



**HAL**  
open science

# Régulation des opérons Maltose/Maltodextrines et Gentiobiose induits en contexte d'infection chez *Enterococcus faecalis*

Maxime Grand

► **To cite this version:**

Maxime Grand. Régulation des opérons Maltose/Maltodextrines et Gentiobiose induits en contexte d'infection chez *Enterococcus faecalis*. Microbiologie et Parasitologie. Normandie Université, 2019. Français. NNT : 2019NORMC226 . tel-03428091

**HAL Id: tel-03428091**

**<https://theses.hal.science/tel-03428091>**

Submitted on 15 Nov 2021

**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.



Normandie Université

## THÈSE

**Pour obtenir le diplôme de doctorat**

**Spécialité ASPECTS MOLECULAIRES ET CELLULAIRES DE LA BIOLOGIE**

**Préparée au sein de l'Université de Caen Normandie**

### **Régulation des opérons Maltose/Maltodextrines et Gentiobiose induits en contexte d'infection chez Enterococcus faecalis**

**Présentée et soutenue par  
Maxime GRAND**

**Thèse soutenue publiquement le 14/11/2019  
devant le jury composé de**

Mme ANNE GALINIER	Directeur de recherche au CNRS, Aix-Marseille Université	Rapporteur du jury
M. PHILIPPE GILOT	Chargé de recherche HDR, UNIVERSITE TOURS UFR MEDECINE	Rapporteur du jury
M. MATTHIEU JULES	Professeur des universités, INRA de JOUY en JOSAS	Membre du jury
Mme PASCALE SERROR	Directeur de recherche à l'INRA, INRA de JOUY en JOSAS	Président du jury
M. NICOLAS SAUVAGEOT	Maître de conférences HDR, Université Caen Normandie	Directeur de thèse

**Thèse dirigée par NICOLAS SAUVAGEOT, UR risques microbiens (Caen)**



UNIVERSITÉ  
CAEN  
NORMANDIE



Normandie de Biologie Intégrative,  
Santé, Environnement





Normandie Université

## THESE

**Pour obtenir le diplôme de doctorat**

**Spécialité** Aspects moléculaires et cellulaires de la biologie

**Préparée au sein de l'Université de Caen, Normandie**

# Régulation des opérons maltose/maltodextrines et gentiobiose induits en contexte d'infection chez *Enterococcus faecalis*

**Présentée et soutenue par  
Maxime GRAND**

**Thèse soutenue publiquement le jeudi 14 novembre 2019  
devant le jury composé de**

Mme Anne GALINIER	Directeur de recherches, CNRS, Université d'aix-Marseille	Rapporteur
M. Philippe GILOT	Chargé de recherches HDR, INRA, Université de Tours	Rapporteur
Mme Pascale SERROR	Directeur de recherches, INRA, AgroParisTech, Université Paris-Saclay	Examineur
M. Matthieu JULES	Professeur, AgroParisTech, Université Paris-Saclay	Examineur
M. Nicolas SAUVAGEOT	Maître de conférences HDR, Université deCaen, Normandie	Directeur de thèse
M. Josef DEUTSCHER	Directeur de recherches, INRA, AgroParisTech, Université Paris-Saclay	Invité

**Thèse dirigée par Nicolas SAUVAGEOT**

**Laboratoire : Unité de Recherches Risques Microbiens (U2RM)**



UNIVERSITÉ  
CAEN  
NORMANDIE



## Remerciements

Avant de commencer ce manuscrit, je souhaiterais adresser mes plus chaleureux remerciements à plusieurs personnes qui ont particulièrement compté pour moi durant ce doctorat.

Je souhaiterais dans un premier temps exprimer ma reconnaissance au directeur de l'Unité de Recherche Risques Microbiens, le professeur Axel Hartke, pour m'avoir permis d'effectuer ce doctorat et pour m'avoir soutenu au cours de cette expérience autant sur le plan scientifique que sur le plan humain.

Je tenais ensuite à remercier le Docteur Anne Galinier et le Docteur Philippe Gilot pour avoir accepté d'évaluer ce travail en tant que rapporteurs. Je souhaiterais aussi adresser mes remerciements au Docteur Pascale Serror ainsi qu'au Professeur Matthieu Jules pour avoir examiné cette thèse.

J'adresse également tous mes remerciements au Docteur Josef Deutscher, au Professeur Christian Magni, au Docteur Victor Blancato, au Docteur Martín Espariz et au Professeur Abdellah Benachour pour leur accompagnement scientifique et leurs avis critiques qui ont grandement contribué à améliorer la qualité de notre travail scientifique.

Je tenais ensuite à exprimer toute ma gratitude à mon directeur de thèse, le Docteur Nicolas Sauvageot pour m'avoir toujours soutenu et guidé durant ces trois années passées ensemble et pour avoir toujours été disponible pour répondre à mes questions. Cette coopération s'est bien entendu déroulée au laboratoire mais également en dehors du cadre de la thèse et c'est pourquoi je te remercie grandement, Nicolas, ainsi que tous les membres de l'équipe "Stress et Virulence" (je ne peux malheureusement pas citer ici le nom de tout le monde mais le cœur y est, vraiment), pour avoir contribué à rendre ce travail bien plus qu'un simple travail de recherche mais une véritable expérience humaine riche en souvenirs.

Enfin, je souhaiterais ajouter quelques lignes pour remercier ma famille et notamment mes parents qui m'ont toujours supporté dans mes études de bien des manières et qui n'ont eu de cesse de m'encourager à persévérer. Je tenais également à remercier mes amis qui m'ont permis de décompresser en dehors du travail et qui ont su montrer un support sans faille. Je remercie enfin ma chérie pour son soutien quotidien tout au long de ce doctorat et dans la vie en général.

Malgré l'évident manque de profondeur que ce mot m'inspire en comparaison de ma réelle gratitude envers toutes les personnes citées précédemment et envers toutes celles qui sauront se reconnaître à travers ces lignes, je ne trouve qu'un seul mot à vous adresser :

# Merci

## Table des matières

Remerciements .....	2
Table des matières .....	3
Liste des abréviations .....	7
Liste des figures .....	9
Liste des tableaux .....	11
Contexte .....	12
Introduction bibliographique.....	17
I.    Le genre <i>Enterococcus</i> .....	18
A.    Historique .....	18
B.    Caractéristiques des entérocoques .....	18
C.    Les souches de référence : V583 et OG1RF.....	24
II.   L'import de substrats .....	25
A.    Les facilitateurs.....	26
B.    Les transporteurs actifs primaires : les transporteurs ABC .....	27
C.    Les transporteurs actifs secondaires .....	32
D.    Les transports par translocation de groupe : les Systèmes PhosphoTransférases PEP dépendent (PTS) .....	35
III.  Les régulateurs transcriptionnels .....	46
A.    Les régulateurs transcriptionnels de la famille LacI/GalR .....	47
B.    Les régulateurs transcriptionnels de la famille RpiR.....	49
IV.  Le métabolisme du maltose et des maltodextrines chez différents organismes modèles .....	51
A. <i>Escherichia coli</i> .....	53
B.    Streptococci .....	56
C. <i>Enterococcus faecalis</i> .....	60
V.   Le métabolisme du gentiobiose et autres $\beta$ -glycosides chez différents organismes modèles.....	62

A.	<i>Escherichia coli</i> .....	62
B.	<i>Streptococcus pneumoniae</i> .....	64
C.	<i>Enterococcus faecalis</i> .....	66
Matériel et méthode.....		67
I.	Matériel biologique .....	68
A.	Conditions de culture et souches bactériennes .....	68
B.	Plasmides .....	68
II.	Cellules compétentes et transformations par électroporation .....	71
A.	<i>Escherichia coli</i> .....	71
B.	<i>Enterococcus faecalis</i> .....	71
III.	Biologie moléculaire .....	72
A.	Réactions de polymérisation en chaîne (PCR) .....	72
B.	Electrophorèse en gel d'agarose.....	72
C.	Restriction et ligation de fragments d'ADN .....	72
D.	Purification des fragments d'ADN.....	73
E.	Extraction de plasmides d' <i>E.coli</i> .....	73
F.	Extraction d'ADN génomique .....	73
G.	Extraction d'ARN totaux .....	73
H.	Réactions de transcription inverse.....	74
I.	RACE-PCR.....	74
J.	Fusion transcriptionnelle .....	74
K.	Empreintes à la DNase I.....	75
L.	Séquençage du transcriptome entier (RNA-seq) .....	75
IV.	Constructions de mutants par génie génétique et complémentations .....	76
A.	Mutants par allèles mutés .....	76
B.	Mutant de délétion.....	77
C.	Complémentations des mutants.....	77

V.	Méthodes biochimiques .....	78
A.	Surexpression et purification de protéines présentant une étiquette poly-histidines .....	78
B.	Phosphorylation de HPr <i>in vitro</i> .....	79
C.	SDS-PAGE.....	80
D.	Native-PAGE.....	80
E.	Retard sur gel.....	81
F.	Microscale Thermophoresis (MST) .....	81
G.	Chromatographie sur couche mince (CCM).....	82
H.	Dosage des sucres dans le surnageant par HPLC .....	82
I.	Thermal Shift Assays (TSA) .....	82
	Résultats .....	84
I.	Objectifs .....	85
II.	Publication scientifique n°1 .....	85
A.	Résumé de la publication scientifique n°1 .....	85
B.	Publication n°1 : The dual role of MalR in induction and a new P-Ser-HPr requiring catabolite repression mechanism of the <i>Enterococcus faecalis mal</i> operons.....	85
C.	Résultats complémentaires .....	138
III.	Publication scientifique n°2 .....	142
A.	Résumé de la publication n°2 .....	142
B.	Publication n°2 : Regulation of maltodextrin genes in <i>Enterococcus faecalis</i> by the combined action of the main regulator of maltose genes MalR and the pleiotropic regulator CcpA .....	142
C.	Résultats complémentaires .....	179
IV.	Publication scientifique n°3 .....	181
A.	Résumé de la publication n°3 .....	181
B.	Publication n°3 : Characterization of the gen locus involved in	

β-1,6-oligosaccharide utilization by <i>Enterococcus faecalis</i> .....	182
C. Résultats complémentaires .....	225
Discussion et perspectives.....	227
I. Le métabolisme du maltose et des maltodextrines chez <i>E. faecalis</i> .....	229
II. Le métabolisme du gentiobiose chez <i>E. faecalis</i> .....	231
Références Bibliographiques.....	237
Annexes.....	261
I. Figures supplémentaires.....	262
II. Compositions des milieux de cultures utilisés dans cette étude.....	263
III. Amorces utilisées .....	265

## Liste des abréviations

<b>μL/μg/μM</b>	: micro – litre, -gramme, -molaire
<b>ABC</b>	: ATP Binding Cassette
<b>ADN</b>	: Acide désoxyribonucléique
<b>ADNc</b>	: ADN complémentaire
<b>AG</b>	: Substance d'Agrégation ou Aggregation Substance
<b>Amp</b>	: Ampicilline
<b>AMPc</b>	: Adénosine triphosphate cyclique
<b>ARN</b>	: Acide ribonucléique
<b>ATP</b>	: Adénosine triphosphate
<b>BLAST</b>	: Basic Local Alignment Search Tool
<b>BSA</b>	: Bovine Serum Albumin (ou Albumine de sérum bovin)
<b>cc</b>	: carbon cured
<b>CCM</b>	: Chromatographie sur Couche Mince
<b>CcpA</b>	: Catabolite Control Protein A
<b>CCR</b>	: Carbon Catabolite Repression (ou répression catabolique)
<b>cf</b>	: Confer
<b>Cm</b>	: Chloramphénicol
<b>CoA</b>	: CoenzymeA
<b>cpm</b>	: count per million
<b>DHAP</b>	: Dihydroxyacétone phosphate
<b>DO</b>	: Densité Optique
<b>DTT</b>	: Dithiothréitol
<b>Eau UP</b>	: Eau Ultra Pure
<b>EDTA</b>	: Acide éthylènediaminetétraacétique
<b>EMSA</b>	: Electrophoretic Mobility Shift Assay
<b>EPS</b>	: Exopolysaccharides
<b>Erm</b>	: Erythromycine
<b>FAD/FADH<sub>2</sub></b>	: Flavine adénine dinucléotide
<b>FAM</b>	: Carboxyfluorescéine
<b>FBP</b>	: Fructose-1,6-bisphosphate
<b>G1P</b>	: Glucose-1-phosphate
<b>G6P</b>	: Glucose-6-phosphate
<b>GA3P</b>	: Glycéraldéhyde-3-phosphate
<b>GalNAc</b>	: N-acétylgalactosamine
<b>GAG</b>	: glycoaminoglycane
<b>GRAS</b>	: Generally Recognized As Safe (ou généralement considéré comme sans danger)
<b>HPLC</b>	: High Performance Liquid Chromatography
<b>InVS</b>	: Institut National de Veille Sanitaire
<b>IPTG</b>	: Isopropyl-β-D-thiogalactopyranoside
<b>IUCPA</b>	: Union Internationale de Chimie Pure et Appliquée
<b>Kan</b>	: Kanamycine
<b>kb</b>	: kilobase
<b>KDPG</b>	: 2-céto-3-déoxy-6-phosphogluconate

<b>L/g/M</b>	: Litre, gramme, molaire
<b>LB</b>	: Lysogenic Broth
<b>LCFA</b>	: Long Chain Fatty Acid (ou Acide gras à longue chaîne)
<b>M6P</b>	: Maltose-6'-phosphate
<b>Man</b>	: Mannose
<b>MBP</b>	: Maltose Binding Protein
<b>MCS</b>	: Multiple Cloning Site (ou Site Multiple de Clonage)
<b>mL/mg/mM/mCi</b>	: milli -litre, -gramme, -molaire, -Curie
<b>MOPS</b>	: acide 3-morpholino-1-propanesulfonique
<b>MST</b>	: MicroScale Thermophoresis
<b>NAD/NADH</b>	: Nicotinamide adénine dinucléotide
<b>NADP/NADPH</b>	: Nicotinamide adénine dinucléotide phosphate
<b>NAG</b>	: N-acétylglucosamine
<b>Ni-NTA</b>	: Nickel-Nitrilotriacetic Acid (ou Acide Nitrilotriacétique De Nickel)
<b>OMS</b>	: Organisation Mondiale de la Santé
<b>PAGE</b>	: Poly Acrylamide Gel Electrophoresis (ou électrophorèse en gel de polyacrylamide)
<b>PAI</b>	: Pathogenicity Island (ou Îlot de pathogénicité)
<b>pb</b>	: paire de bases
<b>PCR</b>	: Polymerase Chain Reaction (ou Réaction de Polymérisation en Chaîne)
<b>PEP</b>	: Phosphoénolpyruvate (ou Acide phosphoénolpyruvique)
<b>PNAG</b>	: poly- $\beta$ -1,6-N-acetylglucosamines
<b>PTS</b>	: PhosphoTransferase System (Système phosphotransférase)
<b>PRD</b>	: PTS Regulation Domain (
<b>RACE</b>	: Rapid Amplification of cDNA Ends
<b>RBS</b>	: Ribosome Binding Site (ou site de fixation des ribosomes)
<b>rpm</b>	: rotation per minute (ou tour par minute)
<b>RT</b>	: Retro Transcription (ou transcription inverse)
<b>SCFA</b>	: Short Chain Fatty Acid (ou Acide gras à courte chaîne)
<b>SDS</b>	: Sodium Dodecyl Sulfate
<b>SIN</b>	: Signalement des Infections Nosocomiales
<b>SIS</b>	: Sugar Isomerase
<b>SOB</b>	: Super Optimal Broth
<b>SOC</b>	: SOB with Carbon Catabolite Repression
<b>TAE</b>	: Tris Acetate EDTA
<b>TBE</b>	: Tris Borate EDTA
<b>Th</b>	: Température d'exposition des surfaces hydrophobes
<b>TMD</b>	: TransMembrane Domain (Domaine Transmembranaire)
<b>TSA</b>	: Thermal Shift Assays
<b>TSS</b>	: Transcription Start Site (Site de début de transcription)
<b>U2RM</b>	: Unité de Recherche Risques Microbiens
<b>VPP</b>	: Voie des Pentoses Phosphates
<b>WT</b>	: Wild Type (sauvage)
<b>x-gal</b>	: 5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside

## Liste des figures

Figure 1 : Grandes fonctions des gènes d' <i>E. faecalis</i> différentiellement exprimés lors d'une infection chez la souris .....	14
Figure 2 : Colonisation de différents organes de souris par la bactérie <i>Enterococcus faecalis</i> .....	15
Figure 3 : Proportion des agents responsables d'infections nosocomiales en France de 2001 à 2017 .....	20
Figure 4 : Exemples de facteurs impliqués dans la virulence intrinsèque d' <i>E. faecalis</i> .....	22
Figure 5 : Schéma simplifié du métabolisme carboné d' <i>E. faecalis</i> .....	24
Figure 6 : GlpF, facilitateur du glycérol chez <i>E. faecalis</i> .....	27
Figure 7 : Structure d'un importeur ABC.....	28
Figure 8 : Structure des protéines NBP .....	29
Figure 9 : Organisation des deux sous-familles CUT1 et CUT2 d'importeurs ABC.....	30
Figure 10 : Importeur ABC EF_0176-0180 chez <i>E. faecalis</i> .....	32
Figure 11 : Fonctionnement des transporteurs actifs secondaires comme symporteurs et antiporteurs .....	33
Figure 12 : Différents états adoptés par les transporteurs actifs secondaires .....	34
Figure 13 : CitH, importeur actif secondaire du citrate chez <i>E. faecalis</i> .....	35
Figure 14 : Schéma du fonctionnement général des PTS .....	36
Figure 15 : Structure de l'enzyme EI. ....	36
Figure 16 : Structure de la protéine HPr .....	37
Figure 17 : Schéma des deux activités kinase et phosphorylase de la protéine HPrK/P .....	41
Figure 18 : Séquence consensus des boîtes cre proposée chez <i>E. faecalis</i> V583.....	42
Figure 19 : Structure des régulateurs à domaines PRD .....	45
Figure 20 : Schéma du réseau de régulation dépendant des PTS .....	46
Figure 21 : Structure des régulateurs transcriptionnels de la famille LacI/GalR .....	47
Figure 22 : Structure quaternaire de la protéine LacI d' <i>E. coli</i> .....	49
Figure 23 : Structure primaire d'un régulateur transcriptionnel de la famille RpiR .....	50
Figure 24 : Représentations du D-glucose.....	51
Figure 25 : Structure générale des maltodextrines .....	52
Figure 26 : Gènes impliqués dans le métabolisme du maltose et des maltodextrines chez <i>E. coli</i> .....	53
Figure 27 : Réactions catalysées par l'amyloamylase MalQ d' <i>E. coli</i> .....	54
Figure 28 : Schéma du métabolisme du maltose et des maltodextrines chez <i>E. coli</i> .....	55
Figure 29 : Modèle de régulation du métabolisme du maltose et des maltodextrines chez <i>E. coli</i> .....	56
Figure 30 : Gènes impliqués dans le métabolisme du maltose et des maltodextrines chez <i>S. pneumoniae</i> ....	57
Figure 31 : Schéma du métabolisme du maltose et des maltodextrines chez <i>S. pneumoniae</i> .....	58
Figure 32 : Modèle de régulation du métabolisme du maltose et des maltodextrines chez <i>S. pneumoniae</i> ...	59
Figure 33 : Gènes impliqués dans le métabolisme du maltose et des maltodextrines chez <i>E. faecalis</i> .....	60
Figure 34 : Schéma du métabolisme du maltose et des maltodextrines chez <i>E. faecalis</i> .....	61
Figure 35 : Représentation des $\beta$ -glycosides gentiobiose et cellobiose en conformation chaise .....	62
Figure 36 : Gènes impliqués dans le métabolisme de différents $\beta$ -glycosides chez <i>E. coli</i> .....	63

Figure 37 : Schéma du métabolisme de $\beta$ -glucosides chez <i>E. coli</i> .....	63
Figure 38 : Modèle de régulation du métabolisme de $\beta$ -glucosides chez <i>E. coli</i> .....	64
Figure 39 : Gènes impliqués dans le métabolisme du cellobiose chez <i>S. pneumoniae</i> .....	65
Figure 40 : Schéma du métabolisme du cellobiose chez <i>S. pneumoniae</i> .....	65
Figure 41 : Modèle de régulation du métabolisme de $\beta$ -glucosides chez <i>S. pneumoniae</i> .....	66
Figure 42 : Région intergénique <i>malT/mapP</i> .....	139
Figure 43 : Promoteur et motif MalR prédits en amont du gène OG1RF_11753 .....	141
Figure 44 : Séquence en amont du gène <i>msmX</i> (OG1RF_10665) .....	142
Figure 45 : Séquence consensus des motifs de reconnaissance proposée pour le régulateur transcriptionnel MalR chez <i>E. faecalis</i> OG1RF .....	179
Figure 46 : Homologies des protéines codées par les opérons maltose chez <i>E. faecalis</i> et <i>E. faecium</i> .....	180
Figure 47 : Promoteur et motif mdxR prédit en amont du gène <i>mdxD</i> (E1162_2125).....	181
Figure 48 : Recherche d'homologies des protéines impliquées dans le métabolisme du cellobiose de <i>S. pneumoniae</i> R6 chez <i>E. faecalis</i> OG1RF .....	225
Figure 49 : Modèle de régulation du métabolisme du maltose et des maltodextrines chez <i>E. faecalis</i> .....	230
Figure 50 : Schéma du métabolisme du gentiobiose et du cellobiose chez <i>E. faecalis</i> .....	232
Figure 51 : Modèle de régulation de l'opéron <i>genBA</i> chez <i>E. faecalis</i> OG1RF .....	233
Figure S.1 : Représentation de $\beta$ -glycosides utilisés dans cette étude. ....	262

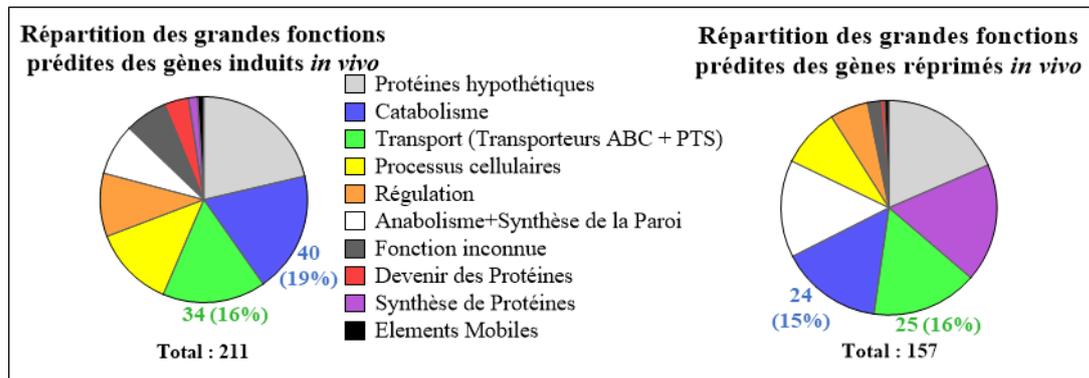
## Liste des tableaux

Tableau 1 : Opérons induits en condition d'infection et caractérisés plus en détails dans la suite de l'étude .	14
Tableau 2 : Caractéristiques communes des entérocoques .....	19
Tableau 3 : Caractéristiques des importeurs ABC CUT1 et CUT2 .....	30
Tableau 4 : Caractéristiques des grandes familles de PTS .....	38
Tableau 5 : Gènes impliqués dans le métabolisme du maltose et des maltodextrines chez différentes espèces de streptocoques .....	57
Tableau 6 : Souches bactériennes utilisées dans cette étude .....	69
Tableau 7 : Plasmides utilisés lors de ce travail .....	70
Tableau 8 : Gènes présentant une expression différentielle sur maltose et/ou chez le mutant <i>malR</i> - par rapport à une croissance sur glucose .....	138
Tableau 9 : Expression des gènes impliqués dans le métabolisme du gentiobiose et/ou cellobiose chez <i>E. faecalis</i> .....	226
Tableau 10 : Couples de monosaccharides liés par des liaisons $\beta$ -1,6 dans les glycosylations de protéines de mammifères.....	235
Tableau S.1 : Composition du milieu de culture LB .....	263
Tableau S.2 : Composition du milieu de culture M17 .....	263
Tableau S.3 : Composition du milieu de culture ccM17 .....	263
Tableau S.4 : Composition du milieu de culture SOB .....	264
Tableau S.5 : Composition du milieu de culture SOC .....	264
Tableau S.6 : Composition du milieu de culture Terrific Broth .....	264
Tableau S.7 : Amorces utilisées pour les purifications de protéines recombinantes .....	265
Tableau S.8 : Amorces utilisées pour les expériences de RT-qPCR .....	265
Tableau S.9 : Amorces utilisées pour la création et la complémentation des mutants .....	266
Tableau S.10 : Amorces utilisées pour les expériences de 5'RACE-PCR .....	267
Tableau S.11 : Autres amorces utilisées dans cette étude .....	267

# Contexte

Le laboratoire "Unité de Recherche Risques Microbiens" (U2RM) dans lequel les travaux présentés dans ce manuscrit ont été réalisés, focalise ses recherches sur l'étude des entérocoques et notamment sur les deux espèces *Enterococcus faecalis* et *Enterococcus faecium*. L'objectif de ce laboratoire est de mieux comprendre le caractère pathogène des entérocoques ainsi que leurs capacités de résistances face à différents stress. Par exemple, un des stress fréquemment rencontré par ces bactéries lors d'infections est le stress oxydatif qui représente une des premières barrières du système immunitaire inné notamment par l'intermédiaire des cellules phagocytaires. La réponse des entérocoques au stress oxydatif représente par conséquent un des axes de recherche majeur du laboratoire (<sup>1-6</sup>). La gestion de ce stress oxydatif apparait centrale chez ces bactéries, des liens avec d'autres aspects de la vie de ce microorganisme comme le métabolisme et notamment celui du glycérol ont pu être établis. En effet, ce dernier peut générer un stress oxydatif interne par la formation au niveau intracellulaire de molécules H<sub>2</sub>O<sub>2</sub> via la voie oxydative de dégradation du glycérol glpK (<sup>7,8</sup>). Une étude menée chez *E. faecalis* semble également faire le lien entre la bactéricidie d'un traitement antibiotique et la génération d'un stress oxydatif (<sup>4</sup>). La résistance de ces deux espèces d'entérocoques à différents agents antibactériens comme les antibiotiques ou les peptides antimicrobiens représente également un axe de recherche important pour le laboratoire (<sup>9-12</sup>). En plus de la résistance de ces bactéries à différents facteurs extérieurs, plusieurs acteurs influençant directement ou non leur virulence ont également été investigués comme des petits ARN (<sup>13</sup>), des protéines cytoplasmiques (<sup>14-16</sup>) ou encore des protéines de surface (<sup>17-19</sup>).

C'est dans ce contexte qu'une étude globale a été entreprise afin de mieux connaître les gènes induits chez la bactérie *E. faecalis* lors d'une infection en modèle murin. Les gènes induits en condition d'infection sont supposés jouer un rôle dans ce processus et par conséquent être importants pour le caractère virulent de la bactérie. Pour cela, une infection du péritoine de la souris a été réalisée suivie d'un séquençage du transcriptome entier (RNA-sequencing). Cette étude a permis de déterminer 368 gènes exprimés différemment en condition *in vivo* comparée à une croissance en milieu BHI servant de référence (<sup>20</sup>). Sur les 368 gènes différemment exprimés dans cette étude, 211 ont été décrits comme induits et 157 comme réprimés lors de l'infection du péritoine chez la souris. L'identification des gènes différemment exprimés a permis de constater que les gènes prédits pour intervenir dans des fonctions cataboliques ou de transport de substrats représentent une forte proportion (**Figure 1**).



**Figure 1 :** Grandes fonctions des gènes d'*E. faecalis* différemment exprimés lors d'une infection chez la souris

Les nombres représentent l'ensemble des gènes concernés par la fonction, les nombres entre parenthèses reflètent leur proportion par rapport au total des gènes induits ou réprimés. D'après<sup>20</sup>.

Au vu de la forte proportion de gènes prédits pour intervenir dans des fonctions métaboliques au sein de ces gènes différemment exprimés *in vivo*, cette étude suggère que le métabolisme occupe une place importante dans la physiologie bactérienne en condition d'infection. Par conséquent, mieux comprendre le métabolisme de cette bactérie s'avère être une étape cruciale pour aboutir à une meilleure compréhension de son comportement vis-à-vis de son hôte. Parmi les gènes retrouvés induits dans cette étude, deux opérons ont retenu notre attention. Ces derniers sont identifiés comme impliqués dans le métabolisme de polysaccharides, présentent un facteur d'induction "*in vivo*" vérifié par RT-qPCR supérieur à 10 par rapport au témoin ayant évolué en milieu BHI et n'ont pas encore été caractérisés. Ces structures génétiques sont présentées dans le **Tableau 1**.

**Tableau 1 :** Opérons induits en condition d'infection et caractérisés plus en détails dans la suite de l'étude

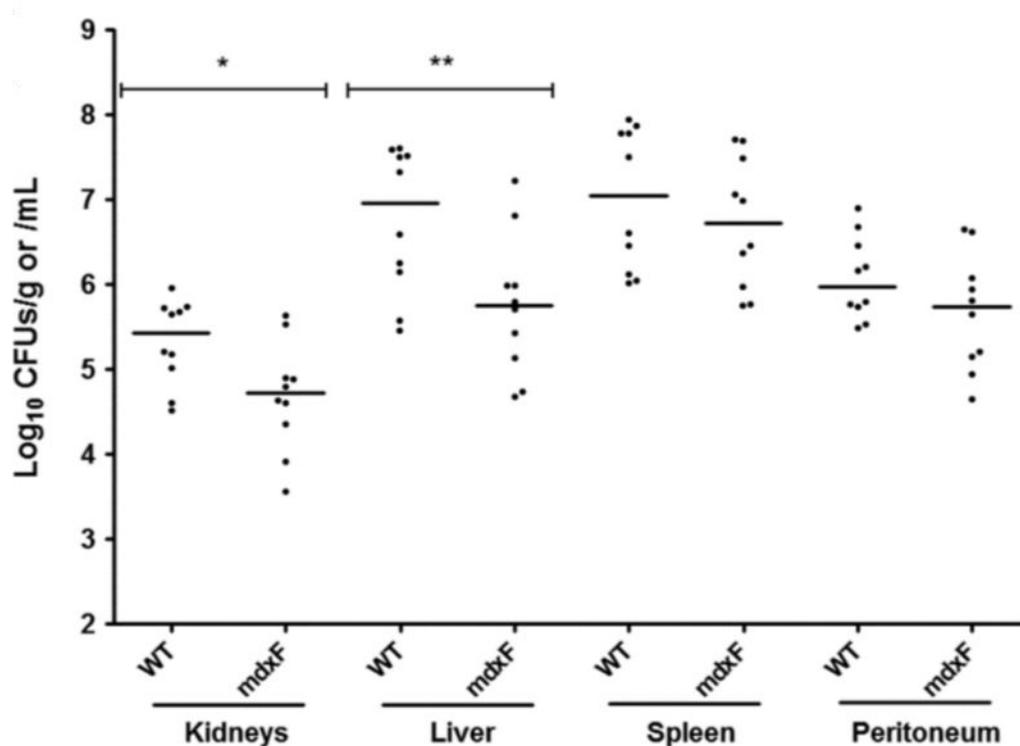
Numéro de gène (V583)	Numéro de gène (OG1RF)	Code à 4 lettres	FI <sup>2</sup> souris/phase exponentielle (RNA-seq)	FI <sup>2</sup> souris/phase exponentielle (RT-qPCR)	FI <sup>2</sup> souris/phase stationnaire (RNA-seq)
<i>ef_0291</i>	<i>OG1RF_10234</i>	<i>celA (genA<sup>1</sup>)</i>	94,9	11,1	5,1
<i>ef_0292</i>	<i>OG1RF_10235</i>	<i>celB (genB<sup>1</sup>)</i>			
<i>ef_1345</i>	<i>OG1RF_11133</i>	<i>mdxE</i>	31,6	10,8	3,7
<i>ef_1344</i>	<i>OG1RF_11134</i>	<i>mdxF</i>			
<i>ef_1343</i>	<i>OG1RF_11135</i>	<i>mdxG</i>			

<sup>1</sup> Le nom de gène entre parenthèses correspond au nouveau nom attribué après sa caractérisation.

<sup>2</sup> FI signifie Facteur d'induction

Le premier opéron (*ef0291-0292* correspondant aux gènes *OG1RF\_10234-10235* chez la souche OG1RF) code pour une  $\beta$ -glucoside hydrolase (renommée GenA) et une protéine EIIC<sup>cel</sup> (renommée GenB) d'un système PTS (voir la partie II.D de l'introduction bibliographique). Une

caractérisation de cet opéron fera l'objet d'une étude plus approfondie dans la suite de ce manuscrit. Le second opéron (*ef1345-1347* correspondant à l'opéron *OG1RF\_11133-11135* chez la souche OG1RF) code pour un transporteur ABC (ATP Binding Cassette). Ce transporteur est composé d'une protéine de liaison du ligand MdxE et de deux protéines transmembranaires MdxF et MdxG. Suite à cette observation, des études récentes menées en partie au laboratoire ont permis de montrer que ce transporteur était impliqué dans l'import de différents types de sucres de la famille des dextrines, notamment le panose, l'isopanose et les maltodextrines dont la longueur de la chaîne est supérieure à 3 résidus glycosyl (<sup>21</sup>). Un mutant délété pour le gène *mdxF* (*ef1346/OG1RF\_11134*), qui fait partie intégrante du transporteur ABC décrit précédemment, possède une capacité réduite de colonisation du foie et du rein chez la souris, comparée à sa souche isogénique sauvage (**Figure 2**).



**Figure 2** : Colonisation de différents organes de souris par la bactérie *Enterococcus faecalis*. La colonisation des organes de souris a été déterminée chez la souche JH2-2 et chez le mutant *mdxF* après injection intrapéritonéale chez des souris Swiss White. D'après (<sup>21</sup>).

Toutes ces informations laissent supposer que ces métabolismes représentent des éléments importants pour permettre à la bactérie *E. faecalis* une colonisation efficace qui est une étape précoce cruciale de l'infection. L'objectif de ce travail de thèse est de présenter un modèle le plus complet possible pour expliquer la régulation du métabolisme du maltose et des maltodextrines ainsi que de définir le rôle de l'opéron *genBA* chez *E. faecalis*. Pour cela, nous présenterons dans un premier temps dans la suite de ce manuscrit le modèle d'étude *E. faecalis* en insistant sur ses

caractéristiques métaboliques, nous détaillerons ensuite des systèmes clefs dans l'import de substrats carbonés ainsi qu'une famille de régulateurs transcriptionnels fréquemment impliqués dans la régulation métabolique procaryote. Nous finirons enfin par une revue du métabolisme du maltose/maltodextrines et des  $\beta$ -glucosides chez différents organismes : d'abord le modèle procaryote *Escherichia coli* puis des bactéries plus proches phylogénétiquement de notre modèle d'intérêt que sont les Streptocoques, nous traiterons ensuite les métabolismes du maltose/maltodextrines et du gentiobiose chez *E. faecalis*.

# **Introduction**

# **bibliographique**

# I. Le genre Enterococcus

## A. Historique

Les entérocoques ont été décrits pour la première fois en 1899 par deux équipes en parallèle. L'une de ces études fait état d'une endocardite chez un jeune patient de 37 ans causée par une nouvelle bactérie que MacCallum et Hastings nommeront *Micrococcus zymogenes* <sup>(22)</sup>. La même année, Thiercelin caractérise une nouvelle bactérie retrouvée dans les fèces humaines qu'il décrit comme des diplocoques parfois organisés en courtes chainettes <sup>(23)</sup>. Il propose le nom d'entérocoque pour rappeler à la fois leur morphologie et leur habitat naturel mais ce nom ne sera retenu que près d'un siècle plus tard. On sait aujourd'hui que ces deux études font vraisemblablement référence à *Enterococcus faecalis*. Les entérocoques seront néanmoins longtemps associés au genre *Streptococcus* où ils seront répertoriés comme "Streptocoques du groupe D" ou encore "Streptocoques de Lancefield". Il faudra attendre 1984 pour que les deux espèces *Streptococcus faecalis* et *Streptococcus faecium* soient phylogénétiquement attribuées au genre *Enterococcus* devenant par la même les deux premiers représentants connus de ce genre bactérien <sup>(24)</sup>. La connaissance du genre *Enterococcus* s'étoffe avec le temps et le nombre d'espèces associées a augmenté avec plus de 30 espèces d'entérocoques différentes décrites en 2012 <sup>(25)</sup>. Actuellement, 57 espèces différentes d'entérocoques ont été décrites d'après le site "List of prokaryotic names with standing in nomenclature" (LPSN) <sup>(26-28)</sup>.

## B. Caractéristiques des entérocoques

### 1. **Caractéristiques générales**

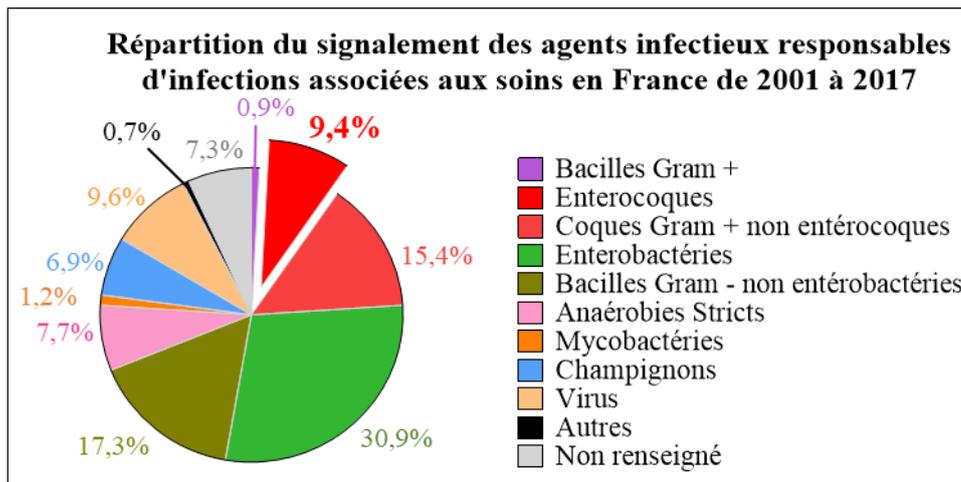
Les entérocoques sont phylogénétiquement classés dans le phylum des firmicutes. Ce sont des bactéries positives à la coloration de Gram qui présentent un faible pourcentage de bases GC dans leur génome. Les entérocoques sont des cellules en forme de coques souvent associés par paires et parfois en courtes chaînettes, non sporulantes et mésophiles, leur température optimale de développement se situe aux environs de 35°C <sup>(25)</sup>. Ces bactéries sont capables de supporter des conditions défavorables comme de fortes concentrations en NaCl, la présence de sels biliaires, l'absence prolongée de nutriments ou encore la dessiccation. Ces microorganismes présentent également de nombreuses résistances aux antibiotiques, qu'elles soient intrinsèques ou acquises <sup>(29)</sup>. Quelques caractéristiques communes partagées par les différentes espèces d'entérocoques sont présentées dans le **Tableau 2**.

**Tableau 2** : Caractéristiques communes des entérocoques

	<b>Caractéristiques communes</b>	<b>Référence</b>
<b>Type trophique</b>	Chimio Organo Hétérotrophe	( <sup>30</sup> )
<b>Type respiratoire</b>	Anaérobie Facultatif	( <sup>30</sup> )
<b>Température de croissance</b>	10 – 45 °C (Optimale ~35°C) Survie > 30 min à 60°C	( <sup>25,30</sup> )
<b>Taille de génome</b>	2,7 à 3,6 Mpb	( <sup>25</sup> )
<b>Pourcentage en GC</b>	37 à 45 %	( <sup>25</sup> )
<b>Catalase</b>	Négative (certaines espèces possèdent une catalase dépendante de l'apport d'hème pour être fonctionnelle)	( <sup>30</sup> )
<b>Tolérance au pH</b>	pH 2,9 – pH 9,6	( <sup>25,31</sup> )
<b>Résistance osmotique</b>	6,5 % NaCl	( <sup>25</sup> )
<b>Résistance à la dessiccation</b>	Plusieurs semaines à plusieurs mois	( <sup>32,33</sup> )
<b>Résistances intrinsèques aux antibiotiques</b>	- $\beta$ -lactamines • pénicilline (résistances de bas niveau) • céphalosporines - Aminosides (résistances de bas niveau)	( <sup>29,34</sup> )

Les entérocoques peuvent être retrouvés de manière ubiquitaire dans la nature mais le plus souvent associés aux animaux (<sup>35</sup>). *E. faecalis* et *E. faecium* sont les espèces les plus documentées à ce jour. Elles sont décrites comme étant les espèces d'entérocoques les plus abondantes dans les fèces humaines (<sup>25</sup>). Ces deux espèces représenteraient typiquement environ 0,1 % de la flore microbienne chez l'Homme en bonne santé (<sup>36,37</sup>) où elles occupent alors une place de microorganismes commensaux. Certaines souches d'*E. faecalis* ont par exemple été décrites comme capables d'assimiler le cholestérol ou encore de produire des acides gras à longues et courtes chaînes (LCFA et SCFA respectivement) reconnus comme bénéfiques pour la santé humaine (<sup>38-41</sup>).

Malgré leur caractère commensal pour l'Homme, les entérocoques sont fréquemment décrits comme des bactéries responsables d'infections nosocomiales associées aux soins ou acquises en milieu hospitalier (<sup>42</sup>). Certaines souches peuvent en effet agir comme des germes opportunistes et être responsables de maladies graves comme des endocardites principalement chez des personnes immunodéprimées ou affaiblies (<sup>43</sup>). D'après un rapport de juillet 2018 de l'Institut National de Veille Sanitaire (InVS) basé sur une surveillance épidémiologique couvrant la période de 2001 à 2017, les entérocoques représentent en France la deuxième cause bactérienne la plus enregistrée dans le cadre de signalement externe des infections nosocomiales (SIN) avec 9,4% derrière l'entérobactérie *Klebsiella pneumoniae* (**Figure 3**).



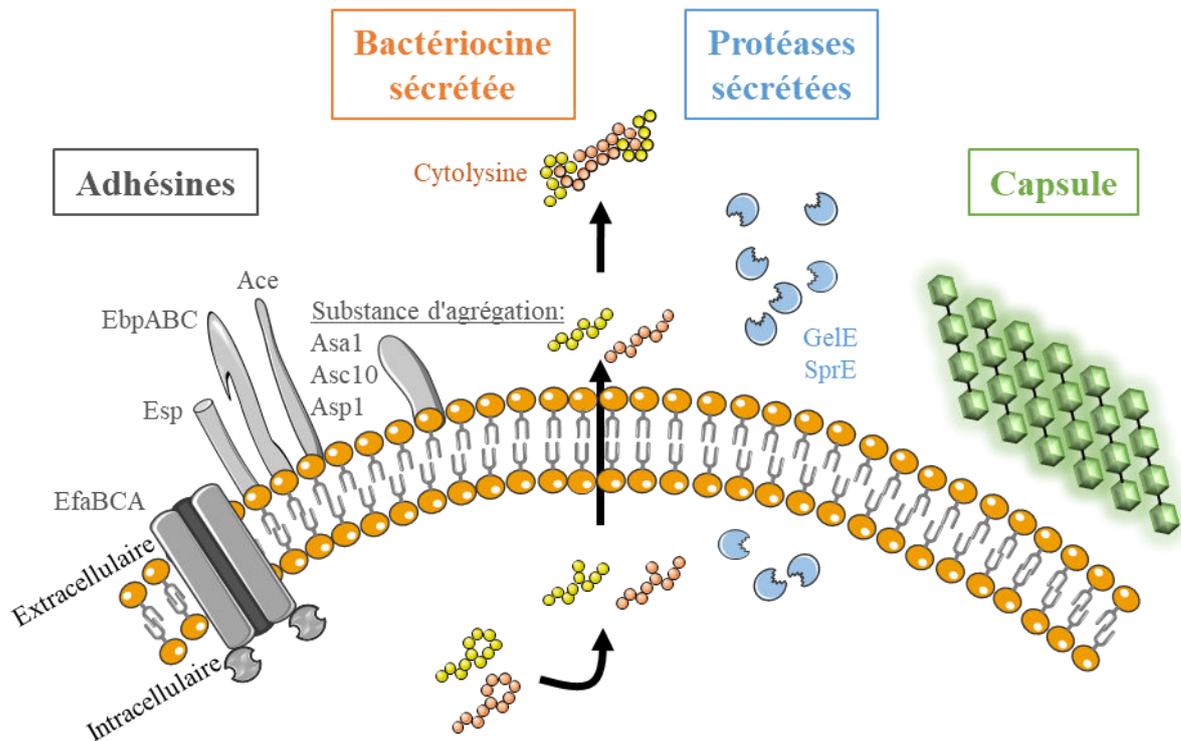
**Figure 3** : Proportion des agents responsables d'infections nosocomiales en France de 2001 à 2017 D'après le rapport de juillet 2018 de l'InVS (<sup>44</sup>).

Cette présence récurrente en milieu hospitalier peut s'expliquer, au moins en partie, par leurs fortes tolérances aux conditions de leur environnement et leurs résistances aux antibiotiques énoncées précédemment. *E. faecium* a été classé par l'Organisation Mondiale de la Santé (OMS) comme un agent pathogène de priorité élevée pour la recherche et le développement de nouvelles méthodes de lutte en raison de ses nombreuses résistances aux antibiotiques, notamment à la vancomycine (<sup>45</sup>).

En plus de leurs fortes capacités d'adaptation et de résistance, les entérocoques présentent une virulence intrinsèque dont plusieurs facteurs ont été identifiés particulièrement chez *E. faecalis* (**Figure 4**) (<sup>46,47</sup>). La première étape du processus d'infection est l'adhésion des cellules bactériennes à celles de l'hôte, pour cela *E. faecalis* possède des protéines de surface regroupées sous le nom d'adhésines. On retrouve parmi elles un groupe de protéines nommées "substance d'agrégation" ou AG. Ces protéines sont ancrées dans la paroi par un motif LPxTG et possèdent deux motifs RGD (Arg-Gly-Asp). Ces protéines permettent l'agrégation bactérienne indispensable à la conjugaison de plasmides. Cette agrégation est le plus souvent stimulée en réponse à des phéromones bactériennes. L'AG peut également favoriser l'adhésion de la bactérie à la matrice extra cellulaire ou directement aux cellules de l'hôte. La protéine Ace représente un autre exemple de protéine de surface permettant une adhésion efficace à la matrice extracellulaire de l'hôte, la classant ainsi dans les constituants MSCRAMM (Microbial Surface Components Recognizing Adhesive Matrix Molecules) (<sup>48</sup>). Un autre acteur, la protéine EfaA, constituant d'un transporteur ABC a également été relié à la virulence d'*E. faecalis* même si la nocivité atténuée d'un mutant pour cette protéine pourrait s'expliquer par une perte de fitness plutôt que par une perte de virulence directe. D'autres protéines comme les pili EbpABC ou la protéine Esp joueraient quant à elles un rôle dans la formation de biofilm qui est un facteur

important dans le processus infectieux et pourrait avoir un rôle dans l'adhésion aux cellules de l'hôte. Le biofilm représente chez les bactéries une organisation communautaire favorisant la mise en place de communication intercellulaire via le système de Quorum Sensing (QS). Ce QS régule via le système Fsr des gènes également impliqués dans la virulence d'*E. faecalis* comme par exemple les protéases sécrétées GeIE et SprE qui possèdent une activité endopeptidase à large spectre. Elles seraient impliquées dans la dégradation des protéines de l'hôte, l'autolyse, le renouvellement des protéines de surface, le relargage d'ADN extracellulaire voire la translocation à travers des épithéliums de l'hôte <sup>(46)</sup>. En plus de ces protéases sécrétées, certaines souches d'*E. faecalis* sont capables de sécréter une bactériocine de la famille des lantibiotiques appelée cytolysine. Cette dernière confère un avantage en limitant la compétition avec d'autres espèces procaryotes mais peut également cibler des cellules eucaryotes et serait potentiellement capable d'affecter les cellules du système immunitaire inné de l'hôte <sup>(46)</sup>. Une autre stratégie pour éviter les défenses immunitaires de l'hôte est la synthèse d'une capsule polysaccharidique via les gènes *cpsC-K* <sup>(48)</sup> qui confère également un avantage en contexte infectieux. Une synthèse des différents facteurs de virulence précédemment décrits est présentée en **Figure 4**.

Tous ces facteurs de virulence peuvent favoriser le caractère pathogène d'*E. faecalis* mais il apparaît que ce comportement opportuniste repose plus sur la nature des interactions avec l'hôte que sur leur simple présence ou absence. Certaines souches sont dépourvues de la majorité de ces facteurs de virulence et sont néanmoins capables de causer des infections chez l'Homme <sup>(47)</sup>. Pour persister dans ces environnements parfois hostiles avec une forte compétitivité entre microorganismes, *E. faecalis* a dans un premier temps besoin de rapidement adapter son métabolisme à son environnement direct.



**Figure 4 :** Exemples de facteurs impliqués dans la virulence intrinsèque d'*E. faecalis* D'après <sup>(46)</sup>.

## 2. Caractéristiques métaboliques d'*E. faecalis*

*E. faecalis* est une bactérie chimio-organo-hétérotrophe aérobie/anaérobie facultative. Elle dégrade ainsi la matière organique via des réactions d'oxydo-réductions pour produire son énergie. Cette bactérie peut utiliser une large gamme de substrats pour supporter son développement. Elle est notamment capable d'utiliser plusieurs dizaines de sucres différents <sup>(30)</sup> mais elle possède également la capacité de métaboliser d'autres molécules comme le glycérol, différents acides organiques comme le citrate, le malate et le lactate <sup>(7,49-51)</sup> ou encore des molécules dérivées d'acide aminés comme l'agmatine <sup>(52)</sup>. Pour utiliser les différents sucres comme source de carbone et d'énergie, *E. faecalis* est capable de réaliser les trois voies de dégradation majeures du glucose que sont la glycolyse (aussi appelée voie d'Embden-Meyerhof-Parnas), la voie des pentoses phosphates (voie de Warburg-Dickens-Horecker ou VPP) et la voie du 2-céto-3-désoxy-6-phosphogluconate (KDPG) (voie d'Entner-Doudoroff) <sup>(30)</sup>. La voie principale de la dégradation du glucose est la glycolyse, elle peut se découper en 2 parties : la voie haute qui consomme de l'ATP pour produire du fructose-1,6-bisphosphate (FBP) et la voie basse qui convertit le glycéraldéhyde-3-phosphate (GA3P) en phosphoénolpyruvate (PEP) en plusieurs étapes. Cette voie basse de la glycolyse permet également de générer des molécules d'ATP par phosphorylation au niveau du substrat

fournissant ainsi de l'énergie à la cellule. Les deux autres voies (VPP et KDPG) aboutissent, quant à elles, à la formation de GA3P qui pourra rejoindre la voie basse de la glycolyse.

Le PEP, formé grâce à la voie basse de la glycolyse, représente un précurseur d'une molécule centrale du métabolisme cellulaire qui est le pyruvate. Le PEP peut être directement transformé en pyruvate par l'action de la pyruvate kinase lors de la dernière étape de la glycolyse produisant une mole d'ATP par mole de produit formé (<sup>53,54</sup>). Cette transformation peut également s'effectuer par le biais de l'enzyme EI commune aux systèmes PhosphoTransferaseSystem (PTS) (voir partie II.D). Dans ce cas, l'énergie libérée lors de l'hydrolyse de la liaison anhydre du PEP permettra indirectement l'import de saccharide. Lors des réactions des voies métaboliques présentées ci-dessus, des cofacteurs oxydés tels que le Nicotinamide adénine dinucléotide (NAD<sup>+</sup>) et le Nicotinamide adénine dinucléotide phosphate (NADP<sup>+</sup>) seront réduits pour former leurs équivalents réduits, respectivement le NADH,H<sup>+</sup> et le NADPH,H<sup>+</sup>. Ces cofacteurs étant essentiels pour de nombreuses réactions cellulaires et n'étant pas consommés, la cellule doit les régénérer en leurs équivalents oxydés. Cette régénération des cofacteurs peut s'effectuer via la respiration, la fermentation ou encore des réactions d'anabolismes.

### **3. La fermentation du pyruvate chez *E. faecalis***

*E. faecalis* est une bactérie aérobie/anaérobie facultative, en l'absence d'oxygène, elle adopte un métabolisme fermentaire pour régénérer ses cofacteurs et éliminer le pyruvate qui risquerait de s'accumuler dans la cellule et bloquer les réactions de catabolisme par excès de produit.

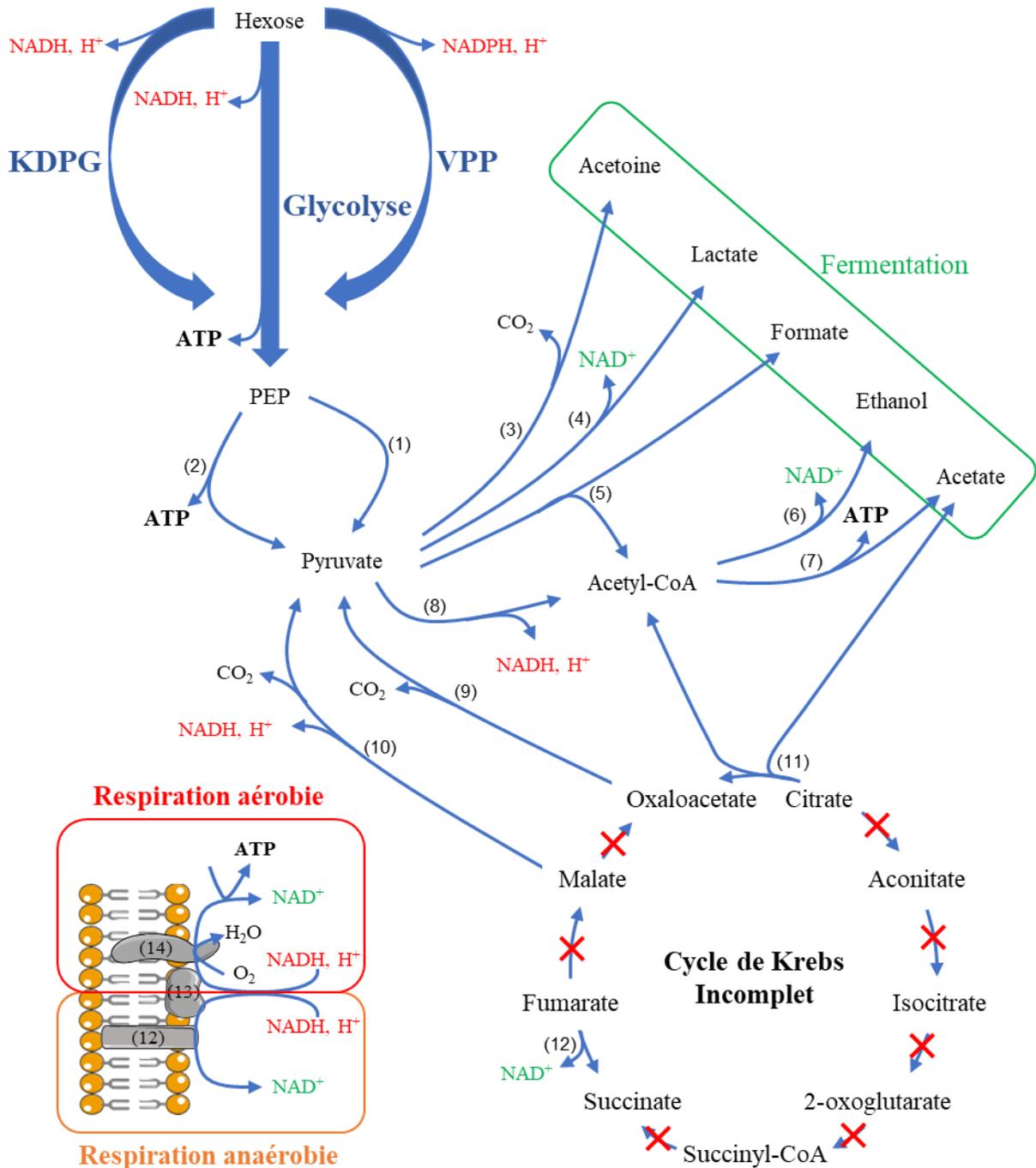
Les entérocoques sont des bactéries lactiques et sont par conséquent capables de convertir le pyruvate en lactate. Elles sont classées comme les Streptocoques ou les Pédicoques dans le groupe des bactéries homofermentaires (<sup>55,56</sup>), c'est-à-dire produisant une large majorité de lactate comme produit de fermentation des sucres. Ce lactate produit par les entérocoques est décrit comme étant optiquement pur en configuration L-(+) (<sup>57,58</sup>). Ces caractéristiques de fermentation s'avèrent intéressantes pour le milieu de l'industrie notamment alimentaire où certaines souches d'entérocoques peuvent être utilisées comme ferments lactiques (<sup>59</sup>) ou posséder des propriétés probiotiques (<sup>40</sup>). Malgré leur caractère homolactique, les entérocoques sont capables de réaliser des fermentations autres que lactique. En plus du lactate, ils peuvent notamment convertir le pyruvate en acétate, en formate ou encore en éthanol selon les conditions de cultures et le substrat utilisé (<sup>50,60,61</sup>). Ces produits de fermentation seront exportés dans le milieu extérieur pour ne pas bloquer les voies de catabolisme (<sup>62</sup>). La fermentation reste

la méthode la plus utilisée par *E. faecalis* pour régénérer ses cofacteurs mais n'est cependant pas sa seule option, cette bactérie peut réaliser la respiration aérobie et anaérobie selon certaines conditions (**Figure 5**).

*E. faecalis* possède en effet une chaîne de transfert d'électrons lui permettant de respirer en utilisant le dioxygène (O<sub>2</sub>) comme accepteur final d'électrons (<sup>63,64</sup>). Cette réaction permet comme la fermentation de ré-oxyder les cofacteurs mais également d'augmenter la production d'énergie pour la cellule si elle est couplée à la phosphorylation oxydative. Ne possédant qu'un cycle de Krebs incomplet, *E. faecalis* n'est cependant pas capable de synthétiser le cytochrome *bd* essentiel au transfert d'électron jusqu'à l'O<sub>2</sub> car la porphyrine, un précurseur essentiel, ne peut pas être produite à partir de succinyl-CoA. Cette bactérie ne peut donc réaliser de respiration aérobie qu'en présence d'un apport extérieur d'hématine (<sup>63,64</sup>). Elle peut néanmoins utiliser le fumarate comme accepteur d'électrons en condition anaérobie via l'intermédiaire de la fumarate réductase. Cela lui permet de régénérer ses cofacteurs réduits mais ne peut pas être couplé à une production supplémentaire d'ATP car le transfert d'électron vers le fumarate ne permet pas l'export de proton dans le milieu extérieur (<sup>65,66</sup>). Cette bactérie possède enfin une NADH oxydase cytosolique permettant en condition aérobie d'oxyder directement le NADH en NAD<sup>+</sup> en réduisant une molécule d'O<sub>2</sub> en H<sub>2</sub>O (<sup>67,68</sup>).

### C. Les souches de référence : V583 et OG1RF

La souche utilisée lors de cette étude est la souche OG1RF alors que celle de référence pour l'étude d'*E. faecalis* est généralement la souche V583. La souche V583 s'est imposée comme souche de référence en étant la première dont le génome a été entièrement séquencé et diffusé en 2003 (<sup>69</sup>). Cette souche représente le premier isolat clinique à se révéler résistant à la vancomycine aux Etats-Unis, cette résistance est due à l'acquisition d'un transposon portant l'opéron VanB. La souche V583 possède beaucoup d'éléments mobiles comptant pour plus de 25% de son matériel génétique total, on y retrouve notamment le transposon précédemment cité mais également 7 prophages, 2 plasmides intégrés, 3 îlots génomiques dont 1 îlot de pathogénicité (PAI) positivement corrélé aux souches cliniques et 3 plasmides réplicatifs. En raison de tous ces éléments mobiles, cette souche possède un chromosome composé d'environ 3,2 Mpb, beaucoup plus large que le core-génome des entérocoques (<sup>69</sup>). Etant une souche clinique, elle est plus virulente que la plupart des autres souches d'*E. faecalis* et possède des résistances à diverses antibiotiques comme la vancomycine et les macrolides. Ces caractéristiques font de cette souche un microorganisme de classe 2. La souche OG1RF a été la deuxième souche dont le génome a été entièrement séquencé et disponible en 2008.



**Figure 5** : Schéma simplifié du métabolisme carboné d'*E. faecalis*

La stœchiométrie et la réversibilité des réactions ne sont pas indiquées pour plus de clarté.

KDPG : Voie du 2-céto-3-désoxy-6-phosphogluconate, VPP : Voie des Pentoses Phosphates

(1) Enzyme EI ; (2) Pyruvate Kinase ; (3) Acétolactate Synthase + Acétolactate Décarboxylase ; (4) Lactate Déshydrogénase ; (5) Pyruvate Formiate Lyase ; (6) Alcool Déshydrogénase ; (7) Phosphoacétyl Transferase + Acétate Kinase ; (8) Pyruvate Déshydrogénase ; (9) Oxaloacétate Décarboxylase ; (10) Malate Décarboxylase ; (11) Citrate Lyase ; (12) Fumarate Réductase ; (13) NADH Oxydo-réductase + Déméthylmenaquinone ; (14) Cytochrome bd + ATP Synthase

Contrairement à la souche V583, elle présente peu d'éléments mobiles et ne possède pas de plasmide. Elle ne possède pas non plus le PAI présent chez d'autres souches comme V583, V586 ou MMH594 (70). La souche OG1RF est par conséquent dépourvue de plusieurs facteurs de virulence décrits précédemment comme l'opéron cytolysine ou encore l'adhésine Esp (69). Malgré leur absence, elle reste cependant capable de causer des infections en modèle animal. La virulence des entérocoques semble donc plus difficile à décrire que la simple présence ou absence de PAI. Son génome d'environ 2,7 Mpb codant pour environ 2700 gènes apparaît beaucoup plus proche du core-génome d'*E. faecalis* que la souche V583 (70). La souche OG1RF se manipule plus facilement en laboratoire et reste moins pathogène pour l'Homme. Toutes ces caractéristiques lui ont permis de s'imposer de plus en plus comme souche de référence pour l'étude du métabolisme chez *E. faecalis*. La souche OG1RF provient initialement de la souche OG1. Cette dernière est une souche commensale qui a déjà été décrite comme pouvant causer des caries chez le rat (71,72). La souche OG1 a ensuite été sélectionnée en laboratoire pour résister à la rifampicine et a été nommée OG1-R (71). La souche OG1RF (initialement nommée OG1-RF1) dérive directement de la souche OG1-R, elle se différencie de sa souche parentale par une résistance à l'acide fusidique (73).

Nous avons vu dans la partie précédente que la croissance bactérienne était supportée par des réactions métaboliques aboutissant à la production d'énergie utilisable par la cellule. Afin de pouvoir assurer rapidement ces réactions cataboliques, la cellule nécessite dans un premier temps un import efficace de différents nutriments de son environnement direct vers le milieu intracellulaire.

## **II. L'import de substrats**

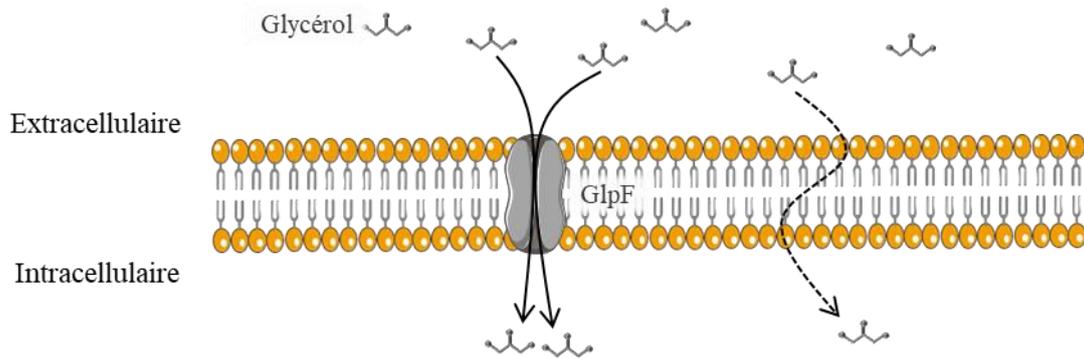
Les bactéries gram positives sont séparées de leur milieu extérieur par une unique membrane plasmique composée d'une bicouche flexible de phospholipides dans laquelle sont ancrées ou insérées plusieurs protéines. Les molécules hydrophobes ainsi que les petites molécules neutres comme les molécules d'eau ou d'éthanol peuvent diffuser à travers la membrane. La membrane plasmique est en revanche hautement imperméable aux molécules chargées. Par exemple, les ions issus de la dissociation des molécules d'eau ( $H^+ + HO^-$ ) ne sont pas capables de diffuser à travers cette membrane malgré leur petite taille. Cette perméabilité relative joue un rôle primordial dans l'établissement de la force proton motrice. Pour les molécules neutres, la vitesse de diffusion est directement reliée à la différence de concentration entre les milieux extra et intracellulaires. Néanmoins, en contexte de croissance, la diffusion des métabolites est un phénomène trop lent pour assurer le bon développement bactérien. Des

acteurs protéiques vont donc intervenir pour accélérer cet import. Ces protéines (ou complexes protéiques) possèdent au moins un domaine transmembranaire ainsi que des domaines intra et extra-cytoplasmiques. Ces transporteurs sont cruciaux pour la cellule et peuvent assurer plusieurs rôles physiologiques comme l'import de nutriments mais également l'expulsion de molécules toxiques ou encore la sécrétion dans le milieu extérieur de molécules d'intérêt. Les transporteurs sont classés en plusieurs grandes familles selon des critères phylogénétiques et fonctionnels (<sup>74,75</sup>). Cette classification peut être simplifiée en considérant 4 grands types de transporteurs : les transporteurs passifs, les transporteurs actifs primaires et secondaires et enfin les transporteurs de groupe. Nous nous concentrerons dans les parties suivantes sur l'import de nutriments particulièrement des substrats carbonés en illustrant chaque cas par un exemple décrit chez *E. faecalis*.

### A. Les facilitateurs

Les facilitateurs apparaissent comme les transporteurs les plus simples. Ils sont généralement constitués d'une seule protéine comportant 12 domaines transmembranaires (TMD pour TransMembrane Domain) formés de structures en hélices  $\alpha$  hydrophobes reliés par des boucles hydrophiles retrouvées au niveau intracellulaire ou extracellulaire (<sup>76</sup>). Ces TMD s'organisent au niveau tridimensionnel pour former un canal aqueux en leur centre et présentent en règle générale une spécificité de substrat assez large (<sup>74,75</sup>). Ces facilitateurs sont impliqués dans le transport de composés comme des hexoses, des polyols ou des ions (<sup>76</sup>). Les substrats traversent la membrane selon le phénomène de diffusion explicité ci-dessus sans modification chimique. La vitesse de transport n'est donc liée qu'aux concentrations intra et extracellulaires des molécules transportées jusqu'à atteindre l'équilibre osmotique pour les composés neutres. Concernant les composés chargés (notamment des ions), c'est le gradient électrochimique qui dictera le sens majoritaire ainsi que la vitesse de transport (<sup>77</sup>). La présence de ces transporteurs augmente grandement la vitesse d'entrée de certaines molécules dans la cellule et permettent ainsi de supporter la croissance bactérienne.

C'est par exemple le cas du facilitateur du glycérol GlpF chez *E. faecalis* (<sup>49,61</sup>). Etant liposoluble, le glycérol peut entrer dans la cellule par diffusion à travers la membrane plasmique mais il peut également entrer de manière plus rapide par ce transporteur sans subir de modification chimique et pourra ensuite être pris en charge pour produire de l'énergie par catabolisme ou être utilisé dans des voies d'anabolisme (**Figure 6**).



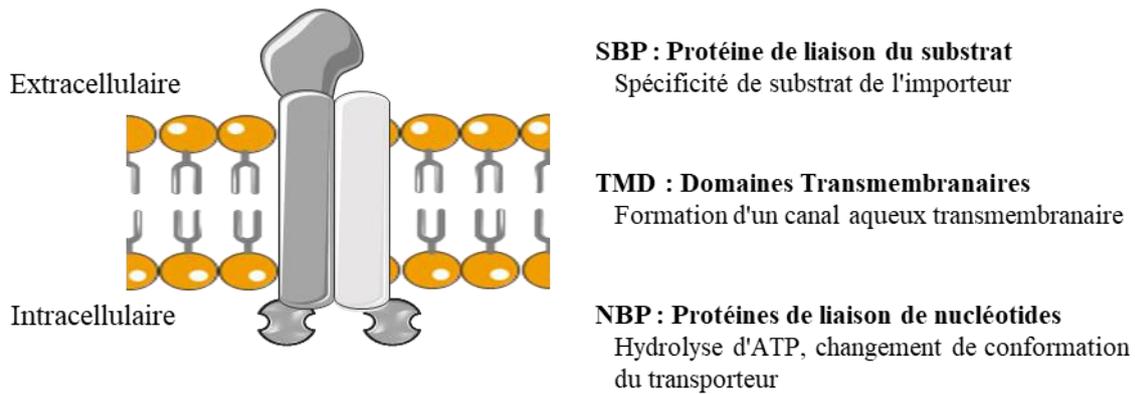
**Figure 6 :** *GlpF, facilitateur du glycérol chez E. faecalis*

*Les flèches noires pleines représentent un transport facilité et efficace. La flèche en pointillé représente la diffusion passive lente du glycérol à travers la membrane plasmique.*

## B. Les transporteurs actifs primaires : les transporteurs ABC

Ces transporteurs ABC (ATP Binding Cassette) sont beaucoup plus gros que les facilitateurs décrits précédemment et forment généralement de larges complexes protéiques. Leur fonction la plus connue est le transport de molécules à travers la membrane plasmique contre leur gradient de concentration en utilisant une source primaire d'énergie par hydrolyse d'une liaison riche en énergie au sein de la molécule d'ATP <sup>(74,75)</sup>. Les transporteurs ABC peuvent être fonctionnellement classés en deux catégories selon le sens du transport des molécules. Les importeurs ABC permettent l'import de nutriments au sein de la cellule et ne sont retrouvés que chez les procaryotes <sup>(78)</sup>. Les exporteurs ABC vont quant à eux permettre un export de molécules dans le milieu extérieur. Ces transporteurs sont retrouvés dans tout le règne du vivant et peuvent jouer un rôle dans la sécrétion de molécules toxiques comme des toxines ou encore dans la résistance à des molécules nocives pour la cellule comme des antibiotiques <sup>(79)</sup>.

En raison de la place primordiale qu'ils occupent au sein du métabolisme bactérien mais également de leur implication dans la présente étude, seuls les importeurs ABC seront traités par la suite. Les importeurs ABC sont quasi systématiquement organisés selon le même schéma : deux protéines de liaison de nucléotides (NBP pour "Nucleotide Binding Protein), deux protéines transmembranaires (TMD pour TransMembrane Domain) et une ou plusieurs protéines de liaison du substrat (SBP pour Solute Binding Protein) (**Figure 7**).



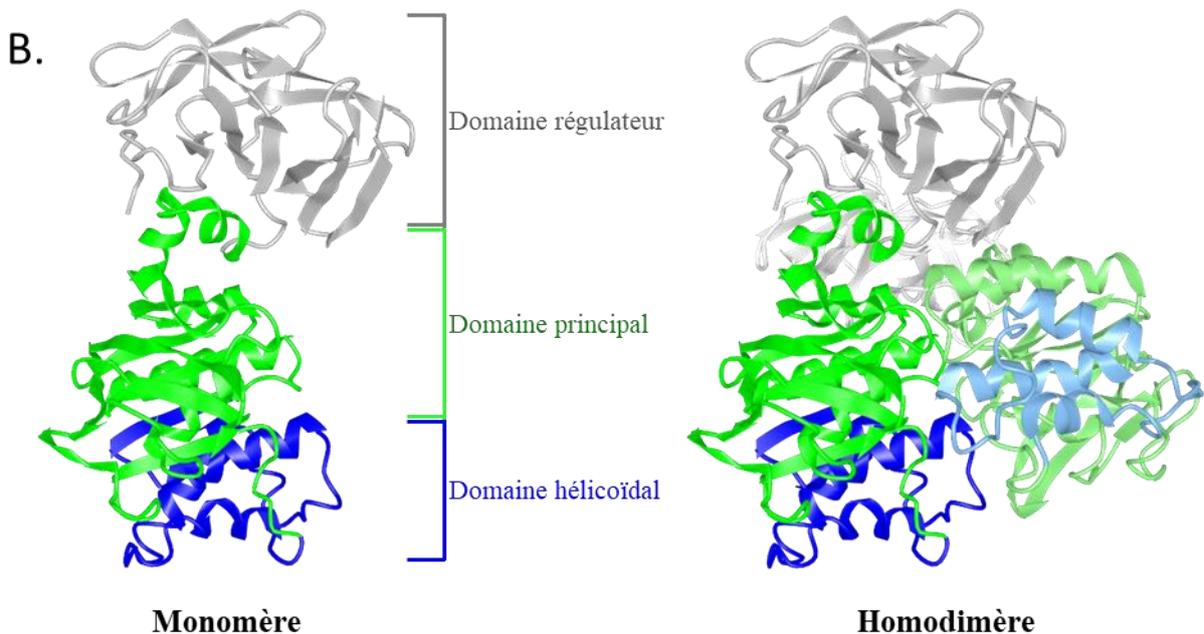
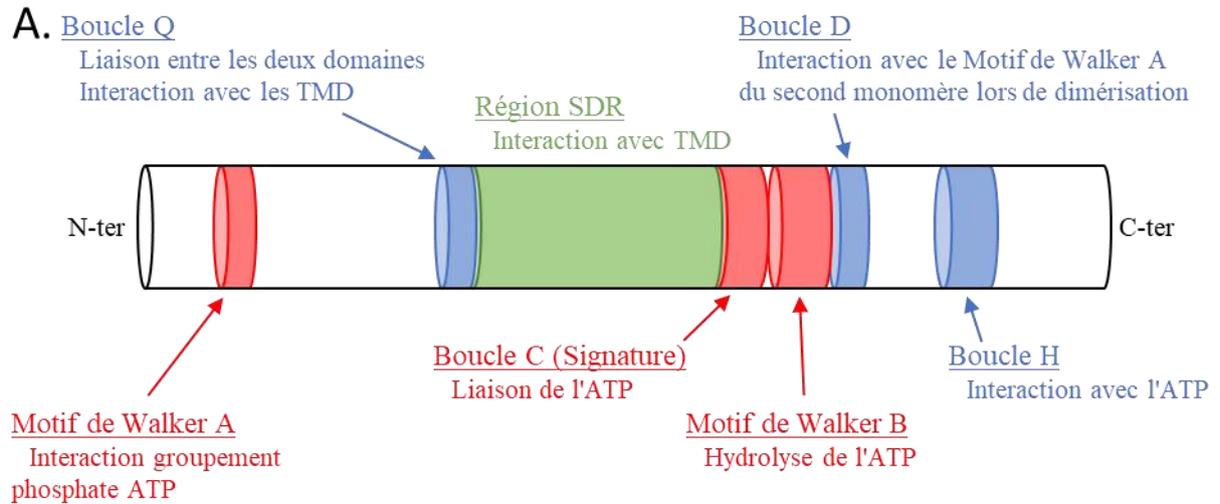
**Figure 7 :** Structure d'un importeur ABC

La fonction principale de chaque constituant est résumée en dessous du nom de ce composant.

### 1. Les protéines de liaison de nucléotides NBP.

Les protéines de liaison de nucléotides NBP ou encore ABC sont capables de lier et d'hydrolyser des molécules d'ATP et sont responsables du nom donné aux transporteurs ABC. Ces protéines à activité ATPase représentent une des plus grandes superfamilles de protéines (<sup>79,80</sup>). Dans le cas des transporteurs ABC, ces protéines NBP apparaissent conservées et sont systématiquement constituées de deux domaines structuraux. Le domaine principal représente la majeure partie de la protéine et contient la plupart des régions d'intérêt détaillées ci-après et représentées en **Figure 8.A**. Ce domaine principal contient notamment deux motifs très conservés : les motifs de Walker A et B. Le motif de Walker A (aussi appelé Boucle P) permet une interaction de la protéine avec les groupements phosphates de l'ATP via un motif GxxGxGKS/T (où x peut être n'importe quel acide aminé) (<sup>81</sup>). Le motif de Walker B contient quant à lui, quatre acides aminés hydrophobes suivis d'un acide aminé acide et va intervenir dans l'hydrolyse de la liaison  $\gamma$ -phosphoanhydride, riche en énergie de la molécule d'ATP (<sup>78</sup>). Ce domaine principal présente également d'autres régions plus variables contenant à chaque fois un acide aminé hautement conservé. Ces régions variables assurent principalement des rôles dans la structure quaternaire des protéines.

Le second domaine est nommé domaine hélicoïdal en raison de la forte proportion d'hélices  $\alpha$  dans son organisation tridimensionnelle. Il contient une unique région très conservée spécifique aux molécules ABC : la boucle C (aussi appelé motif signature). Cette dernière contient un motif LSGGQ et intervient dans l'interaction avec la molécule d'ATP (<sup>79</sup>). Les deux domaines présentés précédemment sont reliés entre eux par la boucle Q qui joue aussi un rôle dans l'interaction avec les protéines à domaines transmembranaires des importeurs ABC. La séquence entre les boucles Q et C forme la région SDR (pour Structurally Diverse Region), cette région variable va jouer un rôle crucial dans l'interaction avec les partenaires



**Figure 8 : Structure des protéines NBP**

**A :** Schéma de la structure primaire "type" d'une protéine NBP. Les régions d'intérêt sont représentées par des cylindres.

Les régions hautement conservées sont symbolisées en rouge, les domaines variables présentant un résidu conservé en bleu et la région SDR très variable en vert.

Une brève description du rôle de chaque région est inscrite en dessous de son nom.

**B.** Structure tridimensionnelle de la protéine MalK d'E. coli disponible sur la banque de structure tridimensionnelle de macromolécules MMDB. ID : 28818 <sup>(82)</sup>. Le domaine principal est représenté en vert et le domaine hélicoïdal en bleu. Un domaine régulateur (représenté en gris) est également présent dans cette protéine mais n'est pas classiquement retrouvé chez les NBP. Dans la forme dimérisée, le second monomère répond au même code avec des couleurs plus claires.

D'après <sup>(79)</sup>

transmembranaires des transporteurs (<sup>79</sup>). Dans le cas des transporteurs ABC, ces protéines NBP sont systématiquement organisées en dimères lors de l'hydrolyse d'ATP, cette dimérisation est assurée par l'interaction de la boucle D avec le motif de Walker A de l'autre monomère (**Figure 8.B**).

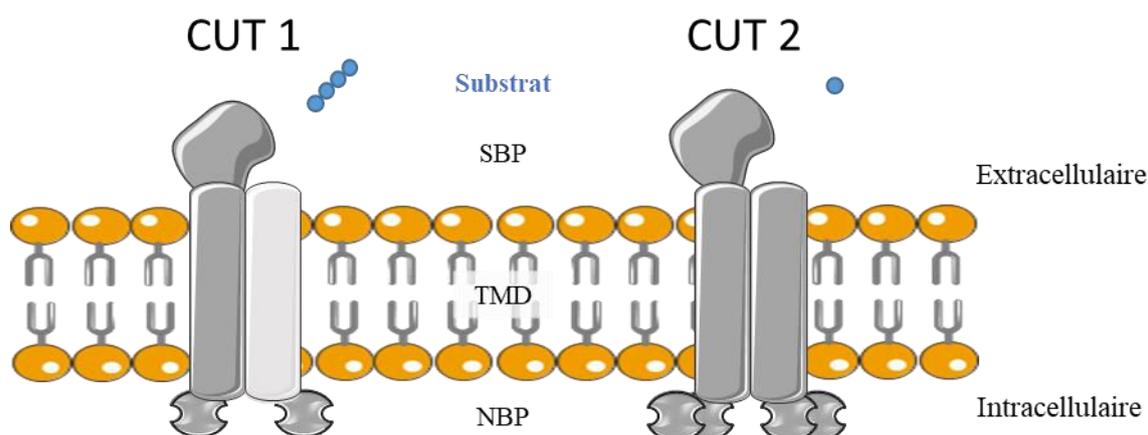
Comme décrit précédemment, les protéines NBP s'organisent en dimère lors de l'hydrolyse d'ATP qui va fournir de l'énergie au système. Le transporteur va alors changer de conformation pour permettre l'import de nutriments contre le gradient de concentration. Il est important de noter que la fixation d'ATP est indispensable pour la dimérisation des NBP et que cette dimérisation est elle-même nécessaire pour l'hydrolyse d'ATP. Ces deux conditions permettront de transformer l'hydrolyse d'ATP en travail mécanique (<sup>78</sup>). L'hydrolyse d'ATP est stimulée par la fixation du ligand sur la protéine de liaison du ligand (SBP pour Solute Binding Protein) dans le cas des importeurs ABC.

## **2. Les protéines de liaison du ligand SBP**

Les importeurs ABC peuvent être impliqués dans l'import de divers substrats du milieu extérieur vers le milieu intracellulaire comme des ions, des vitamines, des peptides ou encore des sucres. Dans le cas des importeurs ABC impliqués dans l'import des sucres, on distingue deux sous-familles, CUT1 et CUT2 (CUT pour Carbohydrate Uptake Transporter). Leur organisation ainsi que leur spécificité de substrat varie mais le principe général d'import reste le même entre ces deux sous-familles (<sup>83</sup>). Les caractéristiques sont regroupées dans le **Tableau 3** et illustrées dans la **Figure 9**.

**Tableau 3** : Caractéristiques des importeurs ABC CUT1 et CUT2

	CUT1	CUT2
<b>Protéine ABC (NBP)</b>	<ul style="list-style-type: none"> <li>• 1 domaine de liaison de nucléotides</li> <li>• Homodimère</li> </ul>	<ul style="list-style-type: none"> <li>• 2 domaines de liaison de nucléotides</li> <li>• Homodimère</li> </ul>
<b>Protéines à Domaines Transmembranaires (TMD)</b>	<ul style="list-style-type: none"> <li>• Hétérodimère</li> </ul>	<ul style="list-style-type: none"> <li>• Homodimère</li> </ul>
<b>Spécificité de substrat</b>	<ul style="list-style-type: none"> <li>• Disaccharides</li> <li>• Polysaccharides</li> <li>• Polyols</li> </ul>	<ul style="list-style-type: none"> <li>• Monosaccharides</li> </ul>



**Figure 9** : Organisation des deux sous-familles CUT1 et CUT2 d'importeurs ABC. Ces transporteurs sont impliqués dans l'import actif de glucides. D'après <sup>(83)</sup>.

La spécificité de substrat est apportée par les protéines SBP, on en retrouve quasi systématiquement au moins une par importeur ABC. Chez les bactéries négatives à la coloration de Gram, ces SBP sont libérées dans le périplasma. Chez les bactéries Gram-positives, ce sont généralement des lipoprotéines ancrées à la membrane et exposées dans le milieu extracellulaire ou des protéines fusionnées aux domaines transmembranaires <sup>(84,85)</sup>. Les protéines SBP monomériques interagissent avec leur ligand par des liaisons faibles la plupart du temps de types liaisons hydrogènes. Leur sélectivité pour leur ligand est variable, certaines SBP peuvent présenter des affinités pour plusieurs ligands apparentés ou au contraire être hautement spécifiques d'un seul ligand <sup>(79)</sup>. Ces SBP présentent généralement des affinités élevées de l'ordre du  $\mu\text{M}$  pour leur(s) ligand(s) et augmentent grandement l'efficacité de transport des importeurs ABC même si certains de ces systèmes ont été décrits comme partiellement fonctionnels même en l'absence de ces SBP <sup>(86,87)</sup>. Certains importeurs ABC sont capables d'importer plusieurs ligands différents mais généralement chimiquement apparentés. Cette pluralité de substrats peut provenir de la flexibilité de spécificité des SBP comme annoncé précédemment ou de la présence de plusieurs SBP différentes pour un même importeur ABC,

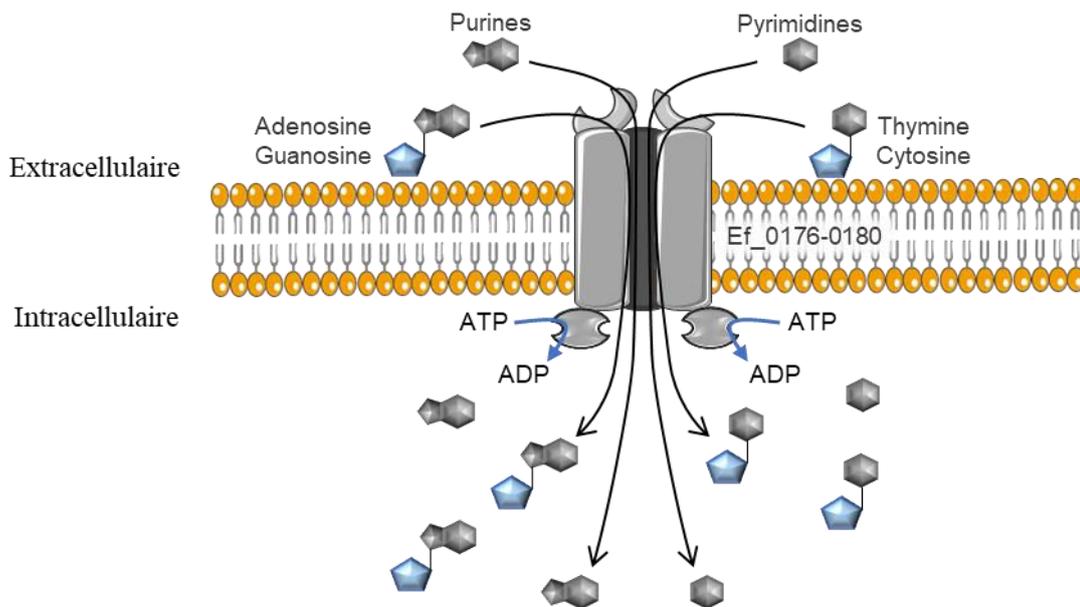
chacune avec une spécificité particulière pour son substrat <sup>(85)</sup>. La liaison du ligand sur les SBP représente la première étape de l'import de molécules. Ce complexe ligand-protéine va ensuite interagir avec les protéines à TMD et stimuler l'activité ATPase au niveau intracellulaire. Cela va provoquer un changement de conformation du système et ainsi permettre l'import de substrat. La transduction du signal apparaît être initiée par l'attachement des SBP au niveau du transporteur et transmise par l'intermédiaire des protéines à TMD <sup>(85)</sup>.

### **3. Les protéines à domaines transmembranaires (TMD)**

Ces protéines constituent une famille de protéines très variables contrairement aux protéines NBP très conservées. Malgré leurs disparités au niveau de leur séquence primaire, ces protéines à TMD présentent une forte conservation de leur structure tridimensionnelle. Elles contiennent systématiquement des domaines hydrophobes formés d'hélices  $\alpha$  permettant d'insérer ces protéines au sein de la membrane plasmique des cellules <sup>(81)</sup>. Ces protéines présentent également entre deux domaines transmembranaires une hélice de couplage exposées du côté intracellulaire contenant un motif EAA très conservé. Cette boucle est impliquée à la fois dans l'interaction des protéines à TMD avec les NBP mais également dans le couplage de l'hydrolyse d'ATP avec le changement de conformation des transporteurs <sup>(79)</sup>. Deux protéines à TMD sont nécessaires pour former un importeur ABC fonctionnel. Ce complexe peut être formé d'un homodimère ou d'un hétérodimère selon la famille du transporteur comme explicité ci-dessus. Les dimères de protéines à TMD forment le canal aqueux du transporteur et vont être responsables du transport du substrat à travers la membrane plasmique. Ces protéines semblent présenter une affinité pour les molécules transportées. Cette affinité apparaît toutefois beaucoup plus faible que leur SBP et les importeurs ABC perdent grandement en efficacité en cas d'absence de SBP <sup>(86,87)</sup>. Lorsque le substrat est absent, le transporteur est "en attente". Quand le ligand est présent, il peut se lier à la SBP et le complexe SBP-ligand va venir se fixer au niveau du pore formé par les protéines à TMD. Comme expliquée précédemment, cette fixation va induire une transduction de signal favorisant l'hydrolyse de l'ATP par les protéines NBP. Cette hydrolyse d'ATP s'accompagne de la séparation des NBP. Ces NBP étant liées aux protéines à TMD, cela va induire un changement de conformation du canal et ainsi convertir l'énergie contenue dans la liaison riche de l'ATP en travail mécanique permettant au complexe de transporter son substrat contre son gradient de concentration <sup>(78)</sup>.

