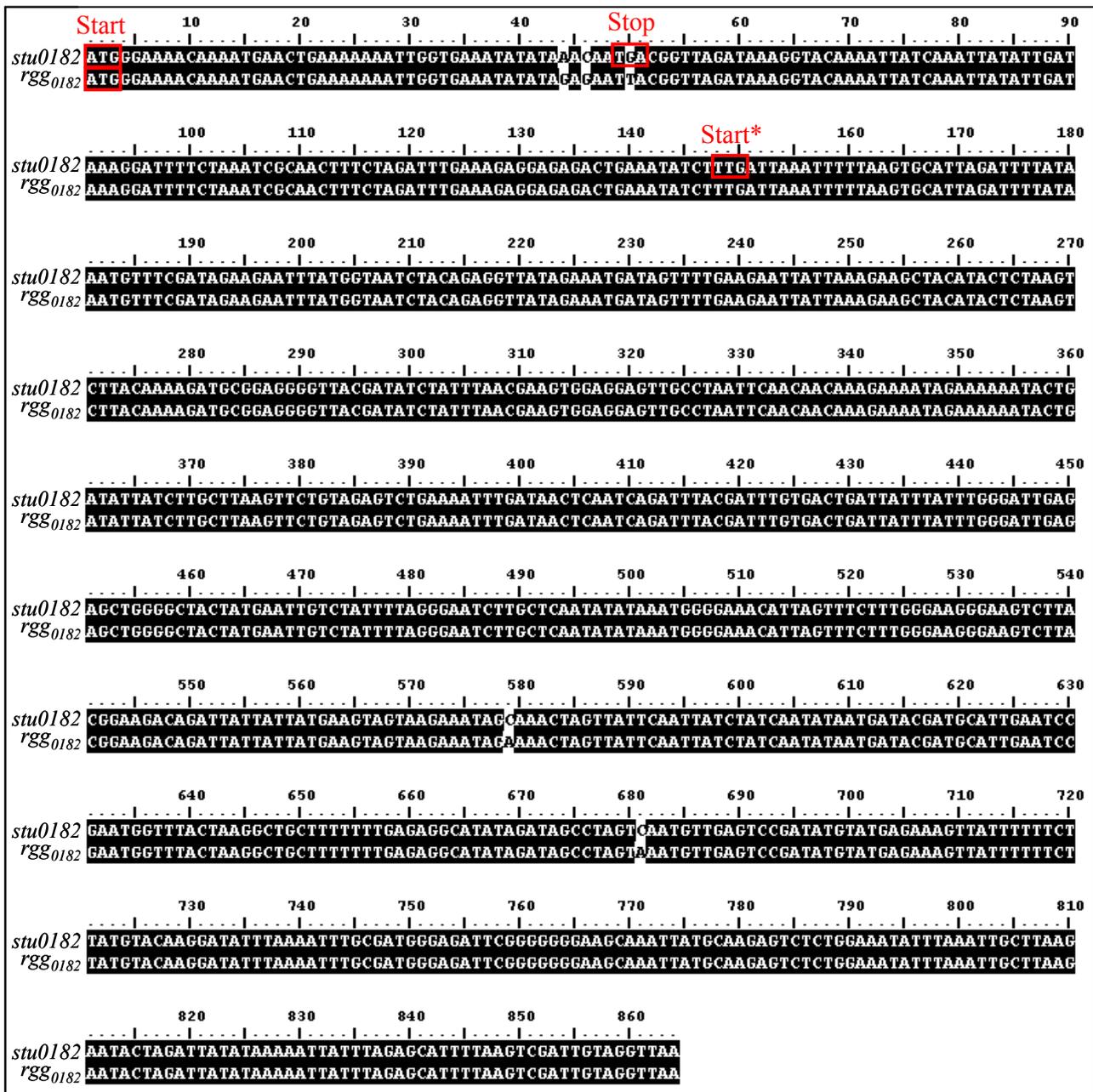


Figure 25 : Alignement des séquences nucléotidiques des ORF *stu0182* et *rgg0182* de *S. thermophilus*. La séquence *stu0182* a été obtenue lors du séquençage complet de la souche LMG18311 (Bolotin *et al.*, 2004). La séquence *rgg0182* a été obtenue lors de ce travail. Le sigle « Start* » correspond au codon d'initiation de la transcription annoté dans les banques de données.

Cet alignement a mis en évidence la présence de 5 substitutions (position 44, 46, 50, 579 et 681). La présence d'une guanine en position 50 de l'ORF *stu0182* génère un codon stop qui a conduit à considérer le codon Start* comme codon d'initiation de la traduction de la protéine Stu0182 (Bolotin *et al.*, 2004). Ce gène se différencie ainsi du gène *rgg0182* par une troncature de 147 nucléotides. Le séquençage du gène *rgg0182* permet de considérer le codon Start comme codon d'initiation de la traduction de la protéine Rgg0182. Ainsi le gène *rgg0182* a une taille de 864 pb et

code une protéine de 288 acides aminés avec une prédiction de masse moléculaire de 33,6 kDa. Cette séquence nucléotidique présente 100% d'identité avec celle du gène *str0182* de la souche CNRZ1066 de *S. thermophilus*. Les divergences de séquence observées entre les gènes *stu0182* et *rgg0182* ainsi que l'homologie avec la souche CNRZ1066 ont conduit à vérifier que la souche utilisée au cours de ce travail était bien LMG18311 et non CNRZ1066. Pour cela, la région variable du gène *cse*, démontrée antérieurement comme étant spécifique de souches (Borges *et al.*, 2005), a été amplifiée par PCR puis digérée par l'enzyme *NheI*(figure 26).

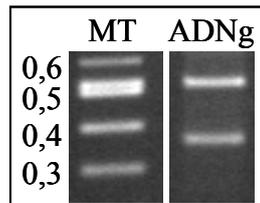


Figure 26 : Profil de restriction obtenu par digestion *NheI* du produit d'amplification par PCR de la région variable du gène *cse* de la souche LMG18311 de *S. thermophilus*. MT = marqueur de taille. ADNg indique que la PCR a été réalisée avec pour matrice l'ADN génomique de la souche LMG18311.

La digestion de la région variable du gène *cse* de la souche CNRZ1066 conduit à l'obtention de 3 fragments (383, 276 et 213pb) alors qu'elle conduit à l'obtention de 2 fragments (520 et 354pb) chez la souche LMG18311. Le profil de digestion ainsi obtenu est cohérent avec les tailles attendues pour la souche LMG18311 de *S. thermophilus* et montre que la souche étudiée est bien *S. thermophilus* LMG18311. La séquence réelle du gène *rgg0182* a été soumise dans les banques de données au numéro d'accession : GenBank Accession JF699754.

3.1.2. L'environnement génétique du gène *rgg0182* est similaire à celui du gène *rgg1358*

L'environnement génétique du gène *rgg0182* de *S. thermophilus* LMG18311 a été analysé et comparé à celui du gène *rgg1358* de *S. thermophilus* LMD9 (figure 27).

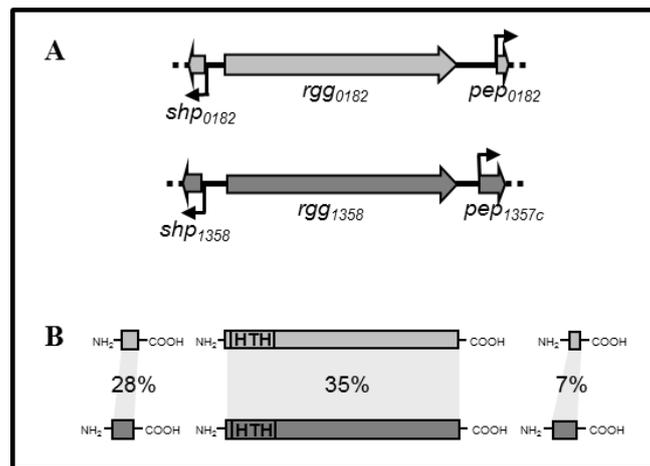


Figure 27 : Représentation schématique des loci *rgg*₀₁₈₂ de la souche LMG18311 et *rgg*₁₃₅₈ de la souche LMD9 de *S. thermophilus* (A) et des protéines correspondantes (B). Les gènes sont représentés par des flèches et les protéines par des rectangles. Les flèches cassées indiquent les promoteurs. L'identité entre les séquences primaires des protéines est indiquée en pourcentage.

La protéine Rgg₀₁₈₂ présente environ 30% d'identité avec les autres protéines de la famille Rgg et 35% d'identité avec la protéine Rgg₁₃₅₈ de *S. thermophilus* LMD9 (figure 27). Cette dernière a été récemment décrite comme étant impliquée dans un mécanisme de quorum sensing (QS) (Fleuchot *et al.*, 2011). La protéine Rgg₀₁₈₂ contient un motif HTH-XRE en position N-terminale (localisé aux niveaux des résidus 11 à 67). Ce gène est entouré par deux ORF, non annotées sur le génome de la souche LMG18311 mais révélées par des analyses réalisées avec le logiciel bactageneSHOW permettant de détecter de petites ORF (Ibrahim *et al.*, 2007a). En amont du gène *rgg*₀₁₈₂, le gène *shp*₀₁₈₂ (63 pb) code potentiellement un petit peptide hydrophobe appartenant au groupe I de la famille des SHP (Fleuchot *et al.*, 2011). En aval du gène *rgg*₀₁₈₂, le gène *pep*₀₁₈₂ (42 pb) code un petit peptide ne présentant aucune similarité avec des peptides présents dans les banques de données. Cette organisation présente des similarités avec l'environnement génétique du gène *rgg*₁₃₅₈ de la souche LMD9. Néanmoins ces deux loci sont distincts comme l'illustre le faible pourcentage d'identité existant entre les différentes séquences primaires des protéines SHP (28%), Rgg (35%) et Pep (7%). De plus, les deux protéines SHP ne sont pas classées dans le même groupe de la famille SHP (Fleuchot *et al.*, 2011) (figure 27).

3.2. Etude du taux de transcrit *rgg*₀₁₈₂ de *S. thermophilus* LMG18311

Dans la littérature, la transcription des gènes *rgg* a été peu étudiée. En effet, seule la transcription du gène *ropB* de *S. pyogenes* a été décrite. La transcription du gène *ropB* est dépendante de la phase stationnaire de croissance (Neely *et al.*, 2003). Ainsi, la transcription du

gène *rgg0182* a été suivie par des expériences de qPCR, utilisant comme matrice les ARN totaux extraits de cellules cultivées en milieu LM17 (milieu complet) ou en milieu MCD (Milieu Chimiquement Défini), à 30 et à 42°C. Dans toutes les expériences, le taux de transcrit obtenu a été normalisé par rapport à celui du gène *ldh* qui code un lactate déshydrogénase. Ce dernier a été considéré comme un bon gène de référence après avoir démontré que celui-ci était transcrit de manière constitutive dans toutes les conditions testées. Les expériences de qPCR ont été réalisées à partir de trois extractions d'ARN totaux indépendantes, réalisées en duplicat. Les taux relatifs de transcrits du gène *rgg0182*, dans les différentes conditions de température et de milieux, et dans trois phases de la croissance (exponentielle, transition et stationnaire) sont présentés dans la figure 28.

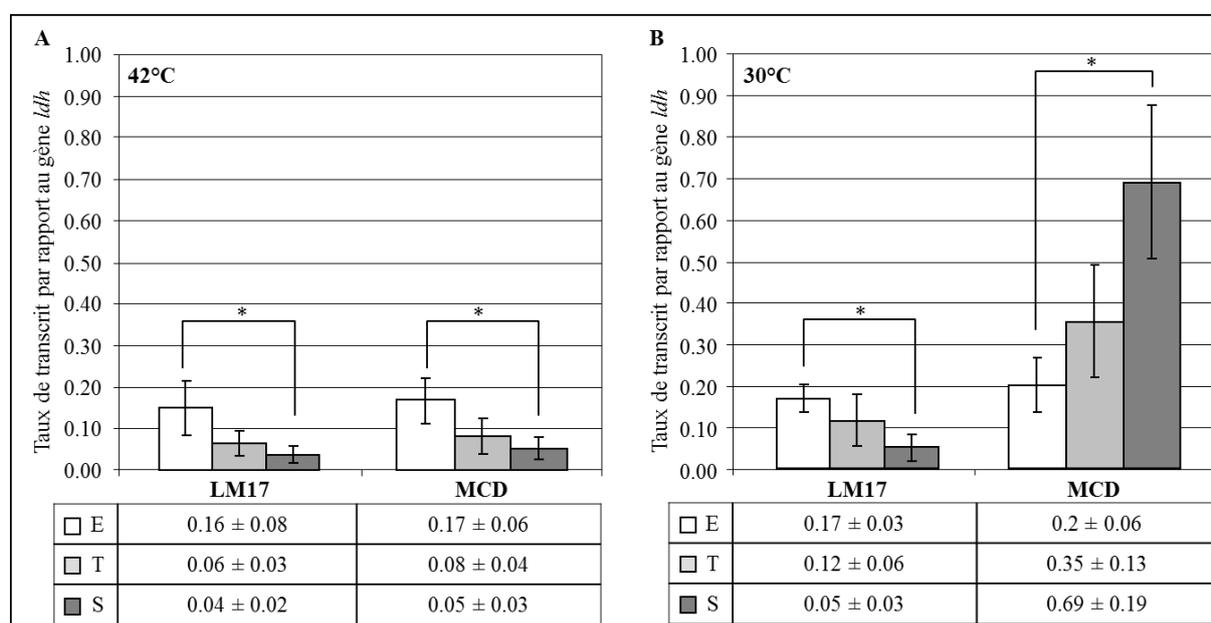


Figure 28 : Taux de transcrit relatif du gène *rgg0182* de *S. thermophilus* LMG18311 cultivée à 42°C (A) et à 30°C (B). Les ARN totaux ont été extraits de la souche sauvage en phase exponentielle (E, barres blanche), en phase stationnaire (T, barres gris clair) et en phase stationnaire de croissance (S, barres gris foncé). Les résultats correspondent à la moyenne ± la déviation standard de trois expériences indépendantes réalisées en duplicat. Test de Student *t*: *, $p < 0.001$.

A 42°C (figure 28A), quel que soit le milieu considéré (LM17 ou MCD) une diminution significative du taux de transcrit au cours de la croissance est observée. En effet, le taux de transcrit est d'environ 0,16 en phase exponentielle et diminue d'un facteur 3 à 4 en phase stationnaire. Ces résultats suggèrent, qu'à 42°C, le niveau de transcription du gène *rgg0182* de *S. thermophilus* est modulé en fonction de la phase de croissance mais qu'il n'est pas influencé par le milieu de culture.

Le taux de transcrit a, ensuite, été étudié à partir de cellules cultivées à 30°C (figure 28B), température à laquelle la bactérie *S. thermophilus* est souvent exposée lors des procédés industriels. En milieu LM17, à 30°C, le taux de transcrit est similaire à celui obtenu lors d'une croissance réalisée à 42°C. Au contraire, le taux de transcrit du gène *rgg0182* de cellules cultivées en milieu

MCD à 30°C est trois fois plus élevé, en phase stationnaire qu'en phase exponentielle de croissance. La comparaison du niveau de transcrits du gène *rgg0182* de cellules en phase stationnaire cultivées en milieu MCD indique qu'il est 14 fois supérieur à 30°C qu'à 42°C. Cela suggère que la transcription du gène *rgg0182* est influencée à la fois par le milieu de culture et par la température à laquelle *S. thermophilus* est soumise.

3.3. Le gène *rgg0182* de *S. thermophilus* LMG18311 code un régulateur transcriptionnel

3.3.1. Implication du gène *rgg0182* de *S. thermophilus* LMG18311 dans la transcription des gènes adjacents

Les données de la littérature indiquent que la plupart des gènes *rgg* codent des régulateurs transcriptionnels et régulent la transcription de gènes qui leur sont adjacents (Sulavik *et al.*, 1992 ; Lyon *et al.*, 1998 ; Fleuchot *et al.*, 2011). Afin de déterminer si le gène *rgg0182* est impliqué dans la régulation transcriptionnelle des gènes qui le flanquent, la mesure par qPCR du niveau de transcription des gènes *shp0182* et *pep0182* a été effectuée dans des contextes génétiques sauvage et mutant $\Delta rgg0182$. Néanmoins, cette stratégie n'a pas fonctionné, la petite taille des ORF *shp0182* et *pep0182* n'ayant pas permis de choisir des amorces efficaces pour les expériences de qPCR. Une stratégie alternative consistant en l'étude de l'activité des promoteurs des gènes *shp0182* et *pep0182* à l'aide de fusions transcriptionnelles, dans des contextes sauvages et $\Delta rgg0182$, a été choisie. Pour ce faire, des plasmides portant le promoteur des gènes *shp0182* et *pep0182* en amont des gènes *luxAB*, codant la luciférase, ont été construits et intégrés dans le gène *blp* des génomes des souches LMG18311 et $\Delta rgg0182$ de *S. thermophilus*. Ainsi, 4 souches ont été obtenues et ont été appelées LMG18311 *blp* :: $P_{pep0182}$ -*luxAB*, LMG18311 *blp* :: $P_{shp0182}$ -*luxAB*, $\Delta rgg0182$ *blp* :: $P_{pep0182}$ -*luxAB* et $\Delta rgg0182$ *blp* :: $P_{pep0182}$ -*luxAB*. Le niveau relatif de l'activité de ces deux promoteurs a ensuite été déterminé dans les conditions de milieu et de température définies précédemment. La figure 29 présente les résultats obtenus en milieu LM17 et la figure 30, ceux obtenus en milieu MCD.

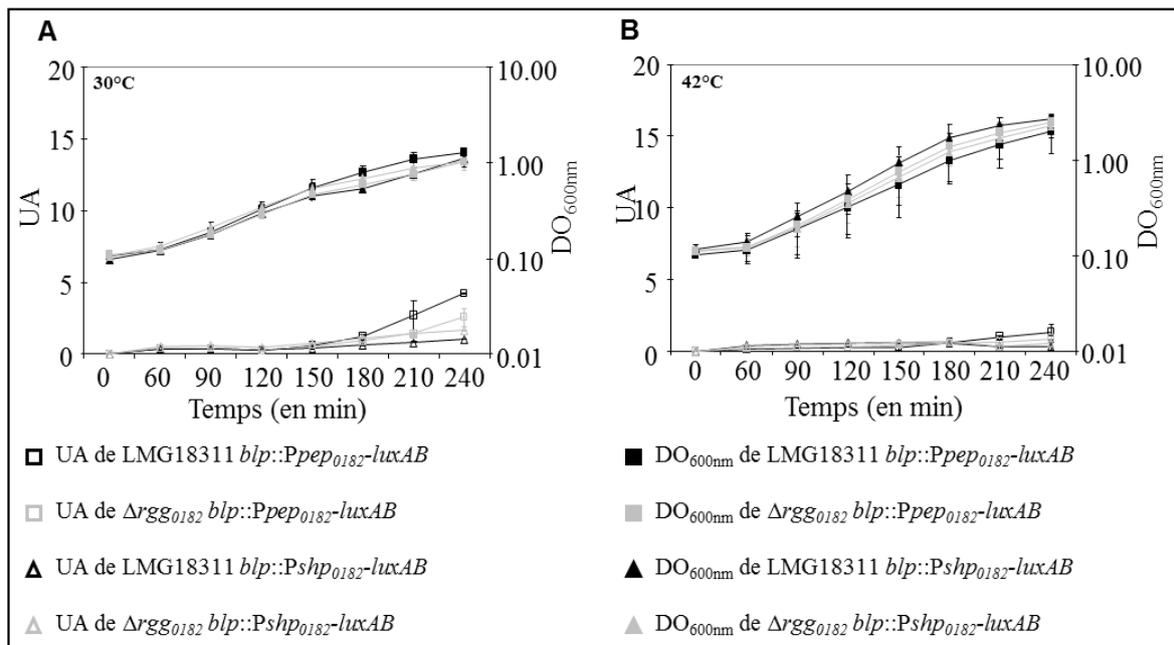


Figure 29 : Représentation de l'activité luciférase et du taux de croissance des souches LMG18311 et Δrgg_{0182} de *S. thermophilus*, contenant les fusions transcriptionnelles $P_{shp0182}$ -*luxAB* et $P_{pep0182}$ -*luxAB*, en milieu LM17. L'expression des fusions a été réalisée sur des cellules cultivées en milieu LM17 à 30°C (A) et à 42°C (B). Les résultats correspondent à la moyenne \pm la déviation standard de trois expériences indépendantes. UA : Unité Arbitraire de la luminescence normalisée par la DO à 600nm.

Lors d'une croissance réalisée en milieu LM17, à 30 ou à 42°C, pas ou peu d'activité luciférase n'est observée dans les contextes sauvage et mutant (figure 29). Ainsi, dans ces conditions, les promoteurs des gènes *shp0182* et *pep0182* sont inactifs, suggérant que ces gènes ne sont pas transcrits.

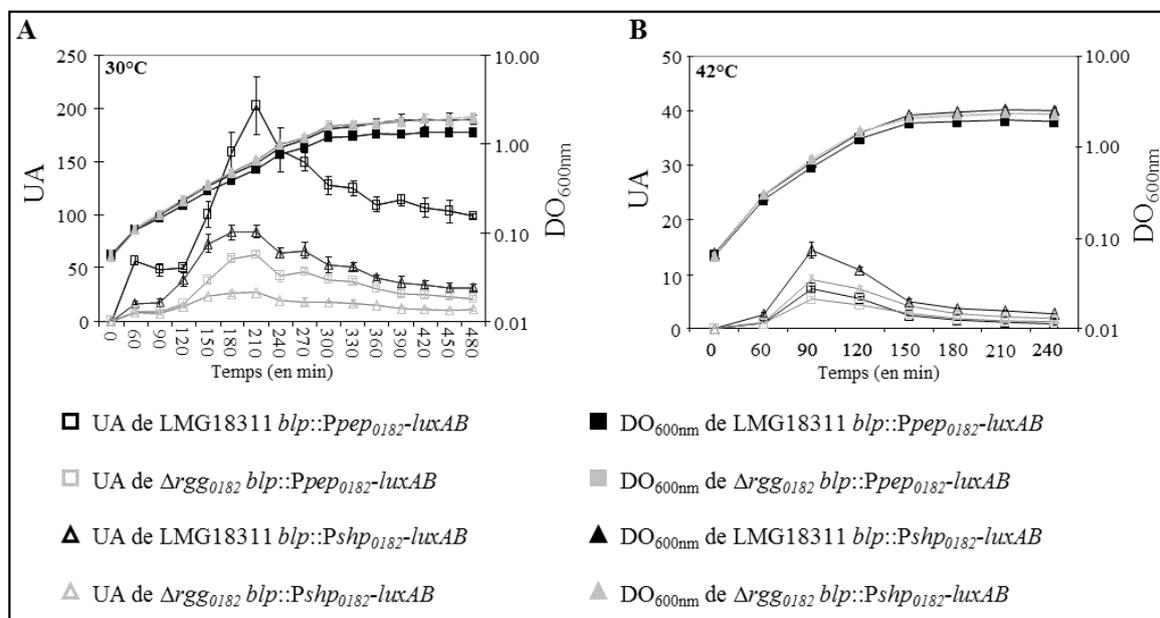


Figure 30: Représentation de l'activité luciférase et du taux de croissance des souches LMG18311 et Δrgg_{0182} de *S. thermophilus*, contenant les fusions transcriptionnelles $P_{shp0182}$ -*luxAB* et $P_{pep0182}$ -*luxAB*, en milieu MCD. L'expression des fusions a été réalisée sur des cellules cultivées en milieu MCD à 30°C (A) et à 42°C (B). Les résultats correspondent à la moyenne \pm la déviation standard de trois expériences indépendantes. UA : Unité Arbitraire de la luminescence normalisée par la DO à 600nm.

Au contraire, lors d'une croissance réalisée en milieu MCD, une activité luciférase est détectée à 30 et à 42°C (figure 30) pour toutes les souches étudiées. A noter que quelles que soient les souches étudiées, l'activité des promoteurs $P_{pep0182}$ et $P_{shp0182}$ est plus élevée à 30°C qu'à 42°C. Le maximum d'activité des fusions $P_{pep0182-luxAB}$ et $P_{shp0182-luxAB}$ est 28 et 6 fois supérieur à 30°C qu'à 42°C, respectivement ($p < 0,001$). De plus, l'activité luciférase de ces fusions transcriptionnelles diffère entre les souches LMG18311 et $\Delta rgg0182$ de *S. thermophilus*. En effet, lorsque les cellules sont cultivées en milieu MCD à 30°C, le maximum d'activité est trois fois moins important dans le contexte mutant que le contexte sauvage ($p < 0,001$) pour les deux constructions. Ces résultats démontrent ainsi que le gène *rgg0182* joue un rôle dans la régulation de la transcription des gènes *shp0182* et *pep0182*.

Par ailleurs, quelles que soient les conditions testées, aucune différence de croissance entre les souches n'est observée. En effet, les souches $\Delta rgg0182$ et LMG18311 présentant des paramètres de croissance (taux de croissance spécifique et biomasse finale) similaires.

3.3.2. La protéine Rgg₀₁₈₂ de *S. thermophilus* LMG18311 se fixe dans la région promotrice des gènes cibles

Les résultats précédents, ainsi que la présence d'un motif HTH dans la région N-terminale de la protéine Rgg₀₁₈₂ suggèrent que la protéine Rgg₀₁₈₂ puisse réguler l'activité transcriptionnelle des promoteurs $P_{shp0182}$ et $P_{pep0182}$ en se fixant sur leur région promotrice. Pour tester cette hypothèse, des expériences de retard de migration sur gel d'électrophorèse ont été réalisées. Pour se faire, la protéine Rgg₀₁₈₂ a été marquée par une queue histidine (His-Tag) et surproduite chez *E. coli*. Parallèlement, un fragment d'ADN contenant la région intergénique *rgg0182-shp0182* et un fragment d'ADN contenant la région intergénique *rgg0182-pep0182* (figure 31) ont été amplifiés et marqués à la biotine.



Figure 31 : Représentation schématique de la position des sondes $P_{shp0182}$ et $P_{pep0182}$, utilisé pour les expériences de retard de migration sur gel. Le fragment $P_{shp0182}$ de 126 pb contient les 72 pb de la région intergénique *rgg0182-shp0182* et une partie de la région 5' du gène *shp0182* et 5' du gène *rgg0182*. Le fragment $P_{pep0182}$ de 165 pb contenant les 150 pb de la région intergénique *rgg0182-pep0182* et une partie de la région 3' du gène *rgg0182* et 5' du gène *pep0182*.

Ensuite, 0,1 pmol de chacune des sondes marquées à la biotine ont été incubées avec 2 pmol de la protéine purifiée His₆-Rgg₀₁₈₂ afin de réaliser des expériences de retard de migration sur gel d'électrophorèse (figure 32).

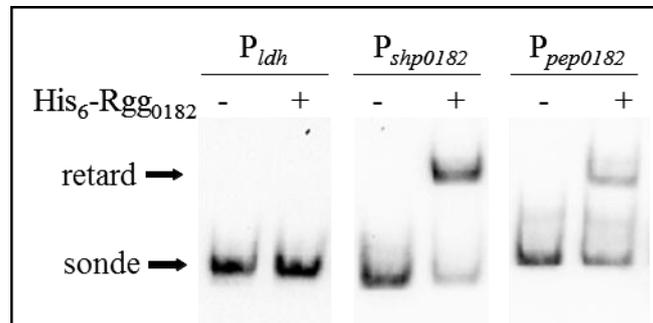


Figure 32 : Retard de migration sur gel de la région promotrice des gènes *shp*₀₁₈₂, *pep*₀₁₈₂ et *ldh* par fixation de la protéine His₆-Rgg₀₁₈₂. L'ADN (0,1 pmol) marqué à la biotine est incubé en absence ou en présence (2 pmol) de la protéine His₆-Rgg₀₁₈₂ purifiée. La sonde P_{ldh} correspond à la région promotrice du gène *ldh* et a été utilisée comme témoin négatif.

La protéine His₆-Rgg₀₁₈₂ retarde la migration des fragments d'ADN P_{shp0182} et P_{pep0182}. Pour vérifier que le retard de migration observé est dû à la fixation de la protéine His₆-Rgg₀₁₈₂ sur des fragments d'ADN, l'expérience a été réalisée avec la région promotrice du gène *ldh*. Ce dernier a été utilisé comme témoin négatif puisque sa transcription n'est pas régulée par la protéine Rgg₀₁₈₂. Dans ces conditions, aucun retard de migration sur gel n'est observé montrant que la fixation de la protéine Rgg₀₁₈₂ sur les régions promotrices des gènes *shp*₀₁₈₂ et *pep*₀₁₈₂ est spécifique.

Les résultats de fusions transcriptionnelles indiquent que la protéine Rgg₀₁₈₂ influence la transcription des gènes *shp*₀₁₈₂ et *pep*₀₁₈₂. De plus, les résultats de retard de migration sur gel d'électrophorèse montrent que le produit du gène *rgg*₀₁₈₂ est capable de se fixer sur la région promotrice de ses deux gènes adjacents, suggérant que celui-ci est impliqué dans la transcription de ceux-ci. Ainsi, l'ensemble de ces résultats montrent que la protéine Rgg₀₁₈₂ est un régulateur transcriptionnel qui peut agir directement sur la transcription de ses gènes cibles.

3.4. Implication du régulateur transcriptionnel Rgg₀₁₈₂ de *S. thermophilus* LMG18311 dans la transcription de gènes codant des chaperons et protéases

L'impact de la température sur la transcription du gène *rgg*₀₁₈₂ suggère un rôle de la protéine Rgg₀₁₈₂ de *S. thermophilus* LMG18311 dans l'adaptation aux changements de température. Ainsi, une hypothèse est que la protéine Rgg₀₁₈₂ puisse contrôler la transcription de gènes codant des

protéines de choc chaud, appelées Hsp (pour Heat shock protein) ou de choc froid, appelées Csp (pour Cold shock protein), incluant ainsi des chaperons et des protéases. Les chaperons et les protéases ATP-dépendantes jouent un rôle majeur dans la survie des bactéries soumises à des conditions où les protéines sont agrégées ou mal repliées. Les chaperons et les protéases interviennent dans la biogénèse des protéines en absence de stress. Cependant, celles-ci sont particulièrement induites en condition de stress froid ou chaud et plus généralement à différents stress.

Au sein du génome de la souche LMG18311 de *S. thermophilus* (Bolotin *et al.*, 2004), plusieurs gènes sont prédits pour coder des chaperons et des protéases majeures, impliquées dans la réponse aux stress thermiques et ont été sélectionnés pour l'analyse : *clpC*, *dnaK*, *dnaJ*, *hsp33*, *groES*, *groEL*, *clpP*, *clpX*, *clpE*, *clpL*, *cspA* et *cspB* (Genbank Accession NC_006448, locus tags *stu0077*, *stu0120-0121*, *stu0180*, *stu0203-0204*, *stu0356*, *stu0581*, *stu0602*, *stu1614*, *stu0837* et *stu0838*, respectivement). Le taux de transcrite de ces gènes a été mesuré par qPCR sur des cellules en phase stationnaire de croissance issues d'une culture réalisée sur les souches sauvage et Δrgg_{0182} en milieu MCD à 30°C (c'est-à-dire quand le gène *rgg_{0182}* est le plus transcrit). L'expérience a été réalisée à partir de 3 cultures indépendantes en duplicat (figure 33).

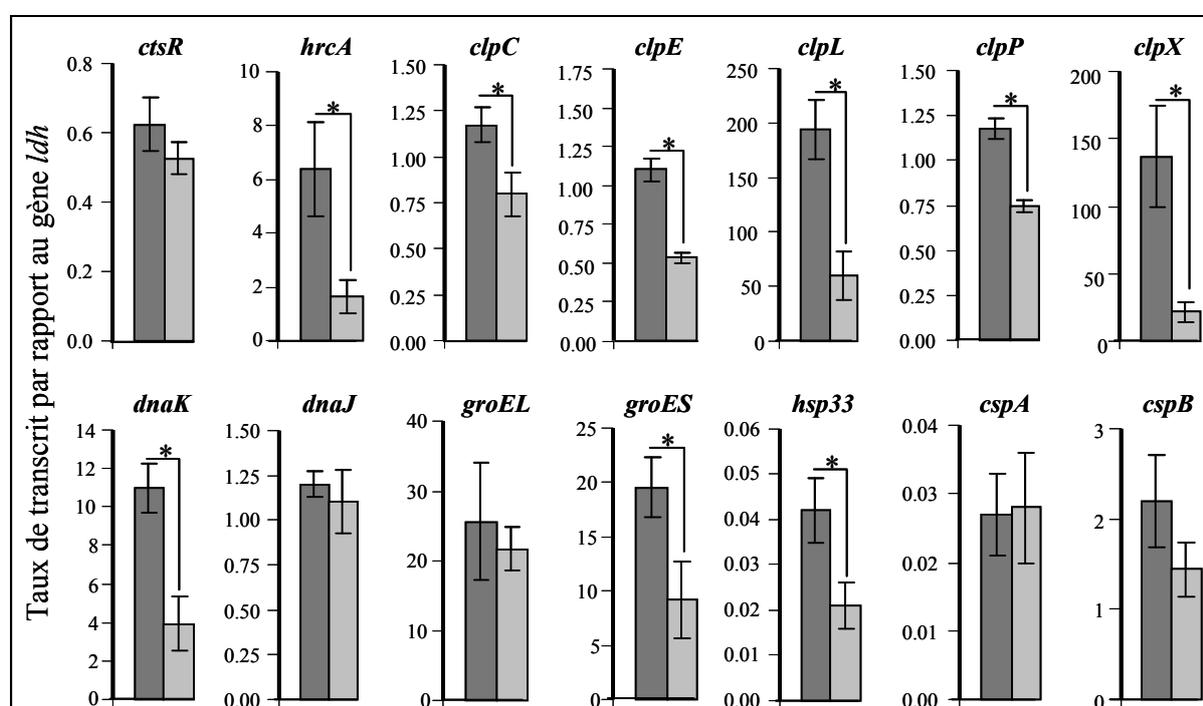


Figure 33 : Taux de transcrite relatif de gènes de *S. thermophilus* sur des cellules en phase stationnaire de croissance dont la croissance a été réalisée en milieu MCD à 30°C. Les ARN totaux ont été extraits de la souche sauvage (barre gris foncé) et du mutant Δrgg_{0182} (barres gris clair). Les résultats correspondent à la moyenne \pm la déviation standard de trois expériences indépendantes réalisées en duplicat. Test de Student *t*: *, $p < 0.001$.

Ces résultats montrent que, dans les conditions testées, le taux de transcrit de tous ces gènes (excepté *dnaJ*, *groEL*, *cspA* et *cspB*) sont sous-transcrit chez le mutant Δrgg_{0182} par rapport à la souche sauvage. En effet, la différence du taux de transcrit entre les deux souches varie de 1,5 à 20 fois, selon le gène considéré, de manière significative ($p < 0.001$). Comme décrit dans d'autres analyses transcriptionnelles chez les streptocoques, une différence de 1,5 fois du taux de transcrit entre deux souches est significative (Fernandez *et al.*, 2004 ; Zheng *et al.*, 2011).