On peut illustrer ces importeurs ABC chez *E. faecalis* avec le système composé des gènes *ef\_0176-0180* chez la souche V583 <sup>(88)</sup>, les gènes *ef\_0179-0180* représentent les deux protéines transmembranaires formant le canal. Le gène *ef\_0178* est prédit pour coder une protéine à

activité ATPase et fournit donc probablement l'énergie au système. Ce transporteur ABC possède deux SBP codées par les gènes *ef\_0176* et *ef\_0177* qui présentent toutes deux des spécificités de substrat proches mais légèrement différentes pour les ribonucléosides ainsi que les bases puriques et pyrimidiques. Ce transporteur permet par conséquent l'import de tous les ribonucléosides et bases azotées qui pourront être utilisés comme source de carbone ou d'azote ou comme précurseurs d'acides nucléiques pour la cellule <sup>(89)</sup> (**Figure 10**).



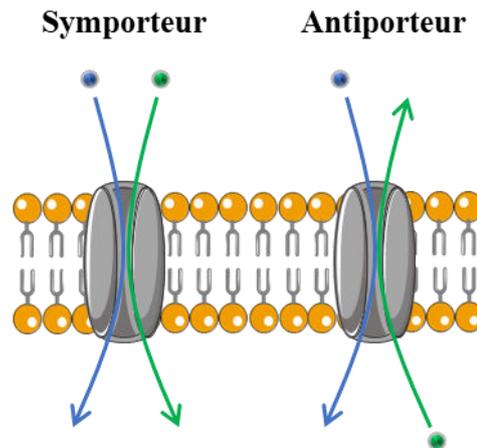
**Figure 10** : Importeur ABC *EF\_0176-0180* chez *E. faecalis*

Schéma de l'import des ribonucléosides et des bases azotées chez *E. faecalis* grâce au transporteur ABC *Ef\_0176-0180*.

Les flèches noires pleines représentent des transports actifs contre le gradient de concentration du milieu extérieur vers le milieu intracellulaire. D'après <sup>88</sup>.

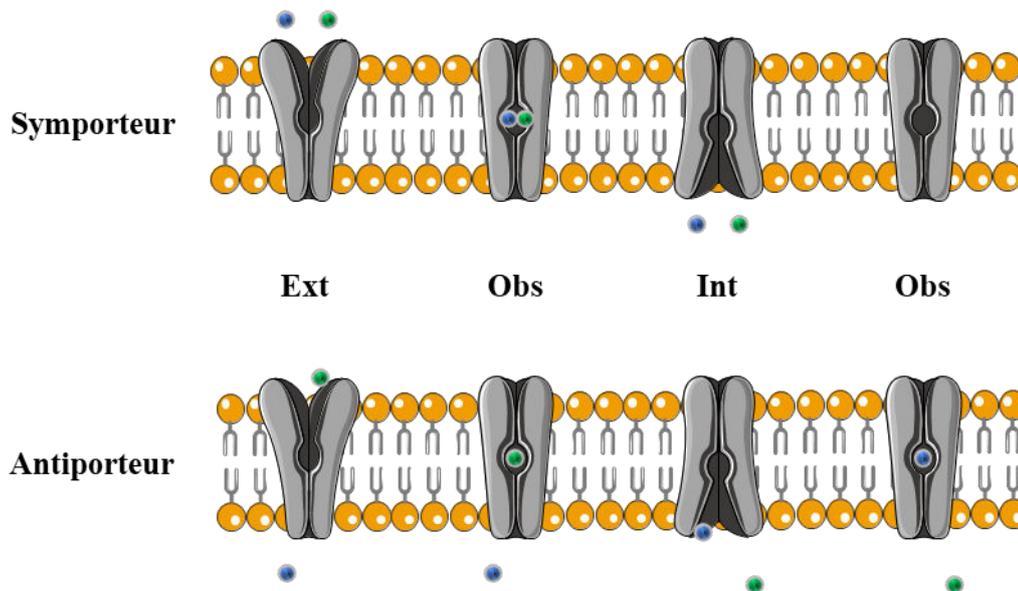
### C. Les transporteurs actifs secondaires

Les transporteurs actifs secondaires nécessitent également une source d'énergie pour transporter leur substrat contre le gradient de concentration. Cependant, contrairement aux transporteurs actifs primaires, cette énergie n'est pas apportée par l'hydrolyse d'une liaison riche en énergie au sein d'une molécule organique mais généralement par un gradient électrochimique d'ions instauré de part et d'autre de la membrane plasmique <sup>(74)</sup>. Ce gradient électrochimique peut être instauré par d'autres systèmes cellulaires qui vont eux nécessiter de l'énergie sous forme d'ATP. Ces transporteurs peuvent être classés en deux grandes catégories indépendamment de leur superfamille selon le sens du transport des deux espèces chimiques. Si les deux éléments sont transportés dans le même sens, on parlera de symporteurs alors que si les deux espèces chimiques sont transportées dans le sens opposé, on désignera ces transporteurs comme des antiporteurs (**Figure 11**).



**Figure 11** : Fonctionnement des transporteurs actifs secondaires comme symporteurs et antiporteurs  
 Les billes bleues et vertes symbolisent les ligands co-transportés. Le transporteur de gauche fonctionne comme un symporteur et celui de droite comme un antiporteur.

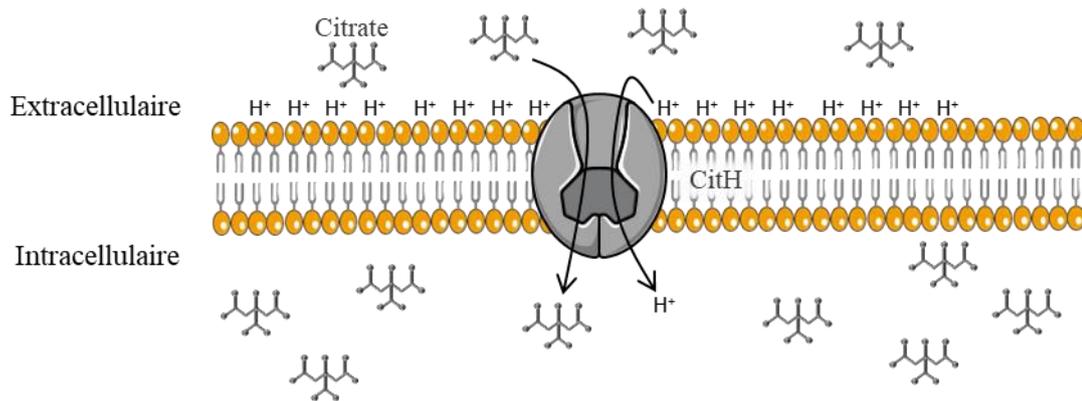
Ces transporteurs actifs secondaires sont assez proches au niveau conceptuel des facilitateurs passifs décrits précédemment. Ils catalysent en effet des réactions entièrement réversibles dont le sens est dicté par la somme des gradients électrochimiques des espèces transportées<sup>(90)</sup>. Dans un contexte d'import de nutriments qui nous intéresse ici, les molécules d'intérêt peuvent être importées contre leur gradient de concentration grâce au gradient électrochimique d'un ion co-transporté et ainsi être concentrées du côté intracellulaire en vue d'être métabolisées. Ces transporteurs possèdent trois états majeurs : ouvert au niveau du milieu extracellulaire, "obstrué" et ouvert au niveau du milieu intracellulaire. Dans un contexte physiologique, un sens de transport global est privilégié en raison des gradients électrochimiques, l'état obstrué d'un symporteur contiendra alors successivement les deux espèces chimiques puis aucune alors que concernant les antiporteurs, l'état obstrué contiendra les molécules seules alternativement<sup>(90,91)</sup> (**Figure 12**).



**Figure 12** : Différents états adoptés par les transporteurs actifs secondaires  
 Les billes vertes et bleues symbolisent les ligands co-transportés.  
 Le transporteur du haut fonctionne comme un symporteur et celui du bas comme un antiporteur.  
 Les différents états sont indiqués entre les deux transporteurs par un code trois lettres. Ext : ouvert vers le milieu extracellulaire, Obs : obstrué, Int : Ouvert vers le milieu intracellulaire.

Concernant leur structure, les transporteurs actifs secondaires représentent des structures beaucoup plus réduites que les ABC transporteurs détaillés dans la partie précédente. Ils sont généralement codés par un seul gène mais peuvent parfois s'organiser en oligomères <sup>(92)</sup>. La chaîne polypeptidique contient systématiquement des domaines transmembranaires (TMD) riches en hélices  $\alpha$  permettant aux protéines de s'insérer dans la membrane plasmique. Le nombre ainsi que l'organisation de ces domaines transmembranaires est variable selon les familles de transporteurs.

Chez *E. faecalis*, on peut citer en exemple le transporteur CitH du citrate <sup>(93)</sup> (**Figure 13**), ce transporteur de la famille 2HCT (pour transporteurs du 2-hydroxycarboxylate) permet l'import du citrate complexé avec un métal divalent, préférentiellement le  $\text{Ca}^{2+}$ . Ce complexe protéique fonctionne comme un symporteur en utilisant la force proton motrice pour coupler l'import du Citrate- $\text{Ca}^{2+}$  avec un proton  $\text{H}^+$ . Aucune structure tridimensionnelle n'est disponible pour ce symporteur chez *E. faecalis* mais le transporteur CitS de *Klebsiella pneumoniae*, pour lequel CitH présente plus de 50% d'identité, est lui très caractérisé. Il couple l'import du citrate avec l'ion  $\text{Na}^+$  <sup>(94,95)</sup>, on peut supposer des modes de fonctionnement similaires chez *E. faecalis*.



**Figure 13** : CitH, importeur actif secondaire du citrate chez *E. faecalis*.

Les flèches noires pleines représentent un transport actif contre le gradient de concentration du milieu extérieur vers le milieu intracellulaire. Un proton  $H^+$  est importé pour chaque molécule de citrate importée.

## D. Les transports par translocation de groupe : les Systèmes PhosphoTransférases PEP dépendent (PTS)

La dernière famille qui va nous intéresser ici est classée sous la détermination de "transports par translocation de groupe dépendant du PEP" également appelés Systèmes PhosphoTransférases (PTS pour PhosphoTransferase Systems) <sup>(74)</sup>. Ces transporteurs ont à ce jour été retrouvés exclusivement chez des procaryotes, principalement chez des bactéries mais également chez quelques archaea <sup>(96-98)</sup>. Ces systèmes sont très largement répandus au sein des différents phyla bactériens et peuvent présenter des différences entre ces microorganismes. Nous nous concentrerons ici sur le fonctionnement des PTS chez les firmicutes. Ces transporteurs ont un rôle double dans la cellule permettant un import très efficace de leur substrat mais jouant également un rôle prépondérant dans la régulation de nombreuses fonctions cellulaires en étant au centre d'un type de régulation appelé la répression catabolique ou CCR (pour Carbon Catabolite Repression).

### 1. Les PTS comme transporteurs

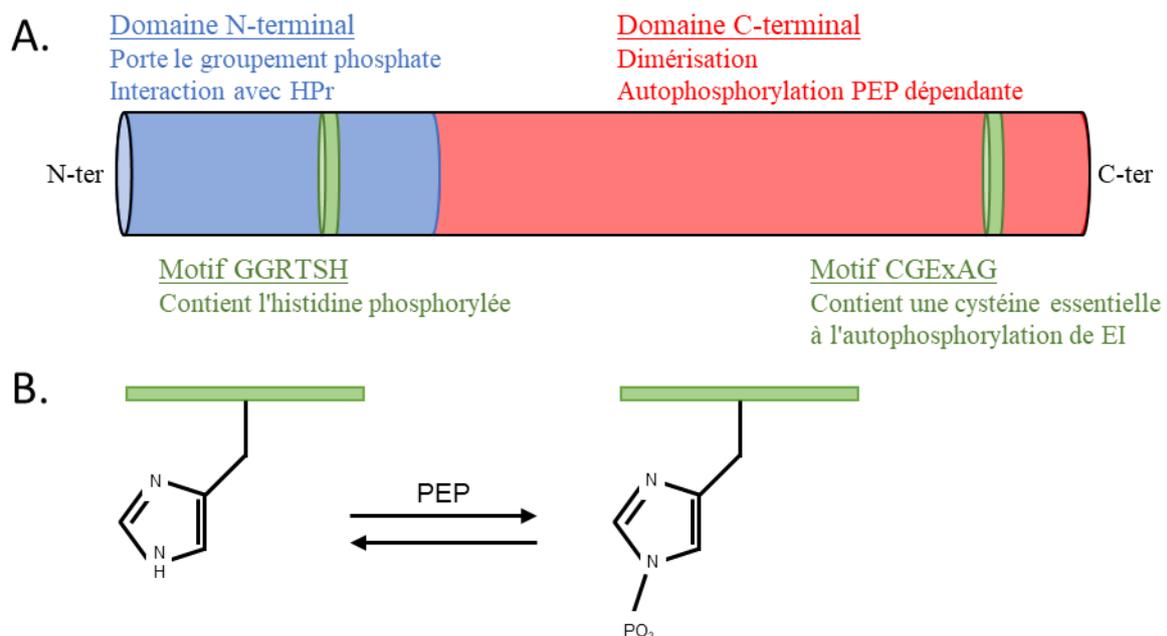
Ces systèmes permettent l'import de sucres ou de molécules dérivées de sucres grâce à l'énergie de la liaison riche du PEP. Cette énergie va être transmise via une cascade de phosphorylations réversibles entre plusieurs acteurs moléculaires <sup>(98)</sup>. Le substrat importé par ces systèmes sera phosphorylé durant le transport. La phosphorylation des sucres représentant généralement une étape préliminaire à leur catabolisme, l'import via les PTS permet un couplage du transport et de la phosphorylation en une étape sans utilisation ultérieure d'énergie. Ce couplage est dans un premier temps assuré par une cascade de transferts de groupements phosphates du PEP vers le substrat importé via plusieurs acteurs protéiques successifs : EI, HPr,

EIIA et EIIB<sup>(99)</sup>, le fonctionnement général des PTS est présenté dans la **Figure 14**.

Les deux premiers acteurs de la cascade de phosphorylation EI et HPr sont communs à tous les PTS.

a. L'enzyme EI

La protéine EI, codée par le gène ptsI, est le premier acteur de la cascade de phosphorylation des PTS<sup>(100)</sup>. Elle peut être séparée en deux domaines fonctionnels. Le domaine N-terminal permet le port du groupement phosphate et l'interaction avec HPr. Le second domaine, appelé C-terminal, permet l'autophosphorylation de EI ainsi que la formation de l'homodimère EI<sub>2</sub> (**Figure 15.A**). On retrouve systématiquement au niveau du domaine N-terminal une histidine conservée (Histidine 191 chez *E. faecalis*) qui va être l'acide aminé phosphorylé en position N-3 de l'anneau imidazole, cette histidine est généralement retrouvée à la fin d'un motif conservé GGRTSH<sup>(101,102)</sup> (**Figure 15.B**).



**Figure 15** : Structure de l'enzyme EI.

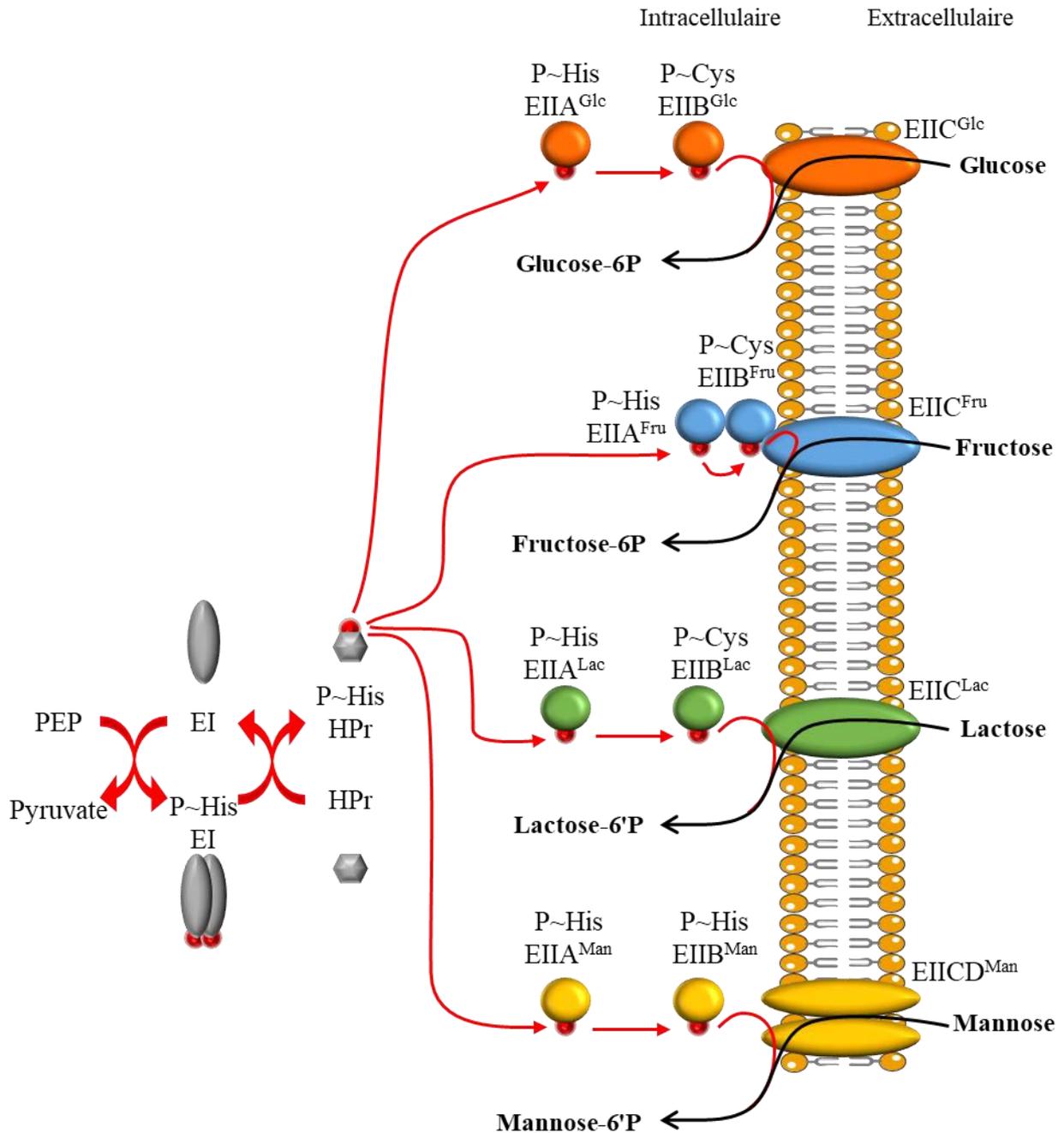
**A** : Schéma de la structure primaire de la protéine EI d'*E. faecalis*. Les domaines d'intérêt sont représentés par des cylindres. Le domaine N-terminal est symbolisé en bleu et le domaine C-terminal en rouge.

Les motifs conservés au sein de ces domaines sont représentés en vert. Le domaine N-terminal comporte environ un tiers des acides aminés totaux de la protéine EI du côté N-terminal. Une brève description du rôle de chaque domaine est inscrite en dessous de son nom.

**B.** Schéma de la phosphorylation de l'histidine 191 de la protéine EI d'*E. faecalis*. Le squelette de la protéine est représenté par un rectangle vert et seule la chaîne latérale de l'histidine 191 est représentée en détail.

D'après<sup>(99,102,103)</sup>.

L'autophosphorylation de la protéine EI nécessite une molécule de PEP comme donneur de



**Figure 14** : Schéma du fonctionnement général des PTS

Les transferts de groupements phosphates sont représentés par des flèches rouges. Les transferts entre EI-HPr et entre EIIA-EIIB sont réversibles mais ne sont pas schématisés pour plus de clarté. Seules les formes phosphorylées des domaines EII sont représentées pour ne pas surcharger le schéma.

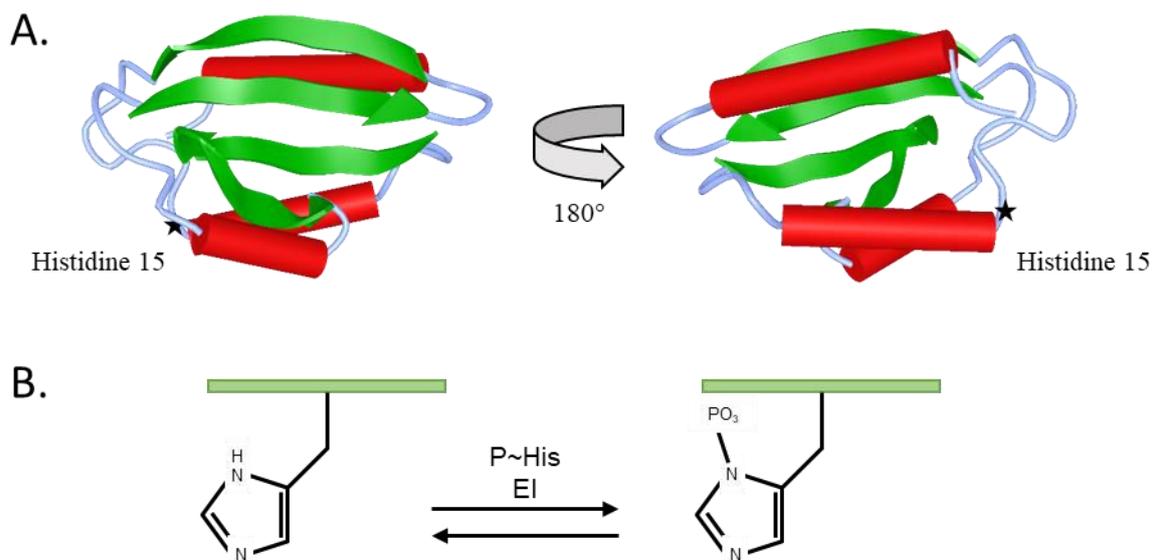
Les transports de substrats sont schématisés par des flèches noires.

Les protéines EI et HPr, communes à tous les PTS, sont schématisées en gris, les domaines EII sont représentés avec un code couleur correspondant au système EII entier.

groupement phosphate ainsi que la présence de cations divalents, préférentiellement le  $Mg^{2+}$  (<sup>99,101</sup>). Une fois phosphorylée, l'enzyme EI va se complexer pour former un homodimère avant de transférer ce groupement phosphate au second acteur cytosolique commun des PTS : la protéine HPr.

### b. La protéine HPr

La protéine HPr est centrale dans le processus de transfert de groupement phosphate lors de l'import de substrats par un PTS. Cette petite protéine d'environ 90 acides aminés codée par le gène *ptsH* possède une région dont la séquence est conservée autour d'une histidine en position 15 (<sup>102</sup>). La structure tridimensionnelle de cette protéine est également très conservée. Elle est constituée de 4 feuillets  $\beta$  antiparallèles formant une structure  $\beta$ -sandwich et de 3 hélices  $\alpha$  dont 2 longues et une plus courte. L'histidine en position 15 qui sera phosphorylée en position N1 de l'anneau imidazole se trouve immédiatement en amont de l'extrémité N-terminale de la première hélice  $\alpha$  dans la séquence protéique (**Figure 16**) (<sup>99,101</sup>).



**Figure 16** : Structure de la protéine HPr

**A.** : Schéma de la structure tridimensionnelle de la protéine HPr chez *E. coli* disponible sur la banque de structure tridimensionnelle de macromolécules MMDB. ID : 56311 (<sup>104</sup>).

Les feuillets  $\beta$  sont représentés par des flèches vertes et les hélices  $\alpha$  par des cylindres rouges.

La position de l'histidine 15 est indiquée par une étoile noire.

**B.** Schéma de la phosphorylation de l'histidine 15. Le squelette de la protéine est représenté par un rectangle vert et seule la chaîne latérale de l'histidine 15 est représentée en détail.

D'après (<sup>105</sup>).

Cette protéine HPr se retrouve au niveau d'un nœud central dans la cascade de transfert de groupement phosphate. Elle va ensuite distribuer ce groupement phosphate présent sur l'histidine en position 15 à l'une des protéines EIIA d'un des PTS présents dans la cellule.

### c. Les protéines EII

Les acteurs suivants sont représentés par les enzymes EII et sont spécifiques à chaque système PTS. Sept grandes familles de PTS peuvent être distinguées chez les firmicutes en fonction de la structure des protéines EII et du substrat transporté. Ces familles sont indiquées par un code à 3 lettres inscrit en exposant. Certaines caractéristiques de ces familles sont indiquées dans le **Tableau 4** <sup>(74)</sup>.

**Tableau 4** : Caractéristiques des grandes familles de PTS

Famille de protéine EII	Exemples de substrats <sup>a</sup>	Organisation possible des différents domaines EII <sup>b</sup>
Glucose/Glucoside	Glucose, NAG <sup>c</sup> , $\alpha$ et $\beta$ -glucosides (maltose, cellobiose, ...), saccharose, ANAM <sup>d</sup>	IIA IIB IIC IIAB IIBC IIABC
Fructose/Mannitol	Fructose, Mannose, Mannitol	IIA IIB IIC IIBC IIABC
Lactose/Diacetylchitobiose/ $\beta$ -glucosides	Lactose, $\beta$ -glucosides (cellobiose, ...), N-N'-diacetylchitobiose	IIA IIB IIC IIBC
Glucitol	Glucitol	IIA IIC IIBC
Galactilol	Galactilol, D-arabitol	IIA IIB IIC
Mannose/Fructose/Sorbose	Glucose, Mannose, Fructose, Sorbose, NAG, Galactosamine	IIA IIB IIC IID IIAB IICD
Ascorbate	L-ascorbate	IIA IIB IIC IIAB IIBC

<sup>a</sup> Ces exemples ne représentent pas une liste exhaustive des substrats possibles des différents PTS concernés.

<sup>b</sup> L'ordre des différents domaines au sein de la séquence protéique n'est pas pris en compte. Ex : EIIAB signifie que l'on peut retrouver une organisation EIIAB et/ou EIIBA.

<sup>c</sup> NAG : N-acetylglucosamine

<sup>d</sup> ANAM = Acide N-acetylmuramique

D'après <sup>(74,106)</sup>.

Ces protéines EII présentent des organisations tridimensionnelles différentes entre les familles, on retrouve cependant 3 grands domaines fonctionnels dans chaque cas (4 pour les EII<sup>Man</sup>) : 2 domaines cytoplasmiques EIIA et EIIB et un domaine transmembranaire EIIC (et EIID<sup>Man</sup>) <sup>(99)</sup>. Ces domaines peuvent provenir de différentes protéines ou être présents au sein d'une même chaîne polypeptidique <sup>(106)</sup>.

Le domaine EIIC (et EIID pour les PTS<sup>Man</sup>) sera le seul acteur de la cascade de phosphorylation à ne pas porter de groupement phosphate. Il est responsable de la spécificité ainsi que de la translocation du substrat, le domaine EIIC va lier ce dernier et le rapprocher de la partie cytoplasmique. Le domaine EIIB phosphorylé va ensuite transférer son groupement phosphate au substrat une fois ce dernier lié au domaine EIIC. La molécule importée présentera alors moins d'affinité pour le domaine EIIC et va être libérée dans le cytosol. Cette

phosphorylation va également augmenter la polarité du substrat et ainsi prévenir sa diffusion à travers la membrane plasmique (<sup>99</sup>). Les domaines EIIB sont généralement phosphorylés au niveau d'une cystéine sauf les EIIB<sup>Man</sup> qui sont phosphorylés sur l'atome N1 de l'anneau imidazole d'une histidine de la même manière que P~His-HPr. Les domaines EIIB peuvent être cytoplasmiques ou ancrés à la membrane et exposés du côté intracellulaire lorsqu'ils sont liés au domaine EIIC dans la même protéine. Le groupement phosphate permettant la phosphorylation des EIIB provient du domaine EIIA phosphorylé sur une histidine au niveau N3 de l'anneau imidazole du PTS correspondant. Un domaine EIIA phosphorylé peut toutefois transférer son groupement phosphate à plusieurs EIIB différents d'une même famille même s'ils ne font pas partie du même système PTS. Ce type de transfert apparaît toutefois moins efficace comparé au transfert de groupement phosphate au sein du même PTS (<sup>107</sup>). En résumé, toutes ces cascades de phosphorylations permettent d'importer de manière très efficace des substrats spécifiques en utilisant comme donneur initial la molécule de PEP.

Le génome d'*E. faecalis* OG1RF contient, d'après son annotation, plus de 80 gènes codant pour des protéines EII entrant dans la composition de PTS dont 36 codant pour des protéines EIIC ou contenant un domaine EIIC. Ce nombre élevé indique que ce type de transporteur doit jouer un rôle crucial dans la physiologie de cette bactérie. D'une manière générale, les bactéries qui dépendent d'un métabolisme anaérobie efficace du glucose via la glycolyse apparaissent comme possédant de nombreux systèmes PTS (<sup>106</sup>). Malgré la bonne compréhension du fonctionnement de ces systèmes, la spécificité de substrat des domaines EIIC reste encore aujourd'hui assez difficile à prédire *in silico* et peu de ces systèmes ont été caractérisés de manière précise chez *E. faecalis*. Le système MalT, dont nous reparlerons dans une future partie de ce manuscrit, est cependant bien décrit (<sup>21,108</sup>). Dans ce système, les trois acteurs EII représentent trois domaines d'une même protéine codée par le gène *malT* et sont disposés selon une organisation EIICBA. La protéine résultante MalT contient par conséquent tous les domaines EII et permet chez cette bactérie l'import du maltose et du maltotriose. L'importance des PTS ne s'arrête cependant pas à un rôle de transporteurs de substrats, ces systèmes jouent en effet un rôle très important de régulation en réponse aux différents nutriments présents dans l'environnement direct de la cellule. Leur implication dans l'import de substrats pourrait même être une acquisition secondaire concernant principalement les entérobactéries et les firmicutes. On connaît en effet plusieurs phyla bactériens ne possédant aucun PTS complet où les protéines liées aux PTS chez ces microorganismes ne possèdent vraisemblablement que des rôles de régulation (<sup>109</sup>).

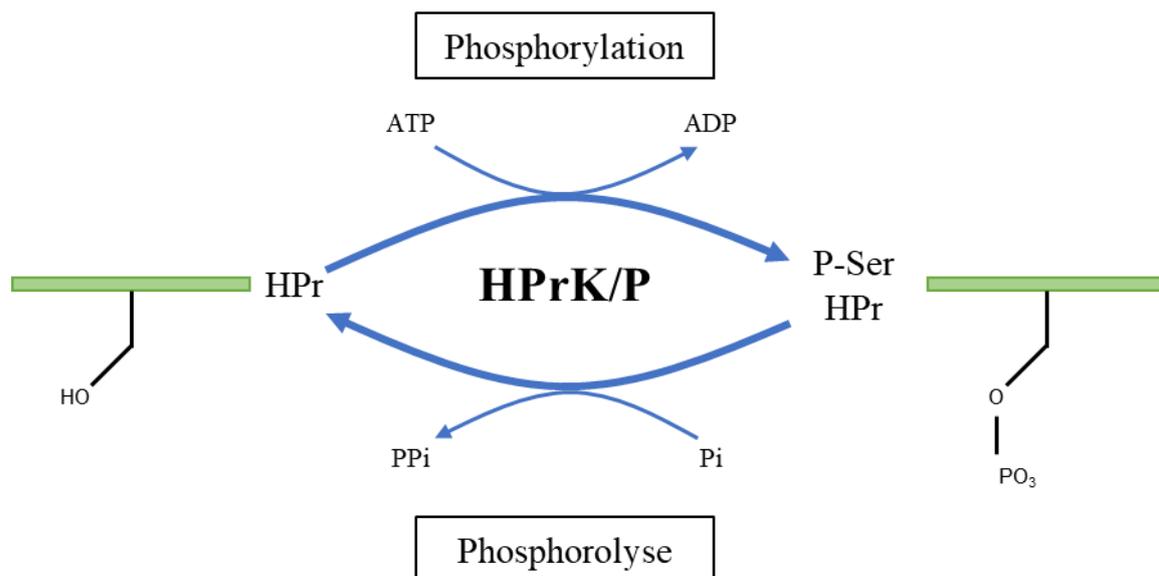
## 2. Le rôle des PTS dans la régulation du métabolisme

Lorsque les bactéries se retrouvent en présence de plusieurs sources de carbones différentes simultanément, elles vont prioritairement métaboliser les plus rentables pour elles. Ce phénomène de hiérarchisation du métabolisme des substrats carbonés est instauré notamment par l'arrêt de l'expression et/ ou du fonctionnement des protéines impliquées dans le métabolisme des sources de carbone moins rentables énergétiquement. Cette hiérarchisation, appelée la répression catabolique (ou CCR pour Carbon Catabolite Repression), permet ainsi à la cellule de prioriser la dégradation de sources de carbones plus rentables et ainsi éviter de gaspiller de l'énergie et du temps <sup>(98,110)</sup>. Ce type de régulation représente une stratégie très répandue au sein des bactéries même si les mécanismes mis en jeu peuvent grandement différer entre phyla bactériens. Chez les firmicutes, c'est la protéine HPr qui va jouer un rôle prépondérant. Comme nous l'avons vu précédemment, cette protéine joue un rôle central dans la cascade de transfert du groupement phosphate du PEP vers le substrat importé lorsqu'elle est phosphorylée sur une histidine conservée en position 15 (P~His-HPr) mais elle possède également un second site de phosphorylation primordial pour son activité de régulation.

### a. P-Ser-HPr et HPrK

La protéine HPr présente en effet un autre site de phosphorylation au niveau d'une sérine conservée généralement en position 46. Cette phosphorylation réversible est assurée par une enzyme bifonctionnelle Kinase/Phosphorylase HPrK/P <sup>(111,112)</sup>. Il est intéressant de noter que la protéine HPr de certaines entérobactéries comme *E. coli* présente une sérine en position 46, ces bactéries ne possèdent en revanche pas la kinase HPrK/P. Par conséquent, la protéine P-Ser-HPr ne peut pas être obtenue chez ces microorganismes. On retrouve en revanche la protéine HPrK/P chez la quasi-totalité des firmicutes connues <sup>(106)</sup>. Cette protéine d'environ 300 acides aminés est constituée de deux domaines : N et C-terminal, le second apparaît très conservé chez les firmicutes alors que le premier présente moins d'homologie entre les espèces. Cette protéine s'organise en homohexamère pouvant fixer jusqu'à 6 protéines HPr <sup>(113)</sup>. Le domaine N-terminal ne possède pas encore de fonction connue. La protéine HPrK/P de *Lactobacillus casei* reste par exemple fonctionnelle même après avoir entièrement délété ce domaine N-terminal. Le domaine C-terminal assure, lui, les deux activités de la protéine, kinase et phosphorylase <sup>(114)</sup>. Ce domaine C-terminal contient notamment un motif de walker A dont nous avons déjà parlé précédemment (cf partie II.B.1) lui permettant de fixer l'ATP. Comme nous venons de le voir, HPrK/P est capable de phosphoryler HPr sur le résidu seryl en position 46 et de retirer cette phosphorylation pour redonner la forme non phosphorylée de HPr.

Cependant ces deux modifications ne représentent pas les deux sens d'une même réaction mais bien deux réactions différentes catalysées par la même protéine. Des mutants pour HPrK/P ont d'ailleurs été obtenus chez *L. casei* où la protéine résultante HPrKV267F présente toujours une activité kinase inchangée mais n'est plus capable de catalyser la réaction de phosphorolyse <sup>(115)</sup>. L'activité kinase utilise l'ATP comme molécule donneuse de groupement phosphate et est stimulée par le premier intermédiaire de la glycolyse qu'est le FBP <sup>(116)</sup>. Cette dernière molécule n'est pas consommée dans la réaction et sert uniquement de stimulus pour réguler l'activité kinase de HPrK/P. De fortes concentrations intracellulaires en FBP sont atteintes lors d'un flux glycolytique important donc lors du catabolisme d'un substrat facilement métabolisable. La proportion de P-Ser-HPr va par conséquent augmenter lors de l'utilisation par la cellule de ces substrats rentables énergétiquement. L'activité de déphosphorylation de P-Ser-HPr est réalisée par phosphorolyse et non par hydrolyse, cette activité est stimulée par le phosphate inorganique (Pi), ce dernier étant un substrat consommé lors de cette réaction (Figure 17) <sup>(113,117)</sup>.



**Figure 17** : Schéma des deux activités kinase et phosphorylase de la protéine HPrK/P  
 La chaîne latérale de la sérine 46 est représentée pour les protéines HPr et P-Ser-HPr. Le squelette de la protéine est représenté par un rectangle vert et seule la chaîne latérale de la sérine 46 est représentée en détail.

La forme phosphorylée P-Ser-HPr n'est pas capable de transférer son groupement phosphate pour permettre l'import de sucres via les PTS. Cette phosphorylation en position 46 ralentit également la phosphorylation de HPr sur l'histidine en position 15 d'un facteur environ 500 *in vitro* <sup>(118)</sup>. Cette forme ne va donc pas intervenir dans le transport de substrats mais uniquement dans la régulation de l'activité et/ou de l'expression de différents acteurs

cellulaires directement ou indirectement. Les rapports des différentes formes de HPr vont ainsi servir pour la cellule de senseur de la disponibilité de sources de carbone dans son environnement.

#### b. CcpA (Catabolite Control Protein A)

Un des rôles de P-Ser-HPr est d'agir en tant que corépresseur essentiel du régulateur pléiotrope CcpA<sup>(119)</sup>. Ce régulateur de la famille LacI/GalR (cf partie III.A pour plus de détails sur cette famille de régulateur) intervient dans la régulation transcriptionnelle de nombreux gènes en fixant des séquences d'ADN spécifiques appelées boîtes *cre* (cis response element). En revanche, la protéine CcpA seule n'est pas capable de fixer efficacement l'ADN et a besoin de former un complexe avec la protéine P-Ser-HPr avec une stœchiométrie 2:2. L'activité régulatrice de CcpA est donc directement reliée à la disponibilité des sources de carbone par l'intermédiaire de P-Ser-HPr<sup>(99,110)</sup>. Chez *Bacillus subtilis*, CcpA est responsable de la régulation de près de 10% des gènes totaux en réponse à la disponibilité des sources de carbones<sup>(120)</sup>. Les boîtes *cre* se caractérisent pas des palindromes imparfaits d'une quinzaine de paires de base hautement dégénérés dont seules quelques nucléotides sont systématiquement retrouvés (**Figure 18**)<sup>(121)</sup>.



**Figure 18** : Séquence consensus des boîtes *cre* proposée chez *E. faecalis* V583

La taille des lettres n'est pas proportionnelle à la fréquence des nucléotides rencontrés. Réalisé à l'aide de l'utilitaire WebLogo v2.8.2<sup>(122)</sup>. D'après<sup>(121)</sup>.

Le complexe CcpA/P-Ser-HPr agit majoritairement comme un répresseur mais il peut également permettre l'induction de l'expression de certains gènes en fonction de l'endroit où se situe la boîte *cre* par rapport au début de transcription du gène cible. Si elle se situe au sein du promoteur ou après le début de transcription, CcpA aura un rôle de répresseur alors que si la boîte *cre* se trouve en amont du promoteur, il peut agir comme un activateur<sup>(110)</sup>.

La répression catabolique peut être instaurée par des mécanismes indépendants du régulateur pléiotrope CcpA mais généralement reliés aux éléments constitutifs des PTS. Une des autres stratégies adoptées par les bactéries est de bloquer les systèmes d'import des sources de carbone non préférentielles.

### c. L'exclusion et expulsion de l'inducteur

La plupart des gènes impliqués dans le catabolisme d'une source de carbone ne sont exprimés de manière efficace qu'en présence du substrat d'intérêt. Cette expression peut être instaurée par une activation directe ou une levée de répression. Une des stratégies adoptées par les bactéries est alors de bloquer l'import de certaines sources de carbones inhibant ainsi rapidement l'expression des gènes liés. Ce blocage de l'import, appelé "exclusion de l'inducteur" est majoritairement dépendant de P-Ser-HPr chez les firmicutes, cette protéine est en effet capable d'inhiber l'activité de certains transporteurs non PTS comme des importeurs ABC<sup>(98,110)</sup>. Dans ce cas, P-Ser-HPr semble interagir avec les protéines NBP présentant une activité ATPase et inhiber l'hydrolyse d'ATP mais le mécanisme exact reste à déterminer<sup>(110)</sup>. C'est notamment le cas pour le transporteur ABC des maltodextrines chez *L. casei*, P-Ser-HPr interagit directement avec l'ATPase MalK vraisemblablement proche du motif de Walker A et inhibe l'activité du transporteur associé<sup>(123)</sup>. Chez les entérobactéries, un système similaire a été décrit chez *E. coli* où EIIA<sup>Glc</sup> interagit avec l'ATPase MalK du transporteur ABC du maltose et inhibe la dimérisation ATP dépendante indispensable à sa fonction<sup>(124,125)</sup>. Comme nous l'avons vu dans une partie précédente, P-Ser-HPr n'est pas un substrat efficace pour la phosphorylation sur l'histidine 15. Par conséquent lors d'une croissance sur glucose, le flux glycolytique est important et la majorité des protéines HPr sont phosphorylées sur leur sérine 46 P-Ser-HPr et très peu sont présentes sous leur forme P~His-HPr. La faible concentration en P~His-HPr peut altérer le bon fonctionnement des PTS spécifiques de sources de carbones alternatives via un phénomène de compétition des EIIA pour P~His-HPr<sup>(99)</sup>. Cette compétition ne semble cependant pas affecter le plein fonctionnement des PTS permettant l'import du glucose.

Une autre stratégie consiste à exporter activement l'inducteur de systèmes impliqués dans le catabolisme de sources de carbone alternatives. Cette stratégie appelée "expulsion de l'inducteur" est adoptée par certains genres bactériens comme les entérocoques, les streptocoques ou les lactocoques<sup>(126,127)</sup>. Ce phénomène reste encore mal caractérisé et fait systématiquement l'objet d'études utilisant des molécules non métabolisables par les bactéries<sup>(126-130)</sup>. Ces conditions d'expérimentations éloignent par conséquent ces travaux du contexte physiologique des microorganismes étudiés. Ce phénomène apparaît toutefois également relié à la protéine P-Ser-HPr. Une étude a cependant montré que P-Ser-HPr n'était pas essentielle pour assurer cette expulsion chez *L. casei*<sup>(131)</sup>. Si les substrats à exporter ont préalablement été importés par des PTS, ils s'accumulent sous leur forme phosphorylée dans le cytoplasme. La

première étape de l'expulsion sera la déphosphorylation de ces sucres par une phosphatase spécifique des sucres phosphorylés associée à la membrane (<sup>132</sup>), les sucres seront par la suite exportés (<sup>99</sup>). L'exclusion de l'inducteur peut également être instaurée plus indirectement comme dans le cas du métabolisme du glycérol chez les firmicutes.

#### d. Le cas de GlpK

La glycerol kinase retrouvée chez les firmicutes représente un exemple original de protéine régulée par phosphorylation via un acteur des systèmes PTS mais ne possédant pas de domaine "PTS-like". Cette enzyme catalyse la phosphorylation du glycérol après son import par transport facilité, cette réaction est la première étape du catabolisme du glycérol via la voie aérobie de phosphorylation du glycérol (ou voie Glp) (<sup>61</sup>). Cette enzyme GlpK possède une région conservée contenant une histidine phosphorylable. Pour posséder une activité efficace, la glycerol kinase nécessite une phosphorylation réalisée par P~His-HPr sur cette histidine conservée (<sup>98</sup>). Si un sucre PTS facilement métabolisable est présent, P~His-HPr transférera préférentiellement son groupement phosphate au domaine EIIA du PTS correspondant et l'enzyme GlpK ne sera plus pleinement active limitant ainsi grandement le catabolisme du glycérol. Ce glycérol non phosphorylé s'accumule dans la cellule et diminue en conséquence sa diffusion à travers la membrane bloquant par la même le transport facilité assuré par GlpF (<sup>61,98</sup>).

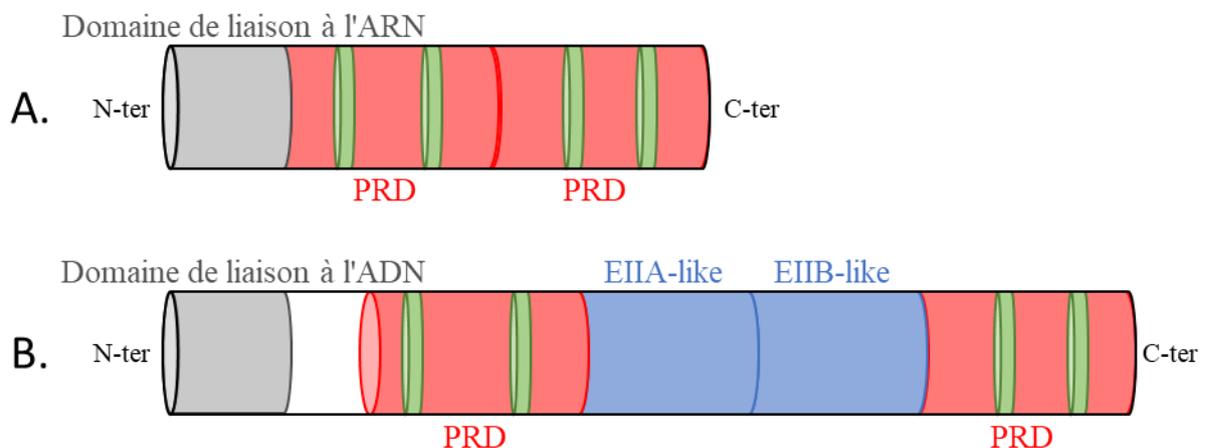
La régulation des gènes impliqués dans le catabolisme de sources de carbone alternatives peut également être assurée par des régulateurs transcriptionnels possédant des domaines particuliers.

#### e. Les régulateurs à domaines PRD (PTS Regulation Domain)

La répression catabolique indépendante de CcpA peut dans certains cas être assurée par certains régulateurs transcriptionnels possédant des domaines particuliers appelés des domaines PRD (PTS Regulation Domain). Ces domaines contiennent généralement deux histidines phosphorylables par différents acteurs protéiques constituant les PTS. L'état de phosphorylation de ces histidines détermine l'activité des régulateurs à domaines PRD. Ces derniers sont en majorité impliqués dans la régulation de quelques gènes du métabolisme notamment des PTS et permettent ainsi une régulation finement ciblée (<sup>110</sup>). Ces régulateurs agissent par activation de la transcription de leurs gènes cibles comme antiterminateurs ou activateurs transcriptionnels (<sup>133</sup>). La phosphorylation d'un PRD effectuée par P~His-HPr stimule l'activité du régulateur cible, lorsqu'un substrat PTS facilement métabolisable est importé, P~His-HPr va préférentiellement transmettre son groupement phosphate au domaine EIIA du PTS concerné

pour permettre l'import du substrat. L'activation de l'expression via le régulateur transcriptionnel à domaine PRD ne pourra, dans ce cas, pas être assurée <sup>(134)</sup>. Les phosphorylations réalisées par des domaines EIIB (ou parfois EIIA) vont quant à elles inhiber l'activité de ces régulateurs. En absence de leur substrat spécifique, les EIIB se retrouvent sous leur forme phosphorylée et vont transmettre leur groupement phosphate au régulateur indiquant ainsi que le métabolite n'est pas présent. A l'inverse, en présence de leur substrat spécifique, les EIIB ne seront plus phosphorylés car ils transmettront préférentiellement leur groupement phosphate pour permettre l'import de ce substrat <sup>(134)</sup>. Pour que ces régulateurs soient actifs, il faut par conséquent réunir plusieurs conditions : la présence du substrat spécifique du PTS concerné et l'absence de sucre PTS facilement métabolisable comme le glucose.

Concernant les régulateurs à domaines PRD agissant comme antiterminateurs, ils possèdent seulement un domaine de fixation à l'ARN au niveau de leur extrémité N-terminale et deux domaines PRD comme décrits précédemment. La fixation de ces régulateurs sur l'ARN au niveau d'un terminateur de transcription permet d'empêcher la formation de la structure tige-boucle et stimule ainsi la transcription du gène cible <sup>(98)</sup>. C'est par exemple le cas de la protéine LicT chez *B. subtilis* qui régule l'expression des gènes *bglPH* et *bglS* impliqués dans le métabolisme des  $\beta$ -glucosides (**Figure 19.A**) <sup>(135)</sup>.



**Figure 19** : Structure des régulateurs à domaines PRD

Les domaines d'intérêt sont représentés par des cylindres. Le domaine de liaison aux acides nucléiques à l'extrémité N-terminale est symbolisé en gris, les domaines PRD sont représentés en rouge et les domaines EIIA-like et EIIB-like en bleu. Les acides aminés phosphorylables au sein des PRD sont indiqués par des cylindres verts. Les motifs conservés au sein de ces domaines sont représentés en vert. **A.** Schéma d'un régulateur à domaines PRD hypothétique fonctionnant comme un antiterminateur. **B.** Schéma d'un régulateur à domaines PRD hypothétique fonctionnant comme un activateur transcriptionnel. L'ordre et le nombre des domaines PTS-like (EIIA et EIIB-like) n'est pas conservé entre les activateurs à domaines PRD. Ces protéines présentent de plus un domaine additionnel selon leur famille dans la région blanche du schéma. D'après <sup>(98)</sup>.

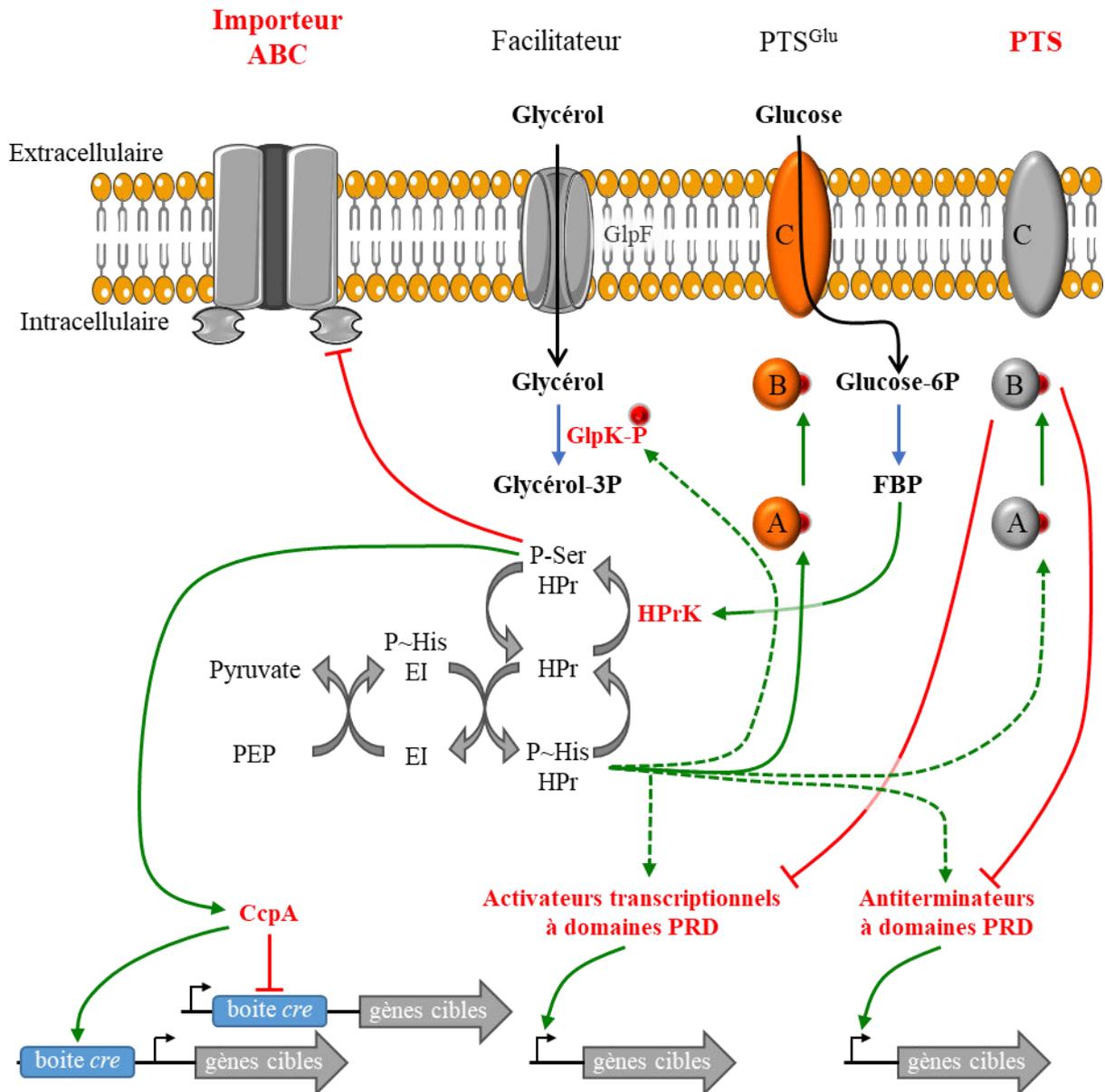
Les régulateurs à domaines PRD agissant comme des activateurs transcriptionnels présentent plus de variabilité. Ils possèdent tous un domaine de fixation à l'ADN au niveau de leur extrémité N-terminale ainsi que deux domaines PRD au sein de leur séquence. Ces régulateurs possèdent également des domaines additionnels EIIA-like et EIIB-like contenant des acides aminés phosphorylables au sein de régions similaires aux domaines EIIA ou EIIB. Ces domaines additionnels peuvent être des cibles de phosphorylations par des acteurs protéiques constituant les PTS ou d'interaction avec d'autres protéines modulant leur activité (**Figure 19.B**) <sup>(98,134)</sup>. On peut citer pour illustrer l'activateur transcriptionnel MtlR chez *B. subtilis*. Ce régulateur assure l'activation transcriptionnelle de l'opéron *mtlAFD* impliqué dans le métabolisme du mannitol. MtlR est composé de deux domaines PRD ainsi qu'un domaine EIIA-like et un EIIB-like. Cette protéine est activée par phosphorylation au niveau du second PRD par P~His-HPr et inhibée par phosphorylation au niveau du domaine EIIB-like par EIIA<sup>Mtl</sup>. Elle nécessite également d'être séquestrée par le domaine EIIB<sup>Mtl</sup> de la protéine MtlA pour être active. Tous ces niveaux de régulation permettent ainsi à la bactérie de finement adapter l'expression des gènes cibles de ce régulateur en fonction de la disponibilité des sources de carbone <sup>(136)</sup>.

#### f. La forme doublement phosphorylée de HPr

Nous avons vu précédemment que la phosphorylation présente sur la sérine en position 46 de HPr chez les firmicutes diminuait l'efficacité de la phosphorylation au niveau de l'histidine en position 15 par EI essentielle au fonctionnement des PTS comme importeurs de substrats. On retrouve néanmoins chez certaines bactéries une forme doublement phosphorylée (P~His)(P-Ser)HPr de la protéine HPr sur les deux acides aminés. La fonction de cette protéine reste très méconnue à ce jour mais est décrite comme peu retrouvée chez les entérocoques <sup>(137,138)</sup>. Toutes ces stratégies représentent une vue d'ensemble non exhaustive des capacités de régulation assurées par les PTS. Ces systèmes permettent à la bactérie d'adopter des réactions rapides et adaptées aux variations de substrats dans son environnement direct via des états de phosphorylation de protéines clefs. Un schéma récapitulatif de ces moyens de régulation est présenté dans la **Figure 20**.

### **III. Les régulateurs transcriptionnels**

La régulation métabolique chez les procaryotes est souvent assurée par des protéines capables de lier l'ADN appelées des régulateurs transcriptionnels. Ces dernières vont moduler l'expression de gènes au sein de leur régulon pour permettre au microorganisme de s'adapter à



**Figure 20** : Schéma du réseau de régulation dépendant des PTS

Les transports de molécules sont symbolisés par des flèches noires et les transformations enzymatiques par des flèches bleues.

Les noms des protéines cibles de la régulation via les PTS sont écrits en rouge. Les domaines EII des PTS sont représentés avec un code couleur correspondant au système EII entier, chaque domaine EII est symbolisé par sa dernière lettre. Les transporteurs en gris sont des transporteurs soumis à la répression catabolique.

Les flèches vertes symbolisent des stimulations d'activités des cibles et les flèches rouges des inhibitions. Les flèches en pointillé symbolisent une compétition pour la protéine centrale P~His-HPr.

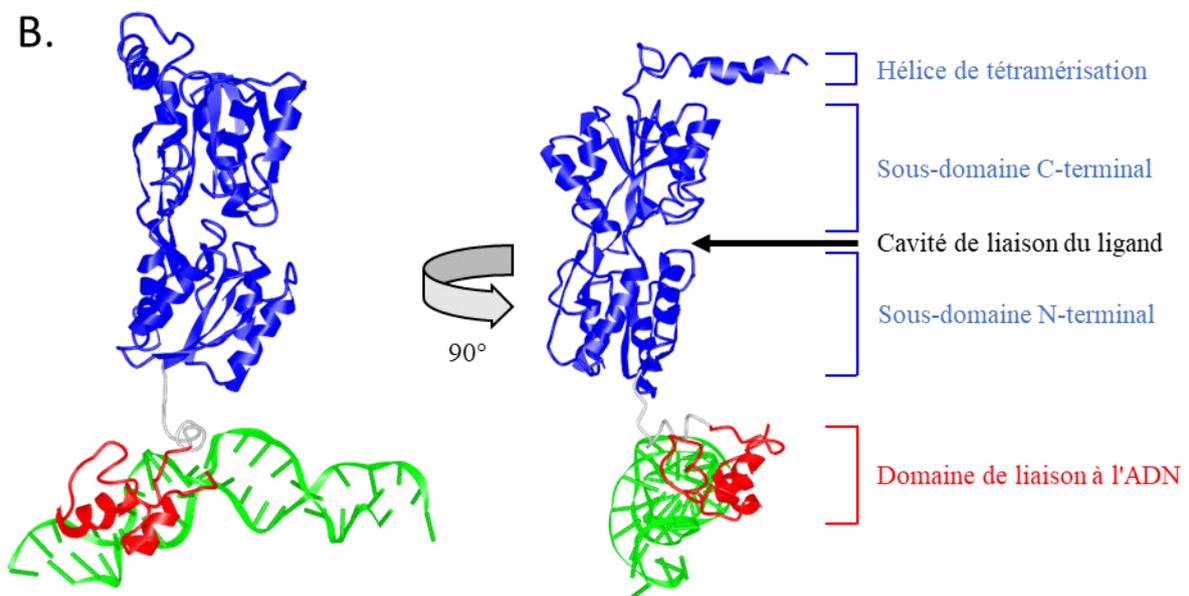
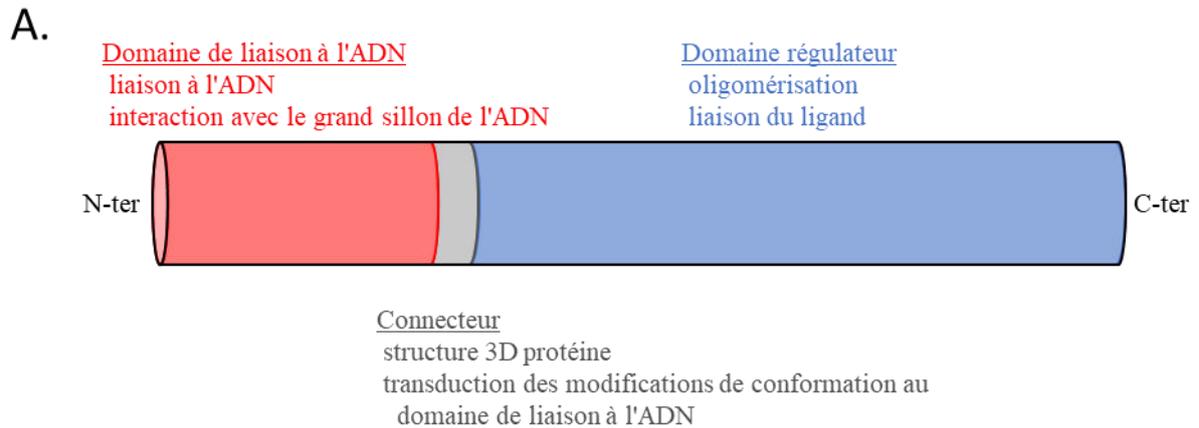
son environnement direct. Deux familles de régulateurs vont nous intéresser pour la suite de ce manuscrit : les régulateurs LacI/GalR et la famille RpiR.

## A. Les régulateurs transcriptionnels de la famille LacI/GalR

Les régulateurs transcriptionnels de la famille LacI/GalR sont retrouvés exclusivement chez les procaryotes. Ils regroupent un nombre important de protéines impliquées dans la régulation des gènes du métabolisme. Cette famille a été nommée en référence aux deux premiers représentants caractérisés chez *E. coli* et impliqués respectivement dans la régulation des métabolismes du lactose et du galactose. Ces protéines présentent des caractéristiques partagées par l'ensemble des représentants de cette famille protéique. Ils reconnaissent des motifs d'ADN composés de séquences palindromiques plus ou moins conservées. Deux grandes sous-familles peuvent être identifiées. D'abord la sous-famille "CcpA-like" qui regroupe la majorité des régulateurs transcriptionnels LacI/GalR chez les firmicutes, les motifs de reconnaissance de ces régulateurs présentent généralement un motif central AANCGNTT (où N peut être n'importe quel nucléotide) <sup>(139,140)</sup>. On retrouve ensuite plus rarement chez les firmicutes des régulateurs de la sous-famille EbgR-like, leur motif de reconnaissance est généralement centré autour d'une séquence TTNASTNAA (où S peut être C ou G) <sup>(140)</sup>. La fixation de ces motifs palindromiques est assurée par l'oligomérisation de ces régulateurs, un monomère reconnaissant la moitié du motif de reconnaissance <sup>(141)</sup>. Ces protéines agissent majoritairement comme des répresseurs mais certains représentants de cette famille peuvent agir comme des activateurs, voire présenter un double rôle répresseur/activateur (par exemple CcpA). Leur activité est généralement modulée par la fixation de petites molécules comme des métabolites qui vont entraîner des changements de conformation du régulateur afin de moduler son activité <sup>(141)</sup>. Ces protéines présentent une séquence primaire relativement conservée formant une structure tridimensionnelle constituée de trois domaines systématiquement organisés dans le même ordre. La partie N-terminale contient le domaine de liaison à l'ADN et la partie C-terminale un domaine dit régulateur. Ces deux domaines sont reliés par un connecteur (linker). La structure de ces régulateurs est présentée en **Figure 21**.

### 1. Le domaine de fixation à l'ADN

Le domaine de fixation à l'ADN, composé d'environ 50-60 acides aminés, est retrouvé à l'extrémité N-terminale des régulateurs transcriptionnels de la famille LacI/GalR. Il est constitué d'un motif en "hélice-coude-hélice" (ou HTH pour Helix-Turn-Helix) lui-même formé de deux hélices  $\alpha$  séparées par un coude de 4 acides aminés contenant généralement une glycine



**Figure 21** : Structure des régulateurs transcriptionnels de la famille LacI/GalR

**A.** Schéma de la structure primaire d'un régulateur transcriptionnel de la famille LacI/GalR. Les domaines d'intérêt sont représentés par des cylindres.

Une brève description du rôle de chaque domaine est indiquée en dessous du nom du domaine.

**B.** Structure tridimensionnelle de la protéine LacI chez *E. coli* en présence d'ADN. Disponible sur la banque de structure tridimensionnelle de macromolécules MMDB. ID : 73984 <sup>(142)</sup>.

Le domaine de liaison à l'ADN est représenté en rouge, le domaine linker en gris et le domaine régulateur en bleu. La molécule d'ADN est représentée en vert.

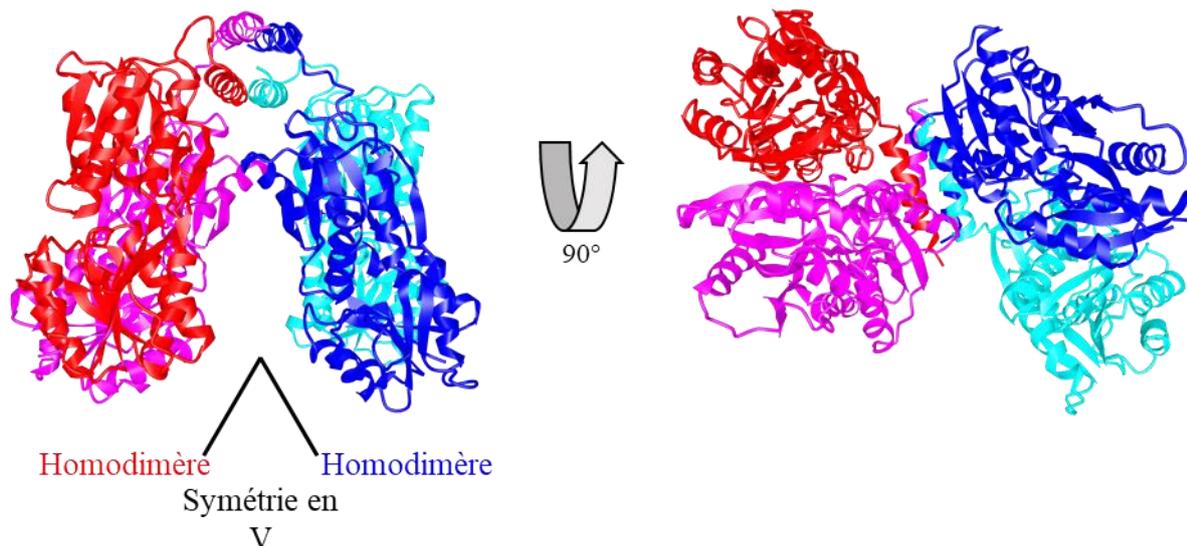
Un seul monomère du régulateur LacI est représenté pour des soucis de clarté. La protéine s'organise normalement en tétramère formé de deux dimères (plus de détails dans le texte).

D'après <sup>(105)</sup>.

(<sup>143</sup>). La seconde hélice  $\alpha$  est généralement appelée l'hélice de reconnaissance (recognition helix) et va interagir directement avec l'ADN via des liaisons faibles (<sup>144</sup>). Cette hélice va interagir avec le grand sillon de l'ADN au niveau des séquences de reconnaissance du régulateur et va être critique pour la spécificité de liaison à l'ADN (<sup>143</sup>). La capacité des régulateurs transcriptionnels LacI/GalR pour lier l'ADN est directement dépendante de la présence ou l'absence de certaines molécules dans la cellule. Cette régulation allostérique est assurée par le domaine régulateur.

## 2. Le domaine régulateur

Le domaine responsable de l'adaptabilité de liaison à l'ADN de cette famille de régulateurs transcriptionnels est situé à l'extrémité C-terminale de la protéine et représente la majorité de la séquence de la protéine. Ce domaine appartient à la superfamille des protéines de liaison périplasmiques (periplasmic binding superfamily) et plus précisément à la famille "Venus flytrap" (littéralement plante attrape mouches). Il se compose de deux sous-domaines globulaires similaires entrelacés chacun formé de deux hélices  $\alpha$  entourant un feuillet  $\beta$  (<sup>145,146</sup>). Ces deux sous-domaines forment une cavité où viendra se loger le ligand du régulateur. La fixation de ce ligand va induire un changement de conformation important au niveau du sous-domaine le plus à l'extrémité N-terminale mais pas au niveau du sous-domaine à l'extrémité C-terminale. Ce domaine régulateur assure également la dimérisation des régulateurs transcriptionnels de la famille LacI/GalR. Dans le cas du répresseur LacI d'*E. coli*, le répresseur s'organise en homotétramère composé de deux dimères organisés en forme de V, cette tétramérisation est assurée par une petite séquence organisée en hélice située à l'extrémité C-terminale du domaine régulateur (**Figure 22**) (<sup>146,147</sup>).



**Figure 22 :** Structure quaternaire de la protéine LacI d'*E. coli*

La structure tridimensionnelle d'un homotétramère de la protéine LacI chez *E. coli* est disponible sur la banque de structure tridimensionnelle de macromolécules MMDB. ID : 50084 <sup>(142)</sup>. Seuls les domaines de régulation sont représentés. Chaque monomère est identifié par un code couleur. Cette structure met en évidence l'organisation en deux dimères organisés en V (V inversé sur ce schéma).

Les deux principaux domaines des régulateurs LacI sont reliés entre eux par un court connecteur.

### 3. Le domaine linker

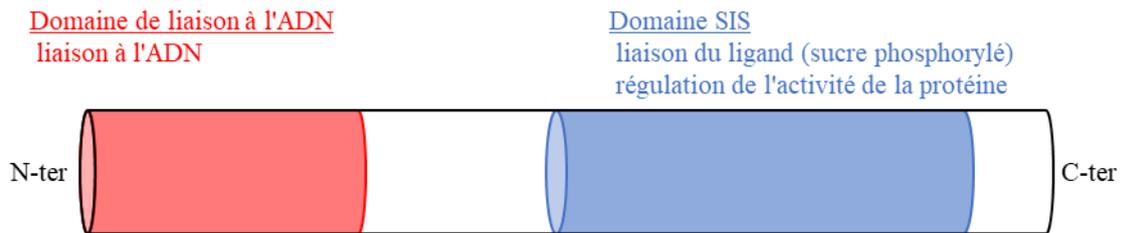
Cet ultime domaine, constitué de moins de 20 acides aminés, assure la liaison entre le domaine de fixation à l'ADN et le domaine régulateur que nous venons de voir précédemment. Une grande proportion de ces régulateurs présentent un motif semi conservé YxPxxxAxxL (où x peut représenter n'importe quel acide aminé) appelé le motif "YPAL" <sup>(148)</sup>. Ce domaine linker forme une hélice centrale coudée entourée de deux extrémités non structurées physiquement reliées aux domaines N-terminal et C-terminal. Il assure un rôle primordial dans la structure tridimensionnelle de la protéine et permet une liberté de mouvement du domaine de liaison à l'ADN <sup>(141)</sup>. Sa fonction n'est cependant pas limitée à un rôle structural, ce connecteur est notamment responsable de la transduction des modifications de conformations lors de la liaison du ligand. Il va ainsi intervenir dans la capacité du régulateur à lier l'ADN en réponse à la présence du ligand <sup>(141,148)</sup>.

### B. Les régulateurs transcriptionnels de la famille RpiR

Cette famille de régulateur RpiR est beaucoup moins documentée que les régulateurs LacI/GalR présentés précédemment. La famille RpiR tient son nom du régulateur du même nom découvert chez *E. coli* <sup>(149)</sup>. La protéine RpiR (aussi nommée AlsR) a d'abord été décrite

chez cette bactérie comme un répresseur sur l'expression du gène *rpiB* (*alsI*) codant pour une protéine à activité ribose-6-P isomérase (<sup>149</sup>). Le régulon de ce régulateur RpiR/AlsR a ensuite été élargi à l'opéron *alsRBACEK* impliqué dans le catabolisme du D-allose (<sup>150</sup>).

Les régulateurs de cette famille RpiR, pouvant agir comme des répresseurs ou des activateurs selon les cas, présentent des similarités, notamment au niveau de l'organisation de leur structure primaire. Une structure primaire typique est présentée dans la **Figure 23**.



**Figure 23** : Structure primaire d'un régulateur transcriptionnel de la famille RpiR  
Les domaines d'intérêt sont représentés par des cylindres. Une brève description du rôle de chaque domaine est indiquée en dessous du nom du domaine.

Comme les LacI/GalR, ils sont composés d'un domaine de liaison à l'ADN de type HTH au niveau de leur extrémité N-terminale. Ils semblent reconnaître des motifs palindromiques dégénérés. En se basant sur les motifs de reconnaissance déterminés pour certains régulateurs de cette famille, un consensus se dégage sous la forme KGAWN<sub>6-8</sub>TKCH (où K peut être T ou G, W peut être T ou A, H peut être C, T ou A et N peut être n'importe quel nucléotide) (<sup>151-153</sup>). Cependant, le nombre d'études caractérisant la spécificité de tels régulateurs est encore limité pour établir avec certitude une réelle séquence consensus. Ces régulateurs se fixent au niveau de régions contenant une ou deux fois le motif consensus proposé ci-dessus. Par homologies avec les régulateurs LacI et au vu de la relative symétrie de leur séquence de reconnaissance, les régulateurs transcriptionnels RpiR sont supposés fonctionner en oligomères. Par exemple, le régulateur MurR d'*E. coli* s'organise vraisemblablement en tétramère pour être pleinement actif. Comme ce régulateur se fixe sur une région contenant deux motifs de liaison, il s'organise probablement en double dimère, chaque dimère fixant un site de reconnaissance (<sup>151</sup>).

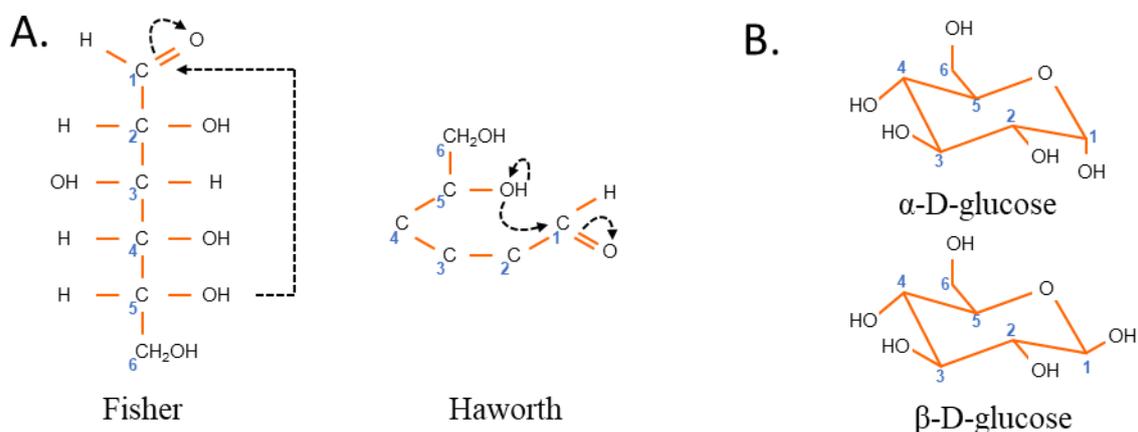
Les régulateurs RpiR présentent également un domaine SIS (pour Sugar Isomérase) au niveau C-terminal (<sup>154</sup>). Ce domaine est également retrouvé chez des protéines à activité "phosphosugar isomérase" d'où il tire son nom. Il est largement distribué dans le monde du vivant aussi bien chez les procaryotes que chez les eucaryotes. La protéine GlnS (Glucosamine-6-P Synthase) d'*E. coli* possède deux domaines SIS et sa structure a été étudiée. Cette protéine catalyse la formation de Glucosamine-6-P par isomérisation du Fructose-6-P en utilisant la glutamine comme source d'azote. La protéine GlnS forme un homodimère pour être

active, une poche est formée au niveau de l'interface entre les domaines SIS situés les plus à l'extrémité N-terminale de chaque monomère. Cette poche va constituer le site de fixation du substrat phosphorylé (<sup>155</sup>). Cette nécessité d'une oligomérisation pour constituer le site de fixation du substrat renforce encore la vraisemblable association en oligomères des régulateurs RpiR. Chez ces régulateurs transcriptionnels, le domaine SIS permet la liaison d'un sucre phosphorylé. La liaison de ce sucre va moduler l'activité du régulateur en modifiant son affinité pour l'ADN (<sup>154</sup>).

#### **IV. Le métabolisme du maltose et des maltodextrines chez différents organismes modèles**

Le métabolisme du maltose et des maltodextrines sera détaillé dans cette partie chez plusieurs organismes modèles. Dans un but de clarté, les réactions seront orientées uniquement dans le sens du catabolisme même si certaines réactions peuvent être réversibles.

Les maltodextrines sont des molécules organiques de la famille des oses. Ces sucres sont composés de plusieurs molécules de glucose liées entre elles par des liaisons glucidiques  $\alpha 1,4$ . Le carbone 1 du glucose possède un pouvoir réducteur du fait de la fonction aldéhyde portée par cet atome (**Figure 24**).



**Figure 24** : Représentations du D-glucose

**A.** Projections de Fisher (gauche) et Haworth sans notion de perspective (droite) d'une molécule de glucose en conformation en chaîne ouverte.

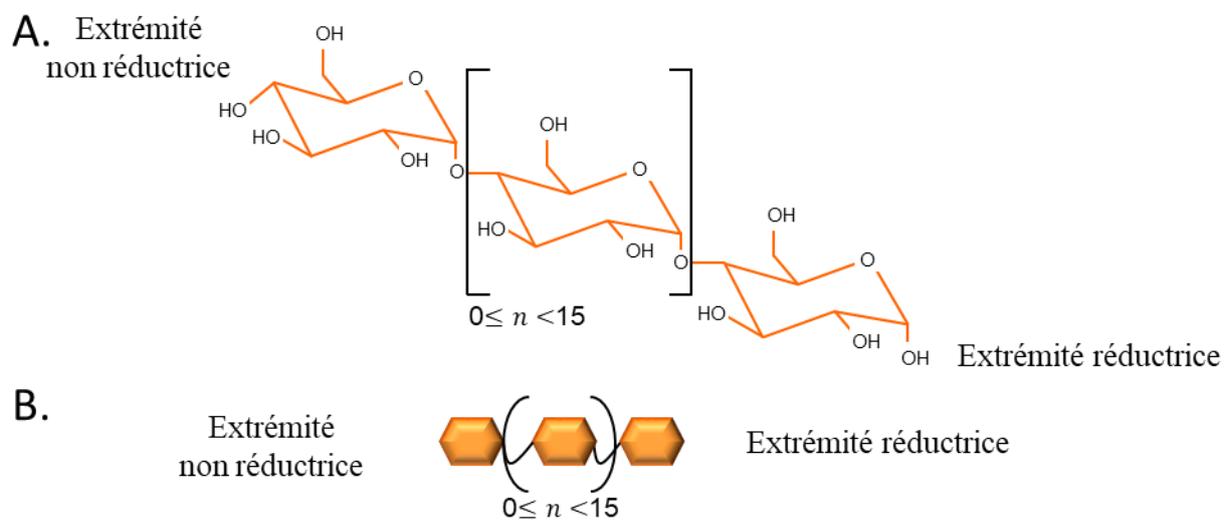
Les chiffres en bleu indiquent la numérotation des carbones selon la nomenclature de l'Union Internationale de Chimie Pure et Appliquée (IUCPA).

Le carbone 1 présente un pouvoir réducteur.

Les flèches en pointillé représentent les transferts d'électrons lors de la réaction de cyclisation.

**B.** Représentation des deux anomères du D-glucopyranose (glucose) en conformation chaise.

Au sein d'un polymère de sucre, ces unités de glucose perdent cependant un groupement hydroxyl et/ou un atome d'hydrogène lors de la formation de la liaison glucidique, on parle alors de résidus glycosyl et non plus de molécules de glucose. Dans le cas des maltodextrines, les liaisons glucosidiques sont engagées entre le carbone 1 du résidu glycosyl  $n$  et le carbone 4 du résidu glycosyl  $n+1$  selon une géométrie  $\alpha$ . Lorsque les résidus glycosyl sont engagés dans une liaison glycosidique, leur carbone anomère ne possède plus de pouvoir réducteur, cela permet par conséquent d'orienter les maltodextrines : l'extrémité présentant un résidu glycosyl dont le carbone 1 n'est pas engagé dans une liaison est appelée extrémité réductrice. L'autre extrémité est donc désignée comme non réductrice (**Figure 25**) <sup>(156)</sup>.



**Figure 25 :** Structure générale des maltodextrines

**A.** Schéma général des maltodextrines linéaires en conformation chaise.

Le  $n$  indique les nombres possibles de résidus glycosyl entre les crochets.

**B.** Représentation simplifiée des maltodextrines. Les résidus glycosyl sont représentés par des hexagones et les liaisons glycosidiques par des courbes noires.

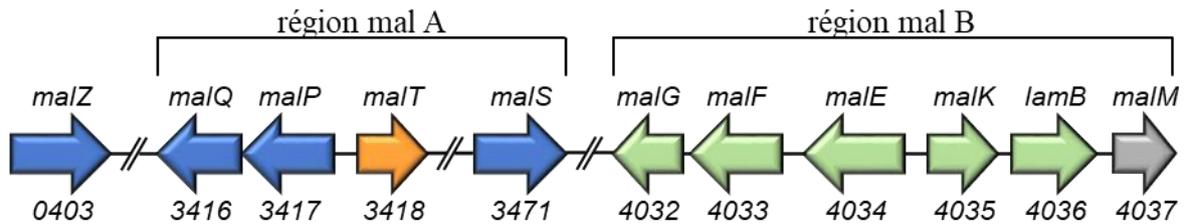
Le  $n$  indique les nombres possibles de résidus glycosyl entre les crochets.

La longueur des maltodextrines peut varier de deux résidus glycosyl à quelques dizaines de résidus. Les maltodextrines sont nommées en fonction de leur degré de polymérisation (DP), c'est à dire du nombre de résidus glycosyl les composant, par l'ajout d'un affixe multiplicateur (maltotriose, maltotétraose, etc.). La plus courte des maltodextrines formée de seulement deux résidus glycosyl est appelée maltose. Ces molécules sont généralement obtenues après dégradation de polymères de stockage que sont le glycogène ou l'amidon par des  $\alpha$ -amylases.

## A. Escherichia coli

### 1. Transport et catabolisme

La bactérie la mieux caractérisée du point de vue du métabolisme du maltose et des maltodextrines est sans aucun doute *Escherichia coli*. Ce microorganisme utilise différents gènes retrouvés dans deux régions distinctes nommées malA et malB pour métaboliser ces sucres (**Figure 26**).



**Figure 26 :** Gènes impliqués dans le métabolisme du maltose et des maltodextrines chez *E. coli*  
Les numéros de gènes inscrits en dessous des flèches correspondent font référence au génome d'*E. coli* K-12 MG1655 (*b\_xxxx*).  
Le sens des flèches reflète le sens de transcription et la couleur la fonction du gène. Les flèches vertes représentent les gènes fonctionnellement liés au transport de substrat, les flèches bleues les gènes liés au catabolisme du substrat, la flèche orange le gène lié à la régulation et la flèche grise représente un gène dont la fonction est inconnue

Une review rédigée par Boos et Shuman (<sup>157</sup>) décrit les principaux gènes impliqués dans ce métabolisme. Dans un premier temps, les substrats doivent franchir la membrane externe. Les maltodextrines d'une longueur inférieure à 7-8 résidus glycosyl sont transportées dans le périplasme grâce à la porine LamB (<sup>158,159</sup>). Le maltose et le maltotriose peuvent également traverser la membrane externe grâce aux porines générales OmpF et OmpC de manière plus lente (<sup>160</sup>). Les longues maltodextrines sont ensuite dégradées par une  $\alpha$ -amylase périplasmique MalS (<sup>159,161</sup>). Cette enzyme hydrolyse préférentiellement la liaison entre le 6 et 7<sup>ème</sup> résidu du côté non réducteur libérant ainsi majoritairement des sucres d'une longueur de 6 résidus et de courtes chaînes constituées de quelques résidus glycosyl. Ces sucres pourront traverser la membrane interne grâce à un importeur ABC MalFGK<sub>2</sub>. Ce transporteur contient deux protéines transmembranaires (MalF et MalG) (<sup>162,163</sup>) associées à deux protéines intracellulaires MalK. Il est rendu fonctionnel par l'hydrolyse d'Adénosine Triphosphate (ATP) assurée par les ATPases MalK (<sup>164,165</sup>). Cette hydrolyse est stimulée par la fixation d'une maltodextrine à la protéine périplasmique MBP (Maltose Binding Protein) codée par le gène malE qui va permettre la transduction d'un signal activant l'activité ATPase (<sup>87,166,167</sup>).

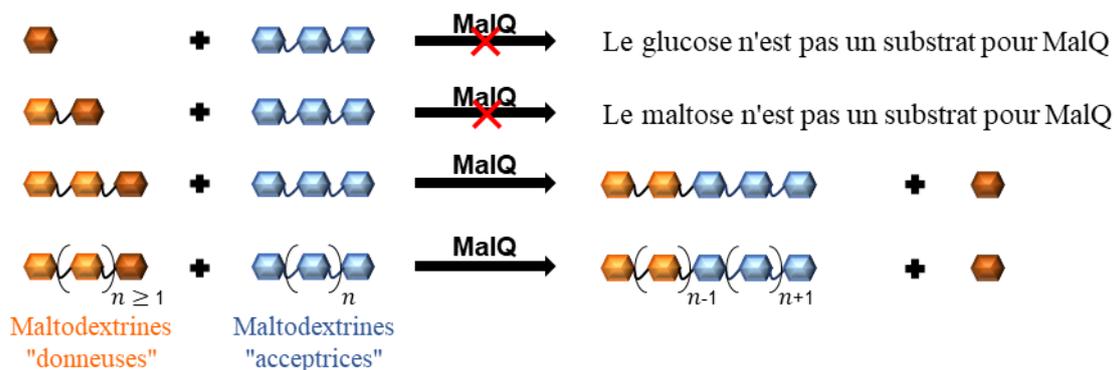
Une fois transportées dans le milieu intracellulaire, les maltodextrines dont la longueur des

chaînes est supérieure à 4 résidus glycosyl seront dégradées par 3 enzymes différentes. La première de ces enzymes, la maltodextrine phosphorylase MalP, permet aux maltodextrines d'être phosphorolysées en clivant la liaison glycosidique  $\alpha$ -1,4 du côté non réducteur des maltodextrines. Cela donnera naissance à du glucose-1-phosphate (G1P) et à une maltodextrine raccourcie d'un résidu glycosyl (<sup>168-171</sup>). Le G1P ainsi formé sera transformé en glucose-6-phosphate (G6P) par l'action de la phosphoglucomutase codée par le gène *pgm* (<sup>172,173</sup>) afin de rejoindre les voies de dégradation classiques comme la glycolyse.

La deuxième enzyme impliquée dans la dégradation des maltodextrines à longues chaînes est une maltodextrine hydrolase nommée MalZ capable de libérer des molécules de glucoses en hydrolysant la liaison  $\alpha$ -1,4 du côté réducteur des maltodextrines. Cette enzyme peut prendre en charge des maltodextrines dont la longueur des chaînes est comprise entre 3 et 7 résidus glycosyl (<sup>171,174-176</sup>).

Etonnamment MalP et MalZ ne permettent pas à *E. coli* de métaboliser le maltose car ce disaccharide n'est pas un substrat pour ces dernières. Par conséquent, lors d'une croissance en présence de maltose uniquement, la bactérie doit posséder une ultime enzyme qui est l'amyломaltase MalQ. Cette dernière permet de transférer un glucose du côté réducteur de la molécule "donneuse" vers l'extrémité non réductrice d'une autre maltodextrine "acceptrice". Ce transfert s'accompagne de la libération d'une molécule de glucose. Toutes les maltodextrines indépendamment de la longueur de leur chaîne peuvent être des molécules "acceptrices" et le maltotriose est la plus petite molécule donneuse (<sup>177-180</sup>), ce mode d'action est représenté en

**Figure 27.**



**Figure 27 :** Réactions catalysées par l'amyломaltase MalQ d'*E. coli*

Les maltodextrines donneuses sont représentées en orange et les acceptrices en bleu. Le résidu glycosyl du côté réducteur des maltodextrines donneuses est symbolisé de couleur plus foncée pour permettre de suivre son devenir au sein des produits.

Le petit  $n$  représente le nombre de résidus glycosyl entre les crochets.