Chez les bactéries à bas pourcentage en bases G et C, le contrôle de la transcription des gènes *clp* et des gènes *dnaK* et *groES* est médié par la fixation des répresseurs CtsR et HrcA, respectivement, à la région promotrice de ces gènes cibles. Ainsi une analyse de la séquence nucléotidique de ces régions promotrices a permis de mettre en évidence l'opérateur CtsR (AGGTCAAANANAGGTCAAA) décrit par Derre *et al.*, (1999) en amont des gènes *clpP*, *clpE*, *clpL*, *ctsR*, *clpC* et *groEL* et le site de fixation HrcA (TTAGCACTC(N)₉GAGTGCTAA) décrit par Zuber et Schuman (1994) seulement en amont des gènes *hrcA*, *groEL* (avec 2 mismatches) et *dnaJ* (6 mismatches). Ainsi, le taux de transcrit des gènes *ctsR* et *hrcA* (locus *stu0076* et *stu0118*, respectivement) a été analysé chez la souche sauvage et la souche Δrgg_{0182} . Les résultats présentés sur la figure 33 ne montrent pas de différence significative du taux de transcrit *ctsR* entre les deux souches. Au contraire, le gène *hrcA* est 4 fois moins transcrit chez le mutant comparé à la souche sauvage suggérant que sa transcription est sous la dépendance de la protéine Rgg₀₁₈₂.

3.5. Implication de la protéine Rgg₀₁₈₂ de *S. thermophilus* LMG18311 dans la réponse au stress thermique

3.5.1. Implication de la protéine Rgg₀₁₈₂ dans la réponse au stress chaud

Les données de la littérature indiquent que certaines protéines Rgg sont impliquées dans la réponse aux stress (Sanders *et al.*, 1998 ; Chaussee *et al.*, 2004 ; Fernandez *et al.*, 2004a ; Pulliainen *et al.*, 2008 ; Bortoni *et al.*, 2009). Par exemple, le gène *ropB* est impliqué dans la réponse aux stress thermique de *S. pyogenes* NZ131 (Chaussee *et al.*, 2004 ; Pulliainen *et al.*, 2008). D'autre part, les résultats précédents ont montré l'influence de la température sur la transcription du gène *rgg₀₁₈₂*. Ainsi, ces données suggèrent que le gène *rgg₀₁₈₂* pourrait être impliqué dans l'adaptation de *S. thermophilus* au choc thermique. En conséquences, la tolérance au stress chaud a été évaluée sur des cellules des souches sauvage et Δrgg_{0182} en phase stationnaire de croissance cultivées pendant 10h en milieu MCD (DO_{600nm} = 1,5) ou en milieu LM17 (DO_{600nm} = 2,5) avant (condition contrôle) et après 15, 30, 45 et 60 minutes d'incubation à 52°C (température limite de croissance de *S. thermophilus* LMG18311 en conditions de laboratoire). Cette expérience a été réalisée 3 fois de

manière indépendante et en triplicat. Les résultats obtenus en milieu LM17 sont présentés sur la figure 34 et ceux obtenus en milieu MCD sur la figure 35.

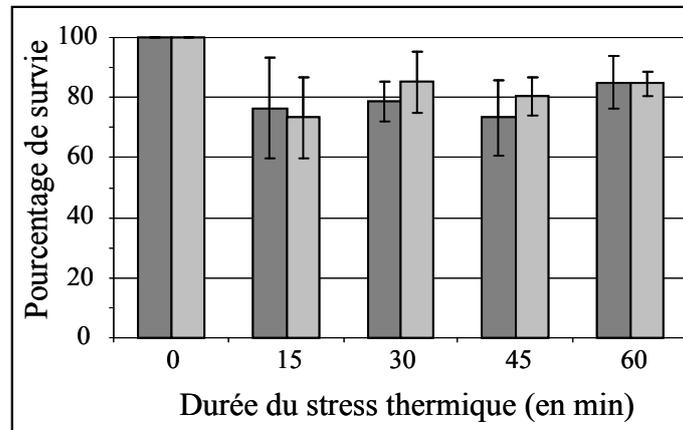


Figure 34 : Survie des souches sauvage et Δrgg_{0182} de *S. thermophilus* soumises à un choc thermique (0, 15, 30, 45 et 60 min à 52°C). La culture de *S. thermophilus* a été réalisée en milieu LM17 à 30°C puis exposé à un choc chaud. Le pourcentage de survie a été calculé par la formule $N / N_0 \times 100$ où N_0 est le nombre d'UFC obtenu dans la condition contrôle et N est le nombre d'UFC après exposition au stress. Les barres gris foncé correspondent à la souche sauvage et les barres gris clair correspondent à la souche Δrgg_{0182} . Les résultats correspondent à la moyenne \pm la déviation standard de 3 expériences indépendantes réalisées en triplicat. Test de Student t : *, $p < 0.001$.

En milieu LM17, quelle que soit la durée d'incubation à 52°C, l'exposition conduit seulement à une réduction d'environ 20% de la survie des deux souches testées. Ces résultats suggèrent donc que dans ces conditions le gène *rgg0182* n'est pas nécessaire à la survie de *S. thermophilus* au stress chaud.

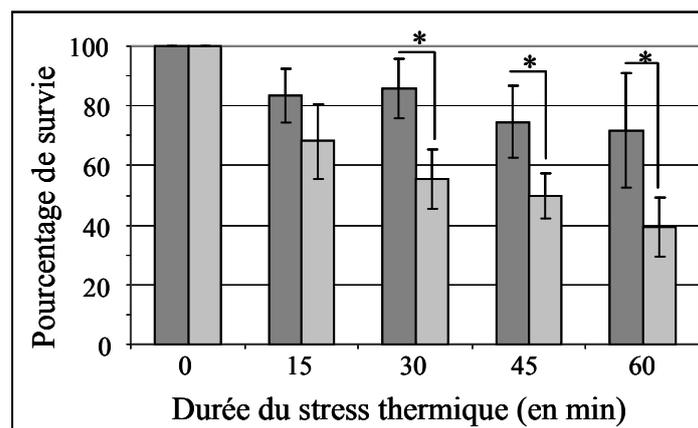


Figure 35 : Survie des souches sauvages et Δrgg_{0182} de *S. thermophilus* après un choc thermique (0, 15, 30, 45 et 60 min à 52°C). La culture de *S. thermophilus* a été réalisée en milieu MCD à 30°C puis exposé à un choc chaud. Le pourcentage de survie a été calculé par la formule $N / N_0 \times 100$ où N_0 est le nombre d'UFC obtenu dans la condition contrôle et N est le nombre d'UFC après exposition au stress. Les barres gris foncé correspondent à la souche sauvage et les barres gris clair correspondent à la souche Δrgg_{0182} . Les résultats correspondent à la moyenne \pm la déviation standard de 3 expériences indépendantes réalisées en triplicat. Test de Student t : *, $p < 0.001$.

Au contraire, les résultats présentés dans la figure 35, obtenus en milieu MCD, montrent qu'à partir de 30 min d'incubation à 52°C, la survie du mutant diffère de celle de la souche sauvage ($p < 0,001$). En effet, la tolérance au stress chaud de la souche Δrgg_{0182} diminue graduellement avec le temps d'exposition (72%, 53%, 46% et 38% de survie à 15, 30, 45 et 60 min, respectivement). De plus, une différence de survie est observée à 30, 45 et 60 min entre les deux souches. En effet, le mutant Δrgg_{0182} est plus sensible que la souche sauvage à ces trois temps d'incubation, suggérant un rôle du gène *rgg0182* dans l'adaptation au stress chaud de *S. thermophilus*.

3.5.2. Implication de la protéine Rgg₀₁₈₂ dans la transcription des gènes *clpE* et *cspB* suite à un stress chaud

La diminution du taux de survie du mutant Δrgg_{0182} suite à un changement de température suggère que la protéine Rgg₀₁₈₂ pourrait réguler la transcription de gènes de *S. thermophilus* impliqués dans la réponse au stress thermique, dans les conditions testées. Pour vérifier cette hypothèse, le taux de transcrit des gènes codant des chaperons et des protéases a été mesuré par qPCR. Les 14 gènes sélectionnés précédemment ont été choisis pour cette analyse. L'expérience a été réalisée trois fois de manière indépendante en duplicat sur des cellules des souches sauvage et Δrgg_{0182} cultivées en milieu MCD à 30°C puis exposées 30 min à 52°C (figure 36).

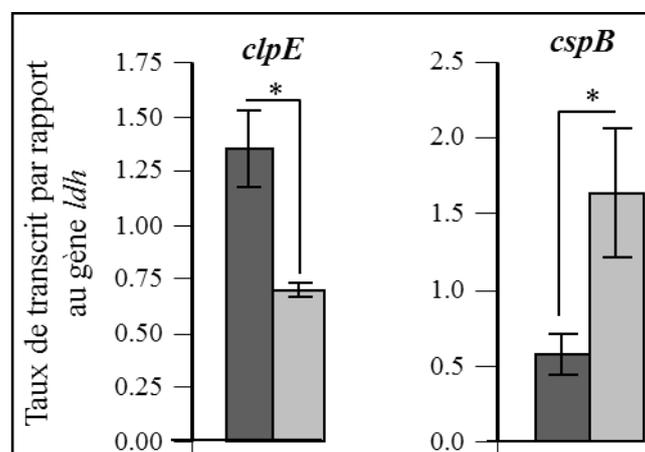


Figure 36 : Taux de transcrit relatif des gènes *clpE* et *cspB* de *S. thermophilus* dans des cellules exposés à un stress chaud. Les ARN totaux ont été extraits de la souche sauvage (barre gris foncé) et du mutant Δrgg_{0182} (barres gris clair) sur des cellules en phase stationnaire de croissance cultivées en milieu MCD à 30°C puis exposées 30 min à 52°C (condition de stress thermique). Les résultats correspondent à la moyenne \pm la déviation standard de trois expériences indépendantes réalisées en duplicat. Test de Student *t*: *, $p < 0,001$.

Les résultats montrent que les gènes *clpE* et *cspB* sont 2 fois moins et trois fois plus transcrits, respectivement, chez le mutant comparé à la souche sauvage ($p < 0,001$). Aucune différence n'a été observée pour les autres gènes étudiés. Ces résultats suggèrent que la protéine Rgg₀₁₈₂ régulerait

positivement la transcription du gène *clpE* et négativement la transcription du gène *cspB* en conditions de stress chaud à 52°C.

3.6. Effet du stress sur le taux de transcrit des chaperons et des protéases

Les données de la littérature indiquent que les gènes des régulateurs HrcA et CtsR sont réprimés par ces répresseurs en conditions normales de croissance. Au contraire, lorsque les bactéries sont soumises à une augmentation de leur température de croissance, l'inactivation des protéines HrcA et CtsR conduit à l'augmentation de la transcription des gènes codant des chaperons et des protéases. Chez *S. thermophilus*, aucune analyse du taux de transcription de tous ces gènes n'a été publiée. Ainsi, le taux de transcrits de ces gènes a été analysé par qPCR avant et après un stress de 30 minutes à 52°C. Les résultats sont présentés dans la figure 37.

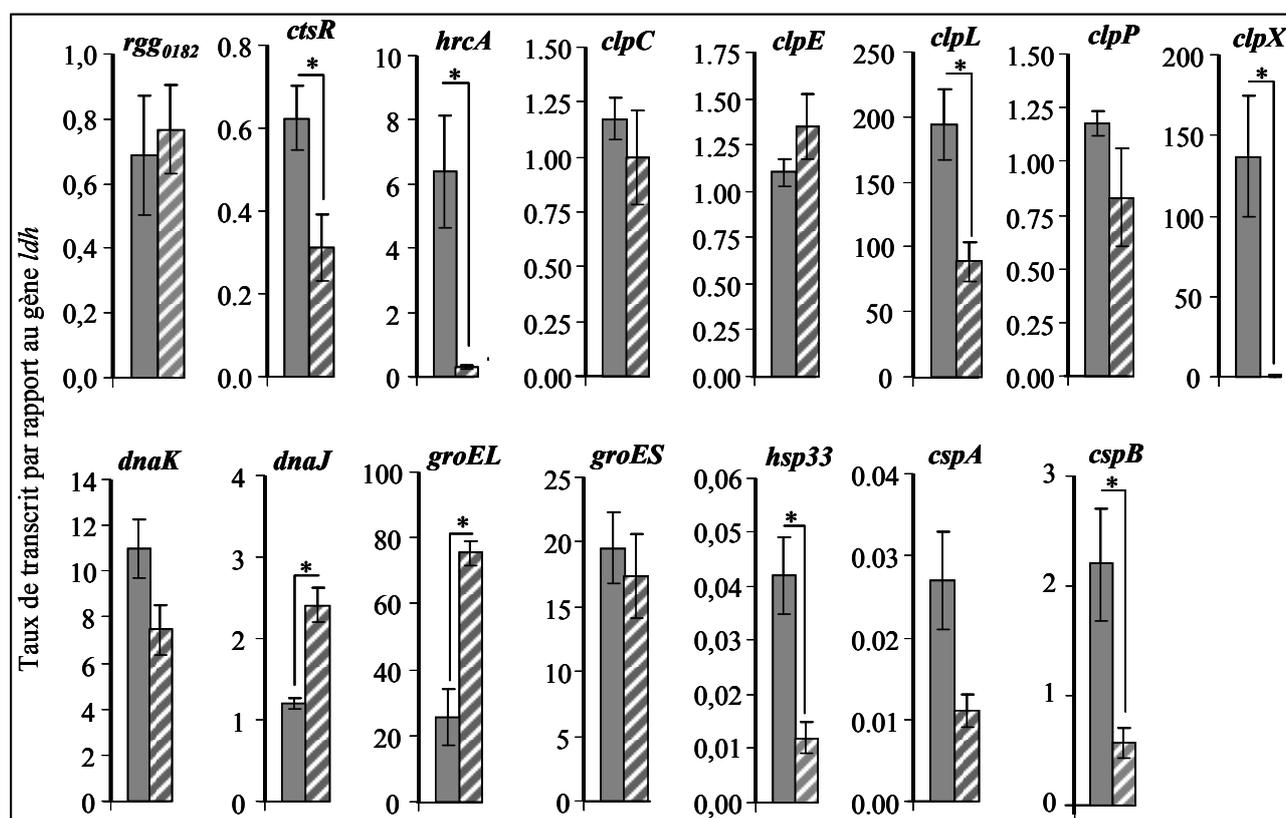


Figure 37 : Taux de transcrit relatif de gènes de *S. thermophilus* avant et après exposition des cellules à une augmentation brutale de la température. Les ARN totaux ont été extraits de la souche sauvage sur des cellules en phase stationnaire de croissance cultivées en milieu MCD à 30°C (barre grises) puis exposées 30 min à 52°C (barres grises hachurées). Les résultats correspondent à la moyenne ± la déviation standard de trois expériences indépendantes réalisées en duplicat. Test de Student *t*: *, $p < 0.001$.

Ces résultats montrent que le taux de transcrit des gènes *ctsR*, *hrcA*, *clpL*, *clpX*, *dnaJ*, *groEL*, *hsp33* et *cspB* sont modifiés après exposition des cellules à 52°C pendant 30 min. En effet, le taux

RESULTATS

de transcrit des gènes *ctsR* et *hrcA* est diminué d'un facteur 2 et 23, respectivement, en condition de stress chaud. Les gènes *dnaJ* et *groEL* sont, au contraire, sur-transcrits d'un facteur 2 et 3, respectivement. Enfin, le taux de transcrit des gènes *clpL*, *clpX*, *hsp33* et *cspB* sont sous-transcrits (de 2 à 240 fois) en condition de stress.

4. Implication du gène *rgg0182* de *S. thermophilus* LMG18311 dans l'hydrophobicité thermo-induite des cellules

Cette partie fait l'objet d'une publication actuellement en préparation.

4.1. Implication de la protéine Rgg₀₁₈₂ dans un phénomène d'adhésion cellulaire thermo-induite

Afin de préciser l'impact de la délétion du gène *rgg0182* sur la croissance de *S. thermophilus*, la croissance des souches sauvage et Δ *rgg0182* en milieu LM17 a été étudiée à 50°C. Les croissances ont été réalisées à l'aide d'un spectrophotomètre-incubateur à plaque (Bioscreen C) permettant de suivre la turbidimétrie de 100 micro-cultures simultanément. Les résultats obtenus à 42 et à 50°C sont présentés dans la figure 38.

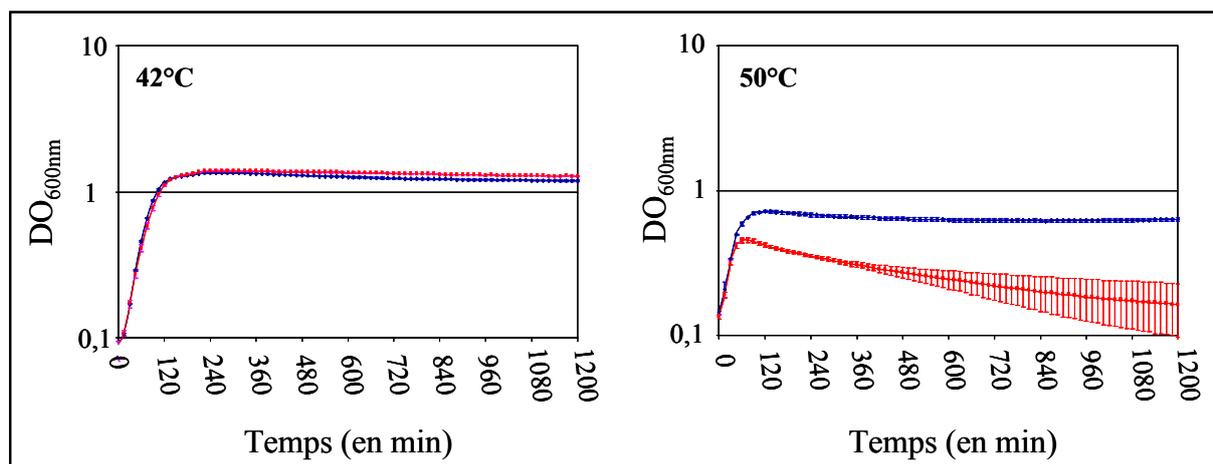


Figure 38 : Suivi de croissance des souches sauvage et Δ *rgg0182* de *S. thermophilus*, réalisé à l'aide du Bioscreen C, en milieu LM17 à 42°C et à 50°C. Les résultats correspondent à la moyenne \pm la déviation standard de 3 expériences indépendantes réalisées en triplicat.

A 42°C, aucune différence n'est observée entre les deux souches étudiées. Au contraire, à 50°C, la courbe de croissance du mutant Δ *rgg0182* se caractérise par un arrêt prématuré de la croissance suivi d'une chute progressive de la DO. Afin de déterminer si la diminution de la densité optique pourrait résulter d'une lyse des cellules du mutant Δ *rgg0182*, des échantillons des cultures des souches Δ *rgg0182* et LMG18311 ont été prélevés à différents temps et étalés sur boîte de pétri. Aucune différence significative de la survie n'a été observée entre les deux souches.

Cependant, l'observation attentive des micro-puits a révélé qu'à 50°C les cellules de la souche Δ *rgg0182* ne se déposaient pas uniformément au fond des micro-puits mais s'accrochaient progressivement sur les parois (figure 39). Ceci a pour conséquence de fausser la lecture de DO,

prise par le faisceau lumineux au centre des micro-puits.

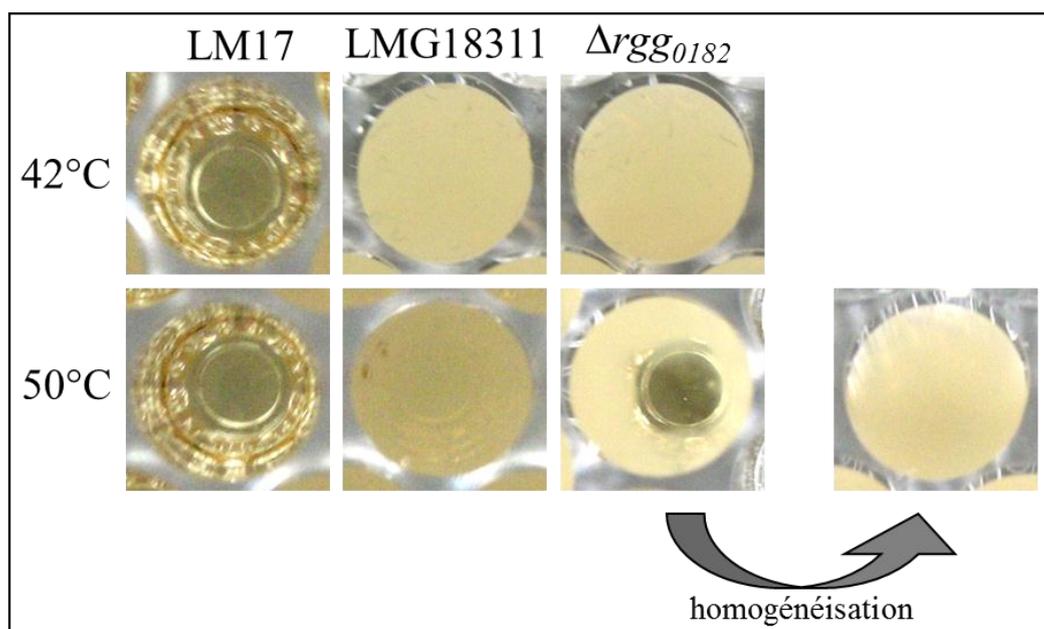


Figure 39 : Photographies prise du fond des micro-puits sur des cultures des souches LMG18311 et Δrgg_{0182} de *S. thermophilus*, réalisée en milieu LM17 à 42°C et à 50°C.

Afin de vérifier cette hypothèse, la densité optique de cultures en phase stationnaire de croissance en microplaques de la souche sauvage et du mutant, cultivées à 50°C, ont été mesurée avant et après homogénéisation ceci afin de disperser les cellules dans le micro-puits. Tandis que la DO de la souche sauvage est constante, celle de la souche Δrgg_{0182} est faible avant homogénéisation puis comparable aux valeurs obtenues pour la souche sauvage après homogénéisation. Cette information indique, d'une part, que la biomasse finale entre les deux souches est similaire lors d'une croissance réalisée en milieu LM17 à 50°C et, d'autre part, que les cellules de la souche Δrgg_{0182} commencent à adhérer aux parois des micro-puits dès qu'elles atteignent la fin de leur phase exponentielle de croissance.

Afin de mieux cerner la température à laquelle ce phénomène est induit des suivis de croissance de la souche sauvage et du mutant ont été réalisés à différentes températures, allant de 30 à 55°C. La DO/DO_0 (mesure permettant d'obtenir une valeur plus juste de la DO sans *à priori* de la DO initiale (DO_0)) obtenue aux différentes températures est représentée sur la figure 40.

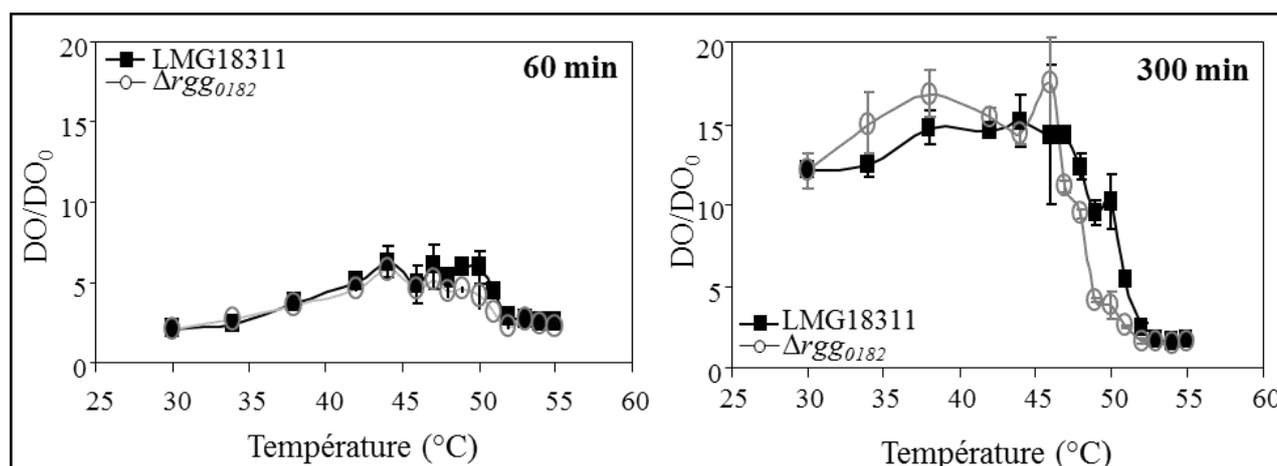


Figure 40 : Courbes représentant la DO/DO_0 obtenue à 60 et à 300 min en fonction des différentes températures étudiées des souches sauvage et Δrgg_{0182} de *S. thermophilus*. Les croissances ont été réalisées à l'aide du Bioscreen C à 30, 34, 38, 42, 44, 46, 47, 48, 49, 50, 51, 52, 53, 54 et 55°C. La DO_0 correspond à la DO initiale des cultures. Les résultats correspondent à la moyenne \pm la déviation standard de 3 expériences indépendantes réalisées en triplicat.

Ces résultats montrent qu'à 60 min, c'est-à-dire en phase exponentielle de croissance, aucune différence significative n'est observée entre les deux souches quelle que soit la température étudiée. En revanche, à 300 min, c'est-à-dire en phase stationnaire de croissance, les courbes présentent des divergences significatives pour des températures comprises entre 47 et 51°C. Au-delà de 52°C, les 2 courbes convergent et aucune croissance cellulaire n'est observée. Lorsque la température de croissance est comprise entre 47 et 51°C, les valeurs du rapport DO/DO_0 de la souche Δrgg_{0182} sont systématiquement plus faibles que celles de la souche sauvage avec une différence maximale observée à 50°C. Ces données indiquent que le phénomène d'adhésion cellulaire sur les parois du micro-puits de la souche Δrgg_{0182} est thermo-induit à des températures supérieures à 47°C.

Afin de vérifier si le phénomène d'adhésion cellulaire de la souche Δrgg_{0182} est également observé dans un plus grand volume, le suivi de la croissance a été étudié en tube de 50 mL (marque Falcon), en milieu LM17 et aux différentes températures choisies précédemment. Les résultats obtenus pour les croissances réalisées aux différentes températures choisies sont présentés dans les figures 41 et 42.

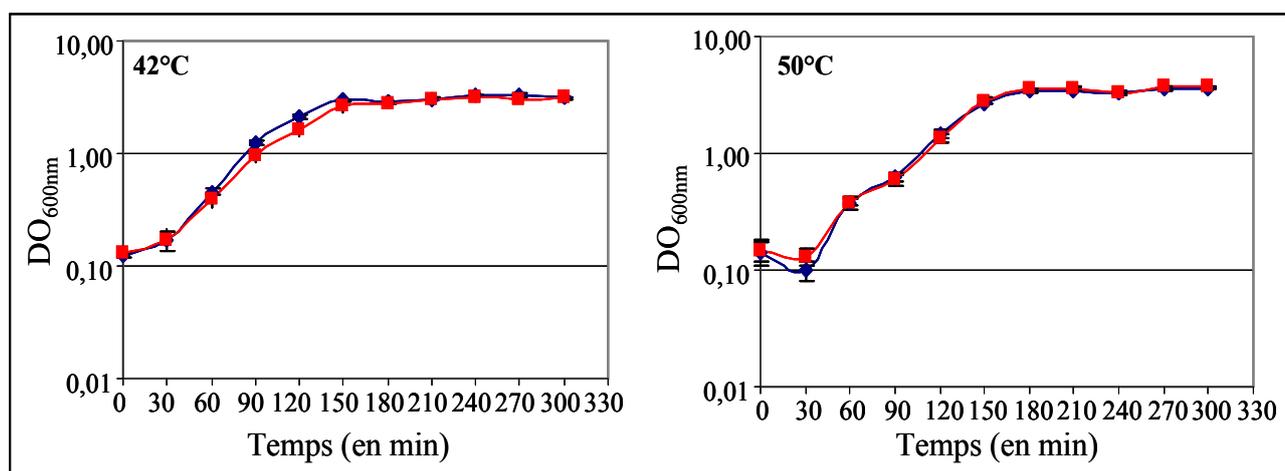


Figure 41 : Suivi de croissance des souches sauvage et Δrgg_{0182} de *S. thermophilus*, réalisé en tube de 50mL (marque Falcon), en milieu LM17 à 42°C et à 50°C. Les croissances ont été réalisées à 42 et à 50°C. Les résultats correspondent à la moyenne \pm la déviation standard de 3 expériences indépendantes réalisées en triplicat.

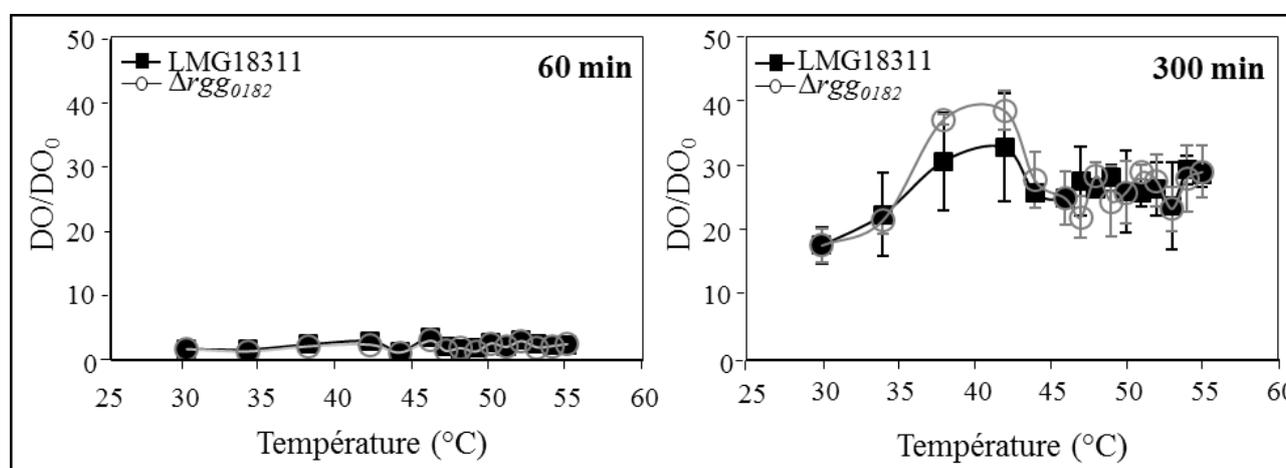


Figure 42 : Courbes représentant la DO/DO_0 obtenue à 60 et à 300 min en fonction des différentes températures étudiées des souches sauvage et Δrgg_{0182} de *S. thermophilus*. Les croissances ont été réalisées à en tube de 50 mL (marque Falcon) à 30, 34, 38, 42, 44, 46, 47, 48, 49, 50, 51, 52, 53, 54 et 55°C. La DO_0 correspond à la DO initiale des cultures. Les résultats correspondent à la moyenne \pm la déviation standard de 3 expériences indépendantes réalisées en triplicat.

Les résultats obtenus montrent qu'aucune différence n'est observée entre la souche sauvage et le mutant Δrgg_{0182} quelle que soit la température étudiée. Ces données indiquent que le phénomène d'adhésion cellulaire de la souche Δrgg_{0182} ne se produit pas dans les conditions testées. Des analyses similaires ont été réalisées en tube de 5 mL (tube à hémolyse) et aucune adhésion cellulaire n'a été observée.

Deux paramètres diffèrent entre les expériences réalisées en Bioscreen C et en tubes (de 5 ou 50 mL) : le volume des cultures et la nature du plastique des contenants. Les microplaques utilisées en Bioscreen C et les tubes de marque Falcon sont constituées de polystyrène de nature différente alors

que les tubes à hémolyse sont faits de polypropylène.

Une hypothèse pour expliquer l'adhésion cellulaire thermo-induite des cellules du mutant sur le polystyrène des parois des microplaques est que la fixation des cellules du mutant Δrgg_{0182} aux parois des micro-puits soit favorisée par des liaisons électrostatiques. Pour le vérifier, des expériences de mobilité électrophorétique ont été réalisées sur des cellules en phase stationnaire, issues d'une croissance réalisée en milieu LM17 à 50°C des souches LMG18311 et Δrgg_{0182} de *S. thermophilus*. Néanmoins, les résultats montrent qu'aucune différence n'est observée entre les deux souches. Les données obtenues suggèrent ainsi que les cellules du mutant Δrgg_{0182} n'adhèrent pas aux parois des microplaques par des interactions électrostatiques. Une autre hypothèse pour expliquer l'adhésion thermo-induite des cellules sur du polystyrène est l'adhésion par interaction hydrophobe.

4.2. Caractérisation des forces hydrophobes sur les parois des microplaques

Afin de caractériser la surface des cellules de la souche sauvage et du mutant Δrgg_{0182} de *S. thermophilus*, la stratégie a consisté en l'utilisation de la microscopie à force atomique (AFM) qui permet de caractériser les propriétés physicochimiques et mécaniques de surfaces et de localiser et de quantifier certaines molécules par interaction spécifique récepteur-ligand (figure 43).

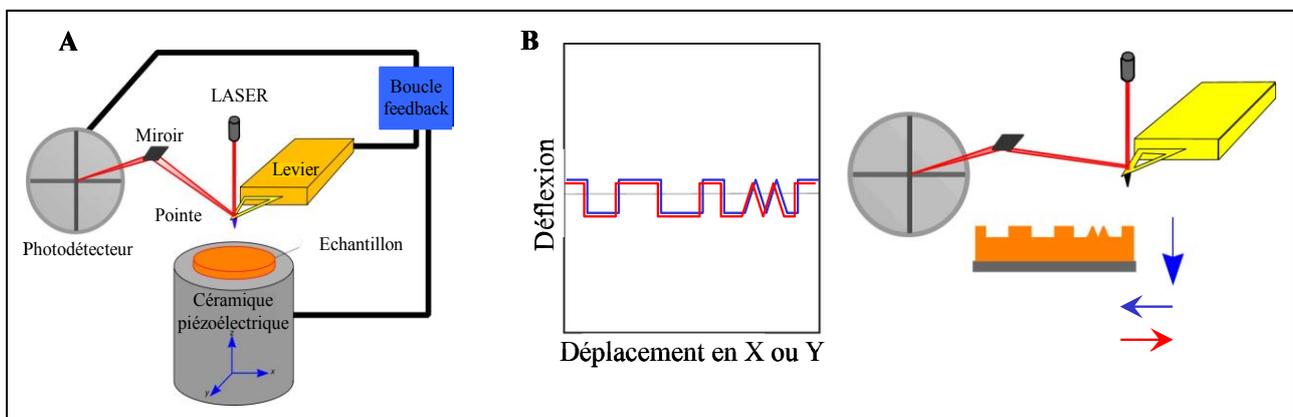


Figure 43 : Représentation schématique de l'AFM. A. Représentation des composants de l'AFM. B. Représentation du fonctionnement de l'AFM (Grégory Francius). La céramique piézoélectrique est l'endroit où l'échantillon est déposé, après préparation, et qui permet son balayage en 3 dimensions (avec une précision de déflexion de l'ordre de 0,1 nm et un déplacement latéral (X, Y) d'une précision de l'ordre de 1 nm). La pointe AFM est ancrée au levier pour scanner l'échantillon. Un faisceau est envoyé au bout de la pointe AFM puis réfléchi, *via* un miroir, vers un photodétecteur qui transmet les informations à un logiciel de traitement des données (Igor6). Ce fonctionnement permet ainsi de réaliser de l'imagerie 2D ou 3D. L'analyse et la précision des images AFM permet la détermination de la rugosité des surfaces et des distances caractéristiques quand la surface est couverte de motifs.