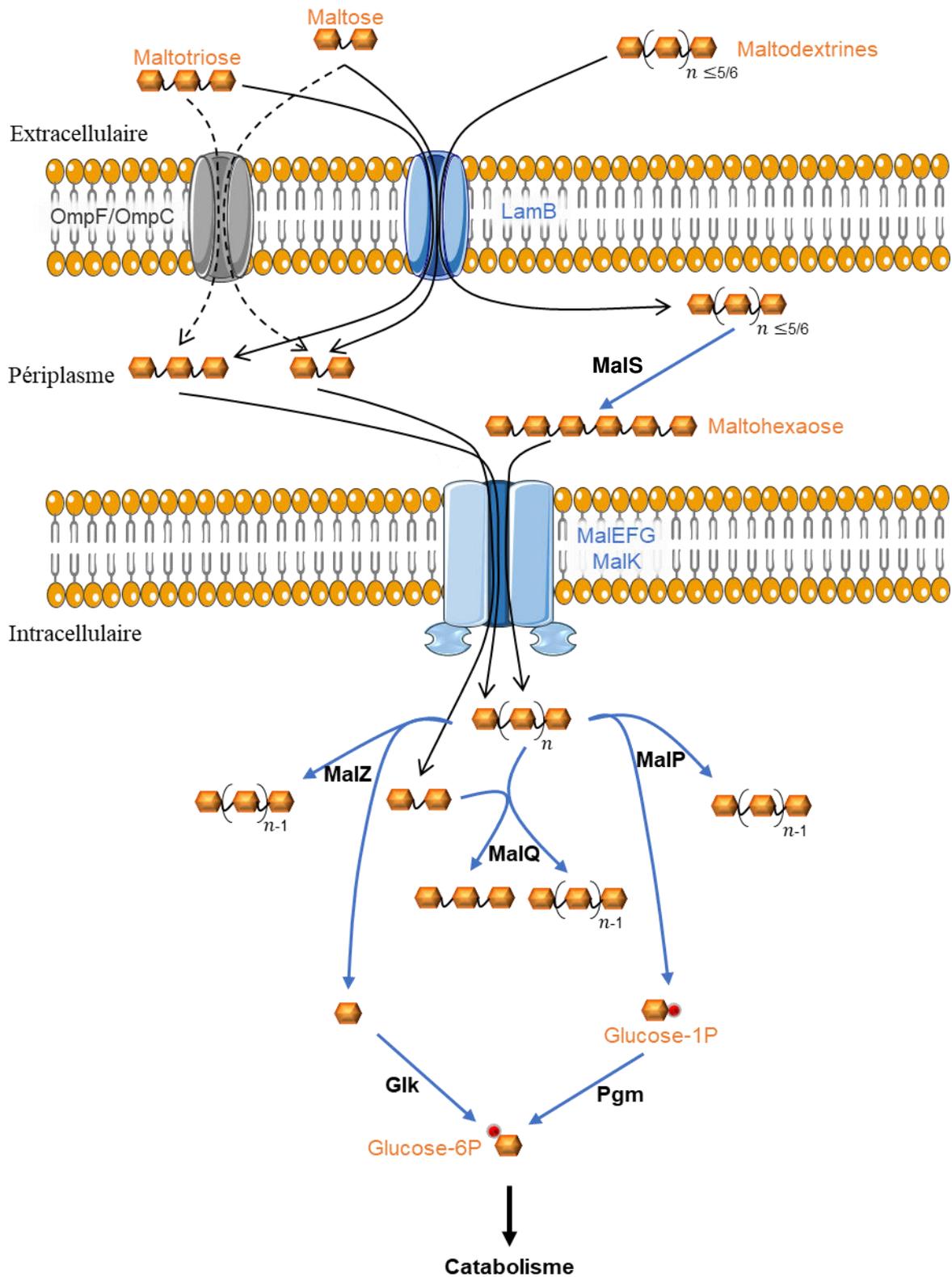
La dégradation du maltose nécessite donc la présence dans la cellule de maltodextrines plus longues servant d'amorces à la réaction. En effet, la cellule est obligée de rallonger la taille du disaccharide en le transformant en maltotriose ou maltotétraose. Ces derniers seront alors pris en charge par la maltodextrine phosphorylase MalP et la maltodextrine hydrolase MalZ.

Un gène nommé *malM* est présent au sein de la région B comme le troisième gène de l'opéron débutant par *malK*. Sa séquence en acides aminés apparaît très conservée (<sup>181,182</sup>) et elle est effectivement exprimée et localisée dans le périplasma (<sup>183</sup>). Cependant son rôle dans la physiologie bactérienne reste à déterminer. Un schéma récapitulatif du métabolisme des maltodextrines chez *E. coli* est présenté dans la **Figure 28**.

L'expression de ces gènes "*mal*" est directement régulée par l'activateur transcriptionnel MalT dont le gène est présent dans la région *malA* (<sup>184-186</sup>).

## 2. Régulation

L'activité du régulateur MalT est contrôlée à la fois au niveau transcriptionnel et post-traductionnel. En effet, la protéine MalT n'est active qu'en présence de maltotriose et d'ATP (<sup>187,188</sup>). Elle va en présence de ces ligands induire l'expression des gènes "*mal*" présentés ci-dessus. Le maltotriose, qui est donc le sucre inducteur du système, est présent dans la cellule à de fortes concentrations lors du catabolisme des maltodextrines via l'action des protéines MalP et MalZ. Cependant ce sucre inducteur peut également être le produit de dégradation du glycogène intracellulaire en l'absence d'apport extérieur de maltodextrines. Lors d'une telle situation, il n'est pas utile pour la bactérie d'induire les gènes du système *mal*. L'amyломaltase MalQ permet dans ce contexte de limiter la formation de maltotriose en rallongeant la longueur des maltodextrines contrôlant ainsi la concentration intracellulaire de l'inducteur. Lors de la croissance sur maltose, c'est au contraire cette enzyme qui permet d'aboutir rapidement à du maltotriose afin d'induire le système *mal*. Néanmoins, plusieurs autres systèmes permettent de limiter l'induction de ce système en cas d'absence de maltodextrines dans le milieu extérieur. Le principal est assuré par l'ATPase MalK de l'importeur ABC qui dans ces conditions est "en attente", c'est-à-dire quand elle n'hydrolyse pas d'ATP pour permettre l'import de maltodextrines. Dans cet état, MalK est capable de séquestrer l'activateur MalT dans le but d'empêcher ce régulateur d'induire le régulon *mal* (<sup>189,190</sup>). Si des maltodextrines sont présentes dans le milieu extérieur, l'importeur ABC MalEFK<sub>2</sub> va devenir fonctionnel et l'ATPase MalK va libérer le régulateur MalT. La présence de maltotriose empêche également cette interaction *in vitro* laissant supposer une compétition entre MalK et le maltotriose pour MalT (<sup>190,191</sup>). D'autres protéines cytoplasmiques, Aes et MalY sont également capables de réduire l'activité



**Figure 28** : Schéma du métabolisme du maltose et des maltodextrines chez *E. coli*

Les flèches noires représentent des transports de molécules et les bleues des transformations enzymatiques.

Le  $n$  indiqué en dessous des maltodextrines représente le nombre total de résidus glucosyl.

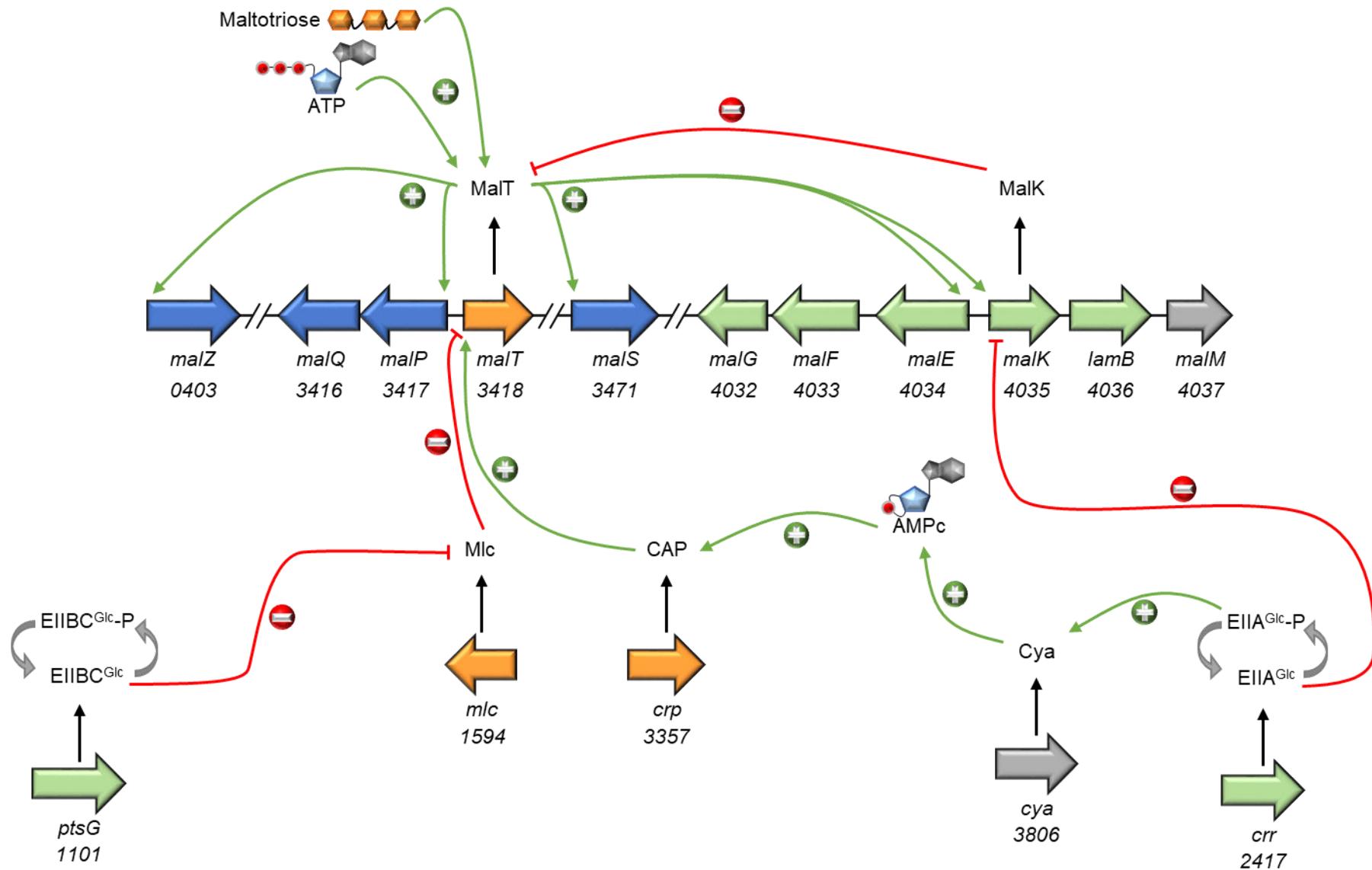
de MalT par des interactions directes mais leur lien avec le métabolisme du maltose reste mal compris (<sup>189,192-194</sup>).

L'activité du régulateur MalT est également régulée au niveau transcriptionnel. En effet l'expression du gène *malT* est sous le contrôle du régulateur global Mlc qui agit dans ce cas comme un répresseur en fixant une région d'ADN située au niveau du début de la transcription du gène *malT* (<sup>195,196</sup>). La protéine Mlc est quant à elle séquestrée par le PTS glucose EIIBC<sup>Glc</sup> lorsqu'il est dans un état non phosphorylé (<sup>197-199</sup>). Cette situation est retrouvée lorsque du glucose est rapidement importé dans la cellule. Ce système indique que les métabolismes des maltodextrines et du glucose sont étroitement liés chez *E. coli*. Cependant en dépit de ce moyen de régulation, le système "mal" reste réprimé lorsque la bactérie utilise du glucose comme source de carbone par des mécanismes de répression catabolique. Dans le cas d'*E. coli*, la hiérarchisation dans l'utilisation des sources de carbone est très largement reliée au taux d'AMPc présent dans la cellule, en effet la protéine CAP (Catabolite Activator Protein) encodée par le gène *crp* agit comme un activateur transcriptionnel et est essentielle pour l'expression du gène *malT* (<sup>200,201</sup>). Ce régulateur CAP nécessite la fixation d'AMPc pour être actif (<sup>202,203</sup>). Le taux d'AMPc dans la cellule est contrôlé par l'AMP cyclase dont l'activité est stimulée par la protéine EIIA du PTS glucose lorsqu'elle est sous sa forme phosphorylée. Lors de l'import du glucose, EIIA<sup>Glc</sup> est majoritairement non phosphorylée et ne pourra donc pas activer l'AMP cyclase. Cela a pour effet d'empêcher CAP d'induire l'expression de MalT qui est à son tour essentielle pour induire tout le système "mal". La protéine EIIA<sup>Glc</sup> dans son état non phosphorylée est également capable d'interagir avec l'ATPase MalK afin d'empêcher l'import de maltodextrines via l'importeur ABC MalFGK<sub>2</sub> (<sup>98,204</sup>). Un schéma du réseau de régulation global est présenté en **Figure 29**.

## B. Streptococci

### 1. Transport et catabolisme

Le métabolisme du maltose et des maltodextrines a été décrit chez plusieurs espèces de Streptocoques. Les voies de dégradation de ces sucres chez ces bactéries apparaissent conservées entre les différentes espèces de Streptocoques même si l'organisation génétique peut être très différente. Les différents gènes impliqués dans ces métabolismes chez *S. pneumoniae* sont représentés dans la **Figure 30**. Les gènes homologues chez *S. pyogenes* et *S. mutans*, deux espèces bien décrites concernant leur métabolisme des maltodextrines, sont présentés dans le **Tableau 5**.



**Figure 29** : Modèle de régulation du métabolisme du maltose et des maltodextrines chez *E. coli*  
 Les flèches rouges représentent des inhibitions de transcription et/ou d'activité et les vertes symbolisent des activations.  
 Les gènes respectent les mêmes codes que la **Figure 26**.

**Tableau 5** : Gènes impliqués dans le métabolisme du maltose et des maltodextrines chez différentes espèces de streptocoques

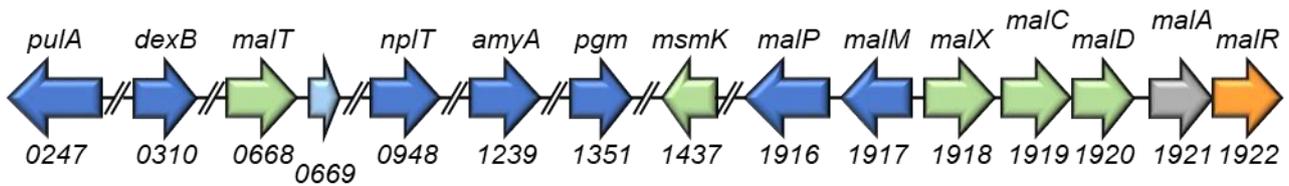
		Gènes		
	Fonctions	<i>S. pneumoniae</i> R6	<i>S. pyogenes</i> M5005	<i>S. mutans</i> UA159
<b>Protéines fonctionnellement liées au transport</b>	PTS maltose/maltotriose	<i>malT</i> (0668)	(1692)	<i>ptsG</i> (2047)
	Maltose-6'-P/Maltotriose-6''-P Phosphatase ?	(0669)	(1691)	(2046)
	Importeur ABC			
	- Protéines à TMD - Protéine SBP - Protéine NBP	<i>malC</i> (1919), <i>malD</i> (1920) <i>malX</i> (1918) <i>msmK</i> (1437)	<i>malF</i> (1059), <i>malG</i> (1060) <i>malE</i> (1058) <i>msmK</i> (1682)	<i>malF</i> (1569), <i>malG</i> (1570) <i>malE</i> (1058) <i>malK</i> (1571), <i>msmK</i> (882)
<b>Protéines impliquées dans le catabolisme</b>	Glycogène phosphorylase	<i>malP</i> (1916)	<i>glgP</i> (1055)	<i>glgP</i> (1564)
	Amylomaltase	<i>malM</i> (1917)	<i>malM</i> (1056)	<i>malQ</i> (1565)
	Amylases	<i>amyA</i> (1239)	<i>amyA</i> (1065) <i>amyB</i> (1066)	<i>amyA</i> (1590)
	Phosphoglucomutase	<i>pgm</i> (1351)	<i>pgmA</i> (0938)	<i>pgm</i> (1077)
	Pululanase	<i>pulA</i> (0247)	<i>pulA</i> (1680)	<i>pulA</i> (1541)
	Glucosidase	<i>dexB</i> (0310)	<i>dexB</i> (1681)	<i>dexB</i> (883)

Le nom usuel des gènes est indiqué s'il est connu. Seul le numéro de référence de chaque gène est indiqué entre parenthèse par soucis de clarté. Les noms de référence de chaque souche suivent la nomenclature suivante :

*S. pneumoniae* R6 (*spr\_xxxx*)

*S. pyogenes* M5005 (*spy\_xxxx*)

*S. mutans* UA159 (*smu\_xxx* ou *smu\_xxxx*)



**Figure 30 :** Gènes impliqués dans le métabolisme du maltose et des maltodextrines chez *S. pneumoniae*. Les numéros de gènes inscrits en dessous des flèches font référence au génome de *S. pneumoniae* R6 (*spr\_xxxx*).

Le sens des flèches reflète le sens de transcription et la couleur la fonction du gène. Les flèches vertes représentent les gènes fonctionnellement liés au transport de substrat, les bleues correspondent aux gènes liés au catabolisme du substrat, la flèche orange représente le gène lié à la régulation et la flèche grise un gène dont la fonction est inconnue.

L'import du maltose et du maltotriose est assuré chez ce genre bactérien par un système PTS MalT<sup>(205-207)</sup>, les oligosaccharides arriveront dans la cellule sous leur forme phosphorylée non métabolisable en l'état par la cellule. Un gène homologue au gène *mapP* décrit chez *E. faecalis* suit systématiquement le gène *malT* et est prédit pour coder une phosphatase dont l'action supposée serait de retirer ce groupement phosphate<sup>(208)</sup>. Les maltodextrines composées de 3 à 7-8 résidus glycosyl sont importées par un importeur ABC codé par les gènes *malXCD* chez *S. pneumoniae*<sup>(209,210)</sup> (*malXFG* chez *S. mutans*<sup>(211)</sup> et *malEFG* chez *S. pyogenes*<sup>(212)</sup>). L'énergie nécessaire pour importer ces substrats est apportée par hydrolyse d'ATP via l'action d'une ATPase souvent codée par un gène distant sur le génome. *S. mutans* fait cependant exception car un gène *malK* qui code pour l'ATPase fournissant l'énergie au système est situé directement en aval du gène *malG*<sup>(211)</sup>. L'activité de cette ATPase MalK (aussi appelée MsmK-like) est interchangeable avec une autre ATPase MsmK dont le gène est situé dans une autre région du chromosome. Au vu des homologies très élevées (>70% d'identité) chez les autres espèces de Streptocoques, l'ATPase MsmK assure très probablement l'apport d'énergie à l'importeur ABC des maltodextrines chez les autres espèces.

Les polysaccharides de grandes tailles ainsi importés seront pris en charge par une glycogène phosphorylase GlgP (également appelée MalP chez certaines souches). Cette protéine avait à l'origine été décrite comme étant capable de phosphoryler le maltose chez la bactérie *S. bovis* (depuis renommée *S. equinus*) car son activité avait été déduite d'un extrait cellulaire et non de la protéine purifiée<sup>(213)</sup>. Par la suite, cette protéine GlgP a été décrite chez *S. mutans* comme uniquement capable de phosphoryler efficacement les maltodextrines d'une longueur égale ou supérieure à 4 résidus glycosyl<sup>(214)</sup>. Son action libère une molécule de G1P et une maltodextrine raccourcie d'un résidu glycosyl.

Pour cataboliser les molécules plus courtes, c'est-à-dire le maltose et le maltotriose, une

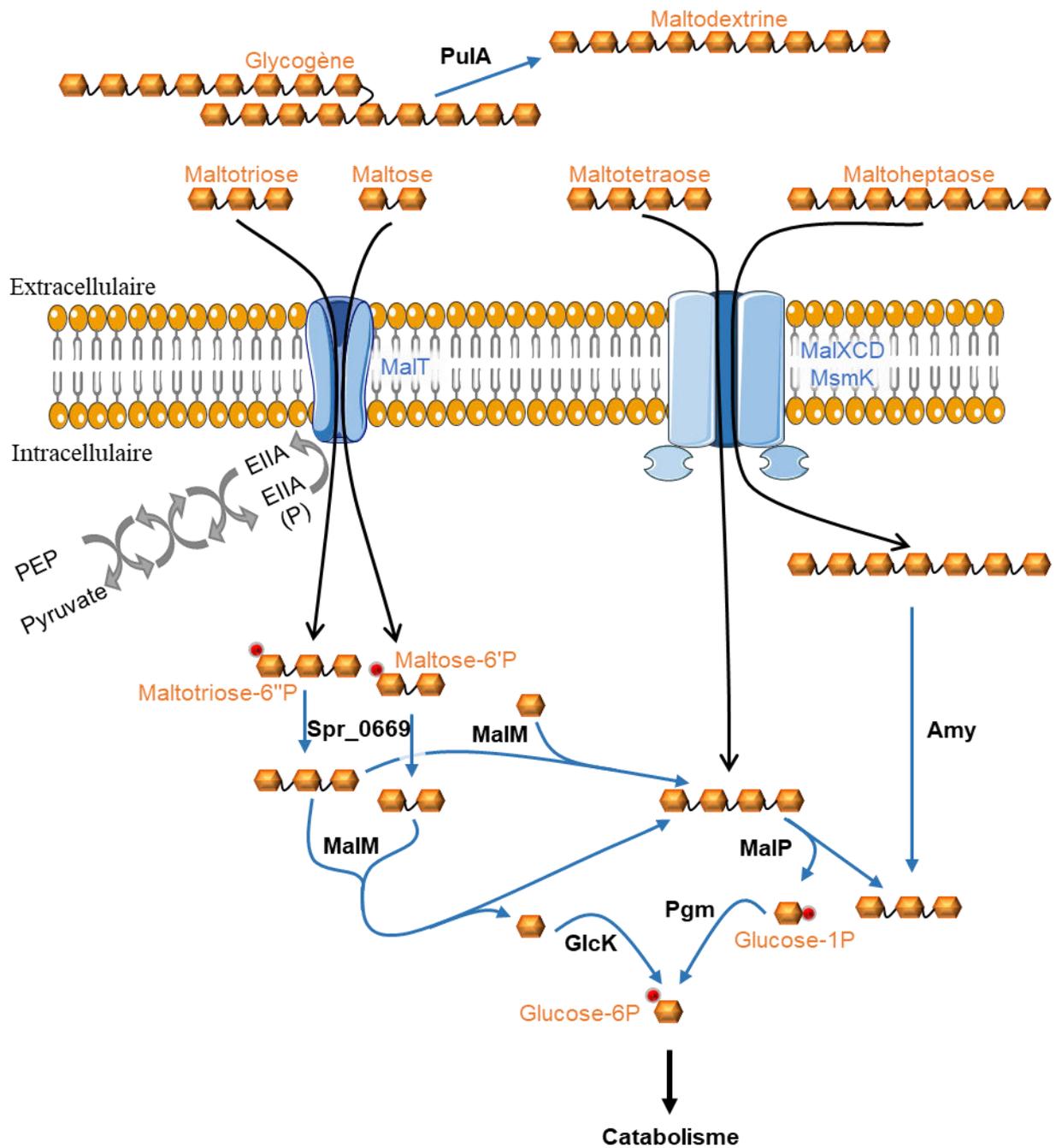
amylomaltase MalM (ou MalQ selon les espèces et les souches) présentant une activité  $\alpha$ -1,4glucanotransferase est nécessaire <sup>(214)</sup>. Cette dernière permet de transférer un résidu glycosyl d'une maltodextrine vers une autre maltodextrine sans faire de distinction de leur longueur. Le maltose ainsi que le glucose peuvent ainsi être utilisés comme molécule "donneuse" ou "acceptrice" <sup>(214,215)</sup>. Cette amylomaltase permet par conséquent la libération de molécules de glucose lors du métabolisme du maltose même si cette activité reste minoritaire. Cette activité a été interprétée à l'origine comme le résultat de la dégradation du maltose par une maltose hydrolase <sup>(213)</sup> mais il semblerait que les Streptocoques ne possèdent pas une telle enzyme. Etant donné la proximité systématique dans le génome des gènes de la glycogène phosphorylase et de l'amylomaltase, on peut supposer que tous les streptocoques dégradent les maltodextrines selon un procédé explicité par Sato *et al* comme décrit ci-dessus <sup>(214)</sup>. Le G1P libéré par l'action de la glycogène phosphorylase ne pourra pas être métabolisé en l'état et sera converti en G6P par l'action d'une phosphoglucomutase Pgm exprimée de manière constitutive <sup>(213)</sup>. Le rôle de cette enzyme est double car elle est essentielle pour connecter le catabolisme des maltodextrines aux voies de dégradation classiques mais a également été décrite comme importante dans la synthèse de la capsule chez certaines souches de Streptocoques impliquant un lien avec la virulence de ces dernières <sup>(216,217)</sup>.

D'autres protéines ont également été décrites dans le contexte du métabolisme des maltodextrines notamment une pullulanase Pula (aussi nommée SpuA) ancrée dans la membrane par un motif LPxTG et exposée au niveau du milieu extérieur <sup>(210,218)</sup>. Cette pullulanase permet d'hydrolyser les liaisons  $\alpha$ 1-6 situées au niveau des ramifications de molécules constituées de résidus glycosyl liés par des liaisons  $\alpha$ -1,4 comme dans le cas du glycogène par exemple. Les maltodextrines linéaires libérées pourront ensuite être importées par l'importeur ABC ou le PTS MalT décrits précédemment en fonction de la longueur de leur chaîne. Une autre protéine d'intérêt, la glucosidase DexB, permet l'hydrolyse des liaisons glycosidiques  $\alpha$ 1-6 au niveau intracellulaire et permet ainsi une dégradation des ramifications des polymères de maltodextrines <sup>(219,220)</sup>. Les Streptocoques possèdent également plusieurs  $\alpha$ -amylases, une première enzyme cytoplasmique produisant majoritairement du maltotriose à partir de chaînes linéaires de maltodextrine a été documentée chez *S. equinus* <sup>(221)</sup> et chez *S. mutans* <sup>(222)</sup>. Une  $\alpha$ -amylase librement sécrétée a également été décrite chez *S. pyogenes* <sup>(223)</sup>, elle est capable de modifier les maltodextrines du milieu extracellulaire pour produire des  $\beta$ -cyclodextrines (maltodextrines cycliques composées de 6 résidus glycosyl) et a été identifiées par plusieurs études comme facteur de virulence <sup>(224,225)</sup> mais n'est peut-être pas directement

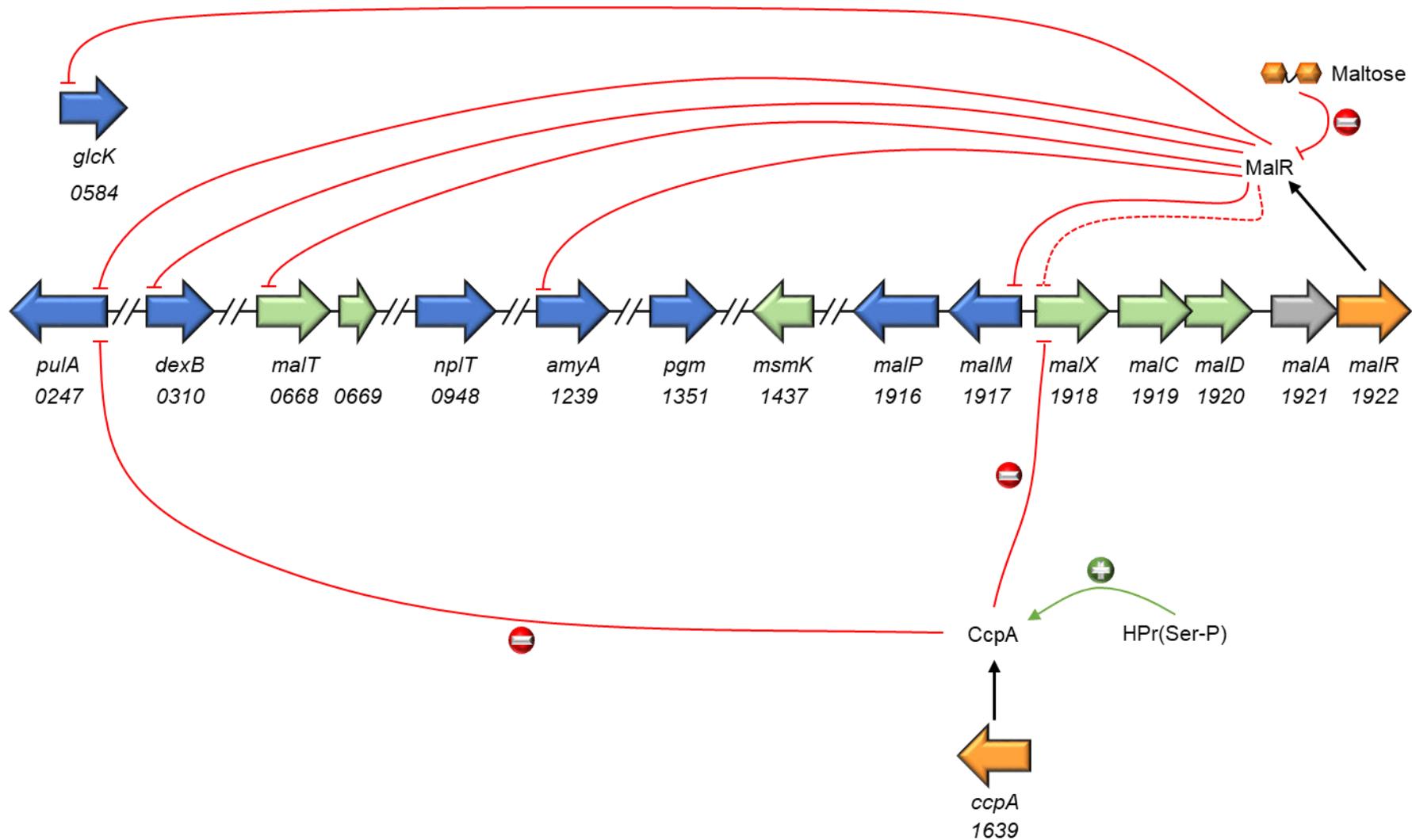
reliée au catabolisme de ces sucres <sup>(223)</sup>. Un schéma récapitulant le métabolisme des maltodextrines chez *S. pneumoniae* est présenté en **Figure 31**.

## 2. Régulation

Pour plus de clarté, nous illustrerons dans cette partie la régulation des gènes des maltodextrines chez les Streptocoques en se référant à l'espèce modèle *S. pneumoniae*. Contrairement à la régulation qui opère chez *E. coli* décrite précédemment, c'est un répresseur transcriptionnel de la famille LacI/GalR qui assure le rôle principal dans la régulation des gènes du maltose et des maltodextrines chez *S. pneumoniae*. La protéine MalR est en effet capable de fixer des régions d'ADN situées en amont des opérons *malXCD* et *malMP*, mais a plus d'affinité pour le promoteur de l'opéron *malMP* que pour celui de l'opéron *malXCD* <sup>(226,227)</sup>. Ces résultats impliquent que l'expression des gènes de l'amylomaltase et de la glycogène phosphorylase impliquées dans le catabolisme des malto-oligosaccharides est plus fortement verrouillée que celle des gènes codants pour le transporteur de ces sucres. La présence de maltose empêche la fixation du répresseur MalR sur l'ADN et agit donc comme l'inducteur de ces deux opérons <sup>(226)</sup> à l'instar de l'allolactose (Galactose-( $\beta$ 1-6)-Glucose) pour le répresseur LacI d'*E. coli* <sup>(228)</sup>. Le régulon de MalR a ensuite été élargi chez *S. pyogenes* avec les gènes de l'opéron débutant par le gène *malT* ainsi qu'aux gènes *pulA* et *dexB* <sup>(229)</sup>. Une étude a également montré que le régulateur pléiotrope CcpA joue un rôle dans la régulation des gènes *mal*, il favoriserait l'induction de ces gènes et exercerait une compétition avec MalR pour les mêmes sites de fixation à l'ADN <sup>(206,230)</sup>. Ce même régulon de MalR a également été décrit chez *S. pneumoniae* <sup>(231)</sup>. MalR est donc capable de réprimer l'expression de tous les gènes présentés dans la partie précédente à l'exception des gènes de la phosphoglucomutase *pgm* et de l'ATPase *msmK*. MalR semble également réguler l'expression du gène *glcK* codant pour la glucokinase chez *S. pneumoniae* <sup>(231)</sup>. Chez cette espèce, le régulateur global CcpA en complexe avec son cofacteur protéique P-Ser-HPr agit comme un répresseur direct sur l'expression des gènes *malXCD* et *pulA* <sup>(231)</sup>. Les auteurs ont déterminé une séquence consensus comme motif de fixation de MalR (CGCAAACGTTTGCR avec R = A ou G) qui apparaît proche des séquences *cre*, les séquences de reconnaissance de CcpA expliquant potentiellement l'implication de CcpA dans la régulation de certains de ces gènes. Ce réseau de régulation est schématisé dans la **Figure 32**.



**Figure 31** : Schéma du métabolisme du maltose et des maltodextrines chez *S. pneumoniae*  
 Les flèches noires représentent des transports de molécules et les bleues des transformations enzymatiques.  
 La cascade de transferts de groupements phosphates a été simplifiée dans un souci de clarté.

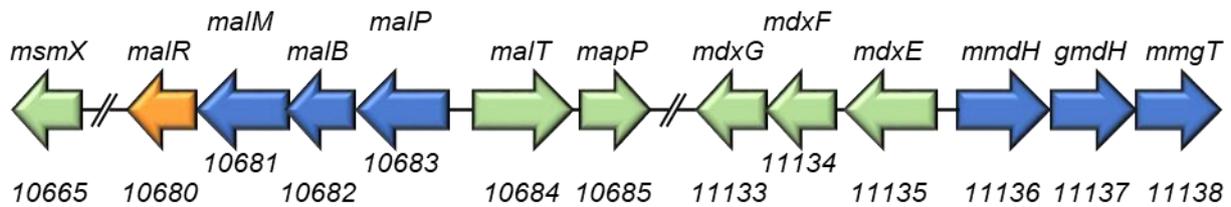


**Figure 32** : Modèle de régulation du métabolisme du maltose et des maltodextrines chez *S. pneumoniae*  
 Les flèches rouges représentent des inhibitions de transcription et/ou d'activité et les vertes symbolisent des activations.  
 Les gènes respectent les mêmes codes que la **Figure 30**.

## C. Enterococcus faecalis

### 1. Transport et catabolisme

Le métabolisme des maltodextrines chez les Entérocoques est très similaire à ce qui est documenté chez les Streptocoques et qui a été décrit ci-dessus mais présente toutefois certaines différences. Afin de métaboliser le maltose, *E. faecalis* utilise deux opérons divergents (**Figure 33**).



**Figure 33** : Gènes impliqués dans le métabolisme du maltose et des maltodextrines chez *E. faecalis*  
Les numéros de gènes inscrits en dessous des flèches correspondent font référence au génome d'*E. faecalis* OG1RF (OG1RF\_xxxxx).

Le sens des flèches reflète le sens de transcription et la couleur la fonction du gène. Les flèches vertes représentent les gènes fonctionnellement liés au transport de substrat, les bleues les gènes liés au catabolisme du substrat et la orange le gène lié à la régulation.

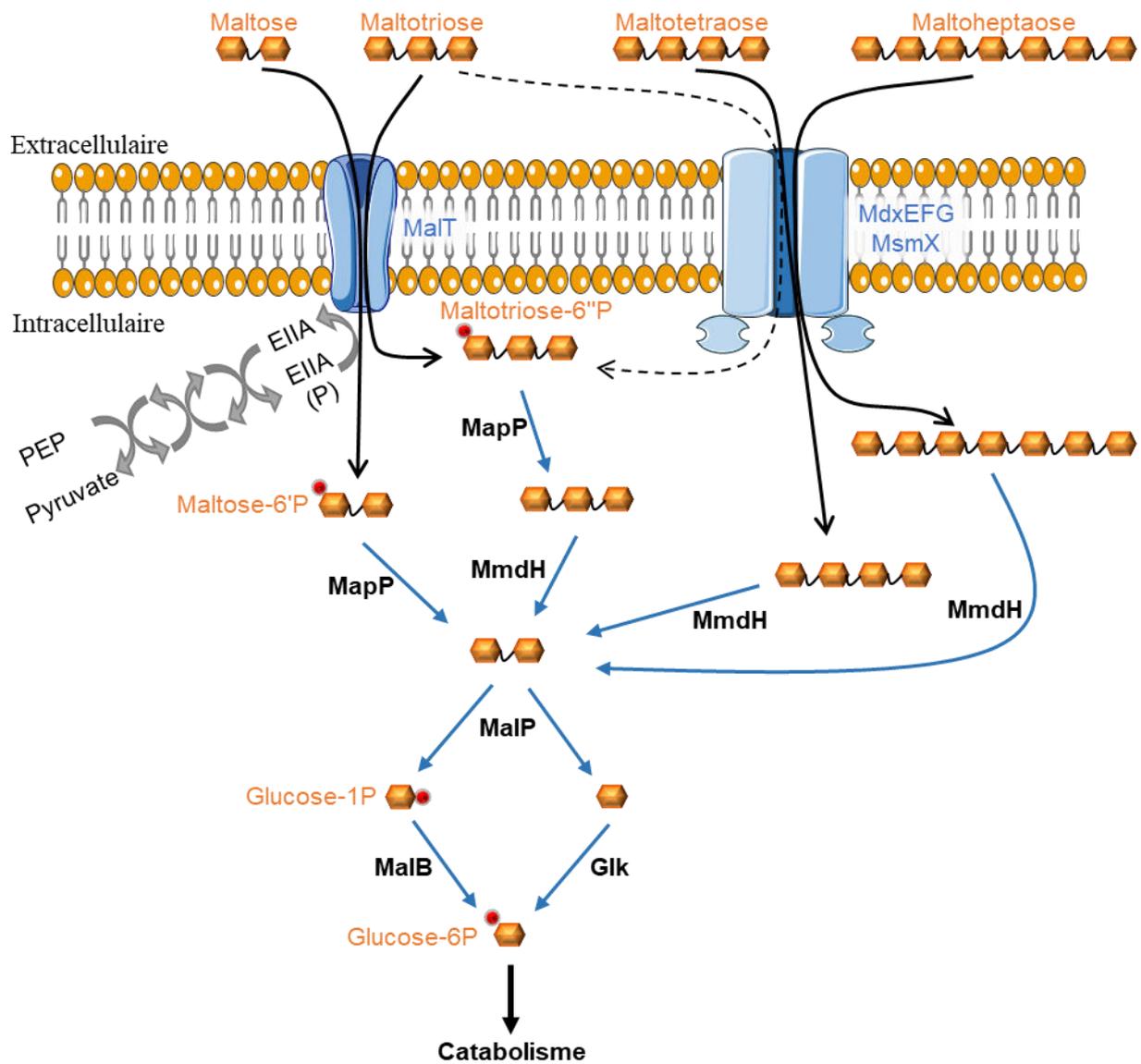
L'un de ces opérons est dédié à l'import du maltose par un système PTS, l'enzyme EIICBA de ce transporteur est codée par le gène *malT* (<sup>108</sup>). Le maltose arrive donc dans la cellule sous la forme de maltose-6'-P (M6P) mais ne peut pas être métabolisé sous cette forme. Une maltose-6'-P phosphatase MapP codée par le gène en aval du gène *malT* va ensuite retirer le groupement phosphate pour reformer du maltose (<sup>208</sup>) qui pourra être phosphorylé par une maltose phosphorylase codée par le premier gène du second opéron (<sup>108</sup>). Cette étape aboutit à la formation d'une molécule de G1P et d'une molécule de glucose. Le G1P devra être converti en G6P par l'action d'une phosphoglucomutase afin d'être métabolisé par la suite. Le gène *malB* en aval du gène *malP* est prédit pour assurer ce rôle au vu des fortes homologues avec d'autres phosphoglucomutases. La molécule de glucose pourra être phosphorylée par la glucokinase GlcK (OG1RF\_12147) afin d'être catabolisée pour produire de l'énergie. Le troisième gène de l'opéron, *malM*, est prédit pour permettre la synthèse d'une aldose épimérase (aussi appelé mutarotase) ayant pour rôle de convertir de manière réversible l' $\alpha$ -D-glucose en son anomère  $\beta$ -D-glucose. Des phosphoglucomutases ont déjà été documentées comme étant spécifiques d'un anomère particulier ( $\alpha$ -D-glucose ou  $\beta$ -D-glucose) (<sup>232-234</sup>), on peut donc supposer que MalB présente également une spécificité stricte pour un anomère du D-glucose, expliquant ainsi la présence du gène *malM*. Le scénario le plus probable est que la phosphorylase MalP produise du  $\beta$ -glucose-1-P à partir de maltose à l'instar de la maltose phosphorylase MalP décrite chez

*Lactobacillus acidophilus* <sup>(235)</sup>. La protéine MalB doit donc permettre la formation d' $\alpha$ -glucose-1-P. Ces gènes n'ont cependant pas fait l'objet d'études précises et leurs fonctions restent donc purement prédictives.

Le système PTS MalT ne permet pas d'importer efficacement les maltodextrines composées de plus de 4 résidus glycosyl. Le maltotétraose et les maltodextrines plus longues sont principalement importées par un importeur ABC de la sous-famille CUT1 codé par l'opéron *mdxEFG*, l'énergie de ce système est apportée par l'hydrolyse d'ATP grâce à l'ATPase MsmX <sup>(21)</sup>. L'opéron placé directement en amont des gènes *mdxEFG* et en orientation opposée code pour trois enzymes impliquées dans la dégradation des maltodextrines. La première de ces enzymes est la maltodextrine hydrolase MmdH <sup>(236)</sup>, elle permet de libérer une unité de maltose du côté non réducteur en hydrolysant la liaison O-glycosidique entre le deuxième et le troisième résidu glycosyl. Le maltose ainsi formé sera pris en charge par la maltose phosphorylase MalP de l'opéron maltose décrite précédemment. Le deuxième gène de l'opéron code pour la protéine GmdH qui possède une activité  $\alpha$ 1-6 glucosidase, elle est capable d'hydrolyser la liaison  $\alpha$ 1-6 entre deux résidus glycosyl <sup>(236)</sup>. La dernière protéine codée par cet opéron est la protéine MmgT dont le rôle physiologique reste pour l'instant inconnu, elle possède *in vitro* une activité glycosyl transférase. Elle est en effet capable de transférer un résidu glycosyl d'une maltodextrine "donneuse" vers une autre maltodextrine "acceptrice" en hydrolysant la liaison glycosidique  $\alpha$ -1,4 entre le premier et le deuxième résidu glycosyl du côté non réducteur <sup>(236)</sup>. Le métabolisme du maltose et des maltodextrines chez *E. faecalis* nécessite par conséquent une utilisation concertée de plusieurs des opérons présentés suivant la longueur des maltodextrines. Le maltotriose représente néanmoins un cas particulier, il peut en effet être importé par le transporteur ABC MdxEFG-MsmX mais aussi de manière plus efficace par le PTS MalT. Une fois à l'intérieur de la cellule, il sera ensuite hydrolysé par MmdH. La molécule de maltose résultante sera finalement prise en charge par la maltose phosphorylase MalP avant de rejoindre l'une des grandes voies de dégradation du glucose. Un schéma illustrant ces métabolismes est proposé en **Figure 34**.

## 2. Régulation

La régulation des opérons présentés ci-dessus impliqués dans le métabolisme du maltose et des maltodextrines chez *E. faecalis* sera discutée plus en détails dans la suite de ce manuscrit. Cette régulation a cependant déjà été documentée chez *E. faecium* E1162. Chez cette souche, le métabolisme du maltose est similaire à celui décrit précédemment pour *E. faecalis* mais le métabolisme des maltodextrines d'une longueur supérieure ou égale à 3 résidus glycosyl est

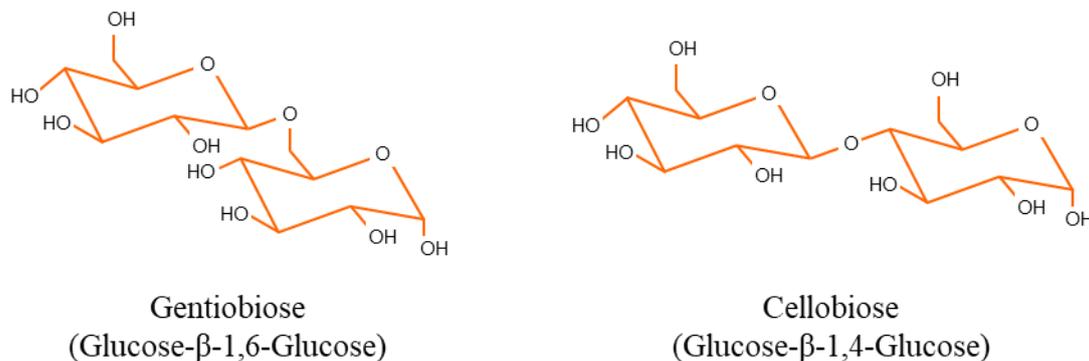


**Figure 34 :** Schéma du métabolisme du maltose et des maltodextrines chez *E. faecalis*  
 Les flèches noires représentent des transports de molécules et les bleues des transformations enzymatiques.

différent. Le métabolisme des maltodextrines chez *E. faecium* E1162 semble homologue au système présent chez *Listeria monocytogenes* (237), un importeur ABC MdxDCB permet l'import des maltodextrines qui seront dégradées par une pullulanase PulA capable d'hydrolyser les liaisons glycosidiques  $\alpha$ -1,4. Ces gènes sont sous le contrôle d'un régulateur transcriptionnel de la famille LacI/GalR appelé MdxR codé par un gène distant sur le génome agissant sur ces gènes comme un activateur (237). Une souche mutante pour le gène *mdxR* voit l'expression des gènes du maltose inchangée, MdxR semble par conséquent seulement impacter l'expression des gènes impliqués dans le métabolisme des maltodextrines plus longues. Après une recherche chez *E. faecalis*, il ne semble pas y avoir de gène homologue à *mdxR*, ce qui laisse supposer un moyen de régulation différent entre les deux espèces.

## V. Le métabolisme du gentiobiose et autres $\beta$ -glycosides chez différents organismes modèles

Le gentiobiose est un disaccharide composé de deux résidus glycosyl à l'image du maltose présenté dans la partie précédente. La liaison glycosidique entre ces deux monosaccharides est cependant très différente. Dans le cas du gentiobiose, les résidus sont reliés par une liaison  $\beta$ -1,6. L'utilisation du gentiobiose par les bactéries est assez peu documentée dans la littérature, son métabolisme semble toutefois régulièrement relié à celui du cellobiose, un autre  $\beta$ -glucoside (238). La structure de ces deux sucres est présentée dans la **Figure 35**.

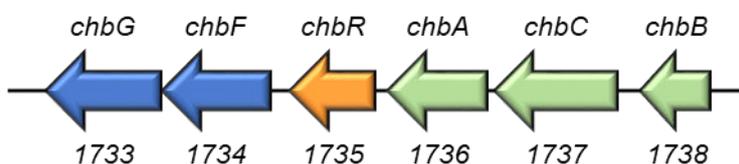


**Figure 35** : Représentation des  $\beta$ -glycosides gentiobiose et cellobiose en conformation chaise

### A. *Escherichia coli*

La bactérie modèle *E. coli* est généralement considérée comme ne pouvant pas utiliser le cellobiose comme seule source de carbone. Des études ont cependant montré que des cellules d'*E. coli* étaient capables d'acquérir la capacité de métaboliser ce disaccharide si elles étaient placées sur un milieu de culture contenant du cellobiose pendant plusieurs semaines (239,240). Ces bactéries possèdent en réalité un opéron leur permettant d'utiliser le cellobiose pour

produire de l'énergie mais ce dernier n'est pas induit par ce sucre. Il a par conséquent longtemps été considéré comme cryptique. Ce type de gènes serait activé par des événements génétiques comme des recombinaisons ou des insertions/délétions/substitutions (<sup>241,242</sup>). L'opéron concerné dans ce cas est l'opéron *chbBCARFG* composé de 6 gènes (**Figure 36**).

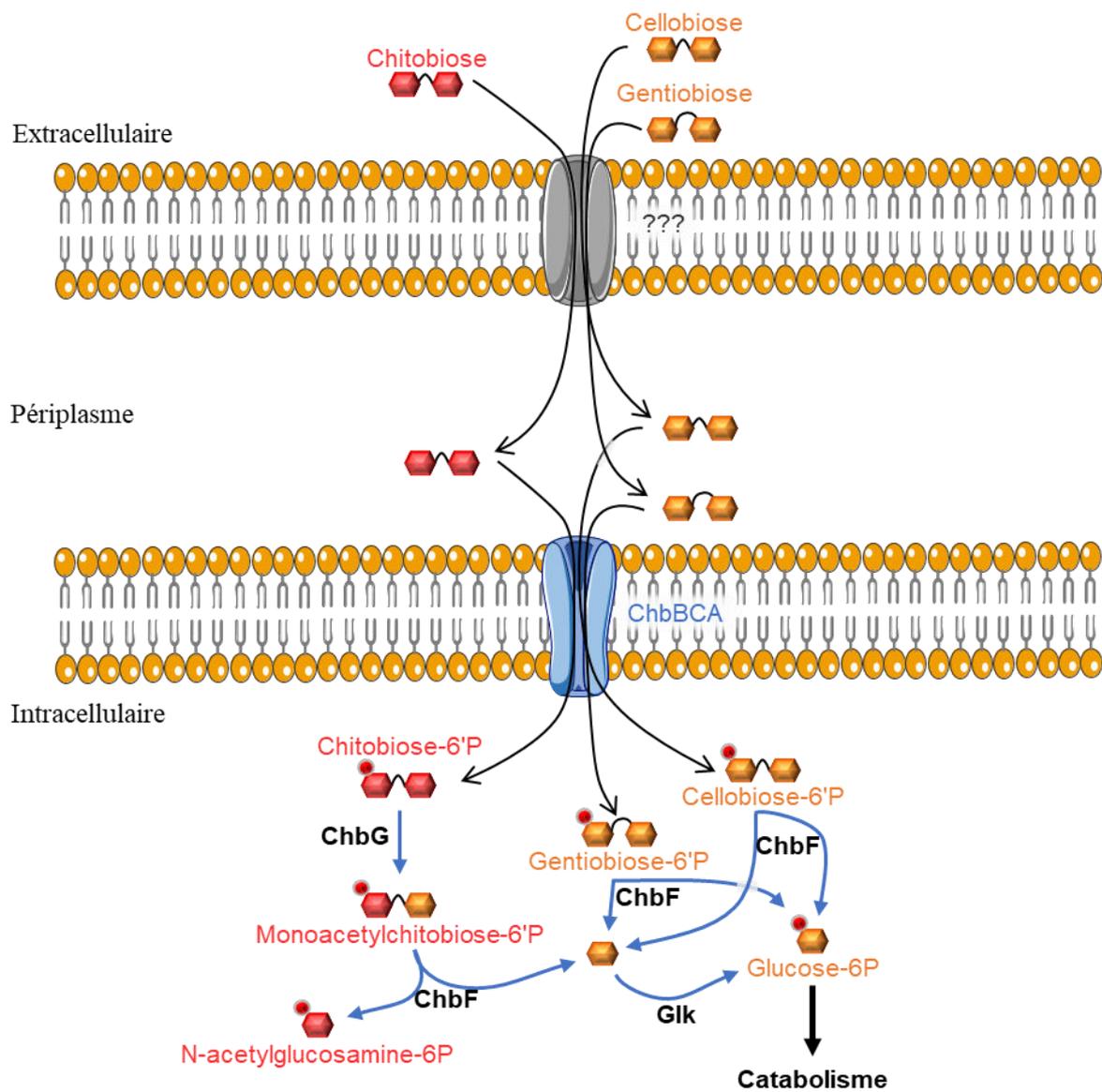


**Figure 36** : Gènes impliqués dans le métabolisme de différents  $\beta$ -glycosides chez *E. coli*  
 Les numéros de gènes inscrits en dessous des flèches correspondantes font référence au génome d'*E. coli* K-12 MG1655 (*b\_*xxxx).

Le sens des flèches reflète le sens de transcription et la couleur la fonction du gène. Les flèches vertes représentent les gènes fonctionnellement liés au transport de substrat, les flèches bleues les gènes liés au catabolisme du substrat et la flèche orange le gène lié à la régulation.

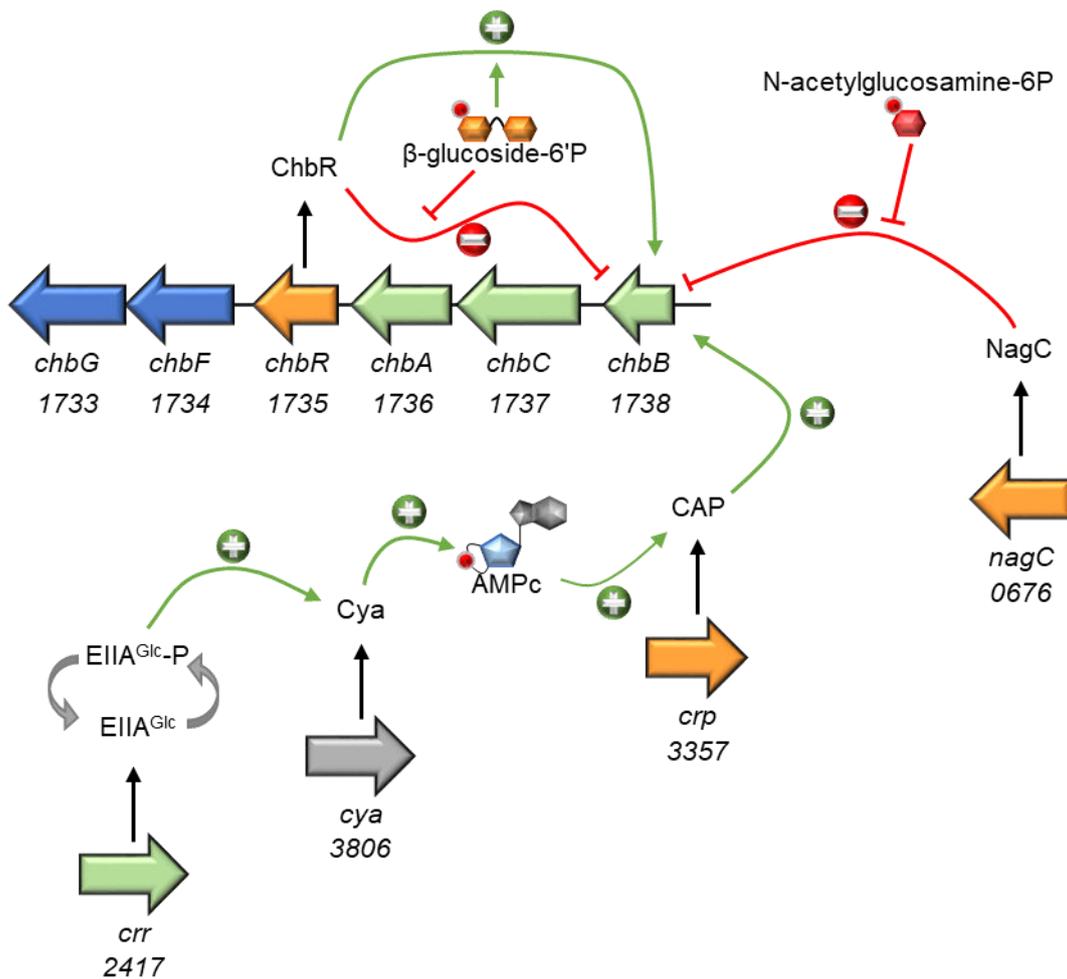
Cet opéron *chb* code chez cette bactérie tous les éléments nécessaires pour prendre en charge plusieurs  $\beta$ -glucosides dont au moins le cellobiose, le gentiobiose, le chitosane (glucosamine  $\beta$ -1,4-glucosamine) et le chitobiose dimère (N-acetylglucosamine  $\beta$ -1,4-N-acetylglucosamine) (la structure de ces  $\beta$ -glucosides est présentée en **Figure S.1** (<sup>243</sup>). Le chitobiose dimère est probablement le substrat naturel de cet opéron car il est le seul à induire son expression. Cet opéron n'est donc pas cryptique, il l'a été considéré à tort car l'inducteur n'avait pas été découvert. Les gènes *chbBCA* codent pour un PTS de la famille cellobiose ( $\beta$ -glucoside) permettant l'import de ces disaccharides sous leur forme phosphorylée. La liaison glucosidique est ensuite clivée par la Phospho- $\beta$ -glucosidase ChbF (<sup>244</sup>). Le dernier gène de l'opéron, *chbG* code pour une déacetylase qui retire le groupement acetyl du résidu du côté réducteur dans le cas du métabolisme du chitobiose (<sup>245</sup>). Un schéma récapitulatif de ce métabolisme est présenté en **Figure 37**.

Cet opéron est directement régulé par trois acteurs protéiques (<sup>246</sup>). Il est dans un premier temps activé par la protéine CAP que nous avons déjà décrite dans la régulation du métabolisme du maltose chez *E. coli*. Pour que CAP soit active, le milieu extracellulaire ne doit pas contenir de glucose disponible. Le deuxième acteur est le régulateur transcriptionnel NagC qui va jouer le rôle d'un répresseur sur l'expression de cet opéron. Cette répression exercée par NagC est levée par la présence de N-acetylglucosamine-6-P qui peut être obtenu par le clivage du chitobiose-6-P par l'action de ChbF. La dernière protéine impliquée dans la régulation de cet opéron est la protéine ChbR codée par le 5<sup>e</sup> gène de l'opéron *chb*. Elle agit comme un répresseur en absence d'inducteur. Cependant, en présence d'un  $\beta$ -disaccharide phosphorylé, ChbR va au



**Figure 37 :** Schéma du métabolisme de  $\beta$ -glucosides chez *E. coli*  
 Les flèches noires représentent des transports de molécules et les bleues des modifications enzymatiques.  
 Pour plus de clarté, la cascade de phosphorylations des acteurs protéiques du PTS n'est pas représentée.

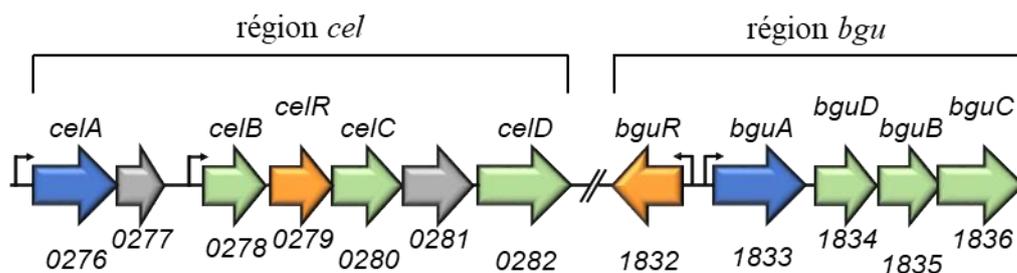
contraire stimuler l'expression de cet opéron. Plusieurs conditions doivent ainsi être réunies pour que cet opéron puisse être exprimé. Tous ces prérequis peuvent expliquer pourquoi cet opéron a longtemps été considéré comme cryptique. Un schéma de régulation de cet opéron est présenté en **Figure 38**.



**Figure 38** : Modèle de régulation du métabolisme de  $\beta$ -glucosides chez *E. coli*  
 Les traits rouges représentent des inhibitions de transcription et/ou d'activité, les flèches vertes symbolisent des activations. Les gènes respectent les mêmes codes que la **Figure 36**.

## B. *Streptococcus pneumoniae*

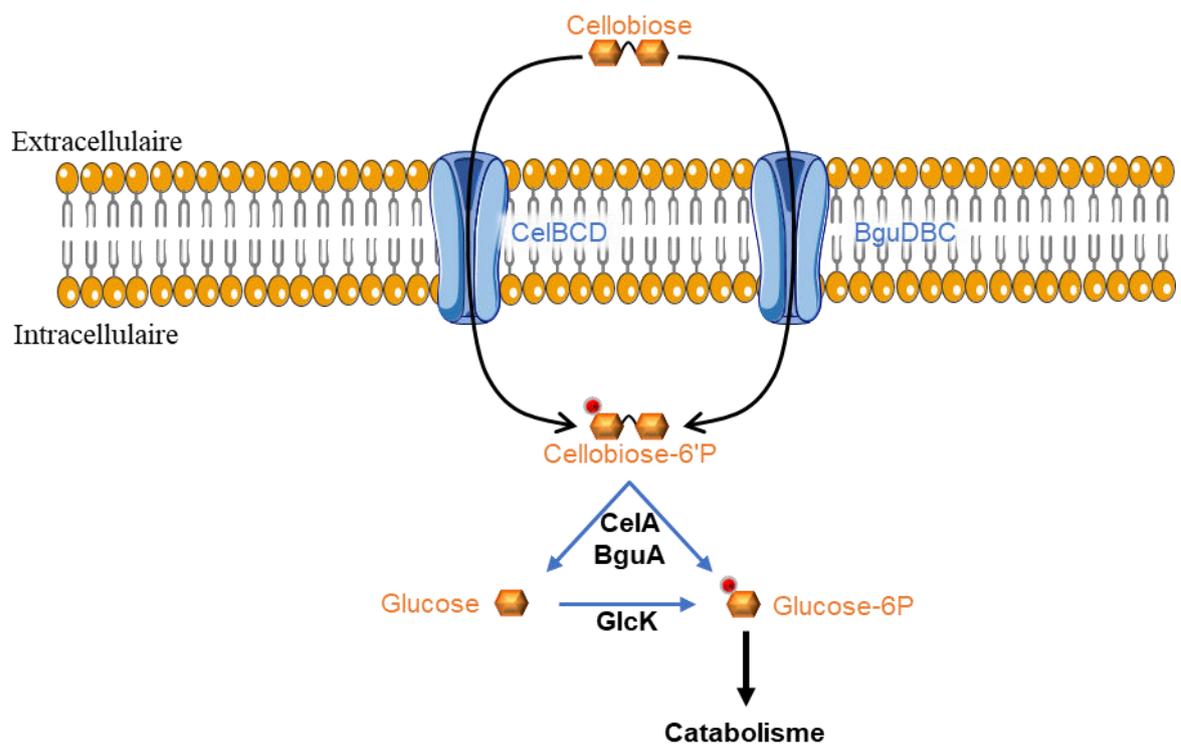
Le métabolisme du cellobiose chez *S. pneumoniae* a été étudié et présente quelques particularités. Deux loci impliqués dans ce métabolisme ont été identifiés chez cette espèce bactérienne. Un premier locus *cel* a été montré comme essentiel pour le métabolisme de ce sucre mais il n'est pas retrouvé chez toutes les souches de cette espèce. Un second locus *bgu* est lui retrouvé chez toutes les souches de *S. pneumoniae* mais ne semble pas essentiel pour ce métabolisme <sup>(247)</sup>. Ces deux régions génétiques sont présentées en **Figure 39**.



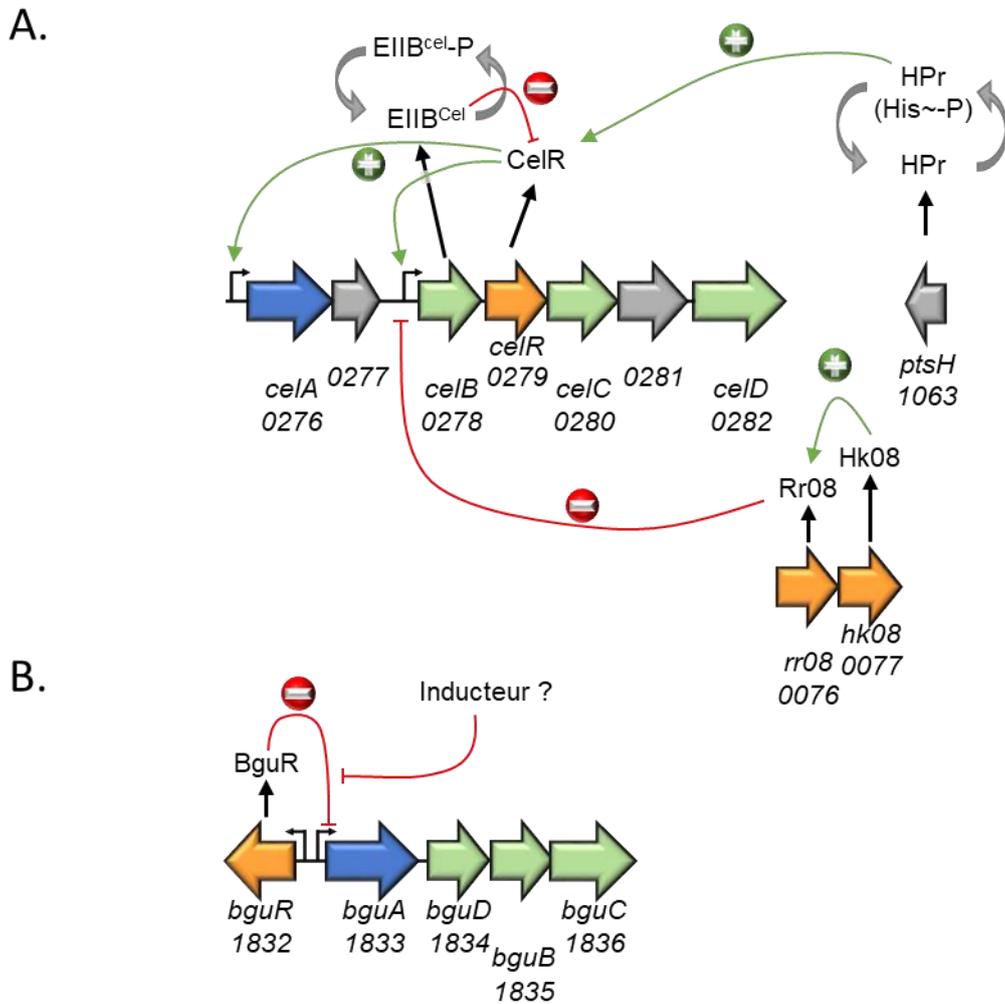
**Figure 39** : Gènes impliqués dans le métabolisme du cellobiose chez *S. pneumoniae*  
 Les numéros de gènes inscrits en dessous des flèches correspondantes font référence au génome de *S. pneumoniae* R6 (*spr\_xxxx*).  
 Le sens des flèches reflète le sens de transcription et la couleur la fonction du gène. Les flèches vertes représentent les gènes fonctionnellement liés au transport de substrat, les flèches bleues les gènes liés au catabolisme du substrat, les flèches orange les gènes liés à la régulation et les flèches grises symbolisent des gènes dont la fonction est inconnue.

Le premier locus est composé de deux opérons constitués d'un total de 7 gènes. Les gènes *celB*, *celC* et *celD* codent pour un PTS complet de la famille cellobiose permettant l'import de ce sucre <sup>(248)</sup>. Le cellobiose-6'-P est ensuite probablement clivé dans la cellule par la 6-Phospho- $\beta$ -glucosidase CelA. Deux gènes *spr0277* et *spr0281* codent pour des protéines transmembranaires hypothétiques dont la fonction reste inconnue. Un mutant pour le gène *spr0281* est cependant incapable d'utiliser le cellobiose comme seule source de carbone <sup>(248)</sup>. Ces deux opérons sont régulés positivement par le régulateur CelR <sup>(249)</sup>. Cette protéine est un régulateur transcriptionnel de la famille Bgl comportant des domaines PRD. Il stimule l'expression des opérons *cel* en agissant comme un antitermineur lorsque le PTS CelBCD est actif <sup>(249)</sup>. Par homologie avec ce même système chez *S. mutans*, une phosphorylation de CelR par CelB inhibe vraisemblablement l'activité de ce régulateur alors que sa phosphorylation par P~His-HPr l'augmente <sup>(250)</sup>. De façon surprenante, l'opéron contenant ce PTS cellobiose semble également régulé négativement par un système à deux composants Hk08/Rr08 même si cette régulation n'est pas entièrement comprise <sup>(248)</sup>.

Le second locus impliqué dans le métabolisme du cellobiose est composé de deux unités de transcriptions pour un total de 5 gènes. D'une manière similaire au système *cel*, l'import du cellobiose est ici assuré par un PTS codé par les gènes *bguDBC* puis le cellobiose-6'-P est probablement hydrolysé par la 6-Phospho- $\beta$ -glucosidase BguA <sup>(247)</sup>. Cet opéron est réprimé par le régulateur transcriptionnel BguR de la famille GntR <sup>(247)</sup>. L'inducteur direct de BguR n'a pas été identifié mais est probablement proche chimiquement du cellobiose. Des schémas récapitulatifs explicitant ce métabolisme ainsi que sa régulation chez *S. pneumoniae* sont présentés respectivement en **Figure 40** et **Figure 41**.



**Figure 40** : Schéma du métabolisme du cellobiose chez *S. pneumoniae*  
 Les flèches noires représentent des transports de molécules et les bleues des transformations enzymatiques.  
 Pour plus de clarté, les cascades de phosphorylations des acteurs protéiques du PTS ne sont pas représentées.



**Figure 41** : Modèle de régulation du métabolisme de  $\beta$ -glucosides chez *S. pneumoniae*  
 Les traits rouges représentent des inhibitions de transcription et/ou d'activité, les flèches vertes symbolisent des activations.  
 Les gènes respectent les mêmes codes que la **Figure 39**.  
**A.** Régulation des gènes de la région *cel*  
**B.** Régulation des gènes de la région *bgu*

### C. *Enterococcus faecalis*

Le métabolisme du cellobiose ou gentiobiose n'est pas décrit chez *E. faecalis*. Cette bactérie apparaît toutefois capable de fermenter ces deux sources de carbones et présente dans son génome plusieurs gènes annotés comme impliqués dans le métabolisme des  $\beta$ -glycosides. La recherche d'homologie par BLASTp à partir de la  $\beta$ -hydrolase CelA de *S. pneumoniae* présentée précédemment permet d'identifier une protéine présentant un fort taux de similitude au sein d'un opéron bicistronique annoté *celBA* chez *E. faecalis*. Le premier gène de cet opéron (*celB*) codant pour une protéine EIIC<sup>cel</sup> présente lui aussi de fortes similitudes avec le gène *celD* de *S. pneumoniae*. La caractérisation de ces gènes fera l'objet d'une analyse plus détaillée dans la suite de ce manuscrit.

# Matériel et méthode

# I. Matériel biologique

## A. Conditions de culture et souches bactériennes

Les compositions des milieux de culture utilisés lors de cette étude sont présentées en annexe dans les **Tableaux S1** à **S6**. Les souches utilisées sont répertoriées dans le **Tableau 6**.

Les cultures de routine d'*Escherichia coli* sont réalisées en milieu LB <sup>(251)</sup> à 37°C sous agitation à 120 rpm. Des antibiotiques peuvent être ajoutés au milieu selon les besoins (Ampiciline 100 µg/mL, Erythromycine 150 µg/mL, Kanamycine 25 µg/mL, Chloramphenicol 15 µg/mL). Les souches Top10 (ThermoFisher Scientific), XL1 blue MRF' Kan Supercompetent (Agilent), NEB5α (New England Biolabs) et EC1000 <sup>(252)</sup> ont été utilisées pour répliquer les plasmides d'intérêt en vue d'une transformation chez *E. faecalis*. Les souches BL21(DE3) <sup>(253)</sup> et M15[pREP4] (Qiagen) ont permis la surexpression des protéines d'intérêt grâce aux vecteurs pET28 (Novagen/Merck-Millipore) et pQE30 (Qiagen) respectivement.

Les cultures de routine d'*Enterococcus faecalis* sont réalisées en milieu M17 <sup>(254)</sup> supplémenté en glucose à 0,5% final (m/v) à 37°C sans agitation. La souche utilisée pour cette étude est la souche OG1RF, elle est présentée plus en détails dans la partie introduction de ce manuscrit. Lorsque l'impact d'une ou plusieurs sources de carbone est étudié sur la physiologie bactérienne, les croissances sont réalisées en milieu ccM17 <sup>(7)</sup> supplémenté avec la source de carbone adaptée.

## B. Plasmides

Les plasmides utilisés ainsi que leurs caractéristiques d'intérêt sont résumées dans le **Tableau 7**.

### 1. pMAD

Le plasmide pMAD <sup>(255)</sup> a été utilisé pour la création de mutants chez la souche OG1RF d'*E. faecalis*. Ce plasmide est documenté comme un vecteur navette d'environ 9,6 kb contenant deux origines de réplication. La première pBR322 permet la réplication du plasmide à un fort nombre de copies chez les bactéries gram négatives et la seconde pET194ts est une origine de réplication thermosensible, elle est fonctionnelle chez les bactéries gram-positives à une température de 30°C mais devient non fonctionnelle lors d'une croissance à plus de 40°C. Le plasmide possède deux gènes de résistances *bla* et *ermC* conférant des résistances aux antibiotiques ampicilline et érythromycine respectivement. Le plasmide possède également le

**Tableau 6 :** Souches bactériennes utilisées dans cette étude

Souche	Caractéristiques ou génotype	Source ou référence
<i>Escherichia coli</i>		
M15 p[REP4]	F-, Φ80ΔlacM15, thi, lac-, mtl-, recA+, Km <sup>R</sup>	Qiagen
NEB5α	fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17	New England Biolabs
BL21(DE3)	fhuA2 [lon] ompT gal (λ DE3) [dem] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5	New England Biolabs
Top10	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ-	Thermo Fisher Scientific
XL1-Blue MRF <sup>+</sup> Kan	Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac [F' proAB lacI <sup>q</sup> ZΔM15 Tn5 (Kan <sup>R</sup> )]	Agilent
<i>Enterococcus faecalis</i>		
OG1RF	Souche sauvage, sans plasmide. Fus <sup>R</sup> , Rif <sup>R</sup>	(73)
<i>malR</i> <sup>-</sup>	Souche mutante pour le gène <i>malR</i> dérivée de la souche parentale OG1RF	Cette étude
Δ <i>ccpA</i>	Souche mutante pour le gène <i>ccpA</i> dérivée de la souche parentale OG1RF	Cette étude
MC	Souche mutante pour les gènes <i>malR</i> et <i>ccpA</i> dérivée de la souche parentale OG1RF	Cette étude
Δ <i>genR</i>	Souche mutante pour le gène <i>genR</i> dérivée de la souche parentale OG1RF	Cette étude

**Tableau 7 : Plasmides utilisés lors de ce travail**

Plasmide	Caractéristiques	Source ou référence
pMAD	Origines de réplication pBR322 et pET194ts. Em <sup>R</sup> , Amp <sup>R</sup> , <i>bgaB</i>	( <sup>255</sup> )
pMAD: <i>malR</i> <sup>-</sup>	Création du mutant <i>malR</i> <sup>-</sup>	Cette étude
pMAD: $\Delta$ <i>ccpA</i>	Création du mutant $\Delta$ <i>ccpA</i>	Cette étude
pUCB300	Origine de réplication pUC. Amp <sup>R</sup> , Erm <sup>R</sup> .	( <sup>256</sup> )
pUCB300: $\Delta$ <i>genA</i>	Création du mutant $\Delta$ <i>genA</i>	Cette étude
pUCB300: <i>genA</i> (V583)	Complémentation du mutant $\Delta$ <i>genA</i>	Cette étude
pUCB300: $\Delta$ <i>genB</i>	Création du mutant $\Delta$ <i>genB</i>	Cette étude
pUCB300: <i>genB</i> (V583)	Complémentation du mutant $\Delta$ <i>genB</i>	Cette étude
pUCB300: $\Delta$ <i>genR</i>	Création du mutant $\Delta$ <i>genR</i>	Cette étude
pUCB300: <i>genR</i> (V583)	Complémentation du mutant $\Delta$ <i>genR</i>	Cette étude
pQE30	Origine de réplication colE1. <i>lacI</i> <sup>q</sup> , Amp <sup>R</sup> , Cm <sup>R</sup> . Etiquette 6xHis à l'extrémité N-terminale pour la purification de protéines	Qiagen
pQE30:HPr	Purification de la protéine 6xHis-HPr	( <sup>115</sup> )
pQE30:GenA	Purification de la protéine 6xHis-GenA	Cette étude
pQE30:GenR	Purification de la protéine 6xHis-GenR	Cette étude
pET28a & pET28b	Origine de réplication pBR322. <i>lacI</i> <sup>q</sup> , Kan <sup>R</sup> . Etiquette 6xHis à l'extrémité N-terminale pour la purification de protéines	Novagen
pET28a:CcpA	Purification de la protéine 6xHis-CcpA	( <sup>257</sup> )
pET28b:MalR	Purification de la protéine 6xHis-MalR	Cette étude
pNZ273	Origine de réplication pSH71 et <i>repA</i> , <i>repC</i> . Cm <sup>R</sup> . Gène rapporteur <i>gusA</i> . Utilisé pour des expériences de fusion transcriptionnelles.	( <sup>258</sup> )
pNZ273: <i>PmalR</i>	Fusion transcriptionnelle, le gène rapporteur <i>gusA</i> sous le contrôle du promoteur du gène <i>malR</i>	Cette étude
pNZ273: <i>PmalR</i> <sup>-</sup>	Fusion transcriptionnelle, le gène rapporteur <i>gusA</i> est précédé de la séquence du promoteur du gène <i>malR</i> en orientation complémentaire.	Cette étude
pRB473	Origine de réplication pE194 et <i>repU</i> . Amp <sup>R</sup> , Cm <sup>R</sup> .	( <sup>259</sup> )
pRB473: <i>malR</i>	Complémentation du mutant <i>malR</i> <sup>-</sup>	Cette étude
pRB473: <i>ccpA</i>	Complémentation du mutant $\Delta$ <i>ccpA</i>	Cette étude
pGEM-T Easy	Origines de réplication pUC et fl. Amp <sup>R</sup> . Compatible avec le TA cloning.	Promega
pGEM-T: <i>PmalP</i>	Amplification du promoteur du gène <i>malP</i>	Cette étude
pGEM-T: <i>PmalT</i>	Amplification du promoteur du gène <i>malT</i>	Cette étude
pGEM-T: <i>PmdxE</i>	Amplification du promoteur du gène <i>mdxE</i>	Cette étude
pGEM-T: <i>PmmdH</i>	Amplification du promoteur du gène <i>mmdH</i>	Cette étude
pGEM-T: <i>gus</i>	Amplification d'un fragment interne du gène <i>gusA</i>	Cette étude

gène *bgaB* codant pour une  $\beta$ -galactosidase exprimée de manière constitutive, cette enzyme permet une discrimination rapide des transformants en présence de X-gal, les colonies possédant le plasmide apparaissant bleues sur le milieu de culture. Le site multiple de clonage est présent en dehors des gènes présentés précédemment, l'addition de fragment d'ADN n'altère donc pas le comportement de ce plasmide. Les événements de recombinaison peuvent être favorisés en faisant varier les températures de croissance ainsi qu'en ajoutant ou en retirant la pression de sélection antibiotique selon les besoins.

## 2. pUCB300

Le plasmide pUCB300 <sup>(256)</sup> est un vecteur uniquement répliatif chez les bactéries gram négatives via une origine de réplication pUC19. Ce petit plasmide suicide de 3,9 kb est également utilisé pour la création de mutants chez *E. faecalis* comme alternative au vecteur présenté précédemment. Deux gènes de résistances *bla* et *ermB* permettent d'exercer une pression de sélection avec les antibiotiques ampicilline et érythromycine respectivement. Le site multiple de clonage (MCS) se trouve au niveau du début du gène *lacZ'* codant pour le peptide  $\alpha$  de la  $\beta$ -galactosidase LacZ permettant ainsi un criblage facilité par  $\alpha$ -complémentation sur un milieu contenant du X-gal.

## 3. pQE30

Le plasmide pQE30 (Qiagen) a été utilisé pour permettre la surexpression des protéines recombinantes HPr, GenA et GenR. Ce vecteur répliatif chez les bactéries gram-négatives possède un gène codant pour une  $\beta$ -lactamase et confère ainsi une résistance à l'ampicilline. Une région codant pour une étiquette poly-histidines (6xHis-tag) est placée sous contrôle d'un promoteur inductible par l'IPTG. Le site multiple de clonage est placé directement en aval du 6xHis-tag et permet ainsi d'insérer la séquence codant la protéine d'intérêt dans le même cadre de lecture que cette étiquette. La protéine ainsi exprimée présentera à son extrémité N-terminale 6 histidines, facilitant ainsi sa purification grâce à une chromatographie d'affinité sur une colonne Ni-NTA.

## 4. pET28a et pET28b

Les plasmides de la série pET28 (Novagen/Merck-Millipore) ont été utilisés pour permettre la surexpression des protéines MalR et CcpA. Ces vecteurs sont répliatifs chez les bactéries gram-négatives grâce à une origine de réplication pBR322. Ces plasmides possèdent un gène *kanR* conférant la résistance à l'antibiotique kanamycine et un gène *lacI* permettant une expression constitutive du répresseur de l'opéron *lac* d'*E. coli*. Le site multiple de clonage est

situé en aval d'une région permettant l'ajout d'une étiquette 6xHis-tag lorsque la séquence codant la protéine d'intérêt est insérée dans le même cadre de lecture. Cette région est sous le contrôle du promoteur du phage T7. La cellule *E. coli* hôte doit donc posséder le gène codant l'ARN polymérase T7 sous contrôle du promoteur lactose pour être inductible par l'IPTG. La protéine d'intérêt ainsi exprimée présentera à son extrémité N-terminale 6 histidines facilitant sa purification grâce à une chromatographie d'affinité sur une colonne Ni-NTA.

## 5. pNZ273

La mesure d'activité du promoteur PmalR a été mesurée par fusion transcriptionnelle en utilisant le plasmide pNZ273<sup>(258)</sup>. Ce plasmide possède une origine pSH71 ainsi que les gènes *repA* et *repC* pour permettre une répllication efficace chez les bactéries gram-positives et gram-négatives. Un gène de résistance au chloramphénicol (*cat*) permet d'assurer une pression de sélection. Le site multiple de clonage est placé en amont du gène *gusA* codant pour une  $\beta$ -glucuronidase servant de gène rapporteur.

## 6. pRB473

Le plasmide pRB473<sup>(259)</sup> a été utilisé pour compléter les souches mutées avec un allèle sauvage. Ce vecteur navette est répllicatif à faible nombre de copies chez les bactéries gram-positives en raison d'une origine de répllication pE194 ainsi que du gène *repU*. Les gènes *bla* et *cat* confèrent respectivement des résistances à l'ampicilline et au chloramphénicol. Le MCS est placé entre deux terminateurs forts rho-indépendants, l'insertion d'un gène fonctionnel avec son promoteur entre ces terminateurs permet de mimer les conditions d'expression d'une souche sauvage et est donc utilisé dans le cadre de complémentations en *trans*.

## 7. pGEM-T Easy

Le plasmide pGEM-T Easy (Promega) utilisé pour le clonage des promoteurs des gènes *malP*, *malT*, *mdxE*, *mmdH* ainsi qu'un fragment interne du gène *gusA* est un vecteur répllicatif chez les bactéries gram-négatives en raison de son origine de répllication pUC ori, ce vecteur possède également une origine de répllication phagique f1 permettant sa répllication en forme simple brin. Ce plasmide confère une résistance aux  $\beta$ -lactamines via le gène *ampR*, il possède aussi le fragment  $\alpha$  du gène *lacZ* au niveau de son site multiple de clonage afin de permettre une discrimination des clones transformés par  $\alpha$ -complémentation. Ce vecteur est commercialisé dans sa forme ouverte avec des extrémités possédant un nucléotide T supplémentaire au niveau des extrémités 3' de chacun des brins afin de faciliter le clonage TA.

## **II. Cellules compétentes et transformations par électroporation**

### **A. *Escherichia coli***

Les cellules d'*E. coli* Top10, XL1 Blue MRF' Kan Supercompetent, NEB-5 $\alpha$  et EC1000 sont rendues électrocompétentes selon la méthode décrite par Dower *et al* <sup>(260)</sup>. Brièvement, après une culture de nuit en milieu LB suivi d'une croissance en milieu SOB à 37°C sous agitation à 120 rpm jusqu'à une DO<sub>600</sub> de 0,8, les cellules subissent quatre lavages successifs avec de l'eau Milli-Q® (Merck-Millipore) en réduisant le volume total de moitié entre chaque lavage. Des étapes de centrifugation permettent de séparer les cellules du milieu liquide. Les cellules sont finalement reprises par du glycérol 10% (v/v) à 4°C. Des aliquotes de 50  $\mu$ l sont conservées à -80°C ou utilisés immédiatement. Les cellules compétentes sont ensuite placées dans des cuvettes "GenePulser® electroporation cuvettes" (Biorad) préalablement refroidies puis un choc électrique est délivré par un appareil GenePulser® (BioRad) selon les paramètres suivants : tension de 2,50 kV, capacitance de 25  $\mu$ F, résistance de 200  $\Omega$ . Immédiatement après le choc, les cellules d'*E. coli* sont reprises dans 1 mL de milieu SOC <sup>(261)</sup> puis régénérées à 37°C pendant 45 min sous agitation à 120 rpm. Les cellules sont finalement étalées sur un milieu LB additionné d'antibiotiques selon les résistances apportées par les plasmides d'intérêt et incubées à 37°C.

### **B. *Enterococcus faecalis***

Les cellules d'*E. faecalis* suivent un traitement visant à fragiliser leur paroi par un traitement aux muramidases avant d'être transformées. Brièvement, un milieu GM17 estensemencé par une culture de nuit réalisée en milieu GM17. Une croissance est réalisée à 37°C jusqu'à une DO<sub>600</sub> de 0,5. Les cellules sont séparées du milieu de culture par une étape de centrifugation puis sont lavées dans du glycérol 10 % (v/v). La paroi bactérienne est fragilisée par 15 minutes d'incubation dans un tampon (10 mM Tris pH 8,0, 50 mM NaCl, 10 mM EDTA, 20 % (m/v) saccharose) comprenant de la mutanolysine (Sigma-Aldrich) et du lysozyme (Sigma-Aldrich) à 10 U/mL et 10  $\mu$ g/mL respectivement. Les cellules sont ensuite lavées par quatre lavages successifs dans un tampon d'électroporation (500 mM saccharose, 10 % glycérol) avant d'être réparties en aliquotes de 50  $\mu$ L puis conservées à -80°C ou utilisées immédiatement. Les cellules compétentes sont ensuite placées dans des "GenePulser® electroporation cuvettes" (Biorad) préalablement refroidies puis un choc électrique est délivré par un appareil GenePulser® (BioRad) selon les paramètres suivants : tension de 2,25 kV,

capacitance de 25  $\mu\text{F}$ , résistance de 400  $\Omega$ . Immédiatement après le choc, les cellules d'*E. faecalis* sont reprises dans 1 mL de milieu GM17 additionné de saccharose (500 mM),  $\text{MgCl}_2$  (20 mM) et  $\text{CaCl}_2$  (2 mM) puis régénérées à 37°C pendant 90 min. Les cellules sont finalement étalées sur un milieu GM17 additionné d'antibiotiques selon les résistances apportées par les plasmides d'intérêt et incubées à 37°C.

### **III. Biologie moléculaire**

#### **A. Réactions de polymérisation en chaîne (PCR)**

Les PCR de vérification sont systématiquement réalisées à l'aide de l'ADN polymérase GoTaq® G2 Flexi DNA Polymerase (Promega) en suivant les recommandations du fournisseur. Les amorces nucléotidiques sont synthétisées chimiquement (Eurofins). Les réactions sont effectuées dans un thermocycleur "Master Cycler" (Eppendorf).

Pour les PCR en amont des clonages nécessitant une grande fidélité, la protéine Phusion High-fidelity DNA polymerase (ThermoFisher Scientific) est utilisée avec le tampon HF en suivant les recommandations du fournisseur.

#### **B. Electrophorèse en gel d'agarose**

La qualité des acides nucléiques est systématiquement contrôlée par électrophorèse en gel d'agarose de 0,8 à 1,2% (m/v), la migration est réalisée dans une cuve Sub-Cell® GT Cell (BioRad) sous un voltage constant de 90 V jusqu'à ce que le front de migration atteigne les 2/3 de la longueur du gel. La taille des fragments d'acides nucléiques est estimée par comparaison avec des marqueurs de masse moléculaire GeneRuler 1 kb DNA Ladder ou GeneRuler 100 pb DNA Ladder (ThermoFisher Scientific) selon les tailles théoriques attendues pour chaque fragment. La présence d'acides nucléiques est révélée par ajout de bromure d'éthidium à 0,35  $\mu\text{g/mL}$ , la lecture est effectuée sur un banc UV.