L'AFM permet d'étudier les caractéristiques physicochimiques de surfaces par une technique appelé force-volume. Cette technique génère des courbes de force basées sur l'interaction entre une pointe AFM très fine et la surface d'un échantillon. Elle permet également de visualiser à l'échelle moléculaire la topographie des surfaces et de mesurer point par point les forces d'adhésion à la surface de l'échantillon étudié. La technique d'imagerie force-volume par AFM permet de dresser des cartographies bidimensionnelles et tridimensionnelles des forces sur une surface mais aussi de localiser des molécules et d'obtenir un grand nombre de propriétés physico-chimiques propres à ces dernières (longueur de contour, élasticité, force spécifique d'adhésion, force de détachement...). Cela permet, par exemple, de différencier des surfaces adhésives de surfaces non adhésives.

Dans un premier temps, la technique AFM utilisant des pointes hydrophobes ou CFM (Chemical Force Microscopy) a été utilisée pour étudier la surface des microplaques et des tubes Falcon. Les images topographiques obtenues ainsi que les mesures d'hydrophobicité sont présentés dans les figures 44 (microplaques) et 45 (tubes Falcon).

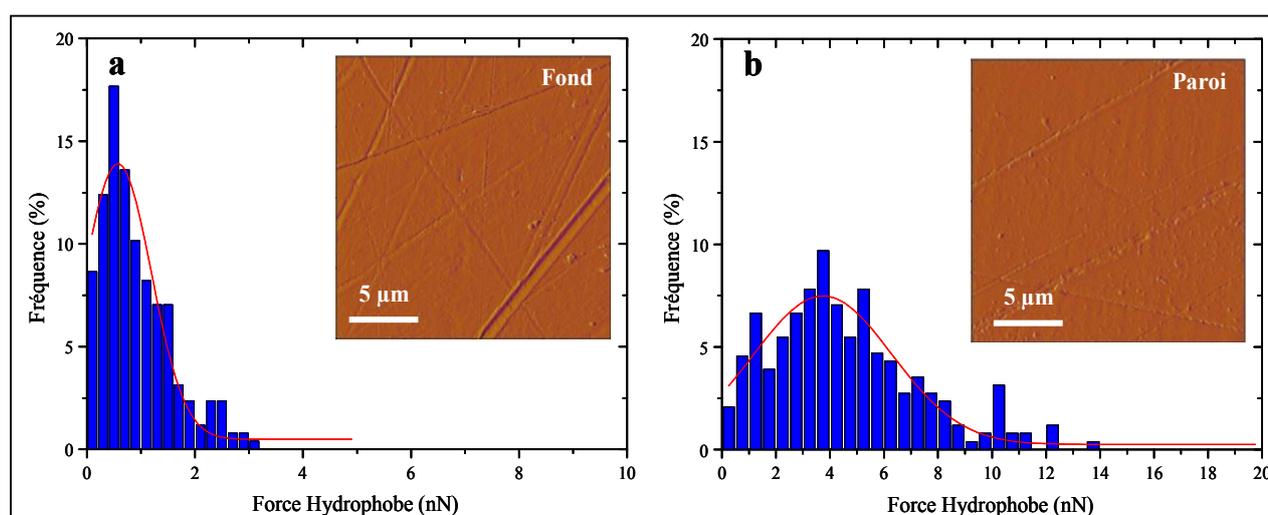


Figure 44 : Imagerie et mesure de la force hydrophobe réalisées à l'aide de l'AFM sur le fond (a) et la paroi (b) des micro-puits des microplaques utilisées pour les croissances en Bioscreen C.

Pour les microplaques, les parois et le fond des micro-puits ont été analysés. Le fond et les parois sont lisses avec des rugosités différentes d'environ 24 nm pour le fond et de de 18 nm pour les parois. Par ailleurs, le fond des micro-puits présente une hydrophobicité faible avec une force d'adhésion de l'ordre de $0,57 \pm 0,06$ nN. En revanche, la paroi possède des caractéristiques hydrophobes plus marquées avec une force d'adhésion moyenne de $3,74 \pm 0,17$ nN. Les différences observées entre le fond et la paroi pourraient expliquer pourquoi, dans certaines conditions, les

cellules du mutant Δrgg_{0182} acquièrent des propriétés d'adhésion aux parois plus importantes.

A partir de ces différents résultats, l'hypothèse qui peut être suggérée pour expliquer ces phénomènes repose sur le fait qu'à partir de la fin de la phase exponentielle, la surface des cellules du mutant Δrgg_{0182} acquièrent des propriétés hydrophobes plus importantes leur permettant d'adhérer de manière préférentielle aux parois plutôt que dans le fond des micro-puits.

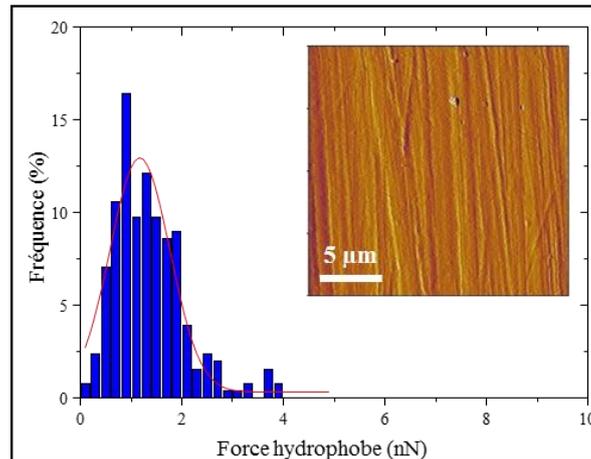


Figure 45 : Imagerie et mesure de la force hydrophobe réalisées à l'aide de l'AFM sur les tubes de marque Falcon.

Les mesures réalisées par AFM sur les tubes Falcon révèlent que ceux-ci sont striés de façon régulière avec une périodicité de l'ordre de 100-200 nm, présentent une rugosité de 29 nm (valeur proche de celle mesurée au fond des microplaques) pour une force d'adhésion hydrophobe de $1,18 \pm 0,06$ nN.

Par ailleurs, la rugosité est plus faible sur les parois des micro-puits par rapport à la surface du tube de marque Falcon ou encore le fond des micro-puits. De plus de nombreuses études ont montré que dans certaines conditions, la rugosité des surfaces a plutôt tendance à favoriser l'adhésion des cellules bactériennes (Kerr et J.Cowling, 2003). Les mesures de rugosité effectuée sur les microplaques et le tube de marque Falcon tendent à montrer que ce paramètre n'a pas d'influence notable sur l'adhésion des cellules puisque celles-ci adhèrent de préférence sur les parois qui sont moins rugueuses que le fond des micro-puits.

4.3. Analyse de l'hydrophobicité des cellules

Afin de caractériser la surface des cellules de la souche sauvage et du mutant Δrgg_{0182} de *S. thermophilus*, des images par AFM ont été réalisées. La figure 46 présente les résultats

d'imagerie des surfaces obtenue par AFM de la souche sauvage et du mutant.

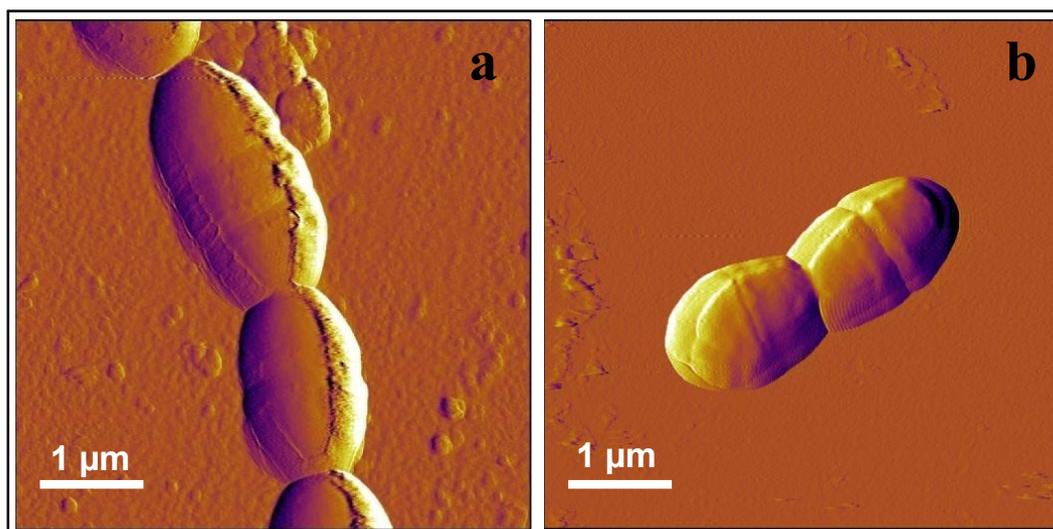


Figure 46 : Imagerie réalisée par AFM sur des cellules de la souche LMG18311 (a) et Δrgg_{0182} (b) de *S. thermophilus*.

Les deux souches ne présentent pas de différences significatives en termes de rugosité de surface. Les cellules sont lisses, ce qui suggère que soit, elles présentent à leur surface très peu ou pas d'exopolymères soit, la frange polymérique qui recouvre la paroi est suffisamment dense et compacte pour maintenir une faible rugosité. Par ailleurs, la mesure de la taille des cellules montre que celles de la souche sauvage ont une taille comprise entre 2 et 2,5 μm alors que celles de la souche Δrgg_{0182} mesurent entre 1,5 et 2 μm . Ainsi, les cellules du mutant Δrgg_{0182} sont légèrement plus petites que celles de la souche sauvage. Ceci suggère que la délétion du gène *rgg0182* aurait pu conduire à un arrêt prématuré de l'élongation de la paroi latérale des cellules.

Afin de caractériser plus en détails la présence de polymères sur la surface des cellules, la technique force-volume réalisée par AFM a été envisagée. Cependant, à l'heure actuelle, cette perspective n'a pas donné de résultats concluant du fait que les cellules bactériennes doivent impérativement être fixées sur un support sans subir de traitement chimique ou thermique pouvant modifier leur état physiologique et par conséquent leurs propriétés physico-chimiques à l'état natif.

Une autre stratégie est envisagée afin de vérifier cette hypothèse. Celle-ci est basée sur la technique d'adhésion microbienne à des solvants (ou MATS pour Microbial Adhesion To Solvent) (Bellon-Fontaine *et al.*, 1996). Ces analyses sont en cours d'expérimentation.

DISCUSSION

III – DISCUSSION ET PERSPECTIVES

1. Comment fonctionnent les protéines Rgg ?

1.1. Comment les protéines Rgg contrôlent-elles l'expression de leurs cibles ?

Dans la littérature, tous les gènes *rgg* décrits codent des régulateurs transcriptionnels et régulent, notamment, les gènes qui leur sont adjacents (Sulavik *et al.*, 1992 ; Lyon *et al.*, 1998 ; Qi *et al.*, 1999 ; Ibrahim *et al.*, 2007b).

Les données obtenues au cours de ce travail ont montré que le gène *rgg₀₁₈₂* n'échappe pas à cette règle puisque la protéine qu'il code contrôle la transcription des ORF *shp₀₁₈₂* et *pep₀₁₈₂* qui lui sont directement adjacents. La protéine Rgg₀₁₈₂ contrôle également la transcription de gènes codant des chaperons et des protéases. Ainsi, le produit du gène *rgg₀₁₈₂* est capable de réguler non seulement les gènes qui lui sont adjacents mais également de nombreux gènes disséminés sur le génome. Cependant, cette étude n'est pas exhaustive et il n'est pas exclu que la protéine Rgg₀₁₈₂ puisse contrôler l'expression d'autres gènes cibles.

Dans la majorité des cas, la délétion du gène *rgg₀₁₈₂* se traduit par une diminution de la quantité de transcrit de ses cibles indiquant ainsi que la protéine Rgg₀₁₈₂ est un activateur de la transcription. Néanmoins, la protéine Rgg₀₁₈₂ est également capable de réprimer la transcription de gènes, comme le gène *cspB* par exemple. La protéine Rgg₀₁₈₂ serait donc à la fois un activateur et un répresseur. Samen *et al.* (2006) ont montré que la protéine RovS de *S. agalactiae* NEM316 était capable de contrôler positivement la transcription des gènes *gbs0230*, *sodA* et *cylE* et négativement celle du gène *fbsA*.

Par ailleurs, des expériences de retard de migration sur gel ont permis de montrer que la protéine Rgg₀₁₈₂ contrôle directement la transcription des gènes *shp₀₁₈₂* et *pep₀₁₈₂* en se fixant sur leur région promotrice. L'hypothèse la plus probable est que les protéines Rgg se fixent à l'ADN *via* leur motif HTH. Loughman et Caparon (2007) ont montré que la substitution du résidu G4 en un résidu A ou du résidu R11 en un résidu A dans la protéine RopB de *S. pyogenes* NZ131 entraîne une diminution de 5 fois et une inactivation, respectivement, de la capacité de la protéine à activer la transcription du gène *speB*. L'étude de la capacité de fixation à l'ADN de la protéine Rgg₀₁₈₂ avec des substitutions des résidus conservés dans le domaine HTH, ou même la délétion de ce domaine, permettrait de valider ou d'invalider cette hypothèse.

DISCUSSION

Plusieurs études ont montré que des protéines Rgg se fixent dans la région promotrice de leurs gènes cibles (Chaussee *et al.*, 2002 ; Vickerman et Minick, 2002 ; Neely *et al.*, 2003 ; Rawlinson *et al.*, 2005 ; Samen *et al.*, 2006 ; Anbalagan *et al.*, 2011 ; Fleuchot *et al.*, 2011). Ces études ont permis de proposer des sites de fixation spécifique à l'ADN. Cependant, ces données indiquent également que chacune des protéines Rgg étudiées reconnaissent une séquence d'ADN qui leur est propre. En effet, le site de fixation de la protéine LasX dans la région promotrice du gène *lasA*, en position -40/-70, est une séquence répétée directe de 7 nucléotides (TTATCCC) (Rawlinson *et al.*, 2005). La protéine RopB se fixe au niveau d'une séquence répétée inversée (GTTAGGCA) dans la région promotrice du gène *speB* (Neely *et al.*, 2003). La séquence de fixation consensus « AWAAWVHTDAW-N6/7-WTKWWAMDWAK » a, par ailleurs, été proposée pour la protéine RovS (Samen *et al.*, 2006). Concernant la protéine Rgg₀₁₈₂ de *S. thermophilus*, sa fixation dans la région promotrice des gènes *shp*₀₁₈₂ et *pep*₀₁₈₂ a permis de rechercher des sites potentielles de fixation (figure 47).

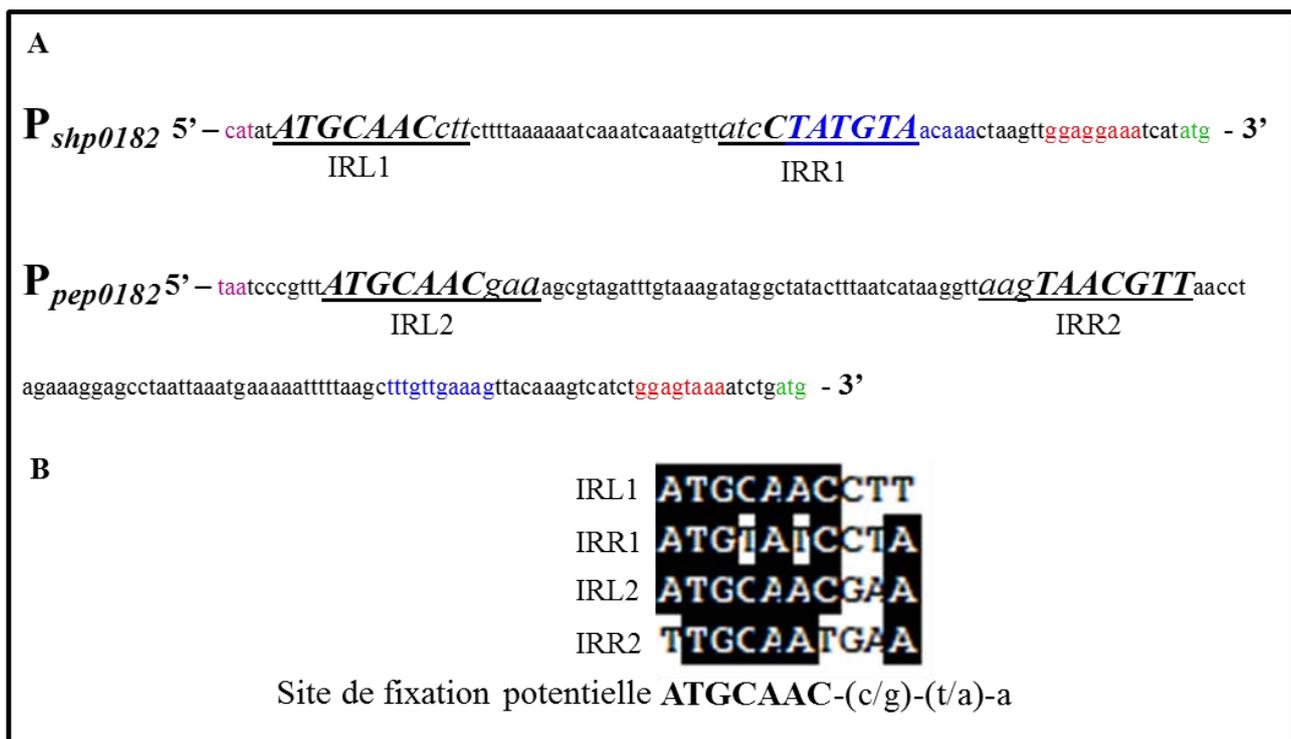


Figure 47 : Schématisation des sites potentiels de fixation de la protéine Rgg₀₁₈₂ dans la région promotrice des gènes *shp*₀₁₈₂ et *pep*₀₁₈₂. (A) Positionnement des IR sur la séquence nucléotidique des régions promotrices P_{*shp0182*} et P_{*pep0182*}. (B) Alignement des IR. Les séquences indiquées en rose correspondent aux codons stop, en bleu les boîtes -10 étendue (GNTTNTATAAT), en rouge les RBS (GGAGGAAA) et en vert les codons initiateurs. Les séquences soulignées correspondent aux IR. Le site potentiel de fixation proposé possède une partie très conservée (en majuscule) et une partie moins conservée (en minuscule). IRL : Inverted Repeat Left ; IRR : Inverted Repeat Right.

Des expériences d'empreinte à la DNase I pourront être réalisées pour vérifier le site de fixation

de la protéine Rgg₀₁₈₂.

Enfin, une hypothèse pour expliquer l'absence d'un site de fixation commun à toutes les protéines Rgg est la diversité des motifs HTH. En effet, bien que le domaine HTH soit présent dans toutes les protéines Rgg, celui présente des divergences d'une protéine Rgg à une autre. La diversité des motifs HTH des protéines Rgg suggère qu'ils puissent reconnaître des sites de fixation différents. De plus, le régulon des protéines Rgg semble être différent. En effet, pour exemple, les résultats obtenus au cours de ce travail montrent que parmi les gènes *rgg* de *S. thermophilus*, seul le gène *rgg₀₁₈₂* est impliqué dans l'organisation des cellules en chaînes. Pour vérifier cette hypothèse, la fixation de différentes protéines Rgg sur différents fragments d'ADN pourra être étudiée par expérience de retard de migration sur gel. Les sites de fixation des différentes protéines Rgg sera ensuite comparée

Par ailleurs, la régulation exercée par les protéines Rgg, positive ou négative, pourrait être directement liée au positionnement de leur site de fixation dans la région promotrice du gène cible comme cela a par exemple été démontré dans le cas du régulateur CcpA de *B. subtilis*. Ce régulateur se fixe au niveau de l'ADN sur un site de fixation appelé *cre*. Un site *cre* localisé en amont de la région -35 engendre une activation de l'expression du gène cible alors que si celui-ci est placé en aval de la boîte -35 ou entre les boîtes -10 et -35 du promoteur, une répression est observée (Kim *et al.*, 2002).

Néanmoins, des expériences préliminaires ne montrent aucune fixation de la protéine Rgg₀₁₈₂ sur le promoteur des gènes *clpE* et *cspB* (communication personnelle, Emmanuelle Bruneau) suggérant alors que la régulation de la transcription de ces gènes par la protéine Rgg₀₁₈₂ serait indirecte. De plus, le site de fixation potentiel proposé dans la figure 47 n'est pas retrouvé dans la région promotrice des gènes *clpE* et *cspB*. Ainsi, la protéine Rgg₀₁₈₂ serait capable de contrôler la transcription de gènes *via* d'autres régulateurs transcriptionnels comme cela a été démontré dans le cas de la protéine RopB de *S. pyogenes* NZ131 (Chaussee *et al.*, 2002). Parmi les régulateurs transcriptionnels contrôlés par la protéine Rgg₀₁₈₂, certains pourraient être des activateurs et permettraient, par exemple, d'induire la transcription du gène *clpE* et d'autres pourraient être des répresseurs et inhiberaient la transcription du gène *cspB*.

L'ensemble de ces données suggère l'existence de régulons Rgg distincts, bien que des chevauchements soient envisageables. Pour vérifier cette hypothèse, la comparaison des régulons, analysés par transcriptomique, de différentes protéines Rgg pourra être effectuée.

1.2. Quelle est la forme active des protéines Rgg ?

Au sein des protéines Rgg, un second domaine, composé de 18 nucléotides, est présent dans la région centrale de la protéine. Celui-ci n'est pas reconnu par les banques de données et a été appelé domaine médian. Loughman et Caparon (2007) ont mis en évidence que la protéine Rgg, appelé RopB, de *S. pyogenes* NZ131 était capable de se dimériser, comme la plupart des régulateurs transcriptionnels (Luscombe *et al.*, 2000). Les auteurs font l'hypothèse que le résidu W₁₅₆, fortement conservé (plus de 95% des séquences primaires des copies Rgg étudiées au cours de ce travail), pourrait intervenir dans l'interaction Rgg-protéine. En effet, la substitution du résidu W₁₅₆ par un résidu A entraîne une instabilité et une diminution de la capacité de régulation de la protéine RopB de *S. pyogenes* (Loughman et Caparon, 2007). De plus, des études ont montré que les résidus tryptophanes sont souvent impliqués dans des interactions protéine-protéine (Ma *et al.*, 2003).

Ainsi, une hypothèse est que le domaine médian pourrait intervenir dans l'activation (ou l'inactivation) de la protéine en permettant sa multimérisation. L'oligomérisation des protéines Rgg pourrait leur permettre d'adopter une conformation spécifique qui favoriserait leur interaction avec l'ADN. La présence d'IR ou de DR dans la région promotrice des gènes cibles des protéines Rgg est un argument en faveur de la fixation de ces protéines sous forme dimérique. En effet, chacune des deux protéines formant le dimère pourraient se fixer à une des séquences répétées. Pour vérifier cette hypothèse, la délétion du domaine médian (ou la substitution des résidus conservés dans ce domaine) pourra être réalisée et la capacité de la protéine Rgg obtenue à se fixer à l'ADN pourra être déterminée par des expériences de retard de migration sur gel.

Par ailleurs, les peptides Shp (dont les gènes sont localisés à proximité immédiate de gène *rgg*) pourraient interagir avec les protéines Rgg. Des analyses récentes ont montré l'implication de la protéine Rgg₁₃₅₈ de *S. thermophilus* LMD9 dans un mécanisme de QS et ont montré que son activité était dépendante de la présence de la phéromone Shp₁₃₅₈ (Fleuchot *et al.*, 2011). Des résultats similaires ont été obtenus avec les protéines Shp3 et Rgg3 de *S. pyogenes* NZ131 (Chang *et al.*, 2011). Fleuchot *et al.* (2011) ont par ailleurs montré que la protéine Rgg₁₃₅₈ interagit spécifiquement, *in vitro*, avec le peptide Shp₁₃₅₈.

Les protéines impliquées dans des mécanismes de QS, notamment la famille RNPP (les régulateurs Rap, NprR, PlcR et PrgX), se caractérisent par une structure TPR (Tétratricopeptide) en position C-terminale impliquée dans des interactions protéine-protéine ou protéine-peptide (Goebel et Yanagida, 1991). Un motif TPR est constitué de deux hélices α antiparallèles séparées par une

boucle. Dans la plupart des cas, ces motifs sont organisés en tandem de 3 à 16 motifs, générant ainsi une super-hélice au sein de laquelle viennent interagir différentes protéines (Das *et al.*, 1998). Fleuchot *et al.* (2011) ont proposé que la protéine Rgg₁₃₅₈ ait une structure TPR qui permettrait l'interaction avec le peptide Shp₁₃₅₈. Cependant, seule l'analyse par cristallographie permettra de déterminer si la protéine Rgg₁₃₅₈ est capable d'adopter de telle structure. L'hypothèse proposée par Fleuchot *et al.* (2011) est que la phéromone Shp₁₃₅₈ favoriserait l'oligomérisation de la protéine Rgg₁₃₅₈ entraînant son activation, indépendamment de sa capacité de fixation à l'ADN. Les résultats obtenus au cours de ce travail sont en accord avec cette hypothèse puisque la protéine Rgg₀₁₈₂ est capable de se fixer à l'ADN en absence du peptide Shp₀₁₈₂.

1.3. Les protéines Rgg sont-elles toutes activées par un mécanisme de QS ?

Comme cela a été démontré pour le couple Shp₁₃₅₈/Rgg₁₃₅₈ de *S. thermophilus* LMD9 et les couples Shp3/Rgg3 de *S. pyogenes* NZ131 (Chang *et al.*, 2011 ; Fleuchot *et al.*, 2011), les résultats obtenus au cours de ce travail permettent de faire l'hypothèse que la protéine Rgg₀₁₈₂ de *S. thermophilus* LMG18311 pourrait également intervenir dans un mécanisme de QS.

En effet, l'environnement génétique du gène *rgg0182* est similaire à celui de l'ORF *rgg1358* de *S. thermophilus* LMD9, bien que le produit des gènes *shp*, *rgg* et *pep* soit différent (faible pourcentage d'identité entre les protéines) et que les peptides Shp₁₃₅₈ et Shp₀₁₈₂ soient classés dans deux groupes distincts (Fleuchot *et al.*, 2011). Un des arguments en faveur de cette hypothèse est l'influence de la protéine Rgg₀₁₈₂ sur la transcription des gènes *shp0182* et *pep0182*. Par ailleurs, lors d'une croissance réalisée en milieu MCD, à 30°C, la transcription des gènes *rgg0182*, *shp0182* et *pep0182* augmente au cours de la phase de croissance. Pour vérifier cette hypothèse, l'influence de la transcription du gène *pep0182* par la protéine Rgg₀₁₈₂ devra être analysée dans différents contextes génomiques : sauvage, Δ *rgg0182*, Δ *shp0182*, Δ *pep* et Δ *ami*.

Le mécanisme de QS pourrait être généralisé à plusieurs des protéines Rgg. En effet, parmi les 11 copies *rgg* différentes présentes au sein des génomes disponibles de *S. thermophilus*, 8 sont situées à proximité d'ORF codant potentiellement un peptide de type Shp. Fleuchot *et al.* (2011) ont par ailleurs montré 61 associations de gènes *shp/rgg* chez les bactéries, suggérant que ce mécanisme de QS puisse être largement répandu.

Par ailleurs, l'originalité du mécanisme de QS décrit chez *S. pyogenes* NZ131 (Chang *et al.*, 2011) (figure 48) suggère que des protéines Rgg distinctes puissent engendrer des effets croisés.

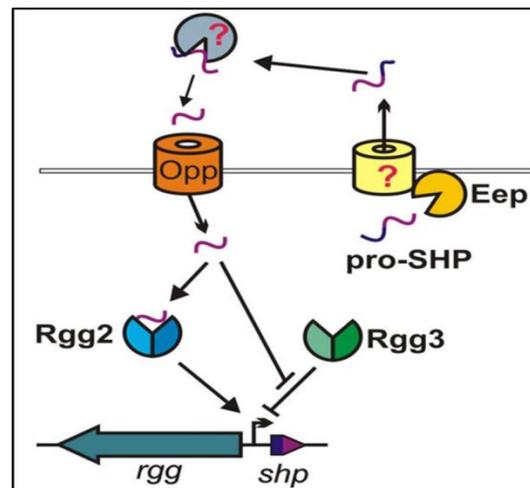


Figure 48 : Modèle du QS faisant intervenir Rgg2/3 et Shp2/3 (d'après Chang 2011). Le peptide pro-Shp est maturé par le système Eep et est sécrété par un système non déterminé. Ensuite, un second traitement conduit à la maturation du peptide Shp sous forme active. La réimportation du peptide Shp mature dans le cytoplasme se fait *via* la perméase Opp. Les peptides Shp interagissent probablement avec la protéine Rgg. En absence du peptide Shp3, la protéine Rgg3 réprime la transcription des gènes *shp2* et *shp3*. Au contraire, en présence du peptide Shp3, la répression par la protéine Rgg3 est inhibée et la transcription des gènes *shp2* et *shp3* par la protéine Rgg2 est induite (Chang 2011).

Chez *S. pyogenes* NZ131, les deux peptides, Shp2 et Shp3, appartiennent tous les deux à la classe I et ont des effets antagonistes (Chang *et al.*, 2011). Ainsi, la présence de plusieurs gènes *rgg* sur les génomes suggère que certains d'entre eux pourraient être impliqués dans des effets croisés. Par exemple, les gènes *rgg* associés aux gènes *shp* codant des peptides Shp de même classe pourraient avoir des effets synergie ou antagonistes. Cependant, des effets croisés entre des protéines Rgg associées à des peptides Shp de classes différentes est tout à fait envisageable. De même que les gènes *rgg* qui ne sont pas associés à des gènes *shp* pourraient être impliqués dans des mécanismes de QS par interaction avec les peptides Shp présents dans le cytoplasme des cellules. Pour vérifier ces hypothèses, l'implication de plusieurs des protéines Rgg dans des mécanismes de QS devra être analysée.

2. Quel est le rôle physiologique des protéines Rgg de *S. thermophilus* ?

2.1. Les protéines Rgg de *S. thermophilus* sont-elles impliquées dans la réponse aux stress ?

2.1.1. Les protéines Rgg ont-elles un effet antagoniste ou compensateur ?

Les génomes de *S. thermophilus* se caractérisent par un grand nombre de gènes *rgg* alors qu'ils évoluent par perte de fonction (Bolotin *et al.*, 2004 ; Siezen *et al.*, 2005 ; Delorme *et al.*, 2009). L'existence d'un grand nombre de gènes *rgg* dans le génome des souches de *S. thermophilus* suggère que leur acquisition et leur maintien ne soient pas contre-sélectionnés et donc que ceux-ci joueraient un rôle important pour la survie de *S. thermophilus*. En effet, plusieurs études ont montré l'implication de protéines Rgg dans la réponse à divers stress (Sanders *et al.*, 1998 ; Chaussee *et al.*, 2004 ; Fernandez *et al.*, 2004a ; Pulliainen *et al.*, 2008 ; Bortoni *et al.*, 2009).

Afin d'étudier l'implication des protéines Rgg de *S. thermophilus* dans la réponse adaptative, les cellules de *S. thermophilus* en phase exponentielle et stationnaire de croissance ont été exposées à divers stress : thermique, oxydant ou acide. Bien que les données obtenues au cours de ce travail aient montré que la protéine Rgg₀₁₈₂ améliore la survie des cellules exposées à 52°C, l'ensemble des expériences réalisées n'a pas permis d'impliquer d'autres gènes *rgg* dans la réponse de *S. thermophilus* aux stress, leur délétion ne conduisant à aucun phénotype particulier. L'hypothèse la plus simple pour expliquer l'absence de phénotype est que ces protéines n'aient pas pour rôle de contribuer à l'adaptation de *S. thermophilus* aux différents stress testés.

Cependant, les conditions de culture pourraient être un paramètre important à prendre en compte lors de l'analyse de la réponse aux stress. Cette dernière hypothèse est à privilégier dans la mesure où le phénotype résultant de la délétion du gène *rgg₀₁₈₂* n'est visible que dans des conditions particulières, c'est-à-dire uniquement lorsque les cellules sont cultivées en milieu MCD et à 30°C. Par ailleurs, la transcription du gène *rgg₀₁₈₂* n'est pas constitutive mais induite en fonction de la phase de croissance et des conditions de croissance. Lors des expériences réalisées, ce dernier paramètre n'a pas été pris en compte pour les autres gènes *rgg* de *S. thermophilus*.

Enfin, la dernière hypothèse est que la présence de plusieurs copies *rgg* au sein du génome de *S. thermophilus* LMG18311 puisse compenser, dans certaines conditions, l'absence d'une des copies. Ainsi, certaines des protéines Rgg pourraient agir en synergie et d'autres auraient des effets antagonistes. Des effets synergiques et antagonistes entre des protéines Rgg a été d'ailleurs récemment décrit chez *S. pyogenes* NZ131 (Chang *et al.*, 2011). En effet, la délétion du gène *ropB* (*rgg1*) ou *rgg3* entraîne une induction (d'un facteur 2 et 3, respectivement) de la production de biofilm par rapport à la souche sauvage. La double délétion de ces deux gènes *rgg* induit, d'un

facteur 5, la production du biofilm, suggérant ainsi un effet synergique de ces deux gènes. Au contraire, la délétion du gène *rgg2* chez un mutant $\Delta rgg1$ ou $\Delta rgg3$ conduit à une production de biofilm similaire à celle de la souche sauvage de *S. pyogenes* NZ131, suggérant ainsi que la protéine Rgg2 ait un effet antagoniste aux protéines Rgg1 et Rgg3.

L'ensemble de ces observations suggère donc que les protéines Rgg soient requises dans des conditions particulières et que certaines d'entre-elles puissent agir en synergie alors que d'autres auraient des effets antagonistes. L'effet croisé, décrit précédemment, entre les protéines Rgg2/3 de *S. pyogenes* NZ131 (Chang *et al.*, 2011) est un argument supplémentaire en faveur de cette hypothèse. Pour le vérifier, l'analyse de la transcription des différents gènes *rgg* de *S. thermophilus* dans différentes conditions de croissance (milieu, température) permettra de déterminer les conditions dans lesquelles ces protéines sont requises. La survie des mutants à différents stress (oxydant, thermique, pH, osmotique,...) sera ensuite étudiée dans les conditions où les gènes *rgg* sont le plus exprimés et dans différents contextes génomiques. En particulier, des mutants présentant la délétion d'un, mais aussi de plusieurs gènes *rgg* (doubles, triples, quadruples mutants,...), permettra de déterminer si oui ou non, les gènes *rgg* de *S. thermophilus* sont, de manière générale, impliqués dans la réponse aux stress et quels sont les gènes *rgg* qui peuvent agir en synergie ou être antagonistes. Enfin, la construction et l'étude d'un mutant présentant la délétion des tous les gènes *rgg* présents sur le génome de la souche LMG18311 permettraient de déterminer si l'ensemble de ces gènes est essentiel à la survie de *S. thermophilus*. A l'heure actuelle, il est cependant difficile de déterminer quels gènes *rgg* pourraient avoir des effets croisés.

2.1.2. La protéine Rgg₀₁₈₂ est-elle impliquée dans les réseaux de régulation de la réponse au stress thermique de *S. thermophilus* ?

Pour répondre à des situations de stress, les organismes induisent un programme complexe d'expression de gènes, impliquant l'augmentation rapide de la concentration d'un jeu de protéines, spécifiques d'un stress donné, permettant à la bactérie d'adapter sa physiologie aux changements environnementaux.