#### **C. Restriction et ligation de fragments d'ADN**

Les endonucléases de restriction (Promega) sont utilisées en respectant leurs conditions optimales de fonctionnement indiquées par le fournisseur, les digestions d'un même fragment d'ADN par deux enzymes différentes sont réalisées simultanément si un même tampon peut être utilisé sans activité star détectable en se basant sur les informations du fournisseur. Les ligations intermoléculaires sont réalisées grâce à la T4 DNA Ligase (ThermoFisher Scientific) en suivant les indications du fabricant.

## **D. Purification des fragments d'ADN**

Les produits PCR ainsi que les fragments d'ADN issus de réactions de digestion ou de ligation sont purifiés grâce au kit "NucleoSpin Gel and PCR clean-up" (Macherey-Nagel) selon les recommandations du fournisseur avant de subir d'autres traitements.

## **E. Extraction de plasmides d'*E.coli***

Afin d'extraire les plasmides, 5 mL de culture de nuit de souches d'*E. coli* possédant les plasmides d'intérêt en milieu LB sont utilisés. Les cellules bactériennes sont d'abord séparées du milieu de culture par centrifugation avant d'être utilisées, l'extraction est réalisée avec le kit "NucleoSpin Plasmid" (Macherey-Nagel) en suivant les recommandations du fournisseur. La qualité des plasmides est systématiquement vérifiée par électrophorèse sur gel d'agarose comme décrit précédemment.

## **F. Extraction d'ADN génomique**

L'ADN génomique a été extrait à l'aide du kit "NucleoSpin Microbial DNA" (Macherey-Nagel) à partir de 5 mL d'une culture de nuit d'*E. faecalis* en milieu GM17 selon les recommandations du fournisseur. La quantité d'ADN est estimée par spectrophotométrie en mesurant l'absorbance à 260 nm en admettant qu'une unité de DO correspond à une concentration de 50 µg/mL. La pureté de l'ADN est mesurée par la mesure des absorbances à 280 et 230 nm.

## **G. Extraction d'ARN totaux**

Les cellules sont préalablement transférées dans un tube "Lysing Matrix B" (MP Biomedicals) en vue d'être lysées à l'aide d'un appareil FastPrep24 (MP Biomedicals) par deux cycles de 30 sec à 6 mouvements par secondes espacés par 5 minutes de repos sur glace. Les ARN sont ensuite extraits à l'aide du kit "Direct-zol" (Zymo Research) selon les recommandations du fabricant. La quantité d'ARN est estimée par mesure d'absorbance à 260 nm à l'aide d'un Nanodrop2000 (ThermoFisher Scientific) avec la correspondance qu'une unité de DO correspondant à une concentration de 40 µg/mL. La pureté des ARN est estimée par la mesure des ratios 260/280 et 260/230 devant être tous deux supérieurs à 2. L'intégrité des ARN est vérifiée par électrophorèse en gel d'agarose additionné de javel pour obtenir une concentration de chlore actif finale de 0,06% <sup>(262)</sup>.

## H. Réactions de transcription inverse

Les ADNc sont synthétisés par une étape de transcription inverse à partir d'1 µg d'ARN totaux à l'aide du kit "Quantitect Reverse Transcription" (Qiagen) en suivant les recommandations du fournisseur.

## I. RACE-PCR

Les réactions de RACE-PCR sont réalisées à l'aide du kit 5'/3' RACE Kit, 2<sup>nd</sup> Generation (Roche Applied Science) à partir de 2 µg d'ARN totaux selon les recommandations du fournisseur. Les amorces utilisées pour les étapes de transcription inverse (\_SP1), d'amplification par PCR (\_SP2) et de séquençage (\_SP3) sont indiquées dans le **Tableau S.10**.

## J. Fusion transcriptionnelle

Les expériences de fusion transcriptionnelles ont été réalisées avec différentes versions du plasmide pNZ273 (<sup>258</sup>). La mesure de l'activité du promoteur PmalR a été mesurée à l'aide du plasmide pNZ273:malR, pour cela la séquence du potentiel promoteur du gène *malR* a été clonée dans le plasmide pNZ273 en amont du gène *gusA* codant pour une β-D-glucuronidase, la mesure de l'activité de cette enzyme est donc directement liée à l'activité du promoteur cloné. Le plasmide vide pNZ273 ainsi qu'un plasmide contenant le même insert que le plasmide pNZ273:PmalR en orientation inverse sont utilisés comme témoins négatifs.

Pour mesurer les activités β-D-glucuronidase, des cellules *d'E. faecalis* OG1RF contenant les versions des plasmides pNZ273 sont cultivées dans les conditions d'intérêt à 37°C sans agitation. Les cellules sont récupérées par centrifugation puis reprises dans un volume de tampon (100 mM Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub> pH 7.0, 10 mM KCl, 1 mM MgCl<sub>2</sub>). La DO<sub>600</sub> est mesurée pour estimer le nombre de cellules présentes dans l'échantillon puis les cellules sont lysées chimiquement par l'ajout d'un tampon avec les concentrations finales (4,65 % (v/v) chloroforme, 0,002 % (v/v) SDS) puis une incubation à 37°C pendant 5 min. Le substrat de la β-D-glucuronidase (p-nitrophenyl-β-D-glucuronide) est ajouté à hauteur de 627,5 µg/mL puis l'extrait cellulaire est incubé à 37°C pendant 10 minutes. La réaction est stoppée par l'ajout de Na<sub>2</sub>CO<sub>3</sub> à une concentration finale de 250 mM. Les débris cellulaires sont sédimentés par une étape de centrifugation puis la DO<sub>420</sub> est mesurée dans le surnageant afin d'estimer la concentration du p-nitrophenol et ainsi évaluer l'activité de la β-D-glucuronidase. L'absence de débris cellulaires dans le surnageant est vérifiée par une mesure de la DO<sub>550</sub>, cette dernière doit être inférieure à 0,05 pour permettre l'exploitation des données déterminées précédemment. L'activité enzymatique est normalisée par la DO<sub>600</sub> déterminée précédemment.

## K. Empreintes à la DNase I

Les régions d'ADN susceptibles de contenir des sites de reconnaissances pour les régulateurs transcriptionnels MalR ou GenR ont été amplifiées par PCR comme décrit précédemment en utilisant une amorce marquée à son extrémité 5' par une molécule fluorescente FAM (Eurofins). Ces produits PCR sont incubés à température ambiante pendant 15 min avec le régulateur transcriptionnel dans des conditions adéquates (cf paragraphe suivant pour plus de détails). Différentes concentrations de DNase RQ1 (Promega) sont ajoutées au mix précédents puis incubés à température ambiante pendant 5 min. Les réactions de digestion sont stoppées par l'ajout d'EDTA à une concentration finale de 2 mM. Les fragments d'ADN résultants sont finalement purifiés et dilués au demi par l'ajout d'un volume de formamide. L'analyse des fragments a été réalisée par la plateforme de l'institut Cochin (Eurofins) en utilisant un standard interne 500 ROX (ThermoFisher Scientific).

Les interactions protéine-ADN ont été réalisées pour la protéine 6xHis-MalR (1  $\mu$ M) dans un tampon (20 mM Tris pH 7,4, 20 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM FBP, 1 mM DTT, 0,01 mg/mL BSA, 62,5  $\mu$ g/mL d'ADN de sperme de harreng, 5% (v/v) glycérol) et pour la protéine 6xHis-GenR (500 nM) dans un tampon (10 mM Tris pH 7,5, 5mM MgCl<sub>2</sub>, 2 mM CaCl<sub>2</sub>, 1 mM DTT, 0,125 mg/mL BSA, 62,5  $\mu$ g/mL d'ADN de sperme de harreng, 5% (v/v) glycérol, 0,25 mM gentiobiose-6'-P).

## L. Séquençage du transcriptome entier (RNA-seq)

Les séquençages de transcriptome entier ont été réalisés à partir d'1  $\mu$ g d'ARN totaux par la plateforme iGE3 de l'Université de Genève (Suisse). Les ARNr ont été éliminés grâce au kit Ribo Zero Magnetic Kit for bacteria (Illumina). Les bibliothèques de séquençages ont été préparées à l'aide du kit Illumina stranded mRNA (Illumina) selon les recommandations du fournisseur. Leur qualité a été validée par Bioanalyzer 2100 (Agilent) et fluorimètre Qubit (ThermoScientific) puis le séquençage a été effectué par un lecteur Illumina HiSeq 2500 en suivant un protocole en lecture simple de 50 cycles.

La qualité des séquençages a été validée par FastQC <sup>(263)</sup>, les séquences obtenues ont ensuite été alignées sur le génome de référence d'*E. faecalis* OG1RF grâce au logiciel TopHat v.2. Le nombre de séquences obtenues par gène a été calculé avec le logiciel HTSeq v.0,6p1 (htseq count), ce nombre a ensuite été normalisé par rapport au nombre total de séquence obtenu afin d'obtenir un niveau d'expression en count per million (cpm). Les gènes dont l'expression était inférieure à 100 cpm dans toutes les conditions testées ont été considérés

comme insuffisamment exprimés ou non induits et n'ont pas été considérés dans les analyses futures.

## **IV. Constructions de mutants par génie génétique et complémentations**

### **A. Mutants par allèles mutés**

#### **1. Mutant *malR*<sup>-</sup>**

Le mutant *malR*<sup>-</sup> a été obtenu en insérant dans le cadre de lecture deux codons stop différents à la place des acides aminés en positions 57 et 60 afin d'obtenir une petite protéine tronquée supposée non fonctionnelle ne possédant que 17,0% de l'extrémité N-terminale de la protéine sauvage correspondante. Pour cela, des mutagenèses dirigées ont été réalisées par PCR en utilisant les amorces dégénérées mut\_malR\_for et mut\_malR\_rev avec les amorces pMAD\_malR\_rev et pMAD\_malR\_for respectivement (**Tableau S.9**). Les deux fragments d'ADN sont ensuite "regroupés" par PCR puis digérés par les endonucléases de restrictions *SalI* et *NcoI* puis le fragment d'ADN obtenu est intégré par une étape de ligation dans le plasmide pMAD préalablement digéré par les mêmes enzymes que précédemment. Le plasmide contenant le fragment d'intérêt est ensuite introduit chez *E. faecalis* par électroporation puis intégré dans le chromosome par recombinaison homologue en cultivant les cellules en présence d'une pression de sélection antibiotique à une température de 42°C ne permettant pas une réplication efficace du plasmide. Le second évènement de recombinaison permettant l'obtention de l'allèle muté est favorisé par des repiquages successifs à 30°C en milieux de cultures ne contenant pas d'antibiotique jusqu'à la perte de résistance. Le bon déroulement de tous les évènements est contrôlé par PCR. Le mutant est ultimement confirmé par un séquençage du génome sur une région d'environ 1,5 kb de part et d'autre de la mutation.

#### **2. Mutants pour les gènes *gen***

Les mutants pour les gènes *gen* ont été obtenus en utilisant une stratégie similaire en insérant également deux codons stop au début de chaque gène afin d'aboutir à des protéines tronquées. Ces codons stop remplacent respectivement les codons 155&160, 113&117 et 60&63 des gènes *genA*, *genB* et *genR*. Les protéines résultantes ne présentent en conséquence que respectivement 32,5%, 25,5% et 23,7% des séquences N-terminales des protéines correspondantes de la souche sauvage.

## B. Mutant de délétion

Le mutant  $\Delta ccpA$  a été obtenu par délétion de la majorité du gène (85%) entre les nucléotides en positions 49 et 899. Des fragments d'ADN correspondant aux régions amont et aval de la zone délétée sont obtenus par PCR en utilisant les couples d'amorces *ccpA<sub>rev</sub>Sal+ccpA<sub>sac</sub>\_rev* et *ccpA<sub>Sac</sub>\_for+ccpA<sub>JH</sub>\_for<sub>Nco</sub>* (**Tableau S.9**). Les deux fragments obtenus ont ensuite été digérés par l'enzyme de restriction *SacI* puis regroupés par une étape de ligation. Le produit obtenu a ensuite été amplifié par PCR avec les amorces situées aux extrémités puis digéré par les endonucléases de restriction *SalI* et *NcoI* avant d'être intégré dans le plasmide pMAD préalablement digéré par les mêmes enzymes. Le plasmide résultant a été par la suite introduit chez *E. faecalis* par électroporation. Les événements de recombinaisons sont favorisés comme décrits précédemment. L'obtention de cellule contenant une région de génome délétée est vérifiée par PCR puis confirmée par un séquençage du génome sur une région d'environ 1,5 kb de part et d'autre de la délétion.

Le double mutant *MalR- $\Delta ccpA$*  (MC) a été obtenu en suivant la procédure décrite précédemment à partir d'un clone de la souche *malR*<sup>-</sup>.

## C. Complémentations des mutants

### 1. Complémentation plasmidique en *trans*

Le plasmide pRB473 a été utilisé pour effectuer des complémentations plasmidiques des mutants *malR*<sup>-</sup> et  $\Delta ccpA$  par l'introduction d'une copie sauvage sur le plasmide dans la souche mutante correspondante. Le plasmide vide a été utilisé pour vérifier que sa seule présence n'influçait pas les phénotypes observés.

#### a. Complémentation du mutant *malR*<sup>-</sup>

Le plasmide utilisé pour effectuer la complémentation de la souche *MalR*<sup>-</sup> est le pRB473:*malR* <sup>(259)</sup>. Le gène *malR* sauvage ainsi que son promoteur interne ont été amplifiés par PCR à partir de l'ADN génomique de la souche OG1RF d'*E. faecalis* puis clonés dans le plasmide pRB473 au niveau de son site multiple de clonage. Le plasmide résultant a ensuite été transformé chez la souche *MalR*<sup>-</sup> d'*E. faecalis*. La présence du plasmide est maintenue par une pression antibiotique en ajoutant du chloramphénicol à une concentration finale de 2  $\mu$ g/mL.

#### b. Complémentation du mutant $\Delta ccpA$

Le plasmide pRB473:*ccpA* <sup>(257)</sup> a été utilisé pour compléter la souche  $\Delta ccpA$ . Le gène *ccpA* ainsi que son promoteur ont été amplifiés par PCR en utilisant l'ADN génomique de la

souche OG1RF d'*E. faecalis* comme matrice. Le fragment obtenu a ensuite été cloné dans le plasmide pRB473 avant d'être transformé chez la souche  $\Delta ccpA$  d'*E. faecalis*. Le plasmide est maintenu dans la cellule par une pression antibiotique en ajoutant du chloramphénicol à une concentration de 2  $\mu\text{g/mL}$ .

## **2. Complémentations génomique en *cis* des mutants *gen*.**

Les complémentations des mutants *gen* ont été réalisées par remplacement allélique des gènes mutés par les gènes correspondants de la souche V583 d'*E. faecalis*. Les protéines résultantes présentent toute une identité de 100% entre les souches OG1RF et V583 mais les gènes présentent de légères variations. Ces quelques substitutions synonymes permettent de vérifier la bonne complémentation par séquençage, la souche complémentée présentant un gène hybride entre les souches OG1RF et V583.

## **V. Méthodes biochimiques**

### **A. Surexpression et purification de protéines présentant une étiquette poly-histidines**

Pour la purification des protéines MalR, CcpA, HPr, GenA et GenR, les gènes d'*E. faecalis* OG1RF codant ces protéines d'intérêt ont été clonés dans des vecteurs d'expression comme indiqué précédemment. Les cellules ont dans un premier temps été cultivées en milieu Terrific Broth (<sup>264</sup>) additionné des antibiotiques correspondant aux résistances apportées par les plasmides à 37°C sous agitation à 120 rpm. Lorsque la  $DO_{600}$  atteint 0.8, de l'IPTG est ajouté à une concentration finale de 0,1 mM pour permettre la surexpression des protéines d'intérêt. Après 3h d'induction à température ambiante sous agitation, les cellules sont récupérées par une étape de centrifugation puis reprise dans un tampon (20 mM  $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HP}_4$  pH 7.6, 1 M NaCl, 100  $\mu\text{g/mL}$  lysozyme, 10 mM imidazole) avant d'être lysées par un appareil FastPrep24 (MP Biomedicals) par deux cycles de 30 sec à 6 mouvements par secondes espacés par 5 minutes de repos sur glace. Les protéines sont séparées des débris cellulaires par une étape de centrifugation à 4500 rpm pendant 30 min à 4°C, le surnageant est ensuite déposé sur une colonne de 0,5 mL de Ni-NTA (Macherey-Nagel) préalablement équilibrée par un tampon (20 mM  $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HP}_4$  pH 7.6, 1 M NaCl, 10 mM imidazole), plusieurs lavages de la colonne sont effectués en augmentant progressivement la concentration en imidazole du tampon jusqu'à 50 mM. Les protéines sont finalement éluées par 2,5 mL de tampon (20 mM  $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HP}_4$  pH 7.6, 1 M NaCl, 500 mM imidazole).

Un échange de tampon (20 mM NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HP<sub>4</sub> pH 7.6, 250 M NaCl, 15 % (v/v) glycérol) ou (50 mM Tris pH7,4, 15% (v/v) glycérol) selon les protéines est finalement effectué grâce à une colonne PG10 Desalting columns (GE Healthcare).

La concentration massique des protéines purifiées est enfin estimée à l'aide du kit Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific) en utilisant une gamme de BSA conformément aux recommandations du fournisseur. Du DTT ainsi que de l'EDTA sont finalement ajoutés selon les besoins à une concentration finale de 1 mM et 0,1 mM respectivement avant de répartir les protéines en aliquotes en vue de leur conservation à -80°C.

## **B. Phosphorylation de HPr *in vitro***

### **1. Phosphorylation de HPr en position sérine46 par HPrK(V267F)**

La protéine HPr a été phosphorylée *in vitro* en position 46 sur l'acide aminé sérine grâce à la protéine HPrK de *L. casei* portant la mutation V267F <sup>(115)</sup>. Cette protéine HPrK(V267F) possède une activité HPr kinase plus importante et ne présente plus l'activité phosphorylase. La phosphorylation se déroule à 37°C pendant 30 min dans un tampon (50 mM Tris pH 7.4, 5 mM ATP, 5 mM MgCl<sub>2</sub>, 1 mM FBP) en présence de 10 µg de HPrK(V267F). La protéine HPrK(V267F) est inactivée thermiquement à 70°C pendant 10 min car la protéine HPr ainsi que sa phosphorylation en position 46 sont thermiquement stables. Le tampon est finalement échangé par une colonne PD10 (GE Healthcare) comme décrit précédemment. La phosphorylation de HPr est vérifiée sur un gel de polyacrylamide en conditions natives comme décrit ultérieurement.

### **2. Phosphorylation de HPr en position histidine15 par EI**

La protéine HPr a été phosphorylée *in vitro* en position 15 sur l'acide aminé histidine grâce à la protéine EI. La phosphorylation se déroule à 37°C pendant 30 min dans un tampon (50 mM Tris pH 7.4, 5 mM PEP, 5 mM MgCl<sub>2</sub>) en présence de 10 µg de protéine EI. Le tampon est échangé par une colonne PD10 (GE Healthcare) comme indiqué précédemment. Une faible portion de la protéine HPr~hP est chauffée à 70°C pendant 10 min pour enlever le groupement phosphate sur l'histidine en position 15. La protéine HPr phosphorylée ainsi que celle chauffée sont déposées sur un gel de polyacrylamide en conditions natives comme décrit ultérieurement dans le but de vérifier le bon déroulement de la phosphorylation.

### C. SDS-PAGE

Suite à la purification des protéines d'intérêt, la qualité des extractions est systématiquement vérifiée par électrophorèse en gel de polyacrylamide additionné de SDS (SDS-PAGE) à l'aide d'un système Mini-Protean® 2 (BioRad). La concentration totale en polyacrylamide est de 12,5 % (v/v) avec un rapport d'acrylamide/bis-acrylamide de 1:37,5 (ITW Reagents). Les protéines sont préalablement dénaturées dans un tampon SDS (75 mM Tris pH 6,8, 7,5 % (v/v)  $\beta$ -mercaptoéthanol, 0,6 % (v/v) SDS, 15 % (v/v) glycérol, 18  $\mu$ g/mL Bleu de Bromophénol) par chauffage à 95°C pendant 10 min. Les protéines sont ensuite chargées dans le gel ainsi qu'un marqueur de poids moléculaire pré-coloré PageRuler Prestained Protein Ladder (TherFisher Scientific) pour permettre une estimation du poids moléculaire des protéines purifiées. Les protéines sont finalement séparées dans un tampon d'électrophorèse (25 mM Tris pH 6,8, 200 mM glycine, 0,1 % (v/v) SDS) sous un ampérage constant de 25 mA. Les bandes correspondant aux protéines sont visualisées en colorant le gel par une solution contenant du Bleu Brillant de Coomassie (30 % (v/v) méthanol, 7,5 % (v/v) acide acétique, 2,5 mg/mL Bleu Brillant de Coomassie R-250).

### D. Native-PAGE

Afin de vérifier la bonne phosphorylation de la protéine HPr soit en position 15 soit en position 46, une électrophorèse en gel de polyacrylamide en conditions natives est réalisée à l'aide d'un système Mini-Protean® 2 (BioRad). La concentration totale en polyacrylamide est de 14 % (v/v) avec un rapport d'acrylamide/bis-acrylamide de 1:37,5 (ITW Reagents). Un tampon de charge (50 mM Tris pH 6,8, 8% (v/v) glycérol, 18  $\mu$ g/mL Bleu de Bromophénol) est ajouté aux protéines pour permettre le dépôt. Les protéines ainsi qu'un marqueur de poids moléculaire pré-coloré PageRuler Prestained Protein Ladder (TherFisher Scientific) sont déposées dans le gel puis séparés dans un tampon d'électrophorèse (25 mM Tris pH 8,3, 192 mM glycine) sous un ampérage constant de 25 mA. Les bandes correspondant aux protéines sont visualisées en colorant le gel par une solution contenant du Bleu Brillant de Coomassie (30% (v/v) méthanol, 7,5% (v/v) acide acétique, 2,5 mg/mL Bleu Brillant de Coomassie R-250). Les formes phosphorylées de HPr (P~His-HPr et P-Ser-HPr) migrent plus rapidement en conditions natives que leur équivalents non phosphorylés. De plus, la liaison phosphoester de la phosphorylation sur la sérine est thermostable contrairement à la liaison phosphoamine de la phosphorylation portée par l'histidine (<sup>265</sup>). Il est donc possible de vérifier si les protéines obtenues sont bien phosphorylées et de déterminer quel acide aminé porte le groupement phosphate.

## **E. Retard sur gel**

### **1. Sonde ADN marquée par un fluorophore**

Pour les expériences de retards sur gel impliquant la protéine MalR, les fragments cibles d'ADN ont été obtenus par PCR en utilisant des amorces marquées au niveau de leur extrémité 5' par une molécule fluorescente Cyanine-5 (Eurofins). Les protéines ainsi que les fragments d'ADN cibles (5 nM) sont mis en contact puis l'interaction est permise par une incubation de 15 minutes à température ambiante dans un tampon (20 mM Tris pH 7,4, 20 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM FBP, 1 mM DTT, 0,1 mM EDTA, 5% (v/v) glycérol, 50 µg/mL ADN de sperme de harreng, 0,01 mg/mL BSA). Les mélanges sont ensuite séparés par une électrophorèse en gel de polyacrylamide à 5 % (v/v) dans une chambre froide à l'obscurité sous une tension constante de 140 V dans un tampon TBE 0,5X. Les bandes d'ADN sont révélées par lecture de fluorescence par un appareil ProXpress (PerkinElmer) avec les réglages (illumination par le dessus, excitation : 625 nm, émission : 680 nm, résolution : 100 µm, temps d'exposition : 120 s).

### **2. Sonde ADN marquée radioactivement**

Les fragments d'ADN cibles utilisés pour les expériences de retards sur gel impliquant la protéine GenR ont été obtenus par PCR en incluant 2 mCi de [ $\alpha$ -<sup>32</sup>P]dATP dans le mélange de réaction. La protéine et les fragments d'ADN (0,1 nM) marqués radioactivement sont mis en contact à température ambiante pendant 15 min dans un tampon (10 mM Tris pH 7,5, 50 mM KCl, 20 mM MgCl<sub>2</sub>, 1 mM DTT, 5% (v/v) glycérol, 100 µg/mL ADN de sperme de harreng, 0,1 mg/mL BSA). Différents sucres phosphorylés ou non ont été ajoutés à ce mélange à 1 mM selon les besoins de l'expérience. L'électrophorèse a ensuite été réalisée en gel d'acrylamide à 8% (v/v) sous une tension constante dans un tampon TBE 0,5X. A la fin de l'électrophorèse, les gels d'acrylamide ont été séchés et la lecture a été effectuée par autoradiographie grâce à un système Cyclone storage phosphor screen (Perkin Elmer).

## **F. Microscale Thermophoresis (MST)**

Pour réaliser les expériences de MST, les protéines 6xHis-MalR et 6xHis-CcpA ont préalablement été marquées de manière covalente avec le fluorophore NT-647-NHS à l'aide du kit Monolith NT™ Protein Labeling Kit RED-NHS (Nanotemper Technologies). Ces protéines ont ensuite été mises en contact à une concentration fixe de 5 nM avec différentes concentrations de protéines testées comme partenaire d'interaction. Une incubation de 15 min à température ambiante dans un tampon (20 mM Tris pH 7,4, 20 mM KCl, 10 mM MgCl<sub>2</sub>,

10 mM FBP, 1 mM DTT, 0,1 mM EDTA, 5% (v/v) glycérol, 0,005% (v/v) Tween20) a été respectée pour permettre les potentielles interactions protéines-protéines. Ce mélange a été chargé dans des capillaires NT.115 premium coated capillaries (Nanotemper Technologies) et les lectures ont été effectuées dans un appareil Monolith NT.115Pico (Nanotemper Technologies) avec les paramètres (LED : 5%, MST : moyen). L'acquisition des signaux de MST a été étudiée entre 4,00 et 5,00 secondes après l'allumage du laser.

## **G. Chromatographie sur couche mince (CCM)**

Les produits de dégradations suite à l'activité de la protéine GenA ont été analysés par chromatographie sur couche mince (CCM). Les différents substrats testés (5 µg) ont été incubés avec 5 µg de la protéine 6xHis-GenA pendant 3h à 37°C dans un tampon (25 mM Tris pH 7,4, 1 mM CaCl<sub>2</sub>), 5 µL de chaque produit a été déposé sur une plaque de silice 60 TLC (Merck). Les dépôts ont ensuite été séparés par la migration par capillarité d'un solvant (60% (v/v) propan-1-ol, 10% (v/v) propan-2-ol, 10% (v/v) éthanol). Après migration, la plaque est pulvérisée par une solution de révélation (0,25% (m/v) α-naphtol, 50% (v/v) éthanol, 10% (v/v) acide sulfurique) et les produits sont révélés après une incubation de 15 min à 100°C.

## **H. Dosage des sucres dans le surnageant par HPLC**

Les concentrations en glucose, maltose et maltopentaose dans les surnageants de cultures ont été déterminées au cours du temps par chromatographie liquide à haute performance (HPLC). Après avoir éliminé les cellules par centrifugation, 20 µL de surnageant ont été injectés dans une colonne Aminex HPX-87H 300x7,8 mm (BioRad). Les différents constituants ont été séparés à 65°C par une élution isocratique avec une solution de 5 mM H<sub>2</sub>SO<sub>4</sub> par un flux constant de 0,6 mL/min. Les molécules d'intérêt ont été détectées par un lecteur 2142 Differential Refractometer (LKB Bromma), les concentrations absolues des différents sucres ont été déterminées par comparaison avec un échantillon standard de concentration connue. Les valeurs déterminées pour chaque sucre ont été standardisées par rapport à la concentration de MOPS (acide 3-morpholino-1-propanesulfonique) non métabolisable.

## **I. Thermal Shift Assays (TSA)**

Pour les expériences de TSA, le fluorophore Sypro™ Orange (Sigma-Aldrich) a été utilisé comme marqueur. La fluorescence (Ex : 515-535 nm, Em : 560-580 nm) est mesurée en continu dans un gradient de température de 25 à 80°C avec une incrémentation de 0,5°C par minute assuré par un appareil CFX96™ Real-Time PCR Detection System (Bio-Rad). Les températures d'exposition des surfaces hydrophobes (*Th*) sont déterminées par la température présentant la

plus forte variation de fluorescence. La variation des  $Th$  a été déterminée pour chaque condition ligand/protéine en soustrayant la valeur de  $Th$  obtenu pour la protéine seule à la valeur de  $Th$  mesurée. Les mesures de TSA ont été réalisées pour la protéine 6xHis-MalR (2  $\mu$ M) dans un tampon (10 mM  $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$  pH 7,6, 125 mM NaCl, 1 mM DTT, 0,1 mM EDTA, 7,5% (v/v) glycérol) et pour la protéine 6xHis-GenR (5  $\mu$ M) dans un tampon (50 mM Tris pH 7,4, 15% (v/v) glycérol).

# Résultats

## **I. Objectifs**

Comme annoncé au début de ce manuscrit, l'objectif de ce doctorat était de caractériser des métabolismes dont l'expression a été déterminée induite en contexte d'infection chez *E. faecalis*. Les métabolismes ayant retenu notre attention correspondent à l'utilisation du maltose et des maltodextrines d'une part et le métabolisme de certains  $\beta$ -glycosides dont principalement le gentiobiose d'autre part.

Les travaux menés au cours de ce doctorat ont pour objectifs de faire l'objet de trois publications scientifiques présentées ci-après. La première publication et la troisième présentées dans ce manuscrit ont été soumises à un comité scientifique de relecture au sein du journal "Molecular Microbiology" et sont en train de subir un processus de révision par les pairs. La deuxième publication est en cours d'écriture et devrait faire l'objet d'une soumission prochainement.

## **II. Publication scientifique n°1**

### **A. Résumé de la publication scientifique n°1**

Cet article scientifique s'intéresse à la régulation des deux opérons divergents *malPBMR* et *malT-mapP* essentiels pour permettre le métabolisme du maltose chez *E. faecalis*. Ces opérons sont induits en présence de maltodextrines indépendamment de leur longueur. Ces opérons sont également soumis à une répression catabolique indépendante de CcpA en présence de glucose. Cette CCR est assurée par un autre régulateur transcriptionnel de la famille LacI/GalR appelé MalR. Cette protéine MalR, codée par le dernier gène de l'opéron *malPBMR*, agit comme un répresseur direct sur l'expression des deux opérons en fixant un motif répondant au consensus CGMAACCGATTGCR (où M peut être C ou A, R peut être A ou G). Cette répression est levée par la présence dans la cellule de l'inducteur maltose. Nous avons pu déterminer *in vitro* une interaction directe de MalR avec la protéine P-Ser-HPr avec une affinité similaire au régulateur pléiotrope CcpA. Cette interaction de MalR avec P-Ser-HPr pourrait permettre d'expliquer son rôle essentiel dans la répression catabolique des opérons maltoses en présence de glucose.

### **B. Publication n°1 : The dual role of MalR in induction and a new P-Ser-HPr requiring catabolite repression mechanism of the *Enterococcus faecalis* mal operons**

1 **The dual role of MalR in induction and a new P-Ser-HPr-requiring**  
2 **catabolite repression mechanism of the *Enterococcus faecalis mal***  
3 **operons**

4  
5 Author names: Maxime Grand<sup>1</sup>, Victor Sebastián Blancato<sup>2</sup>, Martín Espariz<sup>2</sup>, Josef  
6 Deutscher<sup>3,4</sup>, Andreas Pikis<sup>5,6</sup>, Axel Hartke<sup>1</sup>, Christian Magni<sup>2</sup>, Nicolas Sauvageot<sup>1\*</sup>

7  
8 Affiliations:

9 <sup>1</sup>*Normandie Université, UNICAEN, U2RM Stress/Virulence, F-14000 Caen, France*

10 <sup>2</sup>*Instituto de Biología Molecular y Celular de Rosario (IBR-CONICET), Facultad de Ciencias*  
11 *Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Rosario, Argentina*

12 <sup>3</sup>*Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas,*  
13 *France*

14 <sup>4</sup>*UMR 8261, CNRS, Université de Paris, Institut de Biologie Physico-Chimique, 75005 Paris,*  
15 *France*

16 <sup>5</sup>*Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring,*  
17 *Maryland 20993, USA*

18 <sup>6</sup>*Microbial Biochemistry and Genetics Unit, Laboratory of Cell and Developmental Biology,*  
19 *NIDCR, National Institutes of Health, Bethesda, MD 20892, USA*

20  
21 Corresponding Author:

22 \*For correspondence. E-mail [nicolas.sauvageot@unicaen.fr](mailto:nicolas.sauvageot@unicaen.fr);

23 Tel. (+33) 2 31 56 56 61;

24 Fax: (+33) 2 31 56 53 11.

25  
26 Key Words: metabolism, maltose, *Enterococcus*, regulation

27  
28 Running Title: Regulation of maltose metabolism in *Enterococcus faecalis*

## 29 Abstract

30

31 Enterococci are Gram-positive pathogens and leading causes of hospital acquired  
32 infections worldwide. Several studies highlighted various factors related to pathogenicity.  
33 We showed previously that  $\alpha$ -polysaccharides metabolism is important for host  
34 colonization by *Enterococcus faecalis*. Here we show that MalR, a LacI/GalR  
35 transcriptional regulator, is the main factor in the regulation of the two divergent operons  
36 involved in maltose metabolism in this bacterium. The *malR* gene is transcribed from the  
37 *malP* promoter, but also from an internal promoter inside the gene located upstream of  
38 *malR*. In the absence of maltose, MalR acts as a repressor and in the presence of glucose,  
39 it exerts efficient CcpA-independent carbon catabolite repression. The central PTS  
40 protein P-Ser-HPr and the glycolytic intermediate fructose-1,6-bisphosphate interact  
41 directly with MalR and enhance its DNA binding capacity, which allows *E. faecalis* to  
42 adapt its metabolism to environmental conditions.

43

## 44 Introduction

45 Enterococci are Gram-positive bacteria widely distributed in the environment (Lebreton *et al.*,  
46 2014). These common commensal microorganisms inhabit the digestive tract of humans and  
47 animals. Despite their beneficial behaviour for their host, a few decades ago two major species,  
48 *E. faecalis* and *E. faecium*, became leading causes of hospital acquired infections (Robert C.  
49 Moellering, 1992; Edmond *et al.*, 1999; Wisplinghoff *et al.*, 2004; García-Solache and Rice,  
50 2019). Indeed, these bacteria can turn into opportunistic pathogens and cause important health-  
51 care associated diseases like urinary tract infection, endocarditis, bacteremia or septicemia  
52 especially in elderly or immunocompromised persons (Agudelo Higueta and Huycke, 2014;  
53 Ceci *et al.*, 2015). Resistance to antibiotics and to environmental factors like desiccation are  
54 probably the key underlying factors of enterococcal persistence in hospital environments (Bale  
55 *et al.*, 1993; Fisher and Phillips, 2009; Kristich *et al.*, 2014). Several virulence factors have

56 been identified and shown to contribute to the pathogenicity of enterococci, but the infection  
57 process remains poorly understood (Garsin *et al.*, 2014). This is especially true for the first steps  
58 in the process concerning nutrient mobilisation and proliferation in the infected host. Several  
59 studies showed that the metabolism of glycans is an important factor for bacterial host  
60 colonisation (Chang *et al.*, 2004; Sonnenburg *et al.*, 2010; Thomas *et al.*, 2011; Koropatkin *et*  
61 *al.*, 2012; Ng *et al.*, 2013; Xu *et al.*, 2019). A global transcriptomic study revealed that during  
62 peritoneal infection of mice with *E. faecalis*, 123 metabolic genes were differentially expressed  
63 (Muller *et al.*, 2015). Out of those genes, 74 were upregulated suggesting that they play an  
64 important role in host colonisation (Muller *et al.*, 2015). Another recent study showed that a  
65 global transcriptional regulator, MafR, is required for the expression of many genes involved  
66 in carbon metabolism. Interestingly, a *mafR* mutant caused a lower degree of inflammation in  
67 mice, suggesting that the MafR-deficient *E. faecalis* strain was less virulent than its  
68 corresponding wild-type (Ruiz-Cruz *et al.*, 2016). Furthermore, two operons involved in  
69 maltose metabolism in *E. faecalis* were shown to be upregulated during growth in horse blood  
70 (Vebø *et al.*, 2009) and one of these operons has also been reported to be involved in biofilm  
71 formation (Hufnagel *et al.*, 2004; Creti *et al.*, 2006).

72 Maltose is a reducing disaccharide composed of two glucosyl residues connected via an  
73  $\alpha$ -1,4 linkage. It represents the simplest form of maltodextrin obtained after hydrolysis of starch  
74 or glycogen. Maltose metabolism of *E. faecalis* is achieved through two divergent operons  
75 represented in Fig. 1A. In the first metabolic step, the disaccharide is imported by a  
76 phosphoenolpyruvate (PEP):carbohydrate phosphotransferase system (PTS) encoded by *malT*  
77 (Le Breton *et al.*, 2005). The main role of the PTS is to catalyse the uptake and concomitant  
78 phosphorylation of numerous carbohydrates (Deutscher *et al.*, 2006). It is usually composed of  
79 the two general proteins EI and HPr and three or four sugar-specific proteins or domains (EIIA,  
80 EIIB, EIIC, and sometimes EIID). The PTS components EI, HPr, EIIA and EIIB form a

81 phosphorylation cascade with PEP as phosphoryl donor. P~EIIB transfers its phosphoryl group  
82 to the carbohydrate bound to the corresponding membrane-integrated EIIC, and the resulting  
83 P-sugar is subsequently released into the cytoplasm (Deutscher *et al.*, 2006) In *E. faecalis*, the  
84 resulting maltose-6'-phosphate (M6'P) is dephosphorylated by the MapP protein (Mokhtari *et*  
85 *al.*, 2013) before its cleavage into glucose and glucose-1-phosphate (G1P) by the maltose  
86 phosphorylase MalP. G1P can be converted into glucose-6-phosphate (G6P) by MalB encoded  
87 by the second gene of the *malP* operon before entering glycolysis. Maltose metabolism is well  
88 understood in *E. faecalis*, but regulation of the expression of the maltose operons remains  
89 unknown in this species. On the other hand, regulation of maltose metabolism has been studied  
90 in the phylogenetically closely related genus *Streptococcus*. Similar to *E. faecalis*, maltose is  
91 preferentially imported in Streptococci through the PTS permease MalT (Webb *et al.*, 2007;  
92 Shelburne *et al.*, 2008; Bidossi *et al.*, 2012). The M6'P is supposed to be dephosphorylated by  
93 a M6'P phosphatase orthologue of the enterococcal MapP. Maltose is then catabolised by  
94 combined action of a glucanotransferase (Lacks, 1968; Stassi *et al.*, 1981; Sato *et al.*, 2013) and  
95 a glycogen phosphorylase (Martin and Russell, 1987; Sato *et al.*, 2013). Expression of these  
96 genes is repressed by the LacI/GalR family transcriptional regulator MalR in  
97 *Streptococcus pneumoniae*, *Streptococcus pyogenes* and *Streptococcus mutans* (Puyet, 1993;  
98 Nieto *et al.*, 1997; Nieto *et al.*, 2001) and submitted to Carbon Catabolite Repression (CCR).  
99 The catabolite control protein A (CcpA) was also shown to play a role in the regulation of  
100 maltose metabolism in streptococci. Indeed, inactivation of CcpA leads to a decrease in  
101 expression of the maltose genes in *S. pyogenes* (Shelburne *et al.*, 2008) grown in the presence  
102 of glucose. To efficiently bind DNA, CcpA first forms a complex with P-Ser-HPr in low G+C  
103 Gram-positive bacteria (Jones *et al.*, 1997). The P-Ser-HPr/CcpA complex specifically binds  
104 Catabolite-Responsive Elements also named *cre*-boxes to regulate expression of several

105 streptococcal genes including genes involved in metabolism but also in virulence (Iyer *et al.*,  
106 2005; Paluscio *et al.*, 2018).

107 Phosphorylation on Ser-46 of HPr is catalysed by the ATP-dependent kinase/phosphorylase  
108 HPrK/P (Deutscher and Saier, 1983). The kinase activity is stimulated by  
109 fructose-1,6-bisphosphate (FBP) and G6P (Deutscher and Engelmann, 1984; Kravanja *et al.*,  
110 1999; Chaptal *et al.*, 2007), the concentrations of which are elevated in the presence of a rapidly  
111 metabolizable carbon source such as glucose. P-Ser-HPr acts as a signaling molecule and has  
112 several important regulatory functions (Deutscher and Saier, 1983). As mentioned above, it also  
113 functions as co-regulator of CcpA. P-Ser-HPr and is involved in inducer exclusion in firmicutes  
114 (Deutscher, 2008; Homburg *et al.*, 2017).

115 In this study, we analysed the expression of the *malPBMR* and *malT-mapP* operons (Fig. 1.A)  
116 in *E. faecalis* OG1RF using RT-qPCR. These operons are induced in the presence of maltose  
117 and strongly repressed in the presence of glucose. We showed that CCR is independent of  
118 CcpA, but relies on another LacI/GalR family transcriptional regulator, MalR. We confirmed  
119 by Electrophoretic Mobility Shift Assay (EMSA) that MalR acts as a direct repressor of the  
120 expression of the maltose genes and that addition of maltose reduces its DNA binding  
121 capability. By carrying out MicroScale Thermophoresis (MST) we demonstrated a direct  
122 interaction between MalR and P-Ser-HPr. This interaction fine tunes MalR activity and  
123 coordinates the cell metabolism in accordance to carbon source availability.

124

## 125 **Materials and methods**

126 ***Bacterial strains and growth conditions.*** Strains and plasmids used in this study are shown in  
127 Table S1 and Table S2, respectively. Briefly, *Escherichia coli* strains used for cloning  
128 experiments and tagged-protein purification were grown in Lysogeny Broth (LB) (Bertani,

129 1951). When appropriate, cultures were supplemented with ampicillin, kanamycin or  
130 chloramphenicol with final concentrations of 100, 25 and 10  $\mu\text{g mL}^{-1}$ , respectively.

131 *Enterococcus faecalis* OG1RF (Dunny *et al.*, 1978) and its derivative strains (Table S1) were  
132 routinely grown in M17 medium containing 0.5% glucose (w/v) (GM17) at 37°C. For growth  
133 studies and the determination of transcriptional profiles, we used carbon cured M17 MOPS  
134 (ccM17) (Bizzini *et al.*, 2010) supplemented with 0.3% (w/v) of either glucose, maltose, and/or  
135 maltotriose according to experimental needs. When appropriate, cultures were supplemented  
136 with erythromycin or chloramphenicol with final concentrations of 100 and 10  $\mu\text{g mL}^{-1}$ ,  
137 respectively.

138 **Construction of mutants.** The genome sequence assembly of *E. faecalis* OG1RF  
139 (GCF\_000172575.2) was obtained from NCBI Reference Sequence Database (O’Leary *et al.*,  
140 2016). Primers used in this study are listed in Table S3. The *malR* null mutant was obtained by  
141 introducing two stop codons at the beginning of the open reading frame (stop corresponding to  
142 the parental amino acids 57 and 60) in order to produce a small truncated and presumably non-  
143 functional protein. The *ccpA* mutant was constructed by deleting nucleotides 49 to 899 of the  
144 parental gene, presenting approximately 85% of the entire sequence. PCR fragments were  
145 amplified by using *E. faecalis* OG1RF genomic DNA as template, appropriate primers and the  
146 Phusion polymerase (ThermoFisher Scientific) and cloned into the pMAD vector (Arnaud *et*  
147 *al.*, 2004). The modified alleles were used to replace the wild type alleles by the double crossing  
148 over method. Molecular constructs and chromosomal structures of mutants were systematically  
149 validated by Sanger based DNA sequencing (Eurofins).

150 **Complementation of mutants.** To complement the *malR*<sup>-</sup> mutant, the corresponding wild type  
151 gene together with its internal promoter located within the *malM* gene was cloned into the  
152 pRB473 plasmid (Brückner *et al.*, 1993) and electroporated into *E. coli* EC1000. After

153 extraction and purification, the recombinant plasmid was used to transform *E. faecalis malR*.  
154 *E. faecalis* cells containing pRB473-derived plasmids were always grown in the presence of  
155 2  $\mu\text{g mL}^{-1}$  of chloramphenicol.

156 ***Purification of His-tagged proteins.*** To purify MalR, CcpA and HPr, the corresponding genes  
157 were first cloned into expression vectors, pQE30 (Qiagen) for HPr (*ptsH* gene) (Monedero *et*  
158 *al.*, 2001), pET28a(+) (Novagen) for CcpA (Suárez *et al.*, 2011) and pET28b(+) (Novagen) for  
159 MalR (Table S2). For the amplification of the three genes, *E. faecalis* OG1RF genomic DNA  
160 was used as template together with appropriate primer pairs containing specific restriction sites  
161 (Table S3). The DNA fragments were digested with appropriate restriction enzymes and cloned  
162 into one of the above mentioned plasmids restricted with the same enzymes. The resulting  
163 plasmids were transformed into *E. coli* BL21(DE3) (New England Biolabs) or M15[pREP4].  
164 After sequence verification, a transformant for each gene was grown in Terrific Broth (Tartof  
165 and Hobbs, 1987) supplemented with kanamycin and/or ampicillin according to the resistance  
166 carried by the plasmids. Cells were grown until the medium reached an OD<sub>600</sub> above 0.6.  
167 Synthesis of the different proteins was induced with 0.5 mM of isopropyl- $\beta$ -D-1-  
168 thiogalactopyranoside (IPTG) (Eurobio) for 4h at room temperature under vigorous agitation.  
169 Tagged proteins were then purified using the kit Protino Ni-NTA (Macherey-Nagel) under  
170 nondenaturing conditions following the supplier's recommendations. Concentrations of  
171 different proteins were determined using the Pierce™ BCA Protein Assay Kit  
172 (ThermoFisher Scientific) with a BSA standard curve.

173 Part of HPr was phosphorylated *in vitro* using purified EI or HPrK/P(V267F) to obtain  
174 P~His-HPr or P-Ser-HPr, respectively, as previously described (Monedero *et al.*, 2001).

175 ***Isolation of RNA and transcript analysis.*** Overnight cultures were grown in fresh ccM17  
176 medium supplemented with appropriate carbon sources and harvested at mid exponential phase

177 (OD<sub>600</sub> about 0.6) by centrifugation. Total RNA was isolated using Direct-zol™ RNA Miniprep  
178 (Zymo Research) following manufacturer's recommendations. The integrity of RNA was  
179 verified on a bleach agarose gel (Aranda *et al.*, 2012) using 2 µg of RNA. One µg of total RNA  
180 was used to perform a Retro-Transcription (RT) step with the QuantiTect Reverse Transcription  
181 Kit (Qiagen). RT-qPCR analysis of cDNA was performed with the GoTaq® qPCR Master Mix  
182 (Promega). Primers used for RT-qPCR are shown in Table S4. Relative transcript abundances  
183 were determined by a quantification method using a standard genomic DNA curve and the  
184 housekeeping *gyrA* gene as reference.

185 ***β-Glucuronidase assays.*** To determine β-glucuronidase (Gus) activity, exponentially growing  
186 *E. faecalis* OG1RF carrying pNZ273 (Platteeuw *et al.*, 1994) and pNZ273-derived plasmids  
187 were harvested and re-suspended in the same volume of phosphate buffer (50 mM Na<sub>2</sub>HPO<sub>4</sub>-  
188 NaH<sub>2</sub>PO<sub>4</sub> pH 6.6, 10 mM KCl, 1 mM MgSO<sub>4</sub>). OD<sub>600</sub> was measured to standardize the amount  
189 of bacteria. Cells were chemically disrupted by adding lysis buffer (6.5% (v/v) chloroform,  
190 0.002% (v/v) SDS). The solution was vigorously mixed for 1 min and incubated at 37°C for  
191 5 min. The Gus substrate (p-nitrophenyl-β-D-glucuronide) was added to a final concentration  
192 of 627.5 µg mL<sup>-1</sup> to start the reaction. After incubation at 37°C for 10 min, the reaction was  
193 stopped by the addition of Na<sub>2</sub>CO<sub>3</sub> at 250 mM. After removing cellular debris by centrifugation,  
194 p-nitrophenol concentration was calculated by measuring the OD<sub>420</sub>. The β-glucuronidase  
195 activity was determined in arbitrary units (AU) by the following formula:  $Gus = 100 \times$   
196  $OD_{420} / OD_{600}$ .

197 ***Rapid Amplification of cDNA ends by polymerase chain reaction (5' RACE PCR).*** The  
198 transcription start site of the *malR* gene was determined by 5'RACE PCR using the kit 5'/3'  
199 RACE Kit, 2nd Generation (Roche Applied Science). The transcription start site corresponds

200 to the last nucleotide before the poly-A tail in the DNA sequence determined after Sanger  
201 (Eurofins) using the Race primers in Table S3.

202 **Electrophoretic Mobility Shift Assay (EMSA).** DNA fragments containing the *malP* or *malT*  
203 promoter or an internal fragment of the *gusA* gene (used as negative control) were first cloned  
204 into the pGEM-T Easy plasmid (Promega). DNA fragments were amplified by PCR using a  
205 primer labelled with fluorescent cyanine-5 at the 5'end (Table S3) and appropriate pGEMT  
206 derived plasmids as template. To obtain a DNA fragment containing the P2*mal* region, genomic  
207 DNA of *E. faecalis* was used as template with a primer linked to cyanine-5 (Table S3). Labelled  
208 DNA (5 nM) was incubated with MalR or CcpA protein in a final volume of 20  $\mu$ l in binding  
209 buffer (20 mM Tris, pH7.4, 20 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM FBP, 1 mM DTT,  
210 0.1 mM EDTA, 5% (v/v) glycerol, 50 ng  $\mu$ L<sup>-1</sup> poly(dI-dC), 0.01 mg mL<sup>-1</sup> bovine serum  
211 albumin) at room temperature for 20 min before the assay mixture was loaded on a  
212 5% acrylamide gel and separated under native conditions in TBE 0.5X buffer under constant  
213 voltage (140 V) in a dark-chilled room. Fluorescence was read with a ProXpress apparatus  
214 (PerkinElmer) with the following settings (top illumination 625 nm, top reading 680 nm,  
215 100  $\mu$ m resolution, 120 s exposure time).

216 **Synthesis of phosphorylated carbohydrates.** M6P and maltotriose-6''-P were obtained as  
217 previously described (Thompson *et al.*, 2001). G6P and FBP were purchased at Sigma-Aldrich.

218 **Thermal Shift Assay (TSA).** Sypro™ Orange (Sigma-Aldrich) was used as dye, fluorescence  
219 (Ex: 515-535 nm, Em: 560-580 nm HEX™ filter) was monitored in a temperature gradient of  
220 +0.5°C per minute from 25°C to 80°C in a CFX96™ Real-Time PCR Detection System  
221 (Bio-Rad). Temperatures of hydrophobic exposure (*Th*) were determined with Boltzmann  
222 regression on the sigmoid part of each curve using GraphPad Prism 7 software. The variation  
223 of *Th* ( $\Delta Th$ ) was calculated by subtracting *Ths* from *Th* determined without a ligand. Proteins

224 were used at a final concentration of 2  $\mu$ M and interaction assays were carried out in a solution  
225 containing 10 mM NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub> pH 7.6, 125 mM NaCl, 1 mM DTT, 0.1 mM EDTA,  
226 7.5% (v/v) glycerol. Apparent dissociation constants ( $K_D$ ) were calculated using Differential  
227 Scanning Fluorimetry assuming a single binding event with the GraphPad Prism 7 software for  
228 windows, GraphPad Software, La Jolla California USA, [www.graphpad.com](http://www.graphpad.com).

229 ***DNase I footprint analysis.*** The *PmalP* region was amplified as described above with primers  
230 pU (FAM) and pR (Table S3). The PCR product labelled with the FAM fluorophore was first  
231 incubated with 1  $\mu$ M of His-MalR in binding buffer (20 mM Tris HCl, pH7.4, 20 mM KCl,  
232 10 mM MgCl<sub>2</sub>, 10 mM FBP, 1 mM DTT, 5% (v/v) glycerol, 62.5 ng  $\mu$ L<sup>-1</sup> poly(dI-dC),  
233 0.01 mg mL<sup>-1</sup> bovine serum albumin) in a final volume of 65  $\mu$ L at room temperature for 15 min.  
234 Different dilutions (10  $\mu$ L) of RQ1 DNase I (Promega) (from 1/200 to 1/1600) were added to  
235 the mixture and incubated for 5 min at room temperature. The digestion reaction was stopped  
236 by addition of 2 mM of EDTA. The DNA fragments were finally purified with the NucleoSpin  
237 Gel and PCR Clean-up kit (Macherey-Nagel). Aliquots of the digested PCR products (10  $\mu$ L)  
238 were added to 10  $\mu$ L of formamide and 0.1  $\mu$ L of GeneScan-500 ROX size standard  
239 (ThermoFisher Scientific). Analysis of the DNA fragments was performed by the sequencing  
240 platform of the Cochin Institut (Eurofins) and computed using the PeakScanner software v1.0  
241 (ThermoFisher Scientific).

242 ***Microscale Thermophoresis (MST) Binding Assays.*** His-tagged MalR and CcpA were labelled  
243 with the NT-647-NHS dye using the Monolith NT<sup>TM</sup> Protein Labeling Kit RED-NHS  
244 (Nanotemper Technologies). Serially diluted unlabelled proteins tested for interaction and 5 nM  
245 of labelled MalR or CcpA were incubated for 15 min at room temperature in binding buffer  
246 (20 mM Tris pH 7.4, 20 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM FBP, 1 mM DTT, 0.1 mM EDTA,  
247 5% (v/v) glycerol, 0.005% (v/v) Tween20) in a final volume of 20  $\mu$ L. Samples were then  
248 loaded into NT.115 premium coated capillaries (Nanotemper Technologies). Binding

249 experiments were performed in a Monolith NT.115Pico apparatus (Nanotemper Technologies)  
250 with the following parameters: LED power 5%, MST Power medium. MST traces were  
251 analysed between 4.00 and 5.00 seconds after turning on the IR-Laser. Two measurements were  
252 carried out with HPr and P~His-HPr and 3 with P-Ser-HPr. For each experiment, the unlabelled  
253 proteins were produced from independent preparations. Results were obtained with the  
254 MO.Control software version 1.6 and computed with MO.Affinity Analysis software version  
255 2.3 to determine the fraction of the formed complex. Apparent dissociation constants ( $K_D$ ) were  
256 calculated using non-linear fitting assuming one specific binding site with the GraphPad Prism  
257 7 software for Windows, GraphPad Software, La Jolla California USA, [www.graphpad.com](http://www.graphpad.com).

258 **Multiple alignments.** Phyre2 (Kelley *et al.*, 2015) was used to generate a model of the MalR  
259 tridimensional structure. The PDB entry 1ZVV, corresponding to the CcpA-Crh-DNA complex  
260 from *Bacillus subtilis* was listed with as highest scoring template. The resulting alignment was  
261 used as template to perform a multiple alignment with the primary sequence of CcpAs and the  
262 RbsR protein using Clustal Omega (Sievers *et al.*, 2011).

263

## 264 **Results & Discussion**

265 ***Inactivation of malR prevents repression of the maltose genes.*** The genes involved in maltose  
266 metabolism in *E. faecalis* are organized in two divergent operons beginning with the *malP* and  
267 *malT* genes, respectively (Fig. 1A) (Le Breton *et al.*, 2005; Mokhtari *et al.*, 2013). The *malR*  
268 gene, which encodes a transcriptional regulator of the LacI/GalR family, is the last gene of the  
269 *malPBMR* operon. In light of its genomic location and its similarity to members of the  
270 LacI/GalR family, MalR was predicted to be the regulator of *mal* gene expression in *E. faecalis*.  
271 To study the role of MalR in this regulation, we first constructed a markerless *malR* null mutant  
272 by inserting two stop codons at the beginning of the *malR* gene (OG1RF\_10680). Next, the  
273 expression levels of the maltose operons were measured by RT-qPCR using the first gene of  
274 each operon as target (Fig. 2). When the parental strain OG1RF was grown on glucose, these  
275 genes were only weakly expressed. In the presence of maltose or maltotriose, expression of the  
276 genes was induced. When glucose is also present in the medium, these genes are repressed  
277 independently of the presence of maltose or maltotriose. In a *malR*<sup>-</sup> mutant, both *malP* and *malT*  
278 genes were constitutively expressed. This result suggests that MalR is the repressor of the *mal*  
279 operons. By analogy with well characterized LacI/GalR transcriptional regulators, MalR should  
280 bind to its operator under repressing conditions and be released in the presence of the inducer  
281 (Swint-Kruse and Matthews, 2009). To determine the nature of the inducer, we performed  
282 Thermal Shift Assays (TSA) with purified MalR in the presence of different sugars.

283 ***MalR presents affinity for maltose.*** As described above, the *mal* operons appear to be repressed  
284 by glucose and induced by maltose and maltotriose. The two glucans are imported into the cell  
285 through a PTS MalT and therefore enter the cells in phosphorylated form (Le Breton *et al.*,  
286 2005). LacI/GalR type transcriptional regulators are frequently regulated by small ligands  
287 closely related to the function of the regulated proteins. We therefore carried out TSA with  
288 His-tagged MalR and either glucose, maltose, maltotriose or the corresponding

289 phospho-compounds by using the Sypro™ Orange fluorophore (Vivoli *et al.*, 2014;  
290 Grøftehaug *et al.*, 2015). As shown in Fig. 3, the temperature of hydrophobic exposure (*Th*)  
291 of the MalR protein rose with increasing concentrations of maltose (Fig. 3A), suggesting that  
292 MalR presents binding sites for maltose. In contrast, even at very high concentrations of M6'P,  
293 only a slight increase of *Th* was observed (Fig. 3B), demonstrating that MalR has only weak  
294 affinity for M6'P. By assuming a single binding site for maltose, non-linear regression correctly  
295 fitted the data points and allowed to determine an apparent dissociation constant ( $K_D$ ) of  
296  $2.2 \pm 0.4$  mM. This relatively low affinity of the repressor for its inducer should require an  
297 efficient import system in order to fully induce the *mal* genes. Moreover, since M6'P is not an  
298 efficient ligand for MalR, the phosphatase MapP (Mokhtari *et al.*, 2013) might be crucial for  
299 induction of these genes because it dephosphorylates M6'P to maltose. No significant variation  
300 of *Th* over the whole concentration range tested was observed in the presence of glucose,  
301 maltotriose and their phospho-derivatives, suggesting that MalR does not bind these molecules  
302 (Fig. S4). These data indicate that maltose is the inducer of the two *mal* operons in this  
303 bacterium. In contrast to some bacteria, which hydrolyze M6'P formed during PTS-catalyzed  
304 transport of maltose to glucose and G6P by a 6-P- $\alpha$ -glucosidase (Yip *et al.*, 2007), M6'P  
305 arriving in *E. faecalis* via the PTS permease MalT is re-converted to maltose by MapP  
306 (Mokhtari *et al.*, 2013). The unusual dephosphorylation step might have been implemented  
307 during evolution in order to keep *malT* expression under control of the MalR/maltose complex.

308 ***MalR acts as a repressor by direct binding to the promoter regions of malP and malT.*** In  
309 order to prove the direct implication of MalR in the regulation of the maltose operons, we  
310 performed EMSA. Three fluorescently labelled DNA fragments containing either *PmalP*,  
311 *PmalT* or both promoters (the latter DNA fragment is named P2*mal*) (Fig. 4 A) were incubated  
312 in the presence of MalR. Two distinct shifted bands are visible when the entire intergenic region  
313 is used as the template (Fig. 4B, lanes 2 to 5) suggesting a direct specific interaction between

314 MalR and the DNA. In addition, the intensity of the sifted bands decreased when maltose was  
315 present in the reaction buffer (Fig. 4B, lanes 2 to 5). The MalR/DNA interactions are considered  
316 specific because a control DNA was not shifted under the same conditions (Fig. 4C,  
317 lanes 7 to 9), indicating that MalR acts as a transcriptional repressor and that maltose functions  
318 as the inducer of the *mal* operons.

319 The appearance of two shifted bands suggested the presence of two distinct binding sites of  
320 MalR in the P2*mal* DNA fragment. The upper shifted band might correspond to a complex of  
321 DNA with MalR bound to both sites and the lower band to a complex of MalR bound to a single  
322 site. Based on our results that MalR regulates the *malP* and *malT* genes, we hypothesized that  
323 there is a specific MalR operator site for each promoter. Indeed, MalR shifted the two DNA  
324 fragments containing either only *PmalP* or only *PmalT* (Fig. 4C), thus confirming our  
325 hypothesis that each promoter region contains a distinct MalR operator. The shifted band seems  
326 to be more intense and the protein-DNA complex more stable in the case of the *malP* promoter  
327 compared to the *malT* promoter (Fig. 4C, lanes 2 and 5), which might reflect a higher affinity  
328 of MalR for *PmalP* than for *PmalT*. A similar difference in binding affinities of MalR to its  
329 operator sites present upstream of operons implicated in maltose metabolism has been noticed  
330 for *S. pneumoniae* (Nieto *et al.*, 2001). This difference in repression could be of physiological  
331 importance for the bacterium since a too strong repression of the maltose transporter and the  
332 M6P phosphatase genes would slow or even prevent induction of the maltose operons in the  
333 presence of the inducer.

334 ***Determination of MalR binding sites.*** The above data clearly show that *E. faecalis* MalR binds  
335 to DNA fragments harbouring the *malP* or *malT* promoter regions. Therefore, we screened them  
336 for imperfect palindromic sequences with a central AANC motif (where N can be any  
337 nucleotide), which is typical for operators of LacI/GalR type transcriptional regulators  
338 (Weickert and Adhya, 1992; Swint-Kruse and Matthews, 2009). A single short inverted repeat

339 sequence was found in each promoter region (Fig. 5A). To confirm these putative operator  
340 sequences experimentally, DNase I footprinting experiments were carried out. Due to the higher  
341 affinity of MalR for the *malP* promoter, we performed these assays using the *PmalP* region as  
342 template. The predicted MalR binding site was expected to span positions 185 to 198 of this  
343 region. The protected region in Fig. 5B perfectly coincides with the predicted positions.  
344 Interestingly, these motifs show strong similarity to the MalR operators identified in  
345 *S. pneumoniae* (Nieto *et al.*, 1997) and to the *E. faecalis* consensus *cre*-boxes, the binding sites  
346 of CcpA, which also belongs to the LacI/GalR family (Opsata *et al.*, 2010). The resulting MalR  
347 consensus sequence of *E. faecalis* is CGMAACCGATTGCR (where M can be C or A, R can  
348 be A or G). In the *malP* promoter, this sequence is located between the -10 and the putative -  
349 35 box, whereas in the *malT* promoter it is located between the transcription start site and the  
350 Ribosome Binding Site (RBS) of *malT* (Fig. 5.A). These positions are consistent with the  
351 repressor function of MalR, which can either sterically hinder transcription initiation of RNA  
352 polymerase (at *malP*) or form a roadblock (at *malT*). Given that MalR was identified as the  
353 main regulator of the *mal* system, an in-depth characterisation of its gene was carried out.

354 ***The malR gene is expressed from two promoters.*** By carrying out a northern blot, Le Breton  
355 *et al* (Le Breton *et al.*, 2005) had previously shown that the *E. faecalis malP* gene is present in  
356 a 5.2 kb corresponding to the *malPBMR* transcript. Furthermore, co-transcription of *malR* and  
357 *malM* was confirmed by PCR on cDNA using two primers derived from *malR* and *malM*  
358 (Fig. S1). This result suggested that *malR* and *malP* are present in the same transcription unit.  
359 In addition, an *in silico* search did not reveal a transcriptional terminator after the stop codon  
360 of the *malM* gene. However, DNA sequence analysis suggested the presence of a potential  
361 promoter at the end of the *malM* gene. To test the function of this presumed promoter and to  
362 map its transcriptional start site (+1), we performed a Rapid Amplification of cDNA 5' Ends by  
363 Polymerase Chain Reaction (5' RACE PCR) on cells grown on maltose. The results confirmed

364 that this promoter is indeed functional and that its +1 position is located 201 bp upstream of the  
365 predicted start codon of *malR* (Fig. 1B). A probable -10 box (TATTAT) is located upstream of  
366 the +1 extending from -8 to -13.

367 To further characterize the *malR* promoter *PmalR*, we performed transcriptional fusion assays.  
368 A genome region of OG1RF containing *PmalR* was cloned into plasmid pNZ273 upstream of  
369 a *gusA* reporter gene encoding the enzyme  $\beta$ -D-glucuronidase (Fig. 6) (Platteeuw *et al.*, 1994).  
370 After transforming *E. faecalis* OG1RF,  $\beta$ -D-glucuronidase activity was determined at 6 hours  
371 after inoculation in ccM17 medium supplemented with glucose or maltose. At least 5 times  
372 higher  $\beta$ -D-glucuronidase activity was obtained with the *PmalR-gusA* construction compared  
373 with control plasmids (empty pNZ273 and *PmalR* inserted in the opposite direction)  
374 independent of the carbon source used by the bacteria (Fig. 6). These results confirmed that the  
375 promoter present upstream of the *malR* gene is functional and, in view of the low  
376  $\beta$ -D-glucuronidase activity, allows a low and constitutive expression of *malR*. MalR auto  
377 regulates the expression of its gene from *PmalP*. The additional constitutive promoter at the  
378 end of *malM* allows the synthesis of a constant basal level of repressor, which prevents the cell  
379 from wasting energy in the absence of maltose.

380 The *malR* gene has previously been studied in *E. faecalis* and was originally named *bopD*  
381 (Hufnagel *et al.*, 2004). The authors showed its implication in biofilm formation especially in  
382 the presence of glucose and maltose (Creti *et al.*, 2006). They also noticed a residual expression  
383 of the *bopD* gene in a mutant with an insertion of a transposon in the upstream gene *bopB* (here  
384 named *malB*) of the *bopABCD* operon (Hufnagel *et al.*, 2004). This residual expression can  
385 now be explained by the internal promoter *PmalR* identified in our study. Previous attempts to  
386 construct a *malR*<sup>-</sup> mutant in the *E. faecalis* T9 strain had failed (Creti *et al.*, 2006) suggesting  
387 that in certain *E. faecalis* strains this gene might have functions in addition to the regulation of  
388 maltose genes. Moreover, in the three *E. faecalis* strains, JH2-2 (also named TX4000),

389 EnGen0241 and EnGen0299 the well conserved MalR (>99% sequence identity) contain a non-  
390 conservative replacement of threonine-33 in the Helix-Turn-Helix (HTH) domain by a proline  
391 (T33P). We had observed that the JH2-2 strain exhibits constitutive expression of the maltose  
392 genes even in the absence of the inducer similar to the *malR* mutant of OG1RF (data not shown),  
393 suggesting that the MalR protein of JH2-2 might be non-functional. This observation confirms  
394 that the *malR* gene can be inactivated in certain *E. faecalis* strains and that inactivation can  
395 sometimes even occur spontaneously.

396 ***The maltose genes are submitted to CcpA-independent but MalR-dependent CCR.*** As shown  
397 in Fig. 2, induction of the *malP* and *malT* genes was repressed when OG1RF was grown on  
398 glucose and either maltose or maltotriose, demonstrating that these genes are submitted to CCR.  
399 Interestingly, repression by glucose was independent of CcpA (OG1RF\_11453), the major  
400 global CCR regulator in firmicutes, because expression of *malP* and *malT* was still repressed  
401 by glucose in a *ccpA* deletion mutant ( $\Delta ccpA$ ) (Fig. 2). In the *malR* mutant, both *mal* operons  
402 were highly expressed under all tested growth conditions, even when grown on glucose. The  
403 MalR protein therefore appears to be essential for efficient CCR of these operons. To confirm  
404 that the observed deregulation was due to the introduced mutations in the *malR* gene, a wild  
405 type copy of the *malR* gene with its internal promoter was inserted into plasmid pRB473. While  
406 transformation of the *malR* mutant with the empty plasmid had no noticeable effect on maltose  
407 gene expression, a plasmid carrying the *malR* gene restored repression nearly to wild-type level  
408 (Fig S2). The MalR protein is therefore necessary for CCR of the maltose genes by glucose and  
409 presumably other efficiently metabolizable carbon sources.

410 Despite the presence of two MalR operators with the consensus sequence  
411 CGMAACCGATTGCR, which is closely related to *cre*-boxes determined for CcpA of  
412 *E. faecalis* (WTGWAARCGYWWWCW where W=A or T, R=A or G and Y=C or T) (Opsata  
413 *et al.*, 2010), CcpA/P-Ser-HPr does not bind to the P2*mal* region *in vitro* (Fig. S3A). This result

414 and the observed failure of the *ccpA* mutant to cause a relief from CCR suggest that CcpA is  
415 not involved in CCR of the *E. faecalis* maltose genes.

416 ***MalR interacts with P-Ser-HPr in vitro.*** When searching for MalR homologues in *E. faecalis*  
417 using BLASTp (Altschul *et al.*, 1990), CcpA was listed first because it exhibits 30% identity  
418 and 50% similarity to MalR. The finding that in a *malR* mutant the *mal* genes are relieved from  
419 CCR and the observed significant similarity between MalR and CcpA prompted us to study a  
420 direct interaction between MalR and P-Ser-HPr. We carried out MST experiments to detect  
421 interactions of MalR with HPr and its two phospho-forms as described in Experimental  
422 Procedures. Purified His-tagged MalR was first covalently linked to the NT-647-NHS  
423 fluorophore. The amount of NT.647 labelled MalR was maintained constant, while the  
424 concentration of the various non-labelled HPr forms varied from 9.76 nM to 20  $\mu$ M. The known  
425 interaction of His-tagged CcpA with P-Ser-HPr was used as a positive control (Fig. 7C). As  
426 expected, no interaction signal was detected with NT.647-MalR and HPr or P~His-HPr (Fig. 7B  
427 and D). In contrast, an interaction signal was observed with NT.647-MalR or NT.647-CcpA  
428 when P-Ser-HPr was used at concentrations exceeding 1  $\mu$ M (Fig. 7A and C, respectively).  
429 Assuming a single binding site and a specific interaction, a typical binding isotherm curve was  
430 obtained. The approximate apparent dissociation constants for the P-Ser-HPr/CcpA and  
431 P-Ser-HPr/MalR complexes were in the same order ( $K_D$  around 10-20  $\mu$ M), suggesting a similar  
432 affinity for the co-factor. The measured  $K_D$  of the CcpA/P-Ser-HPr interaction appears at least  
433 two fold higher in our experiments than the  $K_D$  previously determined for CcpA/P-Ser-HPr of  
434 *Bacillus subtilis* (Jones *et al.*, 1997) but was still in the same order of magnitude.

435 To get an indication whether MalR uses an interface for the interaction with P-Ser-HPr similar  
436 to that of CcpA and RbsR, which have been shown to interact with P-Ser-HPr, a multiple  
437 sequence alignment was carried out with MalR, CcpA from *B. subtilis*, *Bacillus megaterium*  
438 and *E. faecalis* and RbsR from *B. subtilis*. The crystal structure of the *B. megaterium* CcpA/P-

439 Ser-HPr complex had revealed two CcpA regions which interact with P-Ser-HPr. The first  
440 region extends from amino acids 69 to 99 (Fig. 8.A) and includes three aspartic acids D69, D84  
441 and D99 of subunit II, which interact with R17 of P-Ser-HPr (Schumacher *et al.*, 2004). While  
442 all three Asp are conserved in RbsR, only D84 is replaced by a conservative exchange with E90  
443 in the MalR sequence. The other two Asp are replaced by Arg and Lys. In addition, just  
444 upstream from the first conserved Asp in RbsR and CcpA, MalR carries an insertion of six  
445 amino acids (Fig. 8.A). It is therefore likely that the structure of this region of MalR is  
446 significantly different from that in CcpA proteins and RbsR and other negatively charged amino  
447 acids of this region might be involved in the interaction with R17 of P-Ser-HPr. The second  
448 region of CcpA involved in the interaction with P-Ser-HPr extends from amino acids 296 to  
449 307 in subunit I. D296 forms hydrogen bonds with unphosphorylated H15 and the main chain  
450 of A16, R303 and K307 interact with the phosphate group bound to S46 in P-Ser-HPr (Fig. 8.B)  
451 (Schumacher *et al.*, 2004). D296 is conserved in RbsR and replaced with E297 in MalR. In  
452 MalR and RbsR, R303 is replaced with a Lys, whereas R307 corresponds to a Glu in MalR and  
453 a Gly in RbsR. In summary, the alignment suggests that the RbsR interface involved in the  
454 interaction with P-Ser-HPr largely resembles that of CcpA, whereas the interface of MalR is  
455 probably significantly different. It will therefore be interesting to determine the structure of the  
456 MalR/P-Ser-HPr complex in order to find out how this regulator formed its P-Ser-HPr binding  
457 surface. Insertion of 6 amino acids upstream of R75 probably played an important role in this  
458 adaptation process (Fig. 8.A).

459 The MST experiments established that MalR forms a complex with P-Ser-HPr, which we  
460 suspected to be responsible for the observed CcpA-independent CCR of the *mal* genes. We tried  
461 to support this hypothesis by carrying out *in vivo* experiments with an *E. faecalis* OG1RF-  
462 derived *ptsHI* (*ptsHS46A*) mutant (Eisermann *et al.*, 1988; Deutscher *et al.*, 2006). In this  
463 mutant, the *mal* genes should be relieved from CCR, because the HPrK/P-dependent

464 phosphorylation of HPr is abolished by the replacement of S46 with a non phosphorylatable  
465 alanine. Unfortunately, numerous attempts to obtain the *ptsHI* mutant failed, suggesting that  
466 preventing the phosphorylation of HPr at S46 does not seem to be possible in *E. faecalis*. A  
467 similar behaviour has already been reported for *S. pneumoniae* where the formation of  
468 P-Ser-HPr appears to be essential, although the exact cause remains unknown (Fleming *et al.*,  
469 2015).

470 ***The presence of P-Ser-HPr enhances the DNA binding affinity of MalR in vitro.*** Since MalR  
471 interacts with P-Ser-HPr, we tested whether addition of the co-factor during EMSA would  
472 increase the DNA binding affinity of MalR. Indeed, in the presence of P-Ser-HPr and 25 nM  
473 MalR, the repressor binds DNA more efficiently (compare Fig. 4B, lanes 2 and 6). The upper  
474 shifted band is stronger and the free DNA amount is lower. Moreover, in the presence of  
475 P-Ser-HPr, the maltose-mediated inhibition of MalR/DNA interaction is less efficient (Fig. 4B,  
476 compare lanes 4 and 5 to lanes 8 and 9). This finding might reflect a binding competition  
477 between P-Ser-HPr and maltose for MalR or a P-Ser-HPr induced conformational change of  
478 MalR, which might lower its affinity for maltose. Of note, no enhancement of MalR DNA  
479 binding was observed with other physiologically relevant forms of HPr (unphosphorylated HPr  
480 and P~His-HPr). In addition, HPr and its two P-derivatives do not bind to P2*mal* (Fig. S3B).  
481 The specific P-Ser-HPr-mediated increase of the DNA binding affinity of MalR is probably  
482 responsible for CcpA-independent CCR of the two *mal* operons. The utilization of glucose  
483 increases the amount of P-Ser-HPr (Deutscher and Saier, 1983; Monedero *et al.*, 2001; Ludwig  
484 *et al.*, 2002) and hence MalR affinity for its operator sites. Regarding the strong repression of  
485 the *mal* operons in the  $\Delta$ *ccpA* mutant grown on glucose, two conditions might be responsible.  
486 First, the elevated amount of P-Ser-HPr in the  $\Delta$ *ccpA* mutant (Leboeuf *et al.*, 2000; Ludwig *et*  
487 *al.*, 2002) and, second, the absence of competition between MalR and CcpA for P-Ser-HPr.

488 **MalR exhibits affinity for FBP.** CcpA binds the glycolytic intermediate FBP, which promotes  
489 the formation of the P-Ser-HPr/CcpA complex (Deutscher *et al.*, 1995; Schumacher *et al.*,  
490 2007). We noticed that including FBP in the EMSA binding buffer improved the resolution of  
491 the bands shifted by MalR (data not shown). Because the two proteins, MalR and CcpA, appear  
492 to be functionally closely related in *E. faecalis* we determined the affinity of MalR for FBP by  
493 the same technique used before for MalR and maltose. Indeed, the temperature of hydrophobic  
494 exposure (*Th*) rose with increasing concentrations of FBP (Fig. 3.C), thus supporting an  
495 interaction with MalR. An apparent  $K_D$  of  $6.6 \pm 1$  mM could be determined assuming a single  
496 binding site. In resting *Lactococcus lactis* cells, the FBP concentration was found to increase  
497 from 2 mM to 25 mM, when glucose was added (Thompson and Torchia, 1984). Similarly to  
498 what was observed for CcpA, the interaction with FBP might promote the formation of the  
499 MalR/P-Ser-HPr complex during growth on glucose. FBP is therefore expected to play a dual  
500 role in maltose CCR in *E. faecalis* by stimulating the kinase function of HprK/P (Deutscher and  
501 Engelmann, 1984) and the formation of the MalR/P-Ser-HPr complex.

502

## 503 **Conclusion**

504 In this study we demonstrate that *E. faecalis* MalR functions not only as repressor of the two  
505 *mal* operons, but that it is also involved in their CCR by probably using P-Ser-HPr as  
506 co-repressor. MalR therefore allows *E. faecalis* to specifically respond to the presence of  
507 maltose either as sole carbon source or together with rapidly metabolizable carbohydrates.  
508 Adaptation of bacteria to environmental conditions is crucial for efficient fitness. In  
509 oligotrophic milieu and under harsh competition, optimization of carbohydrate utilization can  
510 be essential for survival. The central sensor of CCR, P-Ser-HPr, is known to play a major role  
511 in metabolic regulation in firmicutes (Deutscher, 2008) by functioning as co-repressor of the  
512 global regulator CcpA or by directly inhibiting non-PTS transporters (Ye *et al.*, 1996;

513 Dossonnet *et al.*, 2000; Viana *et al.*, 2000; Monedero *et al.*, 2008; Homburg *et al.*, 2017).  
514 P-Ser-HPr was also shown to interact *in vitro* with the *B. subtilis* LacI/GalR type transcriptional  
515 regulator RbsR (Müller *et al.*, 2006), but no physiological function was reported for the  
516 complex. We demonstrate in our study that MalR interacts with P-Ser-HPr, but we were not  
517 able to establish an *in vivo* role of the complex in CCR, because the *ptsHI* mutant could not be  
518 obtained. Nevertheless, it is likely that the two proteins also form a complex in the cell to  
519 prevent induction of the *mal* genes in the presence of a preferred carbon source such as glucose.  
520 The interaction of P-Ser-HPr with transcriptional regulators other than CcpA might be a more  
521 frequent CcpA-independent CCR mechanism in Gram-positive bacteria than presently known.  
522 For example, in *Listeria monocytogenes*, glucose caused CcpA-independent repression of the  
523 maltose and maltodextrin genes, which was relieved in an HprK/P mutant (Gopal *et al.*, 2010).  
524 The authors hypothesized an inducer exclusion mechanism but did not rule out another yet  
525 unknown mechanism. In view of our results, the MalR protein of *L. monocytogenes*, a LacI type  
526 repressor, might also form a complex with P-Ser-HPr in order to exert CcpA-independent CCR.  
527 Despite some similarities, there are specific differences between CcpA and MalR. With the  
528 exception of a few examples (Kim *et al.*, 1995), CcpA requires P-Ser-HPr to efficiently interact  
529 with *cre* sites. In contrast, MalR binds efficiently and specifically to its operators already in the  
530 absence of the co-repressor P-Ser-HPr. Both transcription regulators bind FBP, which enhances  
531 their affinity for the co-repressor. However, MalR binds in addition unphosphorylated maltose,  
532 which is required for induction of the *mal* operons and diminishes MalR affinity for its operator.  
533 It is therefore likely that MalR has two distinct binding sites, one for FBP (phosphorylated and  
534 negatively charged) and one for unphosphorylated maltose.

535 When searching for MalR binding sites elsewhere in the *E. faecalis* genome, we found two  
536 putative degenerated sites located in the promoters of physiologically related genes, which code  
537 for proteins required for the uptake and metabolism of maltotetraose and longer maltodextrins

538 (Sauvageot *et al.*, 2017; Joyet *et al.*, 2017). In the closely related *E. faecium* E1162, the  
539 maltodextrin genes are positively regulated by an additional LacI/GalR type transcriptional  
540 regulator, MdxR (Zhang *et al.*, 2013). No homolog of MdxR is present in *E. faecalis*, suggesting  
541 that regulation of maltodextrin gene expression is differently controlled in this species.  
542 Investigations are under way to determinate whether *E. faecalis* MalR is also implicated in the  
543 regulation of the two maltodextrin operons.

544

#### 545 **Acknowledgements**

546 We thank Reinhold Brückner for the generous gift of plasmid pRB473. We also greatly thank  
547 John Thompson for the supply of the phosphorylated sugars. The supports of Olivier  
548 Lesouhaitier for the MST data acquisition and Pierre Soule for the interpretation of these data  
549 were particularly precious. We are particularly grateful to Isabelle Rince for technical  
550 assistance.

551 This work was funded by a chair of excellence for Dr. Magni co-financed by the European  
552 Union and the Normandy County Council in the framework of the ERDF-ESF operational  
553 programme 2014-2020 and by the French Ministry of Education, Research and Innovation.

554

- 555 Agudelo Higueta, N.I., and Huycke, M.M. (2014) Enterococcal Disease, Epidemiology, and  
556 Implications for Treatment. In *Enterococci: From Commensals to Leading Causes of Drug*  
557 *Resistant Infection*. Gilmore, M.S., Clewell, D.B., Ike, Y., and Shankar, N. (eds). Massachusetts  
558 Eye and Ear Infirmary, Boston. <http://www.ncbi.nlm.nih.gov/books/NBK190429/>. Accessed  
559 January 27, 2016.
- 560 Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local  
561 alignment search tool. *J Mol Biol* **215**: 403–410.
- 562 Aranda, P.S., LaJoie, D.M., and Jorcyk, C.L. (2012) Bleach gel: A simple agarose gel for  
563 analyzing RNA quality. *ELECTROPHORESIS* **33**: 366–369.
- 564 Arnaud, M., Chastanet, A., and Debarbouille, M. (2004) New Vector for Efficient Allelic  
565 Replacement in Naturally Nontransformable, Low-GC-Content, Gram-Positive Bacteria. *Appl*  
566 *Environ Microbiol* **70**: 6887–6891.
- 567 Bale, M.J., Bennett, P.M., Beringer, J.E., and Hinton, M. (1993) The survival of bacteria  
568 exposed to desiccation on surfaces associated with farm buildings. *J Appl Bacteriol* **75**: 519–  
569 528.
- 570 Bertani, G. (1951) Studies on lysogenesis: the mode of phage liberation by lysogenic  
571 *Escherichia coli*. *J Bacteriol* **62**: 293–300.
- 572 Bidossi, A., Mulas, L., Decorosi, F., Colomba, L., Ricci, S., Pozzi, G., *et al.* (2012) A functional  
573 genomics approach to establish the complement of carbohydrate transporters in *Streptococcus*  
574 *pneumoniae*. *PloS One* **7**: e33320.
- 575 Bizzini, A., Zhao, C., Budin-Verneuil, A., Sauvageot, N., Giard, J.-C., Auffray, Y., and Hartke,  
576 A. (2010) Glycerol is metabolized in a complex and strain-dependent manner in *Enterococcus*  
577 *faecalis*. *J Bacteriol* **192**: 779–785.
- 578 Brückner, R., Wagner, E., and Götz, F. (1993) Characterization of a sucrase gene from  
579 *Staphylococcus xylosus*. *J Bacteriol* **175**: 851–857.
- 580 Ceci, M., Delpech, G., Sparo, M., Mezzina, V., Sánchez Bruni, S., and Baldaccini, B. (2015)  
581 Clinical and microbiological features of bacteremia caused by *Enterococcus faecalis*. *J Infect*  
582 *Dev Ctries* **9**: 1195–1203.
- 583 Chang, D.-E., Smalley, D.J., Tucker, D.L., Leatham, M.P., Norris, W.E., Stevenson, S.J., *et al.*  
584 (2004) Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proc Natl Acad Sci U S A*  
585 **101**: 7427–7432.
- 586 Chaptal, V., Vincent, F., Gueguen-Chaignon, V., Monedero, V., Poncet, S., Deutscher, J., *et al.*  
587 (2007) Structural analysis of the bacterial HPr kinase/phosphorylase V267F mutant gives  
588 insights into the allosteric regulation mechanism of this bifunctional enzyme. *J Biol Chem* **282**:  
589 34952–34957.
- 590 Creti, R., Koch, S., Fabretti, F., Baldassarri, L., and Huebner, J. (2006) Enterococcal  
591 colonization of the gastro-intestinal tract: role of biofilm and environmental oligosaccharides.  
592 *BMC Microbiol* **6**: 60.

- 593 Deutscher, J. (2008) The mechanisms of carbon catabolite repression in bacteria. *Curr Opin*  
594 *Microbiol* **11**: 87–93.
- 595 Deutscher, J., and Engelmann, R. (1984) Purification and characterization of an ATP-dependent  
596 protein kinase from *Streptococcus faecalis*. *FEMS Microbiol Lett* **23**: 157–162.
- 597 Deutscher, J., Francke, C., and Postma, P.W. (2006) How phosphotransferase system-related  
598 protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev*  
599 *MMBR* **70**: 939–1031.
- 600 Deutscher, J., Küster, E., Bergstedt, U., Charrier, V., and Hillen, W. (1995) Protein kinase-  
601 dependent HPr/CcpA interaction links glycolytic activity to carbon catabolite repression in  
602 gram-positive bacteria. *Mol Microbiol* **15**: 1049–1053.
- 603 Deutscher, J., and Saier, M.H. (1983) ATP-dependent protein kinase-catalyzed phosphorylation  
604 of a seryl residue in HPr, a phosphate carrier protein of the phosphotransferase system in  
605 *Streptococcus pyogenes*. *Proc Natl Acad Sci U S A* **80**: 6790–6794.
- 606 Dossonnet, V., Monedero, V., Zagorec, M., Galinier, A., Pérez-Martínez, G., and Deutscher, J.  
607 (2000) Phosphorylation of HPr by the bifunctional HPr Kinase/P-ser-HPr phosphatase from  
608 *Lactobacillus casei* controls catabolite repression and inducer exclusion but not inducer  
609 expulsion. *J Bacteriol* **182**: 2582–2590.
- 610 Dunny, G.M., Brown, B.L., and Clewell, D.B. (1978) Induced cell aggregation and mating in  
611 *Streptococcus faecalis*: evidence for a bacterial sex pheromone. *Proc Natl Acad Sci U S A* **75**:  
612 3479–3483.
- 613 Edmond, M.B., Wallace, S.E., McClish, D.K., Pfaller, M.A., Jones, R.N., and Wenzel, R.P.  
614 (1999) Nosocomial Bloodstream Infections in United States Hospitals: A Three-Year Analysis.  
615 *Clin Infect Dis* **29**: 239–244.
- 616 Eisermann, R., Deutscher, J., Gonzy-Treboul, G., and Hengstenberg, W. (1988) Site-directed  
617 mutagenesis with the ptsH gene of *Bacillus subtilis*. Isolation and characterization of heat-stable  
618 proteins altered at the ATP-dependent regulatory phosphorylation site. *J Biol Chem* **263**:  
619 17050–17054.
- 620 Fisher, K., and Phillips, C. (2009) The ecology, epidemiology and virulence of *Enterococcus*.  
621 *Microbiology* **155**: 1749–1757.
- 622 Fleming, E., Lazinski, D.W., and Camilli, A. (2015) Carbon catabolite repression by seryl  
623 phosphorylated HPr is essential to *Streptococcus pneumoniae* in carbohydrate-rich  
624 environments. *Mol Microbiol* **97**: 360–380.
- 625 García-Solache, M., and Rice, L.B. (2019) The *Enterococcus*: a Model of Adaptability to Its  
626 Environment. *Clin Microbiol Rev* **32**.
- 627 Garsin, D.A., Frank, K.L., Silanpää, J., Ausubel, F.M., Hartke, A., Shankar, N., and Murray,  
628 B.E. (2014) Pathogenesis and Models of Enterococcal Infection. In *Enterococci: From*  
629 *Commensals to Leading Causes of Drug Resistant Infection*. Gilmore, M.S., Clewell, D.B., Ike,  
630 Y., and Shankar, N. (eds). Massachusetts Eye and Ear Infirmary, Boston.  
631 <http://www.ncbi.nlm.nih.gov/books/NBK190426/>.