Les protéines Hsp et Csp sont impliquées dans la réponse globale aux stress et, plus particulièrement dans la réponse au stress thermique. Certaines des protéines Hsp (comme les protéines GroEL et DnaK et les protéases Clp ATP-dépendantes) sont des molécules chaperonnes qui facilitent le bon repliement des protéines cellulaires et d'autres (les protéases Clp ATP-indépendantes) dégradent les protéines mal repliées (Gottesman *et al.*, 1997). Les protéines Csp, quant à elles, sont des chaperons de l'ARN facilitant la traduction en condition de stress froid (Jiang

et al., 1996 ; Graumann *et al.*, 1997).

Chez *E. coli*, la synthèse de protéines Hsp est contrôlée au niveau transcriptionnel par deux facteurs σ alternatifs appelés σ^{32} et σ^{24} (Yura *et al.*, 1996). Chez *B. subtilis*, la réponse au stress chaud implique 4 principales classes différentes de gènes, inductibles à la chaleur, qui se distinguent par leur mécanisme de régulation (Hecker *et al.*, 1996) : (i) les gènes de la classe I (*groESL*, *dnaK*) sont régulés par le répresseur HrcA (Schulz et Schumann, 1996), (ii) les gènes de la classe II requièrent le facteur alternatif σ^B (Petersohn *et al.*, 2001 ; Price *et al.*, 2001), (iii) les gènes de la classe III (les gènes *clp*) sont contrôlés par le répresseur CtsR (Derre *et al.*, 1999) et (iv) les gènes de la classe IV qui sont régulés par d'autres régulateurs, connus ou à découvrir. Cependant, *B. subtilis* ne reflète pas tous les mécanismes de la régulation de la réponse aux stress chez les bactéries Gram-positives. Ainsi, bien que l'induction de gènes codant des protéines Hsp soit une réponse universelle, les bactéries disposent de plusieurs stratégies pour ajuster la production de protéines Hsp dans des conditions difficiles. Un exemple est la régulation duelle par les répresseurs HrcA et CtsR qui contrôle l'expression des protéines DnaK, GroES et GroEL de *St. aureus* et la protéine ClpP de *S. salivarius* (Chastanet et Msadek, 2003). Ces mécanismes de régulation duelle pourraient agir en synergie pour assurer le maintien d'un faible niveau d'expression en absence de stress et la synthèse finement coordonnée des différentes protéines Hsp dans des conditions défavorables. La figure 49 présente les mécanismes de régulation des gènes *hsp* par les protéines HrcA et CtsR de différents organismes.

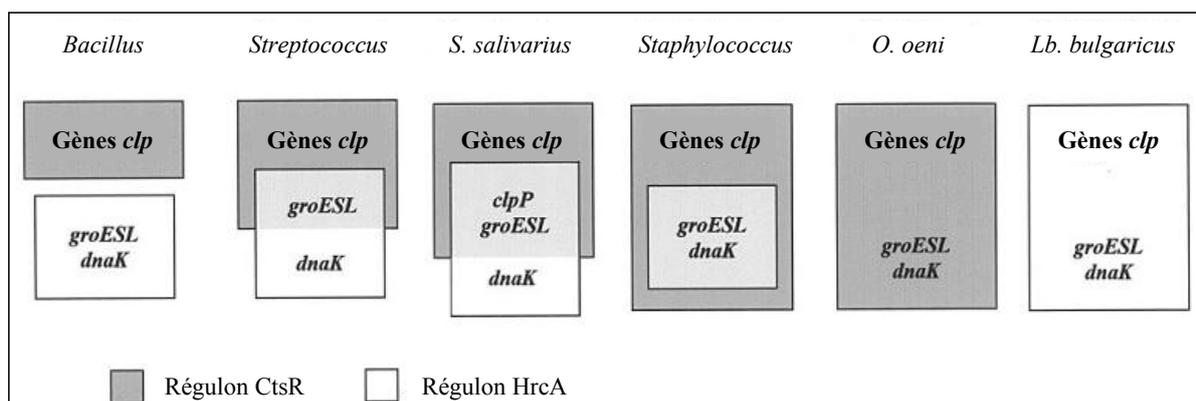


Figure 49 : Régulation duelle par les répresseurs CtsR et HrcA dans différentes bactéries Gram-positives (d'après Chastanet et Msadek, 2003 et Grandvalet *et al.*, 2005). Chez différentes bactéries Gram-positives, des régulons HrcA et CtsR coexistent. Chez *B. subtilis* et les bacilles apparentés (*B. anthracis*, *B. stearothermophilus*, *B. halodurans*, *Clostridium acetobutylicum*, *Clostridium difficile*, *Clostridium perfringens*, *Li monocytogenes* et *Li. innocua*), les deux régulons sont distincts alors que dans le groupe des streptocoques (*S. pneumoniae*, *L. lactis* et *S. salivarius*), ils sont chevauchants. Chez *St. aureus*, le régulon HrcA est complètement inclus dans le régulon CtsR. Les espèces *O. oeni* et *Lb. bulgaricus* ont une régulation des gènes de réponse au stress chaud originale, avec un contrôle prédominant par la protéine CtsR ou la protéine HrcA, respectivement.

DISCUSSION

Les résultats obtenus au cours de ce travail ont permis d'apporter quelques éléments de réponse quant au fonctionnement du réseau de régulation de la réponse au stress thermique de *S. thermophilus* LMG18311. En effet, la régulation duelle de la transcription du gène *groEL* par les répresseurs CtsR et HrcA est envisageable puisqu'une CtsR-box et un élément CIRCE sont retrouvés en amont de ce gène chez *S. thermophilus* LMG18311. De plus, la diminution du taux de transcrite des gènes *ctsR* et *hrcA* s'accompagne d'une augmentation de la transcription du gène *groEL* en condition de stress chaud. Par ailleurs, le gène *dnaJ* pourrait être sous le contrôle de la protéine HrcA puisqu'un élément CIRCE est présent en amont de ce gène et que sa transcription est augmentée en condition de stress. Pour vérifier cette hypothèse, le taux de transcrits des gènes *groEL* et *dnaJ* pourra être déterminé dans des contextes génomique sauvage et mutants $\Delta hrcA$ et $\Delta ctsR$.

Au contraire, la transcription des gènes *clpL* et *clpX* ne semble pas être sous la dépendance des répresseurs HrcA et CtsR. En effet, la diminution de la transcription des gènes *hrcA* et *ctsR* devrait entraîner l'augmentation des transcrits des gènes *clpL* et *clpX*. Or, les résultats obtenus au cours de ce travail montrent l'inverse. L'absence d'une CtsR-box et d'un élément CIRCE en amont du gène *clpX* suggèrent que ce gène n'est pas sous la dépendance de ces répresseurs. Cependant, une CtsR-box est retrouvée en amont du gène *clpL*. L'hypothèse est que la transcription du gène *clpL* ne soit pas sous la dépendance de la protéine CtsR, dans les conditions testées.

Enfin, le stress chaud à 52°C, réalisé sur des cellules en phase stationnaire cultivées en milieu MCD et à 30°C, entraîne la diminution de la transcription du gène *cspB*, suggérant que ce gène ne soit pas requis dans ces conditions et est réprimé par la protéine Rgg₀₁₈₂.

Enfin, dans ces conditions, la protéine Rgg₀₁₈₂ contrôle positivement la transcription du gène *clpE*. Chez *S. pneumoniae*, le gène *clpE* est requis dans la tolérance au stress chaud (Chastanet *et al.*, 2001). Ainsi, la diminution de la survie du mutant Δrgg_{0182} pourrait être due, tout au moins en partie, à la diminution du taux de transcrits du gène *clpE*. Pour vérifier cette hypothèse, un mutant présentant la délétion du gène *clpE* devra être construit et sa survie analysée dans les conditions de stress chaud précédemment décrites. Cependant, la protéine Rgg₀₁₈₂ pourrait également influencer la transcription d'autres cibles, non encore identifiées, dont l'expression pourrait être importante pour la survie des cellules soumises à une augmentation de la température.

Par ailleurs, le contrôle de la transcription des gènes *clpE* et *cspB* par la protéine Rgg₀₁₈₂ serait indirect. Des analyses de transcriptomique dans un contexte sauvage et mutant Δrgg_{0182} permettra de trouver des régulateurs transcriptionnels intermédiaires candidats. Ces données permettront

également de vérifier si la protéine Rgg₀₁₈₂ est ainsi impliquée dans les réseaux de régulation de la réponse au stress thermique.

2.2. La protéine Rgg₀₁₈₂ contrôle-t-elle la synthèse de protéines impliquées dans la division cellulaire ?

La division cellulaire est un processus essentiel consistant à transmettre le patrimoine génétique d'une bactérie à ses descendantes. Ce processus complexe aboutit à la formation de deux cellules filles identiques à partir d'une cellule mère. Quatre grands mécanismes ont lieu : (i) la réplication de l'ADN, (ii) la ségrégation chromosomique, (iii) la scission de la cellule mère en deux cellules filles non disjointes et (iv) la séparation des cellules filles en cellules individuelles (Layec, 2008). Au cours de la division cellulaire, un septum de division est formé au niveau du futur site de séparation. Ce processus nécessite le recrutement de diverses protéines de la division pour promouvoir la formation du septum (Rothfield *et al.*, 1999 ; Buddelmeijer et Beckwith, 2002 ; Errington, 2003). Les protéines de division interviennent, plus précisément, dans l'invagination de la membrane plasmique, dans la synthèse du peptidoglycane latéral permettant l'accroissement de la taille cellulaire et dans la synthèse du peptidoglycane septal permettant la scission de la cellule mère en deux cellules filles. Par ailleurs, une ou plusieurs protéines de division jouerai(en)t un rôle dans la coordination de la formation du septum avec d'autres événements du cycle cellulaire, incluant la réplication chromosomique et la partition des deux chromosomes. Au final, la division cellulaire s'achève par la séparation des deux nouvelles cellules filles en cellules individuelles (Layec, 2008).

Le phénotype « courtes chaînes » du mutant Δ rgg₀₁₈₂ suggère que la protéine Rgg₀₁₈₂ puisse être impliquée dans l'organisation des cellules de *S. thermophilus* en chaînes. Plusieurs hypothèses peuvent être formulées pour expliquer ce résultat.

L'hypothèse la plus simple est que la protéine Rgg₀₁₈₂ active la transcription du gène *cse*. En effet, les enzymes de séparation cellulaire sont assurées par des hydrolases du peptidoglycane et chez *S. thermophilus*, la séparation cellulaire est majoritairement assurée par l'hydrolase Cse, identifiée au laboratoire (Borges *et al.*, 2005 ; Layec *et al.*, 2009). La délétion du gène *cse* entraîne la formation de chaînes extrêmement longues pouvant contenir plus de 3000 cellules (Borges *et al.*, 2005). Cette hypothèse a été vérifiée en comparant le taux de transcrits du gène *cse* dans un contexte génomique sauvage et mutant Δ rgg₀₁₈₂. Cependant, aucune différence du taux de transcrit du gène *cse* n'a été observée entre les deux souches, ce qui invalide cette hypothèse.

Sachant que la longueur des cellules du mutant Δrgg_{0182} est diminuée par rapport à celles de la souche sauvage, l'hypothèse ainsi privilégiée est que la protéine Rgg₀₁₈₂ puisse contrôler la synthèse de protéines de division cellulaire. Cette hypothèse est soutenue par des données de la littérature indiquant que des protéines impliquées dans la formation du peptidoglycane peuvent également influencer la longueur des chaînes de cellules. Par exemple, chez *S. thermophilus*, un phénotype « longues chaînes » a été observé pour des mutants $\Delta rodA$, $\Delta pbp2b$ et $\Delta mreD$ (Thibessard *et al.*, 2002 ; Thibessard *et al.*, 2004). Ces gènes codent des protéines essentielles à la synthèse longitudinal du peptidoglycane permettant l'accroissement de la taille cellulaire au cours de la division bactérienne (Ishino *et al.*, 1986 ; Noirclerc-Savoie *et al.*, 2003). Chez *L. lactis* MG1363, l'inactivation du gène *ponA* conduit à un phénotype « courtes chaînes » (Mercier *et al.*, 2002). Le gène *ponA* code la protéine Pbp1a, une enzyme bifonctionnelle essentielle à la biosynthèse du peptidoglycane lors de la division cellulaire (Goffin et Ghuysen, 1998). Le raccourcissement des chaînes de cellules résulterait de la fragilisation des liaisons peptidiques du peptidoglycane entraînant des cassures (Mercier *et al.*, 2002). Ainsi, la protéine Rgg₀₁₈₂ pourrait inhiber la transcription du gène *ponA* (présent chez *S. thermophilus* LMG18311), par exemple, ce qui entraînerait un raccourcissement de la taille des cellules et des chaînes de cellules courtes.

Un autre argument en faveur de cette hypothèse est l'implication de la protéine Rgg de *S. suis* SS2 dans la morphologie cellulaire. En effet, Zheng *et al.* (2011) ont montré que la délétion du gène *rgg* de *S. suis* SS2 conduit à un phénotype « longues chaînes » et les cellules du mutants sont altérées avec un diamètre plus petit. L'analyse du transcriptome de la souche SS2 dans un contexte génomique sauvage et mutant Δrgg a permis aux auteurs d'émettre l'hypothèse que la protéine Rgg pourrait influencer la transcription des gènes *divIV* (Zheng *et al.*, 2011). La délétion du gène *divIV* chez *S. pneumoniae* conduit à des chaînes de cellules deux à trois fois plus longues que chez la souche sauvage et à des cellules altérées dans leur morphologie (Fadda *et al.*, 2003). Le gène *divIV* est présent sur le génome de *S. thermophilus* LMG18311 et la protéine Rgg₀₁₈₂ pourrait contrôler négativement la transcription de ce gène.

Le taux de transcrits des gènes candidats pourra être analysée par qPCR dans un contexte génomique sauvage et mutant Δrgg_{0182} . Cependant, pour vérifier cette hypothèse et proposer des gènes cibles candidats, une analyse du transcriptome dans un contexte génomique sauvage et mutant Δrgg sera nécessaire au vu du nombre conséquent de protéines impliquées dans la division cellulaire. En effet, chez *S. thermophilus* LMG18311, au moins 81 gènes sont impliqués dans la formation du peptidoglycane et de l'enveloppe cellulaire et au moins 20 gènes dans la division

cellulaire (Bolotin *et al.*, 2004).

Par ailleurs, la délétion du gène *rgg0182* entraîne l'adhésion des cellules sur des surfaces hydrophobes. Cette observation est probablement la conséquence de l'altération de l'allongement des cellules et le raccourcissement des chaînes de cellules. En effet, une corrélation entre la taille des chaînes de cellules et la capacité d'adhésion à une surface a été démontrée pour *S. mutans*, responsable de caries dentaires (Nakano *et al.*, 2005). L'adhésion de *S. mutans* sur les dents est fortement diminuée chez un mutant inactivé du gène *brpA* présentant un phénotype « longues chaînes » (Nakano *et al.*, 2005). Inversement, Nakano *et al.* (2005) ont rapporté que la croissance en longues chaînes de cellules aurait un effet négatif sur l'adhésion de *S. mutans* aux surfaces.

Une hypothèse alternative pour expliquer l'adhésion des cellules du mutant Δ *rgg0182* sur les parois des micro-puits des microplaques du Bioscreen C est la formation de liaisons hydrophobes. Cette hypothèse est supportée par l'analyse des parois des plaques révélant leur hydrophobicité. Cependant, à l'heure actuelle, la stratégie utilisant des pointes hydrophobes, selon la technique force-volume par AFM, n'a pas permis de mesurer l'hydrophobicité de surface des cellules de la souche LMG18311 et du mutant Δ *rgg0182*. Une autre stratégie, basée sur la technique de MATS (Bellon-Fontaine *et al.*, 1996), est actuellement envisagée pour répondre à cette question.

Comme cela a été démontré pour les protéines Rgg1, 2 et 3 de *S. pyogenes* NZ131 (Chang *et al.*, 2011), l'adhésion des cellules à des surfaces hydrophobes chez le mutant Δ *rgg0182* pourrait conduire à la formation d'un biofilm. Pour vérifier cette hypothèse, des expériences de formation de biofilm basée sur la coloration au violet de gentiane (Chang *et al.*, 2011) pourront être réalisées.

2.3. Est-ce que les protéines Rgg sont des régulateurs transcriptionnels globaux ?

Les régulateurs transcriptionnels peuvent réguler un petit nombre de gènes ou au contraire contrôler un large panel de gènes. Ceux qui régulent un petit nombre de gènes sont appelés régulateur locaux ou spécifique et contrôlent directement la transcription de leurs gènes cibles, *via* la fixation du régulateur dans la région promotrice de ceux-ci. Les gènes régulés directement sont appelés gènes spécifiques.

Au contraire, les régulateurs transcriptionnels agissant sur un large panel de gènes sont appelés régulateurs globaux. Selon la définition donnée par Martinez-Antonio et Collado-Vides (2003), un régulateur global est défini par (i) sa capacité à réguler un large nombre de gènes appartenant à des

classes différentes, c'est-à-dire intervenant dans des fonctions bien distinctes ; (ii) le fait qu'il contrôle une cascade de régulation complexe en intervenant directement ou indirectement (par l'intermédiaire de régulateurs variés qui agissent, eux même, sur quelques gènes spécifiques) et (iii) le fait qu'il puisse agir sur des promoteurs de gènes cibles d'autres régulateurs transcriptionnels.

Ainsi, les régulons qui sont sous le contrôle de régulateurs globaux incluent un large nombre de gènes qui sont souvent inclus dans le régulon d'autres régulateurs transcriptionnels. L'ensemble de ces deux types de régulateurs transcriptionnels, spécifiques et globaux, forment, ainsi, un réseau hiérarchisé. Pour exemple, le réseau de régulation d'*E. coli* qui a été bien analysé est présenté dans la figure 50.

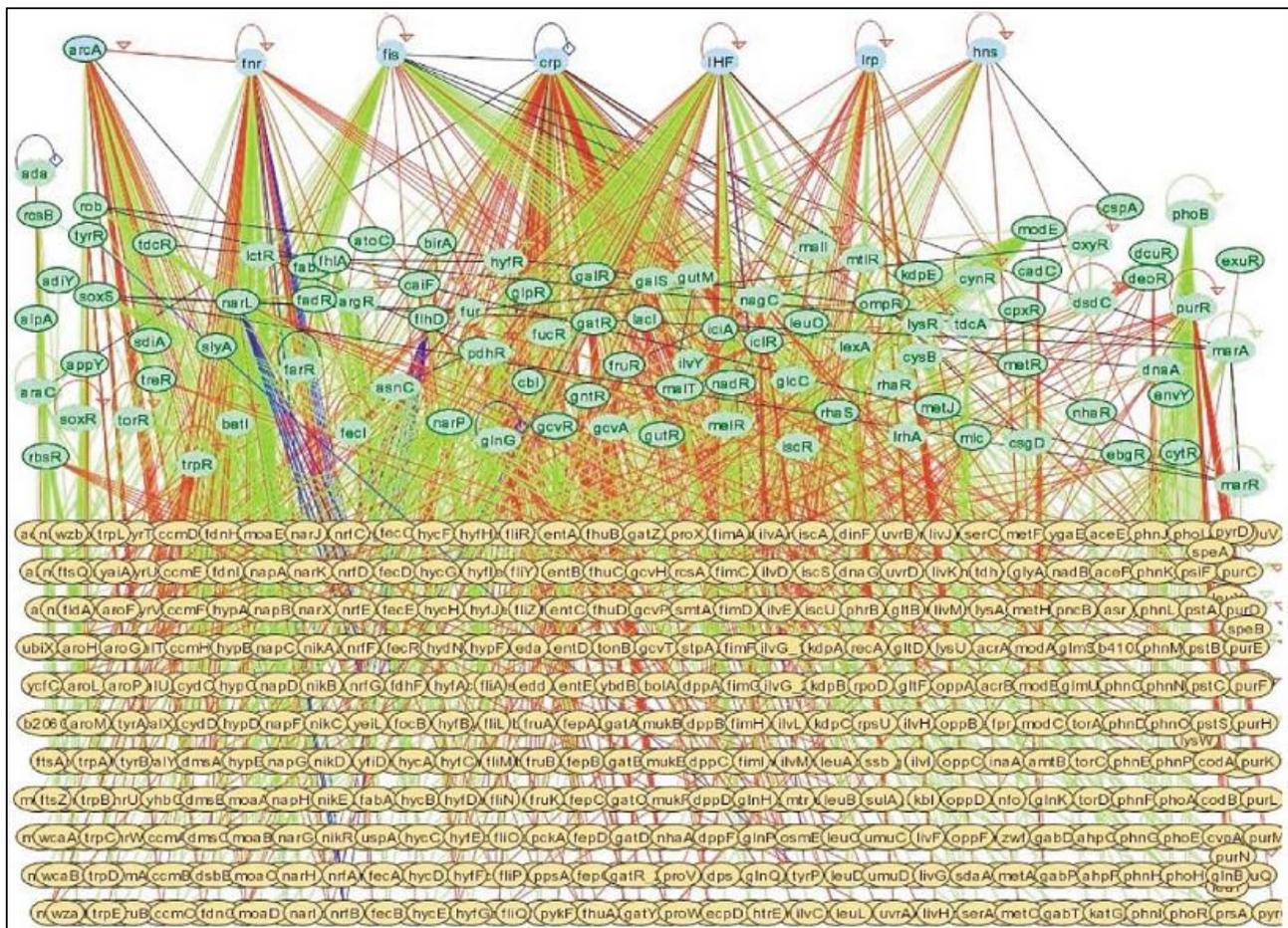


Figure 50 : Vue d'ensemble du réseau de régulation d'*E. coli* (Martinez-Antonio et Collado-Vides, 2003). Les gènes régulés sont indiqués par des ovales jaunes. Les régulateurs transcriptionnels spécifiques sont indiqués par des ovales verts et les régulateurs transcriptionnels considérés comme globaux sont indiqués par des ovales bleus. Les lignes vertes indiquent l'induction, les lignes rouges la répression et les lignes bleues une régulation duelle (activation et répression).

La figure 50 montre que, d'une part, un petit nombre de régulateurs transcriptionnels globaux influence la transcription d'un large nombre de gènes impliqués dans différentes voies métaboliques (Gottesman, 1984 ; Perez-Rueda et Collado-Vides, 2000 ; Gutierrez-Rois *et al.*, 2003) et d'autre

part, un large nombre de régulateurs transcriptionnels locaux ou spécifiques affecte la transcription d'un petit nombre de gènes spécifiques (Martinez-Antonio et Collado-Vides, 2003 ; Wei *et al.*, 2004). Le rôle des régulateurs globaux et spécifiques permet, ainsi, d'adapter la physiologie de l'organisme par activation ou répression des gènes appropriés et prenant en compte, par exemple, des changements environnementaux ou des états métaboliques particuliers (Martinez-Antonio et Collado-Vides, 2003).

Certaines protéines Rgg ont été décrites comme étant des régulateurs transcriptionnels globaux de l'expression génique. En effet, les protéines RopB et Rgg régulent environ 15% des ORF présentes sur les génomes de *S. pyogenes* NZ131 et *S. suis* SS2, respectivement (Chaussee *et al.*, 2003 ; Zheng *et al.*, 2011). Par ailleurs, Samen *et al.* (2006) ont montré que la protéine RovS était capable d'influencer la transcription de gènes de façon indirecte, c'est-à-dire via des régulateurs intermédiaires et font l'hypothèse que cette protéine soient un régulateur transcriptionnel global.

Les résultats obtenus au cours de ce travail permettent de faire l'hypothèse que le gène *rgg₀₁₈₂* de *S. thermophilus* LMG18311 pourrait coder un régulateur transcriptionnel global de l'expression génique. Cette hypothèse se justifie par le fait que : (i) le gène *rgg₀₁₈₂* est impliqué dans des fonctions bien distinctes comme la réponse au stress thermique, la division et l'adhésion cellulaire ; (ii) la protéine Rgg₀₁₈₂ peut agir directement (fixation de la protéine dans la région promotrice des gènes *pep₀₁₈₂* et *shp₀₁₈₂*) et indirectement (absence de fixation sur les promoteurs des gènes *clpE* et *cspB*, Emmanuelle Bruneau, communication personnelle), et (iii) la protéine Rgg₀₁₈₂ agit sur le contrôle de la transcription de gènes inclus dans les régulons HrcA et CtsR. Ainsi, la protéine Rgg₀₁₈₂ serait capable d'influencer la transcription de certains gènes, très probablement, par l'intermédiaire d'autres régulateurs transcriptionnels. Une analyse de transcriptomique et/ou protéomique permettra de vérifier cette hypothèse et d'identifier le régulon de la protéine Rgg₀₁₈₂ de *S. thermophilus* LMG18311.

3. Comment expliquer la répartition des protéines Rgg chez les bactéries ?

Les données de la littérature et les résultats obtenus au cours de ce travail ont permis de montrer que les protéines Rgg sont réparties au sein des génomes appartenant à la famille des *Streptococcaceae* et des genre *Lactobacillus* et *Listeria*. Néanmoins, le faible nombre de génomes disponibles concernant les genres *Aerococcus*, *Carnobacterium*, *Granulicatella*, *Enterococcus*, *Pediococcus*, *Leuconostoc*, *Oenococcus* et *Weissella* n'a pas permis de conclure quant à la répartition des copies *rgg* au sein des génomes appartenant à ces genres bactériens. Ainsi, une hypothèse est que les copies Rgg pourraient être spécifiques des bactéries lactiques. Cependant, pour savoir si les protéines Rgg sont spécifiques des bactéries lactiques ou si elles sont seulement spécifiques de la famille des *Streptococcaceae* et des genres *Lactobacillus* et *Listeria*, seule l'augmentation du nombre de génomes disponibles pourra y répondre efficacement.

La figure 51 présente la parenté phylogénétique existant entre les bactéries lactiques et quelques genres bactériens apparentés.

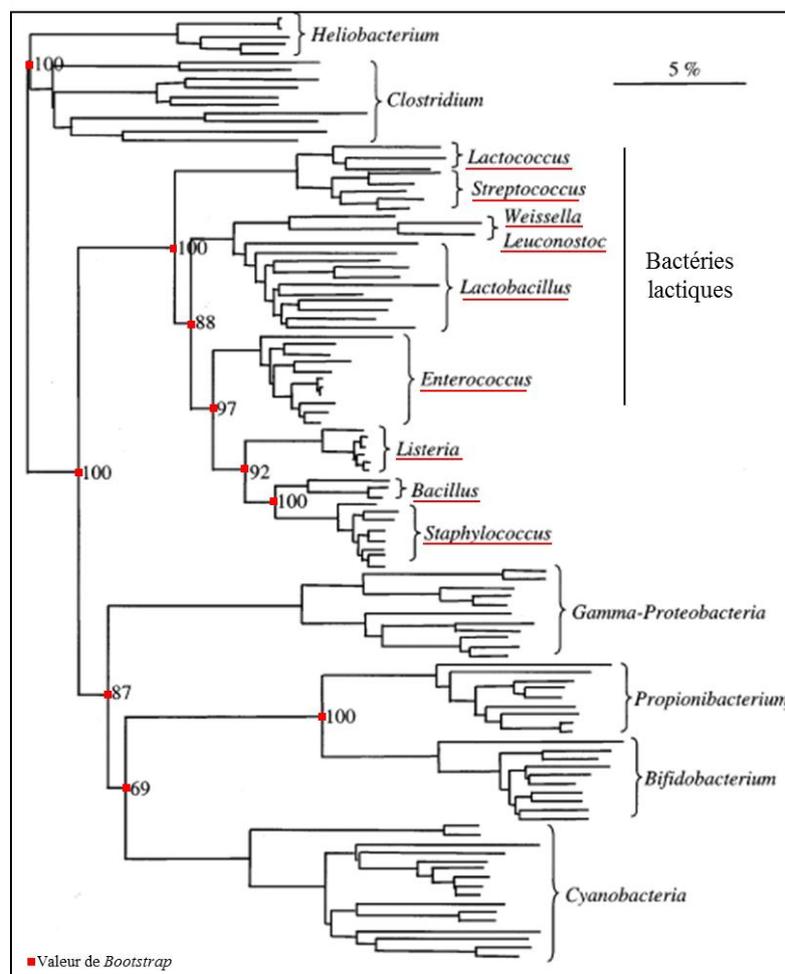


Figure 51 : Arbre phylogénétique montrant la position des bactéries lactiques par rapport aux genres apparentés (*Clostridium*, *Listeria*, *Bacillus*, *Staphylococcus*), aux bactéries Gram positif à haut GC % (*Propionibacterium*, *Bifidobacterium*), et à

certaines bactéries Gram négatif (γ -protéobactéries, cyanobactéries) (d'après Tailleux, 2001). Le dendrogramme est construit par l'alignement des séquences d'ARNr 16S en utilisant la méthode du *Neighbour-Joining* (NJ). Les séquences utilisées sont celles présentes dans les banques de données. Les valeurs de *Bootstrap* indiquées aux différents nœuds importants de l'arbre sont exprimées en pourcentage des résultats obtenus à partir de 1000 répliques. Les genres soulignés en rouge correspondent aux genres où des copies *rgg* sont retrouvées.

La répartition des copies *rgg* au sein des genres *Streptococcus*, *Lactococcus* et *Lactobacillus* pourrait ainsi s'expliquer par la proximité phylogénétique de ces genres. Cependant, la phylogénie ne permet pas, à elle seule, d'expliquer cette répartition qui tient compte également de la proximité écologique des souches. En effet, les souches qui possèdent le plus de copies *rgg* sont celles qui peuplent des niches écologiques communes. Les genres *Lactococcus*, *Lactobacillus* et l'espèce *S. thermophilus* regroupent des bactéries utilisées en industrie alimentaire. Ces bactéries participent, notamment, à la fermentation du lait mais également, pour le genre *Lactobacillus*, à la fermentation de la viande, de certains végétaux, de céréales, etc. Les listérias sont des bactéries ubiquitaires, contaminants des productions alimentaires. Elles peuvent donc coloniser des niches écologiques communes aux genres *Lactococcus*, *Lactobacillus* et de l'espèce *S. thermophilus*. Ainsi, bien que ces critères ne soient exhaustifs, la parenté phylogénétique et la proximité écologique des souches seraient des critères permettant d'expliquer la présence de copies *Rgg* au sein des genres *Lactococcus*, *Lactobacillus*, *Streptococcus* et *Listeria*.

Par ailleurs, ces copies *rgg* auraient pu se mobiliser d'une espèce à une autre par des mécanismes de transferts horizontaux. Eng *et al.* (2011) ont mis au point une méthode permettant de prédire *in silico* quels sont les gènes probablement acquis par transfert horizontal au sein d'un génome bactérien. Cette méthode est basée sur une technique appelé HMM (Hidden Markov Model) qui permet de détecter des régions d'ADN avec une composition en bases atypique par rapport au reste du génome. Cette technique a notamment permis de mettre en évidence que 18% des gènes de la souche CNR1066 de *S. thermophilus* auraient été acquis par transfert horizontal, incluant les 7 gènes *rgg* (Eng *et al.*, 2011).

Plusieurs caractéristiques des gènes *rgg* de *S. thermophilus* sont en faveur de leur acquisition par transfert horizontal : (i) leur pourcentage en bases G et C (allant de 25 à 31 %) plus faible que le pourcentage global du génome de *S. thermophilus* (compris entre 37,2 et 39,8 %), (ii) le polymorphisme de la séquence primaire des protéines *Rgg* ainsi que leur présence/absence d'une souche de *S. thermophilus* à une autre et (iii) la présence d'IS dans l'environnement génétique proche de plusieurs des ORF *rgg* (4 copies sur les 11 de *S. thermophilus*).

DISCUSSION

Ainsi, ces données suggèrent que des gènes *rgg* aient été acquis par transfert horizontal, mais n'exclut pas que certaines des copies puissent être issues de duplication au sein des génomes. Par exemple, les copies RggSt2 et RggSt5 présentent un fort pourcentage d'identité entre-elles et sont présentes dans les 5 génomes disponibles de *S. thermophilus*, suggérant que ces deux copies aient divergées à partir d'une copie ancestrale commune.

Par ailleurs, le motif HTH des protéines Rgg appartient à la famille des motifs HTH-XRE (pour Xenobiotic Response Element) et les membres de la famille XRE incluent des protéines régulatrices des phages Cro et C1. Des études ont montré que la protéine RopB de *S. pyogenes* NZ131 contrôle la transcription de gènes codant des excisionase et des intégrases spécifiques des phages NZ131.1 et NZ131.2 (Dmitriev *et al.*, 2006 ; Anbalagan *et al.*, 2011). La protéine RopB se fixe dans la région promotrice de ces gènes et contrôle ainsi la fréquence d'excisions et d'insertions dans l'ADN des deux phages (Anbalagan *et al.*, 2011). Dmitriev *et al.* (2006) ont rapporté que le pourcentage en bases G et C du gène *ropB* est similaire à celui d'un gène du phage NZ131.1 (32%) et est inférieur à la moyenne global du génome de la souche NZ131 de *S. pyogenes* (39%). Cette caractéristique est partagée par de nombreux gènes *rgg* décrits tels les gènes *lasX*, *gadR*, et *mutR* (Dmitriev *et al.*, 2006) ou les gènes *rgg* de *S. thermophilus* (ce travail). Enfin, la protéine RopB possède des acides aminés similaires à ceux de répresseurs des bactériophages (Dmitriev *et al.*, 2006). Anbalagan *et al.* (2011) font ainsi l'hypothèse que les protéines Rgg puissent être issues de bactériophages. En effet, les gènes codant des systèmes de régulation du bactériophage pourraient conférer des avantages à la bactérie, par exemple, en s'adaptant à des niches écologiques diverses. Ainsi, la participation de ces gènes dans le réseau de régulation globale de la bactérie, comme les gènes *rgg*, lui permettrait d'augmenter sa probabilité d'être maintenu dans le génome par sélection naturelle.

Ces hypothèses sont cohérentes avec la répartition des protéines Rgg dans les bactéries. En effet, les protéines Rgg sont présentes dans des bactéries Gram-positives à bas GC et cela pourrait refléter le spectre d'infection d'un bactériophage ancestral. Par ailleurs, les gènes *rgg* contrôlant l'excision et l'insertion à l'ADN du bactériophage pourraient ainsi se transférer facilement à une bactérie phylogénétiquement proche et présente dans des niches écologiques communes, par des mécanismes de transferts horizontaux.

L'analyse des génomes disponibles, et disposants de copies *rgg*, a permis de montrer que : (i) plusieurs copies *rgg* (1 à 7) peuvent être présentes au sein d'un seul et même génome, et que (ii) le nombre de copies *rgg* par génome est variable selon les souches bactériennes considérées.

Les bactéries lactiques évoluent par perte de gènes et donc par réduction de la taille de leur

DISCUSSION

génomique (Pfeiler et Klaenhammer, 2007). En effet, les génomes des Lactobacillales et des Bacilli divergent d'environ 600 à 1200 gènes, présents chez les Bacilli et absents chez les Lactobacillales (Makarova *et al.*, 2006). Les gènes perdus chez les Lactobacillales concernent des gènes qui ne sont pas nécessaires au développement et à la survie de cellules en milieux riches, milieux au sein desquels les bactéries lactiques sont présentes. De plus, en fonction de leur condition de vie et de leur niche écologique, les génomes des bactéries lactiques se sont spécialisés d'un point de vue métabolique et physiologique, non seulement par la perte mais également par l'acquisition de fonction. Enfin, un génome de petite taille présente moins de gènes redondants au niveau de leur fonction qu'un génome de plus grande taille. L'exemple le plus marquant est celui de *S. thermophilus* dont l'évolution par perte de fonction a été bien étudiée (Bolotin *et al.*, 2004 ; Siezen *et al.*, 2005 ; Delorme *et al.*, 2009). Ainsi, le maintien de ces gènes dans les génomes suggère qu'ils confèrent un avantage adaptatif aux bactéries qui les possèdent.