- 632 Gopal, S., Berg, D., Hagen, N., Schriefer, E.-M., Stoll, R., Goebel, W., and Kreft, J. (2010)  
633 Maltose and Maltodextrin Utilization by *Listeria monocytogenes* Depend on an Inducible ABC  
634 Transporter which Is Repressed by Glucose. *PLoS ONE* **5**: e10349.
- 635 Grøftheauge, M.K., Hajizadeh, N.R., Swann, M.J., and Pohl, E. (2015) Protein–ligand  
636 interactions investigated by thermal shift assays (TSA) and dual polarization interferometry  
637 (DPI). *Acta Crystallogr D Biol Crystallogr* **71**: 36–44.
- 638 Homburg, C., Bommer, M., Wuttge, S., Hobe, C., Beck, S., Dobbek, H., *et al.* (2017) Inducer  
639 exclusion in Firmicutes: Insights into the regulation of a carbohydrate ATP binding cassette  
640 transporter from *Lactobacillus casei* BL23 by the signal transducing protein P-Ser46-HPr. *Mol*  
641 *Microbiol* .
- 642 Hufnagel, M., Koch, S., Creti, R., Baldassarri, L., and Huebner, J. (2004) A putative sugar-  
643 binding transcriptional regulator in a novel gene locus in *Enterococcus faecalis* contributes to  
644 production of biofilm and prolonged bacteremia in mice. *J Infect Dis* **189**: 420–430.
- 645 Iyer, R., Baliga, N.S., and Camilli, A. (2005) Catabolite control protein A (CcpA) contributes  
646 to virulence and regulation of sugar metabolism in *Streptococcus pneumoniae*. *J Bacteriol* **187**:  
647 8340–8349.
- 648 Jones, B.E., Dossonnet, V., Küster, E., Hillen, W., Deutscher, J., and Klevit, R.E. (1997)  
649 Binding of the catabolite repressor protein CcpA to its DNA target is regulated by  
650 phosphorylation of its corepressor HPr. *J Biol Chem* **272**: 26530–26535.
- 651 Joyet, P., Mokhtari, A., Riboulet-Bisson, E., Blancato, V.S., Espariz, M., Magni, C., *et al.*  
652 (2017) The enzymes required for maltodextrin catabolism in *Enterococcus faecalis* exhibit  
653 several novel activities. *Appl Environ Microbiol* .
- 654 Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015) The Phyre2  
655 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**: 845–858.
- 656 Kim, J.H., Guvener, Z.T., Cho, J.Y., Chung, K.C., and Chambliss, G.H. (1995) Specificity of  
657 DNA binding activity of the *Bacillus subtilis* catabolite control protein CcpA. *J Bacteriol* **177**:  
658 5129–5134.
- 659 Koropatkin, N.M., Cameron, E.A., and Martens, E.C. (2012) How glycan metabolism shapes  
660 the human gut microbiota. *Nat Rev Microbiol* **10**: 323–335.
- 661 Kravanja, M., Engelmann, R., Dossonnet, V., Blüggel, M., Meyer, H.E., Frank, R., *et al.* (1999)  
662 The *hprK* gene of *Enterococcus faecalis* encodes a novel bifunctional enzyme: the HPr  
663 kinase/phosphatase. *Mol Microbiol* **31**: 59–66.
- 664 Kristich, C.J., Rice, L.B., and Arias, C.A. (2014) Enterococcal Infection—Treatment and  
665 Antibiotic Resistance. In *Enterococci: From Commensals to Leading Causes of Drug Resistant*  
666 *Infection*. Gilmore, M.S., Clewell, D.B., Ike, Y., and Shankar, N. (eds). Massachusetts Eye and  
667 Ear Infirmary, Boston. <http://www.ncbi.nlm.nih.gov/books/NBK190420/>.
- 668 Lacks, S. (1968) Genetic regulation of maltosaccharide utilization in *Pneumococcus*. *Genetics*  
669 **60**: 685–706.

- 670 Le Breton, Y., Pichereau, V., Sauvageot, N., Auffray, Y., and Rincé, A. (2005) Maltose  
671 utilization in *Enterococcus faecalis*. *J Appl Microbiol* **98**: 806–813.
- 672 Leboeuf, C., Leblanc, L., Auffray, Y., and Hartke, A. (2000) Characterization of the *ccpA* Gene  
673 of *Enterococcus faecalis*: Identification of Starvation-Inducible Proteins Regulated by CcpA. *J*  
674 *Bacteriol* **182**: 5799–5806.
- 675 Lebreton, F., Willems, R.J.L., and Gilmore, M.S. (2014) *Enterococcus* Diversity, Origins in  
676 Nature, and Gut Colonization. In *Enterococci: From Commensals to Leading Causes of Drug*  
677 *Resistant Infection*. Gilmore, M.S., Clewell, D.B., Ike, Y., and Shankar, N. (eds). Massachusetts  
678 Eye and Ear Infirmary, Boston. <http://www.ncbi.nlm.nih.gov/books/NBK190427/>. Accessed  
679 May 26, 2016.
- 680 Ludwig, H., Rebhan, N., Blencke, H.-M., Merzbacher, M., and Stülke, J. (2002) Control of the  
681 glycolytic *gapA* operon by the catabolite control protein A in *Bacillus subtilis*: a novel  
682 mechanism of CcpA-mediated regulation. *Mol Microbiol* **45**: 543–553.
- 683 Martin, S.A., and Russell, J.B. (1987) Transport and phosphorylation of disaccharides by the  
684 ruminal bacterium *Streptococcus bovis*. *Appl Env Microbiol* **53**: 2388–2393.
- 685 Mokhtari, A., Blancato, V.S., Repizo, G.D., Henry, C., Pikis, A., Bourand, A., *et al.* (2013)  
686 *Enterococcus faecalis* utilizes maltose by connecting two incompatible metabolic routes via a  
687 novel maltose 6'-phosphate phosphatase (MapP). *Mol Microbiol* **88**: 234–253.
- 688 Monedero, V., Poncet, S., Mijakovic, I., Fieulaine, S., Dossonnet, V., Martin-Verstraete, I., *et*  
689 *al.* (2001) Mutations lowering the phosphatase activity of HPr kinase/phosphatase switch off  
690 carbon metabolism. *EMBO J* **20**: 3928–3937.
- 691 Monedero, V., Yebra, M.J., Poncet, S., and Deutscher, J. (2008) Maltose transport in  
692 *Lactobacillus casei* and its regulation by inducer exclusion. *Res Microbiol* **159**: 94–102.
- 693 Muller, C., Cacaci, M., Sauvageot, N., Sanguinetti, M., Rattei, T., Eder, T., *et al.* (2015) The  
694 Intraperitoneal Transcriptome of the Opportunistic Pathogen *Enterococcus faecalis* in Mice.  
695 *PloS One* **10**: e0126143.
- 696 Müller, W., Horstmann, N., Hillen, W., and Sticht, H. (2006) The transcription regulator RbsR  
697 represents a novel interaction partner of the phosphoprotein HPr-Ser46-P in *Bacillus subtilis*.  
698 *FEBS J* **273**: 1251–1261.
- 699 Ng, K.M., Ferreyra, J.A., Higginbottom, S.K., Lynch, J.B., Kashyap, P.C., Gopinath, S., *et al.*  
700 (2013) Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric  
701 pathogens. *Nature* **502**: 96–99.
- 702 Nieto, C., Espinosa, M., and Puyet, A. (1997) The maltose/maltodextrin regulon of  
703 *Streptococcus pneumoniae* Differential promoter regulation by the transcriptional repressor  
704 MalR. *J Biol Chem* **272**: 30860–30865.
- 705 Nieto, C., Puyet, A., and Espinosa, M. (2001) MalR-mediated Regulation of the *Streptococcus*  
706 *pneumoniae* malMP Operon at PromoterPM: influence of a proximal divergent promoter  
707 region and competition between MalR and RNA polymerase proteins. *J Biol Chem* **276**: 14946–  
708 14954.

- 709 O’Leary, N.A., Wright, M.W., Brister, J.R., Ciuffo, S., Haddad, D., McVeigh, R., *et al.* (2016)  
710 Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and  
711 functional annotation. *Nucleic Acids Res* **44**: D733-745.
- 712 Opsata, M., Nes, I.F., and Holo, H. (2010) Class IIa bacteriocin resistance in *Enterococcus*  
713 *faecalis* V583: The mannose PTS operon mediates global transcriptional responses. *BMC*  
714 *Microbiol* **10**: 224.
- 715 Paluscio, E., Watson, M.E., and Caparon, M.G. (2018) CcpA Coordinates Growth/Damage  
716 Balance for *Streptococcus pyogenes* Pathogenesis. *Sci Rep* **8**: 14254.
- 717 Platteeuw, C., Simons, G., and Vos, W.M. de (1994) Use of the *Escherichia coli* beta-  
718 glucuronidase (*gusA*) gene as a reporter gene for analyzing promoters in lactic acid bacteria.  
719 *Appl Environ Microbiol* **60**: 587–593.
- 720 Puyet, A. (1993) Expression of bacterial genes involved in maltose metabolism. *World J*  
721 *Microbiol Biotechnol* **9**: 455–460.
- 722 Robert C. Moellering, J. (1992) Emergence of *Enterococcus* as a Significant Pathogen. *Clin*  
723 *Infect Dis* **14**: 1173–1176.
- 724 Ruiz-Cruz, S., Espinosa, M., Goldmann, O., and Bravo, A. (2016) Global Regulation of Gene  
725 Expression by the MafR Protein of *Enterococcus faecalis*. *Front Microbiol* **6**  
726 <http://journal.frontiersin.org/Article/10.3389/fmicb.2015.01521/abstract>. Accessed November  
727 10, 2016.
- 728 Sato, Y., Okamoto-Shibayama, K., and Azuma, T. (2013) The *malQ* gene is essential for starch  
729 metabolism in *Streptococcus mutans*. *J Oral Microbiol* **5**.
- 730 Sauvageot, N., Mokhtari, A., Joyet, P., Budin-Verneuil, A., Blancato, V.S., Repizo, G.D., *et al.*  
731 (2017) *Enterococcus faecalis* uses a PTS permease and a host colonization-related ABC  
732 transporter for maltodextrin uptake. *J Bacteriol* .
- 733 Schumacher, M.A., Allen, G.S., Diel, M., Seidel, G., Hillen, W., and Brennan, R.G. (2004)  
734 Structural Basis for Allosteric Control of the Transcription Regulator CcpA by the  
735 Phosphoprotein HPr-Ser46-P. *Cell* **118**: 731–741.
- 736 Schumacher, M.A., Seidel, G., Hillen, W., and Brennan, R.G. (2007) Structural Mechanism for  
737 the Fine-tuning of CcpA Function by The Small Molecule Effectors Glucose 6-Phosphate and  
738 Fructose 1,6-Bisphosphate. *J Mol Biol* **368**: 1042–1050.
- 739 Shelburne, S.A., Keith, D.B., Davenport, M.T., Horstmann, N., Brennan, R.G., and Musser,  
740 J.M. (2008) Molecular characterization of group A *Streptococcus* maltodextrin catabolism and  
741 its role in pharyngitis. *Mol Microbiol* **69**: 436–452.
- 742 Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., *et al.* (2011) Fast, scalable  
743 generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst*  
744 *Biol* **7**: 539.
- 745 Sonnenburg, E.D., Zheng, H., Joglekar, P., Higginbottom, S.K., Firkbank, S.J., Bolam, D.N., and  
746 Sonnenburg, J.L. (2010) Specificity of polysaccharide use in intestinal bacteroides species  
747 determines diet-induced microbiota alterations. *Cell* **141**: 1241–1252.

- 748 Stassi, D.L., Lopez, P., Espinosa, M., and Lacks, S.A. (1981) Cloning of chromosomal genes  
749 in *Streptococcus pneumoniae*. *Proc Natl Acad Sci U S A* **78**: 7028–7032.
- 750 Suárez, C.A., Blancato, V.S., Poncet, S., Deutscher, J., and Magni, C. (2011) CcpA represses  
751 the expression of the divergent *cit* operons of *Enterococcus faecalis* through multiple *cre* sites.  
752 *BMC Microbiol* **11**: 1.
- 753 Swint-Kruse, L., and Matthews, K.S. (2009) Allosterity in the LacI/GalR family: variations on a  
754 theme. *Curr Opin Microbiol* **12**: 129–137.
- 755 Tartof, K.D., and Hobbs, C.A. (1987) Improved media for growing plasmid and cosmid clones.  
756 *Bethesda Res Lab Focus* **9**: 12.
- 757 Thomas, M., Wrzosek, L., Ben-Yahia, L., Noordine, M.-L., Gitton, C., Chevret, D., *et al.* (2011)  
758 Carbohydrate metabolism is essential for the colonization of *Streptococcus thermophilus* in the  
759 digestive tract of gnotobiotic rats. *PloS One* **6**: e28789.
- 760 Thompson, J., Robrish, S.A., Pikis, A., Brust, A., and Lichtenthaler, F.W. (2001)  
761 Phosphorylation and metabolism of sucrose and its five linkage-isomeric alpha-D-glucosyl-D-  
762 fructoses by *Klebsiella pneumoniae*. *Carbohydr Res* **331**: 149–161.
- 763 Thompson, J., and Torchia, D.A. (1984) Use of <sup>31</sup>P nuclear magnetic resonance spectroscopy  
764 and <sup>14</sup>C fluorography in studies of glycolysis and regulation of pyruvate kinase in  
765 *Streptococcus lactis*. *J Bacteriol* **158**: 791–800.
- 766 Vebø, H.C., Snipen, L., Nes, I.F., and Brede, D.A. (2009) The transcriptome of the nosocomial  
767 pathogen *Enterococcus faecalis* V583 reveals adaptive responses to growth in blood. *PloS One*  
768 **4**: e7660.
- 769 Viana, R., Monedero, V., Dossonnet, V., Vadeboncoeur, C., Pérez-Martínez, G., and Deutscher,  
770 J. (2000) Enzyme I and HPr from *Lactobacillus casei*: their role in sugar transport, carbon  
771 catabolite repression and inducer exclusion. *Mol Microbiol* **36**: 570–584.
- 772 Vivoli, M., Novak, H.R., Littlechild, J.A., and Harmer, N.J. (2014) Determination of protein-  
773 ligand interactions using differential scanning fluorimetry. *J Vis Exp JoVE* 51809.
- 774 Webb, A.J., Homer, K.A., and Hosie, A.H.F. (2007) A phosphoenolpyruvate-dependent  
775 phosphotransferase system is the principal maltose transporter in *Streptococcus mutans*. *J*  
776 *Bacteriol* **189**: 3322–3327.
- 777 Weickert, M.J., and Adhya, S. (1992) A family of bacterial regulators homologous to Gal and  
778 Lac repressors. *J Biol Chem* **267**: 15869–15874.
- 779 Wisplinghoff, H., Bischoff, T., Tallent, S.M., Seifert, H., Wenzel, R.P., and Edmond, M.B.  
780 (2004) Nosocomial Bloodstream Infections in US Hospitals: Analysis of 24,179 Cases from a  
781 Prospective Nationwide Surveillance Study. *Clin Infect Dis* **39**: 309–317.
- 782 Xu, J., Xie, Y.-D., Liu, L., Guo, S., Su, Y.-L., and Li, A.-X. (2019) Virulence regulation of cel-  
783 EIIB protein mediated PTS system in *Streptococcus agalactiae* in Nile tilapia. *J Fish Dis* **42**:  
784 11–19.

- 785 Ye, J.-J., Minarcik, J., and Saier, M.H. (1996) Inducer Expulsion and the Occurrence of an  
786 HPr(Ser-P)-Activated Sugar-Phosphate Phosphatase in *Enterococcus Faecalis* and  
787 *Streptococcus Pyogenes*. *Microbiology* **142**: 585–592.
- 788 Yip, V.L.Y., Thompson, J., and Withers, S.G. (2007) Mechanism of GlvA from *Bacillus*  
789 *subtilis*: a detailed kinetic analysis of a 6-phospho-alpha-glucosidase from glycoside hydrolase  
790 family 4. *Biochemistry (Mosc)* **46**: 9840–9852.
- 791 Zhang, X., Rogers, M., Bierschenk, D., Bonten, M.J.M., Willems, R.J.L., and Schaik, W. van  
792 (2013) A LacI-Family Regulator Activates Maltodextrin Metabolism of *Enterococcus faecium*.  
793 *PLoS ONE* **8**: e72285.
- 794

795 **Figures legends**

796

797 **Fig. 1.** Presentation of the two operons required for maltose metabolism in *E. faecalis* OG1RF.

798 **A.** Large arrows represent ORFs with their orientation indicating the direction of transcription.

799 Gene names are written in the arrows, predicted or proven functions are indicated above and

800 reference loci below the arrows. Promoters are symbolized by thin bent arrows. **B.** Presented

801 are the nucleotide sequence of the chromosomal region between the *malM* and *malR* genes of

802 *E. faecalis* OG1RF and the sequence data obtained with 5'-RACE-PCR. The first nucleotide

803 downstream from the poly-A tail corresponds to the first transcribed nucleotide and is marked

804 +1. The putative Pribnow box and Ribosome Binding Site (RBS) of *malR* are underlined.

805

806 **Fig. 2.** Relative expression of *malP* (**A**) and *malT* (**B**) in the parental strain OG1RF and the

807 *malR*<sup>-</sup> and  $\Delta$ *ccpA* mutants. Expression of the housekeeping *gyrA* gene was used as reference to

808 normalize the total quantity of cDNA. Median values calculated from at least three independent

809 experiments are presented; error bars represent the data range.

810

811 **Fig. 3.** TSA performed with 2  $\mu$ M of MalR. The temperature-induced exposure of hydrophobic

812 surfaces ( $\Delta T_h$ ) in response to increasing concentrations of various compounds is presented:

813 maltose (**A**), maltose-6'-P (**B**) and fructose-1,6-bisphosphate (**C**). The mean values of three

814 independent experiments are presented, errors bars indicate the standard deviations. Full lines

815 represent non-linear regression fitting data points. Horizontal dashed lines represent a  $\Delta T$  of

816 0°C. Dissociation constants were determined using Differential Scanning Fluorimetry assuming

817 a single binding event model with GraphPad Prism version 7.

818

819 **Fig. 4.** EMSA with DNA fragments containing one or both *Pmal*. **A.** Schematic presentation of  
820 DNA regions used as template for EMSA. The assays were performed with 5 nM of Cyanine-5  
821 labelled PCR products. **B.** EMSA results for MalR binding to *P2mal*. Maltose (final  
822 concentration of 5 mM) was added to the samples loaded on lanes 4, 5, 8 and 9 and P-Ser-HPr  
823 was added to the samples loaded on lanes 6 or 9. **C.** EMSA results for MalR binding to *PmalP*  
824 (lanes 1 to 3), *PmalT* (lanes 4 to 6) and an internal fragment of the *gusA* gene (lanes 7 to 9).

825

826 **Fig. 5.** DNase I footprinting with a DNA fragment containing the region upstream from *malP*.  
827 **A** Nucleotide sequence of the chromosomal regions corresponding to the *malP* (upper  
828 sequence) and *malT* promoters (lower sequence) of *E. faecalis* OG1RF. The putative Ribosome  
829 Binding Sites (RBS) and transcription start sites (+1) are underlined. Pribnow boxes and -  
830 35 elements are boxed. The presumed MalR binding sites are in bold and underlined. **B.** DNase  
831 I footprinting experiment carried out with the *PmalP* region as template in the absence (upper  
832 panel) and presence (lower panel) of MalR. The two read outs were aligned based on an internal  
833 size standard (500ROX), the positions of which are indicated by diamonds on the diagonal. The  
834 borders of the region protected by the presence of MalR are indicated by the two broken lines.

835

836 **Fig. 6.**  $\beta$ -Glucuronidase activity in *E. faecalis* strains carrying various pNZ273-derived  
837 plasmids. Schematic presentation of the region upstream from the *gusA* gene in the pNZ273-  
838 derived plasmids and  $\beta$ -glucuronidase activity (in arbitrary units) determined with the  
839 corresponding construction. (1) No promoter in front of *gusA*. (2) Presumed *PmalR* inserted  
840 upstream from *gusA* in the correct orientation and (3) in the opposite orientation.  
841  $\beta$ -glucuronidase activities were measured after growth of the various *E. faecalis* strains

842 transformed with pNZ273-derivative grown in medium containing 0.3% glucose or 0.3%  
843 maltose. The mean values of three independent experiments together with standard deviations  
844 are presented. Asterisks indicate significant differences (\*\*:  $p < 0.01$  ; \*\*\*:  $p < 0.001$ ) determined  
845 with Holm-Sidak multiple comparisons versus the reference group pNZ273 without assuming  
846 equal standard deviations. Statistical analyses were performed with GraphPad Prism version 7  
847 software.

848

849 **Fig. 7.** MST assays were carried out with 5 nM of NT.647-labelled proteins (MalR or CcpA)  
850 and different concentrations of either HPr, P-Ser-HPr or P~His-HPr. The fraction of bound  
851 labelled protein is shown as a function of the concentration of the various HPr forms ranging  
852 from 9.76 nM to 20  $\mu$ M. **A.** MalR was titrated with P-Ser-HPr; the fitting curve is shown as  
853 solid line. **B.** MalR was titrated with HPr. **C.** CcpA was titrated with P-Ser-HPr; the fitting  
854 curve is shown as dashed line. **D.** MalR was titrated with P~His-HPr. For **A** and **C**, analysis  
855 with GraphPad Prism software version 7 allowed the determination of the apparent  $K_D$  by  
856 assuming one specific binding site.

857

858 **Fig. 8.** Alignment of partial primary sequences of different interaction partners of P-Ser-HPr.  
859 The number of amino acids indicated below the sequences are based on the position in each  
860 studied protein. Amino acids identical in MalR and/or RbsR to amino acids of one of the CcpA's  
861 are in bold, conserved amino acids are underlined. Amino acids of CcpA which have been  
862 shown to interact with P-Ser-HPr are marked with a rectangle. Regions identified as CcpA-  
863 specific are underlined in grey, the insertion in MalR is marked in italics. **A.** D69, D87 and D99  
864 of one subunit of CcpA (*B. megaterium*) interact with R17 of P-Ser-HPr. **B.** D296 of the other

865 subunit of CcpA (*B. megaterium*) forms hydrogen bonds with unphosphorylated H15 of P-Ser-  
866 HPr, while R303 and K307 interact with the phosphate group bound to S46.

867

868 **Supplementary Figures**

869 **Fig. S1.** Confirmation that *malR* and *malM* are part of the same operon. Schematic  
870 representation of *malR* and *malM* ORFs (large arrows) and the location of the primers used for  
871 PCR (thin arrows). PCR was performed on cDNA after a growth on maltose as sole carbon  
872 source. Genomic DNA was used as positive control and RNAs as negative control confirming  
873 the absence of contaminating DNA. Based on the DNA ladder, the amplified PCR products  
874 exhibit a length of about 1250 base pairs, which fits well the calculated size of 1282 bp.

875

876 **Fig. S2.** Relative expression of *malP* (A) and *malT* (B) in the parental strain OG1RF, the *malR*<sup>-</sup>  
877 mutant and the complemented *malR*<sup>-</sup> mutant carrying the pRB\_malR plasmid. The *malR*<sup>-</sup> strain  
878 carrying the empty plasmid pRB473 was used to confirm that the expression profile remains  
879 unchanged by the presence of the plasmid. Expression of the housekeeping *gyrA* gene was used  
880 as reference to normalize the total quantity of cDNA. Median values calculated from at least  
881 three independent experiments are presented; error bars indicate the data range.

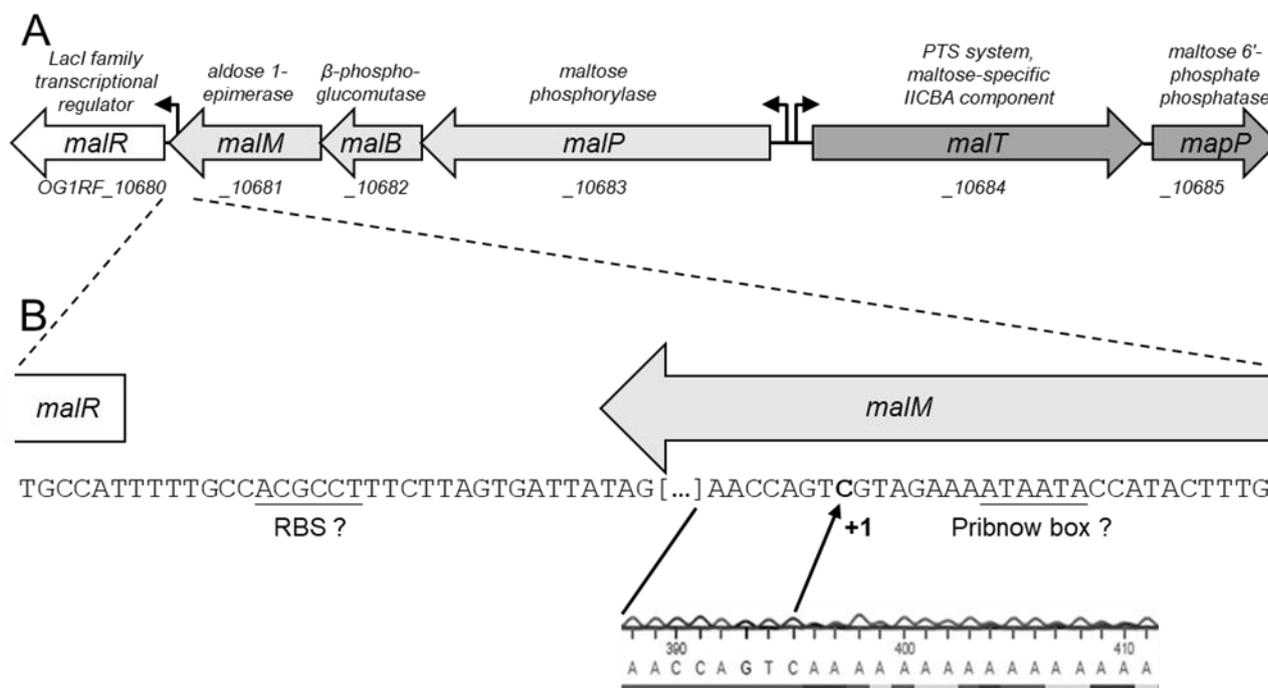
882

883 **Fig. S3.** EMSA with PCR products containing either P2*mal* or an internal *gusA* fragment. The  
884 PCR products were labelled with 5 nM of Cyanine-5. **A.** EMSA was carried out with DNA  
885 fragments containing either P2*mal* (lanes 1 to 5) or an internal *gusA* fragment (lanes 6 to 10), 2  
886 μM P-Ser-HPr and various concentrations of CcpA (1.56 – 12.5 nM). **B.** EMSA carried out  
887 with a P2*mal*-containing PCR fragment and 2 μM of either HPr (lanes 2 and 5), P~His-HPr  
888 (lanes 3 and 6) or P-Ser-HPr (lanes 4 and 7) in the presence (lanes 2, 3 and 4) or absence (lanes  
889 1, 5, 6 and 7) of MalR.

890

891 **Fig. S4.** TSA performed with 2  $\mu\text{M}$  of MalR. The temperature-induced exposure of  
892 hydrophobic surfaces ( $\Delta T_h$ ) in response to increasing concentrations of various compounds is  
893 presented: glucose (**A**), glucose-6-P (**B**), maltotriose (**C**) and maltotriose-6"-P (**D**). The mean  
894 values of three independent experiments are presented; errors bars indicate the standard  
895 deviations. Horizontal dashed lines represent a  $\Delta T$  of 0°C. Dissociation constants were  
896 determined using Differential Scanning Fluorimetry assuming a single binding event model  
897 with GraphPad Prism version 7.

Figure 1



**Fig. 1.** Presentation of the two operons required for maltose metabolism in *E. faecalis* OG1RF. **A.** Large arrows represent ORFs with their orientation indicating the direction of transcription. Gene names are written in the arrows, predicted or proven functions are indicated above and reference loci below the arrows. Promoters are symbolized by thin bent arrows. **B.** Presented are the nucleotide sequence of the chromosomal region between the *malM* and *malR* genes of *E. faecalis* OG1RF and the sequence data obtained with 5'-RACE-PCR. The first nucleotide downstream from the poly-A tail corresponds to the first transcribed nucleotide and is marked +1. The putative Pribnow box and Ribosome Binding Site (RBS) of *malR* are underlined.

Figure 2

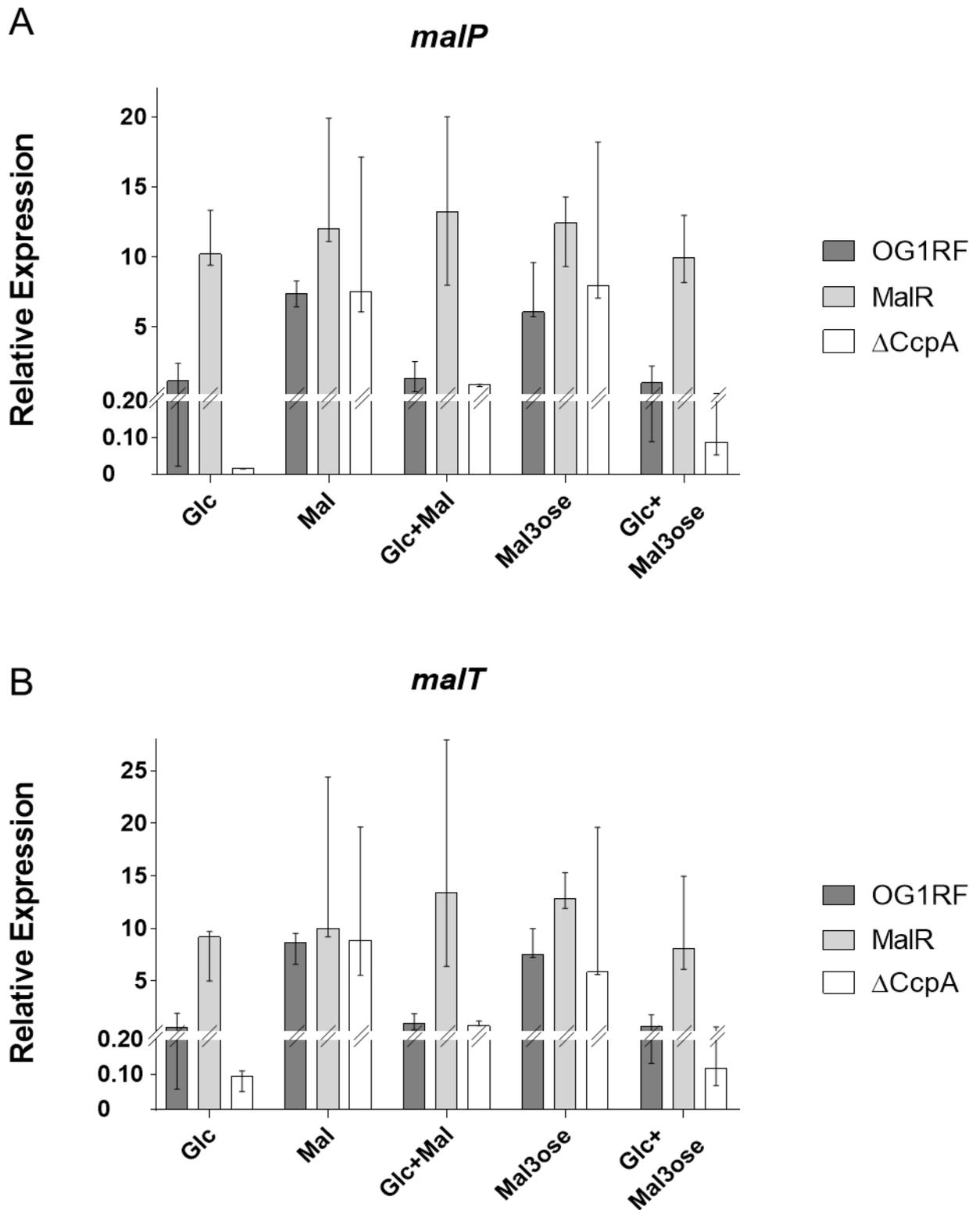
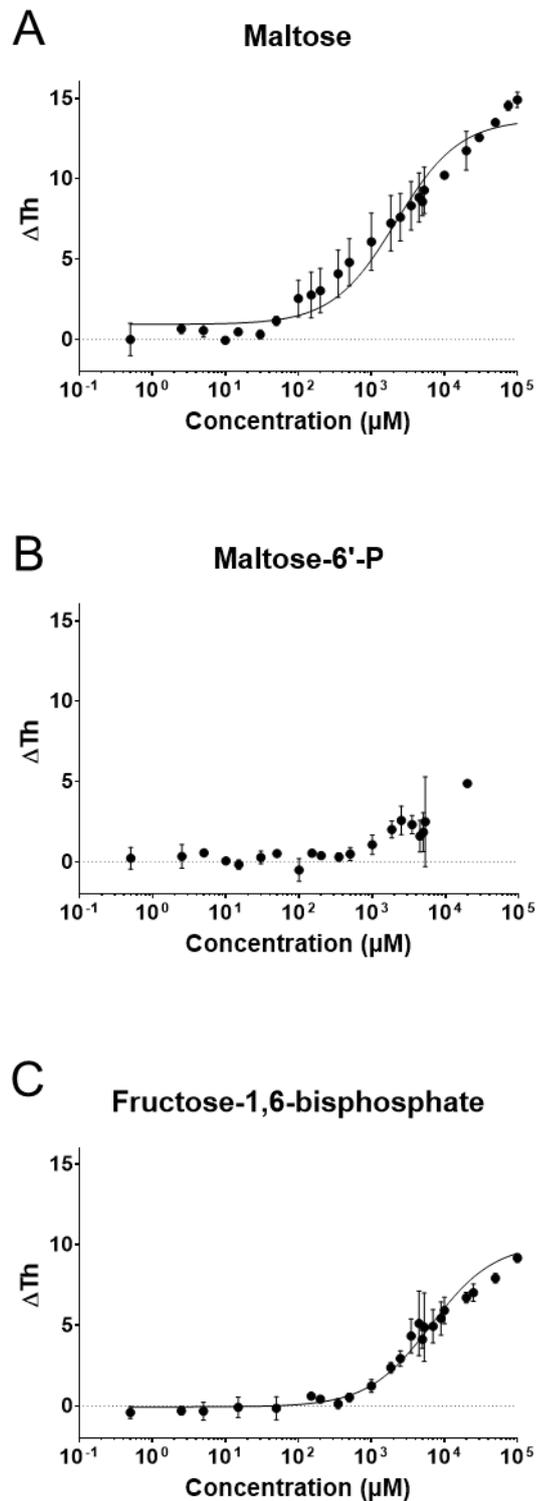


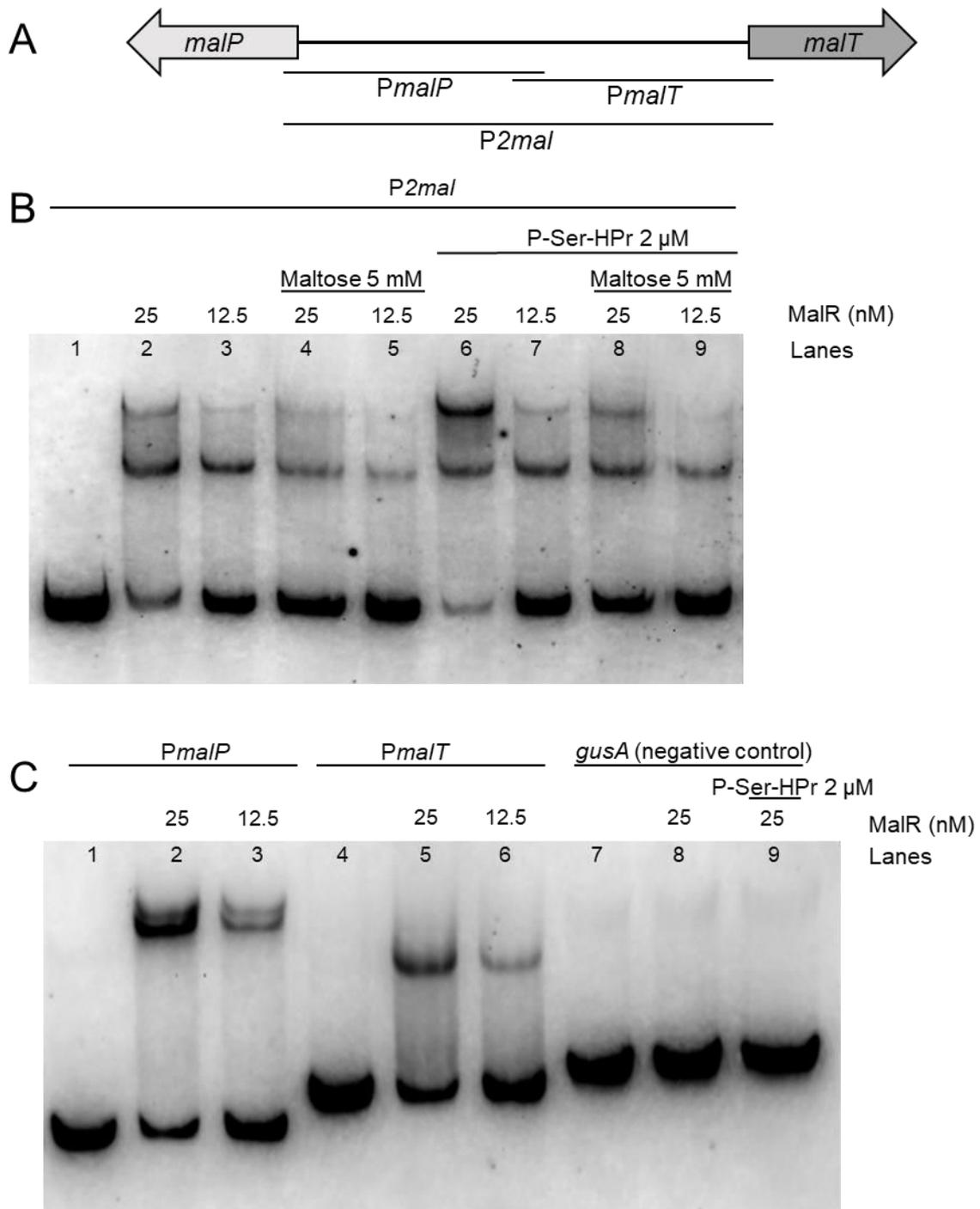
Fig. 2. Relative expression of *malP* (A) and *malT* (B) in the parental strain OG1RF and the *malR* and  $\Delta$ ccpA mutants. Expression of the housekeeping *gyrA* gene was used as reference to normalize the total quantity of cDNA. Median values calculated from at least three independent experiments are presented; error bars represent the data range

Figure 3



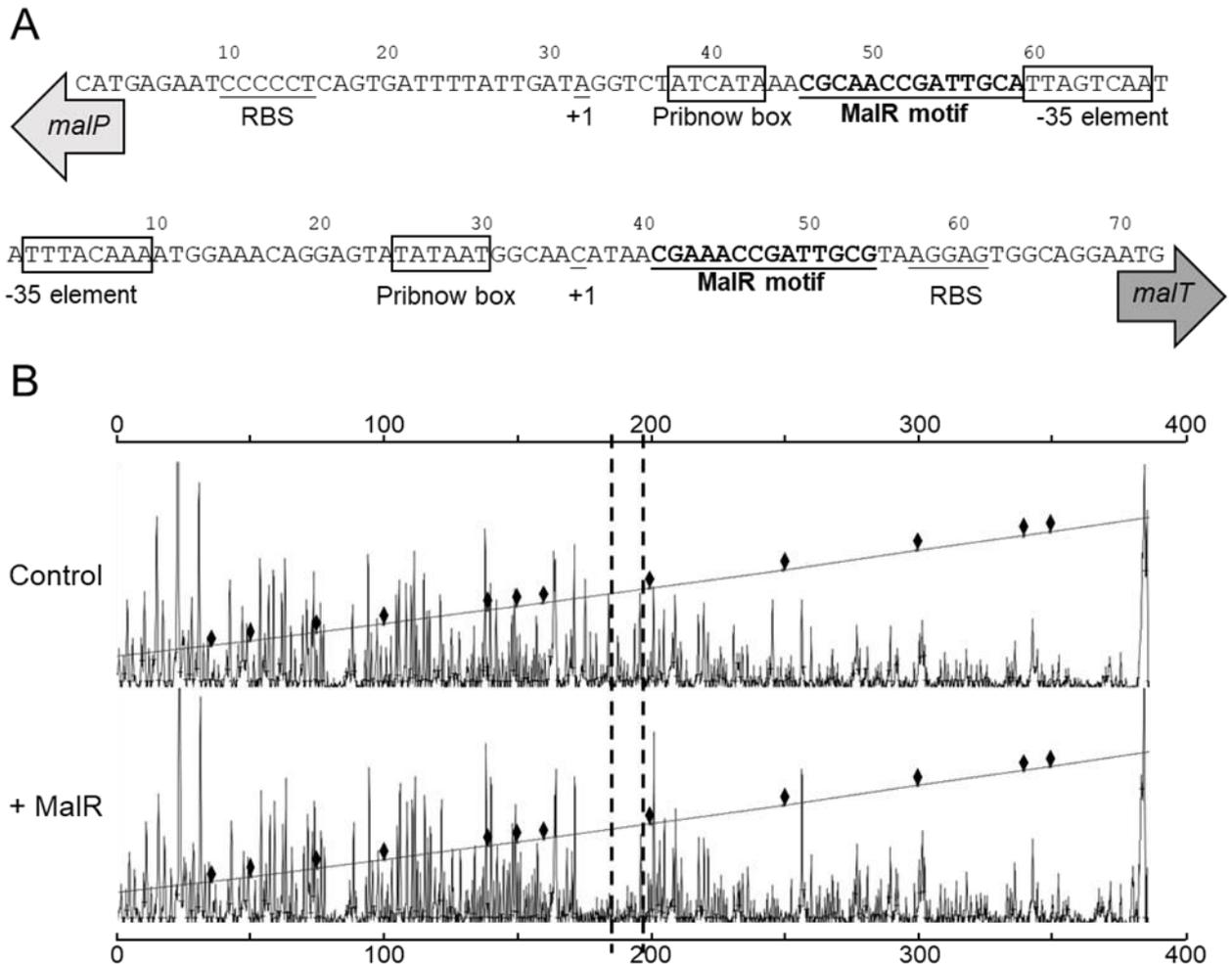
**Fig. 3** TSA performed with 2  $\mu M$  of MalR. The temperature-induced exposure of hydrophobic surfaces ( $\Delta Th$ ) in response to increasing concentrations of various compounds is presented: maltose (A), maltose-6'-P (B) and fructose-1,6-bisphosphate (C). The mean values of three independent experiments are presented, errors bars indicate the standard deviations. Full lines represent non-linear regression fitting data points. Horizontal dashed lines represent a  $\Delta T$  of 0°C. Dissociation constants were determined using Differential Scanning Fluorimetry assuming a single binding event model with GraphPad Prism version 7.

Figure 4



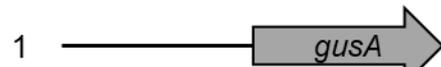
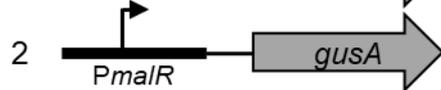
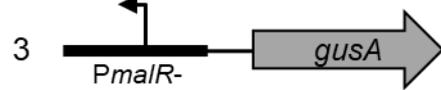
**Fig. 4.** EMSA with DNA fragments containing one or both *Pmal*. **A.** Schematic presentation of DNA regions used as template for EMSA. The assays were performed with 5 nM of Cyanine-5 labeled PCR products. **B.** EMSA results for MalR binding to *P2mal*. Maltose (final concentration of 5 mM) was added to the samples loaded on lanes 4, 5, 8 and 9 and P-Ser-HPr was added to the samples loaded on lanes 6 or 9. **C.** EMSA results for MalR binding to *PmalP* (lanes 1 to 3), *PmalT* (lanes 4 to 6) and an internal fragment of the *gusA* gene (lanes 7 to 9).

Figure 5



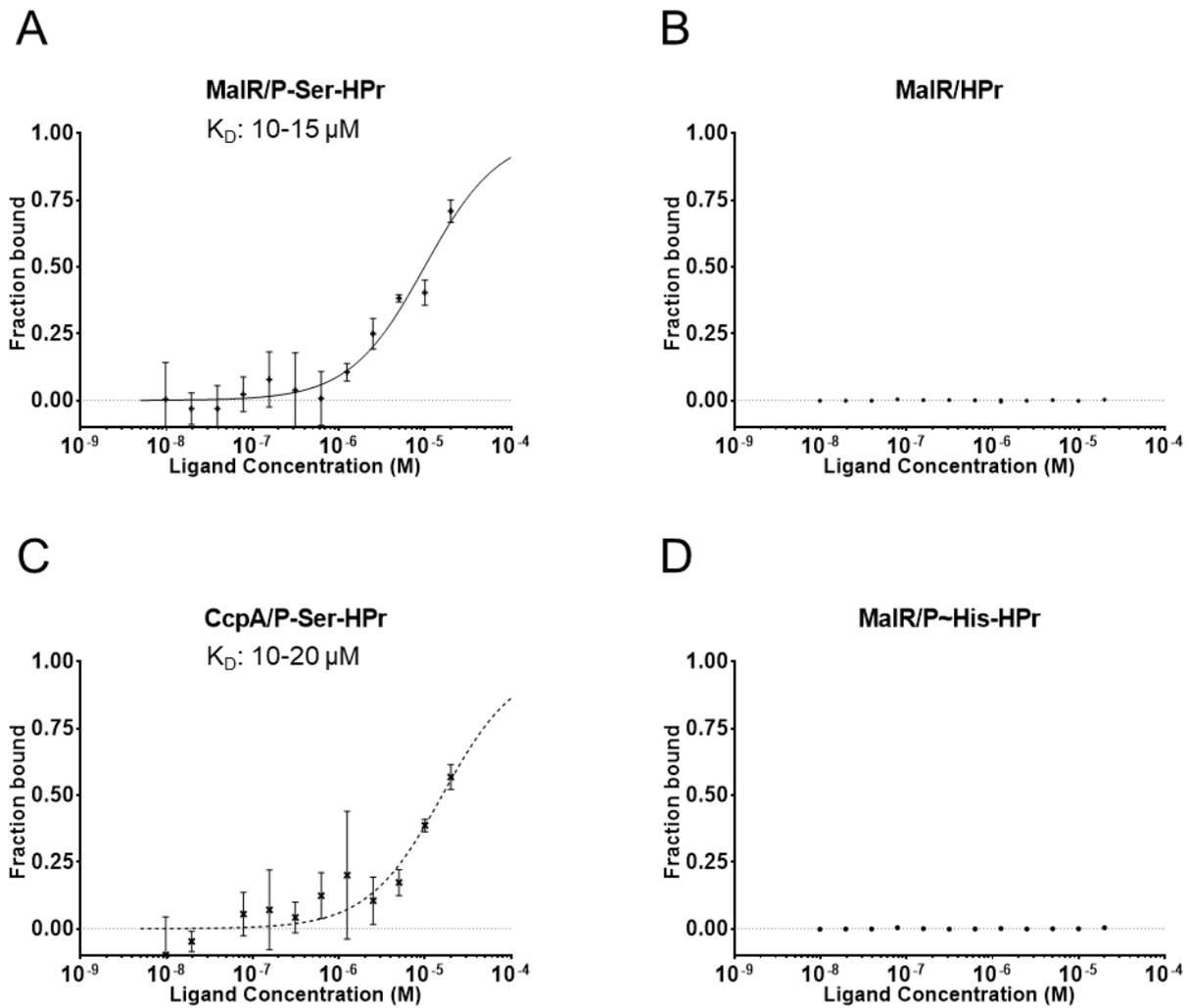
**Fig. 5.** DNase I footprinting with a DNA fragment containing the region upstream from *malP*. **A** Nucleotide sequence of the chromosomal regions corresponding to the *malP* (upper sequence) and *malT* promoters (lower sequence) of *E. faecalis* OG1RF. The putative Ribosome Binding Sites (RBS) and transcription start sites (+1) are underlined. Pribnow boxes and -35 elements are boxed. The presumed MalR binding sites are in bold and underlined. **B.** DNase I footprinting experiment carried out with the *PmalP* region as template in the absence (upper panel) and presence (lower panel) of MalR. The two read outs were aligned based on an internal size standard (500ROX), the positions of which are indicated by diamonds on the diagonal. The borders of the region protected by the presence of MalR are indicated by the two broken lines.

Figure 6

		Glucuronidase activity (AU)		
		Glucose	Maltose	
1		PNZ273	3.12 ± 1.81	4.70 ± 3.39
2		PNZ_PmalR	30.09 ± 3.49 (***)	28.77 ± 3.54 (**)
3		PNZ_MalR-	4.94 ± 3.20	5.49 ± 3.56

**Fig.6.**  $\beta$ -Glucuronidase activity in *E. faecalis* strains carrying various pNZ273-derived plasmids. Schematic presentation of the region upstream from the *gusA* gene in the pNZ273-derived plasmids and  $\beta$ -glucuronidase activity (in arbitrary units) determined with the corresponding construction. (1) No promoter in front of *gusA*. (2) Presumed *PmalR* inserted upstream from *gusA* in the correct orientation and (3) in the opposite orientation.  $\beta$ -glucuronidase activities were measured after growth of the various *E. faecalis* strains transformed with pNZ273-derivative grown in medium containing 0.3% glucose or 0.3% maltose. The mean values of three independent experiments together with standard deviations are presented. Asterisks indicate significant differences (\*\*:  $p < 0.01$  ; \*\*\*:  $p < 0.001$ ) determined with Holm-Sidak multiple comparisons versus the reference group pNZ273 without assuming equal standard deviations. Statistical analyses were performed with GraphPad Prism version 7 software.

Figure 7



**Fig. 7.** MST assays were carried out with 5 nM of NT.647-labeled proteins (MaIR or CcpA) and different concentrations of either HPr, P-Ser-HPr or P~His-HPr. The fraction of bound labelled protein is shown as a function of the concentration of the various HPr forms ranging from 9.76 nM to 20  $\mu$ M. **A.** MaIR was titrated with P-Ser-HPr; the fitting curve is shown as solid line. **B.** MaIR was titrated with HPr. **C.** CcpA was titrated with P-Ser-HPr; the fitting curve is shown as dashed line. **D.** MaIR was titrated with P~His-HPr. For **A** and **C**, analysis with GraphPad Prism software version 7 allowed the determination of the apparent  $K_D$  by assuming one specific binding site.

## Figure 8

A

```

CcpA (E. faecalis)   GVIIP-----DVSNAFFASLARGIDDVATMYKYNIILANSDGDD
67                   72                   87                   102
CcpA (B. subtilis)   GVIIP-----DISSIFYSELARGIEDIATMYKYNIILSNSDQNM
65                   70                   85                   100
CcpA (B. megaterium) GVIIP-----DISNIFYAELARGIEDIATMYKYNIILSNSDQNQ
64                   69                   84                   99
MalR (E. faecalis)   GVILPPLDSKERLGNPFYLEIMEAINEEARLYDMTTAIATAKSFD
64                   75                   90                   105
RbsR (B. subtilis)   GLLLP-----DITNPFFPQLARGAEDELNREGYRLIFGNSDEEL
63                   68                   83                   98
  
```

B

```

CcpA (E. faecalis)   PRLSSITQPLYDIGAVSMRLLTKLMNKE
286                   297                   304   308   313
CcpA (B. subtilis)   PQLSTVVQPTYDIGAVAMRLLTKLMNKE
286                   297                   304   308   313
CcpA (B. megaterium) PQLTSVVQPMYDIGAVAMRLLTKYMNKE
285                   296                   303   307   312
MalR (E. faecalis)   PYLTSIDIDTSELGRMAMKKLYELIHEE
286                   297                   304   308   313
RbsR (B. subtilis)   PPLSTIKQPAYDMGKEAAKLLLGIIKKQ
276                   287                   294   298   303
  
```

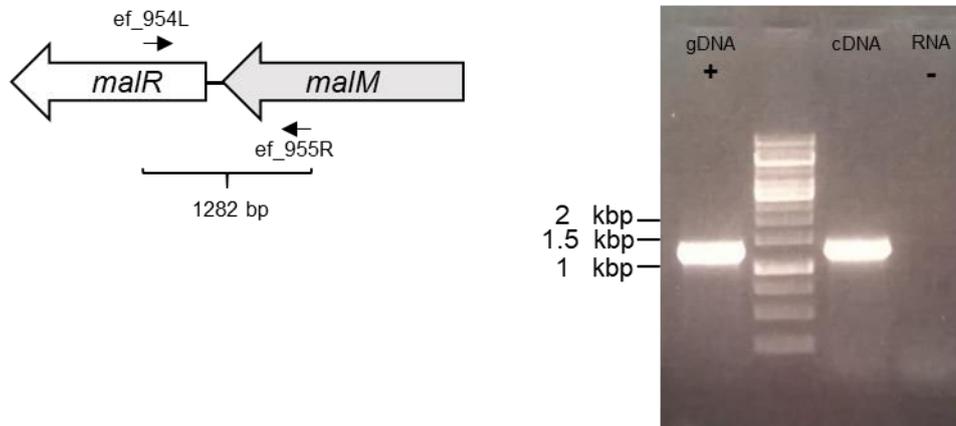
**Fig. 8.** Alignment of partial primary sequences of different interaction partners of P-Ser-HPr. The number of amino acids indicated below the sequences are based on the position in each studied protein.

Amino acids identical in MalR and/or RbsR to amino acids of one of the CcpA's are in bold, conserved amino acids are underlined. Amino acids of CcpA which have been shown to interact with P-Ser-HPr are marked with a rectangle. Regions identified as CcpA-specific are underlined in grey, the insertion in MalR is marked in italics.

**A.** D69, D87 and D99 of one subunit of CcpA (*B. megaterium*) interact with R17 of P-Ser-HPr.

**B.** D296 of the other subunit of CcpA (*B. megaterium*) forms hydrogen bonds with unphosphorylated H15 of P-Ser-HPr, while R303 and K307 interact with the phosphate group bound to S46.

## Figure S1



**Fig. S1.** Confirmation that *malR* and *malM* are part of the same operon. Schematic representation of *malR* and *malM* ORFs (large arrows) and the location of the primers used for PCR (thin arrows). PCR was performed on cDNA after a growth on maltose as sole carbon source. Genomic DNA was used as positive control and RNAs as negative control confirming the absence of contaminating DNA. Based on the DNA ladder, the amplified PCR products exhibit a length of about 1250 base pairs, which fits well the calculated size of 1282 bp.

Figure S2

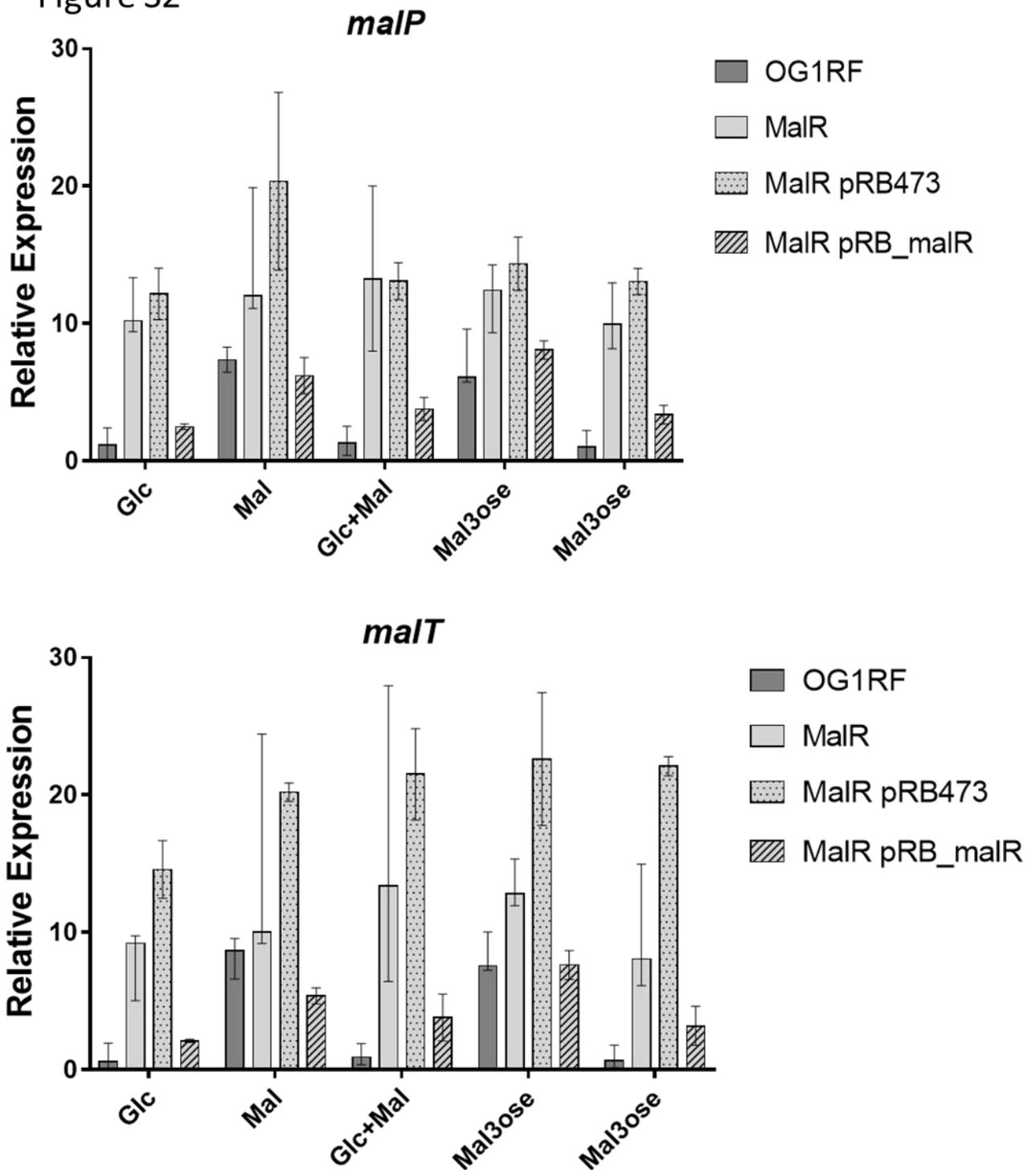
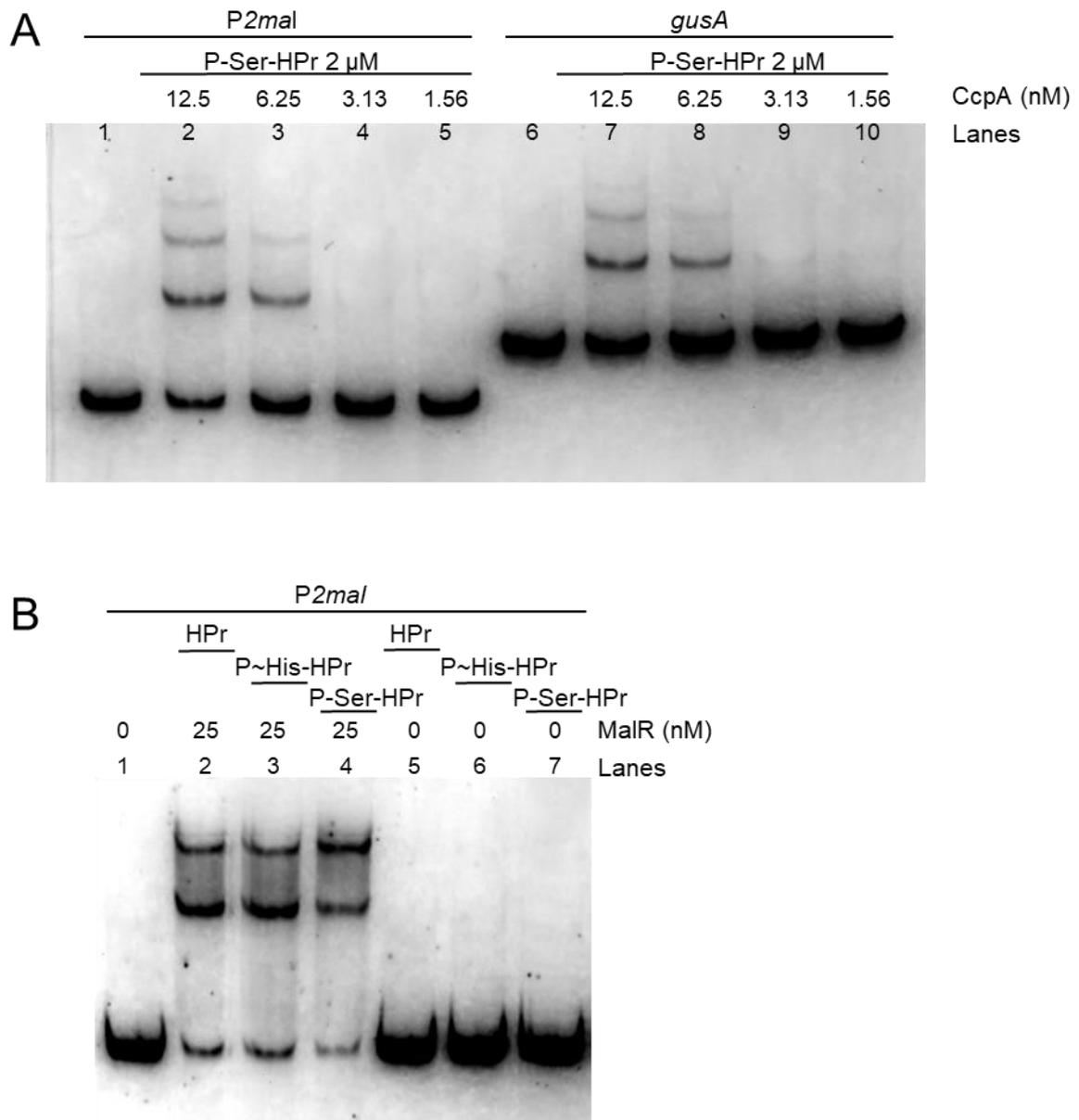


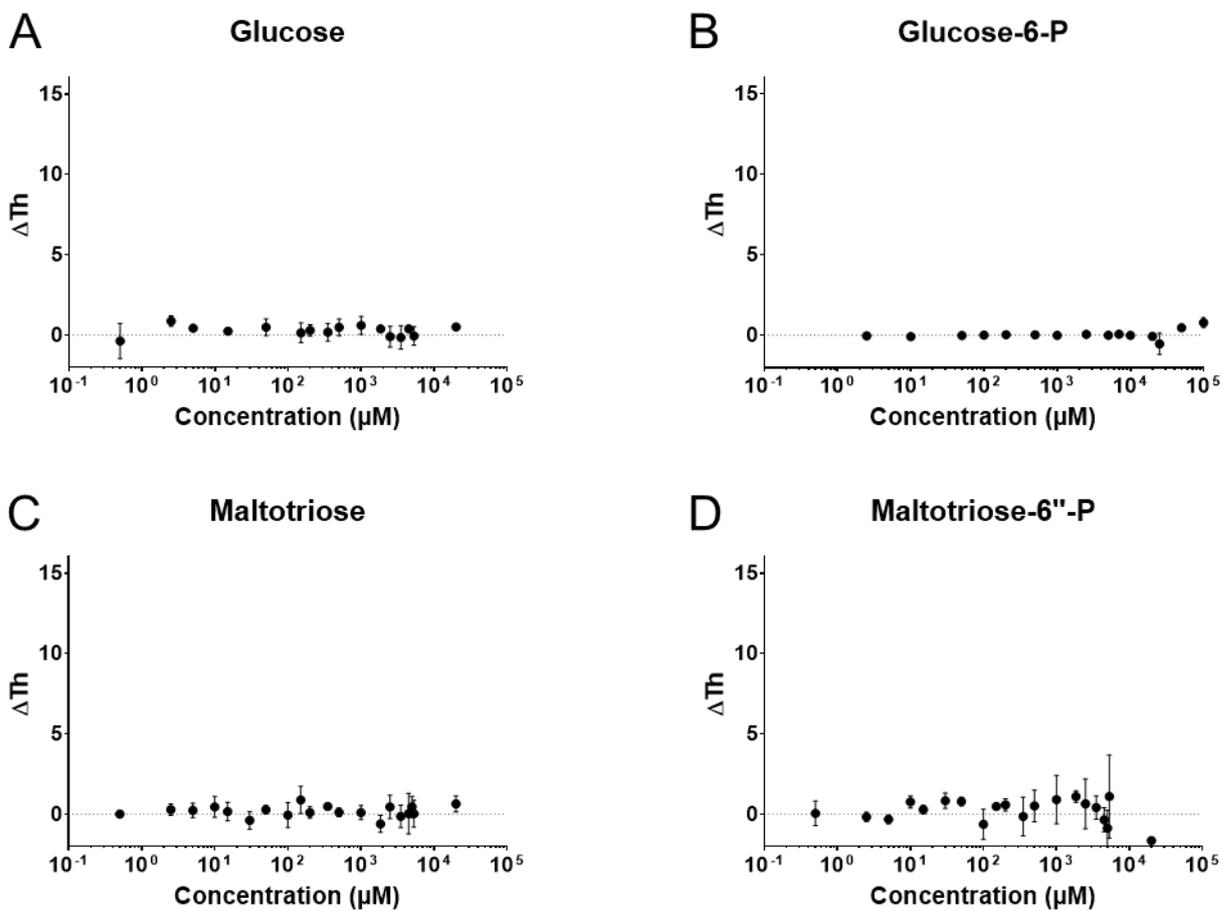
Fig. S2. Relative expression of *malP* (A) and *malT* (B) in the parental strain OG1RF, the *malR* mutant and the complemented *malR* mutant carrying the pRB\_malR plasmid. The *malR*- strain carrying the empty plasmid pRB473 was used to confirm that the expression profile remains unchanged by the presence of the plasmid. Expression of the housekeeping *gyrA* gene was used as reference to normalize the total quantity of cDNA. Median values calculated from at least three independent experiments are presented; error bars indicate the data range.

Figure S3



**Fig.S3.** EMSA with PCR products containing either *P2mal* or an internal *gusA* fragment. The PCR products were labelled with 5 nM of Cyanine-5. **A.** EMSA carried out with a *P2mal*-containing PCR fragment and 2  $\mu$ M of either HPr (lanes 2 and 5), P~His-HPr (lanes 3 and 6) or P-Ser-HPr (lanes 4 and 7) in the presence (lanes 2, 3 and 4) or absence (lanes 1, 5, 6 and 7) of MalR. **B.** EMSA was carried out with DNA fragments containing either *P2mal* (lanes 1 to 5) or an internal *gusA* fragment (lanes 6 to 10), 2 mM P-Ser-HPr and various concentrations of CcpA (1.56 – 12.5 nM).

Figure S4



**Fig. S4** TSA performed with 2  $\mu M$  of MalR. The temperature-induced exposure of hydrophobic surfaces ( $\Delta Th$ ) in response to increasing concentrations of various compounds is presented: glucose (A), glucose-6-P (B), maltotriose (C) and maltotriose-6''-P (D). The mean values of three independent experiments are presented; errors bars indicate the standard deviations. Horizontal dashed lines represent a  $\Delta T$  of 0°C. Dissociation constants were determined using Differential Scanning Fluorimetry assuming a single binding event model with GraphPad Prism version 7.

**Table S1.** Strains used in this study

Strain names	Relevant properties or genotype	Reference or source
<i>Escherichia coli</i>		
<b>M15pRep4</b>	F-, $\Phi$ 80 $\Delta$ <i>lacM15</i> , <i>thi</i> , <i>lac</i> -, <i>mtl</i> -, <i>recA</i> +, Km <sup>R</sup>	Qiagen
<b>NEB5<math>\alpha</math></b>	<i>fhuA2</i> $\Delta$ ( <i>argF-lacZ</i> )U169 <i>phoA glnV44</i> $\Phi$ 80 $\Delta$ ( <i>lacZ</i> )M15 <i>gyrA96</i> <i>recA1 relA1 endA1 thi-1 hsdR17</i>	New England Biolabs
<b>BL21(DE3)</b>	<i>fhuA2</i> [lon] <i>ompT gal</i> ( $\lambda$ DE3) [dcm] $\Delta$ <i>hsdS</i> $\lambda$ DE3 = $\lambda$ sBamHIo $\Delta$ EcoRI-B int::( <i>lacI</i> ::PlacUV5::T7 gene1) <i>i21</i> $\Delta$ <i>nin5</i>	New England Biolabs
<i>Enterococcus faecalis</i>		
<b>OG1RF</b>	Fus <sup>R</sup> , Rif <sup>R</sup> , plasmid-free wild type strain	Dunny, 1978
<b>malR-</b>	<i>malR</i> defective mutant of OG1RF parental strain	This study
<b><math>\Delta</math><i>ccpA</i></b>	<i>ccpA</i> deletion mutant of OG1RF parental strain	This study

**Table S2.** Plasmids used in this study

Plasmids name	Relevant properties	Reference or source
<b>pMAD</b>	Origins of replication pE194ts and pBR322, Em <sup>R</sup> , Amp <sup>R</sup> , <i>bgaB</i>	Arnaud, 2004
<b>pMAD:malR<sup>-</sup></b>	pMAD containing a mutant allele of <i>malR</i> gene (two STOP codons)	This study
<b>pMAD:ΔccpA</b>	pMAD containing a deleted allele of <i>ccpA</i> gene	This study
<b>pQE30</b>	Expression of N-ter-6xHis fusion proteins, Amp <sup>R</sup> , Cm <sup>R</sup> , Ori ColE1	Qiagen
<b>pQE30:HPr</b>	Expression of N-ter-6xHis tagged HPr protein	This study
<b>pET28a (+) &amp; pET28b (+)</b>	Expression of N-ter-6xHis fusion proteins, origin of replication pBR322, <i>lacI<sup>R</sup></i> , Kan <sup>R</sup>	Novagen
<b>pET28a:ccpA</b>	Expression of N-ter-6xHis tagged CcpA protein	This study
<b>pET28b:malR</b>	Expression of N-ter-6xHis tagged MalR protein	This study
<b>pNZ273</b>	Transcriptional fusion experiments, reporter gene <i>gusA</i> , Cm <sup>R</sup> , pSH71 replicon	Platteeuw, 1994
<b>pNZ:PmalR</b>	<i>gusA</i> gene under control of <i>malR</i> promoter	This study
<b>pNZ:PmalR<sup>-</sup></b>	<i>gusA</i> gene preceded by <i>malR</i> promoter in opposite direction	This study
<b>pRB473</b>	Origins of replication pBR322 and pUB110, Amp <sup>R</sup> , Cm <sup>R</sup> , Ble <sup>R</sup>	Brückner, 1993
<b>pRB:malR</b>	pRB473 plasmid containing wild type <i>malR</i> gene of <i>E. faecalis</i> OG1RF strain with its own promoter	
<b>pGEM-T Easy</b>	Cloning of <i>PmalP</i> , <i>PmalT</i> and an internal fragment of the <i>gusA</i> gene to produce a labelled PCR product for EMSA Amp <sup>R</sup>	Promega
<b>pGEM-T PmalP</b>	pGEM-T Easy vector carrying the <i>malP</i> promoter	This study
<b>pGEM-T PmalT</b>	pGEM-T Easy vector carrying the <i>malT</i> promoter	This study
<b>pGEM-T gus</b>	pGEM-T Easy vector containing an internal fragment of the <i>gusA</i> gene	This study

**Table S3.** Primers used in this study

Name	Sequence 5' → 3'	Used for
mut_malR_for	GAATTTGGGTAAACATATGTGAAGTGCCATTG	Construction of <i>malR</i> mutant
mut_malR_rev	CAATGGCACTTCACATATGTAAACCCAAATTC	
pMAD_malR_for	ATCATATCCATGGTGGCTCAAAGG <sup>a</sup>	
pMAD_malR_rev	AAGAGCGTCGACATAGAAAGAAGCGTAAGC	
verif_malR	GGAACCAGCTCAATTTGATGC	verification of pMAD integration
malR_(comp_pRB473)_for	AAAAAATTACGCCCCGCCCTGCCACTCATCGAACCAGC TCAATTTGATGC	<i>malR</i> complementation
malR_(comp_pRB473)_rev	AGAATCCAAGCTTGCATGCCTGCAGGTCGACTCCAACG CTATCCAACCGG	
pRB473_for	GTCGACCTGCAGGCATGCAAGCTTGG	pRB473 PCR amplification
pRB473_rev	GATGAGTGGCAGGGCGGGGCG	
pRB473_for(bis)	TTAAGGGCACCAATAACTGCC	verification of insertion into pRB473 plasmid
pRB473_rev(bis)	TGTTCAGAACGCTCGGTTG	
ccpArevSal	AAGGGGTCGACACGAGCTAAAGAAAGAGCTTG	Construction of <i>ΔccpA</i> mutant
ccpA_Sac_for	GCCATGAGCTCATTTGCCTCTCTAGCAACATC	
ccpA_Sac_rev	ATATCGAGCTCGTGCAATGCGCTTATTAAC	
ccpA_JH_for_Nco	GTAGGATCCATGGCATTAAAGGTTTCGTC	
1F	TCTAGCTAATGTTACGTTACAC	verification of insertion into pMAD plasmid
2R	TCATAATGGGGAAGGCCATC	
pNZ273_MCS_for	ATGAGATAATGCCGACTGTAC	verification of insertion into pNZ273 plasmid
pNZ273_MCS_rev	CGATCCAGACTGAATGCC	
PmalR_BamHI_for	AAAGGATCCTTTAATCCTACGAACCACGC	cloning of <i>PmalR</i> in pNZ273 plasmid
PmalR_PstI_rev	AAACTGCAGAAGGTTAAATTGATAAACCG	
PmalR_BamHI_rev (ctrl-)	AAAGGATCCAAGGTTAAATTGATAAACCG	cloning of negative control for <i>PmalR</i> in pNZ273
PmalR_PstI_for (ctrl-)	AAACTGCAGTTTAAATCCTACGAACCACGC	
PmalP-T_for (Cy5)	ACGTTTGATTTGTTTCATGAGAATC	amplification of labelled <i>P2mal</i>
PmalP-T_rev	CTCAAAACTAAACATTTTTTTCATTC	
pU_(Cy5)	TGTAAAACGACGGCCAGT	amplification of CY5-labelled <i>PmalP</i> , <i>PmalT</i> and <i>gus</i> fragment
pU_(FAM)	TGTAAAACGACGGCCAGT	amplification of FAM labelled <i>PmalP</i> for footprint assays
pR (M13-rev)	CAGGAAACAGCTATGACC	amplification of the multi cloning sequence of the plasmids pGEM-T
Race_0954-1	TCCTTACTATCTGAATACAC	determination of transcription start site of <i>PmalR</i>
Race_0954-2	TCAAAACTTTTGGCAGTTGC	
Race_0954-3	CGTTCTTTGGAATCTAAAGG	

<sup>a</sup>Underlined nucleotides correspond to restriction sites used to digest PCR products

**Table S4.** Primers used for RT-qPCR in this study

Primer name	Sequence 5'→3'	Gene name
ef_0957L	CAGTAAAGCCTTTTGGTCCG	<i>malP</i>
ef_0957R	AGATGATGAAGCCCAACAAG	
ef_0958L	TGATTGAACAAGCTGGCAAA	<i>malT</i>
ef_0958R	CGCAGTGGTGGTTCCTAGTT	
ef_0006L	GATGGGGAAATCAGGGATTC	<i>gyrA</i>
ef_0006R	TCTTTCCATTCGGCATTTC	

## C. Résultats complémentaires

Afin d'étendre le régulon de MalR ou au contraire de le confirmer comme limité aux deux opérons du maltose, nous avons réalisé des expériences de séquençage du transcriptome entier (RNA-seq) dans différentes conditions. Nous avons dans un premier temps cherché à déterminer les gènes différentiellement exprimés chez la souche sauvage OG1RF lorsque les cellules étaient cultivées en présence de maltose par rapport à une condition contrôle après culture sur glucose. Pour déterminer si cette différence d'expression pouvait être liée au régulateur MalR, le transcriptome du mutant *malR*<sup>-</sup> a également été étudié dans ces deux conditions. Le **Tableau 8** regroupe les gènes présentant une expression différentielle lors d'une croissance sur maltose et/ou dans le mutant *malR*<sup>-</sup>.

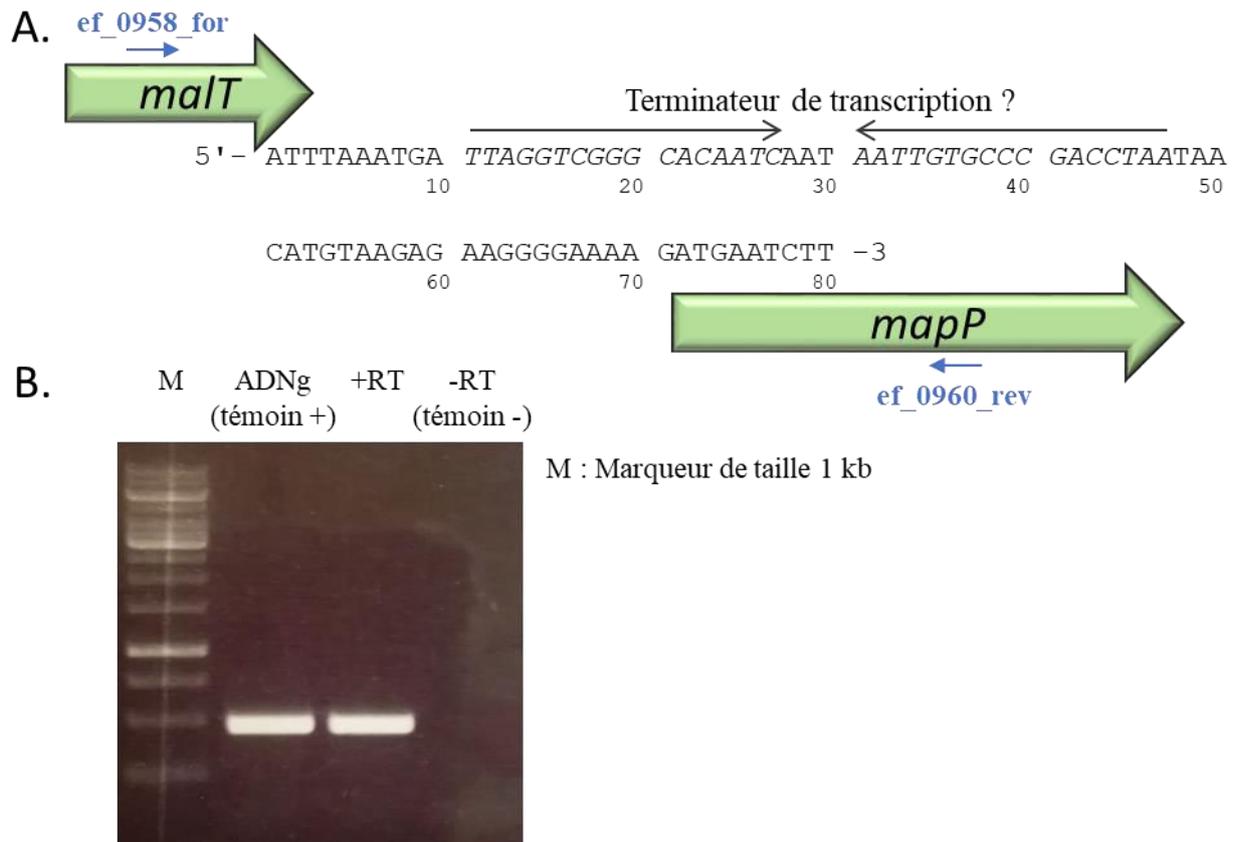
Les opérons débutant par les gènes *malP* et *malT* présentent une induction de leur expression lors d'une croissance sur maltose comme décrit dans la publication présentée ci-dessus. Cette induction peut être expliquée par la levée de répression exercée par MalR. Ce résultat est confirmé par le RNA-seq, ces opérons apparaissant également induits dans la condition *malR*<sup>-</sup>.

Une nette diminution du nombre de cpm correspondants au gène *mapP* comparé au gène *malT* peut être observée dans le tableau précédent. Ces deux gènes sont dans la même orientation et impliqués dans des processus cellulaires étroitement reliés laissant supposer une organisation en opéron. Cependant, un terminateur de transcription peut être prédit *in silico* à la fin du gène *malT* (**Figure 42.A**). Au vu de ces informations, on peut se demander si *malT* et *mapP* sont réellement localisés au sein de la même structure opéronique. Les résultats de RT-PCR présentés en **Figure 42.B** semblent montrer que ces deux gènes sont effectivement présents sur le même transcrit. De plus, d'après les résultats du RNA-seq, l'expression de ces deux gènes apparaît coordonnée. La structure en tige-boucle formée par ce potentiel terminateur de transcription semble tout de même entraîner une atténuation de la transcription du gène *mapP* par rapport au gène *malT*. L'expression de la phosphatase MapP, supposée plus faible en raison du nombre moins important de cpm, semble cependant suffisante pour assurer un bon fonctionnement du métabolisme du maltose. Lors d'une étude antérieure (<sup>108</sup>), les auteurs avaient déterminé par Nother Blot chez la souche JH2-2 que le gène *malT* était présent au sein d'un transcrit de 2,5 kb. Ce résultat est cohérent avec l'expression de *malT* seul mais pas avec la présence de *mapP* au sein du même transcrit. Il est possible que la faible expression du gène *mapP*, probablement due à la présence du terminateur, n'ait pas permis aux auteurs de déceler la présence d'un transcrit plus long contenant les deux gènes.

**Tableau 8 :** Gènes présentant une expression différentielle sur maltose et/ou chez le mutant *malR*- par rapport à une croissance sur glucose

Numéro de gène		Description gène	Count per millions (cpm) normalisés				Log <sub>2</sub> (Fold-Change)	
OG1RF	V583		OG1RF		<i>malR</i> <sup>-</sup>		Mal/Glc (OG1RF)	<i>malR</i> <sup>-</sup> /OG1RF
			Glucose	Maltose	Glucose	Maltose		
10018	0019	PTS EIIB mannose	2706	515	1517	525	-2,39	-0,84
10019	0020	PTS EIIA mannose	6568	1078	3156	943	-2,61	-1,06
10020	0021	PTS EIIC mannose	5111	1192	3167	1248	-2,10	-0,69
10021	0022	PTS EIID mannose	6934	1982	4996	2196	-1,81	-0,47
10619	0891	Aminotransférase	1434	1058	324	378	-0,44	-2,15
10620	0892	Protéine ABC/NBP	399	228	64	65	-0,81	-2,64
10621	0893	Perméase (transporteur ABC)	412	388	121	140	-0,09	-1,76
10623	0895	Oxydoréductase	640	360	148	124	-0,83	-2,11
10665	0938	Protéine ABC/NBP <i>msmX</i>	145	750	457	539	2,37	1,66
10680	0954	Régulateur transcriptionnel <i>malR</i>	586	1306	1879	1730	1,16	1,68
10681	0955	Mutarotase <i>malM</i>	580	1643	2173	2171	1,50	1,90
10682	0956	Phosphoglucomutase <i>malB</i>	775	3052	4684	4376	1,98	2,59
10683	0957	Maltose Phosphorylase <i>malP</i>	2217	6494	12256	10258	1,55	2,47
01684	0958	PTS EIICBA maltose <i>malT</i>	2489	13902	22508	26066	2,48	3,18
10685	0960	Maltose Phosphatase <i>mapP</i>	742	3901	5513	4768	2,39	2,89
11133	1343	Perméase (transporteur ABC) <i>mdxG</i>	90	976	598	810	3,45	2,74
11134	1344	Perméase (transporteur ABC) <i>mdxF</i>	129	1062	689	797	3,04	2,41
11135	1345	SBP (transporteur ABC) <i>mdxE</i>	192	1393	1143	1560	2,86	2,57
11136	1347	Maltodextrin hydrolase <i>mmdH</i>	110	660	618	679	2,58	2,49
11137	1348	Maltodextrin hydrolase <i>gmdH</i>	94	520	340	360	2,46	1,85
11138	1349	Glycosyl transférase <i>mmgT</i>	192	1551	1089	1254	3,01	2,51
11753	2213	PTS EIIBC tréhalose	1375	3466	228	724	1,33	-2,59
12425	3157	Tréhalose phosphorylase <i>trePP</i>	513	2622	528	2643	5,11	1,03
12426	3158	Phosphoglucomutase	161	1141	185	1012	7,10	1,15

Les Log<sub>2</sub>(Fold-Change) supérieurs à 1,5 ou inférieurs à -1,5 dans la condition testée sont mis en évidence respectivement en vert et en rouge..



**Figure 42** : Région intergénique *maltT/mapP*

**A.** Séquence nucléotidique entre les gènes *maltT* et *mapP* chez la souche *OG1RF*. Les séquences complémentaires inversées permettant la formation d'une structure en tige-boucle sont représentées par des flèches fines au-dessus de la séquence.

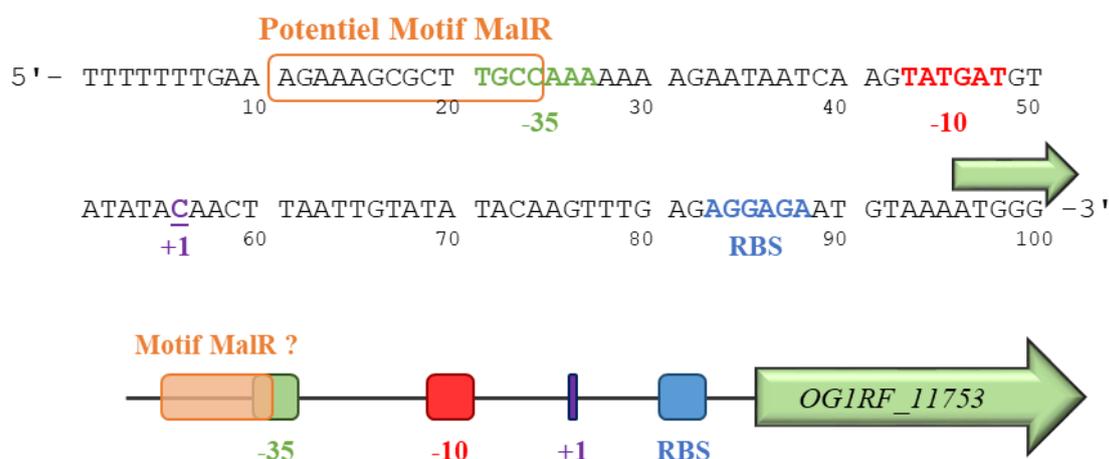
**B.** Résultats de RT-PCR. Les amorces utilisées sont représentées schématiquement en **Figure 42.A**. Les ARN totaux avant l'étape de transcription inverse sont utilisés comme témoin négatifs pour vérifier l'absence d'ADN génomique contaminant.

Un opéron codant pour un PTS complet de la famille mannose (*OG1RF\_10018-21*) apparaît ici réprimé lors d'une croissance sur maltose. Son expression ne semble en revanche pas drastiquement impactée dans un mutant *malR*<sup>-</sup>. Les PTS de cette famille sont régulièrement impliqués dans l'import de glucose (cf partie II.D.1). Cette diminution d'expression lors d'une croissance sur maltose peut donc provenir de l'absence de glucose dans le milieu plutôt et non de la présence de maltose. Cette diminution semble de plus ne pas être liée à MalR. A l'inverse, l'opéron *OG1RF\_12425-26* prédit pour être impliqué dans le catabolisme du tréhalose, apparaît induit par le maltose. Son expression semble également indépendante de MalR. Cette variation peut simplement être due à une levée de la répression catabolique exercée sur cet opéron et non à un effet direct du maltose. Deux potentielles boîtes *cre* légèrement dégénérées ont pu être déterminées, cette induction pourrait donc provenir d'une levée de répression dépendante de CcpA.