Les données de la littérature et celles obtenues au cours de ce travail indiquent que certaines des protéines Rgg sont des régulateurs globaux intervenant dans la transcription de nombreux gènes. En effet, les protéines RopB et Rgg régulent environ 15% des gènes du génome de *S. pyogenes* NZ131 et de celui de *S. suis* SS2, respectivement (Chaussee *et al.*, 2002 ; Zheng *et al.*, 2011). Les gènes régulés codent des protéines aux fonctions très variées, comme la réponse aux stress, le métabolisme, les systèmes de transport, la division cellulaire, la réparation de l'ADN,... Ainsi, les protéines Rgg pourraient être des régulateurs transcriptionnels globaux de l'expression génique. Ces protéines pourraient avoir des régulons plus ou moins larges. De plus certains des gènes de ces régulons pourraient être communs à plusieurs protéines Rgg, comme les protéines Rgg1, 2 et 3 de *S. pyogenes* NZ131 qui sont impliquées dans la formation de biofilm (Chang *et al.*, 2011). De plus, les effets croisés des protéines Rgg2 et 3 de *S. pyogenes* dans l'induction d'un mécanisme de QS (Chang *et al.*, 2011) suggèrent que la présence de plusieurs gènes *rgg* dans un génome pourrait assurer une régulation fine de l'expression génique. Ainsi, le maintien d'un nombre plus ou moins important de copies *rgg* dans de tels génomes suggère que ces copies assureraient un rôle essentiel à la bactérie.

4. Perspectives de travail

4.1. Perspectives à court et moyen terme

Dans un premier temps, le taux de transcrite des différents gènes *rgg* de *S. thermophilus* LMG18311 sera étudié dans différentes conditions de croissance (milieux de culture, température de croissance) afin de déterminer les conditions dans lesquelles ces gènes sont requis. Ensuite, la réponse aux stress (oxydant, acide, thermique,...) des mutants Δrgg sera comparée à la souche sauvage dans les conditions dans lesquelles les gènes *rgg* sont le plus fortement transcrits. Cette étude permettra de déterminer, d'une part, quels sont les gènes *rgg* qui sont impliqués dans la réponse aux stress et, d'autre part, quels sont les gènes *rgg* qui peuvent avoir des effets antagonistes ou synergiques. Par ailleurs, cette étude permettra de déterminer si la majorité des gènes *rgg* sont impliqués dans la réponse à différents stress ou si, au contraire, ils sont spécialisés dans la réponse à un stress particulier.

L'ensemble des résultats obtenus au cours de cette thèse ont permis de mieux caractériser l'implication de la protéine Rgg₀₁₈₂ de *S. thermophilus* LMG18311 et suggère que celle-ci soit un régulateur global de l'expression génique. Une analyse par transcriptomique et/ou par protéomique devra être réalisée et permettra de déterminer les réseaux de régulation dans laquelle la protéine Rgg₀₁₈₂ est impliquée. Les résultats permettront de proposer des gènes cibles candidats impliqués dans les phénotypes observés chez le mutant Δrgg_{0182} et ceux-ci pourront être vérifiés par qPCR dans les contextes génétiques sauvage et mutant.

Par ailleurs, le couple Shp₀₁₈₂/Rgg₀₁₈₂ pourrait agir de concert dans un mécanisme de QS. La délétion des principaux intervenants dans de tels mécanismes (gènes *ami*, *eep*, *shp*) sera réalisée. L'activité du promoteur P_{pep0182} sera ensuite analysée dans ces différents contextes génétiques. De plus, l'interaction de la protéine Shp₀₁₈₂ avec la protéine Rgg₀₁₈₂ pourra être analysée par la technique de Biacore ce qui permettra de mieux comprendre le mode d'activation de la protéine Rgg₀₁₈₂. Enfin, la délétion du gène *pep0182* devra être réalisée afin de déterminer le(s) rôle(s) de celui-ci.

4.2. Perspectives à long terme

A plus long terme, la capacité des protéines Rgg à contrôler la transcription génique sera explorée. Pour cela, le transcriptome de chacun des mutants Δrgg de *S. thermophilus* sera déterminé et comparé à celui de la souche sauvage. Les résultats obtenus permettront, d'une part, de

DISCUSSION

déterminer le régulon de chacune des protéines Rgg et, d'autres part, de montrer un possible chevauchement de ces régulons entre les protéines Rgg. Ce dernier point permettra également de proposer les protéines Rgg qui peuvent avoir des effets antagonistes ou synergiques.

Le mécanisme de QS pourra être étudié pour tous les gènes *rgg* de *S. thermophilus* LMG18311. Cela permettra de savoir si seuls les gènes *rgg* associés à un gène *shp* sont impliqués dans des mécanismes de QS ou si des effets croisés sont possibles. Les effets croisés pourront par ailleurs être analysés en fonction des couples Shp/Rgg pouvant agir en synergie et/ou en antagonisme.

Par ailleurs, les sites de fixation des différentes protéines Rgg pourront être analysés. Pour cela, des expériences de retard de migration sur gels seront réalisées dans la région promotrice de leurs gènes cibles afin de déterminer les séquences d'ADN où ces protéines peuvent se fixer. Des expériences d'empreinte à la DNase I seront ensuite réalisées afin de déterminer avec précision le site de fixation des protéines Rgg. La comparaison des séquences nucléotidiques où se fixent chacune des protéines Rgg permettra de déterminer des sites consensus de fixation.

Enfin, la substitution des acides aminés conservés dans les protéines Rgg pourront être réalisées. Les protéines obtenues, avec des substitutions dans le domaine HTH, seront analysées dans leur capacité à se fixer à l'ADN et celles avec des substitutions dans le domaine médian seront analysées dans leur capacité à se dimériser.

BIBLIOGRAPHIE

BIBLIOGRAPHIE

- Anbalagan, S., W. M. McShan, P. M. Dunman et M. S. Chaussee (2011). "Identification of Rgg Binding Sites in the *Streptococcus pyogenes* Chromosome." J Bacteriol **193**(18): 4933-42.
- Antoine, J. M., F. Adam, A. Fazel et D. Hartley (1994). "Bactéries lactiques en alimentation humaine." Bactéries lactiques : Aspects fondamentaux et technologiques. Loriga. **2**: 419-438.
- Antunes, L. C., R. B. Ferreira, M. M. Buckner et B. B. Finlay (2010). "Quorum sensing in bacterial virulence." Microbiology **156**: 2271–2282.
- Aravind, L., V. Anantharaman, S. Balaji, M. M. Babu et L. M. Iyer (2005). "The many faces of the helix-turn-helix domain: transcription regulation and beyond." FEMS Microbiol Rev **29**(2): 231-62.
- Bassler, B. L. (2002). "Small talk: Cell-to-cell communication in bacteria." Cell **109**: 421–424.
- Béal, C., N. Deschamps, V. Juillard, H. de Roissard, J. Richard et B. Saraux (1994). "Cinétiques de croissance et d'acidification des bactéries lactiques." Bactéries lactiques : Aspects fondamentaux et technologiques. Loriga. **1**: 367- 402.
- Beal, C., P. Louvet et G. Corrieu (1989). "Influence of controlled pH and temperature on the growth and acidification of pure cultures of *Streptococcus thermophilus* 404 and *Lactobacillus bulgaricus* 398." Applied Microbiology and Biotechnology **32**: 148–154.
- Bellon-Fontaine, M. N., J. Rault et C. J. van Oss (1996). "Microbial adhesion to solvents: a novel method to determine the electron-donor/electron-acceptor or Lewis acid-base properties of microbial cells." Colloids and Surfaces B: Biointerfaces **7**(1).
- Benachour, A., N. Sauvageot, V. Pichereau et J. C. Giard (2009). "Les réponses aux stress technologiques et environnementaux." Bactéries lactiques : physiologie, métabolisme, génomique et applications industrielles. Economica. 127- 160.
- Bentley, R. W., J. A. Leigh et M. D. Collins (1991). "Intragenomic structure of *Streptococcus* based on comparative analysis of small-subunit rRNA sequences." Int J Syst Bacteriol **41**(4): 487-94.
- Bentley, S. D., K. F. Chater, A. M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill et D. A. Hopwood (2002). "Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2)." Nature **417**(6885): 141-7.
- Bibal, B., G. Goma, Y. Vayssier et A. Pareilleux (1988). "Influence of pH, lactose and lactic acid on growth of *Streptococcus cremoris*: a kinetic study." Appl. Microbiol. Biotechnol. **28**: 340.

BIBLIOGRAPHIE

- Bolotin, A., B. Quinquis, P. Renault, A. Sorokin, S. D. Ehrlich, S. Kulakauskas, A. Lapidus, E. Goltsman, M. Mazur, G. D. Pusch, M. Fonstein, R. Overbeek, N. Kyprides, B. Purnelle, D. Prozzi, K. Ngui, D. Masuy, F. Hancy, S. Burtreau, M. Boutry, J. Delcour, A. Goffeau et P. Hols (2004). "Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*." Nat Biotechnol **22**(12): 1554-8.
- Bolotin, A., P. Wincker, S. Mauger, O. Jaillon, K. Malarme, J. Weissenbach, S. D. Ehrlich et A. Sorokin (2001). "The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403." Genome Res **11**(5): 731-53.
- Borges, F., S. Layec, A. Thibessard, A. Fernandez, B. Gintz, P. Hols, B. Decaris et N. Leblond-Bourget (2005). "*cse*, a Chimeric and variable gene, encodes an extracellular protein involved in cellular segregation in *Streptococcus thermophilus*." J Bacteriol **187**(8): 2737-46.
- Bortoni, M. E., V. S. Terra, J. Hinds, P. W. Andrew et H. Yesilkaya (2009). "The pneumococcal response to oxidative stress includes a role for Rgg." Microbiology **155**(12): 4123-34.
- Boucher, P. E., F. D. Menozzi et C. Locht. (1994). "The modular architecture of bacterial response regulators. Insights into the activation mechanism of the BvgA transactivator of *Bordetella pertussis*." J Mol Biol. **241**(3): 366-77.
- Bouillanne, C. et M. J. Desmazeaud (1981). "Classement de souches de *Lactobacillus bulgaricus* selon quelques caractères utilisés en fabrication du yoghurt. Association avec *Streptococcus thermophilus*." Sciences Aliments **1**: 7-17.
- Brennan, R. G. (1993). "The winged-helix DNA-binding motif: another helix-turn-helix takeoff." Cell **74**(5): 773-6.
- Buddelmeijer, N. et J. Beckwith (2002). "Assembly of cell division proteins at the *E. coli* cell center." Curr Opin Microbiol. **5**(6): 553-7.
- Chang, J. C., B. Lasarre, J. C. Jimenez, C. Aggarwal et M. J. Federle (2011). "Two Group A Streptococcal Peptide Pheromones Act through Opposing Rgg Regulators to Control Biofilm Development." PLoS Pathog **7**(8): e1002190.
- Chastanet, A. et T. Msadek (2003). "ClpP of *Streptococcus salivarius* is a novel member of the dually regulated class of stress response genes in gram-positive bacteria." J Bacteriol **185**(2): 683-7.
- Chastanet, A., M. Prudhomme, J. P. Claverys et T. Msadek (2001). "Regulation of *Streptococcus pneumoniae* *clp* genes and their role in competence development and stress survival." J Bacteriol **183**(24): 7295-307.
- Chaussee, M. A., E. A. Callegari et M. S. Chaussee (2004). "Rgg regulates growth phase-dependent expression of proteins associated with secondary metabolism and stress in *Streptococcus pyogenes*." J Bacteriol **186**(21): 7091-9.
- Chaussee, M. S., D. Ajdic et J. J. Ferretti (1999). "The *rgg* gene of *Streptococcus pyogenes* NZ131 positively influences extracellular SPE B production." Infect Immun **67**(4): 1715-22.

BIBLIOGRAPHIE

- Chaussee, M. S., G. A. Somerville, L. Reitzer et J. M. Musser (2003). "Rgg coordinates virulence factor synthesis and metabolism in *Streptococcus pyogenes*." J Bacteriol **185**(20): 6016-24.
- Chaussee, M. S., G. L. Sylva, D. E. Sturdevant, L. M. Smoot, M. R. Graham, R. O. Watson et J. M. Musser (2002). "Rgg influences the expression of multiple regulatory loci to coregulate virulence factor expression in *Streptococcus pyogenes*." Infect Immun **70**(2): 762-70.
- Chaussee, M. S., R. O. Watson, J. C. Smoot et J. M. Musser (2001). "Identification of Rgg-regulated exoproteins of *Streptococcus pyogenes*." Infect Immun **69**(2): 822-31.
- Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaiia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M. A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead et B. G. Barrell (1998). "Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence." Nature **393**(6685): 537-44.
- Cunliffe, N. A. et A. J. Jacob (1997). "*Streptococcus vestibularis* bacteraemia." J Infect **34**(1): 85.
- Curk, M. C., J. M. Boeufgras, B. Decaris, F. Gavini, K. Kersters, J. P. Larpent, P. Le Bourgeois, H. de Roissard et C. Rouvier (1994). "Méthodes d'identification des bactéries lactiques." Bactéries lactiques : Aspects fondamentaux et technologiques. Lorica. **1**: 141-168.
- Dalton, T. L. et J. R. Scott (2004). "CovS inactivates CovR and is required for growth under conditions of general stress in *Streptococcus pyogenes*." J Bacteriol **186**(12): 3928-37.
- Das, A. K., P. W. Cohen et D. Barford (1998). "The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions." Embo J **2**(17) : 1192-9.
- de Boer, J. G. et B. W. Glickman (1998). "The *lacI* gene as a target for mutation in transgenic rodents and *Escherichia coli*." Genetics **148**(4): 1441-51.
- de Roissard, H. (1986). "Laits et produits laitiers." Lavoisier: 343.
- Dellaglio, F., H. de Roissard, S. Torriani, M. C. Curk et D. Janssens (1994). "Caractéristiques générales des bactéries lactiques." Bactéries lactiques : Aspects fondamentaux et technologiques. Lorica. **1**: 25-116.
- Delorme, C., C. Bartholini, A. Bolotine, S. D. Ehrlich et P. Renault (2009). "Emergence of a cell wall protease in the *Streptococcus thermophilus* population." Appl Environ Microbiol **76**(2): 451-60.
- Delorme, C., C. Bartholini, M. Luraschi, M. Pons, V. Loux, M. Almeida, E. Guédon, J. F. Gibrat et P. Renault (2011). "Complete genome sequence of pigmented *Streptococcus thermophilus* strain JIM8232." J Bacteriol **193**(19): 5581-2.
- Derre, I., G. Rapoport et T. Msadek (1999). "CtsR, a novel regulator of stress and heat shock

BIBLIOGRAPHIE

- response, controls *clp* and molecular chaperone gene expression in gram-positive bacteria." Mol Microbiol **31**(1): 117-31.
- Derre, I., G. Rapoport et T. Msadek (2000). "The CtsR regulator of stress response is active as a dimer and specifically degraded in vivo at 37 degrees C." Mol Microbiol **38**(2): 335-47.
- Dmitriev, A. V., E. J. McDowell, K. V. Kappeler, M. A. Chaussee, L. D. Rieck et M. S. Chaussee (2006). "The Rgg regulator of *Streptococcus pyogenes* influences utilization of nonglucose carbohydrates, prophage induction, and expression of the NAD-glycohydrolase virulence operon." J Bacteriol **188**(20): 7230-41.
- Dziejman, M. and J. J. Mekalanos (1995). "Two-component signal transduction and its role in the expression of bacterial virulence factors." Two-Component Signal Transduction. A. S. F. Microbiology. 305–317.
- Elsholz, A. K., U. Gerth et M. Hecker (2010). "Regulation of CtsR activity in low GC, Gram+ bacteria." Adv Microb Physiol. **57**: 119-44.
- Eng, C., A. Thibessard, M. Danielsen, T. B. Rasmussen, J. F. Mari et P. Leblond (2011). "In silico prediction of horizontal gene transfer in *Streptococcus thermophilus*." Arch Microbiol **193**(4): 287-97.
- Errington, J. (2003). "Dynamic proteins and a cytoskeleton in bacteria." Nat Cell Biol **5**(3): 175-8.
- Fadda, D., C. Pischedda, F. Caldara, M. B. Whalen, D. Anderluzzi, E. Domenici et O. Massidda (2003). "Characterization of *divIVA* and other genes located in the chromosomal region downstream of the *dcw* cluster in *Streptococcus pneumoniae*." J Bacteriol **185**(20): 6209-14.
- Farrow, J. A. et M. D. Collins (1984). "DNA base composition, DNA-DNA homology and long-chain fatty acid studies on *Streptococcus thermophilus* and *Streptococcus salivarius*." J Gen Microbiol **130**(2): 357-62.
- Federle, M. J., K. S. McIver et J. R. Scott (1999). "A response regulator that represses transcription of several virulence operons in the group A *streptococcus*." J Bacteriol **181**(12): 3649-57.
- Federle, M. J. et J. R. Scott (2002). "Identification of binding sites for the group A streptococcal global regulator CovR." Mol Microbiol **43**(5): 1161-72.
- Fernandez, A. (2004b). "Gènes de la réponse au stress oxydant de *Streptococcus thermophilus* ; implication de la famille multigénique polymorphe codant des régulateurs transcriptionnels Rgg-like." Thèse Université Henri Poincaré, Nancy 1.
- Fernandez, A., A. Thibessard, F. Borges, B. Gintz, B. Decaris et N. Leblond-Bourget (2004a). "Characterization of oxidative stress-resistant mutants of *Streptococcus thermophilus* CNRZ368." Arch Microbiol **182**(5): 364-72.
- Fleuchot, B., C. Gitton, A. Guillot, J. Vidic, P. Nicolas, C. Besset, L. Fontaine, P. Hols, N. Leblond-Bourget, V. Monnet et R. Gardan (2011). "Rgg proteins associated with internalized small hydrophobic peptides: a new quorum-sensing mechanism in *streptococci*." Mol Microbiol **80**(4): 1102-19.

- Fontaine, L., C. Boutry, M. H. de Frahan, B. Delplace, C. Fremaux, P. Horvath, P. Boyaval et P. Hols (2010). "A novel pheromone quorum-sensing system controls the development of natural competence in *Streptococcus thermophilus* and *Streptococcus salivarius*." J Bacteriol **192**(5): 1444-54.
- Fujiwara, T., T. Hoshino, T. Ooshima, S. Sobue et S. Hamada (2000). "Purification, characterization, and molecular analysis of the gene encoding glucosyltransferase from *Streptococcus oralis*." Infect Immun **68**(5): 2475-83.
- Gallegos, M. T., R. Schleif, A. Bairoch, K. Hofmann et J. L. Ramos (1997). "AraC/XylS family of transcriptional regulators." Microbiol Mol Biol Rev **61**(4): 393-410.
- Galperin, M. Y. (2005). "A census of membrane-bound and intracellular signal transduction proteins in bacteria: Bacterial IQ, extroverts and introverts." BMC Microbiol **5**(35).
- Garces, F., F. J. Fernandez, A. M. Gomez, R. Perez-Luque, E. Campos, R. Prohens, J. Aguilar, L. Baldoma, M. Coll, J. Badia et M. C. Vega (2008). "Quaternary structural transitions in the DeoR-type repressor UlaR control transcriptional readout from the L-ascorbate utilization regulon in *Escherichia coli*." Biochemistry **47**(44): 11424-33.
- Giliberti, G., L. Baccigalupi, A. Cordone, E. Ricca et M. De Felice (2006). "Transcriptional analysis of the *recA* gene of *Streptococcus thermophilus*." Microb Cell Fact **5**: 29.
- Giraffa, G. et J. L. Bergère (1987). "Nature du caractère épaississant de certaines souches de *Streptococcus thermophilus*: étude préliminaire." Le lait **67**(3): 285-298.
- Giraffa, G., A. Paris, L. Valcavi, M. Gatti et E. Neviani (2001). "Genotypic and phenotypic heterogeneity of *Streptococcus thermophilus* strains isolated from dairy products." J Appl Microbiol **91**(5): 937-43.
- Goebel, M. et M. Yanagida (1991). "The TPR snap helix: a novel protein repeat motif from mitosis to transcription." Trends Biochem Sci **16**: 173-77.
- Goffin, C. et J. M. Ghuysen (1998). "Multimodular penicillin-binding proteins: an enigmatic family of orthologs and paralogs." Microbiol Mol Biol Rev **62**(4): 1079-93.
- Gottesman, S. (1984). "Bacterial regulation: global regulatory networks." Annu Rev Genet **18**: 415-41.
- Gottesman, S., S. Wickner et M. R. Maurizi (1997). "Protein quality control: triage by chaperones and proteases." Genes Dev. **11**: 815-23.
- Graham, M. R., L. M. Smoot, C. A. Migliaccio, K. Virtaneva, D. E. Sturdevant, S. F. Porcella, M. J. Federle, G. J. Adams, J. R. Scott et J. M. Musser (2002). "Virulence control in group A *Streptococcus* by a two-component gene regulatory system: global expression profiling and in vivo infection modeling." Proc Natl Acad Sci U S A **99**(21): 13855-60.
- Gralla, J. D. et J. Collado-Vides (1996). "Organization and function of transcription regulatory elements." *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology. A. Press.

1232–1245.

- Grandvalet, C., F. Coucheney, C. Beltramo et J. Guzzo (2005). "CtsR is the master regulator of stress response gene expression in *Oenococcus oeni*." J Bacteriol **187**(16): 5614-23.
- Graumann, P., T. M. Wendrich, M. H. Weber, K. Schröder et M. A. Marahiel (1997). "A family of cold shock proteins in *Bacillus subtilis* is essential for cellular growth and for efficient protein synthesis at optimal and low temperatures." Mol Microbiol **25**(4): 741-56.
- Gross, M., S. Hessefort, A. Olin et G. Reddy (1996). "Extensive sequencing of tryptic peptides of a rabbit reticulocyte 66-kDa protein that promotes recycling of Hsp 70. Homology To stress-related proteins." J Biol Chem **271**(28): 16842-9.
- Guédon, G., F. Bourgoïn, M. Pebay, Y. Roussel, C. Colmin, J. M. Simonet et B. Decaris (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?" Mol Microbiol **16**(1): 69-78.
- Gusa, A. A. et J. R. Scott (2005). "The CovR response regulator of group A *streptococcus* (GAS) acts directly to repress its own promoter." Mol Microbiol **56**(5): 1195-207.
- Gutierrez-Rois, R. M., D. A. Rosenblueth, J. A. Loza, A. M. Huerta, J. D. Glasner, F. R. Blattner et J. Collado-Vides (2003). "Regulatory network of *Escherichia coli*: consistency between literature knowledge and microarray profiles." Genome Res **13**: 2435-43.
- Haldenwang, W. G. (1995). "The sigma factors of *Bacillus subtilis*." Microbiol Rev **59**(1): 1-30.
- Hecker, M., W. Schumann and U. Volker (1996). "Heat-shock and general stress response in *Bacillus subtilis*." Mol Microbiol **19**(3): 417-28.
- Hoch, J. A. (1993a). "Regulation of the phosphorelay and the initiation of sporulation in *Bacillus subtilis*." Annu Rev Microbiol **47**: 441-65.
- Hoch, J. A. (1993b). "The phosphorelay signal transduction pathway in the initiation of *Bacillus subtilis* sporulation." J Cell Biochem **51**(1): 55-61.
- Hoch, J. A. (2002). "Control of signal transduction in the sporulation phosphorelay." Signal switches, regulons and cascades: control of bacterial gene expression. S. Symposium61. Cambridge: 199-212.
- Hols, P., F. Hancy, L. Fontaine, B. Grossiord, D. Prozzi, N. Leblond-Bourget, B. Decaris, A. Bolotin, C. Delorme, S. Dusko Ehrlich, E. Guédon, V. Monnet, P. Renault et M. Kleerebezem (2005). "New insights in the molecular biology and physiology of *Streptococcus thermophilus* revealed by comparative genomics." FEMS Microbiol Rev **29**(3): 435-63.
- Hoskisson, P. A. et S. Rigali (2009). "Chapter 1: Variation in form and function the helix-turn-helix regulators of the GntR superfamily." Adv Appl Microbiol **69**: 1-22.
- Ibrahim, M., A. Guillot, F. Wessner, F. Algaron, C. Besset, P. Courtin, R. Gardan et V. Monnet

BIBLIOGRAPHIE

- (2007b). "Control of the transcription of a short gene encoding a cyclic peptide in *Streptococcus thermophilus*: a new quorum-sensing system?" J Bacteriol **189**(24): 8844-54.
- Ibrahim, M., P. Nicolas, P. Bessieres, A. Bolotin, V. Monnet et R. Gardan (2007a). "A genome-wide survey of short coding sequences in *streptococci*." Microbiology **153**(Pt 11): 3631-44.
- Idigoras, P., A. Valiente, L. Iglesias, P. Trieu-Cout et C. Poyart (2001). "Meningitis due to *Streptococcus salivarius*." J Clin Microbiol **39**(8): 3017.
- Ikebe, T., M. Ato, T. Matsumura, H. Hasegawa, T. Sata, K. Kobayashi et H. Watanabe (2010). "Highly frequent mutations in negative regulators of multiple virulence genes in group A *streptococcal* toxic shock syndrome isolates." PLoS Pathog **6**(4): e1000832.
- Ishihama, A. (2000). "Functional modulation of *Escherichia coli* RNA polymerase." Annu Rev Microbiol **54**: 499-518.
- Ishino, F., W. Park, S. Tomioka, S. Tamaki, I. Takase, K. Kunugita, H. Matsuzawa, S. Asoh, T. Ohta, B. G. Spratt et M. Matsushashi (1986). "Peptidoglycan synthetic activities in membranes of *Escherichia coli* caused by overproduction of penicillin-binding protein 2 and rodA protein." J Biol Chem **261**(15): 7024-31.
- Jamet, E. (2009). "Les bactéries lactiques : une composante de l'écosystème microbien des fromages." Bactéries lactiques : physiologie, métabolisme, génomique et applications industrielles. Economica: 319- 348.
- Jeon, Y. H., T. Yamazaki, T. Otomo, A. Ishihama et Y. Kyogoku (1997). "Flexible linker in the RNA polymerase alpha subunit facilitates the independent motion of the C-terminal activator contact domain." J Mol Biol **267**(4): 953-62.
- Jiang, S. M., N. Ishmael, J. Dunning Hotopp, M. Puliti, L. Tissi, N. Kumar, M. J. Cieslewicz, H. Tettelin et M. R. Wessels (2008). "Variation in the group B *Streptococcus* CsrRS regulon and effects on pathogenicity." J Bacteriol **190**(6): 1956-65.
- Jiang, W., L. Fang et M. Inouye (1996). "Complete growth inhibition of *Escherichia coli* by ribosome trapping with truncated cspA mRNA at low temperature." Genes Cells **1**(11): 965-76.
- Johnsborg, O. et L. S. Havarstein (2009). "Regulation of natural genetic transformation and acquisition of transforming DNA in *Streptococcus pneumoniae*." FEMS Microbiol Rev **33**(3): 627-42.
- Kawamura, Y., X. G. Hou, F. Sultana, H. Miura et T. Ezaki (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): 406-8.
- Kerr A et M. J. Cowling (2003). "The effects of surface topography on the accumulation of biofouling." Philos Mag **83**:2779-95.
- Kim, H. J., C. Jourlin-Castelli, S. I. Kim et A. L. Sonenshein (2002). "Regulation of the *Bacillus subtilis* *ccpC* gene by *ccpA* and *ccpC*." Mol Microbiol **43**(2): 399-410.

- Klug, A. et D. Rhodes (1987). "Zinc fingers: a novel protein fold for nucleic acid recognition." Cold Spring Harb Symp Quant Biol. **52**: 473-82.
- Koonin, E. V., L. Aravind et M. Y. Galperin (2000). A comparative genomic view of the microbial stress response. Bacterial Stress Responses. A. press. 417-444.
- Kreikemeyer, B., K. S. McIver et A. Podbielski (2003). "Virulence factor regulation and regulatory networks in *Streptococcus pyogenes* and their impact on pathogen-host interactions." Trends Microbiol **11**(5): 224-32.
- Lancefield, R. (1993). "A serological of human and other groups of hemolytic *streptococci*." J Exp Med **57**(4): 571-95.
- Landschulz, W. H., P. F. Johnson et S. L. McKnight (1988). "The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins." Science **240**(4860): 1759-64.
- Laub, M. T. et M. Goulian (2007). "Specificity in two-component signal transduction pathways." Annu Rev Genet **41**: 121-45.
- Layec, S. (2008). "Le gène *cse*, de création récente, code une hydrolase du peptidoglycane impliquée dans la séparation des cellules de *Streptococcus thermophilus*." Thèse Université Henri Poincaré, Nancy 1.
- Layec, S., J. Gerard, V. Legue, M. P. Chapot-Chartier, P. Courtin, F. Borges, B. Decaris et N. Leblond-Bourget (2009). "The CHAP domain of Cse functions as an endopeptidase that acts at mature septa to promote *Streptococcus thermophilus* cell separation." Mol Microbiol **71**(5): 1205-17.
- Leblond-Bourget, N. et G. Guédon (2009). "Organisation et évolution des génomes des bactéries lactiques agroalimentaires." Bactéries Lactiques. Economica. 199-214.
- Lee, M. S. et D. A. Morrison (1999). "Identification of a new regulator in *Streptococcus pneumoniae* linking quorum sensing to competence for genetic transformation." J Bacteriol **181**(16): 5004-16.
- Lee, S. F., G. D. Delaney et M. Elkhateeb (2004). "A two-component covRS regulatory system regulates expression of fructosyltransferase and a novel extracellular carbohydrate in *Streptococcus mutans*." Infect Immun **72**(7): 3968-73.
- Lefebure, T. et M. J. Stanhope (2007). "Evolution of the core and pan-genome of *Streptococcus*: positive selection, recombination, and genome composition." Genome Biol **8**(5): R71.
- Li, M., Y. Lai, A. E. Villaruz, D. J. Cha, D. E. Sturdevant et M. Otto (2007). "Gram-positive three-component antimicrobial peptide-sensing system." Proc Natl Acad Sci U S A **104**(22): 9469-74.
- Linton, K. J. et C. F. Higgins (2007). "Structure and function of ABC transporters: the ATP switch provides flexible control." Pflugers Arch **453**: 555-567.

BIBLIOGRAPHIE

- Liu, M., R. J. Siezen et A. Nauta (2009). "In Silico Prediction of Horizontal Gene Transfer Events in *Lactobacillus bulgaricus* and *Streptococcus thermophilus* Reveals Protocooperation in Yogurt Manufacturing." Appl Environ Microbiol **75**(12): 4120-4129.
- Loones, A. (1994). "Laits fermentés par les bactéries lactiques." Bactéries lactiques : Aspects fondamentaux et technologiques. Loriga. **2**: 135- 154.
- Loughman, J. A. et M. G. Caparon (2007). "Contribution of invariant residues to the function of Rgg family transcription regulators." J Bacteriol **189**(2): 650-5.
- Luscombe, N. M., S. E. Austin, H. M. Berman et J. M. Thornton (2000). "A overview of the structures of protein-DNA complexes." Genome Biol **1**: REVIEWS001.
- Lyon, W. R., C. M. Gibson et M. G. Caparon (1998). "A role for trigger factor and an rgg-like regulator in the transcription, secretion and processing of the cysteine proteinase of *Streptococcus pyogenes*." Embo J **17**(21): 6263-75.
- Ma, B., T. Elkayam, H. Wolfson et R. Nussinov (2003). "Protein-protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces." Proc Natl Acad Sci U S A **100**: 5772-5777.
- Maddocks, S. E. et P. C. Oyston (2008). "Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins." Microbiology **154**(Pt 12): 3609-23.
- Makarova, K., A. Slesarev, Y. Wolf, A. Sorokin, B. Mirkin, E. Koonin, A. Pavlov, N. Pavlova, V. Karamychev, N. Polouchine, V. Shakhova, I. Grigoriev, Y. Lou, D. Rohksar, S. Lucas, K. Huang, D. M. Goodstein, T. Hawkins, V. Plengvidhya, D. Welker, J. Hughes, Y. Goh, A. Benson, K. Baldwin, J. H. Lee, I. Diaz-Muniz, B. Dosti, V. Smeianov, W. Wechter, R. Barabote, G. Lorca, E. Altermann, R. Barrangou, B. Ganesan, Y. Xie, H. Rawsthorne, D. Tamir, C. Parker, F. Breidt, J. Broadbent, R. Hutkins, D. O'Sullivan, J. Steele, G. Unlu, M. Saier, T. Klaenhammer, P. Richardson, S. Kozyavkin, B. Weimer et D. Mills (2006). "Comparative genomics of the lactic acid bacteria." Proc Natl Acad Sci U S A **103**(42): 15611-6.
- Marijuan, P. C., J. Navarro et R. del Moral (2010). "On prokaryotic intelligence: Strategies for sensing the environment." BioSystems **99**: 94-103.
- Martinez-Antonio, A. et J. Collado-Vides (2003). "Identifying global regulators in transcriptional regulatory networks in bacteria." Curr Opin Microbiol. **6**(5): 482-9.
- Martirani, L., R. Raniello, G. Naclerio, E. Ricca et M. De Felice (2001). "Identification of the DNA-binding protein, HrcA, of *Streptococcus thermophilus*." FEMS Microbiol Lett **198**(2): 177-82.
- Mercier, C., C. Durrieu, R. Briandet, E. Domakova, J. Tremblay, G. Buist et S. Kulakauskas (2002). "Positive role of peptidoglycan breaks in lactococcal biofilm formation." Mol Microbiol **46**(1): 235-43.
- Miller, A. A., N. C. Engleberg et V. J. DiRita (2001). "Repression of virulence genes by phosphorylation-dependent oligomerization of CsrR at target promoters in *S. pyogenes*."

- Mol Microbiol **40**(4): 976-90.
- Mogk, A. et W. Schumann (1997). "Cloning and sequencing of the *hrcA* gene of *Bacillus stearothermophilus*." Gene **194**(1): 133-6.
- Morrison, D. A. et M. S. Lee (2000). "Regulation of competence for genetic transformation in *Streptococcus pneumoniae*: a link between quorum sensing and DNA processing genes." Res Microbiol **151**(6): 445-51.
- Nair, S., I. Derre, T. Msadek, O. Gaillot et P. Berche (2000). "CtsR controls class III heat shock gene expression in the human pathogen *Listeria monocytogenes*." Mol Microbiol **35**(4): 800-11.
- Nakano, K., K. Fujita, K. Nishimura, R. Nomura et T. Ooshima (2005). "Contribution of biofilm regulatory protein A of *Streptococcus mutans*, to systemic virulence." Microbes Infect **7**(11-12): 1246-55.
- Neely, M. N., W. R. Lyon, D. L. Runft et M. Caparon (2003). "Role of RopB in growth phase expression of the SpeB cysteine protease of *Streptococcus pyogenes*." J Bacteriol **185**(17): 5166-74.
- Noirclerc-Savoie, M., C. Morlot, P. Gérard, T. Vernet et A. Zapun (2003). "Expression and purification of FtsW and RodA from *Streptococcus pneumoniae*, two membrane proteins involved in cell division and cell growth, respectively." Protein Expr Purif **30**(1): 18-25.
- Nystrom, T., et F. C. Neidhardt. (1994). "Expression and role of the universal stress protein, UspA, of *Escherichia coli* during growth arrest." Mol Microbiol **11**:537-544.
- Ortiz de Orué Lucana, D. et M. R. Groves (2009). "The three-component signalling system HbpS-SenS-SenR as an example of a redox sensing pathway in bacteria." Amino Acids **37**(3): 479-86.
- Pabo, C. O. et R. T. Sauer (1992). "Transcription factors: structural families and principles of DNA recognition." Annu Rev Biochem **61**: 1053-95.
- Parente, E. et T. Cogan (2004). Starter Cultures: General Aspects. London: Elsevier Academic Press. E. Ltd. 123- 148.
- Partridge, S. M. (2000). "Prosthetic valve endocarditis due to *Streptococcus vestibularis*." J Infect **41**(3): 284-5.
- Pavlovic, G., V. Burrus, B. Gintz, B. Decaris et G. Guedon (2004). "Evolution of genomic islands by deletion and tandem accretion by site-specific recombination: ICES_{St1}-related elements from *Streptococcus thermophilus*." Microbiology **150**(Pt 4): 759-74.
- Perez-Rueda, E. et J. Collado-Vides (2000). "The repertoires of DNA-binding transcription regulators in *Escherichia coli* K-12." Nucleic Acids Res **28**: 1838-47.
- Petersohn, A., M. Brigulla, S. Haas, J. D. Hoheisel, U. Volker et M. Hecker (2001). "Global analysis of the general stress response of *Bacillus subtilis*." J. Bacteriol **183**: 5617-31.