D'une manière plus surprenante, on retrouve également des structures génétiques qui apparaissent réprimées dans un mutant *malR*<sup>-</sup> mais dont l'expression semble inchangée lors du

changement de source de carbone comme les opérons *OGIRF\_10621-19* et le gène *OGIRF\_10623*. Ce comportement laisserait penser que MalR pourrait jouer le rôle d'activateur pour ces opérons. Ces gènes sont présents au sein de deux unités transcriptionnelles divergentes. Aucun motif de fixation de MalR n'a pu être prédit *in silico* au sein de la région intergénique de 170 pb entre les codons ATG. De plus, l'ajout de maltose ne semble pas affecter significativement l'expression de ces gènes alors que ce disaccharide module l'affinité de MalR pour l'ADN. Il apparaît donc assez peu probable que MalR joue un rôle direct dans la régulation de ces opérons mais un effet indirect encore non déterminé ne peut pas être exclu.

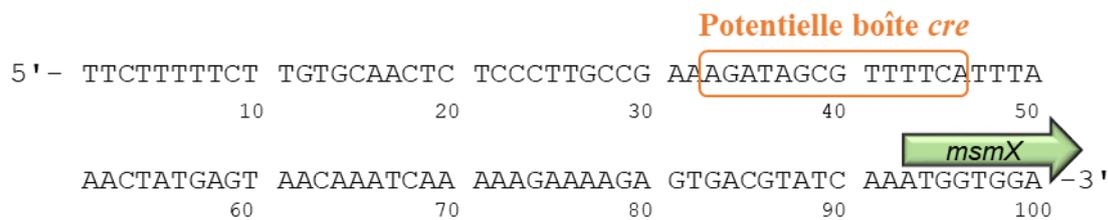
Le gène *OGIRF\_11753* codant pour une protéine EIIBC d'un PTS prédit pour importer le tréhalose apparaît lui aussi réprimé chez un mutant *malR* mais son expression est seulement peu induite en présence de maltose. Un potentiel motif typique des régulateurs transcriptionnels LacI (cf partie III.A.) a pu être prédit dans la région d'ADN située en amont de ce gène. Ce motif peut correspondre à un motif MalR légèrement dégénéré (AGAAAGCGCTTGCC, les nucléotides soulignés correspondent au consensus de MalR). Ne connaissant pas le site de début de transcription de ce gène, il nous est seulement possible d'émettre des hypothèses quant à son promoteur. Après une étude *in silico*, des séquences correspondant aux consensus des séquences classiquement retrouvés au sein des promoteurs procaryotes ont pu être déterminées et sont présentées dans la **Figure 43**. Le potentiel motif MalR est prédit pour être situé directement en amont de l'élément -35 avec un léger chevauchement. Ce positionnement peut être compatible avec une action comme activateur direct de MalR, aucun autre motif potentiel n'a pu être déterminé. En raison de la légère variation du site présent dans cette séquence avec son consensus, MalR devrait présenter une affinité assez faible pour ce site. Deux scénarii sont envisageables, MalR pourrait agir comme un activateur direct sur l'expression de ce gène en l'absence de maltose mais on peut également imaginer une implication plus indirecte de cette protéine dans cette régulation. En effet MalR pourrait être capable de fixer faiblement ce site et ainsi entrer en compétition avec un autre régulateur LacI qui agirait lui directement comme un activateur pour ce gène. L'effet activateur de MalR observé lors du RNA-seq résulterait dans ce cas de l'absence de compétition exercée par MalR chez le mutant. Cependant ces résultats de RNA-seq n'ont pas été validés par RT-qPCR et toutes ces hypothèses restent purement spéculatives.



**Figure 43 :** Promoteur et motif MalR prédits en amont du gène *OG1RF\_11753*

Les différents éléments prédits du promoteur ainsi qu'un potentiel motif MalR sont mis en évidence par un code couleur. Tous ces éléments sont uniquement prédictifs et n'ont fait l'objet d'aucune vérification. Une représentation graphique est proposée en dessous de la séquence.

Trois derniers opérons (*OG1RF\_10665*, *OG1RF\_11135-33* et *OG1RF\_11136-38*) sont présents dans ce tableau et apparaissent induits à la fois par la présence de maltose et par l'absence de MalR. Ces opérons ont déjà été présentés dans la partie IV.C de l'introduction, ils interviennent dans le métabolisme des maltodextrines chez *E. faecalis*. Le gène *msmX* (*OG1RF\_10665*) code pour une protéine NBP fournissant l'énergie au transporteur MdxEFG. Malgré la visible implication de MalR dans l'expression de ce gène d'après le RNA-seq, aucun motif MalR n'a pu être prédit ni en amont ni dans le début de ce gène. L'implication de ce régulateur transcriptionnel n'est par conséquent pas certaine. Cette protéine MsmX est homologue à la protéine du même nom décrite chez *B. subtilis* (<sup>266</sup>). Elle permet chez cette bactérie de fournir de l'énergie à plusieurs transporteurs ABC. Si la protéine MsmX d'*E. faecalis* assure elle aussi une fonction aussi centrale, il n'est pas surprenant que son expression ne soit pas régulée par un régulateur répondant à un stimulus aussi spécifique que la présence de maltose. Une implication indirecte de MalR reste cependant envisageable. Le promoteur de ce gène *msmX* n'a pas pu être prédit avec certitude *in silico*. Il est cependant possible d'identifier en amont de ce gène, une potentielle boîte *cre*, ce gène pourrait alors être sous le contrôle du régulateur pléiotrope CcpA (**Figure 44**). De plus, le gène *msmX* homologue chez *B. subtilis* est effectivement sous contrôle direct de CcpA (<sup>267</sup>).



**Figure 44** : Séquence en amont du gène *msmX* (*OG1RF\_10665*)  
 La potentielle boîte cre identifiée a été mise en évidence par un cadre orange.

La région intergénique entre les opérons *mdxEFG* et *mmdH-gmdH-mmgT* présente, elle, deux potentiels motifs de fixation de MalR. La régulation de ces structures opéroniques fera l'objet d'une étude plus approfondie dans la deuxième publication présentée dans ce manuscrit.

### III. Publication scientifique n°2

#### A. Résumé de la publication n°2

Comme annoncé précédemment, les opérons débutant par les gènes *mdxE* et *mmdH* sont essentiels au métabolisme des maltodextrines d'une longueur supérieure à 3 résidus glycosyl chez *E. faecalis*. Ces opérons sont induits par le maltose et par l'absence du régulateur MalR. Ce régulateur transcriptionnel du métabolisme du maltose est en effet également directement impliqué dans la régulation des opérons maltodextrines. MalR agit en effet sur l'expression des opérons maltodextrines de manière similaire à ce qui a été rapporté dans la publication précédente pour les opérons maltose. La fixation *in vitro* de MalR sur le promoteur de l'opéron *mdxEFG* a pu être clairement établie. En revanche, ce régulateur semble présenter une affinité très faible pour le promoteur de l'opéron débutant par *mmdH*.

MalR seul n'explique pas intégralement le profil d'expression de ces opérons maltodextrines. Ces gènes présentent en effet un second niveau d'induction lors du métabolisme de maltodextrines à longues chaînes comme le maltopentaose indépendamment de MalR. Le régulateur transcriptionnel pléiotrope CcpA est responsable de cette seconde induction. La protéine P-Ser-HPr semble assurer une régulation fine et une coordination de ces deux régulateurs transcriptionnels en fonction de la disponibilité en source de carbone. L'importeur ABC MdxEFG est lui aussi au centre de la régulation de ces opérons via un phénomène d'exclusion de l'inducteur probablement par une inhibition de l'activité de l'ATPase MsmX par la protéine P-Ser-HPr.

#### B. Publication n°2 : Regulation of maltodextrin genes in *Enterococcus faecalis* by the combined action of the main regulator of maltose genes MalR and the pleiotropic regulator CcpA

1 **Regulation of maltodextrin genes in *Enterococcus faecalis* by the**  
2 **combined action of the maltose genes regulator MalR and the**  
3 **pleiotropic regulator CcpA**

4

5 Author names: Maxime Grand<sup>a</sup>, Josef Deutscher<sup>b,c</sup>, Axel Hartke<sup>a</sup>, Nicolas Sauvageot<sup>a#</sup>

6 Affiliations:

7 <sup>a</sup>*Normandie Université, UNICAEN, U2RM Stress/Virulence, F-14000 Caen, France*

8 <sup>b</sup>*Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France*

9 <sup>c</sup>*UMR 8261, CNRS, Université de Paris, Institut de Biologie Physico-Chimique, 75005 Paris,*  
10 *France*

11

12 Corresponding Author:

13 #For correspondence. E-mail: nicolas.sauvageot@unicaen.fr

14

15 Tel. (+33) 2 31 56 56 61;

16 Fax: (+33) 2 31 56 53 11.

17

18 Keywords: metabolism, maltodextrin, *Enterococcus*, regulation

19 Running title: Regulation of maltodextrin metabolism in *Enterococcus faecalis*

20

## 21 ABSTRACT

22 **Enterococci are Gram-positive bacteria members of healthy human microbiota but also**  
23 **represent a leading cause of nosocomial infections in the world. Maltodextrin metabolism of**  
24 ***Enterococcus faecalis* was already described as an important factor for mammals host**  
25 **colonization. Here we show that the LacI/GalR transcriptional regulator MalR, the maltose**  
26 **genes regulator, is also the main regulator of the operons responsible for the metabolism of**  
27 **maltodextrins longer than maltotriose. These two metabolisms consequently are coordinated**  
28 **and induced by the presence of the disaccharide maltose. The two operons involved in the**  
29 **metabolism of long maltodextrins are also regulated by the pleiotropic regulator CcpA, this**  
30 **global regulator exerts a moderate repression and allows a modulation of expression**  
31 **complementary of the MalR mediated regulation, especially in presence of maltose. The**  
32 **activity of the maltodextrin ABC transporter is also rapidly abolished by inducer exclusion.**  
33 **All these strategies of regulation allow *E. faecalis* to rapidly adapt its metabolism to the**  
34 **environmental variations.**

35

## 36 IMPORTANCE

37 ***Enterococcus faecalis* represents a leading cause of hospital acquired infection worldwide.**  
38 **Several studies highlighted the importance of carbohydrates metabolism in the infection**  
39 **process of this bacterium and maltodextrin metabolism was reported as particularly induced**  
40 **in mice infection context. No data were available concerning the regulation of this metabolism**  
41 **important for the infection process of this opportunistic pathogen.**

## 42 Introduction

43 Enterococci are Gram-positive commensal bacteria predominantly encountered in the digestive  
44 tract of animals and humans (1). These bacteria usually represent members of the commensal  
45 human microbiota but two species, *Enterococcus faecalis* and *Enterococcus faecium* are also  
46 leading causes of health care associated infections worldwide (2–4). Indeed, these opportunistic  
47 pathogens can be responsible for urinary tract infections, endocarditis or bacteremia in  
48 immunocompromised subjects (3). Several characteristics of these microorganisms can explain  
49 their persistence in hospitals like their tolerance to desiccation (5, 6) or their high antibiotic

50 resistances (7). In addition, enterococci can possess several intrinsic virulence factors often linked  
51 to their host pathogenicity like secreted proteases or toxins (8). The infection process of  
52 enterococci remains however poorly characterised. The nutrients acquisition and metabolism in  
53 the host is one of the very first steps in the colonisation process, some studies have already shown  
54 in different bacterial models that glycans mobilisation represents an important factor for host  
55 colonisation (9–14). A global transcriptomic study was performed in *E. faecalis* to identify the  
56 genes differently expressed during a mice intraperitoneal infection (15). Among the 211 genes  
57 determined as upregulated in this infection model, about 35 % (74/211) were directly linked to the  
58 metabolism or nutrients mobilisation. Moreover, a global transcriptional regulator MafR, shown to  
59 be essential for the expression of many genes involved in carbon metabolism in *E. faecalis* (16,  
60 17), was already reported to be important for pathogenicity. Indeed, a *mafR* mutant was shown to  
61 cause a lower degree of inflammation in mice than the isogenic wild type strain (16). All these data  
62 support the importance of glycan metabolism in the infection process of *E. faecalis*.

63 An operon formed by the genes *mdxEFG* (*efl345-1343* in the V583 strain) appears of particular  
64 interest. Indeed, this operon, encoding the only known efficient maltodextrin uptake  
65 ABC transporter in *E. faecalis* (18), was reported as upregulated in the transcriptomic study  
66 described above (15). A null mutant for MdxF, an essential trans-membrane protein of the previous  
67 ABC transporter, presents a defect in organ colonization in mouse (18). Some studies also reported  
68 that maltodextrins binding proteins, component of ABC transporter, were important for  
69 colonization of *Streptococcus pneumoniae* (19) and *Streptococcus pyogenes* (20), two  
70 gram-positive human pathogens. Maltodextrins are carbohydrates composed of glycosyl residues  
71 linked with  $\alpha$ -1,4-O-bonds. These glycans are generally obtained after incomplete hydrolysis of  
72 storage polymers like starch and glycogen. To metabolise maltodextrins of three (maltotriose) to at  
73 least seven residues (maltoheptaose), *E. faecalis* uses two tricistronic divergent operons shown in  
74 Fig. 1A. The uptake of maltodextrins from maltotetraose to at least maltoheptaose is mostly  
75 achieved by the MdxEFG ABC transporter cited above (18). After being imported, these

76 carbohydrates are cleaved by the "maltose producing maltodextrin hydrolase" MmdH (21) which  
77 hydrolyses the  $\alpha$ -1,4-bond between the second and third residues from the non-reducing end of  
78 maltodextrins. The complete hydrolysis of even numbered maltodextrins by MmdH consequently  
79 results in production of maltose molecules and odd numbered maltodextrins in production of  
80 maltose and one glucose. The genes present in the two operons beginning with the *mdxE* and *mmdH*  
81 are from now on named "maltodextrin genes" throughout this study. The produced maltose is then  
82 phosphorolysed in glucose-1-phosphate (G1P) and glucose by the MalP protein encoded by the  
83 first gene of the *malPBMR* operon located elsewhere on the genome (22, 23). The metabolisms of  
84 maltose and maltodextrins in *E. faecalis* therefore appear closely related. In a previous study ((24),  
85 unpublished data), we showed that maltose genes were regulated in this bacterium by MalR, a  
86 LacI/GalR family transcriptional regulator encoded by the last gene of the *malPBMR* operon. These  
87 maltose genes were shown to be submitted to Carbon Catabolite Repression (CCR) when cells  
88 were grown in presence of the preferential sugar glucose ((24), unpublished data). This CCR was  
89 independent of the Catabolite control protein A (CcpA), the main regulator of CCR in *firmicutes*  
90 (25) but dependent on MalR in *E. faecalis*. To efficiently bind DNA on specific recognition  
91 sequences named *cre*-boxes located upstream of target genes (26), CcpA requires its co-repressor  
92 P-Ser-HPr. P-Ser-HPr is synthesized by the ATP-dependent phosphorylation of the general  
93 PhosphoTransferase System (PTS) protein HPr on Ser-46 by the HPr kinase/phosphorylase  
94 (HPrK/P) (27). The PTS catalyse the concomitant uptake and phosphorylation of numerous  
95 carbohydrates through a phosphorylation cascade from the phosphoenolpyruvate (PEP) to the  
96 imported glycan. This cascade is mediated by the two general proteins EI and HPr and  
97 sugar-specific proteins or domains EIIA and EIIB. P~EIIB transfers its phosphoryl group to the  
98 carbohydrate bound to the corresponding membrane spanning EIIC (and EIID for the mannose  
99 family). The resulting P-sugar is finally released into the cytoplasm (28). In this phosphorylation  
100 cascade, the HPr protein is phosphorylated on His-15 (P~His-HPr) but this protein also carries an  
101 additional phosphorylatable seryl residue (Ser-46) as mentioned above. The P-Ser-HPr is not a

102 functional phosphoryl donor for the uptake cascade but serves as a central signalling molecule in  
103 *firmicutes* and carries several regulatory functions. Besides acting as the essential co-repressor of  
104 CcpA, P-Ser-HPr is also implicated in inducer exclusions in low %GC Gram-positive bacteria (25,  
105 29). The kinase activity of HPrK/P is stimulated by the presence of the two glycolysis precursors  
106 fructose-1,6-bisphosphate (FBP) and glucose-6-phosphate (G6P) (30–32), the concentration of  
107 such metabolites is increased when a rapidly metabolizable carbon source like glucose is used. The  
108 main regulator of the maltose genes, MalR, described as a partner of P-Ser-HPr, was also essential  
109 for the CCR of these genes. This transcriptional regulator was shown to specifically bind motifs  
110 with the following consensus CGMAACCGATTGCR. Moderately degenerated motifs are present  
111 upstream of maltodextrin genes and an implication of the MalR regulator was supposed. In the  
112 phylogenetically related species *E. faecium* E1162, the genes involved in the maltodextrins  
113 metabolism are positively regulated by a LacI/GalR family transcriptional regulator MdxR (33).  
114 When searching for MdxR homologies in *E. faecalis* genomes, no homologous protein was found  
115 suggesting different regulation pathways in the two species. The regulator of the maltose genes,  
116 MalR, was however in *E. faecalis* the best hit against MdxR.

117 In the present study, we determined using RT-qPCR that the genes involved in maltodextrins  
118 metabolism were slightly induced in presence of maltose and highly induced in presence of a long  
119 maltodextrin like maltopentaose. The regulator of the maltose genes, MalR, and the CCR global  
120 regulator, CcpA, appear to jointly regulate these maltodextrin operons. Indeed, these genes are  
121 repressed in presence of glucose in a MalR dependent manner while the CcpA protein only  
122 conditionally appears essential for this CCR. Electrophoretic Mobility Shift Assays (EMSA)  
123 confirmed the direct interaction of both MalR and CcpA in the DNA region upstream of the  
124 maltodextrin genes. In addition to this direct transcriptional regulation, the maltodextrin genes are  
125 regulated in *E. faecalis* by an inducer exclusion phenomenon which rapidly abolish the activity of  
126 the ABC importer in presence of glucose and therefore prevent the formation of the inducer maltose  
127 in the cell through the MmdH activity.

## 128 Results

129 **Maltodextrin genes are repressed by both the regulator of maltose genes MalR and the global**  
130 **regulator CcpA.** An efficient maltodextrins metabolism in *E. faecalis* requires the *malPBMR*  
131 operon involved in maltose metabolism (22) and two additional divergent operons located  
132 elsewhere in the genome (Fig. 1A). The LacI/GalR family transcriptional regulator MalR was  
133 previously described as the regulator of the maltose genes and its implication in regulation of the  
134 maltodextrin genes was supposed ((24), unpublished data). To verify this hypothesis, the OG1RF  
135 wild type and its isogenic *malR*<sup>-</sup> mutant were cultivated on different carbon sources and expression  
136 of the first genes of the maltodextrin operons (*mdxE* and *mmdH*, respectively) was studied by  
137 RT-qPCR. The overall expression profiles of *mdxE* and *mmdH* operons were similar in every  
138 conditions tested but the last operon appeared less expressed compared to the *mdxEFG* operon.  
139 These genes were strongly repressed in the OG1RF parental strain in the presence of glucose  
140 (Fig. 2, dark grey bars). In the presence of maltose as the sole carbon source, the expression level  
141 of these genes was increased compared to glucose, but an even higher expression of both operons  
142 was obtained with maltopentaose in the growth medium. However, expression levels of both  
143 operons were significantly repressed when glucose and maltose or glucose and maltopentaose was  
144 simultaneously present during growth. These two operons therefore appears to be subjected to  
145 CCR.

146 In a *malR*<sup>-</sup> mutant (grey bars), both operons also presented similar expression profiles under the  
147 different conditions tested. On maltose or maltopentaose, their expression levels were comparable  
148 to those observed in the parental strain although maltopentaose triggered higher expression of both  
149 operons than maltose. In the presence of glucose sole or in combination with maltose or  
150 maltopentaose, comparable expression as on maltose-growing wild type cells are observed. We  
151 concluded that the MalR protein seems to be indeed involved in the regulation of the maltodextrin  
152 operons and that it is essential for maximal repression of *mdxE* and *mmdH* operons.

153 Given that expression of the two operons in the *malR*<sup>-</sup> mutant is only partially de-repressed  
154 suggesting that *mdxE* and *mmdH* operons should be subjected to at least one other type of  
155 regulation. In order to determine if the CCR observed was dependent or independent of the major  
156 global regulator CcpA, expression profiles of the operons in a  $\Delta$ *ccpA* mutant were studied (Fig. 2.  
157 light grey bars). This mutant still presented the same fully induced state as the wild type when cells  
158 were grown on maltopentaose as the sole carbon source. When grown on maltose only, the deletion  
159 of the *ccpA* gene led to a fully induced expression of the two operons similar to the level obtained  
160 after a growth on maltopentaose. The CcpA protein was therefore responsible of the partial  
161 repression observed in the wild type when the maltose was used as the sole carbon source and  
162 modulate the expression of the maltodextrin genes when maltose is present. When the maltose was  
163 absent and the glucose was present, the  $\Delta$ *ccpA* mutant showed a similar expression profile as the  
164 wild type. However, in the condition where both carbohydrates were present, the  $\Delta$ *ccpA* mutant  
165 showed a partially induced state comparable with the *malR*<sup>-</sup> mutant. In the light of these data, the  
166 maltose repressor MalR seems to be the major regulator of the maltodextrin genes but the  
167 pleiotropic regulator CcpA also acts as a repressor on these genes especially when MalR is absent  
168 or inactive (in the *malR*<sup>-</sup> mutant or in the presence of maltose, respectively). The fully induced  
169 expression level of the maltodextrin operons should consequently require the inactivation of the  
170 two previous actors. To elucidate this point, the double mutant MC (for *malR*<sup>-</sup> $\Delta$ *ccpA*) carrying both  
171 mutations described above was studied. This mutant presented a fully induced expression profile  
172 for *mdxE* and *mmdH* genes under all tested conditions (Fig. 2, white bars). These data confirm that  
173 both CcpA and MalR are involved in the regulation of the maltodextrin operons.

174 The complemented strains carrying a pRB473 derivative plasmid with a wild type copy of the  
175 corresponding gene presented expression levels near to the wild type in almost every conditions  
176 (Fig. S1).

177

178 **Both MalR and CcpA regulate the maltodextrin genes by direct interaction with the promoter**  
179 **region.** Mutations in *malR* and/or *ccpA* gene affected expression profiles of the maltodextrin genes.  
180 Direct interactions of these two regulators with the intergenic region between the two maltodextrin  
181 operons were therefore suggested. Upstream of *mdxE* and *mmdH* operon, a similar imperfect  
182 palindromic motif (TGCAACCGATTGCG and TGCAATCGTTTGCG, respectively) was  
183 identified. Both are comparable to the binding site of MalR determined previously ((24),  
184 unpublished data) and the recognition sequence of CcpA *cre*-boxes of *E. faecalis* (34). EMSA were  
185 performed in order to elucidate if the MalR and CcpA proteins are able to bind these promoter  
186 regions. Different fluorescently labeled regions containing either the *mdxE* promoter ( $P_E$ ), the  
187 *mmdH* promoter ( $P_H$ ) or both ( $P_{EH}$ ) were used as DNA templates. EMSA results obtained with the  
188 MalR protein with the  $P_{EH}$  region are shown in Fig. 3A. Direct specific interactions for the MalR  
189 protein with the full length intergenic region was observed. Indeed two distinct shifted bands were  
190 visible when this region was used as a template, suggesting that MalR binds to the two binding  
191 sequences present in the  $P_{EH}$ . When  $P_E$  or  $P_H$  were used (Fig. 3B lanes 1 to 5 and 6 to 10  
192 respectively), a single shifted band was visible for each promoter. However, the shifted bands are  
193 clearly stronger with  $P_E$  compared to  $P_H$  strongly suggesting a better affinity of MalR for the former  
194 promoter. Concerning the  $P_H$  region, a smear was observed rather than a clear shifted band  
195 suggesting that the MalR/DNA complex was unstable and split off during electrophoresis. We  
196 previously showed that MalR was able to form a complex with P-Ser-HPr supposed to improve its  
197 affinity for its target ((24), unpublished data). In the present study, this enhancement of affinity  
198 was confirmed and especially visible by the significant decrease in the amount of free DNA (for  
199 example, compare lanes 6 and 2 or 7 and 3 in Fig. 3A, lanes 4 and 2 or 9 and 7 in Fig. 3B). No  
200 shifted complex were visible using the internal fragment of the *gusA* gene as DNA template  
201 indicating that MalR did not present non-specific DNA binding under these experimental  
202 conditions (Fig. S2A). Unphosphorylated HPr and P-His-HPr did not affect the DNA binding  
203 affinity of MalR and different forms of HPr alone did not bind DNA (Fig. S2B).

204 For the experiments with CcpA, its essential cofactor P-Ser-HPr was systematically added at 2  $\mu$ M  
205 for optimal DNA binding conditions (Fig. 4) (26). The CcpA/P-Ser-HPr complex presents a low  
206 selectivity for DNA and was able to bind the negative control DNA at concentrations  $\geq 6.25$  nM  
207 despite absence of *cre*-boxes like sequences (Fig. 4A lanes 7 and 8). Results obtained with such  
208 concentrations were therefore interpret as non-specific binding and our attention was only focused  
209 on results obtained with CcpA used at concentration of 3.13 nM or below. Results shown in Fig. 4A  
210 suggest that the CcpA/P-Ser-HPr complex was able to specifically bind the  $P_{EH}$  region at least once  
211 (lanes 4 and 5). Tested on individual promoters, the CcpA/P-Ser-HPr complex seemed able to bind  
212 the  $P_E$  region correctly (Fig. 4B lanes 3 and 4) but only poorly the  $P_H$  region (Fig. 4B lanes 7 and  
213 8). In light of these results, both MalR and CcpA proteins, seems to acts on regulation of  
214 maltodextrin genes by direct interactions with their promoter regions. Both regulators presents a  
215 better affinity for  $P_E$  upstream of the *mdxE* gene compared to the *mmdH* promoter.

216 Following the direct interaction evidence between MalR and DNA regions carrying maltodextrin  
217 genes promoters, the experimental confirmation of MalR binding sites by footprinting assays were  
218 undertaken. The predicted MalR binding sites and their relative positions in the  $P_E$  and  $P_H$  regions  
219 are presented in Fig. 1B. The MalR binding motif present upstream the *mdxE* gene was  
220 experimentally confirmed and situated in the protected region in the footprint (Fig. 1C). In contrast,  
221 concerning the  $P_H$  region, we failed to confirm the MalR binding motif and no protected region  
222 was visible through the whole sequence (Fig. S3). This MalR protection defect could be explained  
223 by the poor apparent *in vitro* affinity of the regulator for this latter region which is in accordance  
224 with the EMSA results (Fig. 3B).

225 ***E. faecalis* produces less P-Ser-HPr when grown on maltopentaose compared to glucose or**  
226 **maltose.** We showed that CcpA directly repressed the maltodextrin operons and was responsible  
227 for the difference between the fully induced state observed on maltopentaose and the partially  
228 induced expression level observed on maltose. As the ability of CcpA to efficiently bind DNA is  
229 directly correlated with the presence of its cofactor P-Ser-HPr, we wondered if a difference in P-

230 Ser-HPr concentration would be at the basis of the differential expression levels on maltose and  
231 maltopentaose. The relative abundance of the different forms of HPr was estimated by  
232 Western-Blotting assays as previously described (35). Briefly, after the cells were grown, each  
233 crude extract was divided into two fractions. One of them was directly loaded onto a Native  
234 Polyacrylamide Gel and the second was boiled for 5 minutes before loading to eliminate the  
235 temperature-labile phosphorylation on H15 (36). The remaining band in these samples correspond  
236 to the temperature-stable P-Ser-HPr. A double phosphorylated form of HPr was generally detected  
237 in cells grown on maltose but only rarely detected in bacteria grown on maltopentaose. After heat  
238 treatment, an intense band corresponding to P-Ser-HPr was detected for cultures grown on maltose  
239 whereas only a faint band of this form was observed in cells grown on maltopentaose (Fig. 5). The  
240 combined results suggest that the content of the CcpA cofactor P-Ser-HPr is less abundant in  
241 cultures grown on maltopentaose than in cells grown on maltose.

242 **Addition of glucose in growth medium rapidly prevents the import of maltopentaose**  
243 **independently of transcriptional regulation.** As described above, both maltodextrin operons are  
244 submitted to CCR by glucose in the wild type strain. These results showed also that the repression  
245 observed on glucose and maltose was dependent on CcpA whereas it appeared CcpA-independent  
246 on glucose and maltopentaose (Fig. 2). An inhibition of the MdxEFG ABC transporter by an  
247 intracellular actor in response to the presence of glucose could explain this CcpA-independent  
248 regulation. Indeed, this phenomenon known as inducer exclusion would prevent maltodextrins  
249 longer than four glucosyl residues from being imported in the presence of glucose. Consequently,  
250 no MmdH mediated hydrolysis of maltodextrin could occur preventing the synthesis of the inducer  
251 maltose. The MalR repressor should therefore be active and efficiently repress the expression of  
252 both maltodextrin operons. To verify this hypothesis, the concentration of maltopentaose in the  
253 growth medium was monitored over time by HPLC. The wild type OG1RF strain was grown in  
254 ccM17 supplemented with 0.2% of maltopentaose for 3h at 37°C and glucose was then added to a  
255 final concentration of 0.15%. A control was systematically performed without addition of glucose.

256 When cells were grown in presence of maltopentaose only, the maltodextrin was continuously  
257 consumed over time (Fig. 6A.I). When glucose was added after three hours of growth on  
258 maltopentaose, consumption of the maltodextrin decreased immediately whereas the added glucose  
259 was rapidly consumed (Fig. 6A.II). Two mechanisms could be at the basis of this fast glucose effect  
260 on maltopentaose consumption. It could be due either to transcriptional repression by MalR and/or  
261 CcpA/P-Ser-HPr (see Fig. 2) or by an inducer exclusion mechanism. However, due to the rapid  
262 kinetic, we favoured the last possibility. If right, the double mutant MC which constitutively  
263 express the maltodextrin transporter MdxEFG should also demonstrate a decrease in the  
264 maltopentaose consumption after addition of glucose. As shown in Fig. 6B.I, this behaviour was  
265 indeed observed immediately after the addition of glucose. After 270 min, a rapid decrease of  
266 maltopentaose was noticed in the MC mutant which coincided with exhaustion of glucose in the  
267 medium (Fig. 6B.II). This phenomenon was not observed in the OG1RF strain although glucose  
268 was consumed at a similar rate. The wild type strain probably required time to induce the *mdxEFG*  
269 operon while the MC mutant did not cause of its constitutively induced state. Altogether, these data  
270 strongly support that the ABC transporter involved in maltodextrin uptake is submitted to CCR  
271 exerted by glucose through an inducer exclusion phenomenon.

272 Similar experiments were conducted monitoring the maltose concentration instead of  
273 maltopentaose (Fig. S4). Maltose is imported into the cell mainly by the PTS MalT. In the wild  
274 type strain OG1RF, the maltose concentration decreased over time during the growth. After  
275 addition of glucose, the maltose concentration still decreased even if we noticed a moderate  
276 reduction of the maltose consumption compared to condition without glucose. Similar to the  
277 *mdxEFG* operon, the *malT* gene is repressed by glucose ((24), unpublished data) which could be  
278 responsible for the observed decrease in maltose consumption. We therefore performed similar  
279 measures in a *malR* mutant which is relieved from CCR for maltose genes. In this background, the  
280 decrease of maltose uptake was almost abolished meaning that this previously noticed reduction in  
281 the OG1RF strain was mainly due to transcriptional repression. The marginal decrease in

282 consumption of maltose induced by the addition of glucose in the *malR*<sup>-</sup> mutant might be explained  
283 by the competition for P~His-HPr between the EIIA domain of MalT and several glucose PTS  
284 EIIA. This putative competition appears to impact only slightly the MalT function.

285 Taken together, these data strongly suggest that the repressed expression of the *mdxEFG*, observed  
286 when cells were grown on glucose and maltopentaose, not only involves transcriptional repression  
287 but also an inhibition of the maltopentaose uptake via inducer exclusion. The addition of glucose  
288 therefore seemed to only affect the maltose uptake by the PTS permease MalT by transcriptional  
289 repression. A *malR*<sup>-</sup> mutant, in which the *malT* gene is relieved from CCR, actually did not show a  
290 significantly affected uptake rate after addition of glucose.

## 291 Discussion

292 Adaptation of bacteria to environmental conditions is crucial to provide them an efficient fitness  
293 essential for survival under harsh competition for nutrients. The regulation of carbohydrate  
294 utilization therefore represents a field of major interest to understand the bacterial lifestyle. In the  
295 present communication, we showed that the maltodextrin genes in *E. faecalis* were directly  
296 regulated by the joint action of both MalR and CcpA. In *S. pneumoniae*, the MalXCD  
297 ABC transporter involved in maltodextrins uptake was also described as repressed by both the  
298 maltose genes regulator MalR and the global regulator CcpA (37–39). MalR of *E. faecalis* was  
299 previously shown to present affinity for maltose but not for glucose or other maltodextrins ((24),  
300 unpublished data). This disaccharide was therefore the only described inducer of the maltose genes.  
301 Given that MalR also regulates the maltodextrin genes, the maltose appears to be the central inducer  
302 of the four operons. The maltodextrin hydrolase MmdH activity results in formation of maltose  
303 (21), in consequences, this glycan effectively represents a central molecule of both maltose and  
304 maltodextrin systems in *E. faecalis*. In light of the essential part of the MmdH activity in the  
305 induction of both maltose and maltodextrin system, the lower affinity of the two repressors MalR  
306 and CcpA could be of physiological interest. A too strong repression of the *mmdH* gene could slow

307 down the formation of the inducer maltose and impact the cellular fitness in presence of  
308 maltodextrins. On the contrary, in *S. pneumoniae*, MalR seems to have a better affinity for the  
309 operator of maltodextrin phosphorylase than for the operator of the maltodextrins uptake system  
310 (37, 38). This difference might allow a basal expression of the maltodextrin uptake system in order  
311 to not definitely lock the system.

312 Inducer exclusions have already been reported to regulate the maltose/maltodextrin metabolism in  
313 other bacteria. In the Gram-negative enterobacteria *E. coli* and *Salmonella typhimurium* the  
314 ATP-binding subunit MalK of the maltose/maltodextrin ABC transporter MalE-FGK<sub>2</sub> was reported  
315 to be the target of the inducer exclusion exerted by the glucose. The unphosphorylated EIIA<sup>Glc</sup>  
316 protein actually interacts with MalK and inhibit the ATP dependent closure of MalK dimer causing  
317 the inhibition of the ABC transporter activity (40, 41). The EIIA<sup>Glc</sup> is encountered in the cell when  
318 the corresponding glucose PTS is used to import glucose. This strategy allows these bacteria to  
319 favour the easily metabolisable glucose over more complex polymers. The inhibition mechanism  
320 of the maltodextrin ABC transporter MalE1-F1G1K1<sub>2</sub> of *Lactobacillus casei* has been resolved. In  
321 this *firmicute*, P-Ser-HPr directly prevents the MalK1 mediated ATP hydrolysis by direct  
322 interaction (29). This inducer exclusion strategy consisting of the inhibition of the ATPase subunit  
323 of ABC transporters by direct interaction with a central signalling molecule therefore appears  
324 relatively conserved throughout bacteria. In *E. faecalis*, the nucleotide binding protein encoded by  
325 the gene *OG1RF\_10665*, homologous to the MsmX ATPase of *Bacillus subtilis*, has been  
326 suggested as the ATPase of the MdxEFG ABC transporter (18, 42). The ATP hydrolysis activity  
327 of the *E. faecalis* MsmX homologue might also be inhibited by P-Ser-HPr in a similar manner than  
328 in *L. casei*.

329 In light of the RT-qPCR assays, MalR seems to be the major regulator of the maltodextrin system  
330 and its deletion led to the relief of CCR of both *mdxE* and *mmdH* operons in every tested conditions.  
331 On the other hand, MalR presented *in vitro* only a moderate affinity for the motif in the *PmdxE*  
332 region and a low affinity for the *PmmdH* region. The MalR-mediated repression exerted on these

333 genes seems to be effective in absence of inducer but should be rapidly relieved in presence of  
334 maltose. The inhibition of the MdxEFG ABC transporter by inducer exclusion prevents the  
335 maltodextrin from being imported in the presence of glucose. The inducer maltose consequently  
336 will not be formed and the weakly bound MalR might be sufficient to exert an efficient repression.  
337 Moreover, the pleiotropic regulator CcpA was also shown to directly repress the two maltodextrin  
338 operons. Its repressive action is thought to be important in presence of maltose when the DNA  
339 affinity of MalR decrease. Indeed, in the condition where both glucose and maltose are present, the  
340 maltose could enter the cell mainly by MalT (22) but another study showed that maltose was also  
341 able to be taken up by a still unidentified transporter (18). Some maltose should therefore be present  
342 in the cell and the low affinity of MalR for the operators might no longer be sufficient to efficiently  
343 repress the maltodextrin genes expression. The CcpA protein seems to be, in that case, essential to  
344 keep repressed expression of these genes.

345 In *E. faecium* E1162, maltodextrins are metabolised by a single operon formed by the operon  
346 *pula-mdxDCBA* encoding a hydrolase and an ABC transporter (33). This operon is activated by  
347 the MdxR protein, a LacI/GalR transcriptional regulator encoded by a gene located elsewhere in  
348 the genome. In the E1162 wild type strain, the expression of the *mdxB* gene was reported as highly  
349 induced in presence of maltose although in a *mdxR* mutant, this operon was not expressed (33).  
350 Moreover, a *mdxR* mutant was unable to grow on maltotetraose as sole carbon source while the  
351 wild type grew correctly. The authors therefore concluded that MdxR was the main regulator of  
352 the maltodextrin operon in *E. faecium* E1162. Interestingly, no homologue of MdxR could be  
353 identified in *E. faecalis*. The synteny of the maltose genes (*malPBMR* and *malT-mapP* operons) is  
354 conserved in both species. Moreover the MalR motifs determined in *E. faecalis* OG1RF are 100%  
355 conserved in *E. faecium* while the intergenic region between the start codon of *malP* and *malT*  
356 presents only 62.1% of identity between the two species. In view of these data, we can speculate  
357 that MalR should also be the main regulator of the maltose genes in *E. faecium*. However, no MalR  
358 binding site or *cre*-box could be identified upstream of the *pula* gene. This suggests that the

359 maltodextrin metabolism in these two species are regulated differently, *E. faecium* possesses a  
360 specific regulator for the maltodextrin metabolism but in *E. faecalis* the maltodextrin operons are  
361 also under control of MalR.

362

## 363 Materials and methods

364 **Bacterial strains and growth conditions.** Strains and plasmids used in this study are listed in  
365 Table S1. *Escherichia coli* cells were grown in Lysogeny Broth (LB) (43). Following antibiotics  
366 (Sigma-Aldrich) were added when appropriate: Ampicillin, Kanamycin and Chloramphenicol with  
367 final concentration of 50, 25 and 10 µg/mL respectively.

368 *Enterococcus faecalis* strains (Table S1) were routinely grown in M17 medium (44) supplemented  
369 with glucose at 0.5% (w/v) (GM17) at 37°C. Growth and transcriptional studies were performed  
370 using carbon cured M17 MOPS (ccM17) (45) supplemented with 0.3% (w/v) glucose, maltose,  
371 and/or maltopentaose according to experimental needs. When appropriate, cultures were  
372 supplemented with Erythromycine or Chloramphenicol (Sigma-Aldrich) with final concentrations  
373 of 100 and 2 µg/mL respectively.

374 **Construction of mutant.** The *malR*<sup>-</sup> $\Delta$ *ccpA* markerless double mutant (MC) was obtained after  
375 transformation of the pMAD: $\Delta$ *ccpA* plasmid (Table S1) in a strain carrying a mutant allele for the  
376 *malR* gene (*malR*<sup>-</sup>). For more details concerning the construction of the pMAD: $\Delta$ *ccpA*, please refer  
377 to ((24), unpublished data).

378 **Complementation of mutants.** For complementation of the *ccpA* gene, corresponding wild type  
379 genes and its promoter were first cloned into the pRB473 plasmid (46) and introduced in *E. faecalis*  
380 carrying a deletion in the *ccpA* gene. The complementation of the *malR* gene was performed as  
381 previously described ((24), unpublished data).

382 **Purification of His-tagged proteins and *in vitro* phosphorylation of the HPr protein.** For the  
383 purification of His-tagged proteins, the plasmids containing the genes of interest were used (Table  
384 S1) as previously described ((24), unpublished data). Briefly the His tagged proteins were purified  
385 with the kit Protino Ni-NTA (Macherey Nagel). The imidazole was removed from the buffer using  
386 a PD-10 Desalting Column (GE Healthcare). Concentrations of proteins were determined with the  
387 kit Pierce™ BCA Protein Assay (ThermoFisher Scientific) with a BSA standard curved and

388 immediately stored at -80°C.

389 HPr was subsequently phosphorylated *in vitro* using purified EI enzyme or HPrV267F in order to  
390 obtain P~His-HPr or P-Ser-HPr respectively as previously described (47)(.

391 **Isolation of total RNA and transcript analysis.** Cells were grown in ccM17 medium  
392 supplemented with appropriate carbon sources, total RNA were extracted using Direct-zol™ RNA  
393 Miniprep (Zymo Research), cDNA were synthesized with the QuantiTect Reverse Transcription  
394 Kit (Qiagen) and RT-qPCR analysis were performed using GoTaq® qPCR Master Mix (Promega)  
395 as previously described ((24), unpublished data). Primers used for RT-qPCR are shown in Table S2.

396 **Electrophoretic Mobility Shift Assays (EMSA).** DNA fragments containing *mdxE* promoter or  
397 *mmdH* promoter were first cloned into the pGEM-T Easy plasmid (Promega) following  
398 manufacturer's instructions. DNA sequences were fluorescently labeled by PCR using a primer  
399 linked to a cyanin-5 fluorescent protein at the 5'end (Table S2) and appropriate pGEM-T derivative  
400 plasmids as template. A cyanin-5 labeled PCR product containing an internal fragment of the *gusA*  
401 gene was used as negative control to check for non-specific binding. To produce the Pmaltodextrin,  
402 genomic DNA of *E. Faecalis* OG1RF was used as template with a primer containing a cyanin-5  
403 (Table S2). Assays were performed as previously described ((24), unpublished data) in a binding  
404 buffer (20 mM Tris, pH7.4, 20 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM FBP, 1 mM DTT, 0.1 mM EDTA,  
405 5% (v/v) glycerol, 50 ng/μL poly(dI-dC), 0.01 mg/mL Bovine Serum Albumin (BSA)).

406 **Supernatant carbohydrates quantification.** Glucose, maltose and maltopentaose contents of  
407 culture supernatants were quantified using High Performance Liquid Chromatography (HPLC).  
408 Cells were removed by centrifugation and 20 μl of supernatants were analysed after injection into  
409 an Aminex HPX-87H column 300x7.8 mm (BioRad) and eluted at 65°C with by 5 mM H<sub>2</sub>SO<sub>4</sub> at a  
410 flow rate of 0.6 mL min<sup>-1</sup> for 25 min. Carbohydrates were detected using a  
411 2142 Differential Refractometer (LKB Bromma), absolute concentrations of raw data collected  
412 were estimated against a standard curve and standardized by the non metabolizable

413 3-(N-morpholino)propanesulfonic acid (MOPS) amount.

414 **Quantification of different phosphorylated forms of HPr by Western-Blot.** Cells were pellet  
415 by centrifugation after grown to OD<sub>600</sub> of 0.6 and resuspended in 0.05 volume of extraction buffer  
416 (10 mM Tris pH 7.4, 1mM EDTA, 0.5% (v/v) β-mercaptoethanol, 0.05% (v/v) Tween20). Cells  
417 were disrupted mechanically using a FastPrep-24 (MP Biomedicals) with 0.1 mm silica glass beads,  
418 cell debris were removed by subsequent centrifugation. The proteins were separated by  
419 nondenaturing polyacrylamide gel electrophoresis (Native PAGE) on 14% total acrylamide gel  
420 with a Tris-Glycine buffer (25 mM Tris pH 8.3, 192 mM glycine) in a 4°C chilled room and  
421 transferred to a polyvinylidene difluoride membrane (PVDF) using the iBlot® Gel Transfer Stacks  
422 PVDF, Mini kit (Invitrogen) following supplier's recommendations. Different forms of HPr were  
423 detected with a rabbit polyclonal antiserum raised against HPr of *Staphylococcus carnosus* (35).  
424 Rabbit HPr antibodies were finally visualized using a goat anti-rabbit IgG alkaline phosphatase  
425 conjugate (Sigma-Aldrich) and the ECL Western blot ... kit (GE Healthcare) according to  
426 manufacturer's indications.

427 **Footprinting Assays.** DNA regions were amplified as described above with primers pU (FAM)  
428 and pR (Table S2). Interaction with MalR and FAM-labeled PCR products and the DNase  
429 digestion were performed as previously described ((24), unpublished data).

430

## 431 Acknowledgements

432 We thank Reinhold Brückner for the generous gift of plasmid pRB473. We greatly thank Isabelle  
433 Rince for technical assistance.

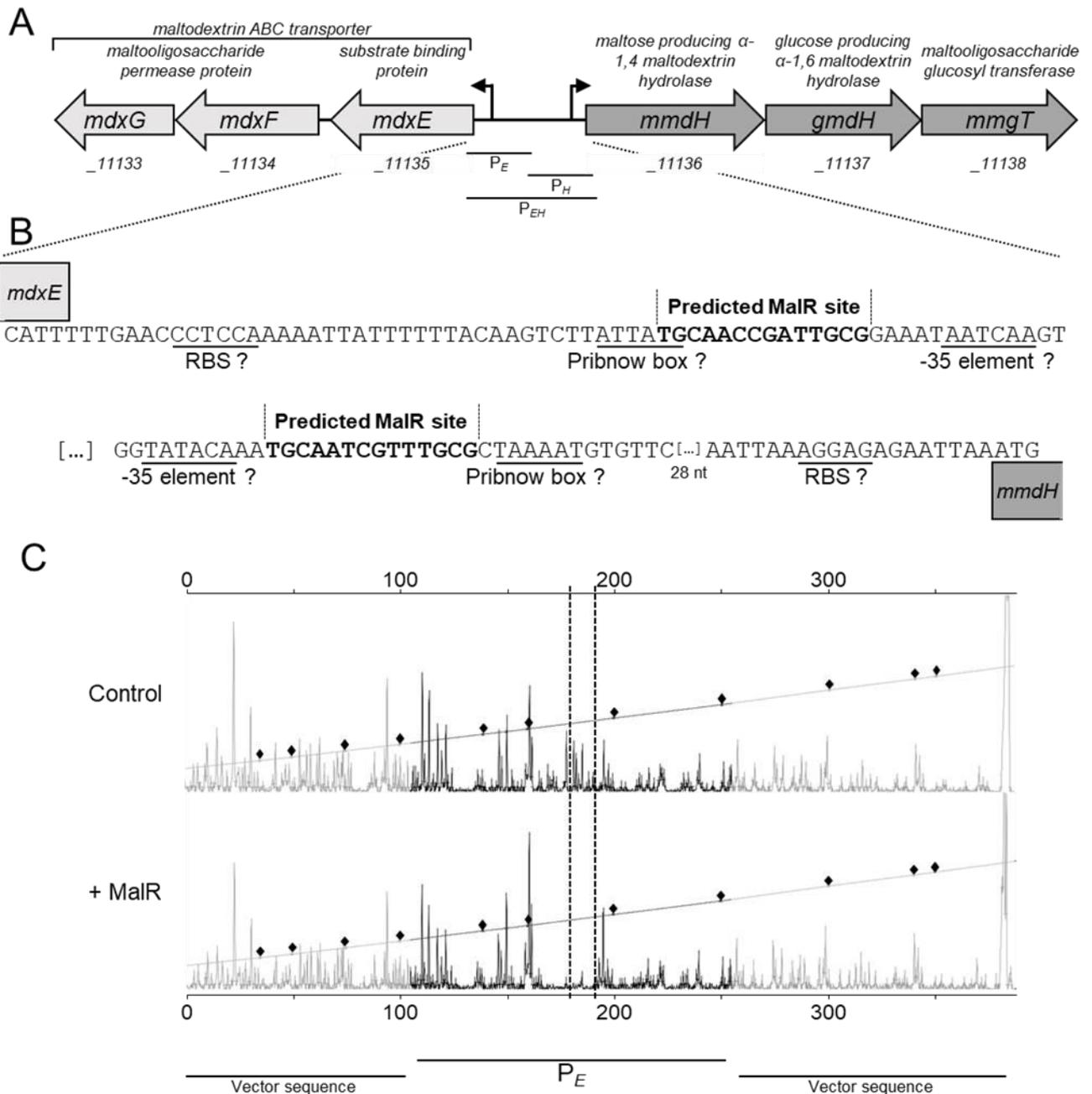
434

- 435 1. Lebreton F, Willems RJL, Gilmore MS. 2014. Enterococcus Diversity, Origins in Nature, and Gut  
436 Colonization, p. . *In* Gilmore, MS, Clewell, DB, Ike, Y, Shankar, N (eds.), *Enterococci: From Commensals to Leading*  
437 *Causes of Drug Resistant Infection*. Massachusetts Eye and Ear Infirmary, Boston.
- 438 2. Wisplinghoff H, Bischoff T, Tallent SM, Seifert H, Wenzel RP, Edmond MB. 2004. Nosocomial Bloodstream  
439 Infections in US Hospitals: Analysis of 24,179 Cases from a Prospective Nationwide Surveillance Study. *Clin Infect*  
440 *Dis* 39:309–317.
- 441 3. Agudelo Higueta NI, Huycke MM. 2014. Enterococcal Disease, Epidemiology, and Implications for  
442 Treatment, p. . *In* Gilmore, MS, Clewell, DB, Ike, Y, Shankar, N (eds.), *Enterococci: From Commensals to Leading*  
443 *Causes of Drug Resistant Infection*. Massachusetts Eye and Ear Infirmary, Boston.
- 444 4. García-Solache M, Rice LB. 2019. The Enterococcus: a Model of Adaptability to Its Environment. *Clin*  
445 *Microbiol Rev* 32.
- 446 5. Bale MJ, Bennett PM, Beringer JE, Hinton M. 1993. The survival of bacteria exposed to desiccation on  
447 surfaces associated with farm buildings. *J Appl Bacteriol* 75:519–528.
- 448 6. Fisher K, Phillips C. 2009. The ecology, epidemiology and virulence of Enterococcus. *Microbiology*  
449 155:1749–1757.
- 450 7. Kristich CJ, Rice LB, Arias CA. 2014. Enterococcal Infection—Treatment and Antibiotic Resistance, p. . *In*  
451 Gilmore, MS, Clewell, DB, Ike, Y, Shankar, N (eds.), *Enterococci: From Commensals to Leading Causes of Drug*  
452 *Resistant Infection*. Massachusetts Eye and Ear Infirmary, Boston.
- 453 8. Garsin DA, Frank KL, Silanpää J, Ausubel FM, Hartke A, Shankar N, Murray BE. 2014. Pathogenesis and  
454 Models of Enterococcal Infection, p. . *In* Gilmore, MS, Clewell, DB, Ike, Y, Shankar, N (eds.), *Enterococci: From*  
455 *Commensals to Leading Causes of Drug Resistant Infection*. Massachusetts Eye and Ear Infirmary, Boston.
- 456 9. Chang D-E, Smalley DJ, Tucker DL, Leatham MP, Norris WE, Stevenson SJ, Anderson AB, Grissom JE,  
457 Laux DC, Cohen PS, Conway T. 2004. Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proc Natl Acad Sci*  
458 *U S A* 101:7427–7432.
- 459 10. Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firbank SJ, Bolam DN, Sonnenburg JL. 2010.  
460 Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell*  
461 141:1241–1252.
- 462 11. Thomas M, Wrzosek L, Ben-Yahia L, Noordine M-L, Gitton C, Chevret D, Langella P, Mayeur C, Cherbuy  
463 C, Rul F. 2011. Carbohydrate metabolism is essential for the colonization of *Streptococcus thermophilus* in the  
464 digestive tract of gnotobiotic rats. *PLoS One* 6:e28789.
- 465 12. Koropatkin NM, Cameron EA, Martens EC. 2012. How glycan metabolism shapes the human gut microbiota.  
466 *Nat Rev Microbiol* 10:323–335.
- 467 13. Ng KM, Ferreyra JA, Higginbottom SK, Lynch JB, Kashyap PC, Gopinath S, Naidu N, Choudhury B, Weimer  
468 BC, Monack DM, Sonnenburg JL. 2013. Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric  
469 pathogens. *Nature* 502:96–99.
- 470 14. Xu J, Xie Y-D, Liu L, Guo S, Su Y-L, Li A-X. 2019. Virulence regulation of cel-EIIB protein mediated PTS  
471 system in *Streptococcus agalactiae* in Nile tilapia. *J Fish Dis* 42:11–19.
- 472 15. Muller C, Cacaci M, Sauvageot N, Sanguinetti M, Rattei T, Eder T, Giard J-C, Kalinowski J, Hain T, Hartke  
473 A. 2015. The Intraperitoneal Transcriptome of the Opportunistic Pathogen *Enterococcus faecalis* in Mice. *PLoS One*  
474 10:e0126143.
- 475 16. Ruiz-Cruz S, Espinosa M, Goldmann O, Bravo A. 2016. Global Regulation of Gene Expression by the MafR  
476 Protein of *Enterococcus faecalis*. *Front Microbiol* 6.
- 477 17. Ruiz-Cruz S, Moreno-Blanco A, Espinosa M, Bravo A. 2019. Transcriptional activation by MafR, a global  
478 regulator of *Enterococcus faecalis*. *Sci Rep* 9:6146.
- 479 18. Sauvageot N, Mokhtari A, Joyet P, Budin-Verneuil A, Blancato VS, Repizo GD, Henry C, Píkis A, Thompson  
480 J, Magni C, Hartke A, Deutscher J. 2017. *Enterococcus faecalis* uses a PTS permease and a host colonization-related  
481 ABC transporter for maltodextrin uptake. *J Bacteriol*.
- 482 19. Hava DL, Camilli A. 2002. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence  
483 factors. *Mol Microbiol* 45:1389–1406.
- 484 20. Shelburne SA, Sumbly P, Sitkiewicz I, Okorafor N, Granville C, Patel P, Voyich J, Hull R, DeLeo FR, Musser  
485 JM. 2006. Maltodextrin utilization plays a key role in the ability of group A *Streptococcus* to colonize the oropharynx.  
486 *Infect Immun* 74:4605–4614.

- 487 21. Joyet P, Mokhtari A, Riboulet-Bisson E, Blancato VS, Espariz M, Magni C, Hartke A, Deutscher J, Sauvageot  
488 N. 2017. The enzymes required for maltodextrin catabolism in *Enterococcus faecalis* exhibit several novel activities.  
489 *Appl Environ Microbiol*.
- 490 22. Le Breton Y, Pichereau V, Sauvageot N, Auffray Y, Rincé A. 2005. Maltose utilization in *Enterococcus*  
491 *faecalis*. *J Appl Microbiol* 98:806–813.
- 492 23. Mokhtari A, Blancato VS, Repizo GD, Henry C, Pikis A, Bourand A, de Fátima Álvarez M, Immel S,  
493 Mechakra-Maza A, Hartke A, Thompson J, Magni C, Deutscher J. 2013. *Enterococcus faecalis* utilizes maltose by  
494 connecting two incompatible metabolic routes via a novel maltose 6'-phosphate phosphatase (MapP). *Mol Microbiol*  
495 88:234–253.
- 496 24. Grand M, Blancato VS, Espariz M, Deutscher J, Pikis A, Hartke A, Magni C, Sauvageot N. 2019. The dual  
497 role of MalR in induction and a new P Ser HPr requiring catabolite repression mechanism of the *Enterococcus faecalis*  
498 *mal* operons. *Mol Microbiol* ??
- 499 25. Deutscher J. 2008. The mechanisms of carbon catabolite repression in bacteria. *Curr Opin Microbiol* 11:87–  
500 93.
- 501 26. Jones BE, Dossonnet V, Küster E, Hillen W, Deutscher J, Klevit RE. 1997. Binding of the catabolite repressor  
502 protein CcpA to its DNA target is regulated by phosphorylation of its corepressor HPr. *J Biol Chem* 272:26530–26535.
- 503 27. Deutscher J, Saier MH. 1983. ATP-dependent protein kinase-catalyzed phosphorylation of a seryl residue in  
504 HPr, a phosphate carrier protein of the phosphotransferase system in *Streptococcus pyogenes*. *Proc Natl Acad Sci U S*  
505 *A* 80:6790–6794.
- 506 28. Deutscher J, Francke C, Postma PW. 2006. How phosphotransferase system-related protein phosphorylation  
507 regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev MMBR* 70:939–1031.
- 508 29. Homburg C, Bommer M, Wuttge S, Hobe C, Beck S, Dobbek H, Deutscher J, Licht A, Schneider E. 2017.  
509 Inducer exclusion in Firmicutes: Insights into the regulation of a carbohydrate ATP binding cassette transporter from  
510 *Lactobacillus casei* BL23 by the signal transducing protein P-Ser46-HPr. *Mol Microbiol*.
- 511 30. Deutscher J, Kessler U, Alpert CA, Hengstenberg W. 1984. Bacterial phosphoenolpyruvate-dependent  
512 phosphotransferase system: P-Ser-HPr and its possible regulatory function? *Biochemistry (Mosc)* 23:4455–4460.
- 513 31. Kravanja M, Engelmann R, Dossonnet V, Blüggel M, Meyer HE, Frank R, Galinier A, Deutscher J, Schnell  
514 N, Hengstenberg W. 1999. The *hprK* gene of *Enterococcus faecalis* encodes a novel bifunctional enzyme: the HPr  
515 kinase/phosphatase. *Mol Microbiol* 31:59–66.
- 516 32. Chaptal V, Vincent F, Gueguen-Chaignon V, Monedero V, Poncet S, Deutscher J, Nessler S, Morera S. 2007.  
517 Structural analysis of the bacterial HPr kinase/phosphorylase V267F mutant gives insights into the allosteric regulation  
518 mechanism of this bifunctional enzyme. *J Biol Chem* 282:34952–34957.
- 519 33. Zhang X, Rogers M, Bierschenk D, Bonten MJM, Willems RJJ, van Schaik W. 2013. A LacI-Family  
520 Regulator Activates Maltodextrin Metabolism of *Enterococcus faecium*. *PLoS ONE* 8:e72285.
- 521 34. Opsata M, Nes IF, Holo H. 2010. Class IIa bacteriocin resistance in *Enterococcus faecalis* V583: The mannose  
522 PTS operon mediates global transcriptional responses. *BMC Microbiol* 10:224.
- 523 35. Leboeuf C, Leblanc L, Auffray Y, Hartke A. 2000. Characterization of the *ccpA* Gene of *Enterococcus*  
524 *faecalis*: Identification of Starvation-Inducible Proteins Regulated by CcpA. *J Bacteriol* 182:5799–5806.
- 525 36. Vadeboncoeur C, Brochu D, Reizer J. 1991. Quantitative determination of the intracellular concentration of  
526 the various forms of HPr, a phosphocarrier protein of the phosphoenolpyruvate: Sugar phosphotransferase system in  
527 growing cells of oral streptococci. *Anal Biochem* 196:24–30.
- 528 37. Nieto C, Espinosa M, Puyet A. 1997. The maltose/maltodextrin regulon of *Streptococcus pneumoniae*  
529 Differential promoter regulation by the transcriptional repressor MalR. *J Biol Chem* 272:30860–30865.
- 530 38. Nieto C, Puyet A, Espinosa M. 2001. MalR-mediated Regulation of the *Streptococcus pneumoniae* malMP  
531 Operon at PromoterPM : influence of a proximal divergent promoter region and competition between MalR and RNA  
532 polymerase proteins. *J Biol Chem* 276:14946–14954.
- 533 39. Afzal M, Shafeeq S, Manzoor I, Kuipers OP. 2015. Maltose-Dependent Transcriptional Regulation of the mal  
534 Regulon by MalR in *Streptococcus pneumoniae*. *PLoS One* 10:e0127579.
- 535 40. Blüschke B, Volkmer-Engert R, Schneider E. 2006. Topography of the surface of the signal-transducing  
536 protein EIIA(Glc) that interacts with the MalK subunits of the maltose ATP-binding cassette transporter (MalFGK2)  
537 of *Salmonella typhimurium*. *J Biol Chem* 281:12833–12840.
- 538

- 539 41. Wuttge S, Licht A, Timachi MH, Bordignon E, Schneider E. 2016. Mode of Interaction of the Signal-  
540 Transducing Protein EIIA(Glc) with the Maltose ABC Transporter in the Process of Inducer Exclusion. *Biochemistry*  
541 (Mosc) 55:5442–5452.
- 542 42. Ferreira MJ, Sá-Nogueira I de. 2010. A multitask ATPase serving different ABC-type sugar importers in  
543 *Bacillus subtilis*. *J Bacteriol* 192:5312–5318.
- 544 43. Bertani G. 1951. Studies on lysogenesis: the mode of phage liberation by lysogenic *Escherichia coli*. *J*  
545 *Bacteriol* 62:293–300.
- 546 44. Terzaghi BE, Sandine WE. 1975. Improved medium for lactic streptococci and their bacteriophages. *Appl*  
547 *Microbiol* 29:807–813.
- 548 45. Bizzini A, Zhao C, Budin-Verneuil A, Sauvageot N, Giard J-C, Auffray Y, Hartke A. 2010. Glycerol is  
549 metabolized in a complex and strain-dependent manner in *Enterococcus faecalis*. *J Bacteriol* 192:779–785.
- 550 46. Brückner R, Wagner E, Götz F. 1993. Characterization of a sucrase gene from *Staphylococcus xylosus*. *J*  
551 *Bacteriol* 175:851–857.
- 552 47. Monedero V, Poncet S, Mijakovic I, Fieulaine S, Dossonnet V, Martin-Verstraete I, Nessler S, Deutscher J.  
553 2001. Mutations lowering the phosphatase activity of HPr kinase/phosphatase switch off carbon metabolism. *EMBO*  
554 *J* 20:3928–3937.
- 555 48. Dunny GM, Brown BL, Clewell DB. 1978. Induced cell aggregation and mating in *Streptococcus faecalis*:  
556 evidence for a bacterial sex pheromone. *Proc Natl Acad Sci U S A* 75:3479–3483.
- 557 49. Suárez CA, Blancato VS, Poncet S, Deutscher J, Magni C. 2011. CcpA represses the expression of the  
558 divergent *cit* operons of *Enterococcus faecalis* through multiple *cre* sites. *BMC Microbiol* 11:1.
- 559

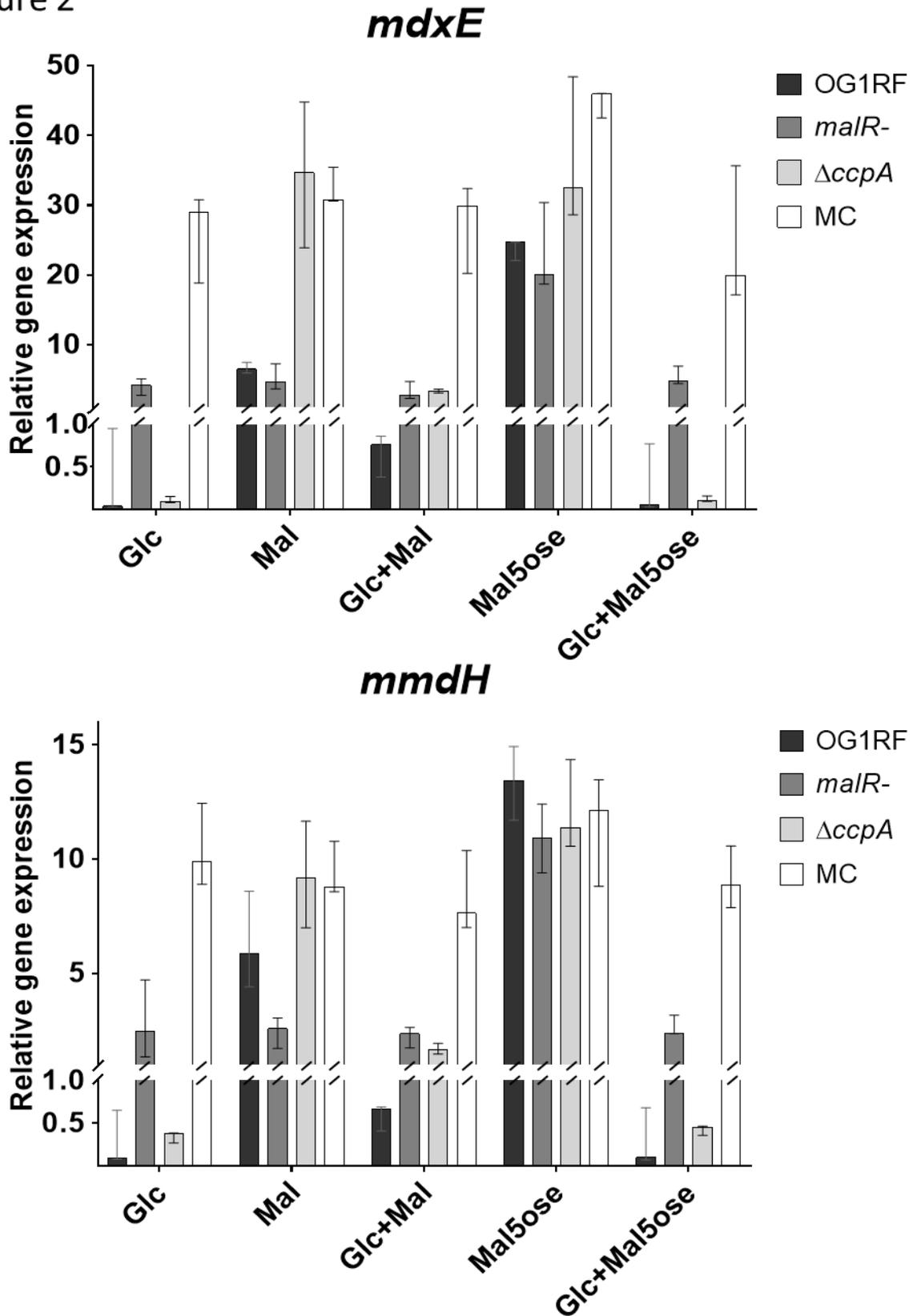
Figure 1



**FIG 1** (A) Genetic region containing genes used for maltodextrins metabolism in *E. faecalis* OG1RF. Large arrows represent ORFs with their orientation corresponding to transcription direction. Promoters are indicated by thin bent arrows. Proven function of each gene is indicated above, name of gene in the arrow and reference loci below. (B) Nucleotide sequence of promoters of the genes *mdxE* (upper sequence) and *mmdH* (lower sequence). Putatives Ribosome Binding Sites (RBS), Pribnow box and -35 elements are underlined. Predicted MalR binding sites are marked in bold. (C) Dnase I footprinting assays carried out for MalR on the  $P_E$  region. The read outs were aligned based on an internal 500ROX size standard which positions are marked by diamonds on the diagonal. The positions of predicted MalR binding sites are marked out with broken lines.

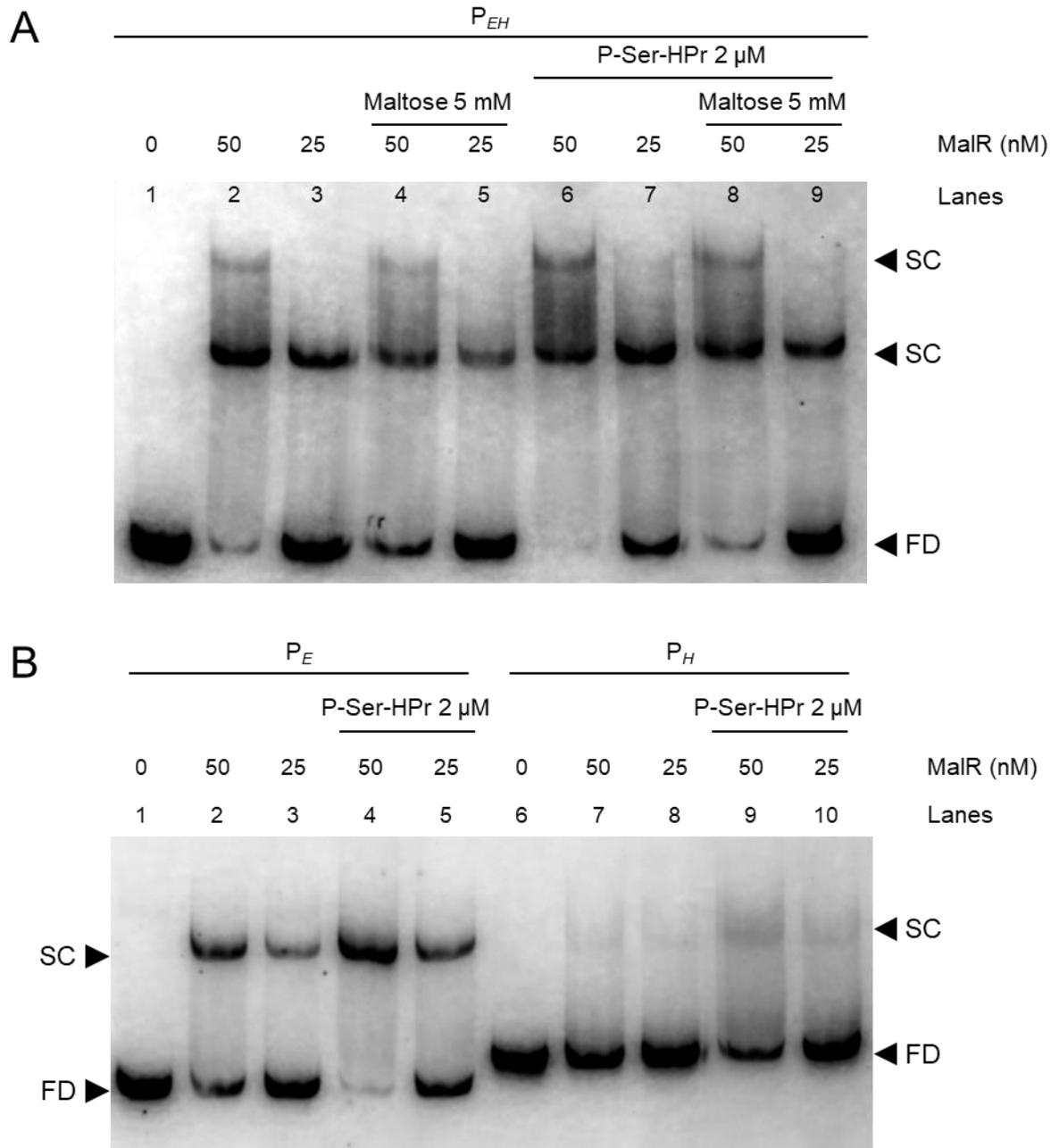
Nucleotides in positions 103 to 256 correspond to the  $P_E$  region, the external positions (represented by transparent regions) correspond to the pGEM-T Easy vector sequence.

Figure 2



**FIG 2** Relative expression of *mdxE* (upper half) and *mmdH* genes (lower half) in the wild type strain OG1RF (dark grey■), *malR*- (grey■),  $\Delta ccpA$  (light grey■) and MC (white□) mutants. Expression of the housekeeping *gyrA* gene is used as reference to normalize the total quantity of cDNA. Median values calculated from at least three independent experiments are presented in bars, error bars represent ranges of data.

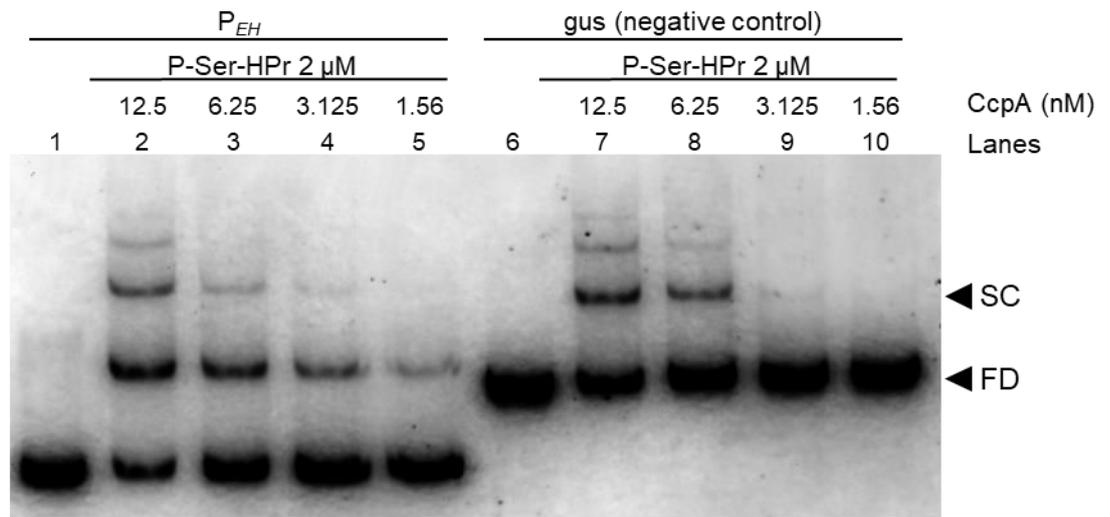
Figure 3



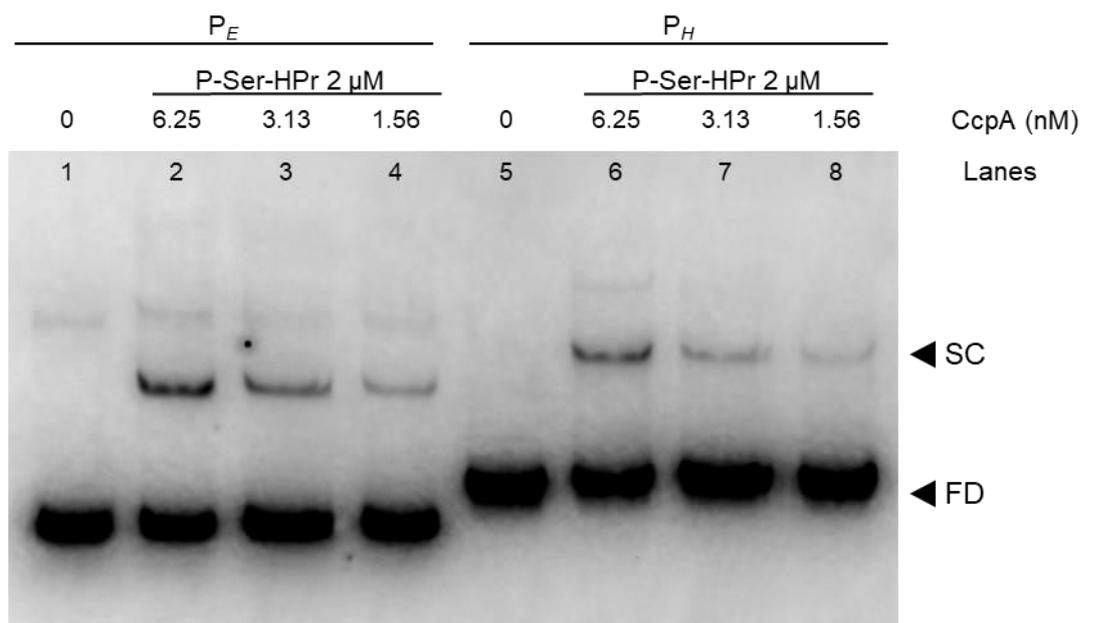
**FIG 3** EMSA results for MaIR on DNA fragments containing either the  $P_E$ , the  $P_H$  or both regions ( $P_{EH}$ ). The assays were performed on 5 nM of Cyanine-5 labeled PCR products. Proteins or maltose concentrations used are indicated above. Free DNA (FD) and shifted complex (SC) are indicated by arrowheads. (A) EMSA results for MaIR binding to the  $P_{EH}$  region. Maltose was added to a final concentration of 5 mM to samples loaded in lanes 4, 5, 8, and 9. P-Ser-HPr (2  $\mu$ M final concentration) was added to samples loaded on lanes 6 to 9. (B) EMSA results for MaIR binding on  $P_E$  (lanes 1 to 5) or  $P_H$  (lanes 6 to 10). P-Ser-HPr (2  $\mu$ M final concentration) was added to samples loaded on lanes 4, 5, 9, and 10.

Figure 4

A

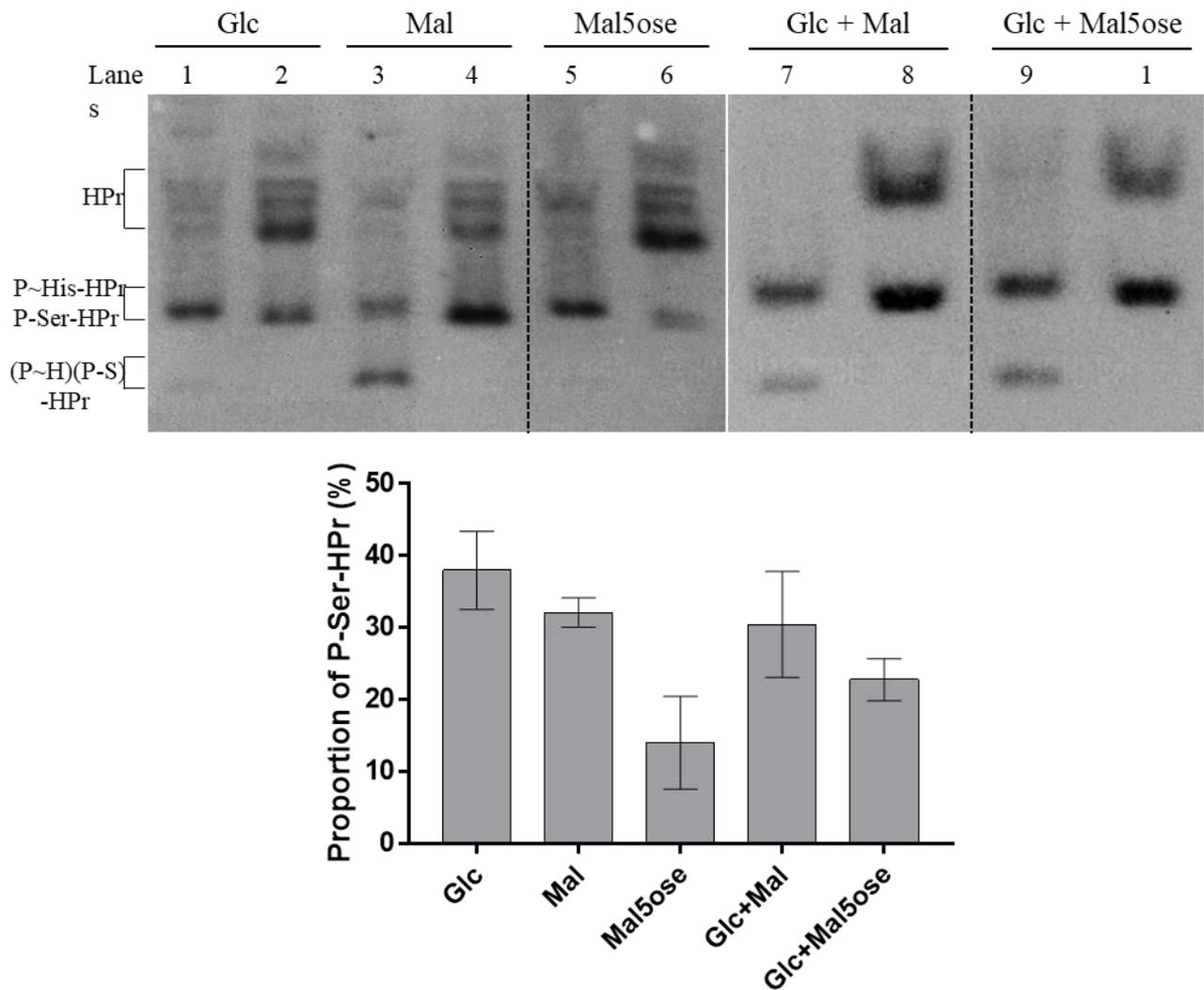


B



**FIG 4** EMSA results for CcpA on DNA fragments containing either the  $P_E$ , the  $P_H$ , both regions ( $P_{EH}$ ) or an internal fragment of the *gusA* gene. The assays were performed on 5 nM of Cyanine-5 labeled PCR products. Proteins concentrations used are indicated above. Free DNA (FD), shifted complex (SC) or non-specific amplification (NS) are indicated by arrowheads. (A) EMSA results for CcpA/P-Ser-HPr binding to the  $P_{EH}$  (lanes 1 to 5) or an internal fragment of the *gusA* gene (lanes 6 to 10). (B) EMSA results for CcpA/P-Ser-HPr binding on  $P_E$  (lanes 1 to 4) or  $P_H$  (lanes 5 to 8).

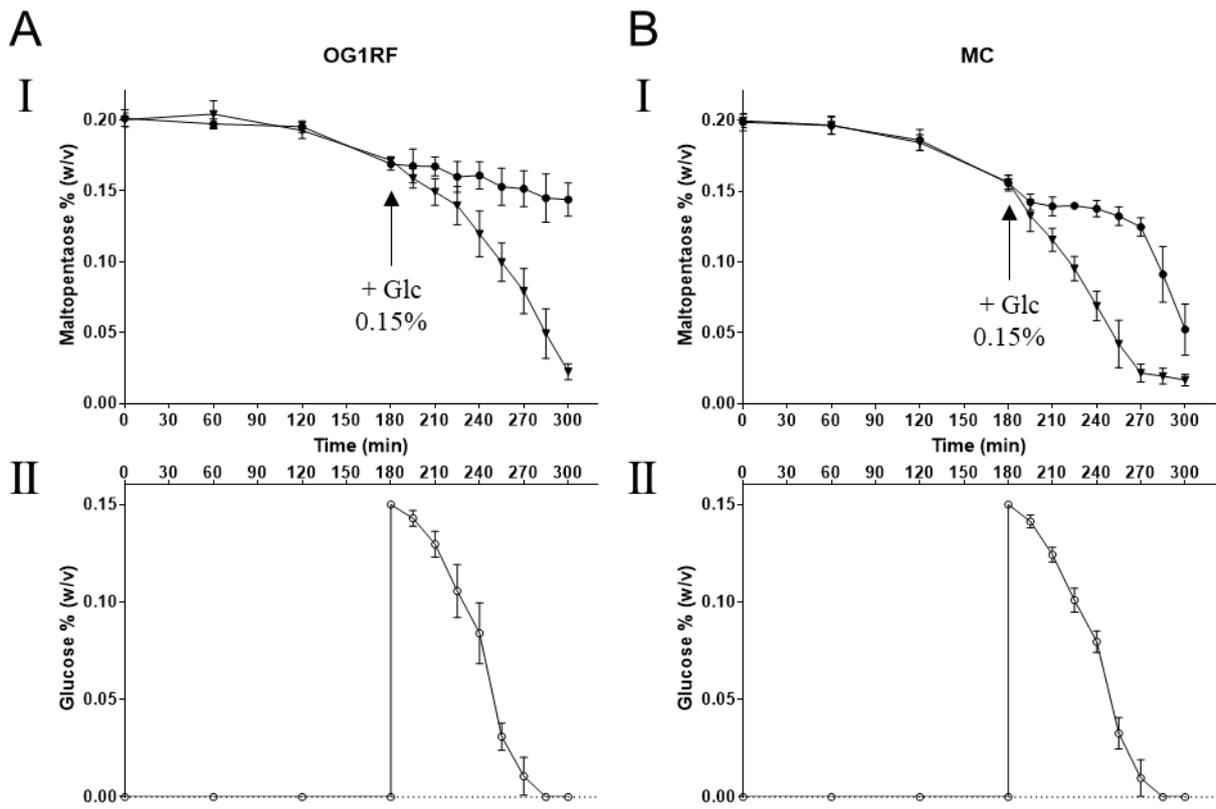
Figure 5



**FIG 5** Western Blot analysis on HPr of *E. faecalis* OG1RF strain. Extracts loaded in odd numbered lanes correspond to total native proteins. Samples in even numbered lanes correspond to the same sample loaded on the previous lane boiled for 5 min to remove phosphorylations on H15 but not on S46. Positions of different forms of HPr were indicated with braces. Images were obtained from results on the same gel but not on continuous wells, the vertical broken lines indicate the cropped region. Cells were grown on ccM17 supplemented carbon source indicated above. The proportion of the P-Ser-HPr, measured after three independent experiments are indicated on bars below (mean  $\pm$  standard error).

564

Figure 6



**FIG 6** Maltopentaose and glucose concentrations (upper and lower graph respectively) in the growth medium. Measures obtained when glucose was added in the medium at 3h was represented with circles (●), points corresponding to the condition without addition of glucose was symbolised by triangles (▼). The addition of glucose is indicated by an arrow for more convenience. (A) Results obtained with the wild type strain on the left and (B) with the MC mutant on the right.