- Peterson, S. N., C. K. Sung, R. Cline, B. V. Desai, E. C. Snesrud, P. Luo, J. Walling, H. Li, M. Mintz, G. Tsegaye, P. C. Burr, Y. Do, S. Ahn, J. Gilbert, R. D. Fleischmann et D. A. Morrison (2004). "Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays." Mol Microbiol **51**(4): 1051-70.
- Pette, J. et H. Lolkema (1950). "Symbiosis and antibiosis in mixed cultures *Lb. bulgaricus* and *S. thermophilus*." Neth Milk Dairy Journal **4**: 197-208.
- Pfeiler, E. A. et T. R. Klaenhammer (2007). "The genomics of lactic acid bacteria." Trends Microbiol **15**(12): 546-53.
- Phillips, K. et S. E. Phillips (1994). "Electrostatic activation of *Escherichia coli* methionine repressor." Structure **2**(4): 309-16.
- Poyart, C., G. Quesne, S. Coulon, P. Berche et P. Trieu-Cuot (1998). "Identification of *streptococci* to species level by sequencing the gene encoding the manganese-dependent superoxide dismutase." J Clin Microbiol **36**(1): 41-7.
- Price, C. W., P. Fawcett, H. Ceremonie, N. Su, C. K. Murphy et P. Youngman (2001). "Genome-wide analysis of the general stress response in *Bacillus subtilis*." Mol. Microbiol. **41**: 757-74.
- Pulliainen, A. T., J. Hytonen, S. Haataja et J. Finne (2008). "Deficiency of the Rgg regulator promotes H₂O₂ resistance, AhpCF-mediated H₂O₂ decomposition, and virulence in *Streptococcus pyogenes*." J Bacteriol **190**(9): 3225-35.
- Qi, F., P. Chen et P. W. Caufield (1999). "Purification of mutacin III from group III *Streptococcus mutans* UA787 and genetic analyses of mutacin III biosynthesis genes." Appl Environ Microbiol **65**(9): 3880-7.
- Ramos, J. L., M. Martinez-Bueno, A. J. Molina-Henares, W. Teran, K. Watanabe, X. Zhang, M. T. Gallegos, R. Brennan et R. Tobes (2005). "The TetR family of transcriptional repressors." Microbiol Mol Biol Rev **69**(2): 326-56.
- Rampersaud, A., S. L. Harlocker et M. Inouye (1994). "The OmpR protein of *Escherichia coli* binds to sites in the ompF promoter region in a hierarchical manner determined by its degree of phosphorylation." J Biol Chem **269**(17): 12559-66.
- Rasmussen, T. B., M. Danielsen, O. Valina, C. Garrigues, E. Johansen et M. B. Pedersen (2008). "*Streptococcus thermophilus* core genome: comparative genome hybridization study of 47 strains." Appl Environ Microbiol **74**(15): 4703-10.
- Rawlinson, E. L., I. F. Nes et M. Skaugen (2002). "LasX, a transcriptional regulator of the lactocin S biosynthetic genes in *Lactobacillus sakei* L45, acts both as an activator and a repressor." Biochimie **84**(5-6): 559-67.
- Rawlinson, E. L., I. F. Nes et M. Skaugen (2005). "Identification of the DNA-binding site of the Rgg-like regulator LasX within the lactocin S promoter region." Microbiology **151**(Pt 3): 813-23.

- Ritossa, F. (1962). "A new puffing pattern induced by temperature shock and DNP." Drosophila Experientia **18**: 571-573.
- Rothfield, L., S. Justice et J. García-Lara (1999). "Bacterial cell division." Annu Rev Genet **33**: 423-48.
- Roussel, Y., M. Pebay, G. Guedon, J. M. Simonet et B. Decaris (1994). "Physical and genetic map of *Streptococcus thermophilus* A054." J Bacteriol **176**(24): 7413-22.
- Samen, U. M., B. J. Eikmanns et D. J. Reinscheid (2006). "The transcriptional regulator RovS controls the attachment of *Streptococcus agalactiae* to human epithelial cells and the expression of virulence genes." Infect Immun **74**(10): 5625-35.
- Sanders, J. W., K. Leenhouts, J. Burghoorn, J. R. Brands, G. Venema et J. Kok (1998). "A chloride-inducible acid resistance mechanism in *Lactococcus lactis* and its regulation." Mol Microbiol **27**(2): 299-310.
- Schleifer, K., M. Ehrmann, U. Krusch et H. Neve (1991). "Revival of the species *Streptococcus thermophilus* (ex Orla-Jensen, 1919) nom. rev." Syst Appl Microbiol **14**: 386-388.
- Schmidt, J. L., C. Tourneur et J. Lenoir (1994). "Fonctions et choix des bactéries lactiques en technologies laitières." Bactéries lactiques : Aspects fondamentaux et technologiques. Loriga. **2**: 37- 54.
- Schulz, A. et W. Schumann (1996). "*hrcA*, the first gene of the *Bacillus subtilis dnaK* operon encodes a negative regulator of class I heat shock genes." J Bacteriol **178**(4): 1088-93.
- Selye, H. (1975). "Implications of stress concept." N Y State J Med **75**(12): 2139-45.
- Siezen, R. J., B. Renckens, I. van Swam, S. Peters, R. van Kranenburg, M. Kleerebezem et W. M. de Vos (2005). "Complete sequences of four plasmids of *Lactococcus lactis* subsp. *cremoris* SK11 reveal extensive adaptation to the dairy environment." Appl Environ Microbiol **71**(12): 8371-82.
- Skaugen, M., E. L. Andersen, V. H. Christie et I. F. Nes (2002). "Identification, characterization, and expression of a second, bicistronic, operon involved in the production of lactocin S in *Lactobacillus sakei* L45." Appl Environ Microbiol **68**(2): 720-7.
- Storz, G. and R. Hengge-Aronis (2000). "Preface". Bacterial Stress Responses. A. Press.1
- Stover, C. K., X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrener, M. J. Hickey, F. S. Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L. Goltry, E. Tolentino, S. Westbrook-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter, K. R. Folger, A. Kas, K. Larbig, R. Lim, K. Smith, D. Spencer, G. K. Wong, Z. Wu, I. T. Paulsen, J. Reizer, M. H. Saier, R. E. Hancock, S. Lory et M. V. Olson (2000). "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen." Nature **406**(6799): 959-64.
- Sulavik, M. C. et D. B. Clewell (1996). "Rgg is a positive transcriptional regulator of the *Streptococcus gordonii gtfG* gene." J Bacteriol **178**(19): 5826-30.

- Sulavik, M. C., G. Tardif et D. B. Clewell (1992). "Identification of a gene, *rgg*, which regulates expression of glucosyltransferase and influences the Spp phenotype of *Streptococcus gordonii* Challis." J Bacteriol **174**(11): 3577-86.
- Sun, Z., X. Chen, J. Wang, W. Zhao, Y. Shao, L. Wu, Z. Zhou, T. Sun, L. Wang, H. Meng, H. Zhang et W. Chen (2011). "Complete Genome Sequence of *Streptococcus thermophilus* Strain ND03." J Bacteriol **193**(3): 793-4.
- Tailleux, P. (2001). "Les bactéries lactiques, ces êtres vivants apparus il y a près de 3 milliards d'années." Lait **81**: 1- 11.
- Tapp, J., M. Tholleson et B. Herrmann (2003). "Phylogenetic relationships and genotyping of the genus *Streptococcus* by sequence determination of the RNase P RNA gene, *rnpB*." Int J Syst Evol Microbiol **53**(Pt 6): 1861-71.
- Tayeb, J., C. Bouillanne et M. J. Desmazeaud (1984). "Computerized control of growth with temperature in a mixed culture of lactic acid bacteria." J. Ferment. Technol. **62**: 461-70.
- Tettelin, H., V. Maignani, M. J. Cieslewicz, J. A. Eisen, S. Peterson, M. R. Wessels, I. T. Paulsen, K. E. Nelson, I. Margarit, T. D. Read, L. C. Madoff, A. M. Wolf, M. J. Beanan, L. M. Brinkac, S. C. Daugherty, R. T. DeBoy, A. S. Durkin, J. F. Kolonay, R. Madupu, M. R. Lewis, D. Radune, N. B. Fedorova, D. Scanlan, H. Khouri, S. Mulligan, H. A. Carty, R. T. Cline, S. E. Van Aken, J. Gill, M. Scarselli, M. Mora, E. T. Iacobini, C. Brettoni, G. Galli, M. Mariani, F. Vegni, D. Maione, D. Rinaudo, R. Rappuoli, J. L. Telford, D. L. Kasper, G. Grandi et C. M. Fraser (2002). "Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V *Streptococcus agalactiae*." Proc Natl Acad Sci U S A **99**(19): 12391-6.
- Thibessard, A., F. Borges, A. Fernandez, B. Gintz, B. Decaris et N. Leblond-Bourget (2004). "Identification of *Streptococcus thermophilus* CNRZ368 genes involved in defense against superoxide stress." Appl Environ Microbiol **70**(4): 2220-9.
- Thibessard, A., A. Fernandez, B. Gintz, B. Decaris et N. Leblond-Bourget (2002). "Transposition of pGh9:ISS1 is random and efficient in *Streptococcus thermophilus* CNRZ368." Can J Microbiol **48**(5): 473-8.
- Thibessard, A., A. Fernandez, B. Gintz, N. Leblond-Bourget et B. Decaris (2001). "Hydrogen peroxide effects on *Streptococcus thermophilus* CNRZ368 cell viability." Res Microbiol **152**(6): 593-6.
- Thompson, J. D., D. G. Higgins et T. J. Gibson (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic Acids Res **22**(22): 4673-80.
- Tsakalidou, E., E. Zoidou, B. Pot, L. Wassill, W. Ludwig, L. A. Devriese, G. Kalantzopoulos, K. H. Schleifer et K. Kersters (1998). "Identification of *streptococci* from Greek Kasseri cheese and description of *Streptococcus macedonicus* sp. nov." Int J Syst Bacteriol **48 Pt 2**: 519-27.
- Vickerman, M. M. et P. E. Minick (2002). "Genetic analysis of the *rgg-gtfG* junctional region and

BIBLIOGRAPHIE

- its role in *Streptococcus gordonii* glucosyltransferase activity." Infect Immun **70**(4): 1703-14.
- Waters, C. M. et B. L. Bassler (2005). "Quorum sensing: cell-to-cell communication in bacteria." Annu Rev Cell Dev Biol **21**: 319–346.
- Wei, G. H., D. P. Liu et C. C. Liang (2004). "Charting gene regulatory networks: strategies, challenges and perspectives." Biochem J **381**: 1–12.
- Wigneshweraraj, S. R., A. Ishihama et M. Buck (2001). "In vitro roles of invariant helix-turn-helix motif residue R383 in sigma(54) (sigma(N))." Nucleic Acids Res **29**(5): 1163-74.
- Wilkinson, S. P. et A. Grove (2006). "Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins." Curr Issues Mol Biol **8**(1): 51-62.
- Yuan, G. et S. L. Wong (1995). "Isolation and characterization of *Bacillus subtilis* *groE* regulatory mutants: evidence for *orf39* in the *dnaK* operon as a repressor gene in regulating the expression of both *groE* and *dnaK*." J Bacteriol **177**(22): 6462-8.
- Yura, T., K. Nakahigashi et M. Kanemori (1996). "Transcriptional regulation of stress-inducible genes in procaryotes." EXS **77**: 165-81.
- Zhang, G. et S. A. Darst (1998). "Structure of the *Escherichia coli* RNA polymerase alpha subunit amino-terminal domain." Science **281**(5374): 262-6.
- Zheng, F., H. Ji, M. Cao, C. Wang, Y. Feng, M. Li, X. Pan, J. Wang, Y. Qin, F. Hu et J. Tang (2011). "Contribution of the Rgg transcription regulator to metabolism and virulence of *Streptococcus suis* serotype 2." Infect Immun **79**(13): 1319-28.
- Zotta, T., K. Asterinou, R. Rossano, A. Ricciardi, M. Varcamonti et E. Parente (2009). "Effect of inactivation of stress response regulators on the growth and survival of *Streptococcus thermophilus* Sfi39." Int J Food Microbiol **129**(3): 211-20.
- Zuber, U. et W. Schumann (1994). "CIRCE, a novel heat shock element involved in regulation of heat shock operon *dnaK* of *Bacillus subtilis*." J Bacteriol **176**(5): 1359-63.

ANNEXES

ANNEXES

Article

Publication :

The *rgg₀₁₈₂* gene encodes a transcriptional regulator required for the full *Streptococcus thermophilus* LMG18311 thermal adaptation.

Romain Henry, Emmanuelle Bruneau, Rozenn Gardan, Stéphane Bertin, Betty Fleuchot, Bernard Decaris, Nathalie Leblond-Bourget.

BMC Microbiology. Septembre 2011, en presse.

The *rgg₀₁₈₂* gene encodes a transcriptional regulator required for the full *Streptococcus thermophilus* LMG18311 thermal adaptation.

Romain Henry^{1, 2}, Emmanuelle Bruneau^{1, 2}, Rozenn Gardan³, Stéphane Bertin^{1, 2}, Betty Fleuchot³, Bernard Decaris^{1, 2} and Nathalie Leblond-Bourget^{1, 2*}

¹ INRA, UMR1128 Génétique et Microbiologie, F-54506 Vandœuvre-lès-Nancy Cedex, France.

² Université de Lorraine, UMR1128 Génétique et Microbiologie, F-54506 Vandœuvre-lès-Nancy Cedex, France.

³ INRA, UMR1319 MICALIS, Equipe Peptides et Communication Bactérienne, F-78352 Jouy en Josas, France.

*Corresponding author: Phone: (33) 3 83 68 42 10

Fax: (33) 3 83 68 44 99

Email: bourget@nancy.inra.fr

E-mail addresses:

Romain Henry Romain.Henry@scbiol.uhp-nancy.fr

Emmanuelle Bruneau Emmanuelle.Bruneau@scbiol.uhp-nancy.fr

Rozenn Gardan rozenn.gardan@jouy.inra.fr

Stéphane Bertin stephane.bertin@scbiol.uhp-nancy.fr

Betty Fleuchot Betty.Fleuchot@jouy.inra.fr

Bernard Decaris decaris@nancy.inra.fr

Nathalie Leblond-Bourget bourget@nancy.inra.fr

Abstract

Background: *Streptococcus thermophilus* is an important starter strain for the production of yogurt and cheeses. The analysis of sequenced genomes of four strains of *S. thermophilus* indicates that they contain several genes of the *rgg* family potentially encoding transcriptional regulators. Some of the Rgg proteins are known to be involved in bacterial stress adaptation.

Results: In this study, we demonstrated that *Streptococcus thermophilus* thermal stress adaptation required the *rgg₀₁₈₂* gene which transcription depends on the culture medium and the growth temperature. This gene encoded a protein showing similarity with members of the Rgg family transcriptional regulator. Our data confirmed that Rgg₀₁₈₂ is a transcriptional regulator controlling the expression of its neighboring genes as well as chaperones and proteases encoding genes. Therefore, analysis of a Δ *rgg₀₁₈₂* mutant revealed that this protein played a role in the heat shock adaptation of *Streptococcus thermophilus* LMG18311.

Conclusions: These data showed the importance of the Rgg₀₁₈₂ transcriptional regulator on the survival of *S. thermophilus* during dairy processes and more specifically during changes in temperature.

Background

The species *Streptococcus thermophilus* is a Lactic Acid Bacterium (LAB) used as a starter of fermentation in yogurt and cheese production. In nature and during dairy fermentation processes, *S. thermophilus* is subjected to sudden changes in its environment and its industrial performance is conditioned by its ability to successfully adapt to harsh conditions. To survive, like many other bacteria, this species must develop appropriate physiological responses by modifying gene expression appropriately.

One of the stresses, that *S. thermophilus* commonly encounters, is the modification of the temperature. For instance, during the production of dairy products, temperature shifts are applied to regulate the bacterial growth and, thus, control the lactic acid production [1]. *S. thermophilus* survival against thermal stress is conditioned by its ability to sense and quickly adapt its physiology mainly by the synthesis of adequate proteins at the right moment. For example, adaptation of *S. thermophilus* to a lowering of temperature required the synthesis of a set of chaperones called cold shock proteins (Csp) that is strongly induced in response to a rapid decrease in growth temperature [2, 3]. As in other Gram positive bacteria, *S. thermophilus* also responds to thermal stress by synthesizing a conserved set of heat-shock proteins (Hsp), including both chaperones and proteases [4]. Their role during heat stress is to rescue, or to scavenge, heat-denatured proteins. In LAB, the adaptation to heat shock required the two negative transcriptional factors HrcA [5] and CtsR [6] which negatively control the expression of major molecular chaperones (DnaK, DnaJ, GroEL and GroES) [7] and proteases (Clp family) [6], respectively. Furthermore, Zotta *et al.* (2009) have shown the involvement of the HrcA and CtsR proteins in the heat stress response of *S. thermophilus* Sfi39 [8]. Apart from these data, little is known about the network of regulation controlling *S. thermophilus* adaptation to temperature changes.

Among bacterial transcriptional regulators is the wide conserved family of Rgg regulators encoded by genes, exclusively found in the order of *Lactobacillales* and the family *Listeriaceae* [9]. Rgg regulators act by binding to the promoter region of their target genes [10-13]. At their N-terminal end, they carry a Helix-Turn-Helix (HTH) XRE DNA-binding domain demonstrated to be important for their activity as transcriptional regulators [14]. They are positive regulator [15, 16] or act both as activator and repressor [17, 18]. Most of the Rgg regulators control the transcription of their neighboring genes [9, 16, 19, 20]. However, Rgg from *S. pyogenes* NZ131, *S. agalactiae* NEM316 or *S. suis* SS2 are considered as global regulators since controlling highly diverse genes scattered on the genome [12, 13, 21, 22]. In these cases, Rgg proteins are involved in a network of regulation and modulate the expression of other transcriptional regulators, including several two-component regulatory systems, which are important in the transcriptional response to changing

environments [12, 13, 21]. Several Rgg proteins contribute to bacterial stress response. For instance, the Rgg protein of *Lactococcus lactis*, also known as GadR, is associated with glutamate-dependent acid tolerance [15]. Within *Streptococcus*, several Rgg proteins have been involved in oxidative- and/or to thermal-stress responses [23-25].

The high number of *rgg* genes observed in the genomes of *S. thermophilus* strains (7 in strains LMG18311 and CNRZ1066, 6 in LMD-9 and 5 in ND03) [26-28] suggests that their acquisition and their preservation are advantageous for *S. thermophilus*. However, the involvement of these genes in *S. thermophilus* LMG18311 stress response is still hypothetical and none of the 7 *rgg* genes of LMG18311 has been studied at the molecular level. To determine whether any of the *rgg* genes of *S. thermophilus* LMG18311 are involved in adaptation to changes in environmental conditions, Δ *rgg* deletion mutant was constructed and its tolerance to different stresses was tested. In this study, we demonstrate that (i) the transcription of *rgg*₀₁₈₂ gene from *S. thermophilus* LMG18311 is influenced by culture medium and growth temperature, (ii) Rgg₀₁₈₂ is a transcriptional regulator that modulate not only the transcription of its proximal target genes but is also involved in the network of regulation of the transcription of genes coding chaperones and proteases, (iii) this gene is involved in heat shock response.

Results

Analysis of the *rgg*₀₁₈₂ locus.

The *rgg*₀₁₈₂ gene corresponds to the *stu0182* gene of the complete genome sequence of *S. thermophilus* LMG18311 [26]. However, the sequencing of the *rgg*₀₁₈₂ gene (GenBank Accession JF699754) followed by its sequence comparison with the *stu0182* gene revealed that this latter contained sequencing discrepancies (substitutions in positions 44, 46, 50, 579 and 681 compared to the *rgg*₀₁₈₂ sequence). The *rgg*₀₁₈₂ gene (864 bp) potentially encodes a protein of 288 amino acids with a predicted molecular mass of 35.6 kDa. This protein exhibited an identity of about 30% with other streptococcal proteins belonging to the Rgg family of transcriptional regulators and 35% identity (e-value = $8e^{-48}$) with Rgg₁₃₅₈ from *S. thermophilus* LMD-9 which was recently shown to be involved in a quorum sensing (QS) mechanism [9]. Rgg₀₁₈₂ contained a HTH-XRE motif from amino acid 11 to 67 typical of Rgg regulators and a Rgg-C-terminal motif from amino acid 70 to 288 (Fig. 1). Therefore, the *rgg*₀₁₈₂ gene was predicted to encode a transcriptional regulator.

The gene *rgg*₀₁₈₂ was surrounded by two ORFs (Fig. 1), not annotated in the genome of the strain LMG18311, but revealed using the software bactgeneSHOW designed for small-gene detection [29]. Indeed, upstream of the *rgg*₀₁₈₂ gene was the *shp*₀₁₈₂ gene (63 nucleotides long), potentially encoding a small hydrophobic peptide belonging to the group I of the SHP family [9]. Downstream of *rgg*₀₁₈₂ was the *pep*₀₁₈₂ gene (42 nucleotides long), encoding a small peptide with no similarity with peptides found in databases. Although, the genetic organization of the *rgg*₀₁₈₂ locus was similar to that of the *rgg*₁₃₅₈ of the LMD-9 strain from *S. thermophilus*, these two loci were distinct as illustrated by the low sequence identity between the proteins encoded by them (Fig. 1). The two *shp* genes were classified in two distinct groups from the SHP family [9]. Finally, the *rgg*₀₁₈₂ locus and its flanking genes were also found in the genome of CNRZ1066 strain but missing in the genome of ND03 and LMD-9 strains.

Transcription analysis of the *rgg*₀₁₈₂ gene.

In the literature, studies of *rgg* genes transcription are scarce. Indeed, only the *ropB* transcription from *Streptococcus pyogenes* has been studied [10]. Thus, it was of interest to determine whether transcription of *rgg* was constitutive or not. To do so, the *rgg*₀₁₈₂ transcription was studied by qPCR from cells cultivated in LM17 or CDM medium at 30 or 42°C. In all qPCR experiments, the values were normalized to the expression of the *ldh* gene encoding the lactate dehydrogenase. This gene was considered as a relevant reference since it was demonstrated to be constitutively expressed in all tested conditions (data not shown). The qPCR experiments were realized from three independent RNA extracts and done in duplicate.

Fig. 2A showed the relative transcription levels of the *rgg0182* gene in the LMG18311 strain. When cultivated at 42°C in LM17 medium, the wild-type strain showed a significant decrease in its *rgg0182* mRNA levels during growth. Indeed, the *rgg0182* mRNA level was highest in the exponential phase (0.16 +/- 0.08) and was down-regulated 4-fold in stationary phase (p = 0.01). Similar results were obtained in CDM medium at 42°C where the transcription of *rgg0182* was found to be more than 3-fold higher in the exponential phase than in stationary phase (p<0.001). Whatever the medium tested, the transcription of *rgg0182* was found to be growth-phases dependent.

We then investigated whether *rgg0182* was transcribed at other temperatures and chose to work at 30°C, temperature at which *S. thermophilus* can be exposed during industrial processes (Fig. 2B). When cells were cultivated at 30°C in LM17, the profile of *rgg0182* transcripts was similar with that observed at 42°C. In contrast when cells were grown at 30°C in CDM medium, an increase of *rgg0182* transcription was observed during the growth, i.e. the *rgg0182* mRNA level was more than 3-fold higher (p<0.001) in stationary phase than in exponential phase. The *rgg0182* transcripts level of stationary phase cells grown in CDM medium was 14-fold higher (p<0.001) at 30 than at 42°C indicating it was on the influence of the growth temperature. Taken together, these results revealed that the kinetics of *rgg0182* transcription was medium and temperature dependent and that the transcript level of *rgg0182* was the highest in stationary phase cells cultivated in CDM at 30°C.

Effects of the Rgg₀₁₈₂ protein on the transcription of its flanking genes.

Data from the literature indicate that several products of *rgg* genes regulate adjacent genes [9, 16, 19, 20]. To determine whether the product of the *rgg0182* gene was involved in the transcriptional regulation of its flanking genes, we designed primers and used them in qPCR to measure the level of transcription of the *shp0182* and *pep0182* genes. However, the qPCR failed probably because of the small size of the *pep0182* and *shp0182* genes that did not allow the design of effective primers. As an alternative strategy, the activity of the *pep0182* and *shp0182* promoters was studied with transcriptional fusion in a wild-type and a Δ *rgg0182* background. To do so, plasmids carrying transcriptional fusions coupling the intergenic region of each flanking gene to a *luxAB*-reporter fusion were constructed and named pGICB004::P_{*pep0182*} and pGICB004::P_{*shp0182*}, as well as the Δ *rgg0182* strain carrying a chromosomal deletion of the *rgg0182* gene. Both plasmids were integrated in the wild-type or Δ *rgg0182* mutant chromosome. Relative levels of activity of the *pep0182* and *shp0182* promoters were determined in both strains either grown in LM17 or CDM (Fig. 3), at 30 or 42°C. Whatever the conditions, no significant difference in the growth rate or yield was observed between the wild type and the mutant. In LM17 at 30 and 42°C, almost no luciferase activity was detected with both promoters in the wild type or the Δ *rgg0182* background (data not shown). This suggests that the promoters P_{*pep0182*} and P_{*shp0182*} are not active in these experimental conditions. In contrast in CDM

medium, for both promoters in a wild type background, a luciferase activity was detected at 30 °C and 42°C (Fig. 3). Nevertheless, the maximum of $P_{pep0182-luxAB}$ and $P_{shp0182-luxAB}$ activity were 28- and 6-fold higher ($p < 0.001$) at 30°C than at 42°C, respectively. In addition, the level of activity of the $P_{pep0182-luxAB}$ and $P_{shp0182-luxAB}$ fusions differed between the wild-type and the mutant strains. Indeed, in cells cultivated in CDM at 30°C, in the $\Delta rgg0182$ mutant the $P_{pep0182-luxAB}$ and the $P_{shp0182-luxAB}$ showed both a maximum activity that was 3-fold lower ($p < 0.001$) than in the LMG18311 strain (Fig. 3). These results demonstrated that $rgg0182$ played a role in the regulation of the transcription of both $P_{pep0182-luxAB}$ and $P_{shp0182-luxAB}$ fusions and supported the hypothesis that Rgg_{0182} may, directly or not, regulate the transcription of pep_{0182} and shp_{0182} genes. Moreover, the growth medium, as described above and by Ibrahim *et al.* (2007b), and, in an original way, the temperature were parameters that influenced the levels of activity of the promoters $P_{pep0182}$ and $P_{shp0182}$.

Study of the binding of Rgg₀₁₈₂ protein of *S. thermophilus* upstream of *shp₀₁₈₂* and *pep₀₁₈₂* genes.

The results above and the presence of the HTH motif on the N-terminal part of Rgg_{0182} protein suggested that Rgg_{0182} may regulate the transcriptional activities of the $P_{shp0182}$ and the $P_{pep0182}$ promoters by binding to their DNA sequences. To test this hypothesis, DNA electrophoretic mobility shift assay were carried out. To do so, the His₆- Rgg_{0182} protein was overproduced in *E. coli* C41(DE3), verified by SDS-PAGE and Western blot (data not shown). Immobilized Metal ion Affinity Chromatography (IMAC) purification of the His₆- Rgg_{0182} protein was performed. The purity of the Rgg_{0182} protein was assessed by SDS-PAGE using Coomassie blue protein staining, i.e. only one band of the expected molecular mass (35.7 kDa) was revealed (data not shown).

A 126 bp PCR amplified DNA fragment (Fig. 1), including the entire 72 bp intergenic rgg_{0182} - shp_{0182} region and part of the 5' end of the shp_{0182} and rgg_{0182} genes, was incubated with the purified His₆- Rgg_{0182} protein. As can be seen in Fig 4, the Rgg_{0182} protein retarded the shp_{0182} promoter DNA fragment. The same experiment was realized with a 165 bp PCR amplified fragment, covering the entire 150 bp intergenic rgg_{0182} - pep_{0182} region including the pep_{0182} promoter, and analogous results were obtained (Fig. 4). The P_{ldh} probe corresponding to the promoter region of the *ldh* gene was chosen as a negative control in EMSA experiments since its expression was not under the control of Rgg_{0182} . Using P_{ldh} as a probe, no DNA retardation was observed, demonstrating that Rgg_{0182} binds specifically to the promoter of its target genes. Thus, these results demonstrated conclusively that Rgg_{0182} activated the shp_{0182} and pep_{0182} genes transcription by binding to their promoter regions.

Effects of the Rgg₀₁₈₂ protein on the transcription of genes encoding protease and chaperone proteins.

The impact of temperature on the *rgg₀₁₈₂* gene transcription suggested a role for the Rgg₀₁₈₂ protein on *S. thermophilus* LMG18311 adaptation to thermal changes. Thus, we hypothesized that Rgg₀₁₈₂ might control the transcription of genes encoding a set of heat- and cold-shock proteins including chaperones and proteases. Chaperones and ATP-dependent proteases play a major role for bacterial survival under conditions of heat stress where proteins tend to unfold and aggregate. Based upon the *S. thermophilus* LMG18311 genome sequence [26], genes predicted to encode the major chaperones and proteases involved in heat shock responses were selected for analysis: *clpC*, *dnaK*, *dnaJ*, *hsp33*, *groES*, *groEL*, *clpP*, *clpX*, *clpE*, *clpL* (Genbank Accession [NC_006448](#), locus tags *stu0077*, *stu0120-0121*, *stu0180*, *stu0203-0204*, *stu0356*, *stu0581*, *stu0602*, *stu1614*, respectively). The two *cspA* and *cspB* genes (locus tags, *stu0837-0838* respectively), involved in *S. thermophilus* cold stress response, were also included in this study. The transcript levels of these genes were measured by qPCR on stationary phase cells of the wild-type and the Δ *rgg₀₁₈₂* mutant grown in CDM medium at 30°C (i.e. when *rgg₀₁₈₂* was the most transcribed) from 3 independent experiments done in duplicate (Fig. 5). In these conditions, the transcript level of almost all genes encoding protease and chaperone proteins (except that of *dnaJ*, *groEL*, *cspA* and *cspB*) was under-expressed in the Δ *rgg₀₁₈₂* mutant compared to the wild type strain suggesting a role for Rgg₀₁₈₂ in the control of their transcription. The difference in the transcript abundance between the wild type and Δ *rgg₀₁₈₂* mutant strains ranged from 1.5- to 20-fold and were statistically significant ($P < 0.001$). As described in other *Streptococcus* transcriptional analysis, a 1.5-fold difference in transcript level was interpreted as a significant difference in expression between the strains [21, 23].

In low-GC Gram positive bacteria, the control of the transcription of the *clp* family genes and of *dnaK* and *groES* genes is primarily mediated by binding of the CtsR and HrcA repressors, respectively, to promoter region of target genes. In *S. thermophilus* LMG18311, we found CtsR operators (AGGTCAAANANAGGTCAAA) [6] upstream of *clpP*, *clpE*, *clpL*, *ctsR*, *clpC* and *groEL* genes and HrcA binding sites (GCACTC(N)₉GAGTGCTAA) [30] only upstream of *hrcA*, *groEL* (with 2 mismatches) and *dnaJ* (6 mismatches). These results prompted us to evaluate the level of *ctsR* and *hrcA* transcripts (locus tags, *stu0076* and *stu0118* respectively) in the wild-type and the Δ *rgg₀₁₈₂* mutant. These data revealed no significant difference for *ctsR* gene whereas the *hrcA* transcript level was nearly 4-fold reduced in the absence of *rgg₀₁₈₂* suggesting that Rgg₀₁₈₂ positively controls *hrcA* transcription.

These results indicate that Rgg₀₁₈₂ is a positive transcriptional regulator of heat shock proteins encoding genes in particular of *hrcA*, *clpC*, *clpE*, *clpL*, *clpP*, *clpX*, *dnaK*, *groES* and *hsp33* genes.

Role of the *rgg0182* gene in the heat shock response of *S. thermophilus*.

Knowing that several *rgg* genes from pathogenic streptococci are involved in stress response and taking into account the above data, we checked whether *rgg0182* could be involved in the *S. thermophilus* adaptation to heat shock. The heat tolerance was evaluated on stationary phase cells grown for 10 h in CDM medium ($OD_{600nm} = 1.5$) or in LM17 medium ($OD_{600nm} = 2.5$) at 30°C (where *rgg0182* was found to be higher or lower transcribed, respectively) before (control condition) and after a 15, 30, 45 and 60 minutes incubation at 52°C (temperature limit for growth of *S. thermophilus* LMG18311 in our laboratory conditions). The experiments were realized 3 times independently in triplicate. Using the LM17 medium (data not shown), no significant difference was observed between the strains. An exposure at 52°C, whatever its duration, resulted in a 20% decrease of the survival of both strains.

On the contrary, when stationary phase cells grown in CDM were exposed to a 52°C heat stress for up to 30 min, the mutant showed a significant increase of the sensibility compared to the wild type ($p < 0.001$) (Fig. 6). The heat tolerance of the $\Delta rgg0182$ mutant decreased gradually with the heat exposure time (72%, 53%, 46% and 38% of survival at 15, 30, 45 and 60 minutes, respectively). Between both strains, a difference of survival was observed at 30, 45 and 60 minutes where the mutant was up to 1.75 fold less resistant than the wild type strain. Thus, the decreased of survival of the mutant show that *rgg0182* plays a role in *S. thermophilus* adaptation to heat stress.

The Rgg₀₁₈₂ protein of *S. thermophilus* LMG18311 is involved in the transcription regulation of *clpE* and *cspB* genes in heat stress condition.

The impairment of the survival of the $\Delta rgg0182$ mutant cells following a sudden increase in temperature suggested that the *rgg0182* gene may act to regulate the transcription of *S. thermophilus* genes involved in the heat shock response. To investigate a possible role for Rgg₀₁₈₂ in changes of the transcription of heat shock genes, the transcript level of genes encoding chaperones and proteases were measured by qPCR. The transcript levels of the 14 selected stress-responsive genes were studied, in three independent experiments done in duplicate, on stationary cells of the wild-type and the $\Delta rgg0182$ mutant grown in CDM and exposed 30 minutes at 52°C. Our results showed that *clpE* and *cspB* genes were about 2-fold less and 3-fold more transcribed, respectively, in the mutant strain compared to wild-type ($p < 0.001$) (Fig. 7). No significant difference was observed for the other genes studied (data not shown). This observation suggests that at high temperature, Rgg₀₁₈₂ is a positive regulator of *clpE* transcription and a negative regulator of *cspB* transcription.