565

Figure S1

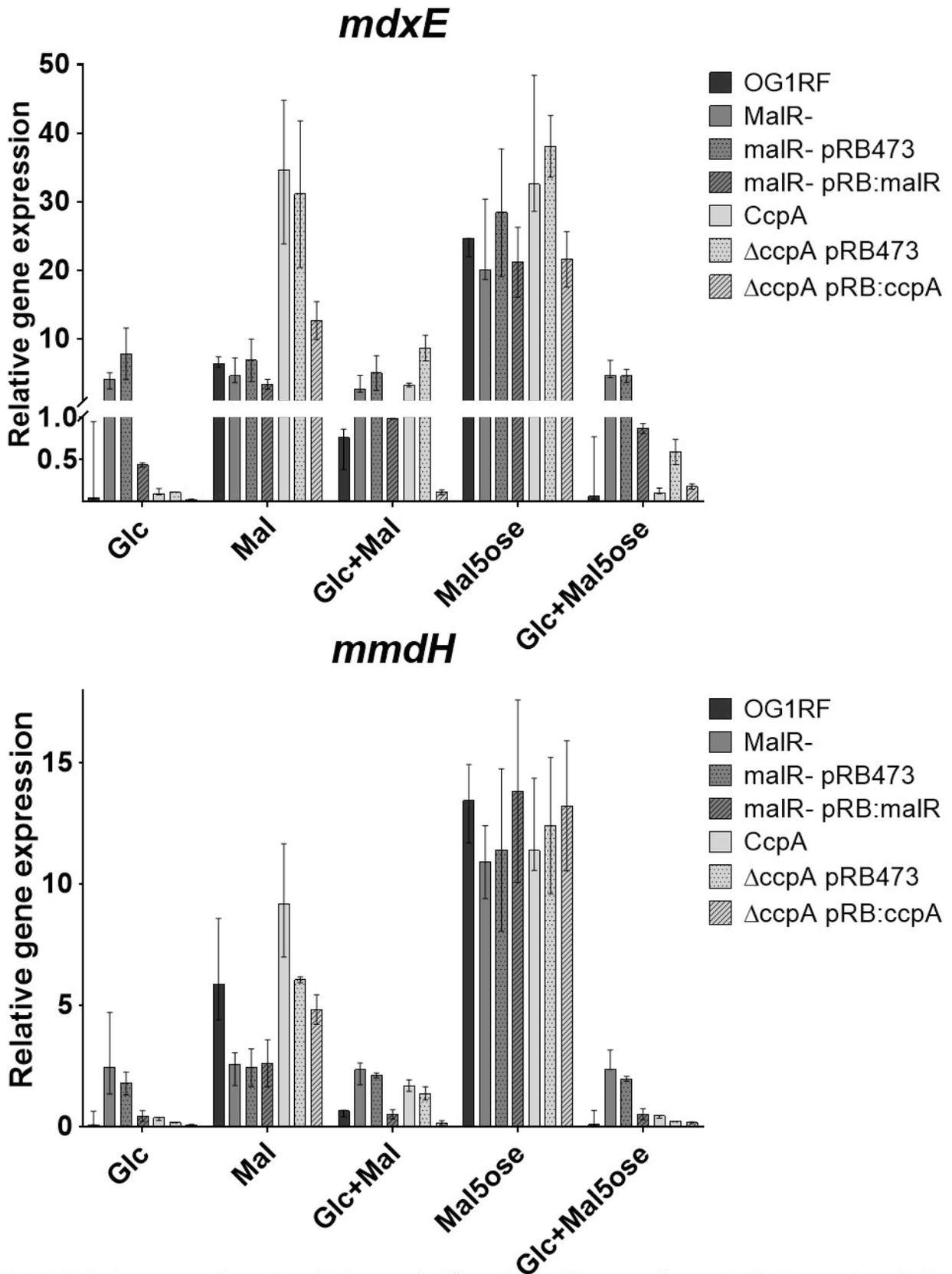
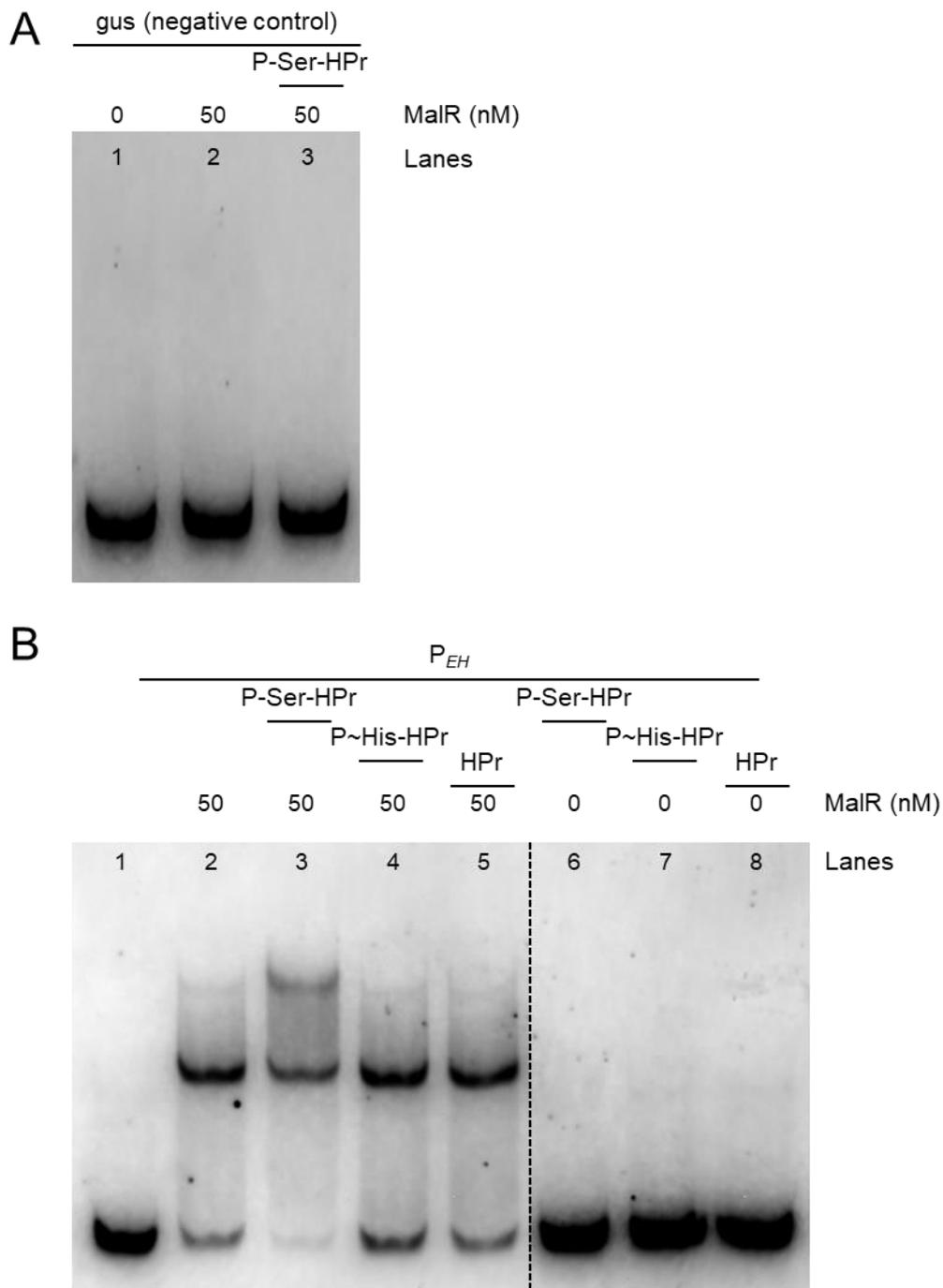


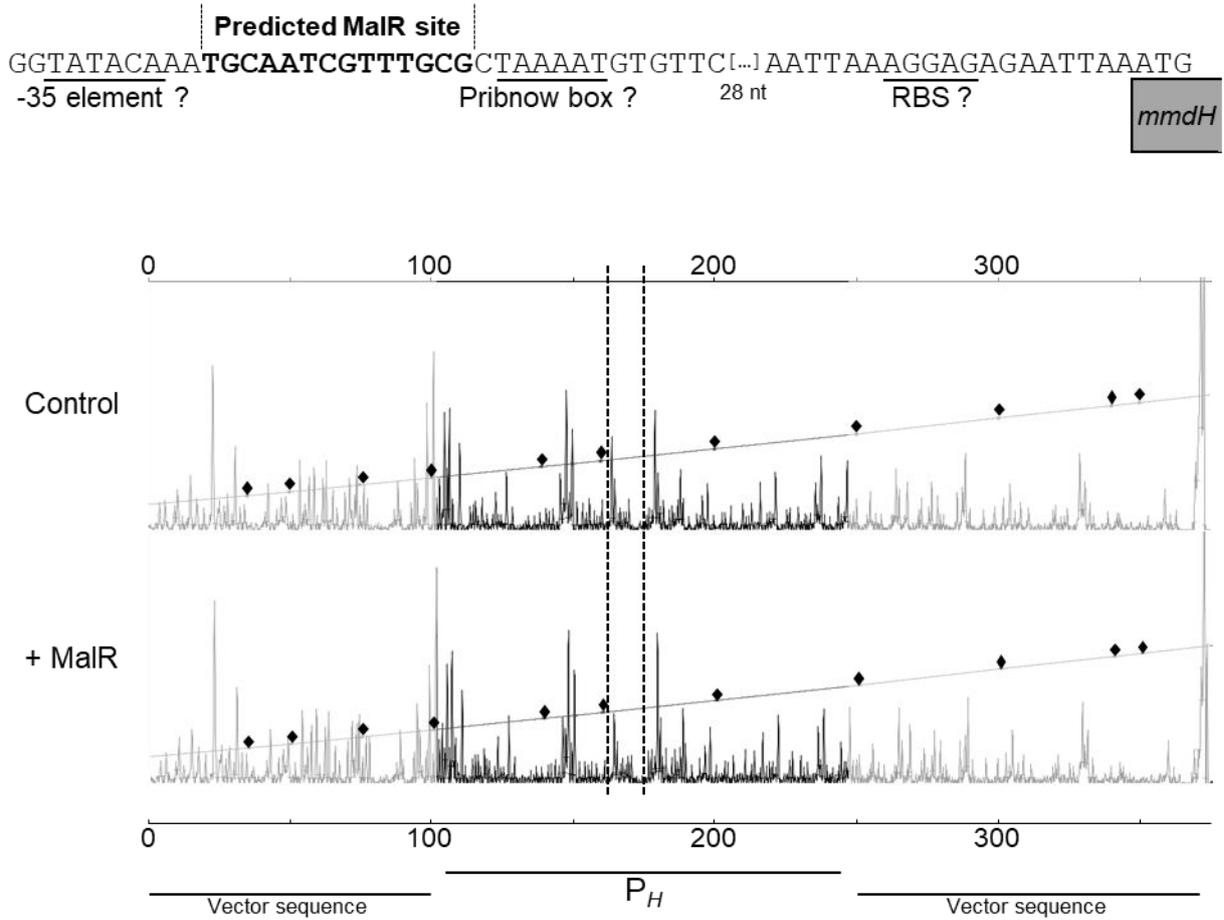
FIG S1 Relative expression of *mdxE* (upper half) and *mmdH* genes (lower half). Expression of the housekeeping *gyrA* gene is used as reference to normalize the total quantity of cDNA. Median values calculated from at least three independent experiments are presented in bars, error bars represent ranges of data.

## Figure S2



**FIG S2** (A) EMSA with MaIR on an internal fragment of the *gusA* gene. The assays were performed with 5 nM of Cyanine-5 labeled PCR products. 50 nM of MaIR was added in lanes 2 and 3. 2  $\mu$ M of P-Ser-HPr was added in lane 3. (B) EMSA with the  $P_{EH}$  region. The assays were performed with 5 nM of Cyanine-5 labeled PCR products. 2  $\mu$ M of different forms of HPr were added in presence (lanes 2 to 5) or absence of MaIR (lanes 6 to 8). Images were obtained from results on the same gel but not on continuous wells, the vertical broken lines indicate the cropped region.

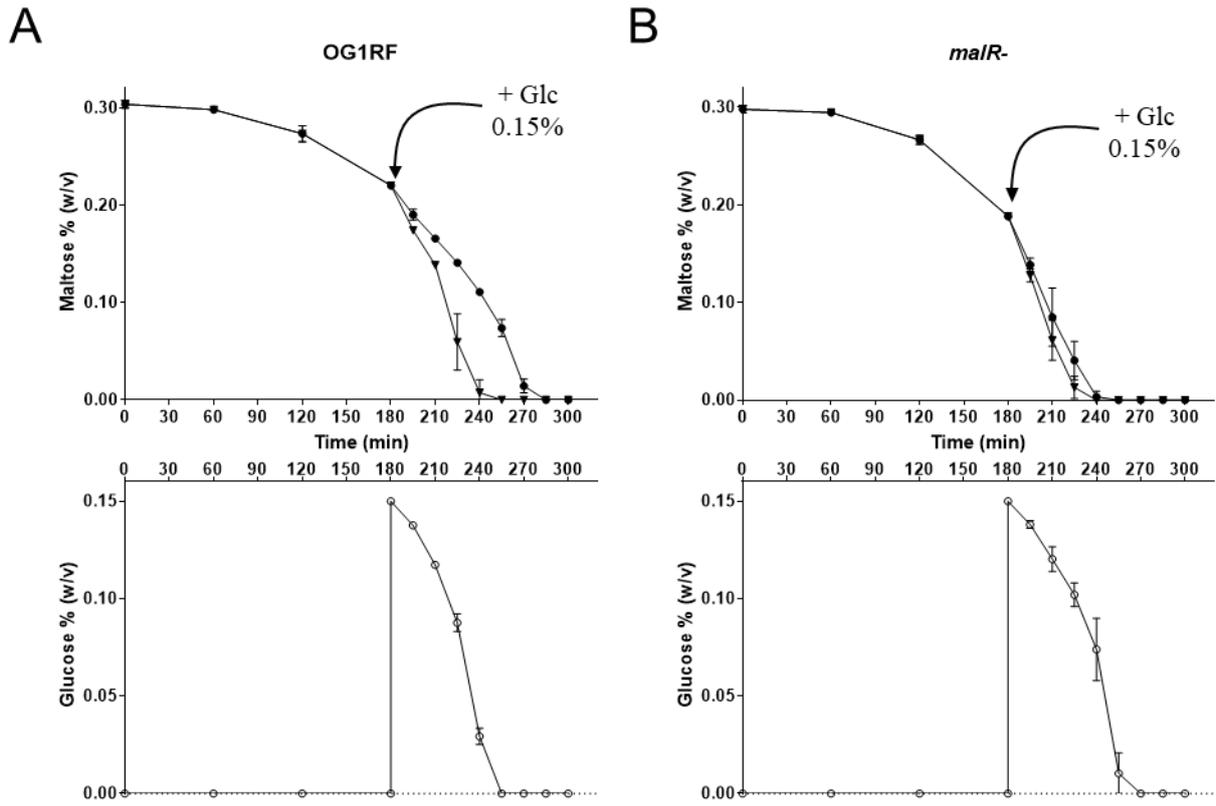
Figure S3



**FIG S3** Nucleotide sequence of the *mmdH* promoter. Putatives Ribosome Binding Sites (RBS), Pribnow box and -35 elements are underlined. Predicted MalR binding sites is marked in bold. The Dnase I footprinting assay carried out for MalR on the P<sub>H</sub> region is shown below. The read outs were aligned based on an internal 500ROX size standard which positions are marked by diamonds on the diagonal. The positions of predicted MalR binding sites are marked out with broken lines. Nucleotides in positions 103 to 247 correspond to the P<sub>H</sub> region, the external positions (represented by transparent regions) correspond to the pGEM-T Easy vector sequence.

568

Figure S4



**FIG S4** Maltose and glucose concentrations (upper and lower graph respectively) in the growth medium. Measures obtained when glucose was added in the medium at 3h was represented with circles (●), points corresponding to the condition without addition of glucose was symbolised by triangles (▼). The addition of glucose is indicated by an arrow for more convenience. (A) Results obtained with the wild type are presented on the left and (B) with the *malR* mutant on the right.

569

Strains and plasmids	Relevant properties or genotype	Ref or source
<i>Enterococcus faecalis</i>		
<b>OG1RF</b>	FusR, RifR, plasmid-free wild type strain	Dunny, 1978 (48)
<i>malR</i> <sup>-</sup>	<i>malR</i> defective mutant of OG1RF parental strain	Grand, 2019 (24)
$\Delta$ <i>ccpA</i>	<i>ccpA</i> deletion mutant of OG1RF parental strain	Grand, 2019 (24)
<b>MC</b>	Double mutant for the <i>malR</i> and <i>ccpA</i> genes	This Study
<i>Plasmids</i>		
<b>pMAD:malR</b>	pMAD containing a mutant allele of <i>malR</i> gene	Grand, 2019 (24)
<b>pMAD:<math>\Delta</math><i>ccpA</i></b>	pMAD containing a deleted allele of <i>ccpA</i> gene	Grand, 2019 (24)
<b>pQE30:HPr</b>	Expression of N-ter-6xHis tagged HPr protein	Monedero, 2001(47)
<b>pET28a:ccpA</b>	Expression of N-ter-6xHis tagged CcpA protein	Suárez, 2011 (49)
<b>pET28b:malR</b>	Expression of N-ter-6xHis tagged MalR protein	Grand, 2019 (24)
<b>pRB473</b>	Origins of replication pBR322 and pUB110, AmpR, CmR, Ble R	Brückner, 1993 (46)
<b>pRB:malR</b>	pRB473 plasmid containing wild type <i>malR</i> gene of <i>E. faecalis</i> OG1RF strain with its intern promoter	Grand, 2019 (24)
<b>pRB:ccpA</b>	pRB473 plasmid containing wild type <i>ccpA</i> gene of <i>E. faecalis</i> OG1RF strain with its own promoter	This study
<b>pGEM-T P<sub>E</sub></b>	pGEM-T Easy vector carrying the <i>mdxE</i> promoter	This study
<b>pGEM-T P<sub>H</sub></b>	pGEM-T Easy vector carrying the <i>mmdH</i> promoter	This study
<b>pGEM-T gus</b>	pGEM T Easy vector carrying an internal fragment of the <i>gusA</i> gene	Grand, 2019(24)

Name	Sequence 5' → 3'	Description
<b>PmdxE-mH_for (Cy5)</b>	AATCTTTTTCCAATTCGCTTTC	Amplification of labeled Pmaltodextrin
<b>pmdxE-mH_rev</b>	GCAGTGTTCAATTAATTCTCTCC	
<b>pU (Cy5)</b>	TGTAAAACGACGGCCAGT	Amplification of labeled PmdxE, PmmdH and gus fragment
<b>pR (M13-rev)</b>	CAGGAAACAGCTATGACC	
<b>RT-qPCR primers</b>		
<b>ef_1345L</b>	ATGCCTTCCTTTCTGGCCCT	<i>mdxE</i>
<b>ef_1345R</b>	CCATTGCCAAAATCAATGGT	
<b>ef_1347L</b>	CAACCAGGTGTCCCTTGCT	<i>mmdH</i>
<b>ef_1347R</b>	TGCTGATTTTCCTCCCAAAC	
<b>ef_0006L</b>	GATGGGGAAATCAGGGATTC	<i>gyrA</i>
<b>ef_0006R</b>	TCTTTCCATTCGGCATTTC	

574 **Figure Legends**

575 **FIG 1** (A) Genetic region containing genes used for maltodextrins metabolism in *E. faecalis*  
576 OG1RF. Large arrows represent ORFs with their orientation corresponding to transcription  
577 direction. Promoters are indicated by thin bent arrows. Proven function of each gene is indicated  
578 above, name of gene in the arrow and reference loci below. (B) Nucleotide sequence of promoters  
579 of the genes *mdxE* (upper sequence) and *mmdH* (lower sequence). Putatives Ribosome Binding  
580 Sites (RBS), Pribnow box and -35 elements are underlined. Predicted MalR binding sites are  
581 marked in bold. (C) Dnase I footprinting assays carried out for MalR on the P<sub>E</sub> region. The read  
582 outs were aligned based on an internal 500ROX size standard which positions are marked by  
583 diamonds on the diagonal. The positions of predicted MalR binding sites are marked out with  
584 broken lines. Nucleotides in positions 103 to 256 correspond to the P<sub>E</sub> region, the external positions  
585 (represented by transparent regions) correspond to the pGEM-T Easy vector sequence.

586 **FIG 2** Relative expression of *mdxE* (upper half) and *mmdH* genes (lower half) in the wild type  
587 strain OG1RF (dark grey ■), *malR*-(grey ■), *AccpA* (light grey ■) and MC (white□) mutants.  
588 Expression of the housekeeping *gyrA* gene is used as reference to normalize the total quantity of  
589 cDNA. Median values calculated from at least three independent experiments are presented in bars,  
590 error bars represent ranges of data.

591 **FIG 3** EMSA results for MalR on DNA fragments containing either the P<sub>E</sub>, the P<sub>H</sub> or both regions  
592 (P<sub>EH</sub>). The assays were performed on 5 nM of Cyanine-5 labeled PCR products. Proteins or maltose  
593 concentrations used are indicated above. Free DNA (FD) and shifted complex (SC) are indicated  
594 by arrowheads. (A) EMSA results for MalR binding to the P<sub>EH</sub> region. Maltose was added to a final  
595 concentration of 5 mM to samples loaded in lanes 4, 5, 8, and 9. P-Ser-HPr (2 μM final  
596 concentration) was added to samples loaded on lanes 6 to 9. (B) EMSA results for MalR binding  
597 on P<sub>E</sub> (lanes 1 to 5) or P<sub>H</sub> (lanes 6 to 10). P-Ser-HPr (2 μM final concentration) was added to  
598 samples loaded on lanes 4, 5, 9, and 10.

599 **FIG 4** EMSA results for CcpA on DNA fragments containing either the P<sub>E</sub>, the P<sub>H</sub>, both regions  
600 (P<sub>EH</sub>) or an internal fragment of the *gusA* gene. The assays were performed on 5 nM of Cyanine-5  
601 labeled PCR products. Proteins concentrations used are indicated above. Free DNA (FD), shifted  
602 complex (SC) or non-specific amplification (NS) are indicated by arrowheads.(A) EMSA results  
603 for CcpA/P-Ser-HPr binding to the P<sub>EH</sub> (lanes 1 to 5) or an internal fragment of the *gusA* gene  
604 (lanes 6 to 10). (B) EMSA results for CcpA/P-Ser-HPr binding on P<sub>E</sub> (lanes 1 to 4) or P<sub>H</sub> (lanes 5  
605 to 8).

606 **FIG 5** Western Blot analysis on HPr of *E. faecalis* OG1RF strain. Extracts loaded in odd numbered  
607 lanes correspond to total native proteins. Samples in even numbered lanes correspond to the same  
608 sample loaded on the previous lane boiled for 5 min to remove phosphorylations on H15 but not  
609 on S46. Positions of different forms of HPr were indicated with braces. Images were obtained  
610 from results on the same gel but not on continuous wells, the vertical broken lines indicate the  
611 cropped region. Cells were grown on ccM17 supplemented carbon source indicated above. The  
612 proportion of the P-Ser-HPr, measured after three independent experiments are indicated on bars  
613 below (mean  $\pm$  standard error).

614 **FIG 6** Maltopentaose and glucose concentrations (upper and lower graph respectively) in the  
615 growth medium. Measures obtained when glucose was added in the medium at 3h was represented  
616 with circles (●), points corresponding to the condition without addition of glucose was symbolised  
617 by triangles (▼). The addition of glucose is indicated by an arrow for more convenience. (A)  
618 Results obtained with the wild type strain on the left and (B) with the MC mutant on the right.

619 **Supplementary Figure legends**

620 **FIG S1** Relative expression of *mdxE* (upper half) and *mmdH* genes (lower half). Expression of the  
621 housekeeping *gyrA* gene is used as reference to normalize the total quantity of cDNA. Median  
622 values calculated from at least three independent experiments are presented in bars, error bars  
623 represent ranges of data.

624 **FIG S2** (A) EMSA with MalR on an internal fragment of the *gusA* gene. The assays were  
625 performed with 5 nM of Cyanine-5 labeled PCR products. 50 nM of MalR was added in lanes 2  
626 and 3. 2  $\mu$ M of P-Ser-HPr was added in lane 3. (B) EMSA with the  $P_{EH}$  region. The assays were  
627 performed with 5 nM of Cyanine-5 labeled PCR products. 2  $\mu$ M of different forms of HPr were  
628 added in presence (lanes 2 to 5) or absence of MalR (lanes 6 to 8). Images were obtained from  
629 results on the same gel but not on continuous wells, the vertical broken lines indicate the cropped  
630 region.

631 **FIG S3** Nucleotide sequence of the *mmdH* promoter. Putatives Ribosome Binding Sites (RBS),  
632 Pribnow box and -35 elements are underlined. Predicted MalR binding sites is marked in bold. The  
633 Dnase I footprinting assay carried out for MalR on the  $P_H$  region is shown below. The read outs  
634 were aligned based on an internal 500ROX size standard which positions are marked by diamonds  
635 on the diagonal. The positions of predicted MalR binding sites are marked out with broken lines.  
636 Nucleotides in positions 103 to 247 correspond to the  $P_H$  region, the external positions (represented  
637 by transparent regions) correspond to the pGEM-T Easy vector sequence.

638 **FIG S4** Maltose and glucose concentrations (upper and lower graph respectively) in the growth  
639 medium. Measures obtained when glucose was added in the medium at 3h was represented with  
640 circles (●), points corresponding to the condition without addition of glucose was symbolised by  
641 triangles (▼). The addition of glucose is indicated by an arrow for more convenience. (A) Results  
642 obtained with the wild type are presented on the left and (B) with the *malR*<sup>-</sup> mutant on the right.

643

## C. Résultats complémentaires

644 Les résultats présentés dans cette deuxième publication nous permettent d'élargir un peu le  
645 motif de fixation de MalR en incluant le motif de reconnaissance placé en amont du gène *mdxE*  
646 tout en gardant à l'esprit que le régulateur MalR présente une affinité plus faible pour ce dernier  
647 site que pour les sites placés dans la région maltose. Le potentiel site en amont du gène *mmdH* n'a  
648 pas pu être confirmé expérimentalement et MalR semble présenter une affinité très faible pour ce  
649 site *in vitro*. Par conséquent, il n'a pas été utilisé pour proposer un consensus de fixation pour MalR.  
650 Le nouveau motif identifié devient alors YGMAACCGATTGCR (où Y peut être C ou T, M peut  
651 être A ou C et R peut être A ou G) (**Figure 45**).



652

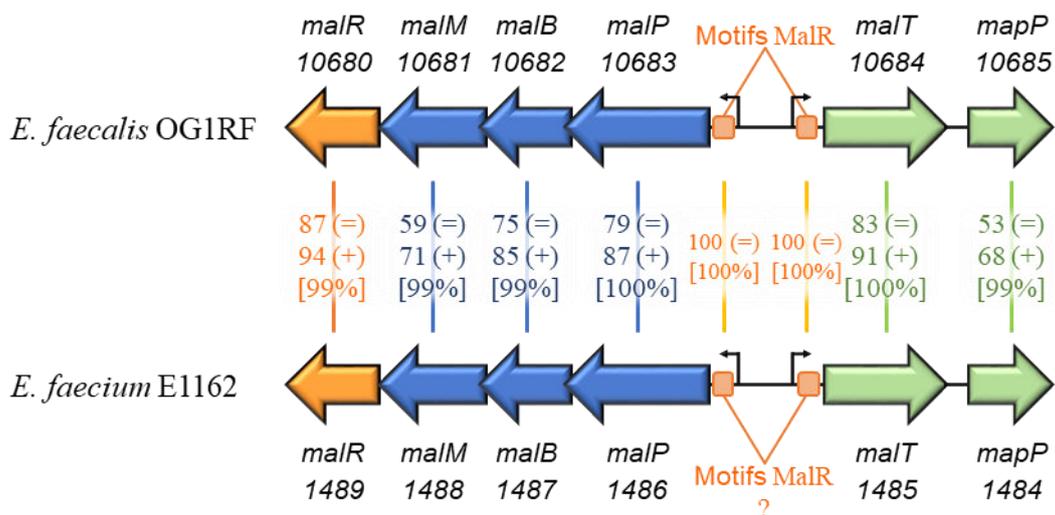
653 **Figure 45** : Séquence consensus des motifs de reconnaissance proposée pour le régulateur transcriptionnel  
654 MalR chez *E. faecalis* OG1RF

655 La taille des lettres est proportionnelle à la fréquence des nucléotides rencontrés.

656 Réalisé à l'aide de l'utilitaire WebLogo v2.8.2 <sup>(122)</sup>.

657

658 Comme décrit dans la partie IV.C de la partie introduction, les gènes *pulA-mdxDCBA*  
659 impliqués dans le métabolisme des maltodextrines chez *E. faecium* sont sous contrôle d'un  
660 activateur transcriptionnel MdxR <sup>(237)</sup>. Les gènes impliqués dans le métabolisme du maltose chez  
661 cette espèce sont eux homologues au système retrouvé chez *E. faecalis*. Les motifs de fixation de  
662 MalR dans la région intergénique entre les gènes *malP* et *malT* peuvent être retrouvés avec une  
663 similarité de 100% dans la région intergénique entre les gènes *mal* chez *E. faecium* E1162, une  
664 régulation MalR-dépendante similaire à celle décrite chez *E. faecalis* peut alors être proposée  
665 (**Figure 46**).



666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

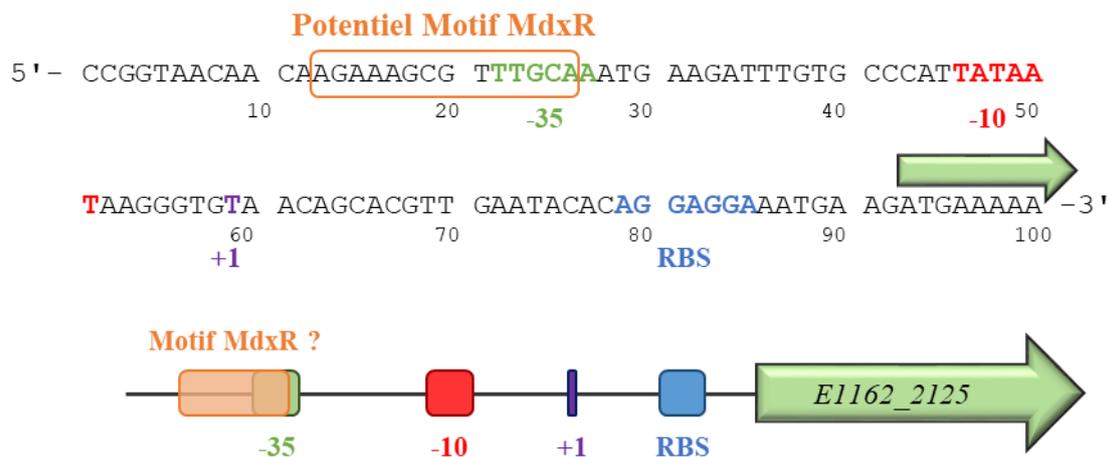
693

**Figure 46 :** Homologies des protéines codées par les opérons maltose chez *E. faecalis* et *E. faecium*

Les homologies des protéines impliquées dans le métabolisme du maltose chez *E. faecalis* OGIRF (en haut) et *E. faecium* E1162 (en bas). Les gènes homologues entre les deux espèces sont reliés par des traits de couleur. L'identité (=), la similarité (+) ainsi que le pourcentage de couverture [xx%] lors de la recherche d'homologies sont indiqués pour chaque comparaison.

En revanche, aucune information n'est disponible concernant le potentiel rôle de MalR dans la régulation des gènes *pulA-mdxDCBA* chez *E. faecium* E1162. Les gènes *pulA* et *mdxDCBA* sont tous disposés dans la même orientation, impliqués dans une même voie métabolique et aucun terminateur rho-indépendant n'a pu être prédit entre ces gènes. Toutes ces données laissent supposer une organisation en opéron avec une transcription débutant en amont du gène *pulA*. Au vu de la régulation par MdxR décrite dans la littérature (237), un motif de reconnaissance des régulateurs transcriptionnels LacI/GalR a été recherché en amont du gène *pulA*. Comme MdxR est prédit pour appartenir à la sous-famille "CcpA-like" (cf partie III.A de l'introduction) des régulateurs LacI, un motif de reconnaissance centré en AANCGNTT a été recherché (140). Aucun motif répondant à ces critères n'a pu être prédit jusqu'à 600 pb en amont de l'ATG du gène *pulA*, aucun motif de fixation pour MdxR ne peut donc être proposé pour cette région.

Il est cependant intéressant de noter qu'un potentiel promoteur a été identifié *in silico* en amont du gène *mdxD*. Cela peut signifier que les gènes *pulA* et *mdxDCBA* ne sont peut-être pas présents au sein de la même structure opéronique ou qu'un promoteur interne additionnel est présent en amont du gène *mdxD*. Un motif de fixation pour un régulateur transcriptionnel de la famille LacI a pu être prédit eu sein de ce promoteur putatif (*PmdxD*). Ce motif chevauche une partie du potentiel élément -35 et pourrait par conséquent être compatible avec un effet activateur assuré par un régulateur LacI (Figure 47). La séquence (TGCAAACGCTTTCT) de ce motif apparaît trop dégénérée par rapport au consensus de fixation de MalR déterminé chez *E. faecalis* pour prédire une éventuelle implication directe de MalR. Il se pourrait alors que ce motif corresponde à la région de fixation de MdxR.



694  
695  
696  
697  
698  
699

**Figure 47** : Promoteur et motif mdxR prédit en amont du gène mdxD (E1162\_2125)

Les différents éléments prédits du promoteur ainsi qu'un potentiel motif MdxR sont mis en évidence par un code couleur. Tous ces éléments sont uniquement prédictifs et n'ont fait l'objet d'aucune vérification. Une représentation graphique est proposée en dessous de la séquence.

## 700 IV. Publication scientifique n°3

### 701 A. Résumé de la publication n°3

702 L'opéron *OGIRF\_10235-34* (initialement nommé *celBA*) induit en condition d'infection chez  
703 la souris était prédit pour être impliqué dans le métabolisme du cellobiose chez *E. faecalis* par  
704 homologie avec le système *cel* de *S. pneumoniae*. Il s'est avéré que cet opéron n'est en réalité pas  
705 requis chez *E. faecalis* pour métaboliser le cellobiose mais s'avère essentiel pour le métabolisme  
706 du gentiobiose. Cette structure bicistronique a par conséquent été renommée *genBA* pour mieux  
707 refléter son rôle dans la physiologie bactérienne. La protéine GenB codée par le premier gène de  
708 l'opéron contient un domaine EIIC d'un PTS de la famille  $\beta$ -glucoside. Ce transporteur semble  
709 présenter des affinités pour le gentiobiose permettant ainsi l'entrée de ce disaccharide dans la cellule  
710 sous sa forme phosphorylée (gentiobiose-6'-P). Le second gène de l'opéron *genA* code pour une  
711 hydrolase qui va cliver la liaison  $\beta$ -1,6 entre les deux résidus glycosyl au sein de ce disaccharide  
712 phosphorylé afin de former du glucose et du G6P. Cet opéron est régulé par un régulateur  
713 transcriptionnel de la famille RpiR codé par le gène *genR* (*OGIRF\_10236*). La protéine GenR va  
714 pour cela former un complexe avec le gentiobiose-6'-P importé par GenB et activer l'expression de  
715 l'opéron *genBA*. Cet opéron est également soumis à une répression catabolique exercée par le  
716 complexe CcpA/P-Ser-HPr.

717 Pour former un système PTS fonctionnel, la protéine GenB nécessite la présence de protéines  
718 EIIA et EIIB codées par des gènes présents ailleurs sur le chromosome. Des gènes candidats ont  
719 été identifiés pour permettre la formation d'un PTS complet et fonctionnel (*OGIRF\_10745* et/ou  
720 *10750* pour la protéine EIIB et *OGIRF\_10751* pour la protéine EIIA).

721 L'opéron *genBA* a également été caractérisé comme requis pour le métabolisme de  
722 l'amygdaline. Cette molécule est un hétéroside, sa partie glucidique est le gentiobiose sur lequel est  
723 fixé un groupement mandelonitrile. Cette molécule est également capable d'induire l'expression de  
724 l'opéron *genBA*. La protéine GenR présente également des affinités *in vitro* pour des holosides  
725 composés d'un  $\beta$ -glucose lié à un groupement aglycone comme l'arbutine et la salicine (tous ces  
726 hétérosides sont présentés en **Figure S.1**). Cet opéron pourrait présenter ainsi une gamme de  
727 substrat plus large que le simple gentiobiose.

728 **B. Publication n°3 :** Characterization of the gen locus involved in  
729  $\beta$ -1,6-oligosaccharide utilization by *Enterococcus faecalis*

1        **Characterization of the *gen* locus involved in  $\beta$ -1,6-**  
2 **oligosaccharide utilization by *Enterococcus faecalis*.**

3 Maxime Grand<sup>1</sup>, Marion Aubourg<sup>2</sup>, Andreas Pikis<sup>3,4</sup>, John Thompson<sup>4</sup>, Josef Deutscher<sup>5,6</sup>, Axel  
4 Hartke<sup>1</sup>, Nicolas Sauvageot<sup>1\*</sup>

5 Running title: Gentiobiose degradation by *Enterococcus faecalis*

6 <sup>1</sup>*Normandie Université, UNICAEN, U2RM Stress/Virulence, F-14000 Caen, France.*

7 <sup>2</sup>*Normandie Université, UNICAEN, U2RM Antibio-résistance et Pathologies équine, F-14000*  
8 *Caen, France.*

9 <sup>3</sup>*Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring,*  
10 *Maryland 20993, USA*

11 <sup>4</sup>*Microbial Biochemistry and Genetics Unit, Laboratory of Cell and Developmental Biology,*  
12 *NIDCR, National Institutes of Health, Bethesda, MD 20892, USA.*

13 <sup>5</sup>*Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France.*

14 <sup>6</sup>*UMR 8261, CNRS, Université de Paris, Institut de Biologie Physico-Chimique, 75005 Paris,*  
15 *France.*

16 \*For correspondence. E-mail [nicolas.sauvageot@unicaen.fr](mailto:nicolas.sauvageot@unicaen.fr); Tel. (+33) 2 31 56 56 61; Fax: (+33)  
17 2 31 56 53 11.

18

19

## 20 **Abstract**

21 **The bicistronic *genBA* operon (formely named *celBA*) of the opportunistic pathogen**  
22 ***Enterococcus faecalis*, encodes a 6-phospho- $\beta$ -glucosidase (GenA) and a phosphotransferase**  
23 **system permease EIIC (GenB). It resembles the *cel* operon of *Streptococcus pyogenes*, which**  
24 **is implicated in the metabolism of the cellobiose. However, *genBA* mutants grew normally on**  
25 **cellobiose, but not (*genA*) or only slowly (*genB*) on gentiobiose and amygdalin. The two**  
26 **glucosides were also found to be the main inducers of the operon, confirming that the encoded**  
27 **proteins are involved in the utilization of  $\beta$ -1,6- rather than  $\beta$ -1,4-linked oligosaccharides.**  
28 **Expression of the *genBA* operon is regulated by the transcriptional activator GenR, which is**  
29 **encoded by the gene upstream from *genB*. Thermal shift analysis showed that it binds**  
30 **gentiobiose-6'-P with a Kd of 0.04 mM and with lower affinity also other phospho-sugars.**  
31 **The GenR/gentiobiose-6'-P complex binds to the promoter region upstream from *genB*. The**  
32 **presence of a *cre* box in the *genBA* promoter region and gel-shift experiments demonstrated**  
33 **that the operon is under negative control of the global carbon catabolite regulator CcpA. We**  
34 **also show that the orphan EIIC (GenB) protein uses the EIIA, components of the putative**  
35 **OG1RF\_10750-OG1RF\_10755 operon situated elsewhere on the chromosome to form a**  
36 **complete and functional PTS transporter.**

## 37 **Plain Language Summary**

38 Infections by enterococci, major health-care associated pathogens, are difficult to treat due to their  
39 increasing resistance to clinically relevant antibiotics and new treatment strategies are urgently  
40 needed. A largely unexplored aspect is how these pathogens proliferate and which substrates they  
41 use to grow inside infected hosts.

42

## 43 **Introduction**

44 *E. faecalis* is a commensal bacterium inhabiting the digestive tract of humans and many animals  
45 (Agudelo Higueta and Huycke, 2014). It has become a major health problem in humans because it  
46 can cause a wide variety of diseases such as infectious endocarditis, urinary tract infection,  
47 bacteremia, and septicemia (Wisplinghoff *et al.*, 2004; Diederich *et al.*, 2014; Ceci *et al.*, 2015).  
48 The emergence of enterococcal strains resistant to antibiotics observed in recent decades represents  
49 a serious public health issue. Mechanisms involved in the pathogenicity of these microorganisms  
50 are not well understood, although important work has been undertaken to identify virulence factors  
51 (Garsin *et al.*, 2014). Numerous studies have highlighted the link between metabolism and  
52 virulence in enterococci (Ruiz-Cruz *et al.*, 2015; Paganelli *et al.*, 2016; Zhang *et al.*, 2017). A  
53 recent transcriptomic study with *E. faecalis* strain V583 allowed the identification of 211 genes  
54 significantly induced during an intraperitoneal infection in mice (Muller *et al.*, 2015). Among them  
55 were several genes encoding  $\alpha$ - and  $\beta$ - glucosidases, suggesting that bacteria adapt their metabolism  
56 to nutrient conditions inside the host and that these activities are important for organ colonization,  
57 survival in the host and likely virulence. The glucosidases seemed to be implicated in the utilization  
58 and degradation of maltose and maltodextrin ( $\alpha$ -linked) and cellobiose ( $\beta$ -linked). The maltose and  
59 maltodextrin utilization pathways play a role in the colonization of mice organs by *E. faecalis*  
60 (Sauvageot *et al.*, 2017; Joyet *et al.*, 2017a), but no information is hitherto available for the  
61 degradation pathways of  $\beta$ -glucosides.

62 Among the genes most strongly induced during infection in mice, 2 genes annotated in the V583  
63 strain as *ef0291* and *ef0292* encode a 6-phospho- $\beta$ -glucosidase and a  
64 phosphoenolpyruvate:carbohydrate phosphotransferase system (PTS) EIIC, respectively. The  
65 corresponding proteins show strong homology to those encoded by the *cel* system present in  
66 streptococci and were therefore assumed to catalyze cellobiose utilization. Cellobiose is a  $\beta$ -1,4-  
67 linked diholoside produced from the degradation of cellulose, a linear form of  $\beta$ -glucoside  
68 polysaccharides presents in plants, but not found in mammalian tissues. Many bacteria, including  
69 streptococci, are able to catalyze the degradation of cellobiose. The *cel* genes required for  
70 cellobiose utilization by *Streptococcus mutans* and *Streptococcus pneumoniae* have been  
71 identified as virulence factors even if it seems to be strain-specific (Zeng and Burne, 2009; Shafeeq  
72 *et al.*, 2011). They include the gene *celA* coding for 6-phospho- $\beta$ -glucosidase, *celB*, *celC*, *celD*,  
73 which code for the components of a cellobiose-specific PTS, *celR*, which codes for a MtlR/LicR-  
74 like transcription activator, and one or two genes of unknown function. The permease CelD also  
75 catalyzes the transport of various  $\beta$ -oligoglucosides found in plants, such as gentiobiose, arbutin,  
76 amygdalin and aesculin (Bidossi *et al.*, 2012). In *S. mutans*, inactivation of the *cel* locus prevents

77 growth on cellobiose, suggesting that it is the essential operon for the utilization of this holoside  
78 (Zeng and Burne, 2009). However, another operon, named *bgu* is highly induced by cellobiose. It  
79 exhibits strong homology to the *cel* locus and comprises 5 genes organized in two transcriptional  
80 units. One operon encodes a glycosyl hydrolase (*bguA*) and three PTS components, EIIC, B and A  
81 (*bguDBC*). The second transcript encodes the GntR-like transcriptional regulator *BguR*, which acts  
82 as a repressor of the adjacent *bgu* operon (Shafeeq *et al.*, 2013).

83 Numerous other streptococcal genes have been linked to the utilization of cellobiose and other  $\beta$ -  
84 glucosides, including the *bgl*, *esc*, *bfd* and *gom* operons (Kiliç *et al.*, 2004; Terra *et al.*, 2015). They  
85 are involved in nutrient metabolism, biofilm formation, colonization of mammalian environments  
86 and virulence. This diversity and the complexity of pathways observed for the utilization of  $\beta$ -  
87 glucosides highlight the importance of this disaccharide for bacteria.

88 In this communication, we describe the characterization of the *gen* locus in *E. faecalis* induced  
89 during an intraperitoneal infection in mice. We show that the *genBA* operon is the main operon for  
90 the assimilation and degradation of gentiobiose or more generally involved in  $\beta$ -1,6 glucosides  
91 metabolism. The catabolite response regulator CcpA and the transcriptional activator GenR  
92 regulate the expression of *genBA* and gentiobiose-6'-phosphate seems to be the main inducer.

93

## 94 **Experimental Procedures**

### 95 **Bacterial strains and growth conditions**

96 *E. faecalis* wild-type strain OG1RF and mutants derived from it were routinely grown at 37°C  
97 without shaking in M17 medium containing 0.5% w/v glucose. Growth studies were carried out  
98 with carbon-depleted M17MOPS medium, which was prepared as previously described (Bizzini *et*  
99 *al.*, 2010). It was supplemented with 0.3% w/v of glucose, cellobiose, maltose, gentiobiose,  
100 amygdalin or salicin. Erythromycin was added when appropriate at a concentration of 50 µg mL<sup>-1</sup>.

101 *E. coli* strains M15[pREP4] (Qiagen) and Top10 (Thermo Fisher Scientific) used for protein  
102 purification and cloning experiments were grown aerobically with gyratory shaking at 250 rpm in  
103 LB medium at 37°C. Growth was followed by measuring the absorption at 600 nm in a Novaspec  
104 II spectrophotometer. Ampicillin, kanamycin and erythromycin were added when appropriate at  
105 concentrations of 100, 25 and 150 µg mL<sup>-1</sup>, respectively. Results depicted in the growth figures are  
106 representatives of experiments that were repeated at least three times and the maximal standard  
107 deviation calculated is less than 10 %.

108

### 109 **General molecular methods**

110 Restriction endonucleases, shrimp alkaline phosphatase and T4 DNA ligase were obtained from  
111 Promega and used according to the manufacturers' instructions. PCR experiments were performed  
112 with the thermocycler "Master Cycler Gradient" (Eppendorf) using GoTaq or Phusion DNA  
113 polymerase (Promega). The primers used for PCR experiments are listed in TABLE S1. If needed,  
114 PCR products were purified using the Nucleospin® Gel and PCR clean-up kit (Macherey-Nagel).  
115 *E. coli* was transformed by electroporation with the Gene Pulser Apparatus (Bio-Rad Laboratories)  
116 (Dower *et al.*, 1988) and *E. faecalis* and *L. lactis* as previously described (Holo and Nes, 1989).  
117 Plasmids were extracted from *E. coli* by using the Nucleospin® Plasmid kit (Macherey-Nagel).  
118 RNA extractions were performed as previously described (Sauvageot *et al.*, 2017). The 5' end of  
119 the mRNA corresponding to the *genBA* operon and *genR* gene (Fig. 1) were mapped on a 5'RACE  
120 (Rapid Amplification of cDNA Ends) PCR product obtained by using the 3'/5' RACE kit (Thermo  
121 Fisher Scientific) and primers *genB\_sp1* and *genR\_sp1* (TABLE S1) for the reverse transcriptase  
122 reaction and primer *genB\_sp2* and *genR\_sp2* for the PCR after poly A tailing. The PCR products  
123 were subsequently purified and sequenced with the primer *genB\_sp3* and *genR\_sp3* for *genB* and  
124 *genR*, respectively.

125 Total RNA was extracted by using the Direct-zol™ RNA MiniPrep Kit (Zymo research). Isolated  
126 RNA was spectrophotometrically quantified before 2 µg of the total RNA was used for reverse  
127 transcription with the QuantiTect® Reverse Transcription kit (Qiagen). Quantitative real-time PCR  
128 (RT-qPCR) was realized with SYBR Green fluorescence with a Cfx96 Real Time PCR Machine  
129 (Bio Rad). The primers used for the RT-qPCR experiments were 0291L and 0291R for *genA*,  
130 0292L and 0292R for *genB* and 0293L and 0293R for *genR* (TABLE S1). Raw data were converted  
131 into expression data by the absolute quantification method using the standard curve and then  
132 normalized using the housekeeping *gyr* gene as reference. Mean values and standard deviations  
133 were calculated from 3 independent replicates.

134

### 135 **Construction of *E. faecalis* *genA*, *genB*, *genR* and OG1RF\_10751 mutants and complemented** 136 **strains**

137 *E. faecalis* mutants were constructed by introducing two stop codons at the beginning of each ORF  
138 in order to produce truncated proteins. For each mutant, two PCR products of approximately 1000  
139 bp on both sides of the mutation site were amplified by using chromosomal DNA as template and  
140 the primer pairs pMAD-genX-for/Mut-genX-rev and pMAD-genX-rev/Mut-genX-for,  
141 respectively (X represents A, B or R, TABLE S1). The 5' ends of primers Mut-genX-for and Mut-  
142 genX-rev contain *Bgl*III restriction sites. After amplification, PCR products were digested with the  
143 corresponding enzyme and ligated. A second PCR was performed using the reaction mixture of the  
144 ligation as DNA template and primers pMAD-genX-for / pMAD-genX-rev. The resulting 2 kb  
145 fragment was cloned into plasmid pUBC300 (Frère *et al.*, 1993) digested with *Bam*HI and *Sal*I and  
146 the recombinant plasmid was used to transform *E. coli* Top10. It was purified and then transferred  
147 into *E. faecalis* OG1RF strain according to a previously described protocol (Mokhtari *et al.*, 2013).  
148 After electroporation, cells were plated on GM17 agar containing 50 µg/mL erythromycin and  
149 incubated at 37°C. The presence of the plasmid in clones resistant to the antibiotic was verified by  
150 PCR using oligonucleotides *verif\_X* and plasmid-based primers (PU or PR). To allow the second  
151 crossing over, integrants were grown for 150 generations on GM17 broth at 37°C without  
152 antibiotics. For each gene, several antibiotic-susceptible clones were isolated and the presence of  
153 the introduced *Bgl*III site and the stop codons were confirmed by *Bgl*III restriction and DNA  
154 sequencing, respectively. For the construction of the OG1RF\_10751 mutant strain, 91% of the  
155 central part of the gene was deleted using the same strategy as for the construction of the *gen*  
156 mutants with primers pMAD1017\_for/Mut\_1018\_rev and EF1019\_rev/Mut\_1018\_for. The  
157 deletion was realized in frame between the start and the end of the gene in order to avoid polar  
158 effects on the expression of the downstream genes.

159 Mutants were functionally complemented by an orthologous knock-in strategy. Therefore, the  
160 mutated copy in strain OG1RF was replaced with the intact gene of *E. faecalis* strain V583. The  
161 amino acid sequences of the orthologous genes are entirely identical but the nucleotide sequences  
162 differ between strain OG1RF and V583 and were used as a marker for correct complementation.  
163 The V583 genes with adjacent regions were cloned into the plasmid pUCB300 and the allelic  
164 exchange was realized as described above.

165

### 166 **Synthesis of phosphorylated sugars**

167  $\beta$ -Glycosidic sugars were phosphorylated with the  $\beta$ -glucoside kinase (BglK) from *Klebsiella*  
168 *pneumoniae* (Thompson *et al.*, 2002) and  $\alpha$ -linked maltose was prepared using the  $\alpha$ -glucoside-  
169 specific PTS present in palatinose-grown cells of *K. pneumoniae* (Thompson *et al.*, 2001).  
170 Gentiobiose, arbutin and salicin were phosphorylated at the O-6 position of the glucose residue and  
171 cellobiose and maltose at the same position of the non-reducing moiety of the disaccharides.

172

### 173 **Purification of His-tagged *E. faecalis* proteins**

174 In order to purify GenR and GenA with an N-terminal His-tag, we first amplified the corresponding  
175 genes by PCR using *E. faecalis* DNA as template and appropriate primer pairs creating restriction  
176 sites at the ends of the amplified DNA (TABLE S1). We subsequently digested the amplified DNA  
177 fragments with appropriate restriction enzymes and cloned them into the His-tag expression vector  
178 pQE30 (Qiagen) restricted with the same enzymes. After verification of the correct sequence of the  
179 amplified genes, the resulting plasmids were used to transform *E. coli* strain M15[pREP4]. A  
180 transformant for each amplified gene was subsequently grown in 300 mL of Terrific medium  
181 supplemented with ampicillin until the medium reached an OD600 of about 0.5. Expression of the  
182 different *E. faecalis* genes was subsequently induced by growing the cells for 3 h at 30°C in the  
183 presence of 0.1 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG). The cells were harvested by  
184 centrifugation. The preparation of crude extracts and the purification of His-tagged GenR and  
185 GenA on Ni-NTA columns under non-denaturing conditions were carried out by following the  
186 protocol described previously (Joyet *et al.*, 2017a). After purification the identity of the protein  
187 was verified by mass spectrometry (Sauvageot *et al.*, 2002).

188 His-CcpA, His-HPr and His-HprK were purified as described previously (Galinier *et al.*, 1997).

189 HprK was used for phosphorylation of HPr on the seryl residue (Monedero *et al.*, 2001).

190

## 191 **Protein-Ligand interactions**

192 The interaction between GenR and different (phospho-)sugars was studied with Thermal Shift  
193 Assays (TSA). Sypro Orange (Sigma) was used as dye and the fluorescence was monitored while  
194 the solution was stepwise heated by 1°C from 25°C to 60°C. Fluorescence data were collected on  
195 a Cfx96 Real Time PCR Machine (Bio Rad) adjusted to HEX<sup>TM</sup> as fluorescent dye. The temperature  
196 of hydrophobic exposure (*Th*), at which half of the protein molecules are unfolded, was calculated  
197 with the formula  $-d(\text{RFU})/dT$  (RFU = relative fluorescence units). The *Th* of the control (without  
198 ligand) was subtracted from the *Th* measured for each condition in order to calculate the binding  
199 constants. GenR was used at a concentration of 5 μM and the concentration range of the tested  
200 sugars was between 0 and 10 mM. The assays were carried out in 50mM Tris/HCl buffer, pH7.4,  
201 containing 15% glycerol.

202

## 203 **Hydrolysis of phospho-sugars and separation of the products by thin layer chromatography**

204 To detect whether GenA was able to hydrolyze phosphorylated sugars, 5 μl of a 1 μg/μL substrate  
205 solution in 50 mM Tris/HCl, pH 7.4, was mixed with 5 or 20 μg of GenA. The assay mixture,  
206 which also contained 1 mM CaCl<sub>2</sub>, was incubated for 3 h at 37°C. Aliquots of 5 or 10 μl were  
207 spotted on thin layer silica gel 60 TLC plates (Merck) and the products formed from the various  
208 phospho-sugars were subsequently separated as previously described (Joyet *et al.*, 2017b).

209

## 210 **Electrophoretic Mobility Shift Assays**

211 A DNA fragment of 234 bp corresponding to the *gen* operon promoter region was PCR  
212 amplified in the presence of 2 mCi [ $\alpha$ -<sup>32</sup>P]dATP with primers Promo\_genB\_for and  
213 Promo\_genB\_rev (TABLE S1) and subsequently purified using NucleoSpin columns (Macherey-  
214 Nagel). Purified His-GenR protein (500 nM - 4 nM final concentration) was added to 10 μL of 2x  
215 dilution buffer, composed of 20 mM Tris-HCl pH 7.5, 2 mM DTT, 2mM EDTA, 100 mM KCl, 40  
216 mM MgCl<sub>2</sub>, 10% glycerol, 0.2 mg/mL BSA, 2 mM of ligand (sugars or phospho-sugars) and 2  
217 mg/mL herring sperm as nonspecific competitor DNA. The mixture was incubated for 15 min at  
218 room temperature and samples were separated on 8% acrylamide/bis-acrylamide non-denaturing  
219 gels, in 0.5X Tris borate-EDTA buffer, with a pH 7.5 (TBE). Electrophoresis was performed during

220 3 hours. Gels were dried and analyzed by autoradiography using a Cyclone storage phosphor screen  
221 (Perkin Elmer). A 204 pb internal fragment of the *gusA* gene was used as negative control.

222

### 223 **DNase I footprint Analysis**

224 DNase I footprinting was performed according to Zianni *et al.* (Zianni *et al.*, 2006). The promotor  
225 region of the *genBA* operon was amplified with primers Promo\_genB\_for and Promo\_genB\_rev  
226 and then cloned into the pTOPO Zero Blunt plasmid (Thermofisher). The DNA fragment used for  
227 footprinting was generated by PCR with the reverse primer labelled with the FAM fluorophore and  
228 Promo\_genB\_for. First, 500 nM of His-GenR were incubated in binding buffer (10 mM Tris pH7.5,  
229 1 mM DTT, 5 mM MgCl<sub>2</sub>, 2 mM CaCl<sub>2</sub>, 0.125 mg/mL bovine serum albumin (BSA), 62.5 µg/mL  
230 herring sperm, 0.25 mM gentiobiose-6'-P and 5% glycerol). Subsequently, 500 ng of fluorescently  
231 labeled probe was added to the reaction mixture to a final volume of 80 µL and incubated for 20  
232 min at room temperature. Finally, various amounts (0.00041 to 0.016 units) of RQ1-DNase I  
233 (Promega) were added and the sample was incubated for another 5 min at room temperature. The  
234 reaction was stopped by adding 5 µL of 20 mM EGTA (pH 8.0) and heating for 10 minutes at  
235 65°C. The DNA fragments were purified with the NucleoSpin columns kit (Macherey Nagel) and  
236 eluted in 50 µL H<sub>2</sub>O to eliminate salts that can interfere with capillary electrophoresis. 10 µL of  
237 digested DNA were added to 9.9 µL formamide and 0.1 µL of GeneScan-500 rox size standards  
238 (Thermo Fisher Scientific). DNA fragment analysis was performed by the sequencing platform of  
239 the Cochin Institut / Eurofins (Paris).

240

### 241 **RNA sequencing**

242 The RNA sequencing experiments were performed at the iGE3 genomics platform of the University  
243 of Geneva. A total of 1 µg of total RNA was ribo depleted using the Ribo Zero Magnetic Kit for  
244 bacteria from Epicentre (Illumina, Wisconsin, USA). Libraries were then prepared using the  
245 Illumina TruSeq stranded mRNA kit according to manufacturer's recommendations. Libraries  
246 were validated on the Bioanalyzer 2100 (Agilent Technologies, Santa Clara, USA) and the Qubit  
247 fluorimeter (Thermo Scientific). Samples were multiplexed by 8 and loaded at 8 pmoles on one  
248 lane of an Illumina HiSeq 2500 according to the single read, 50 cycle protocol.

249

### 250 **Bioinformatic analysis**

251 The sequencing quality control was done with FastQC (Wang *et al.*, 2012). The reads were mapped  
252 with the TopHat v.2 software to the *Enterococcus faecalis* OG1RF reference (Bourgogne *et al.*,  
253 2008). The table of counts with the number of reads mapping to each gene feature of the Aus0004

254 reference genome was prepared with HTSeq v0.6p1 (htseq count). The differential expression  
255 analysis was performed with the R/Bioconductor statistical analysis package EdgeR v. 3.4.23  
256 (Robinson *et al.*, 2010). The counts were normalized according to the library size and filtered. The  
257 genes with a count above 1 count per million reads (cpm) in at least 2 samples were kept for further  
258 analysis. The differentially expressed gene tests were statistically assessed with the exact test in  
259 EdgeR. The differentially expressed gene p-values were corrected for multiple testing error with a  
260 5% FDR (false discovery rate) by using the Benjamini Hochberg (BH) procedure.

261

## 262 **Results**

### 263 **Organization of the *gen* operon in *E. faecalis* OG1RF**

264 A transcriptome study had revealed that many enterococcal genes, including *ef0291-ef0292*, are  
265 strongly induced during *in vivo* intraperitoneal infection of mice with *E. faecalis* strain V583. We  
266 became interested in these genes and the upstream gene *ef0293*, because they presumably encode  
267 a 6-phospho- $\beta$ -glucosidase, a cellobiose-specific EIIC PTS component, and a MurR/RipR-like  
268 transcription regulator. To study the physiological roles of these genes, we used the *E. faecalis*  
269 strain OG1RF which is easier to manipulate than the V583 clinical isolate. The corresponding genes  
270 in this strain are *OG1RF\_10234*, *OG1RF\_10235* and *OG1RF\_10236*, which we renamed *genA*,  
271 *genB* and *genR*, respectively (Fig. 1A). RT-PCR experiments revealed that these three genes are  
272 transcribed from two different promoters located upstream from *genB* and *genR* (P1 and P2).  
273 Transcriptional start sites were mapped by 5'-RACE PCR with RNA extracted from cells harvested  
274 during exponential growth on cellobiose as the substrate (Fig. 1B). A transcription terminator is  
275 located just downstream from *genR* (Fig. 1A), suggesting that the *genBA* operon and the *genR* gene  
276 belong to two different transcription units. This was confirmed by RT-PCR with RNAs isolated  
277 from cells grown on cellobiose (FIG. S1).

278

### 279 **Construction of *gen* mutants and gene expression analysis.**

280 To carry out a functional analysis of the *gen* genes, markerless mutants were constructed by  
281 introduction of stop codons in the first part of each gene. These mutants were tested for their ability  
282 to grow in M17cc medium complemented with 0.3% cellobiose (Fig. 2A). Growth of all mutants  
283 on this sugar was comparable to the wild-type strain suggesting either that these genes are not  
284 essential for cellobiose metabolism or that another efficient *cel* PTS and 6-phospho- $\beta$ -glucosidase  
285 are functional in *E. faecalis*. Rapid screening with a ready-to-use medium, which allows the

286 fermentation of 49 carbohydrates (API 50 CH, Biomerieux), showed that the three mutants ferment  
287 gentiobiose and amygdalin significantly slower than OG1RF. Gentiobiose is a disaccharide  
288 composed of two units of D-glucose joined by a  $\beta$ -1,6 linkage. Amygdalin, which is present in  
289 some fruit seeds, is a gentiobiose derivative carrying the cyanogenic mandelonitrile group bound  
290 to the C1 (FIG. S2). Such compounds have been reported as present in the gut because not  
291 hydrolysed in the stomach or small intestine (Crittenden and Playne, 1996). The slowed  
292 fermentation of these disaccharides was confirmed by growth studies (Fig. 2B and C). We observed  
293 that the wild-type strain grew efficiently on gentiobiose and amygdalin. In contrast, the three  
294 mutants were not able to grow during the first 10 hours although slow growth of the 3 mutants was  
295 observed thereafter (Fig. 2B), which might be due to activation of an alternative system able to  
296 compensate the mutations. These results suggest that EF0291 and EF0292 do not catalyze  
297 cellobiose but gentiobiose transport and catabolism. Despite their annotation as *celA* and *celB*  
298 in the V583 genome sequence, we renamed the corresponding genes *genA* and *genB*. *ef0293*, the  
299 putative transcriptional regulator of the *genBA* operon, was named *genR*. Concerning amygdalin, a  
300 slightly different growth behavior was observed. The  $\Delta$ *genA* mutant did not grow on this sugar. In  
301 contrast, the  $\Delta$ *genB* and  $\Delta$ *genR* mutants did grow on amygdalin although more slowly compared to  
302 the wild type strain (Fig 2C). These results indicate that amygdalin, although structurally similar  
303 to gentiobiose, is taken up not only by GenB, but also by another transporter. However, the only  
304 *E. faecalis* enzyme able to hydrolyze this disaccharide is GenA. Growth is restored in all  
305 complemented strains (Fig. 2). On gentiobiose, the three complemented strains grew as the wild-  
306 type strain (Fig. 2B). On amygdalin, the *genR* and *genB* complemented strains grew as the control  
307 whereas the complemented *genA* strain grows more slowly than the wild-type (Fig. 2C).

308

### 309 **The *genBA* operon is induced by both gentiobiose and amygdalin**

310 The expression of the three genes was assayed by RT-qPCR with cells cultivated on M17cc + 0.3%  
311 glucose, cellobiose, gentiobiose or amygdalin. When grown on glucose alone, expression of *genA*  
312 and *genB* was very low. Growth on gentiobiose and amygdalin caused strong expression of *genA*  
313 and *genB*, which was between 270- and 640-fold stronger than in glucose-grown cells (Fig. 3, A  
314 and B). The *genA* and *genB* genes are also induced by cellobiose, although to a lesser extent (80-  
315 fold increased expression compared to growth on glucose). These results confirmed that the *gen*  
316 operon is involved in the metabolism of  $\beta$ -1,6-linked disaccharides, but a contribution to the  
317 metabolism of cellobiose cannot be excluded although the mutants showed no growth defect on  
318 this sugar (Fig. 3A and B). Compared to the *genBA* genes, expression of *genR* is much lower and

319 similar on all carbon sources used (Fig. 3C), suggesting that, unlike the *genBA* operon, *genR* is a  
320 barely expressed constitutive transcription unit.

321

### 322 **GenR is a transcriptional activator of the *gen* operon**

323 Even if GenR protein exhibits no homology with the streptococcal CelR regulator, the localization  
324 of *genR* just upstream from the *genBA* operon and its sequence similarity with phosphosugar-  
325 binding transcriptional regulators strongly suggested that GenR is the regulator of the adjacent  
326 operon. The failure of the *genR* mutant to grow on gentiobiose (Fig. 2B) suggested that the GenR  
327 protein functions as a positive regulator of the *genBA* operon. Indeed, expression of *genA* and *genB*  
328 in cellobiose-grown *genR* mutant cells was even lower than in glucose-grown cells (Fig. 4A and  
329 B), whereas *genR* expression was similar under both conditions (Fig. 4C). To confirm that GenR  
330 directly regulates the *genBA* operon, it was overproduced in *E. coli* and purified. Since GenR has  
331 a sugar isomerase (SIS) domain, its capacity to bind different phosphorylated sugars was first tested  
332 using Thermal Shift Assays (TSA). While cellobiose and gentiobiose had no effect on the stability  
333 of the protein, important shifts of the denaturation spectra were observed with phosphorylated  
334 sugars (Fig. 5). Gentiobiose-6'-P had the highest affinity with a binding constant of 0.04 mM,  
335 further supporting the assumption that GenA and GenB mainly transport and catabolize sugars with  
336  $\beta$ -1,6 linkages. Interestingly, arbutin-6-P and salicin-6-P, two glucosides which carry phenyl  
337 derivatives bound to the C-1 via a  $\beta$ -linkage, also exhibited significant affinity to GenR (binding  
338 constants of 0.1 and 0.38 mM, respectively). In contrast, cellobiose-6-P with its  $\beta$ -1,4 linkage binds  
339 to GenR with much lower affinity (1.7 mM), a result in agreement with the lower induction of the  
340 *gen* operon by cellobiose compared to gentiobiose.

341 We subsequently studied binding of GenR to a DNA fragment containing the promoter region of  
342 the *gen* operon by carrying out EMSA. Because GenR requires a phosphorylated sugar as co-  
343 activator, interaction assays were performed in the presence of gentiobiose-6-P as inducer and  
344 unphosphorylated gentiobiose as a negative control. Based on the results obtained with the TSA, a  
345 final concentration of 1 mM sugar or P-sugar was used. In the presence of unphosphorylated  
346 gentiobiose, GenR was able to shift part of the DNA fragment when high protein concentrations  
347 were used (Fig. 6A, lanes 6 and 7). When gentobiose-6-P was added, the same amount of GenR  
348 was able to shift the entire DNA fragment (Fig. 6, lanes 2 and 3).

349 TSA had shown that GenR can bind several P-sugars. We therefore tested the effect of other P-  
350 sugars on GenR activation. In order to avoid non-specific binding, the GenR concentration was  
351 lowered to 100 nM. In addition to gentiobiose-6'-P (Fig. 6B, lane 2), arbutin-6-P significantly

352 increased the binding capacity of GenR (Fig. 6B, lane 8), a result in agreement with the TSA results.  
353 In the presence of several other P-sugars (salicin-6-P, maltose-6'-P), thin bands shifted band were  
354 observed suggesting a weak binding. However compared to the amount of free DNA it indicates  
355 that GenR did not significantly shift the DNA fragment containing the *genBA* promoter in the  
356 conditions tested.

357 To map the GenR binding site on the *genBA* promoter region, DNaseI footprinting experiments  
358 were performed (Fig. 1A). GenR with bound gentiobiose-6'-P protected a 20 bp region (5'-  
359 TCTTGGTACATGTTTCTCAT-3') extending from positions -48 to -68 relative to the *genB*  
360 transcriptional start site. This protected region is located just downstream the *genR* terminator and  
361 upstream the -10 sequence of *genB*. This position is in agreement with the transcriptional activator  
362 function of GenR.

363

#### 364 **The CcpA/Ser-phosphorylated HPr complex represses *genBA* expression.**

365 To determine whether the *gen* genes are submitted to catabolite repression (CR), qRT experiments  
366 were carried out with RNA isolated from cells grown in media containing glucose and one of the  
367 three oligosaccharides (Fig 3). Compared to cells grown only on gentiobiose or amygdalin the  
368 additional presence of glucose significantly repressed the expression of *genA* and *genB* (Fig. 3A  
369 and B), whereas *genR* was only slightly affected (Fig. 3C). With respect of error bars, the repressive  
370 effect of glucose was not obvious in combination with cellobiose. These results are in agreement  
371 with the presence of a potential catabolite response element (*cre*) (5'-ATGTATCCGTTTACA-3')  
372 upstream from *genB* overlapping the transcriptional start site (C in bold case in Fig. 1B) (Miwa *et*  
373 *al.*, 2000). CcpA, the central regulator of carbon metabolism in Gram positive bacteria, interacts  
374 with the general PTS protein HPr phosphorylated at Ser-46 (P-Ser-HPr). The complex binds to *cre*  
375 sites and inhibits or stimulates the expression of the downstream genes. We therefore carried out  
376 EMSA with a 234 bp DNA fragment covering the intergenic region between *genB* and *genR* and  
377 various concentrations of CcpA in the presence or absence of P-Ser-HPr. As expected, P-Ser-HPr  
378 alone did not interact with the DNA fragment (Fig. 7, lane 2). A weak shifted band appeared with  
379 CcpA alone, which was strongest at the highest CcpA concentration (13.5 nM) (Fig. 7, lanes 7 to  
380 10). In the presence of both proteins nearly all DNA was shifted with two bands appearing at low  
381 CcpA concentrations and three at 13.5 nM CcpA (Fig. 7, lanes 3 to 6). The binding is specific as  
382 no shift was observed with control DNA without a *cre* box. These results strongly suggest that CR  
383 of *genBA* is mediated by the P-Ser-HPr/CcpA complex.

384

385 **Comparison of the transcriptome of gentiobiose- and glucose-grown cells identifies the**  
386 **lacking EIIA<sup>Gen</sup>.**

387 Bacteria often possess several PTSs, which catalyze the transport and concomitant phosphorylation  
388 of their carbohydrate substrates (Deutscher *et al.*, 2006). The last proteins in this cascade are the  
389 EII complexes, which consist of an integral membrane domain (EIIC) and cytoplasmic domains or  
390 proteins (EIIA and EIIB). GenB is homologous to EIIC proteins but the genes encoding EIIA and  
391 EIIB are absent from the *gen* operon and in its neighboring chromosomal region. To catalyze the  
392 observed transport of gentiobiose, GenB must cooperate with EIIA and EIIB proteins encoded by  
393 genes located elsewhere on the chromosome. To identify these two EII proteins we performed a  
394 transcriptome analysis of bacteria grown on gentiobiose by RNA-sequencing. We identified a total  
395 of 140 genes with significantly higher transcript levels and 34 repressed genes compared to glucose  
396 condition (FIG S2 and S3). As expected, the two predominantly induced genes are *genA* and *genB*.  
397 The majority of the other induced genes is involved in metabolism and encodes many PTS, ABC  
398 transporters and glycosyl hydrolases. The increased expression relative to cultures grown on  
399 glucose of these metabolic genes is probably due to the relief of catabolite repression in cells grown  
400 on gentiobiose. Many EIIA or EIIB encoding genes were shown to be induced on gentiobiose  
401 relative to glucose (OG1RF11235 and OG1RF12262 induced 26.74x and 16.87 x in gentiobiose,  
402 respectively) but most of them do not belong to the same PTS family than GenB. Only the genes  
403 encoding EIIA and EIIB present in a region of the genome spanning from the gene OG1RF\_10745  
404 to 10755 belong to the cellobiose PTS family (Fig. 8A). These PTS components were suspected to  
405 be the partners of GenB.

406 The genomic region contains 11 genes among which 5 genes encode two EIICs, two EIIBs and one  
407 EIIA belonging to the cellobiose family (Fig. 8A). Of note, the amino acid sequences of the two  
408 EIIB<sup>Cel</sup> are 100% identical but the enzyme encoded by OG1RF\_10750 is three amino acids longer.  
409 Because of this redundancy, we decided to inactivate the only EIIA<sup>Cel</sup> encoding gene of this region  
410 (OG1RF\_10751) and the resulting mutant was then tested for growth on gentiobiose, amygdalin as  
411 well as cellobiose (Fig. 8B). Growth of the mutant on glucose is similar to that of the parental  
412 strain. However, the mutant does not grow on gentiobiose and amygdalin as well as cellobiose.  
413 This strongly suggests that the EIIA<sup>Cel</sup> encoded by OG1RF\_10751 is the functional partner of GenB  
414 for the transport of the  $\beta$ -1,6-oligosaccharides but also that this PTS component is shared with  
415 transporter involved in cellobiose metabolism. Furthermore we speculate that one of the two  
416 EIIB<sup>Cel</sup> or both encoded in this region complete the PTS<sup>Gen</sup>. The identified EIIA enzyme is likely  
417 common to several PTS complexes because the mutant also shows a growth defect on cellobiose  
418 contrary to the *genB* mutant. The construction of the EIIA<sup>Cel</sup> complemented strain was realized and  
419 the growth was restored for all substrates for which growth defect had been observed (Fig. 8B).

420

421 **GenA exhibits 6-phospho- $\beta$ -glucosidase activity.**

422 Sugars transported by a PTS arrive as phospho-compounds inside the cell. 6-P-Glucosides formed  
423 during PTS uptake are usually hydrolyzed by a 6-phospho-glucosidase to glucose-6-P and the  
424 remaining part of the glucoside. Indeed, GenA exhibits significant similarity to 6-phospho- $\beta$ -  
425 glucosidases. To confirm its presumed activity, purified GenA was incubated with gentiobiose-6'-  
426 P, cellobiose-6'-P and maltose-6'-P as well as the unphosphorylated disaccharides and the resulting  
427 products were analyzed by thin layer chromatography (Fig. 9A). No hydrolysis was observed with  
428 the unphosphorylated sugars. A spot migrating to the position of glucose appeared with  
429 gentiobiose-6'-P as the substrate demonstrating that GenA hydrolyzes  $\beta$ -1,6 linkages. However, a  
430 very thin band corresponding to glucose was also observed with cellobiose-6'-P suggesting that  
431 GenA can also slowly hydrolyze  $\beta$ -1,4 linkages. The weak  $\beta$ -1,4 hydrolase activity of GenA was  
432 confirmed by using 2-times more protein and substrate (Fig. 9B), which increased the intensity of  
433 the glucose spot. These results do not allow excluding completely the implication of the products  
434 of the operon *gen* of the degradation of cellobiose. Indeed, this operon is slightly induced by  
435 cellobiose and GenA is capable of degrading cellobiose-6'-P. Although it is not the main sugar, it  
436 is difficult to determine the role of these proteins in the cellobiose pathway. No degradation was  
437 observed with maltose-6'-P suggesting that GenA is specific for  $\beta$ -glycosides.

438

439

440 **Conclusion**

441 *E. faecalis* is an important health-care associated pathogen. After infection, it has to adapt its  
442 metabolism to the host environment. A metabolic pathway, including genes annotated as *genA* and  
443 *genB*, was induced *in vivo* and allowed the cell to degrade gentiobiose (Muller *et al.*, 2015).  
444 Utilization of this sugar has been reported only in few microorganisms such as *Lactobacilli* or  
445 *Bifidiobacteria* (34–36). Gentiobiose and cellobiose are unlikely to be present in the host, which  
446 means that it is probably not the nature of the sugar that is of interest but the type of linking ( $\alpha$  or  
447  $\beta$ ) between the glucose residues. *E. faecalis* might induce these glucosidases to hydrolyze complex  
448 sugars. This was confirmed by RNA-Seq analysis where numerous glycosyl hydrolyzes were  
449 expressed in parallel with proteins exhibiting Endo-N acetylgluco/galactosaminidase or lyase  
450 activities. Within the host, two possible carbon sources may be suspected, glycoaminoglycans  
451 (GAG) and glycosylated eukaryotic proteins. It has been shown that *S. pneumoniae* express a set  
452 of glucosidases involved in the degradation of these compounds and thus are important for

453 virulence (King, 2010). GAGs are a major component of the extracellular matrix composed of linear  
454 polysaccharides with various repeating disaccharide units. We tried to cultivate *E. faecalis* on some  
455 of these GAGs (hyaluronic acid, chondroitin sulfate and heparin) but no growth was observed on  
456 these substrates (data not shown). It is therefore likely that these enterococcal enzymes are used for  
457 the degradation of the glycosylation part of host proteins. This hypothesis is supported by recent  
458 studies. Indeed, it has been shown that this microorganism is able to degrade glycoproteins and N-  
459 glycans via the secretion of the endo- $\beta$ -N-acetylglucosaminidase EndoE (38–41).

460 Furthermore, the gene *OGIRF\_10935* coding for a putative glycosylasparaginase has been  
461 identified in the *E. faecalis* genome. This enzyme may catalyze the hydrolysis of the  $\beta$ -N-  
462 glycosidic bond between asparagine and N-acetylglucosamine of asparagine linked N-glycans  
463 (Guan *et al.*, 1996). The gene coding this protein is present in an operon structure with genes coding  
464 for a PTS permease suggesting the uptake and the degradation of alternative carbon sources from  
465 protein deglycosylation.

466 Another insight supporting this hypothesis, is that the gene *ef0361* encoding a chitinase, a putative  
467 N-linked glycosyl hydrolase, is also induced during infection (Muller *et al.*, 2015). The next step  
468 in the understanding of carbon sources and energy utilization within a host will be to determine the  
469 contribution of these enzymes to the infection process of the bacteria.

470

## 471 **Acknowledgement**

472 We are grateful to Agathe Vigier of the Eurofins company for her technical assistance in the  
473 development and determination of the GenR binding site by footprint. We also thank Isabelle Rincé  
474 for technical assistance and the iGE3 genomics platform of the University of Geneva for RNA  
475 sequencing experiments ([http://www.ige3.unige.ch/genomics\\_platform.php](http://www.ige3.unige.ch/genomics_platform.php)).

476

## 477 **References**

- 478 Agudelo Higueta, N.I., and Huycke, M.M. (2014) Enterococcal Disease, Epidemiology, and  
479 Implications for Treatment. In *Enterococci: From Commensals to Leading Causes of Drug*  
480 *Resistant Infection*. Gilmore, M.S., Clewell, D.B., Ike, Y., and Shankar, N. (eds). Massachusetts  
481 Eye and Ear Infirmary, Boston. <http://www.ncbi.nlm.nih.gov/books/NBK190429/>. Accessed  
482 September 28, 2018.
- 483 Andersen, J.M., Barrangou, R., Abou Hachem, M., Lahtinen, S.J., Goh, Y.J., Svensson, B., and  
484 Klaenhammer, T.R. (2013) Transcriptional analysis of oligosaccharide utilization by  
485 *Bifidobacterium lactis* BI-04. *BMC Genomics* **14**: 312.
- 486 Bidossi, A., Mulas, L., Decorosi, F., Colomba, L., Ricci, S., Pozzi, G., *et al.* (2012) A functional  
487 genomics approach to establish the complement of carbohydrate transporters in *Streptococcus*  
488 *pneumoniae*. *PLoS One* **7**: e33320.
- 489 Bizzini, A., Zhao, C., Budin-Verneuil, A., Sauvageot, N., Giard, J.-C., Auffray, Y., and Hartke, A.  
490 (2010) Glycerol is metabolized in a complex and strain-dependent manner in *Enterococcus*  
491 *faecalis*. *J Bacteriol* **192**: 779–785.
- 492 Bøhle, L.A., Mathiesen, G., Vaaje-Kolstad, G., and Eijsink, V.G.H. (2011) An endo- $\beta$ -N-  
493 acetylglucosaminidase from *Enterococcus faecalis* V583 responsible for the hydrolysis of high-  
494 mannose and hybrid-type N-linked glycans. *FEMS Microbiol Lett* **325**: 123–129.
- 495 Bourgogne, A., Garsin, D.A., Qin, X., Singh, K.V., Sillanpaa, J., Yerrapragada, S., *et al.* (2008) Large  
496 scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF.  
497 *Genome Biol* **9**: R110.
- 498 Ceci, M., Delpech, G., Sparo, M., Mezzina, V., Sánchez Bruni, S., and Baldaccini, B. (2015) Clinical  
499 and microbiological features of bacteremia caused by *Enterococcus faecalis*. *J Infect Dev Ctries*  
500 **9**: 1195–1203.
- 501 Collin, M., and Fischetti, V.A. (2004) A novel secreted endoglycosidase from *Enterococcus*  
502 *faecalis* with activity on human immunoglobulin G and ribonuclease B. *J Biol Chem* **279**: 22558–  
503 22570.
- 504 Crittenden, R.G., and Playne, M.J. (1996) Production, properties and applications of food-grade  
505 oligosaccharides. *Trends Food Sci Technol* **7**: 353–361.
- 506 Daughtry, K.V., Johanningsmeier, S.D., Sanozky-Dawes, R., Klaenhammer, T.R., and Barrangou,  
507 R. (2018) Phenotypic and genotypic diversity of *Lactobacillus buchneri* strains isolated from  
508 spoiled, fermented cucumber. *Int J Food Microbiol* **280**: 46–56.
- 509 Deutscher, J., Francke, C., and Postma, P.W. (2006) How phosphotransferase system-related  
510 protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev*  
511 *MMBR* **70**: 939–1031.
- 512 Diederich, A.-K., Wobser, D., Spiess, M., Sava, I.G., Huebner, J., and Sakıncı, T. (2014) Role of  
513 Glycolipids in the Pathogenesis of *Enterococcus faecalis* Urinary Tract Infection. *PLOS ONE* **9**:  
514 e96295.

- 515 Dower, W.J., Miller, J.F., and Ragsdale, C.W. (1988) High efficiency transformation of *E. coli* by  
516 high voltage electroporation. *Nucleic Acids Res* **16**: 6127–6145.
- 517 Frère, J., Benachour, A., Novel, M., and Novel, G. (1993) Identification of the theta-type minimal  
518 replicon of the *Lactococcus lactis* subsp. *lactis* CNRZ270 lactose protease plasmid pUCL22. *Curr*  
519 *Microbiol* **27**: 97–102.
- 520 Galinier, A., Haiech, J., Kilhoffer, M.-C., Jaquinod, M., Stülke, J., Deutscher, J., and Martin-  
521 Verstraete, I. (1997) The *Bacillus subtilis* crh gene encodes a HPr-like protein involved in carbon  
522 catabolite repression. *Proc Natl Acad Sci* **94**: 8439–8444.
- 523 Garbe, J., Sjögren, J., Cosgrave, E.F.J., Struwe, W.B., Bober, M., Olin, A.I., *et al.* (2014) EndoE  
524 from *Enterococcus faecalis* hydrolyzes the glycans of the biofilm inhibiting protein lactoferrin  
525 and mediates growth. *PLoS One* **9**: e91035.
- 526 Garsin, D.A., Frank, K.L., Silanpää, J., Ausubel, F.M., Hartke, A., Shankar, N., and Murray, B.E.  
527 (2014) Pathogenesis and Models of Enterococcal Infection. In *Enterococci: From Commensals to*  
528 *Leading Causes of Drug Resistant Infection*. Gilmore, M.S., Clewell, D.B., Ike, Y., and Shankar, N.  
529 (eds). Massachusetts Eye and Ear Infirmary, Boston.  
530 <http://www.ncbi.nlm.nih.gov/books/NBK190426/>. Accessed November 7, 2018.
- 531 Guan, C., Cui, T., Rao, V., Liao, W., Benner, J., Lin, C.-L., and Comb, D. (1996) Activation of  
532 Glycosylasparaginase FORMATION OF ACTIVE N-TERMINAL THREONINE BY INTRAMOLECULAR  
533 AUTOPROTEOLYSIS. *J Biol Chem* **271**: 1732–1737.
- 534 Holo, H., and Nes, I.F. (1989) High-Frequency Transformation, by Electroporation, of *Lactococcus*  
535 *lactis* subsp. *cremoris* Grown with Glycine in Osmotically Stabilized Media. *Appl Environ*  
536 *Microbiol* **55**: 3119–3123.
- 537 Joyet, P., Mokhtari, A., Riboulet-Bisson, E., Blancato, V.S., Espariz, M., Magni, C., *et al.* (2017a)  
538 Enzymes Required for Maltodextrin Catabolism in *Enterococcus faecalis* Exhibit Novel Activities.  
539 *Appl Environ Microbiol* **83**.
- 540 Joyet, P., Mokhtari, A., Riboulet-Bisson, E., Blancato, V.S., Espariz, M., Magni, C., *et al.* (2017b)  
541 Enzymes Required for Maltodextrin Catabolism in *Enterococcus faecalis* Exhibit Novel Activities.  
542 *Appl Environ Microbiol* **83**.
- 543 Kiliç, A.O., Tao, L., Zhang, Y., Lei, Y., Khammanivong, A., and Herzberg, M.C. (2004) Involvement  
544 of *Streptococcus gordonii* beta-glucoside metabolism systems in adhesion, biofilm formation,  
545 and in vivo gene expression. *J Bacteriol* **186**: 4246–4253.
- 546 King, S.J. (2010) Pneumococcal modification of host sugars: a major contributor to colonization  
547 of the human airway? *Mol Oral Microbiol* **25**: 15–24.
- 548 Koutsioulis, D., Landry, D., and Guthrie, E.P. (2008) Novel endo- $\alpha$ -N-acetylgalactosaminidases  
549 with broader substrate specificity. *Glycobiology* **18**: 799–805.
- 550 Miwa, Y., Nakata, A., Ogiwara, A., Yamamoto, M., and Fujita, Y. (2000) Evaluation and  
551 characterization of catabolite-responsive elements (cre) of *Bacillus subtilis*. *Nucleic Acids Res* **28**:  
552 1206–1210.

- 553 Mokhtari, A., Blancato, V.S., Repizo, G.D., Henry, C., Pikis, A., Bourand, A., *et al.* (2013)  
554 *Enterococcus faecalis* utilizes maltose by connecting two incompatible metabolic routes via a  
555 novel maltose 6'-phosphate phosphatase (MapP). *Mol Microbiol* **88**: 234–253.
- 556 Monedero, V., Poncet, S., Mijakovic, I., Fieulaine, S., Dossonnet, V., Martin-Verstraete, I., *et al.*  
557 (2001) Mutations lowering the phosphatase activity of HPr kinase/phosphatase switch off  
558 carbon metabolism. *EMBO J* **20**: 3928–3937.
- 559 Muller, C., Cacaci, M., Sauvageot, N., Sanguinetti, M., Rattei, T., Eder, T., *et al.* (2015) The  
560 Intraperitoneal Transcriptome of the Opportunistic Pathogen *Enterococcus faecalis* in Mice. *PLoS*  
561 *One* **10**: e0126143.
- 562 Paganelli, F.L., Huebner, J., Singh, K.V., Zhang, X., Schaik, W. van, Wobser, D., *et al.* (2016)  
563 Genome-wide Screening Identifies Phosphotransferase System Permease BepA to Be Involved in  
564 *Enterococcus faecium* Endocarditis and Biofilm Formation. *J Infect Dis* **214**: 189–195.
- 565 Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010) edgeR: a Bioconductor package for  
566 differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.
- 567 Ruiz-Cruz, S., Espinosa, M., Goldmann, O., and Bravo, A. (2015) Global Regulation of Gene  
568 Expression by the MafR Protein of *Enterococcus faecalis*. *Front Microbiol* **6**: 1521.
- 569 Sauvageot, N., Mokhtari, A., Joyet, P., Budin-Verneuil, A., Blancato, V.S., Repizo, G.D., *et al.*  
570 (2017) *Enterococcus faecalis* Uses a Phosphotransferase System Permease and a Host  
571 Colonization-Related ABC Transporter for Maltodextrin Uptake. *J Bacteriol* **199**.
- 572 Sauvageot, N., Muller, C., Hartke, A., Auffray, Y., and Laplace, J.-M. (2002) Characterisation of  
573 the diol dehydratase pdu operon of *Lactobacillus collinoides*. *FEMS Microbiol Lett* **209**: 69–74.
- 574 Shafeeq, S., Kloosterman, T.G., and Kuipers, O.P. (2011) CelR-mediated activation of the  
575 cellobiose-utilization gene cluster in *Streptococcus pneumoniae*. *Microbiol Read Engl* **157**:  
576 2854–2861.
- 577 Shafeeq, S., Kuipers, O.P., and Kloosterman, T.G. (2013) Cellobiose-mediated gene expression in  
578 *Streptococcus pneumoniae*: a repressor function of the novel GntR-type regulator BguR. *PLoS*  
579 *One* **8**: e57586.
- 580 Terra, V.S., Zhi, X., Kahya, H.F., Andrew, P.W., and Yesilkaya, H. (2015) Pneumococcal 6-  
581 Phospho- $\beta$ -Glucosidase (BglA3) Is Involved in Virulence and Nutrient Metabolism. *Infect Immun*  
582 **84**: 286–292.
- 583 Thompson, J., Lichtenthaler, F.W., Peters, S., and Pikis, A. (2002) Beta-glucoside kinase (BglK)  
584 from *Klebsiella pneumoniae*. Purification, properties, and preparative synthesis of 6-phospho-  
585 beta-D-glucosides. *J Biol Chem* **277**: 34310–34321.
- 586 Thompson, J., Robrish, S.A., Pikis, A., Brust, A., and Lichtenthaler, F.W. (2001) Phosphorylation  
587 and metabolism of sucrose and its five linkage-isomeric alpha-D-glucosyl-D-fructoses by  
588 *Klebsiella pneumoniae*. *Carbohydr Res* **331**: 149–161.
- 589 Tong, H., and Dong, X. (2005) *Lactobacillus concavus* sp. nov., isolated from the walls of a  
590 distilled spirit fermenting cellar in China. *Int J Syst Evol Microbiol* **55**: 2199–2202.