Discussion

The aim of the present study was to determine if Rgg₀₁₈₂ functioned as a transcriptional regulator. First, we showed that it was transcribed in a growth phase dependent manner i.e., in LM17 (at 30°C and 42°C) or CDM (at 42°C), a higher expression level was observed in exponential phase than in stationary phase. Interestingly, using CDM medium, it was found that the *rgg₀₁₈₂* transcripts were more abundant at 30°C than at 42°C suggesting that *rgg₀₁₈₂* transcription was also influenced by temperature. Because of their immediate vicinity with the *rgg₀₁₈₂* gene, the transcription of *shp₀₁₈₂* and *pep₀₁₈₂* genes was hypothesized to be under the control of Rgg₀₁₈₂. This was confirmed by the use of transcriptional fusions showing that the activation of the P_{*shp₀₁₈₂*} and P_{*pep₀₁₈₂*} promoters required the presence of Rgg₀₁₈₂ and that their activity was optimal under the conditions where transcription of the *rgg₀₁₈₂* gene was mostly expressed (i.e. in CDM medium at 30°C in stationary phase growth). Finally, to confirm the probable interaction of Rgg₀₁₈₂ with DNA, EMSA experiments were carried out and demonstrated conclusively that Rgg₀₁₈₂ binds to the promoter region of the *shp₀₁₈₂* and *pep₀₁₈₂* target genes. Together these results were in coherence with Rgg₀₁₈₂ being a transcriptional regulator, positively and directly, controlling the expression of *shp₀₁₈₂* and *pep₀₁₈₂* genes.

The *rgg₀₁₈₂* locus combined a gene encoding a transcriptional regulator of the Rgg family with another gene encoding a small hydrophobic peptide of the SHP family. Recently, one of these *shp/rgg* loci, named *shp/rgg₁₃₅₈* in LMD-9 has been demonstrated to encode two components of a novel QS mechanism [9]. This system involves a Rgg transcriptional regulator and a SHP pheromone that is detected and reimported into the cell by the Ami oligopeptide transporter. The target gene of the *shp₁₃₅₈/rgg₁₃₅₈* pair, called *pep_{1357C}*, is located just downstream of the *rgg₁₃₅₈* gene, and encodes a secreted cyclic peptide [31]. By analogy with the Shp₁₃₅₈/Rgg₁₃₅₈ locus, we hypothesize that the SHP₀₁₈₂/Rgg₀₁₈₂ pair would also be involved in a QS mechanism with Shp₀₁₈₂ being a pheromone possibly controlling the activation of the Rgg₀₁₈₂. Thus, one hypothesis to explain the high *rgg₀₁₈₂* transcript level in CDM is to consider that the *rgg₀₁₈₂* transcription would be under the control of the Rgg₀₁₈₂/SHP₀₁₈₂ pair and to consider that in a medium free of peptide, the internalization of *shp₀₁₈₂* by Ami transporter would be facilitated. Moreover, the mechanism of *rgg₀₁₈₂* expression seemed to be more complex than that of *rgg₁₃₅₈* since not only influenced by the culture medium but also by the temperature. Further experiments will be done (i) to determine whether the QS mechanism involving the SHP₁₃₅₈ and the Rgg₁₃₅₈ can be generalized to other SHP/Rgg pairs, including SHP₀₁₈₂/Rgg₀₁₈₂ pair and (ii) to understand the mechanism by which temperature could influence the *rgg₀₁₈₂* expression.

On the other hands, induction of the *rgg₀₁₈₂* expression at 30°C suggests that this gene might

participate in the physiological adaptation of *S. thermophilus* to this temperature. When cells were cultivated in CDM at 30°C, the inactivation of *rgg0182* was associated with a reduce expression of genes encoding chaperone and protease proteins. In *Bacillus subtilis*, the DnaKJ complex facilitates substrates folding to the native state and the GroESL complex provides an isolated environment for the proper folding of small protein substrates [32]. The degradation of unfolded proteins and small peptides is ensured by a protease complex composed of the protease subunit ClpP and several ATPases of the Clp family [32]. Thus, the Rgg₀₁₈₂ is a transcriptional regulator whose biological roles would be to control the homeostasis of chaperone and protease proteins in cells grown at 30°C in CDM. This is in concordance with data obtained in *S. pyogenes* where Rgg is found (at the protein level) to control the expression of ClpL, ClpP, GroEL and DnaK in stationary phase (4). Furthermore, it was shown that ClpL protein of *S. thermophilus* Sfi39 is necessary for correct response to both heat and cold stresses [4].

Results of qPCR experiments also showed an effect of Rgg0182 on *hrcA* expression. However, preliminary EMSA results (data not shown) indicated that the Rgg0182 protein did not bind to the *hrcA* promoter region. This suggests that the transcription of *hrcA* obviously is stimulated by Rgg₀₁₈₂ indirectly, perhaps by influencing the expression of another regulatory protein. Such indirect regulation has already been reported for other Rgg proteins [12, 13, 21] and, in the present study, might be extended to, at least, some of the *rgg0182* distal target genes.

Finally, to assess the significance of Rgg-associated changes in the expression of genes involved in the heat shock response, we checked whether the deletion of *rgg0182* had an impact on the survival of the strains under heat stress conditions (shift from 30°C to 52°C for 15 min to 60 min). Interestingly, an impaired survival of the mutant was observed but only when the cells were cultivated in the CDM medium, i.e. in conditions where the difference in the level of *rgg0182* transcripts was maximal between both strains. In the mutant cultivated in CDM, the percent of survival decreased with the duration of the heat exposure. These results demonstrated a role for Rgg₀₁₈₂ in *S. thermophilus* fitness in response to sudden increased of the temperature. As observed in other streptococcal strains [24, 25], the deletion of the *rgg0182* gene is not associated with a drastic modification of the survival to stress suggesting that this regulator is not essential but important for heat stress adaptation. Furthermore, our results showed that *cspB* and *clpE* genes were 2-fold lower and 3-fold higher, respectively, in the mutant compared to the wild-type strain after the heat stress. Data from literature indicate that most Csp proteins are required when cells are grown at low growth temperature [2, 3]. Thus, the Rgg₀₁₈₂ would negatively control the production of CspB when the latter is not required. Moreover, in *S. pneumoniae*, the *clpE* gene has been demonstrated to be required for thermo-tolerance [33], therefore we hypothesize that the heat sensitivity of the *S. thermophilus* Δ *rgg0182* mutant would result, at least partially, from a reduced level of ClpE

expression. Alternatively, it is also conceivable that Rgg₀₁₈₂ regulates the transcription of other genes encoding proteins involved in the *S. thermophilus* heat stress response. A transcriptomic analysis would identify all targets of this regulator within *S. thermophilus* LMG18311.

Conclusions

In conclusion, our study gave a better understanding of the thermal adaptation of the important dairy starter, *S. thermophilus*. These data showed the importance of the Rgg₀₁₈₂ transcriptional regulator on the survival of *S. thermophilus* during industrial processes and more specifically during changes in temperature.

Methods

Bacterial strains, media and reagents.

Streptococcus thermophilus LMG18311 and its derivatives are presented in Table 1. *S. thermophilus* strains were grown at 30 or 42°C in M17 medium with lactose (10 g/l) (LM17, a classical medium for *S. thermophilus* growth) [34] or in a chemically defined medium (CDM, a peptide free-medium) [35]. Pre-cultures were incubated at 42°C in milk medium except for the luciferase assays as mentioned below. For numeration, agar was added to the medium (15 g/l) and cells were incubated under anaerobic conditions using GENbox anaer in Generbox jars (bioMérieux SA, Marcy-l'Etoile, France). *S. thermophilus* strains containing the pG⁺host9 vector [36] were cultivated in the presence of erythromycin (final concentration 2 µg/ml) at 30°C when plasmid self-maintenance was required and at 42°C for selection of clones with the chromosome's integrated plasmid.

E. coli strains were grown in Luria-Bertani (LB) medium [37] at 30 or 37°C with shaking (250 rpm). *E. coli* EC101 [38] was used as a host for recombinant plasmids derived from pG⁺host9 [36] and *E. coli* C41(DE3) [39] for recombinant plasmids derived from pET15b (Novagen). When required, erythromycin (150 µg/ml) or ampicillin (150µg/ml) was added.

DNA and RNA manipulations.

Conventional techniques for DNA manipulation, such as preparation of chromosomal and plasmidic DNA, restriction enzymes digests, PCR experiments, transformation by electroporation and Southern blotting were performed as described [37]. Plasmids and primers used in this study are listed in Table 1 and Table 2, respectively. Oligonucleotides were purchased from Eurogentec (Liège, Belgique). Sequencing reactions were done from PCR products by the Beckman Coulter Genomics Society (Grenoble, France) using an ABI 3730XL Sanger sequencing platform.

Plasmids and strains construction.

The Δ rgg₀₁₈₂ mutant was constructed following the experimental procedure described by Layec *et al.* (2009) [40] using the following primers pairs: rgg₀₁₈₂I1/rgg₀₁₈₂I2 and rgg₀₁₈₂II1/rgg₀₁₈₂II2 to amplify fragments I and II, respectively. The resulting Δ rgg₀₁₈₂ mutant was checked by sequencing.

The construction of pGICB004::P_{pep0182} and pGICB004::P_{shp0182} vectors was done as described in [9]. Briefly, the *pep0182* and *shp0182* promoters were amplified by PCR using the couple of oligonucleotides Pep0182up/Pep0182down and SHP0182up/SHP0182down, respectively. Following a digestion with restriction enzymes *SpeI/EcoRI*, both fragments were independently inserted downstream from *luxAB* in the pGICB004 plasmid previously digested with the same restriction enzymes. The pGICB004 plasmid was used since it allows the integration of

transcriptional fusions to the *luxAB* reporter genes at the *blp* locus in *S. thermophilus* [9]. The final pGICB004::P_{pep0182} and pGICB004::P_{shp0182} plasmids were used to transform the LMG18311 and in its derivative Δ *rgg0182* strain. The strains LMG18311 *blp*::P_{pep0182}-*luxAB*, LMG18311 *blp*::P_{shp0182}-*luxAB*, Δ *rgg182* *blp*::P_{pep0182}-*luxAB* and Δ *rgg182* *blp*::P_{shp0182}-*luxAB* were obtained by gene replacement at the *blp* locus as described in [9].

For the construction of the pET15b::*rgg0182* plasmid necessary for the Rgg₀₁₈₂-His₆-tagged protein overexpression, the *rgg0182* (864 bp) gene was amplified by PCR using the primers Rgg₀₁₈₂SE1 and Rgg₀₁₈₂SE2 which contained *Nde*I and *Bam*HI restriction sites, respectively, with *S. thermophilus* LMG18311 chromosomal DNA as template. The amplified DNA fragment was double digested with restriction enzymes *Nde*I/*Bam*HI and fused in-frame to the 3'end His₆-tagged of pET15b vector (Novagen), generating the pET15b::*rgg0182* plasmid. The in-frame fusion was confirmed by DNA sequencing.

Luciferase assays.

To perform luciferase assays, pre-cultures were grown overnight at 30 or 42°C in CDM or LM17 medium. Pre-cultures were then diluted to an OD_{600nm} of 0.05, in 50 ml of respective appropriate medium and temperature. A volume of 1 ml of the culture was sampled at regular intervals during the growth until the stationary phase and analyzed as follows: OD_{600nm} was measured, 10 μ L of a 0.1% nonyl-aldehyde solution was added to the sample and the luminescence was measured with a Luminoskan TL (Labsystems). Results are reported in relative luminescence divided by the OD_{600nm} (AU). Three independent experiments were realized.

Overexpression, purification of Rgg₀₁₈₂-His₆-tagged protein and Western blotting.

Expression of the His₆-tagged protein was induced in *E. coli* C41(DE3) containing pET15b::*rgg0182* for 4h at 30°C by adding Isopropyl β ,D-thiogalactopyranoside (IPTG, 1mM final concentration) to the OD_{600nm} = 0.5 culture. Cells were harvested by centrifugation at 14,000 rpm, at 4°C for 30 min. The supernatant was discarded and cells were suspended in lysis buffer (50 mM phosphate sodium pH 8.0, 300 mM NaCl, and 10 mM imidazol) and stored at -20°C. The cells were disrupted on ice with a microtip of Sonifier 250 (Branson Ultrasonics). The soluble fraction including the recombinant His₆-tagged protein was collected by centrifugation at 20,000 rpm for 45 min at 4°C and loaded on an affinity chromatography column equilibrated with lysis buffer. When the UV absorbance at 280 nm had fallen to the zero baseline, the recombinant Rgg₀₁₈₂ protein was eluted by elution buffer (50 mM phosphate sodium pH 8.0, 300 mM NaCl, 250 mM imidazol). The eluted fraction was collected and finally concentrated in Tris EDTA buffer pH 8.0. The purity of the His₆-tagged proteins was confirmed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis

(SDS-PAGE) using 15% acrylamide resolving.

For Western blot experiments, proteins were size separated by SDS-PAGE 12% acrylamide resolving gel and electroblotted onto polyvinylidene difluoride (PVDF) membrane (Roche Applied Science) using a semi-dry blotting system (Bio-Rad). After transfer, the PVDF membrane was blocked with 5% skim milk in Tris-buffered saline containing 0.1% tween 20 (TBS-T) for 1 h. The membrane was subsequently incubated for 1 h with penta-His antibodies (1:10,000) (Qiagen), washed three times with TBS-T and incubated for 1 h with conjugated goat anti-mouse immunoglobulin G (H + L)-horseradish peroxidase (1:10,000) (Bio-Rad). The membrane was washed three times with TBS-T. Finally, the antibodies-antigen complexes were visualized by using the Bio-Rad Immun-Star Western system as described by the manufacturer and were detected by chemiluminescence, using the Chemi-Doc from Bio-Rad.

DNA electrophoretic mobility shift assay (EMSA).

The DNA binding of the His6-tagged Rgg₀₁₈₂ protein to the *shp*₀₁₈₂ and *pep*₀₁₈₂ promoter regions was tested by EMSA using the LightShift Chemiluminescent EMSA Kit (Thermo Scientific). The promoter regions of *ldh* (P_{ldh} , 110pb), *shp*₀₁₈₂ ($P_{shp0182}$, 126 bp) and *pep*₀₁₈₂ ($P_{pep0182}$, 165 bp) were amplified by PCR using the P_{ldh}-5'/P_{ldh}-3', P_{shp}-3'/P_{shp}-5' and P_{pep}-3'/ P_{pep}-5' primers, respectively. These were 3'-end biotin labelled with Biotin 3' End DNA Labeling Kit (Thermo Scientific) and used in EMSA according to the manufacturer's instructions. Chemiluminescent detection of biotin DNA on membranes was realised with the Chemi-Doc apparatus (Bio-Rad).

RNA extraction and quantitative RT-PCR (qPCR) experiments.

RNA extractions were adapted from Kieser *et al.* (1999) [41]. RNAs were extracted from cultures grown in CDM or LM17 medium in exponential, transition, or stationary phase at 30 or 42°C. RNAs were also extracted from stationary phase cells exposed to a 30 min temperature shift from 30 to 52°C. The RNAs were treated with amplification grade DNase I (Euromedex). The quantity and quality of the RNA samples were verified by agarose gel electrophoresis and by measuring their absorbance at 260 and 280 nm (NanoDrop-1000). Reverse transcription was performed according to the manufacturer's instructions (MMLV-reverse transcriptase, Invitrogen). cDNA was generated from 1.25 µg of DNA-free RNA and used for qPCR analysis of transcription of *rgg*₀₁₈₂ gene and its potential target genes transcript levels. Gene transcripts quantification was done using the CFX96 manager software (Bio-Rad) with the following program: 1 cycle at 98°C for 3 min and 40 cycles at 95°C for 10 s and at 58°C for 45 s. The amplification reactions were carried out with SYBR Green Supermix (Bio-Rad). Melting curve analysis was performed with 0.5°C increments every 10 s from 55 to 95°C to check that the cDNA amplification did not lead to secondary products. The primers

used for qPCR are listed in Table 2. The efficiency of all primers pairs was checked in qPCR using serial dilutions of cDNA, and ranged from 90 to 100%. The level of gene transcript was calculated with *ldh* gene as the internal control gene for normalization [23].

Physiological characterization of the Δ *rgg*₀₁₈₂ mutant.

Stationary phase cells were harvested from cultures grown in CDM at 30°C by centrifugation at 4,500 rpm for 10 min. Cells were washed twice and resuspended in 10 mM sterile phosphate buffer, pH 7.0 with a final OD_{600nm} of 1.0. Then, for heat stress, cells were treated by incubation at 52°C during 15, 30, 45 and 60 min (heat stress condition) or not (control condition). Cultures were then diluted to appropriate concentrations, spread on LM17 agar plates and incubated overnight at 42°C under anaerobic conditions. A percentage of survival was calculated as $N / N_0 \times 100$ where N_0 is the CFU number of the control condition and N the CFU number after heat stress condition. Three independent experiments done in triplicate were realized.

Statistical analysis.

Data are expressed as mean +/- standard deviation (SD). Statistical analysis was performed with Student's *t* test. A *p* value <0.05 was considered statistically different.

Nucleotide sequence accession number.

The DNA sequence reported in this paper has been deposited in GenBank under accession number JF699754.

Authors' contributions

Conceived and designed the experiments: RH EB BD NL. Performed the experiments: RH EB RG SB BF. Analyzed the data: RH EB RG BF NL. Wrote the paper: RH EB NL. All authors read and approved the final manuscript.

Acknowledgments

This study was supported by the Institut National de la Recherche Agronomique (INRA) and the Ministère de l'Éducation Nationale de la Recherche et de la Technologie (MENRT). We thank N. Rouhier for his technical advices and his technical supports. We thank S. Payot-Lacroix and M. Genay-Bernard for critical reading of the manuscript.

References

1. Kosikoski, F.V. and V.V. Mistry, **Volume 1: Origins and Principles**, in Cheese and Fermented Milk Foods, r.e. Westport, Editor. 1997.
2. Wouters, J.A., F.M. Rombouts, W.M. de Vos, O.P. Kuipers, and T. Abee, **Cold shock proteins and low-temperature response of *Streptococcus thermophilus* CNRZ302**. Appl Environ Microbiol, 1999. **65**(10): p. 4436-42.
3. Perrin, C., C. Guimont, P. Bracquart, and J.L. Gaillard, **Expression of a new cold shock protein of 21.5 kDa and of the major cold shock protein by *Streptococcus thermophilus* after cold shock**. Curr Microbiol, 1999. **39**(6): p. 342-0347.
4. Varcamonti, M., S. Arsenijevic, L. Martirani, D. Fusco, G. Naclerio, and M. De Felice, **Expression of the heat shock gene *clpL* of *Streptococcus thermophilus* is induced by both heat and cold shock**. Microb Cell Fact, 2006. **5**: p. 6.
5. Martirani, L., R. Raniello, G. Naclerio, E. Ricca, and M. De Felice, **Identification of the DNA-binding protein, HrcA, of *Streptococcus thermophilus***. FEMS Microbiol Lett, 2001. **198**(2): p. 177-82.
6. Derre, I., G. Rapoport, and T. Msadek, **CtsR, a novel regulator of stress and heat shock response, controls *clp* and molecular chaperone gene expression in gram-positive bacteria**. Mol Microbiol, 1999. **31**(1): p. 117-31.
7. Kilstrup, M., S. Jacobsen, K. Hammer, and F.K. Vogensen, **Induction of heat shock proteins DnaK, GroEL, and GroES by salt stress in *Lactococcus lactis***. Appl Environ Microbiol, 1997. **63**(5): p. 1826-37.
8. Zotta, T., K. Asterinou, R. Rossano, A. Ricciardi, M. Varcamonti, and E. Parente, **Effect of inactivation of stress response regulators on the growth and survival of *Streptococcus thermophilus* Sfi39**. Int J Food Microbiol, 2009. **129**(3): p. 211-20.
9. Fleuchot, B., C. Gitton, A. Guillot, J. Vidic, P. Nicolas, C. Besset, L. Fontaine, P. Hols, N. Leblond-Bourget, V. Monnet, and R. Gardan, **Rgg proteins associated with internalized small hydrophobic peptides: a new quorum-sensing mechanism in *streptococci***. Mol Microbiol, 2011. **80**(4): p. 1102-19.
10. Neely, M.N., W.R. Lyon, D.L. Runft, and M. Caparon, **Role of RopB in growth phase expression of the SpeB cysteine protease of *Streptococcus pyogenes***. J Bacteriol, 2003. **185**(17): p. 5166-74.
11. Vickerman, M.M. and P.E. Minick, **Genetic analysis of the *rgg-gtfG* junctional region and its role in *Streptococcus gordonii* glucosyltransferase activity**. Infect Immun, 2002. **70**(4): p. 1703-14.
12. Samen, U.M., B.J. Eikmanns, and D.J. Reinscheid, **The transcriptional regulator RovS controls the attachment of *Streptococcus agalactiae* to human epithelial cells and the expression of virulence genes**. Infect Immun, 2006. **74**(10): p. 5625-35.
13. Chaussee, M.S., G.L. Sylva, D.E. Sturdevant, L.M. Smoot, M.R. Graham, R.O. Watson, and J.M. Musser, **Rgg influences the expression of multiple regulatory loci to coregulate virulence factor expression in *Streptococcus pyogenes***. Infect Immun, 2002. **70**(2): p. 762-

14. Loughman, J.A. and M.G. Caparon, **Contribution of invariant residues to the function of Rgg family transcription regulators**. J Bacteriol, 2007. **189**(2): p. 650-5.
15. Sanders, J.W., K. Leenhouts, J. Burghoorn, J.R. Brands, G. Venema, and J. Kok, **A chloride-inducible acid resistance mechanism in *Lactococcus lactis* and its regulation**. Mol Microbiol, 1998. **27**(2): p. 299-310.
16. Sulavik, M.C., G. Tardif, and D.B. Clewell, **Identification of a gene, *rgg*, which regulates expression of glucosyltransferase and influences the Spp phenotype of *Streptococcus gordonii* Challis**. J Bacteriol, 1992. **174**(11): p. 3577-86.
17. Chaussee, M.S., R.O. Watson, J.C. Smoot, and J.M. Musser, **Identification of Rgg-regulated exoproteins of *Streptococcus pyogenes***. Infect Immun, 2001. **69**(2): p. 822-31.
18. Rawlinson, E.L., I.F. Nes, and M. Skaugen, **LasX, a transcriptional regulator of the lactocin S biosynthetic genes in *Lactobacillus sakei* L45, acts both as an activator and a repressor**. Biochimie, 2002. **84**(5-6): p. 559-67.
19. Chaussee, M.S., D. Ajdic, and J.J. Ferretti, **The *rgg* gene of *Streptococcus pyogenes* NZ131 positively influences extracellular SPE B production**. Infect Immun, 1999. **67**(4): p. 1715-22.
20. Lyon, W.R., C.M. Gibson, and M.G. Caparon, **A role for trigger factor and an *rgg*-like regulator in the transcription, secretion and processing of the cysteine proteinase of *Streptococcus pyogenes***. Embo J, 1998. **17**(21): p. 6263-75.
21. Zheng, F., H. Ji, M. Cao, C. Wang, Y. Feng, M. Li, X. Pan, J. Wang, Y. Qin, F. Hu, and J. Tang, **Contribution of the Rgg transcription regulator to metabolism and virulence of *Streptococcus suis* serotype 2**. Infect Immun, 2010. **79**(13): p. 1319-28.
22. Chaussee, M.S., G.A. Somerville, L. Reitzer, and J.M. Musser, **Rgg coordinates virulence factor synthesis and metabolism in *Streptococcus pyogenes***. J Bacteriol, 2003. **185**(20): p. 6016-24.
23. Fernandez, A., A. Thibessard, F. Borges, B. Gintz, B. Decaris, and N. Leblond-Bourget, **Characterization of oxidative stress-resistant mutants of *Streptococcus thermophilus* CNRZ368**. Arch Microbiol, 2004a. **182**(5): p. 364-72.
24. Bortoni, M.E., V.S. Terra, J. Hinds, P.W. Andrew, and H. Yesilkaya, **The pneumococcal response to oxidative stress includes a role for Rgg**. Microbiology, 2009. **155**(Pt 12): p. 4123-34.
25. Chaussee, M.A., E.A. Callegari, and M.S. Chaussee, **Rgg regulates growth phase-dependent expression of proteins associated with secondary metabolism and stress in *Streptococcus pyogenes***. J Bacteriol, 2004. **186**(21): p. 7091-9.
26. Bolotin, A., B. Quinquis, P. Renault, A. Sorokin, S.D. Ehrlich, S. Kulakauskas, A. Lapidus, E. Goltsman, M. Mazur, G.D. Pusch, M. Fonstein, R. Overbeek, N. Kyprides, B. Purnelle, D. Prozzi, K. Ngui, D. Masuy, F. Hancy, S. Burteau, M. Boutry, J. Delcour, A. Goffeau, and P. Hols, **Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus***. Nat Biotechnol, 2004. **22**(12): p. 1554-8.

27. Makarova, K., A. Slesarev, Y. Wolf, A. Sorokin, B. Mirkin, E. Koonin, A. Pavlov, N. Pavlova, V. Karamychev, N. Polouchine, V. Shakhova, I. Grigoriev, Y. Lou, D. Rohksar, S. Lucas, K. Huang, D.M. Goodstein, T. Hawkins, V. Plengvidhya, D. Welker, J. Hughes, Y. Goh, A. Benson, K. Baldwin, J.H. Lee, I. Diaz-Muniz, B. Dosti, V. Smeianov, W. Wechter, R. Barabote, G. Lorca, E. Altermann, R. Barrangou, B. Ganesan, Y. Xie, H. Rawsthorne, D. Tamir, C. Parker, F. Breidt, J. Broadbent, R. Hutkins, D. O'Sullivan, J. Steele, G. Unlu, M. Saier, T. Klaenhammer, P. Richardson, S. Kozyavkin, B. Weimer, and D. Mills, **Comparative genomics of the lactic acid bacteria**. Proc Natl Acad Sci U S A, 2006. **103**(42): p. 15611-6.
28. Sun, Z., X. Chen, J. Wang, W. Zhao, Y. Shao, L. Wu, Z. Zhou, T. Sun, L. Wang, H. Meng, H. Zhang, and W. Chen, **Complete Genome Sequence of *Streptococcus thermophilus* Strain ND03**. J Bacteriol, 2011. **193**(3): p. 793-4.
29. Ibrahim, M., P. Nicolas, P. Bessieres, A. Bolotin, V. Monnet, and R. Gardan, **A genome-wide survey of short coding sequences in *streptococci***. Microbiology, 2007a. **153**(Pt 11): p. 3631-44.
30. Zuber, U. and W. Schumann, **CIRCE, a novel heat shock element involved in regulation of heat shock operon *dnaK* of *Bacillus subtilis***. J Bacteriol, 1994. **176**(5): p. 1359-63.
31. Ibrahim, M., A. Guillot, F. Wessner, F. Algaron, C. Besset, P. Courtin, R. Gardan, and V. Monnet, **Control of the transcription of a short gene encoding a cyclic peptide in *Streptococcus thermophilus*: a new quorum-sensing system?** J Bacteriol, 2007b. **189**(24): p. 8844-54.
32. Moliere, N. and K. Turgay, **Chaperone-protease systems in regulation and protein quality control in *Bacillus subtilis***. Res Microbiol, 2009. **160**(9): p. 637-44.
33. Chastanet, A., M. Prudhomme, J.P. Claverys, and T. Msadek, **Regulation of *Streptococcus pneumoniae* *clp* genes and their role in competence development and stress survival**. J Bacteriol, 2001. **183**(24): p. 7295-307.
34. Terzaghi, B.E. and W.E. Sandine, **Improved medium for lactic streptococci and their bacteriophages**. Appl Microbiol, 1975. **29**(6): p. 807-13.
35. Letort, C. and V. Juillard, **Development of a minimal chemically-defined medium for the exponential growth of *Streptococcus thermophilus***. J Appl Microbiol, 2001. **91**(6): p. 1023-9.
36. Maguin, E., P. Duwat, T. Hege, D. Ehrlich, and A. Gruss, **New thermosensitive plasmid for gram-positive bacteria**. J Bacteriol, 1992. **174**(17): p. 5633-8.
37. Sambrook, J., E.F. Fritsch, and T. Maniatis, **Molecular cloning : a laboratory manual**. 1989.
38. Leenhouts, K., **Integration strategies and vectors**. Dev Biol Stand, 1995. **85**: p. 523-30.
39. Miroux, B. and J.E. Walker, **Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels**. J Mol Biol, 1996. **260**(3): p. 289-98.

40. Layec, S., J. Gerard, V. Legue, M.P. Chapot-Chartier, P. Courtin, F. Borges, B. Decaris, and N. Leblond-Bourget, **The CHAP domain of Cse functions as an endopeptidase that acts at mature septa to promote *Streptococcus thermophilus* cell separation.** Mol Microbiol, 2009. **71**(5): p. 1205-17.
41. Kieser, T., M. Bibb, M. Buttner, K. Chater, and D. Hopwood, **Practical *Streptomyces* Genetics.** John Innes Foundation ed. 1999.

Table 1. Bacterial strains and plasmids used in this study.

Strains and plasmids	Genotype/phenotype/ source	Origin or reference
<i>Streptococcus thermophilus</i>		
LMG18311	Wild-type; isolated from yogurt.	BCCM LMG
Δrgg_{0182}	Derivative of LMG18311 carrying a 699 bp deletion in the <i>rgg_{0182}</i> gene	This study
LMG18311 <i>blp::P_{pep0182}-luxAB</i>	LMG18311 containing a <i>P_{pep0182}-luxAB</i> fusion at the <i>blp</i> locus	This study
LMG18311 <i>blp::P_{shp0182}-luxAB</i>	LMG18311 containing a <i>P_{shp0182}-luxAB</i> fusion at the <i>blp</i> locus	This study
Δrgg_{0182} <i>blp::P_{pep0182}-luxAB</i>	Δrgg_{0182} containing a <i>P_{pep0182}-luxAB</i> fusion at the <i>blp</i> locus	This study
Δrgg_{0182} <i>blp::P_{shp0182}-luxAB</i>	Δrgg_{0182} containing a <i>P_{shp0182}-luxAB</i> fusion at the <i>blp</i> locus	This study
<i>Escherichia coli</i>		
EC101	<i>supE hsd-5 thi Δ(lac-proAB) F⁺(traD6 proAB⁺ lacI^q lacZΔM15) repA⁺</i> , derivative of TG1 strain	[37, 38]
EC101 pG9 Δrgg_{0182}	Derivative of EC101 containing pG9 Δrgg_{0182}	This study
C41(DE3)	Derived from BL21(DE3) (<i>F⁻ ompT hsdS_B(r_B⁻ m_B⁻) gal dcm met</i> (DE3); Novagen); at least one uncharacterized mutation	[39]
C41(DE3) pET15b:: <i>rgg_{0182}</i>	Derivative of C41(DE3) containing pET15b:: <i>rgg_{0182}</i>	This study
Plasmid		
pG ⁺ host9	Thermosensitive plasmid replication origin from pVE6002, used for gene replacement; Ery ^R	[36]
pG9 Δrgg_{0182}	Derivative of pG ⁺ host9 carrying a 699 bp (from nucleotide 82 to 783) deletion for <i>rgg_{0182}</i>	This study
pGICB004	Thermosensitive plasmid, <i>Em^r</i> ; used for the construct of transcriptional fusions. Allow the integration of the <i>luxAB</i> genes of <i>Photobacterium luminescens</i> at the <i>blp</i> locus of the <i>S. thermophilus</i> chromosome	[9]
pGICB004:: <i>P_{pep0182}</i>	Derivative of pGICB004 used to introduce a <i>P_{pep0182}-luxAB</i> transcriptional fusion at the <i>blp</i> locus	This study
pGICB004:: <i>P_{shp0182}</i>	Derivative of pGICB004 used to introduce a <i>P_{shp0182}-luxAB</i> transcriptional fusion at the <i>blp</i> locus	This study
pET15b	Expression vector for N-terminal His ₆ -tagged fusion; Amp ^R	Novagen
pET15b:: <i>rgg_{0182}</i>	Derivative of pET15b carrying a 864 bp insert coding a Rgg ₀₁₈₂ His ₆ -tagged protein	This study

BCCM Belgian Coordinated Collections of Microorganisms, LMG Laboratory of Microbiology and Genetics, University of Gent

Table 2. Primers used for PCR and qPCR.

Name	Primer sequence (5'–3')		
Primers for vectors construction			
Rgg ₀₁₈₂ I1	CTGGA ACTGCAGG AGCAGC		
Rgg ₀₁₈₂ I2	ATAATTTGAG AATTC TGTACCTT		
Rgg ₀₁₈₂ II1	TCTCTG GAATTC TTTAAATTG C		
Rgg ₀₁₈₂ II2	CTTGT CCTGCAG TCTCACTCCC		
Rgg ₀₁₈₂ SE1	CCCCCCCC ATATGGG AAAACAAAATGAACTG		
Rgg ₀₁₈₂ SE2	CCCC GGATCCT TAACTACAATCGACTTAAA		
Pep0182up	GA AGAATTC GTAAAAAGCATCAGATTTTAC		
Pep0182down	AACT ACTAGT GCATTTTAAAGTCGATTGTAGG		
SHP0182up	GA AGAATTC AAACAATTT CATATG ATTTCCTCC		
SHP0182down	AACT ACTAGTTC AGTCTCTCCTCTTTC		
Primers for DNA EMSA			
Pldh-5'	ACGCTTTC ACTTAATAATTC		
Pldh-3'	TGGTCTAAACATCTCCTTA		
Pshp-5'	GTAAATACATGTCAATAGGAC		
Pshp-3'	TTTGTTTTCC CATATG CAACC		
Ppep-5'	AGCATCAGATTTT ACTCCAGATG		
Ppep-3'	TTGTAGGTTAATCCCGTTTATGC		
Primers for gene expression analysis			
Gene	Locus	Forward	Reverse
<i>ldh</i>	stu1280	AAGCTATCCTTGACGATGAA	AATAGCAGGTTGACCGATAA
<i>rgg₀₁₈₂</i>	stu0182	GAAGTGGAGGAGTTGCCTAA	CCCAGCTCTCAATCCCAA
<i>hrcA</i>	stu0118	ACACCTCTTCAAGGAACTGAT	GTCACTTCATCATCGGAGATA
<i>dnaK</i>	stu0120	GACATTGACGAAGTCATCCT	GCACCCATAGCAACTACTTC
<i>dnaJ</i>	stu0121	CGTGAAGTGACATGTAAGACA	ACCAAGTGGTGTTTGTGTATC
<i>groES</i>	stu0203	GAAGGTACCCG ACTCTTACTG	AACGTAATCTTCTCCGTCCTTA
<i>groEL</i>	stu0204	ATTGCTTATAATGCCGGTTA	AGCGTTAAATCCTGTACCAA
<i>hsp33</i>	stu0180	TTTAGTAGGTCCTTTCATG	CGATTTACCAGAAATAAGC
<i>cspA</i>	stu0837	ATTGGTTTAA CGCTGACAAAG	TAACCTTTTGACCTTCATCGT
<i>cspB</i>	stu0838	TATGGCAAATGGAACAGTAAA	CAAGTGATTTGAATCCATCAG
<i>ctsR</i>	stu0076	AGATCAGCTCAGCGAACA	AGAATACGCGAACGAATG
<i>clpC</i>	stu0077	ACTGGCAGATTATA CCAAAGAC	CACCTACCAAGACAGGATTATT
<i>clpP</i>	stu0356	CTTGCTCAAGACTCGTAATAACT	AAGCCATATTCAAGTGTTTCTT
<i>clpX</i>	stu0581	TGGACTTATCCCTGAATTTATC	AAGACAAGAGGGTTTGATACTG
<i>clpE</i>	stu0602	CCGTACCAAGAACAATCCT	ACGGATAACTTGCTTGCTT
<i>clpL</i>	stu1614	CGTTTCGACGCAGTTATT	TTCGCTAACTGCCAAGTC

Restriction sites are indicated with bold letters: *Pst*I: **CTGCAG**; *Eco*RI: **GAATTC**; *Spe*I: **ACTAGT**; *Nde*I: **CATATG**; *Bam*HI: **GGATCC**

Figure 1. Schematic representation of the *rgg0182* and *rgg1358* loci (A) and of the corresponding proteins (B). Although the *rgg0182* and *rgg1358* loci present analogies (A), they encoded distinct proteins (B). Numbers in panel A indicate the position of nucleotides, with the +1 position being that of the first nucleotide of the *rgg0182* gene. The “deletion fragment” corresponds to the deleted portion of the *rgg0182* gene in the Δ *rgg0182* mutant. The broken arrows indicate the promoters. *Pshp0182* and *Ppep0182* materialized the position of the 126 bp and 165 bp PCR fragment respectively used in EMSA. In panel B, amino acids sequence identities are indicated in percent. HTH indicated the Helix-Turn-Helix-XRE motif.

Figure 2. Relative *rgg0182* gene transcript level from *S. thermophilus* LMG18311 cells, grown at 42°C (A) and at 30°C (B).

Total RNAs from the wild type strain were extracted from exponential (E, white bars), transition (T, light gray bars) and stationary (S, dark gray bars) phase cells. Data are presented as the mean \pm standard deviation from three independent experiments performed in duplicate. Student's *t* test: *, $p < 0.001$.

Figure 3. Luciferase activity and growth of the LMG18311 and the Δ *rgg0182* strains containing the *Pshp0182-luxAB* and *Ppep0182-luxAB* transcriptional fusions, in CDM medium.

The expression of the fusions was followed in strains cultivated in CDM medium, at 30°C (A) or at 42°C (B). Data are presented as the mean \pm standard deviation of three independent experiments. AU: luminescence arbitrary units normalized against the OD_{600nm} of the cultures.

Figure 4. Analysis of the Rgg₀₁₈₂ binding to DNA.

Electrophoretic mobility shift assay (EMSA) of the promoter regions of the two target genes (*shp0182* and *pep0182*) of Rgg₀₁₈₂ in the absence or in the presence of the purified His₆-Rgg₀₁₈₂ protein. DNA probes labelled with biotin (0.1 pmol each) were incubated with 2 pmol of Rgg₀₁₈₂. The *P_{ldh}* probe is an *ldh* promoter fragment used as a negative control.

Figure 5. Relative genes transcript level of *S. thermophilus* stationary phase cells grown in CDM medium at 30°C.

Total RNAs were extracted from stationary phase cells of *S. thermophilus* LMG18311 (dark gray bars) and its isogenic Δ *rgg0182* mutant (light gray bars) grown in CDM at 30°C. Data are presented as the mean \pm standard deviation of the gene transcript levels measured from 3 independent experiments done in duplicate. Student's *t* test: *, $p < 0.001$.

Figure 6. Survival of the *S. thermophilus* strain LMG18311 and the Δrgg_{0182} mutant after heat shock (0, 15, 30, 45 and 60 min at 52°C).

S. thermophilus was cultivated in CDM medium at 30°C and then exposed to heat stress. The percentage of survival was calculated as $N / N_0 \times 100$ where N_0 is the CFU number of the control condition and N the CFU number in heat stress condition. Dark gray bars correspond to wild type strain and light gray bars correspond to Δrgg_{0182} strain. Data are presented as the mean \pm standard deviation of 3 independent experiments done in triplicate. Student's *t* test: *, $p < 0.001$.

Figure 7. Relative genes transcript level of *S. thermophilus* cells exposed to a heat stress. Total RNAs were extracted from stationary phase cells of *S. thermophilus* LMG18311 (dark gray bars) and its isogenic Δrgg_{0182} mutant (light gray bars) grown in CDM at 30°C until stationary phase and then exposed 30 min at 52°C (heat stress condition). Data are presented as the mean \pm standard deviation of the gene transcript levels measured from 3 independent experiments done in duplicate. Student's *t* test: *, $p < 0.001$.

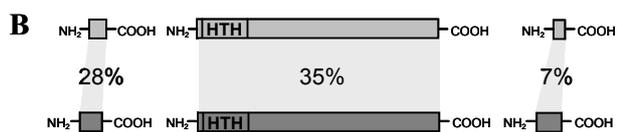
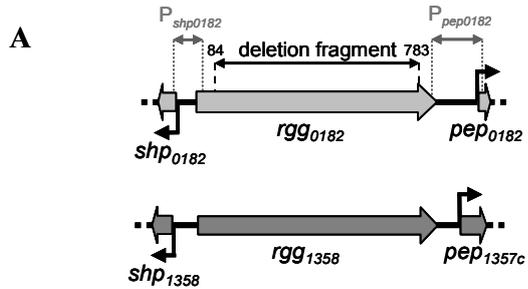


Figure 1.

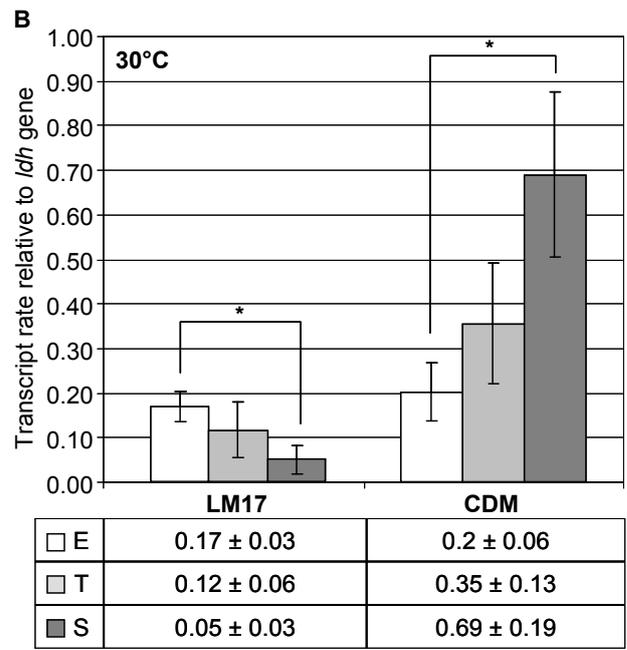
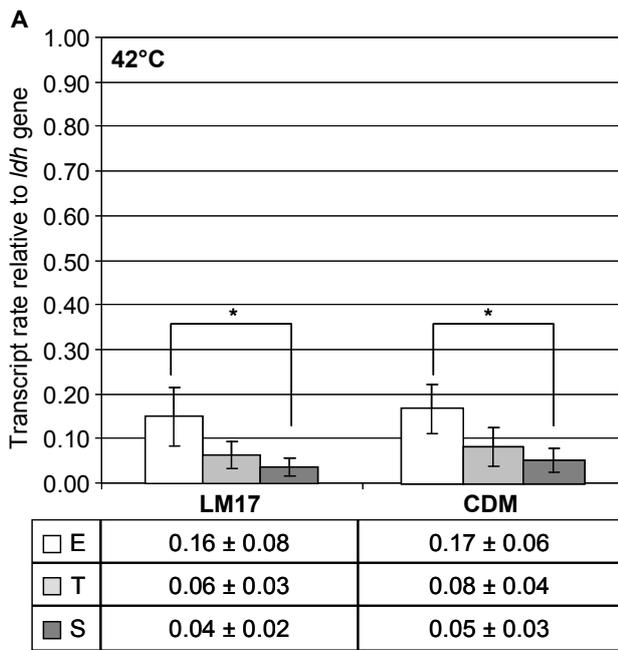


Figure 2.

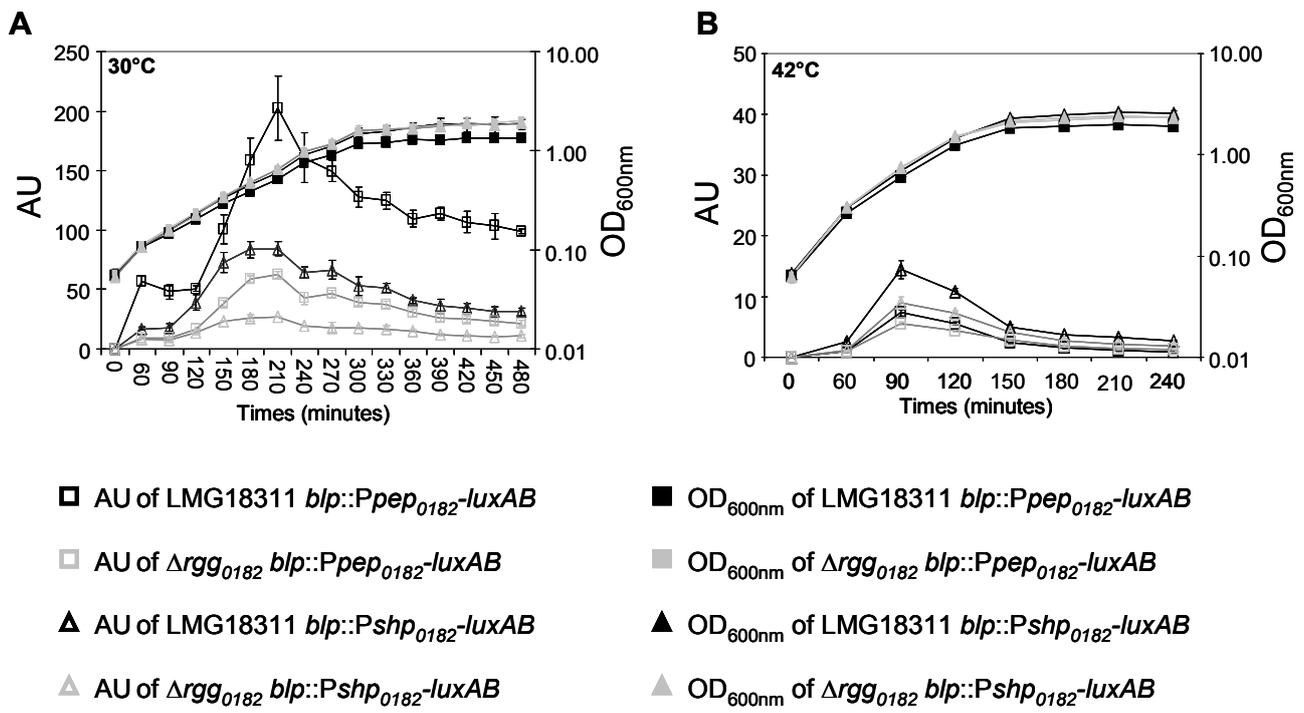


Figure 3.

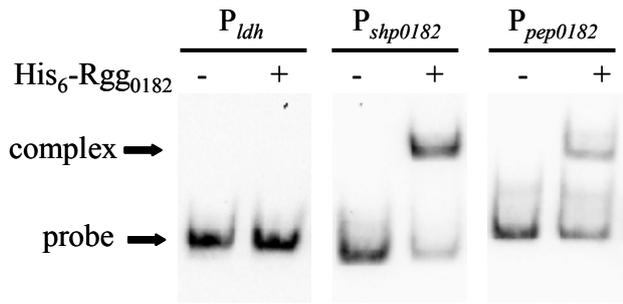


Figure 4.

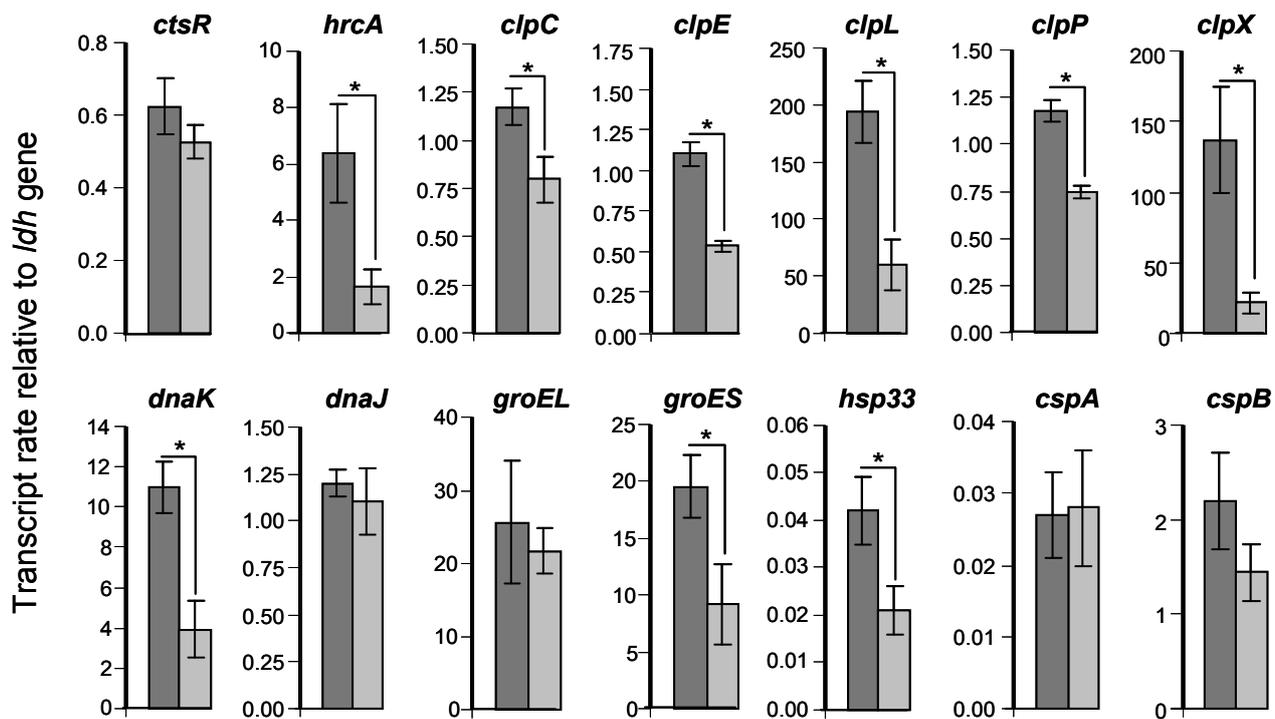


Figure 5.

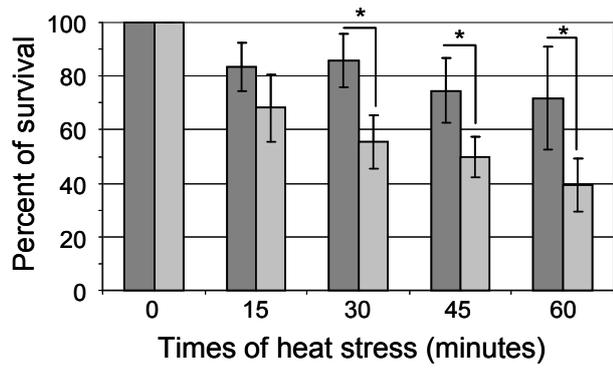


Figure 6.

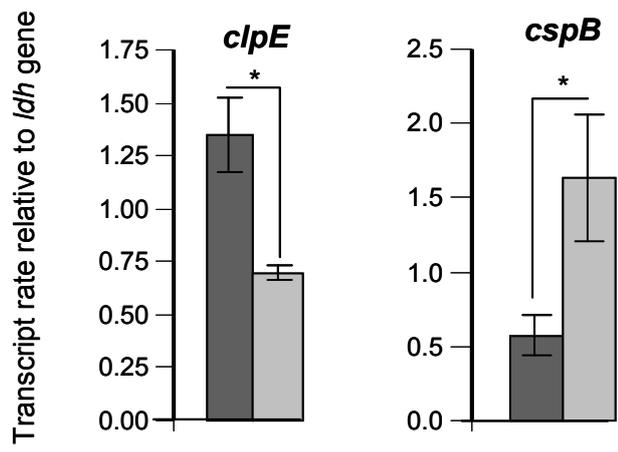


Figure 7.

Actes de congrès

Publication affichée n°1 : Congrès nationale.

Influence des gènes *rgg 0182*, *rgg 1947*, *stk 0183* et *stk 1948* dans la croissance de *Streptococcus thermophilus* LMG18311.

Romain Henry, Bernard Decaris, Nathalie Leblond-Bourget.

15ème colloque du Club des Bactéries Lactiques à Rennes. Novembre 2007.

Publication affichée n°2 : Congrès nationale.

Implication du régulateur transcriptionnel Rgg_{STK1} dans l'adaptation de *Streptococcus thermophilus* LMG18311 à son environnement.

Romain Henry, Julien Hiblot, Stéphane Bertin, Bernard Decaris, Nathalie Leblond-Bourget.

16ème colloque du Club des Bactéries Lactiques à Toulouse. Mai 2009.

Publication affichée n°3 : Congrès nationale.

Caractérisation des sites de fixation du régulateur transcriptionnel Rgg_{stk1} de la souche LMG18311 de *Streptococcus thermophilus* au niveau du promoteur de ses gènes cibles caractérisés.

Emmanuelle Bruneau, Romain Henry, Stéphane Bertin, Betty Fleuchot, Rozenn Gardan, Nathalie Leblond-Bourget.

17ème colloque du Club des Bactéries Lactiques à Nancy. Novembre 2010.

Communication orale : Congrès nationale.

Implication du régulateur transcriptionnel Rgg_{stk1} dans la réponse au stress de *Streptococcus thermophilus* LMG18311

Romain Henry, Emmanuelle Bruneau, Stéphane Bertin, Bernard Decaris, Nathalie Leblond-Bourget.

17ème colloque du Club des Bactéries Lactiques à Nancy. Novembre 2010.

Influence des gènes *rgg 0182*, *rgg 1947*, *stk 0183* et *stk 1948* dans la croissance de *Streptococcus thermophilus* LMG18311.

POSTER

Romain Henry, Bernard Decaris et Nathalie Leblond-Bourget*

* Laboratoire de génétique et microbiologie UMR-INRA 1128 – IFR 110, Faculté des Sciences et Techniques - Université Henri Poincaré Nancy 1, BP 239 54506, VANDOEUVRE-LES-NANCY, France, bourget@nancy.inra.fr

Résumé :

Au sein du genre *Streptococcus*, *Streptococcus thermophilus* est l'espèce dont le génome séquencé présente le plus grand nombre de gènes *rgg*-like. Ces gènes sont décrits dans la littérature comme codant des régulateurs transcriptionnels, et modulant notamment l'expression de gènes adjacents [1]. Chez les streptocoques pathogènes, les protéines Rgg sont impliquées dans la régulation de gènes de virulence [2] et de biosynthèse d'acides aminés et dans l'adaptation aux changements environnementaux [3].

Le génome de *S. thermophilus* LMG18311 présente sept ORF *rgg*-like. Parmi elles, deux sont adjacentes à une ORF codant potentiellement une kinase à sérine/thréonine (*stk*). Afin de déterminer si les protéines Rgg peuvent réguler la transcription de ces kinases (et/ou inversement) et si le couple Rgg-kinase pourrait constituer un système à deux composants, la stratégie consiste en l'obtention de mutants, délétés des gènes *rgg 0182* et *rgg 1947*, ainsi que de leur cible potentielle, *stk 0183* et *stk 1948*. Deux mutants ont été actuellement obtenus, l'un présente la délétion du gène *rgg 0182* et l'autre la délétion de *stk 0183*.

Une analyse de la croissance de ces mutants en milieu chimiquement défini a ensuite été réalisée. Le mutant Δ *stk 0183* présente une biomasse finale plus importante, augmentée de DO=0,4, et une phase de latence plus courte, par rapport à la souche sauvage. Ainsi, l'ORF *stk 0183* est impliquée dans les mécanismes de croissance de *S. thermophilus* LMG18311. Ceci suggère que les kinases pourraient être impliquées dans la régulation globale des métabolismes cellulaire de *S. thermophilus* LMG18311. Au contraire, le mutant Δ *rgg 0182* ne semble pas être affecté dans sa croissance. L'absence de différence observée chez ce mutant pourrait s'expliquer par la présence d'autres gènes *rgg*-like au sein de *S. thermophilus* LMG18311.

En perspectives, ce travail consistera en l'étude du rôle des ORF *rgg*-like et des kinases dans l'adaptation de *S. thermophilus* LMG18311 à son environnement.

Références : [1] Sulavik, M.C., Tardif, G., et Clewell, D.B. Identification of a gene, *rgg*, which regulates expression of glucosyltransferase and influences the Spp phenotype of *Streptococcus gordonii* Challis. *J Bacteriol*, 1992, **174**: 3577-3586. [2] Chaussee, M.S., Ajdic, D., et Ferretti, J.J. The *rgg* gene of *Streptococcus pyogenes* NZ131 positively influences extracellular SPE B production. *Infect Immun*, 1999, **67**: 1715-1722. [3] Sanders, J.W., Leenhouts, K., Burghoorn, J., Brands, J.R., Venema, G., et Kok, J. A chloride-inducible acid resistance mechanism in *Lactococcus lactis* and its regulation. *Mol Microbiol*, 1998, **27**: 299-310.

Mots clés : *Streptococcus*, protéines Rgg, kinase, croissance cellulaire, adaptation

Titre : Implication du régulateur transcriptionnel Rgg_{STK1} dans l'adaptation de *Streptococcus thermophilus* LMG18311 à son environnement.

Auteur(s) : Henry Romain, Hiblot Julien, Bertin Stéphane, Decaris Bernard et Leblond-Bourget Nathalie

Adresse(s) : Laboratoire de génétique et microbiologie UMR-INRA 1128 / IFR 110, Faculté des Sciences et Techniques - Université Henri Poincaré Nancy 1, BP 239 54506, VANDOEUVRE-LES-NANCY, France

Résumé :

Au sein du genre *Streptococcus*, *Streptococcus thermophilus* est l'espèce dont le génome séquencé présente le plus grand nombre de gènes *rgg*. Ces gènes sont décrits dans la littérature comme codant des régulateurs transcriptionnels, et modulant notamment l'expression de gènes adjacents. Chez les streptocoques pathogènes, les protéines Rgg sont impliquées dans la régulation de gènes de virulence et de biosynthèse d'acides aminés et dans l'adaptation aux changements environnementaux.

Le génome de *S. thermophilus* LMG18311 présente sept ORF *rgg*. Afin de déterminer le rôle de ces gènes au sein de la cellule, la stratégie consiste en l'obtention de mutants, délétés de gènes *rgg*. Cinq simples mutants ont été actuellement obtenus. Parmi eux, la souche \square *rgg*_{STK1} présente un phénotype courte-chaîne. En effet, *S. thermophilus* se développe sous forme de chaînettes d'une vingtaine de cellules en moyenne. Le mutant présentant la délétion du gène *rgg*_{STK1} présente des chaînes de 2 cellules en moyenne suggérant que Rgg_{STK1} serait impliqué dans l'organisation des cellules en chaînes.

Une analyse de ce mutant en réponse à un stress froid (-4°C) a ensuite été réalisée. Le mutant \square *rgg*_{STK1} présente une chute de DO lors de la phase stationnaire de croissance, contrairement à la souche sauvage alors que dans les conditions optimales de croissance, aucune différence n'est observée. Ces données suggèrent que Rgg_{STK1} pourrait être impliqué dans un système de maintien de la phase stationnaire de croissance. Des analyses de qPCR sont actuellement en cours afin de déterminer les cibles du régulateurs Rgg_{STK1} et comprendre les mécanismes mis en jeu lors de la réponse au stress froid.

En perspectives, ce travail consistera en l'étude du rôle des ORF *rgg*-like et des kinases dans l'adaptation de *S. thermophilus* LMG18311 à son environnement.

Caractérisation des sites de fixation du régulateur transcriptionnel Rgg_{stk1} de la souche LMG18311 de *Streptococcus thermophilus* au niveau du promoteur de ses gènes cibles caractérisés

BRUNEAU Emmanuelle¹, HENRY Romain¹, BERTIN Stéphane¹, GARDAN Rozenn² et LEBLOND-BOURGET Nathalie¹

¹ Laboratoire de Génétique et Microbiologie, UMR UHP-INRA 1128/ IFR 110 Université Henri Poincaré Nancy 1, Faculté des Sciences et Techniques, 54506 Vandœuvre-lès-Nancy, France

² Unité de Biochimie Bactérienne, UR477, INRA, 78350 Jouy en Josas, France

emmanuelle.bruneau@scbiol.uhp-nancy.fr

Streptococcus thermophilus est une bactérie utilisée dans l'industrie laitière pour la fabrication de yaourts et de fromages. Lors de son utilisation industrielle, cette bactérie subit de nombreuses variations des conditions de cultures (choc thermique, modification de l'oxygénation et de l'acidité). En conséquences des stress subits, la population bactérienne doit modifier son métabolisme par l'induction de l'expression de nombreux gènes. *S. thermophilus* LMG18311 possède au sein de son génome 7 copies de gènes *rgg* codant des régulateurs transcriptionnels. Chez les streptocoques pathogènes, ces protéines ont été décrites dans la littérature comme participant à la régulation de gènes impliqués dans la virulence, dans le métabolisme et dans l'adaptation aux changements environnementaux (Chaussee *et al.*, 2002 ; Chaussee *et al.*, 2004).

Des analyses réalisées au Laboratoire de Génétique et Microbiologie ont permis de mettre en évidence l'implication du gène *rgg_{stk1}* de la souche LMG18311 de *S. thermophilus* dans la réponse aux stress thermique et acide. Ainsi, Rgg_{stk1} pourrait participer à l'adaptation de *S. thermophilus* à son environnement.

Les gènes de la famille *rgg* sont décrits dans la littérature comme régulant, dans la plupart des cas, l'expression d'un ou plusieurs gènes qui leur sont adjacents par la fixation de la protéine Rgg sur l'ADN de la région promotrice des gènes régulés (Vickerman *et al.*, 2003 ; Nelly *et al.*, 2003).

Ainsi Rgg_{stk1} pourrait réguler les gènes qui sont adjacents à son locus. Afin de vérifier cette hypothèse, des fusions avec la luciférase des promoteurs des gènes *shp_{stk1}* et *pep_{stk1}*, encadrant le gène *rgg_{stk1}*, ont été réalisées. L'étude transcriptionnelle de ces dernières a montré que Rgg_{stk1} régule l'expression des gènes *shp_{stk1}* et *pep_{stk1}*. Dans le but d'étudier les sites de fixation à l'ADN de la protéine Rgg_{stk1}, celle-ci a été exprimée chez *Escherichia coli* à l'aide du plasmide pET15b. Rgg_{stk1} a ensuite été purifiée pour réaliser des expériences de retard sur gel. De plus, cette technique permettra de déterminer si la régulation des gènes *shp_{stk1}* et *pep_{stk1}*, par la protéine Rgg_{stk1}, est directe ou indirecte.

Les expériences de retard sur gel permettront également d'étudier d'autres cibles potentielles qui sont actuellement en cours de caractérisation.

Références

- Chaussee M. S., Sylva G. L., Sturdevant D. E., Smoot L. M., Graham M. R., Waston R. O. and Musser J. M., **2002**, Infect. Immun., (70): 762-770.
Chaussee M. A., Callegari E. A. and Chaussee M. S., **2004**, J. Bacteriol., (186): 7091-7099.
Neely M. N., Lyon W. R., Runft D. L. and Caparon M., **2003**, J. Bacteriol. (185): 5166-5174.
Vickerman M. M., Wang M. and Baker L. J., **2003**, Microbiology, (149): 399-406.

Implication du régulateur transcriptionnel Rgg_{stk1} dans la réponse au stress de *Streptococcus thermophilus* LMG18311

HENRY Romain, BRUNEAU Emmanuelle, BERTIN Stéphane, DECARIS Bernard, LEBLOND-BOURGET Nathalie

Laboratoire de Génétique et Microbiologie UMR-INRA 1128 / IFR 110, Faculté des Sciences et Techniques – Campus Victor Grignard, BP 239, 54506 Vandoeuvre-les-Nancy, France
Romain.Henry@scbiol.uhp-nancy.fr

Les gènes *rgg* sont décrits dans la littérature comme codant des régulateurs transcriptionnels. Chez les streptocoques pathogènes, les protéines Rgg sont impliquées dans la régulation de gènes de virulence et de biosynthèse d'acides aminés et dans l'adaptation aux changements environnementaux (Sulavik *et al.*, 1992 ; Chausse *et al.*, 2002).

Au sein du genre *Streptococcus*, *S. thermophilus* est l'espèce dont le génome séquencé présente le plus grand nombre d'ORF *rgg* : 7 ORF *rgg* au sein du génome de la souche LMG18311 contre 2 au maximum pour les streptocoques pathogènes. Afin de déterminer le rôle de ces gènes au sein de la cellule, la stratégie a consisté en l'obtention de mutants, présentant la délétion de gènes *rgg*. Cinq simples mutants ont été obtenus dont le mutant présentant la délétion du gène *rgg_{stk1}*.

La transcription du gène *rgg_{stk1}* a été analysée par PCR quantitative en temps réel dans différentes conditions de croissance : milieu chimiquement défini MCD et en milieu complet M17, différentes températures (30°C et 42°C). Les résultats obtenus montrent que le gène *rgg_{stk1}* est transcrit de manière phase dépendante dans toutes les conditions testées. De plus le niveau de transcription dépend du milieu de culture.

Le gène *rgg_{stk1}* présente le plus fort taux de transcription en milieu MCD à 30°C et en phase stationnaire. Dans ces conditions, des analyses phénotypiques en réponse à un stress thermique et un stress acide été réalisées et ont permis de montrer l'implication de la protéine Rgg_{stk1}. En effet, lors d'un stress thermique (à 52°C) et lors d'un stress acide (à pH 3), réalisé pendant 30 min, le mutant Δ *rgg_{stk1}* est, respectivement, 100 fois plus sensible et 100 fois plus résistant que la souche sauvage. Ces données suggèrent que la protéine Rgg_{stk1} pourrait réguler des gènes impliqués dans la réponse au stress thermique et acide.

Afin de vérifier cette hypothèse, des analyses de PCR quantitative en temps réel sont actuellement en cours afin de comparer la transcription de gènes cibles potentiels entre le mutant Δ *rgg_{stk1}* et la souche sauvage. Les premiers résultats obtenus semblent suggérer que les gènes *hsp33*, *cspA* et *cspB*, tous trois impliqués dans la réponse au stress thermique, pourraient être des cibles de la protéine Rgg_{stk1}.

References

Chaussee MS, Sylva GL, Sturdevant DE, Smoot LM, Graham MR, Watson RO, Musser JM. **2002**, Infect Immun. Feb;70(2):762-70.

Sulavik MC, Tardif G, Clewell DB. **1992**, J Bacteriol. Jun;174(11):3577-86.

Abstract

Streptococcus thermophilus, bacterium used in dairy industry, is subjected to many environmental changes (oxygen concentration modification, temperature shift, pH variation ...) during manufacturing of fermented products. To adapt itself to these modifications, bacterium has defense systems which are controlled by regulation networks involving, in particular, transcriptional regulators. Among them, they are transcriptional regulators of the Rgg family. *rgg* genes code transcriptional regulators which are very polymorph. Several studies showed their involvement in stress response.

The originality of *S. thermophilus* is the presence of many different *rgg* copies on genomes of this bacterial species (between 5 and 7 copy according to strain). Objective of this study is the characterization of the Rgg family and study of the *S. thermophilus* *rgg* gene function, and, in particular, involvement of these genes in *S. thermophilus* adaptation to environmental changes.

In silico analyses has allowed to show that *rgg* genes are specific to the Streptococcaceae family of and the *Lactobacillus* and the *Listeria* genus. Proteins belonging to this family are polymorphic and don't form a monophyletic group. Analyses nevertheless showed that these proteins were characterized by two domains: a "HTH domain" in N-terminal position, very probably involved in the DNA interaction, and a "median domain" in central position which the role remains to determine.

Analyses realized during this work show that *S. thermophilus* LMG18311 *rgg*₀₁₈₂ gene codes a transcriptional regulator which participates, in particular, in the adaptive response of *S. thermophilus* exposed to an increase of its growth temperature. Furthermore, the deletion of the *rgg*₀₁₈₂ gene confers on cells the capacity of a hydrophobic surface adhesion and leads to a reduction of cells chains size. All these phenotypes suggest that Rgg₀₁₈₂ protein is a global regulator of *S. thermophilus* LMG18311 genic expression. Target genes of this regulator were looked for among genes of thermal stress response and indicate that Rgg₀₁₈₂ protein is involved in the chaperons and proteases homeostasis.

Keywords: *Streptococcus thermophilus*, Rgg family, stress, global transcriptional regulator, polymorphism, adaptive response

Résumé

Streptococcus thermophilus, bactérie utilisée en industrie agro-alimentaire, est soumise à de nombreuses modifications environnementales (modification de la concentration en oxygène, de la température, du pH,...) lors de la fabrication de produits fermentés. Pour s'adapter à ces modifications, la bactérie dispose de systèmes de défense qui sont contrôlés par des réseaux de régulation impliquant notamment des régulateurs transcriptionnels. Parmi eux, se trouvent les régulateurs de la famille *rgg*. Les gènes *rgg* codent des régulateurs transcriptionnels très polymorphes. Plusieurs études ont montré leur implication dans la réponse aux stress.

L'originalité de *S. thermophilus* est que de nombreuses copies *rgg* différentes sont présentes sur les génomes de cette espèce bactérienne (entre 5 et 7 copie selon les souches). Cette étude a pour objectif la caractérisation de la famille Rgg et l'étude du rôle des gènes *rgg* de *S. thermophilus*, et, notamment, l'implication de ces gènes dans l'adaptation de *S. thermophilus* à son environnement.

Les analyses réalisées *in silico* ont permis de montrer que les gènes *rgg* sont spécifiques de la famille des *Streptococcaceae* et des genres *Lactobacillus* et *Listeria*. Les protéines appartenant à cette famille sont polymorphes et ne forment pas un groupe monophylétique. Les analyses ont néanmoins montré que ces protéines se caractérisaient par deux domaines : un « domaine HTH » en position N-terminale, très probablement impliqué dans la liaison à l'ADN et un « domaine médian » en position centrale, dont le rôle reste à déterminer.

Les analyses réalisées au cours de ce travail montrent que le gène *rgg₀₁₈₂* de *S. thermophilus* LMG18311 code un régulateur transcriptionnel qui participe, notamment, à la réponse adaptative de *S. thermophilus* exposée à une augmentation de sa température de croissance. De plus, la délétion du gène *rgg₀₁₈₂* confère aux cellules une capacité d'adhésion à une surface hydrophobe et induit une réduction de la taille des chaînes. L'ensemble de ces phénotypes suggère que la protéine Rgg₀₁₈₂ est un régulateur global de l'expression génique de *S. thermophilus* LMG18311. Des gènes cibles de ce régulateur ont été recherchés parmi les gènes de réponses au choc thermique et indique que la protéine Rgg₀₁₈₂ est impliquée dans l'homéostasie des chaperons et des protéases.

Mots clés : *Streptococcus thermophilus*, famille Rgg, stress, régulateur transcriptionnel global, polymorphisme, réponse adaptative