- 591 Wang, L., Wang, S., and Li, W. (2012) RSeQC: quality control of RNA-seq experiments.  
592 *Bioinforma Oxf Engl* **28**: 2184–2185.
- 593 Wisplinghoff, H., Bischoff, T., Tallent, S.M., Seifert, H., Wenzel, R.P., and Edmond, M.B. (2004)  
594 Nosocomial Bloodstream Infections in US Hospitals: Analysis of 24,179 Cases from a Prospective  
595 Nationwide Surveillance Study. *Clin Infect Dis* **39**: 309–317.
- 596 Zeng, L., and Burne, R.A. (2009) Transcriptional regulation of the cellobiose operon of  
597 *Streptococcus mutans*. *J Bacteriol* **191**: 2153–2162.
- 598 Zhang, X., Maat, V. de, Guzmán Prieto, A.M., Prajsnar, T.K., Bayjanov, J.R., Been, M. de, *et al.*  
599 (2017) RNA-seq and Tn-seq reveal fitness determinants of vancomycin-resistant *Enterococcus*  
600 *faecium* during growth in human serum. *BMC Genomics* **18**: 893.
- 601 Zianni, M., Tessanne, K., Merighi, M., Laguna, R., and Tabita, F.R. (2006) Identification of the  
602 DNA Bases of a DNase I Footprint by the Use of Dye Primer Sequencing on an Automated  
603 Capillary DNA Analysis Instrument. *J Biomol Tech JBT* **17**: 103–113.
- 604

## 605 **Figure legends**

606 **Fig. 1. A.** Genetic organization of the *E. faecalis* OG1RF *gen* region. Large arrows represent the  
607 ORFs and their orientation indicates the transcriptional direction. The hairpin structure represents  
608 the putative transcriptional terminator located 11 nucleotides downstream from the *genR* stop  
609 codon (\*). Its sequence is delimited by convergent arrows. P1 and P2 indicate the transcriptional  
610 start sites. Chromatograms of footprint experiments were obtained in the absence (control) and  
611 presence of His-GenR. The binding sequence is underlined **B.** Nucleotide sequence of the *genBA*  
612 and *genR* promoter regions. The last base upstream from the poly A tail corresponds to the  
613 transcription initiation site (+1), C and T for P1 (*genBA*) and P2 (*genR*), respectively (in bold case).  
614 -10 boxes are underlined and the *cre* box sequence is written in italics. Letters in upper case indicate  
615 amino acid sequences of GenB and GenR proteins.

616

617 **Fig. 2.** Growth studies with the *E. faecalis* wild-type OG1RF strain (diamond), *genA* (square), *genB*  
618 (triangle) and *genR* (circle) mutant strains (filled symbols) and complemented strains (open  
619 symbols) on 0.3% cellobiose (A), 0.3% gentiobiose (B) and 0.3% amygdalin (C).

620

621 **Fig. 3.** Relative expression of the *genA* (A), *genB* (B) and *genR* genes (C) in the wild-type OG1RF  
622 strain. The wild-type strain was grown in M17cc medium containing either 0.3% of glucose (Glc),  
623 cellobiose (Cel), gentiobiose (Gen) or amygdalin (Amy) with or without 0.3% glucose. Induction  
624 factors were determined by the ratio of the *genA*, *genB* or *genR* expressions in the different growth  
625 conditions over the expression of these genes in glucose-alone condition (calibrator sample). The  
626 mean values calculated from three independent experiments are presented. Standard deviations are  
627 indicated.

628

629 **Fig. 4.** Relative expression of the *genA* (A), *genB* (B) and *genR* genes (C) in the  $\Delta$ *genR* mutant and  
630 complemented strains. The mutants were grown in M17cc medium containing either 0.3% glucose  
631 or 0.3% cellobiose. Induction factors were determined by the ratio of the *genA*, *genB* or *genR*  
632 expressions in the different growth conditions over the expression of these genes in glucose-alone  
633 condition (calibrator sample). The mean values calculated from three independent experiments are  
634 presented. Standard deviations are indicated.

635

636 **Fig. 5.** Thermal shift assays with GenR in the presence of various concentrations of different  
637 ligands (gentiobiose (▲), cellobiose (◆), gentiobiose-6'-P (△), cellobiose-6'-P (◇), arbutin-6-P  
638 (○) and salicin-6-P (□)) on the thermal stability  $\Delta Th$  of GenR.

639

640 **Fig. 6.** EMSA of the promoter region of the *genB* gene. EMSAs were performed with 0.1 nM of  
641 labelled DNA and varying concentrations of GenR. **A.** Binding of GenR to the *genB* promoter with  
642 1 mM gentiobiose-6'-P (lanes 2 to 5 and 11 to 14) or 1 mM gentiobiose (lanes 6 to 9). The  
643 concentrations of GenR used were from 500 nM to 4 nM and are indicated above the gel. DNA-  
644 protein interactions were performed on the *genB* promoter region (234 bp, lanes 1 to 9) or an  
645 internal *gusA* fragment (204 pb, lanes 10 to 14) as negative DNA control. Brackets on the right  
646 indicate the positions of the retarded complex and free DNA. Lanes 1 and 10 correspond to DNA  
647 without protein. **B.** Binding of GenR to the *genB* promoter with 1 mM of different ligands,  
648 gentiobiose-6'-P (Gen-6P), gentiobiose (Gen), cellobiose-6'-P (Cel-6P), cellobiose (Cel), salicin-  
649 6-P (Sal-6P), salicine (Sal), arbutin-6-P (Arb-6P), maltose-6'-P (Mal-6P) and maltose (Mal).  
650 DNA-protein interactions were performed on the *genB* promoter region (234 bp, lanes 1 to 11) or  
651 an internal *gusA* fragment (204 pb, lanes 12 to 15) as negative DNA control. Lane 1 and 11  
652 correspond to DNA without protein.

653

654 **Fig 7.** Binding of the His-CcpA/P-ser-HPr complex to the promoter region of the *genBA* operon.  
655 EMSA was performed with 0.1 nM of labelled DNA, containing the promoter region of the *genBA*  
656 operon (234 bp) or an internal *gusA* fragment (204 pb) as negative control. A fixed concentration  
657 of P-Ser-HPr (5.4  $\mu$ M; lanes 2 to 6 and 11 to 15) and different amounts of purified His<sub>6</sub>-CcpA were  
658 used (from 13.5 nM to 1.68 nM). Brackets on the right indicate the positions of the retarded complex  
659 and free DNA. Lanes 1 and 11 correspond to DNA without protein.

660

661 **Fig. 8. A.** Genetic organization of the *E. faecalis* OG1RF\_10745 to 10755 region. Arrows represent  
662 the ORFs and their orientation indicates the transcriptional direction. **B.** Growth studies with the  
663 *E. faecalis* wild-type OG1RF strain (diamond), the EIIA 10751 mutant (square) and EIIA10751  
664 complemented strains (triangle) on 0.3% glucose, 0.3% cellobiose, 0.3% gentiobiose and 0.3%  
665 amygdalin.

666

667 **Fig. 9.** Hydrolysis of disaccharides by GenA. **A.** 5  $\mu\text{g}$  of gentiobiose (lane 8) , gentiobiose-6'-P  
668 (lane 9), cellobiose (lane 10), cellobiose-6'-P (lane 11), maltose (lane (12) and maltose-6'-P (lane  
669 13) were incubated with 1  $\mu\text{g}/\text{mL}$  of purified His-tagged GenA at 37°C for 3 h in a final volume of  
670 10  $\mu\text{L}$ . Aliquots (5  $\mu\text{L}$ ) of the reaction mixture were subsequently spotted on silica gel plates  
671 (loading line: L) and the reaction products were separated by chromatography as described in  
672 experimental procedures. After drying, the plate was sprayed with Molisch reagent and heated at  
673 100°C for 10 minutes to visualize the reaction products. Migration of reaction products was  
674 compared to standards for glucose (lane 1), gentiobiose (lane 2), gentiobiose-6'-P (lane 3),  
675 cellobiose (lane 4), cellobiose-6'-P (lane 5), maltose (lane (6) and maltose-6'-P (lane 7). **B.** 5  $\mu\text{g}$   
676 of gentiobiose (lane 14) , gentiobiose-6'-P (lane 15), cellobiose (lane 16), cellobiose-6'-P (lane 17)  
677 were incubated with purified 2  $\mu\text{g}/\text{mL}$  of His-tagged GenA at 37°C for 3 h in a final volume of 20  
678  $\mu\text{L}$  and aliquots (10  $\mu\text{L}$ ) of the reaction mixture were spotted on silica gel plates. Glucose released  
679 by the action of GenA is indicated by the large arrows.

680

## 681 **Supplementary figures**

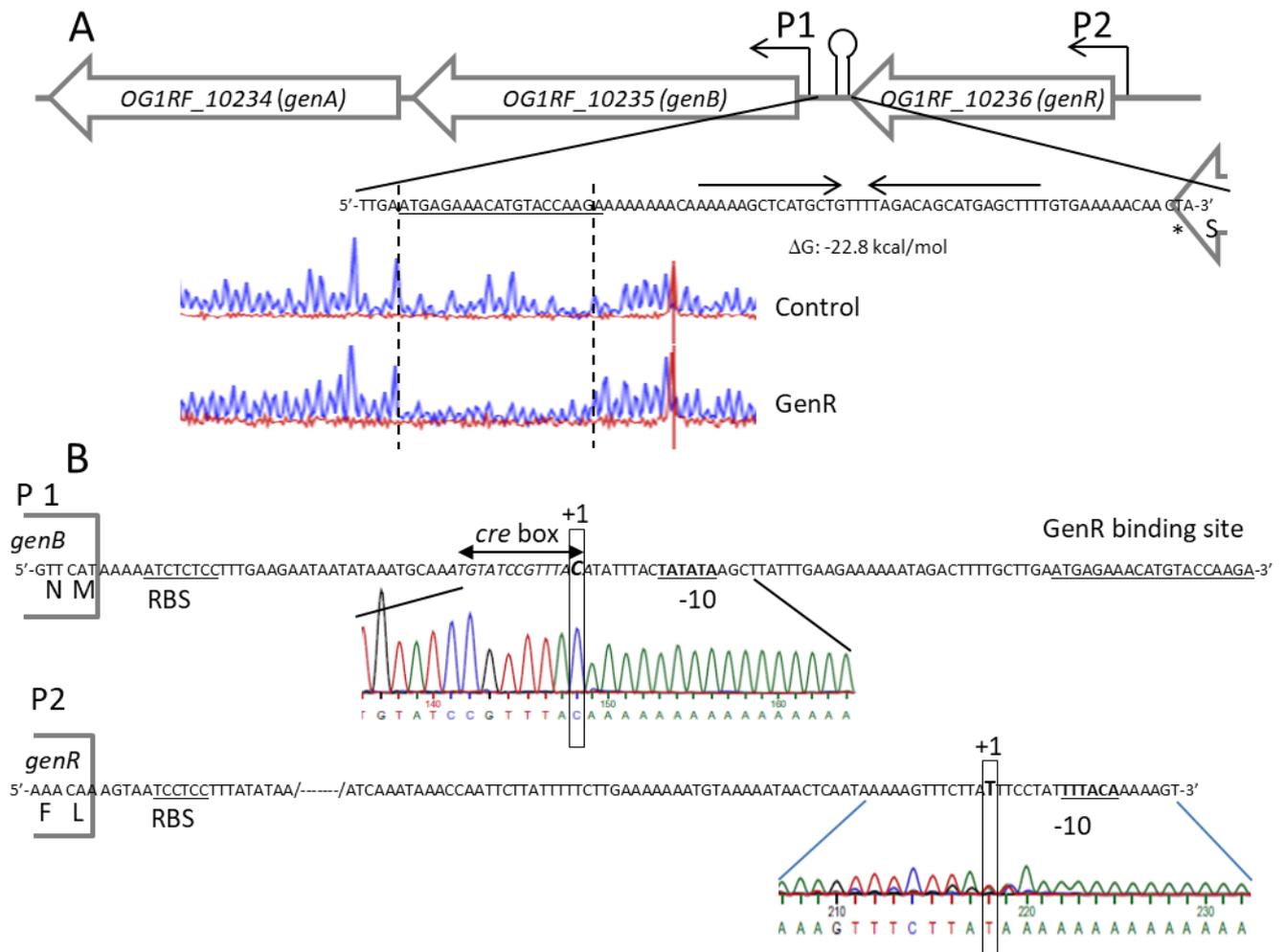
682

683 **Fig. S1.** Characterization of transcriptional units of the locus *gen* by RT-PCR on RNAs. **A.** Genetic  
684 organization of the *E. faecalis* OG1RF *gen* region. Amplified regions are indicated by arrows. **B.**  
685 Amplification products obtained on chromosomal DNA (control, lanes 1 and 4), on cDNA  
686 synthesised with reverse transcriptase (lanes 2 and 5) or without reverse transcriptase (lanes 3 and  
687 6).

688

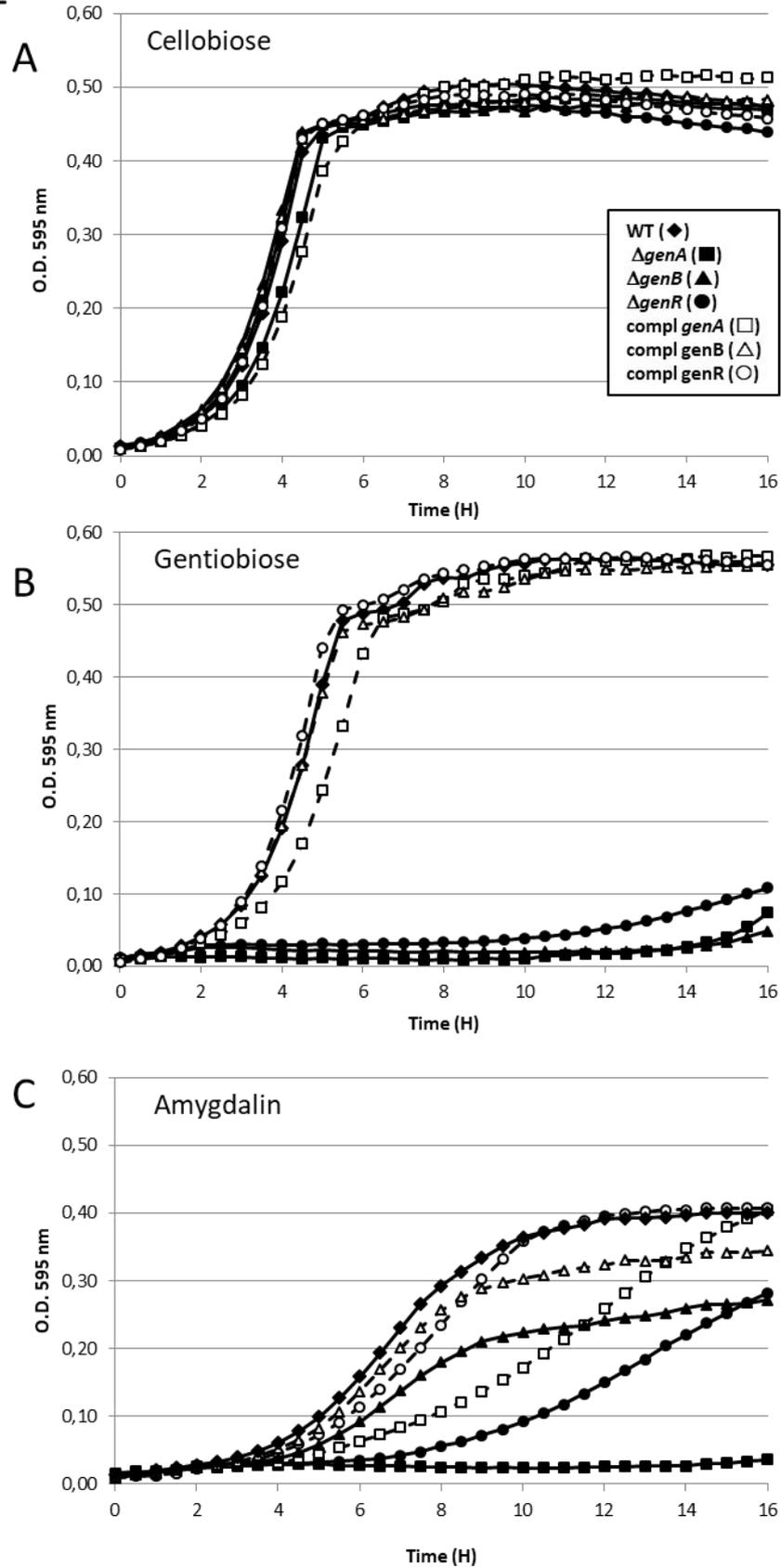
689 **Fig. S2.** The structure of gentiobiose, amygdalin, arbutin and salicin molecules.

Figure 1



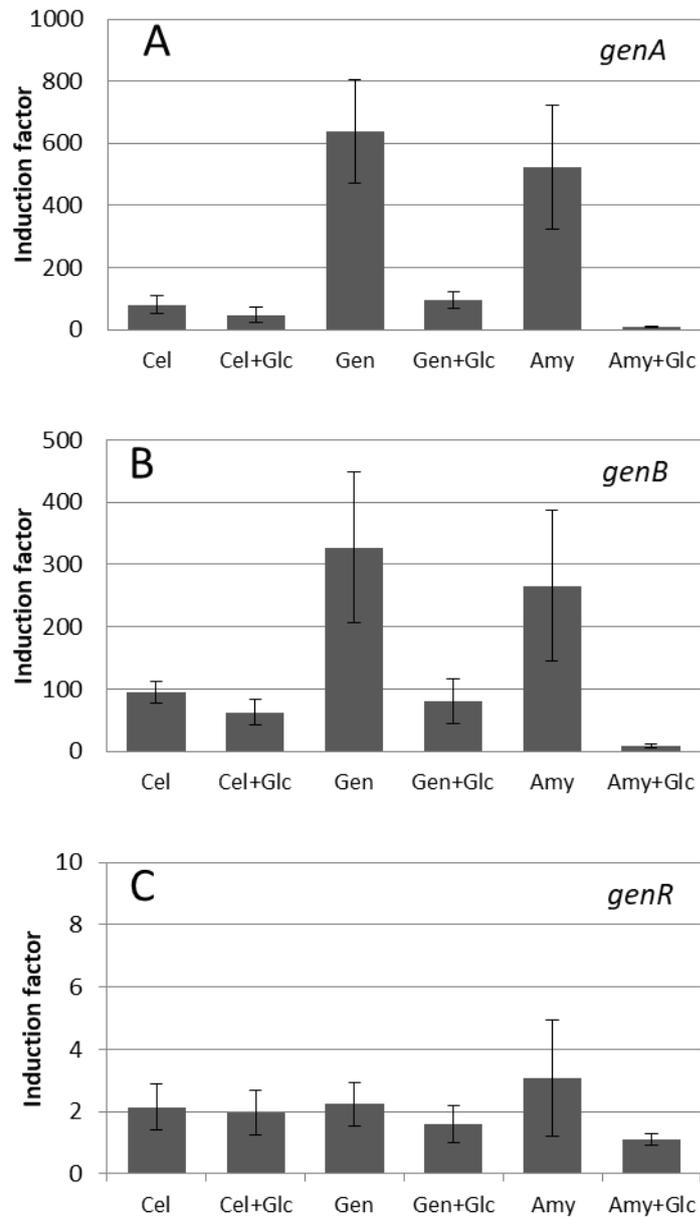
**Fig. 1. A.** Genetic organization of the *E. faecalis* OG1RF *gen* region. Large arrows represent the ORFs and their orientation indicates the transcriptional direction. The hairpin structure represents the putative transcriptional terminator located 11 nucleotides downstream from the *genR* stop codon (\*). Its sequence is delimited by convergent arrows. P1 and P2 indicate the transcriptional start sites. Chromatograms of footprint experiments were obtained in the absence (control) and presence of His-GenR. The binding sequence is underlined **B.** Nucleotide sequence of the *genBA* and *genR* promoter regions. The last base upstream from the poly A tail corresponds to the transcription initiation site (+1), C and T for P1 (*genBA*) and P2 (*genR*), respectively (in bold case). -10 boxes are underlined and the *cre* box sequence is written in italics. Letters in upper case indicate amino acid sequences of GenB and GenR proteins.

Figure 2



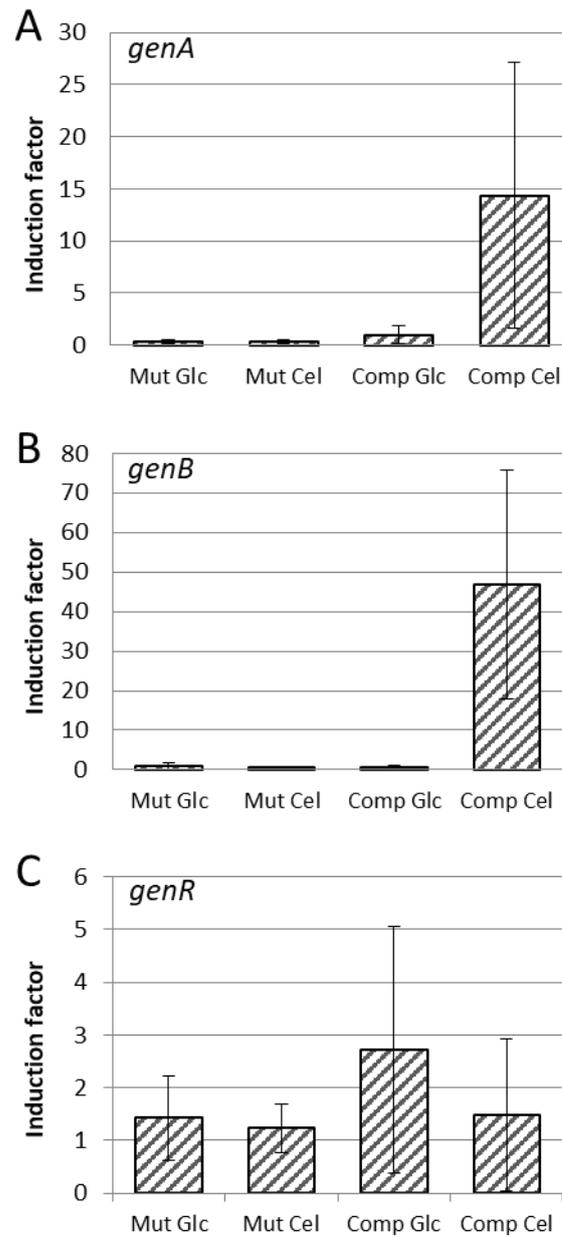
**Fig. 2.** Growth studies with the *E. faecalis* wild-type OG1RF strain (diamond), *genA* (square), *genB* (triangle) and *genR* (circle) mutant strains (filled symbols) and complemented strains (open symbols) on 0.3% cellobiose (A), 0.3% gentiobiose (B) and 0.3% amygdalin (C).

Figure 3



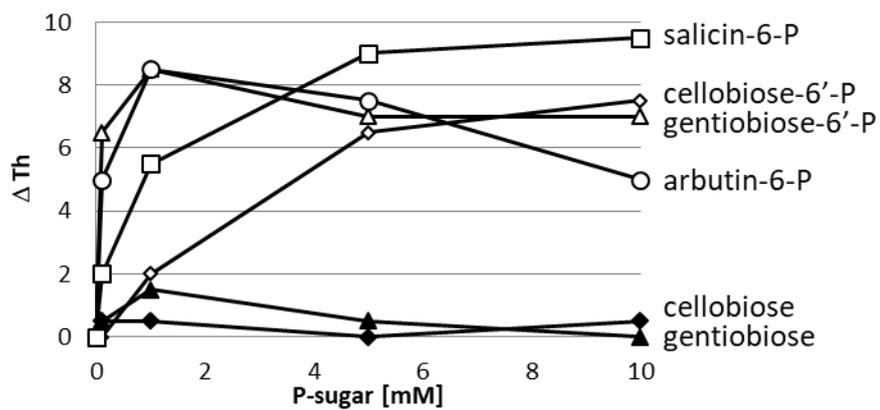
**Fig. 3.** Relative expression of the *genA* (A), *genB* (B) and *genR* genes (C) in the wild-type OG1RF strain. The wild-type strain was grown in M17cc medium containing either 0.3% of glucose (Glc), cellobiose (Cel), gentiobiose (Gen) or amygdalin (Amy) with or without 0.3% glucose. Induction factors were determined by the ratio of the *genA*, *genB* or *genR* expressions in the different growth conditions over the expression of these genes in glucose-alone condition (calibrator sample). The mean values calculated from three independent experiments are presented. Standard deviations are indicated.

Figure 4



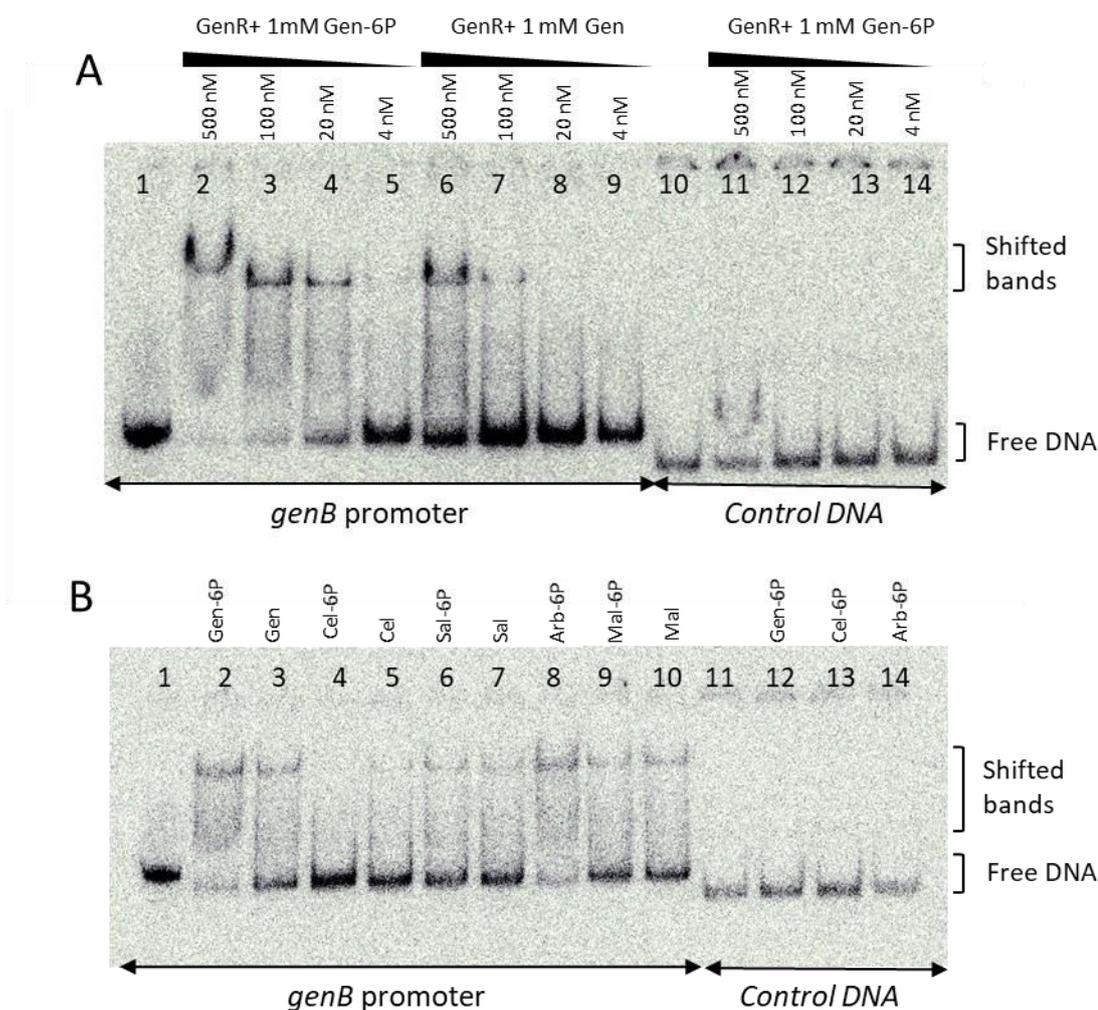
**Fig. 4.** Relative expression of the *genA* (A), *genB* (B) and *genR* genes (C) in the *DgenR* mutant and complemented strain. The mutant was grown in M17cc medium containing either 0.3% glucose or 0.3% cellobiose. Induction factors were determined by the ratio of the *genA*, *genB* or *genR* expressions in the different growth conditions over the expression of these genes in glucose-alone condition (calibrator sample). The mean values calculated from three independent experiments are presented. Standard deviations are indicated.

Figure 5



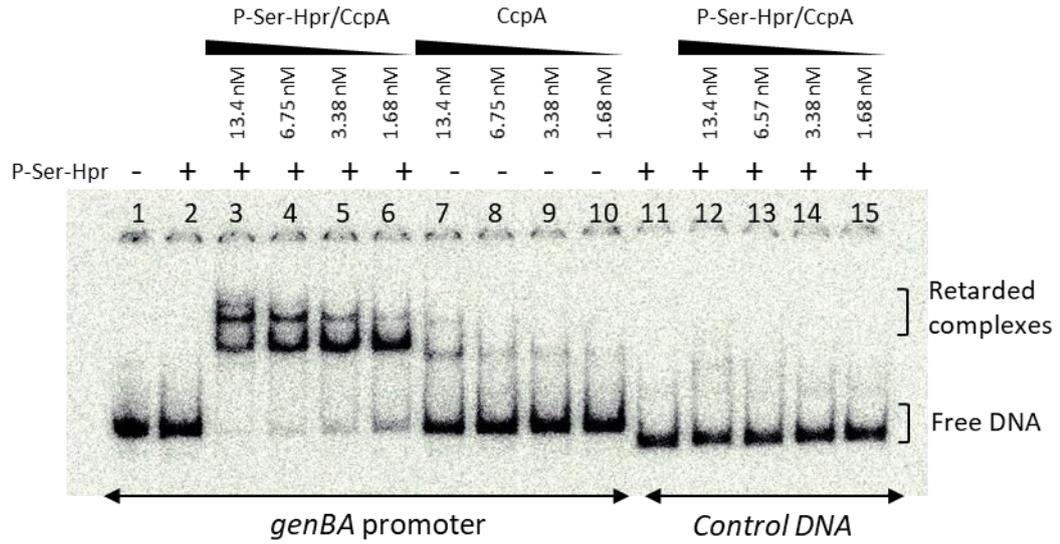
**Fig. 5.** Thermal shift assays with GenR in the presence of various concentrations of different ligands (gentiobiose (▲), cellobiose (◆), gentiobiose-6'-P (△), cellobiose-6'-P (◇), arbutin-6-P (○) and salicin-6-P (□)) on the thermal stability  $\Delta Th$  of GenR.

Figure 6



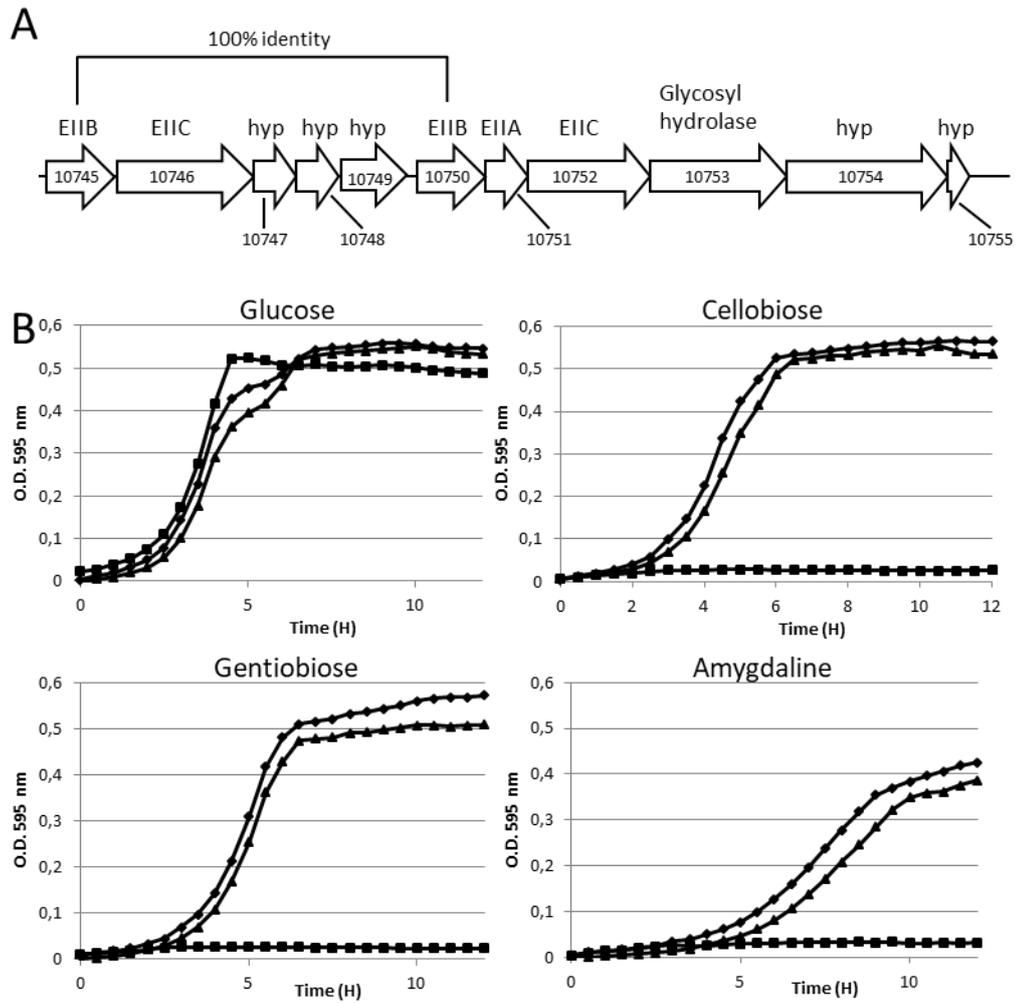
**Fig. 6.** EMSA of the promoter region of the *genB* gene. EMSAs were performed with 0.1 nM of labelled DNA and varying concentrations of GenR. **A.** Binding of GenR to the *genB* promoter with 1 mM gentiobiose-6'-P (lanes 2 to 5 and 11 to 14) or 1 mM gentiobiose (lanes 6 to 9). The concentrations of GenR used were from 500 nM to 4 nM and are indicated above the gel. DNA-protein interactions were performed on the *genB* promoter region (234 bp, lanes 1 to 9) or an internal *gusA* fragment (204 pb, lanes 10 to 14) as negative DNA control. Brackets on the right indicate the positions of the retarded complex and free DNA. Lanes 1 and 10 correspond to DNA without protein. **B.** Binding of GenR to the *genB* promoter with 1 mM of different ligands, gentiobiose-6'-P (Gen-6P), gentiobiose (Gen), cellobiose-6'-P (Cel-6P), cellobiose (Cel), salicin-6'-P (Sal-6P), salicine (Sal), arbutin-6'-P (Arb-6P), maltose-6'-P (Mal-6P) and maltose (Mal). DNA-protein interactions were performed on the *genB* promoter region (234 bp, lanes 1 to 11) or an internal *gusA* fragment (204 pb, lanes 12 to 15) as negative DNA control. Lane 1 and 11 correspond to DNA without protein.

Figure 7



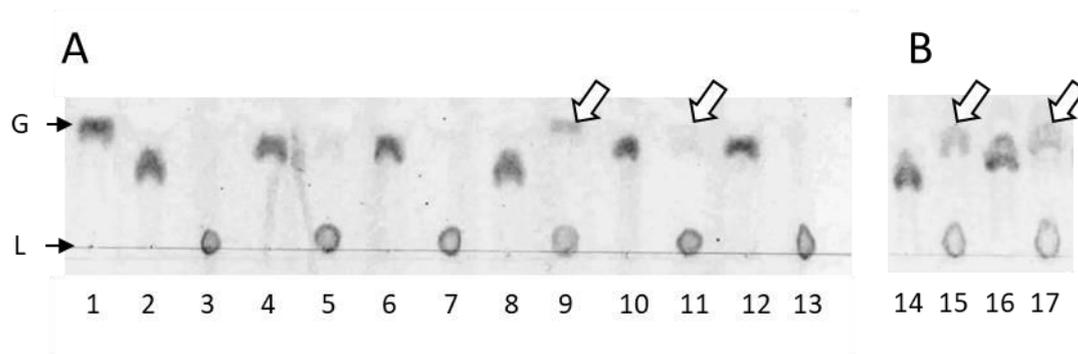
**Fig 7.** Binding of the His-CcpA/P-ser-HPr complex to the promoter region of the *genBA* operon. EMSA was performed with 0.1 nM of labelled DNA, containing the promoter region of the *genBA* operon (234 bp) or an internal *gusA* fragment (204 pb) as negative control. A fixed concentration of P-Ser-HPr (5.4  $\mu$ M; lanes 2 to 6 and 11 to 15) and different amounts of purified His<sub>6</sub>-CcpA were used (from 13.5 nM to 1.68 nM). Brackets on the right indicate the positions of the retarded complex and free DNA. Lanes 1 and 11 correspond to DNA without protein.

Figure 8



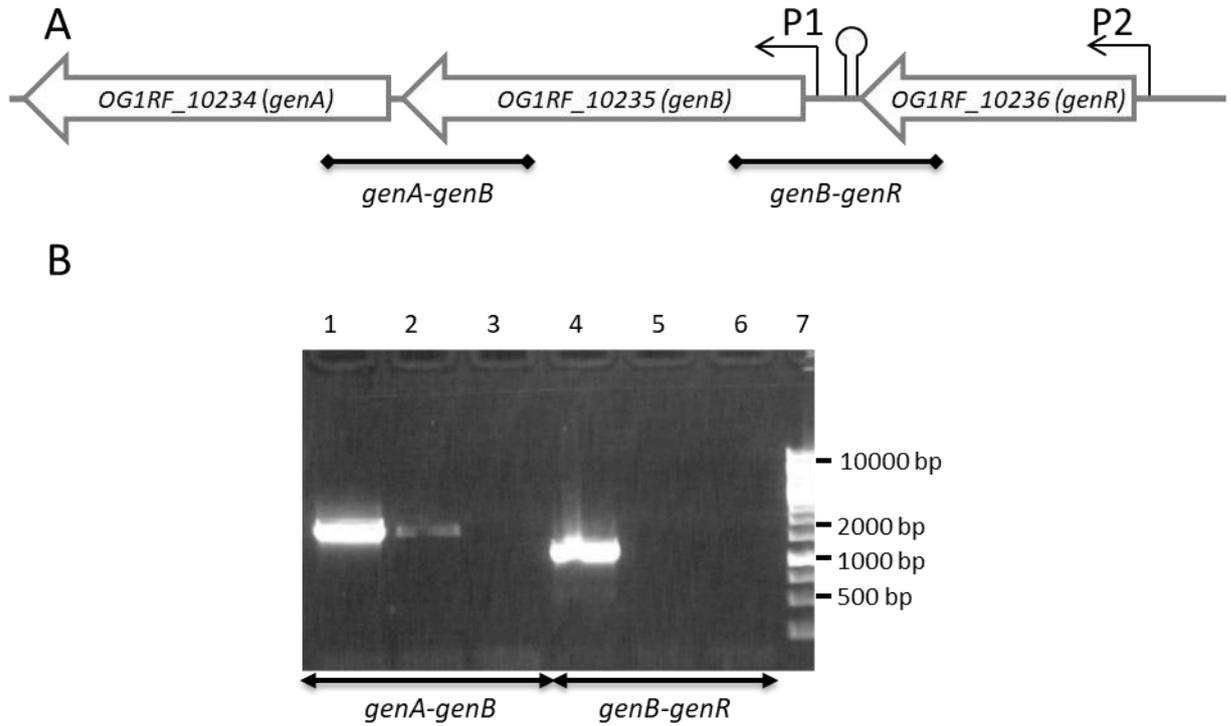
**Fig. 8. A.** Genetic organization of the *E. faecalis* OG1RF\_10745 to 10755 region. Arrows represent the ORFs and their orientation indicates the transcriptional direction. **B.** Growth studies with the *E. faecalis* wild-type OG1RF strain (diamond), the EIIA 10751 mutant (square) and EIIA10751 complemented strains (triangle) on 0.3% glucose, 0.3% cellobiose, 0.3% gentiobiose and 0.3% amygdalin.

Figure 9



**Fig. 9.** Hydrolysis of disaccharides by GenA. **A.** 5  $\mu$ g of gentiobiose (lane 8), gentiobiose-6'-P (lane 9), cellobiose (lane 10), cellobiose-6'-P (lane 11), maltose (lane 12) and maltose-6'-P (lane 13) were incubated with 1 mg/mL of purified His-tagged GenA at 37°C for 3 h in a final volume of 10  $\mu$ L. Aliquots (5  $\mu$ L) of the reaction mixture were subsequently spotted on silica gel plates (loading line: L) and the reaction products were separated by chromatography as described in experimental procedures. After drying, the plate was sprayed with Molisch reagent and heated at 100°C for 10 minutes to visualize the reaction products. Migration of reaction products was compared to standards for glucose (lane 1), gentiobiose (lane 2), gentiobiose-6'-P (lane 3), cellobiose (lane 4), cellobiose-6'-P (lane 5), maltose (lane 6) and maltose-6'-P (lane 7). **B.** 5  $\mu$ g of gentiobiose (lane 14), gentiobiose-6'-P (lane 15), cellobiose (lane 16), cellobiose-6'-P (lane 17) were incubated with purified 2 mg/mL of His-tagged GenA at 37°C for 3 h in a final volume of 20  $\mu$ L and aliquots (10  $\mu$ L) of the reaction mixture were spotted on silica gel plates. Glucose released by the action of GenA is indicated by the large arrows.

Figure S1



**Fig. S1.** Characterization of transcriptional units of the locus *gen* by RT-PCR on RNAs. **A.** Genetic organization of the *E. faecalis* OG1RF *gen* region. Amplified regions are indicated by arrows. **B.** Amplification products obtained on chromosomal DNA (control, lanes 1 and 4), on cDNA synthesised with reverse transcriptase (lanes 2 and 5) or without reverse transcriptase (lanes 3 and 6).

Figure S2

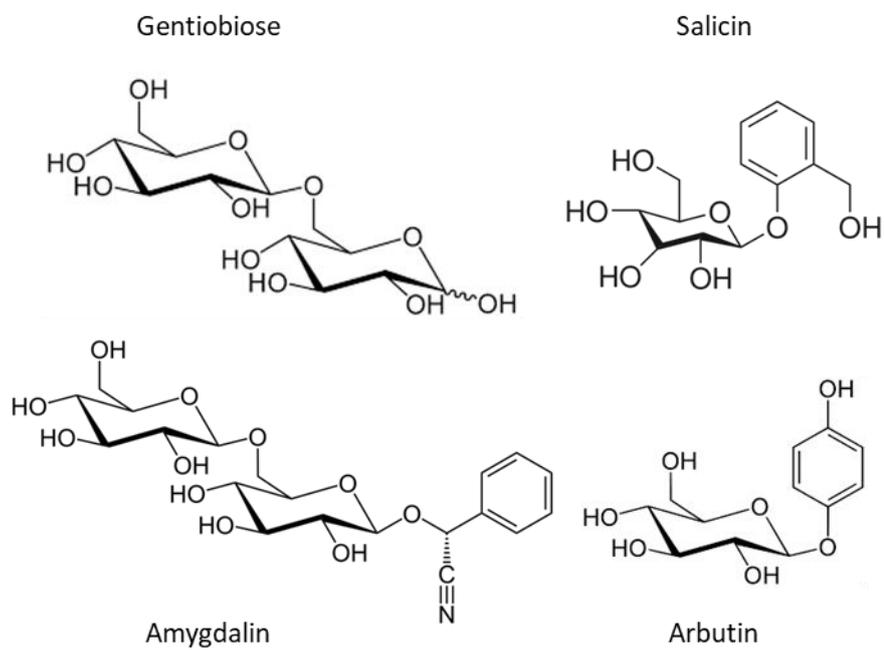


Fig. S2. The structure of gentiobiose, amygdalin, arbutin and salicin molecules.

**TABLE S1.** Primers used in this study.

Primers	Sequence (5' to 3') <sup>a</sup>	Orientation <sup>b</sup>
genB_sp1	GCCAGCTGTCCAATCTTTGT	-
genB_sp2	TAGGCAATACCGATGACCGC	-
genB_sp3	AAACACGGAACCAACAATGG	-
genR_sp1	TGAACCGTGGCATCATAAAA	-
genR_sp2	TCGGAAAAACCAGCATAGCC	-
genR_sp3	GGTGGAGACATGACAGGAAGTAG	-
pMAD_genA_for	TTTT <u>GGATCC</u> ATTTTTAACGGTCAC	+
pMAD_genA_rev	GGCC <u>GTCGAC</u> TTTAGACGAAGGCC	-
Mut_genA_for	TTGCT <u>AGATCT</u> TGATTCACTCGTTACCAG	+
Mut_genA_rev	ATCA <u>AGATCT</u> AGCAAATTAGACAAAAAATC	-
pMAD_genB_for	CCAG <u>GGATCC</u> AAATGATTGTTCAAGC	+
pMAD_genB_rev	ATCG <u>GTCGAC</u> CGTTTCTTCTGCGTAG	-
Mut_genB_for	GATCT <u>AGATCT</u> TAAAGATGCGTTTCTGGC	+
Mut_genB_rev	CTTA <u>AGATCT</u> AGATCACATAAACAAAATG	-
pMAD_genR_for	GCTT <u>GGATCC</u> TCTCGCCCTATGAG	+
pMAD_genR_rev	TTGC <u>GTCGAC</u> CCATTGAGAAATAGC	-
Mut_genR_for	TTGA <u>AGATCT</u> TAATATGCAGTGAAGAAG	+
Mut_genR_rev	ATTA <u>AGATCT</u> TCAAAAACCAGCATAGCC	-
verif_genA	TTGTCTAATTTGCTAGATCTTGA	+
verif_genB	TTTTGTTTATGTGATCTAGATCTT	+
verif_genR	TATGCTGGTTTTGAAGATCTT	+
EF0291L	AGAAGACCGCGAAGCAATTA	+
EF0291R	TGGAAGTCGGGATTGATAGC	-
EF0292L	GCCATGGTGGTTTTCATTTT	+
EF0292R	GCCAGCTGTCCAATCTTTGT	-
EF0293L	TGAACCGTGGCATCATAAAA	-
EF0293R	CAAAATGGGCTATGCTGGTT	+
GenA_pQE30_for	CTGAG <u>GGATCC</u> ATGTCAATTTTGAAAAATGATTTTC	+
GenA_pQE30_rev	GGCCA <u>AGCTTT</u> TATAATCTTCCCATTCGATAGC	-
GenR_pQE30_for	GAGGG <u>GATCC</u> TTGTTTGATTTAGAGAAGGTTTCCAG	+
GenR_pQE30_rev	TGAAA <u>AGCTT</u> CTAGAAAGGTGTTTCATCTAGTTC	-
Promo_genB_for	AGGAACTAGATGAAACACCTTCC	+
Promo_genB_rev	GGAGGTAATAATTTCTCGTTAATCC	-
pMAD_1017for	TTAATGAATTCTGTTATTTTGGATATC	+
EF1019_rev	AAATGGATCCAAGAAAGGTTTGTTTAC	-
mut_1018for	GTATTGCTAGCGAGTAATCTCGTTGTC	+
mut_1018rev	CCCATGCTAGCTTCTAAGTTCTGTTGTC	-
verif_1017	AGCAGATGAATTTTACACGC	+
Verif_1019rev	TTGCCCTGTTCCACCTAAG	-

<sup>a</sup>Underlined sequences correspond to restriction sites and bold letters indicate stop codons.

<sup>b</sup>The + means oriented in and the - oriented against the direction of transcription.

**TABLE S2.** List of genes up-regulated during growth on cellobiose and/or gentiobiose.

gene	Induction factor on cellobiose	Induction factor on gentiobiose	name	description
OG1RF_10063		3.27		N-acylglucosamine-6-phosphate 2-epimerase
OG1RF_10085	2.65			diacylglycerol kinase catalytic domain protein
OG1RF_10100		17.44	arcB	ornithine carbamoyltransferase
OG1RF_10101	3.33	22.91	arcC	carbamate kinase
OG1RF_10102	3.27	16.08	ntcA	global nitrogen regulator NtcA
OG1RF_10103		10.58		UIT3 family protein
OG1RF_10107		8.33	EndoS	family 20 glycosyl hydrolase
OG1RF_10136	2.55			iron (Fe3+) ABC superfamily transporter binding protein
OG1RF_10156		4.86	rplV	50S ribosomal protein L22
OG1RF_10160		4.88	rpsQ	30S ribosomal protein S17
OG1RF_10215	4.44		arb	6-phospho-beta-glucosidase
OG1RF_10216	2.42	5.10	bglA	6-phospho-beta-glucosidase
OG1RF_10234	98.49	1944.76	celA	6-phospho-beta-glucosidase
OG1RF_10235	168.90	1194.88	celB	Gentiobiose-specific PTS component EIIC
OG1RF_10265	3.54	3.49		ankyrin repeat family protein
OG1RF_10273		50.44		major facilitator family transporter
OG1RF_10274		38.79	arcC2	carbamate kinase
OG1RF_10310		5.95	eda-1	2-dehydro-3-deoxy-phosphogluconate aldolase
OG1RF_10311		3.72		PfkB family carbohydrate kinase
OG1RF_10312		4.21	kdu1	4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase
OG1RF_10314		7.25	rhiN	RhiN protein
OG1RF_10315		7.65		TRAP-T family transporter binding protein
OG1RF_10316		5.79		TRAP-T family transporter membrane protein
OG1RF_10319		16.18	rhaB	rhamnulokinase
OG1RF_10320		14.39	rhaA	L-rhamnose isomerase
OG1RF_10321		33.74	rhaD	rhamnulose-1-phosphate aldolase
OG1RF_10322		18.45	rhaU	L-rhamnose 1-epimerase
OG1RF_10323		12.02		AraC family transcriptional regulator
OG1RF_10328	3.44			peptidoglycan binding protein
OG1RF_10336	10.40			putative glucosamine-6-phosphate deaminase
OG1RF_10340		3.97		PTS family mannose/fructose/sorbose porter component IIC
OG1RF_10342	2.83		sorB	PTS family mannose/fructose/sorbose porter component IIB
OG1RF_10360		3.53	feoB	FeoB family ferrous iron (Fe2+) uptake protein
OG1RF_10365		17.28		hypothetical protein
OG1RF_10404		10.50	csn	csn1 family CRISPR-associated protein
OG1RF_10405		9.92	cas1	CRISPR-associated protein cas1
OG1RF_10406		7.40	cas2	CRISPR-associated protein cas2
OG1RF_10407		10.92	csn2	csn2 family CRISPR-associated protein
OG1RF_10412	2.10			ABC superfamily ATP binding cassette transporter membrane protein
OG1RF_10415		4.21		GNAT family acetyltransferase
OG1RF_10445		146.86		hypothetical protein
OG1RF_10446		7.18		hypothetical protein
OG1RF_10449		7.07		hypothetical protein
OG1RF_10456		4.04	lacR	lactose PTS family porter repressor

OG1RF_10503		4.34		PEP phosphonmutase family protein
OG1RF_10503	2.81			PEP phosphonmutase family protein
OG1RF_10508	4.43	3.22		von Willebrand factor domain LPTXG domain protein
OG1RF_10550		21.21		family 8 polysaccharide lyase
OG1RF_10565		6.40		PTS family porter component IIC
OG1RF_10566		6.39		hypothetical protein
OG1RF_10567		7.23		outer surface protein
OG1RF_10612	2.16			hypothetical protein
OG1RF_10651	2.09			hypothetical protein
OG1RF_10722	2.02		mraW	S-adenosyl-methyltransferase MraW
OG1RF_10747		16.00		hypothetical protein
OG1RF_10748	5.52	47.62		hypothetical protein
OG1RF_10749	5.76	53.47		hypothetical protein
OG1RF_10750	2.90		celA3	PTS family porter,C enzyme I
OG1RF_10751	3.18		lacF	PTS family lactose/cellobiose porter component IIA
OG1RF_10752	5.92			PTS family lactose/cellobiose (lac) porter component IIC
OG1RF_10753	14.77	26.92	bglA2	6-phospho-beta-glucosidase
OG1RF_10754	17.63	42.28		GNAT family acetyltransferase
OG1RF_10808		8.87		M protein trans-acting positive regulator
OG1RF_10810		5.85		hypothetical protein
OG1RF_10825		7.90		hypothetical protein
OG1RF_10826		4.11		hypothetical protein
OG1RF_10827		4.61		hypothetical protein
OG1RF_10849		8.45	galK	galactokinase
OG1RF_10850		8.26	galE2	UDP-glucose 4-epimerase
OG1RF_10857		4.13		GNAT family acetyltransferase
OG1RF_10870		7.72		cell wall surface anchor family protein
OG1RF_10904		3.07		PTS family porter component IIA
OG1RF_10905		3.73	ulaA	PTS family ascorbate porter component IIC
OG1RF_10907		8.06	ulaD	hexulose-6-phosphate synthase
OG1RF_10908		8.30	sgbU	L-ribulose-5-phosphate 3-epimerase
OG1RF_10927	2.16			YpsA like protein
OG1RF_10965	2.19			response regulator
OG1RF_11004		8.44		ABC superfamily transporter membrane protein
OG1RF_11005		27.00		ABC superfamily transporter substrate-binding protein
OG1RF_11006		17.64		hypothetical protein
OG1RF_11007		14.41		acetyl xylan esterase
OG1RF_11008		20.15		endonuclease/exonuclease/phosphatase
OG1RF_11009		26.19	bglX	putative beta-glucosidase
OG1RF_11010		7.84		cellobiose-phosphorylase
OG1RF_11014	11.08		yckE	beta-glucosidase
OG1RF_11015	11.64			gfo/ldh/MocA family oxidoreductase
OG1RF_11037	2.16			cell wall surface anchor family protein
OG1RF_11082	2.80			serine/threonine protein phosphatase
OG1RF_11123		7.43		hypothetical protein
OG1RF_11140	9.44		mgtA2	magnesium-importing ATPase
OG1RF_11141	4.30		pdhA	pyruvate dehydrogenase complex E1 component alpha subunit
OG1RF_11142	3.93		pdhB	pyruvate dehydrogenase complex E1 component beta subunit

OG1RF_11143	3.10	3.69	aceF	pyruvate dehydrogenase complex E2 dihydrolipoamide acetyltransferase
OG1RF_11144	2.53	3.39	lpd	dihydrolipoyl dehydrogenase
OG1RF_11146	2.05		gldA	glycerol dehydrogenase
OG1RF_11182		6.77	moaA	molybdenum cofactor biosynthesis protein A
OG1RF_11184		8.25	moaB	molybdenum cofactor biosynthesis protein B
OG1RF_11185		10.93		molybdopterin-binding domain protein
OG1RF_11186		16.62	modA	molybdenum ABC superfamily transporter binding protein
OG1RF_11187		16.78	modB	molybdenum (Mo2+) ABC superfamily transporter membrane protein
OG1RF_11188		11.76		ABC superfamily transporter ABC protein
OG1RF_11199	94.01	93.71	licH	6-phospho-beta-glucosidase
OG1RF_11201		5.08	msrA	ABC superfamily transporter ABC protein
OG1RF_11235		26.74		PTS family porter component IIA
OG1RF_11280	2.54		aroE	shikimate dehydrogenase
OG1RF_11305	2.91		ndh	NADH dehydrogenase
OG1RF_11315	3.79		phrB	deoxyribodipyrimidine photolyase
OG1RF_11317		3.64	scrA	PTS family beta-glucosides porter IIAABC component
OG1RF_11318		5.49	malL2	oligo-1,6-glucosidase
OG1RF_11325	5.41			integral membrane protein
OG1RF_11348		7.62	pduU	propanediol utilization protein PduU
OG1RF_11373	3.06		lpd2	dihydrolipoyl dehydrogenase
OG1RF_11418		6.48		family 38 glycosyl hydrolase
OG1RF_11419		6.56		conserved protein UCPO28846
OG1RF_11470	2.32		pstS2	phosphate ABC superfamily transporter, binding protein
OG1RF_11499	3.27		purE	phosphoribosylaminoimidazole carboxylase catalytic subunit PurE
OG1RF_11509		35.95	EngEF	hypothetical protein
OG1RF_11511		14.16		PTS family mannose/fructose/sorbose porter component IID
OG1RF_11512		10.13		PTS family mannose/fructose/sorbose porter component IIC
OG1RF_11513		10.34	levE	PTS family mannose/fructose/sorbose porter component IIB
OG1RF_11514		5.65	bgaL2	beta-galactosidase
OG1RF_11519	6.46		gspA1	general stress protein A
OG1RF_11520	10.67		gspA2	general stress protein A/glycosyltransferase family 8
OG1RF_11530		28.78	amiD4	N-acetylmuramoyl-L-alanine amidase
OG1RF_11531		29.53		glycosyl hydrolase
OG1RF_11590		10.98	glpF2	MIP family glycerol uptake facilitator protein GlpF
OG1RF_11698	2.57			protein of hypothetical function DUF147
OG1RF_11716		4.65	rgpB	group 2 glycosyl transferase
OG1RF_11769		7.40		hypothetical protein/DNRLRE domain-containing protein
OG1RF_11770		10.78		hypothetical protein
OG1RF_11772		4.44		sugar ABC superfamily transporter membrane protein
OG1RF_11773		10.72		ABC superfamily transporter membrane protein
OG1RF_11774		49.07		sugar ABC superfamily transporter sugar-binding protein
OG1RF_11775		20.72		glucuronoyl hydrolase
OG1RF_11784		19.10	iolG2	inositol 2-dehydrogenase
OG1RF_11786		13.12	iolB	myo-inositol catabolism protein IolB
OG1RF_11787		6.83	iolD	3D-(3,5/4)-trihydroxycyclohexane-1,2-dione hydrolase
OG1RF_11788		3.99	iolC	5-dehydro-2-deoxygluconokinase
OG1RF_11800	3.46	3.32		family 4 N-acetylmuramoyl-L-alanine amidase
OG1RF_11801	2.72			hypothetical protein

OG1RF_11802	2.49		hypothetical protein
OG1RF_11822		7.95	ATP-binding protein
OG1RF_11835	3.31		TrmH family RNA methyltransferase
OG1RF_11846	2.26	phoH	PhoH family protein
OG1RF_11924	3.25	6.06	cell wall surface anchor family protein
OG1RF_11940		18.80	nifJ2 pyruvate:ferredoxin oxidoreductase
OG1RF_11944		23.57	YqeB family selenium-dependent molybdenum hydroxylase system protein
OG1RF_11945		18.27	yqeC hypothetical protein
OG1RF_11946		18.92	hypothetical protein
OG1RF_11947		15.84	yedF2 selenium metabolism protein YedF
OG1RF_11948		18.31	selD selenide water dikinase
OG1RF_11949		13.56	putative cysteine desulfurase
OG1RF_11953		14.04	molybdenum transport domain protein
OG1RF_11954		7.37	xanthine/uracil permease
OG1RF_11955		38.48	endoribonuclease inhibitor of translation
OG1RF_11956		49.67	arcC3 carbamate kinase
OG1RF_11957		41.21	ygeW carbamoyltransferase YgeW
OG1RF_11958		38.02	ygeY M20/DapE family protein YgeY
OG1RF_11959		53.44	dpaL diaminopropionate ammonia-lyase
OG1RF_11960		35.46	hydA dihydropyrimidinase
OG1RF_11961		11.41	ygfK selenate reductase subunit YgfK
OG1RF_11975	4.74		bglH beta-glucosidase
OG1RF_11993		3.94	smpB SsrA-binding protein
OG1RF_12026	2.23		alcohol dehydrogenase
OG1RF_12041	2.98	relA2	GTP diphosphokinase
OG1RF_12058		4.20	sbcC exonuclease SbcC
OG1RF_12077		3.44	amino acid permease
OG1RF_12098	2.27		gfo/ldh/MocA family oxidoreductase
OG1RF_12167		4.10	endOF3 Mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase
OG1RF_12189	2.08		MarR family transcriptional regulator
OG1RF_12207	2.14		ABC superfamily ATP binding cassette transporter, ABC protein
OG1RF_12247	2.25	ulaA2	PTS family ascorbate porter, component IIC
OG1RF_12257		20.62	4-hydroxy-2-oxoglutarate aldolase
OG1RF_12258		17.27	L-seryl-tRNA(Sec) selenium transferase
OG1RF_12259		13.56	PTS family mannose/fructose/sorbose porter component IID
OG1RF_12260		14.91	PTS family mannose/fructose/sorbose porter component IIC
OG1RF_12261	3.70	14.69	PTS family porter, enzyme I
OG1RF_12262	6.32	16.87	PTS system fructose component IIA
OG1RF_12288		3.76	hypothetical protein
OG1RF_12289		3.50	hypothetical protein
OG1RF_12290		5.87	fdrA FdrA protein
OG1RF_12300	2.68		sigma-70 family protein
OG1RF_12303		13.25	family 8 polysaccharide lyase
OG1RF_12352		7.25	ABC superfamily transporter ABC protein
OG1RF_12356		11.13	hypothetical protein
OG1RF_12379		4.73	ABC superfamily transporter ABC protein
OG1RF_12432	3.90	msrB	peptide-methionine (R)-S-oxide reductase
OG1RF_12450	2.61		hypothetical protein

OG1RF_12455	2.03		hypothetical protein
OG1RF_12464	4.07		ABC superfamily transporter, binding protein
OG1RF_12531	3.84	ctsR	transcriptional regulator CtsR

---

**TABLE S3.** List of genes down-regulated during growth on cellobiose and/or gentiobiose.

gene	Repression factor on cellobiose	Repression factor on gentiobiose	name	description
OG1RF_10013	4.66	11.69	purA	adenylosuccinate synthase
OG1RF_10018	2.64	4.53	manX	PTS family mannose/fructose/sorbose porter component IIB
OG1RF_10019	3.23	3.53	manX2	PTS family mannose porter, IIB component
OG1RF_10020	3.46	3.93	manY	PTS family mannose/fructose/sorbose porter component IIC
OG1RF_10021	3.48	3.20		PTS family mannose/fructose/sorbose porter component IID
OG1RF_10022	2.00	3.57		hypothetical protein
OG1RF_10057	2.71			oligopeptide ABC superfamily transporter, binding protein
OG1RF_10077	4.66		atpF	ATP synthase FO sector subunit B
OG1RF_10078	2.97			hypothetical protein
OG1RF_10079	4.76			major facilitator family transporter
OG1RF_10080		50.90		hypothetical protein
OG1RF_10081		3221.89	hupB	DNA-binding protein
OG1RF_10126	2.39			ABC superfamily transporter ABC protein
OG1RF_10127	2.73			ABC superfamily, transporter membrane protein
OG1RF_10128	2.74			ABC superfamily transporter membrane protein
OG1RF_10148	5.75	5.31	thiD	phosphomethylpyrimidine kinase
OG1RF_10180	2.12			hypothetical protein
OG1RF_10181	2.51			M20D family peptidase
OG1RF_10187	2.04		brnQ	LIVCS family branched chain amino acid:cation symporter
OG1RF_10286	2.17			membrane protein
OG1RF_10305	2.79			hypothetical protein
OG1RF_10391	2.45			MATE efflux family protein
OG1RF_10538	2.26			ABC superfamily transporter ABC protein
OG1RF_10619		4.66	aspB	aspartate aminotransferase
OG1RF_10620		5.63		ABC superfamily transporter ABC protein
OG1RF_10623	2.77	7.97		putative glycerol dehydrogenase
OG1RF_10774		18.99		xanthine/uracil permease family protein
OG1RF_10781		2.62	rpmF	50S ribosomal protein L32
OG1RF_10798		13.23		hypothetical protein
OG1RF_10804	2.30			ATP-binding protein
OG1RF_10895	2.09	4.09		glutamine ABC superfamily transporter membrane protein
OG1RF_10896		3.94		amino acid ABC superfamily transporter membrane protein
OG1RF_10991	6.34	23.57		spermidine/putrescine ABC superfamily transporter permease protein
OG1RF_10992	8.52	20.69		ABC superfamily transporter ABC protein
OG1RF_10993	9.20	11.37		spermidine/putrescine ABC superfamily transporter
OG1RF_10994	13.83		ade	adenine deaminase
OG1RF_10995	17.93		mtaD	putative S-adenosylhomocysteine deaminase
OG1RF_11301	2.78			major facilitator family transporter
OG1RF_11489	14.99		purD	phosphoribosylamine-glycine ligase
OG1RF_11490	13.20		purH	bifunctional purine biosynthesis protein PurH
OG1RF_11491	10.25		purN	phosphoribosylglycinamide formyltransferase
OG1RF_11492	10.81		purM	phosphoribosylaminoimidazole synthetase
OG1RF_11493	7.36		purF	amidophosphoribosyltransferase
OG1RF_11494	4.56		purL	phosphoribosylformylglycinamide synthase

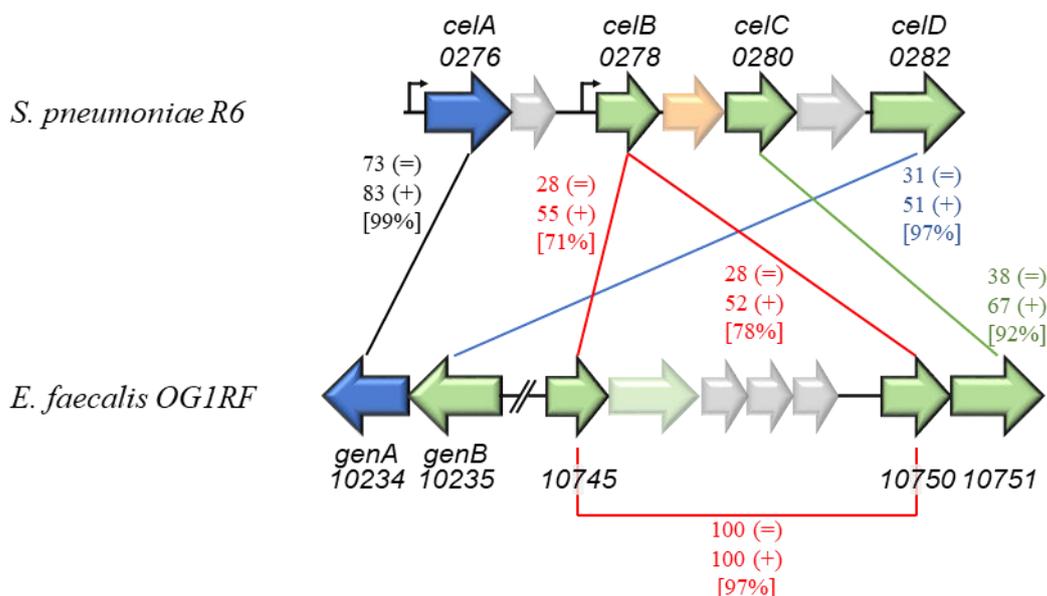
OG1RF_11495	3.01		purL2	phosphoribosylformylglycinamide synthase
OG1RF_11496	2.63	7.03	purS	phosphoribosylformylglycinamide synthase subunit PurS
OG1RF_11497		4.36	purC	phosphoribosylaminoimidazole-succinocarboxamide synthetase
OG1RF_11521	2.53			phosphatidylglycerol--membrane-oligosaccharide glycerophosphotransferase
OG1RF_11629	6.24	8.68	ywIG	YwIG protein
OG1RF_11630	5.79	4.99		hypothetical protein
OG1RF_11631	4.13			putative phosphomethylpyrimidine kinase
OG1RF_11692	2.18		glmS	D-fructose-6-phosphate amidotransferase
OG1RF_11693	2.08			cobalt (Co2+) ABC superfamily transporter membrane protein
OG1RF_11701		2.70	glnA	glutamate--ammonia ligase
OG1RF_11860	4.20	10.23	guaC	GMP reductase
OG1RF_11861	3.37	10.74		NCS2 family xanthine/uracil permease
OG1RF_11862		10.75	guaD	guanine deaminase
OG1RF_11999		4.68		P-ATPase superfamily P-type ATPase cadmium transporter
OG1RF_12006	2.21		groEL	chaperonin GroEL
OG1RF_12014		4.14		glycine betaine/carnitine/choline ABC superfamily transporter
OG1RF_12075		3.88		integral membrane protein
OG1RF_12135	2.21		thiD2	phosphomethylpyrimidine kinase
OG1RF_12233	2.09			CoA-binding domain protein
OG1RF_12235	2.49			S1 family extracellular protease
OG1RF_12238	2.28	7.75		NCS2 family xanthine/uracil permease family protein
OG1RF_12284	2.19	4.14		sulfate transporter/STAS domain protein
OG1RF_12476		11.79		PTS family fructose/mannitol (fru) porter component IIA

---

## C. Résultats complémentaires

La recherche d'homologie par BLASTp à partir de la  $\beta$ -hydrolase CelA de *S. pneumoniae* présentée dans la partie V.B de l'introduction permet d'identifier une protéine présentant un fort taux de similitude au sein de l'opéron bicistronique *genBA* d'*E. faecalis* présenté dans le papier précédent. Le premier gène de cet opéron (*genB*) présente lui aussi de fortes similitudes avec le gène *celD* de *S. pneumoniae* codant pour une protéine EIIC<sup>cel</sup>. La recherche d'homologie des protéines CelC (EIIC<sup>cel</sup>) et CelB (EIIB<sup>cel</sup>) de *S. pneumoniae* a permis d'identifier les protéines codées par trois gènes distants (*OG1RF\_10745*, *10750* et *10751*) sur le chromosome chez *E. faecalis*. Ces gènes homologues sont présentés en **Figure 48**.

Il est intéressant de noter que les deux gènes homologues au gène *celB* de *S. pneumoniae* présentent 100 % d'identité au niveau de la séquence protéique résultante chez *E. faecalis*. La protéine codée par le gène *OG1RF\_10750* possède simplement 3 acides aminés supplémentaires (SSN) au niveau de l'extrémité C-terminale. Ces deux protéines sont vraisemblablement fonctionnellement interchangeables au vu d'un tel niveau de similarité. Comme ces deux gènes se trouvent au sein de deux structures transcriptionnelles différentes, ils sont potentiellement régulés différemment. Leur rôle exact dans la physiologie bactérienne n'a pas encore été déterminé, ces protéines doivent être impliquées dans l'import de molécules chimiquement apparentées mais légèrement différentes.



**Figure 48** : Recherche d'homologies des protéines impliquées dans le métabolisme du cellobiose de *S. pneumoniae* R6 chez *E. faecalis* OG1RF

Les gènes homologues entre les deux espèces sont reliés par des traits de couleur. L'identité (=), la similarité (+) ainsi que le pourcentage de couverture [xx%] lors de la recherche d'homologies sont indiqués pour chaque comparaison.

Comme explicité dans la publication n°3 présente dans ce manuscrit, le gène *OG1RF\_10751*,

codant pour une protéine EIIA, est modérément induit par la présence du cellobiose et du gentiobiose. Un mutant pour cette protéine est de plus incapable de se développer efficacement en utilisant comme source de carbone l'un de ces deux disaccharides. Il semble donc que cette protéine EIIA forme un complexe avec la protéine GenB pour former un PTS fonctionnel. Il est probable qu'elle constitue également un autre complexe avec les protéines codées par les gènes voisins *OGIRF\_10751* et *10753* codant respectivement des protéines EIIB et EIIC. Ce dernier PTS est vraisemblablement impliqué dans l'import du cellobiose. L'induction de l'expression des gènes *OGIRF\_10751-52* codant pour les protéines complémentaires du PTS gentiobiose était prédite pour être indépendante de l'activateur transcriptionnel GenR. En effet un mutant  $\Delta genR$  ne présente pas de défaut de croissance sur cellobiose alors qu'il est essentiel pour la croissance sur gentiobiose. Cette induction indépendante de GenR a été confirmée par des expériences de RNA-seq dont les résultats sont présentés dans le **Tableau 9**. Ces données montrent clairement que GenR agit comme un activateur sur l'expression de l'opéron *genBA* mais n'a aucun effet sur l'expression des autres gènes potentiellement impliqués dans le métabolisme du cellobiose.

L'expression du gène *OGIRF\_10745* apparaît très faible dans toutes les conditions testées. La protéine EIIB résultante n'est par conséquent vraisemblablement pas impliquée dans la formation du PTS gentiobiose avec la protéine GenB. Ce PTS gentiobiose fonctionnel est donc prédit pour être constitué de la protéine GenB (EIIC) ainsi que des protéines EIIB et EIIA codées respectivement par les gènes *OGIRF\_10750* et *10751*.

**Tableau 9 :** Expression des gènes impliqués dans le métabolisme du gentiobiose et/ou cellobiose chez *E. faecalis*

Numéro de gène		Description gène	Count per millions (cpm) normalisés				Log <sub>2</sub> (Fold-Change)		
OG1RF	V583		OG1RF			$\Delta$ genR	Cel/Glc (OG1RF)	Gen/Glc (OG1RF)	$\Delta$ genR/ OG1RF (Cel)
			Glucose (Glc)	Cellobiose (Cel)	Gentiobiose (Gen)	Cellobiose (Cel)			
10234	0291	6-Phospho- $\beta$ -glucosidase <i>genA</i>	21	2061	7822	10	6,6	8,6	-7,7
10235	0292	PTS EIIC gentiobiose <i>genB</i>	21	3607	6546	23	7,4	8,3	-7,3
10236	0293	Régulateur transcriptionnel <i>genR</i>	32	38	24	30	-	-	-
10745	1012	PTS EIIB cellobiose/gentiobiose ?	17	28	48	20	-	-	-
10746	1013	PTS EIIC cellobiose ?	98	231	474	175	1,2	2,3	-0,4
10747	1014	Protéine hypothétique	7	28	64	23	-	-	-
10748	1015	Protéine hypothétique	4	23	76	17	-	-	-
10749	1016	Protéine hypothétique	7	40	122	32	2,6	4,2	-0,4
10750	1017	PTS EIIB cellobiose/gentiobiose ?	104	308	562	277	1,6	2,4	-0,2
10751	1018	PTS EIIA cellobiose/gentiobiose	299	969	1292	863	1,7	2,1	-0,2
10752	1019	PTS EIIC cellobiose ?	987	5963	7531	5628	2,6	2,9	-0,1
10753	1020	6-Phospho- $\beta$ -glucosidase	96	1449	2657	1533	3,9	4,8	0,1
10754	1021	GNAT acetyl transferase	107	1917	2971	2011	4,2	4,8	0,1

Les Log<sub>2</sub>(Fold-Change) supérieurs à 1,5 ou inférieurs à -1,5 dans la condition testée sont mis en évidence respectivement en vert et en rouge.

Les gènes pour lesquels l'expression ne dépasse 100 cpm dans aucune condition testée sont considérés comme trop faiblement exprimés, les Log<sub>2</sub>(Fold-Change) ne sont pas calculés pour ces derniers.

# **Discussion et perspectives**

Nous avons étudié au cours de ce travail divers gènes de la bactérie *E. faecalis* dont l'expression a été déterminée comme induite en condition d'infection chez la souris. Les opérons qui ont retenu notre attention sont tous impliqués dans le métabolisme des sucres. La plus forte expression de ces gènes retrouvée lorsque la bactérie est présente au sein d'un hôte laisse supposer que le métabolisme énergétique occupe une place importante dans la persistance et le développement de la cellule bactérienne dans un tel environnement. En effet, l'acquisition et l'utilisation de sources d'énergie est une étape cruciale pour permettre la survie des microorganismes chimio-organo-hétérotrophes au sein d'un hôte, et par conséquent pour déclencher un processus infectieux <sup>(268)</sup>. Même si la production d'énergie utilisable pour la cellule est indispensable à la survie des microorganismes, il ne faut pas réduire le métabolisme à la simple production d'ATP. En effet, le métabolisme énergétique occupe une place centrale dans la physiologie bactérienne et permet d'obtenir, en plus de la production d'énergie, les précurseurs moléculaires permettant la réalisation de toutes les réactions anaboliques. Ces précurseurs sont représentés, chez les bactéries Gram-positives par 13 molécules qui, grâce aux différentes réactions d'anabolisme, permettent de produire toutes les molécules biologiques <sup>(269)</sup>. On retrouve par exemple au sein de ces précurseurs des molécules comme le FBP, le PEP, l'acetyl-CoA ou encore le ribose-5-P. En ce sens, le métabolisme régit directement ou indirectement les stratégies d'adaptation adoptées par les microorganismes.

Les entérocoques sont des microorganismes ubiquitaires rencontrés dans l'environnement ou associés à d'autres organismes, principalement animaux mais également parfois végétaux <sup>(35)</sup>. Ils occupent ainsi des niches écologiques variées et ont la nécessité de posséder un métabolisme diversifié et de bonnes capacités d'adaptation pour faire face à leur environnement changeant. Cette régulation minutieuse est d'autant plus importante quand ces bactéries se retrouvent au sein du tractus digestif, un environnement complexe où la disponibilité en sources de carbone est très variable et où la compétition inter-espèces est forte <sup>(268)</sup>. Mieux comprendre comment les microorganismes modulent leur métabolisme en fonction des diverses conditions rencontrées dans leur environnement peut représenter une première étape clef pour prédire leurs capacités ainsi que leurs stratégies d'adaptation et ainsi nous permettre d'ajuster en conséquence nos façons d'agir pour éviter les comportements à risque. A titre d'illustration, on peut citer le cas des ribotypes RT027 et RT078 hyper-virulents de *Clostridioides difficile* (anciennement *Clostridium difficile* <sup>(270)</sup>). Ces deux ribotypes épidémiques résultent d'une sensibilité augmentée des bactéries à la présence de tréhalose leur conférant ainsi un avantage compétitif en présence de faibles concentrations de tréhalose <sup>(271)</sup>. Ce disaccharide non réducteur a connu une augmentation importante de son utilisation dans le domaine agro-alimentaire depuis le début des années 2000 après son obtention du statut "GRAS" (Generally Recognized As Safe). L'utilisation abondante de ce sucre apparaît

directement corrélée à l'augmentation des signalements des ribotypes RT027 et RT078 au sein des épidémies de *C. difficile* en milieu hospitalier. Une meilleure compréhension des mécanismes moléculaires sous-jacent à la dispersion épidémique de cette bactérie pourrait ainsi permettre de remettre en question le statut GRAS du tréhalose et/ou d'envisager une utilisation plus raisonnée de son utilisation dans le milieu agro-alimentaire.

Lors de ce travail, nous avons dans un premier temps focalisé notre attention sur le métabolisme du maltose et des maltodextrines chez *E. faecalis*, reporté comme important pour cette bactérie en contexte d'infection chez un hôte mammifère.

## **I. Le métabolisme du maltose et des maltodextrines chez *E. faecalis***

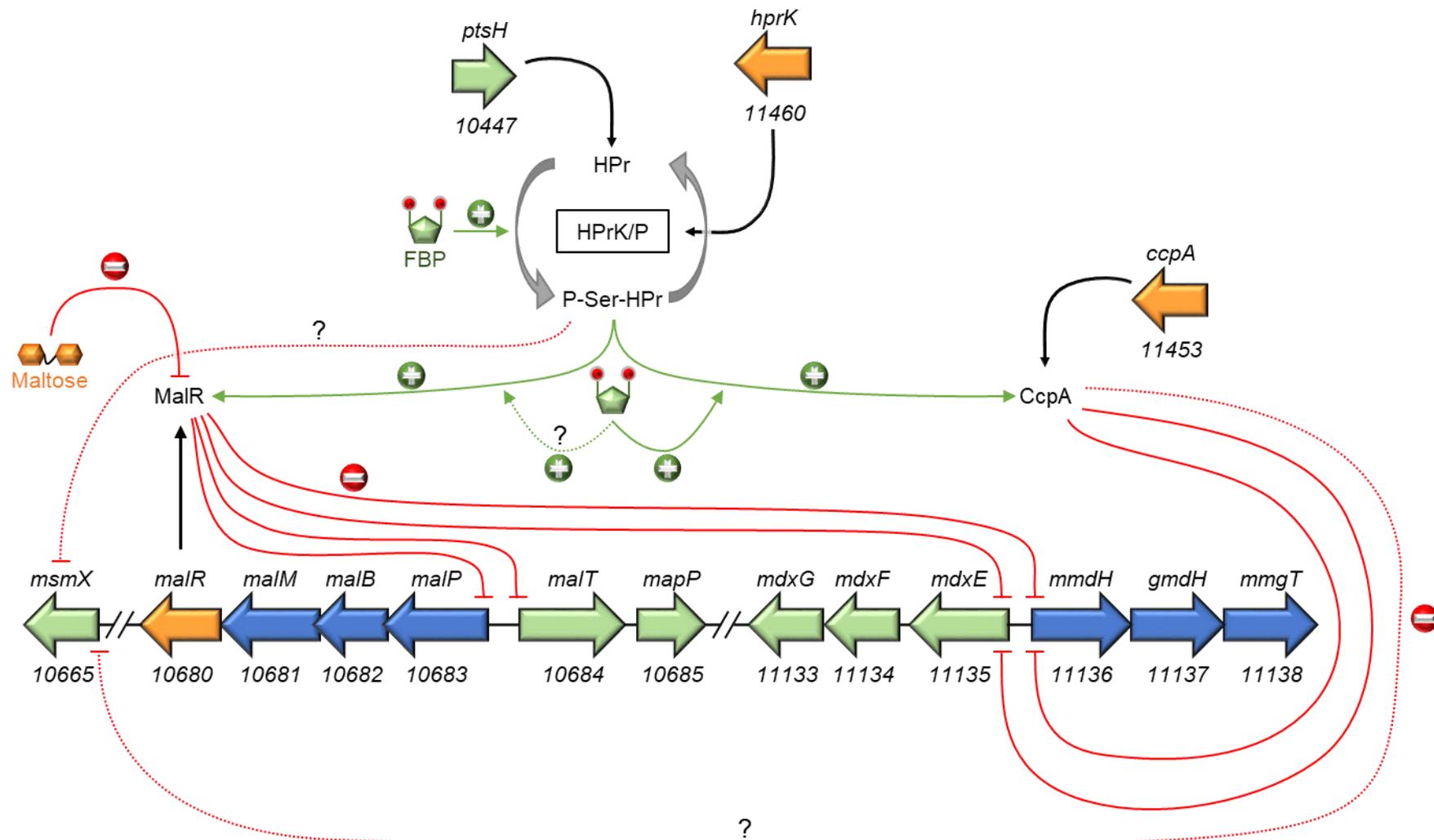
Les maltodextrines proviennent généralement de l'hydrolyse partielle de polymères plus longs comme l'amidon ou le glycogène. Ces grosses molécules de stockage proviennent majoritairement d'apports extérieurs et représentent une source importante de l'alimentation humaine principalement dans les sociétés sédentaires à l'agriculture développée (<sup>272</sup>). On retrouve en conséquence beaucoup d'enzymes impliquées dans la dégradation de l'amidon à la fois dans le génome humain (<sup>272</sup>) ou dans le microbiome humain (<sup>273</sup>). La première étape de dégradation de l'amidon est assurée par des  $\alpha$ -amylases et pullulanases salivaires de l'hôte mammifère qui vont libérer des chaînes plus courtes de polysaccharides présentant des liaisons glycosidiques  $\alpha$ -1,4 et  $\alpha$ -1,6 (<sup>273</sup>). On considère généralement que les courtes chaînes glucidiques résultantes sont efficacement absorbées par les cellules de l'hôte au niveau du petit intestin (<sup>274</sup>). Au niveau du colon, où sont majoritairement retrouvés les entérocoques, il est donc peu probable que des courtes chaînes de maltodextrines comme le maltose ou le maltotriose provenant de l'alimentation soient présentes en grande quantité. On sait en revanche qu'une proportion non négligeable de l'amidon (appelée amidon résistant) provenant de l'alimentation n'est pas dégradée durant le transit du bol alimentaire et atteint le colon (<sup>273</sup>). *E. faecalis* n'est pas capable de se développer en utilisant comme seule source de carbone le glycogène ou l'amidon car cette bactérie ne possède pas les enzymes indispensables (<sup>275</sup>). Cette bactérie peut en revanche, comme décrit dans la partie IV.C de l'introduction, dégrader les maltodextrines jusqu'à une longueur d'au moins 7 résidus glycosyl et certains oligosaccharides de glucoses contenant à la fois des liaisons  $\alpha$ -1,4 et/ou  $\alpha$ -1,6 (<sup>276,277</sup>). L'utilisation de ces sucres par *E. faecalis* nécessite par conséquent une hydrolyse préalable des macromolécules d'amidon ou de glycogène. Cette première dégradation de l'amidon résistant peut être assurée dans le colon par des amylases produites par d'autres microorganismes présents dans le tractus digestif comme *Ruminococcus bromii*, *Bacteroides thetaiotaomicron*,

*Roseburia inulinivorans* ou encore certains *Bifidobacterium* (278).

Le travail réalisé durant ce doctorat nous a permis de dresser un modèle de régulation des gènes impliqués dans le métabolisme du maltose et des maltodextrines chez *E. faecalis*. La synthèse de ces résultats est présentée sous la forme d'un schéma récapitulatif (**Figure 49**).

Nous avons pu déterminer que le maltose représentait indéniablement l'inducteur du système maltose/maltodextrines chez *E. faecalis* en levant la répression exercée par le répresseur MalR. Ce disaccharide occupe une place centrale dans le métabolisme du maltose et des maltodextrines de cette bactérie (visible sur la **Figure 34**). Toutes les maltodextrines métabolisées par *E. faecalis* aboutiront en effet à la formation intracellulaire de l'inducteur maltose via l'activité de la maltodextrine hydrolase MmdH. La régulation des opérons maltodextrines par le maltose via MalR n'est donc pas surprenante. Toutefois, l'affinité de 6xHis-MalR pour le maltose mesurée *in vitro* apparaît assez faible avec une constante de dissociation de l'ordre du mM. Les constantes de dissociation déterminées pour des régulateurs transcriptionnels LacI/GalR disponibles dans la littérature se situent généralement autour de quelques  $\mu\text{M}$  dans une gamme allant de 0,2 à 300  $\mu\text{M}$  (279-284). Même si ces constantes ont majoritairement été déterminées chez la bactérie gram négative *E. coli*, très éloignée phylogénétiquement d'*E. faecalis*, les régulateurs LacI/GalR semblent en règle général présenter plus d'affinité pour leur inducteur que MalR. Il est possible que cette faible affinité apparente de MalR pour son inducteur résulte d'une surestimation de cette constante à cause des conditions *in vitro*. En effet, la protéine utilisée présente une étiquette 6-histidines à son extrémité N-terminale. Or, toutes les étiquettes peuvent affecter le comportement de leur fusion protéique même si cette petite étiquette est généralement considérée comme ayant peu d'impact au niveau de la structure tridimensionnelle de la protéine originelle (285,286). Il est par conséquent possible que dans la cellule, MalR présente une meilleure affinité pour le maltose.

Cependant, une faible affinité de ce répresseur pour son inducteur pourrait également faire partie intégrante du système de régulation. Nous savons que les opérons maltose/maltodextrines sont soumis à la répression catabolique et réprimés indirectement par la présence de glucose via la protéine P-Ser-HPr. Cette répression est assurée par MalR indépendamment de la présence ou non de maltose. Pour expliquer cette observation, différentes expériences semblent indiquer que la présence de P-Ser-HPr augmente l'affinité de la protéine MalR *in vitro* pour son site opérateur, la rendant de fait plus difficile à séparer de l'ADN et renforçant ainsi la répression exercée sur les systèmes maltose/maltodextrines. En revanche, la présence de fortes concentrations de maltose inhibe toujours la fixation du complexe P-Ser-HPr/MalR sur l'ADN même si cet effet est plus limité sur ce complexe que sur la protéine MalR seule. Ainsi, si MalR présentait une affinité trop élevée pour le maltose, la répression exercée par ce régulateur pourrait être levée trop rapidement ne



**Figure 49 :** Modèle de régulation du métabolisme du maltose et des maltodextrines chez *E. faecalis*  
 Les courbes rouges représentent des inhibitions de transcription et/ou d'activité, les flèches vertes symbolisent des activations.  
 Les activités prédites mais non confirmées expérimentalement sont indiquées par des courbes en pointillé et sont marquées d'un point d'interrogation (?).

permettant pas à la cellule de prioriser le métabolisme du glucose par rapport à celui du maltose. Une plus faible affinité du répresseur pour cet inducteur implique alors un système d'import très efficace capable de rapidement faire augmenter la concentration intracellulaire en maltose afin de permettre une expression efficace des opérons impliqués dans ce métabolisme. L'augmentation intracellulaire en maltose est assurée, chez *E. faecalis*, par l'action combinée du PTS MalT et de la M6P phosphatase MapP. Les PTS sont connus pour être des transporteurs très efficaces présentant de fortes affinités pour leur ligand et capables de rapidement faire augmenter la concentration intracellulaire de leur substrat. D'après les données disponibles concernant l'import du glucose par des PTS chez *E. coli*, ce sucre peut atteindre des concentrations intracellulaires de plusieurs mM en quelques secondes <sup>(287)</sup>. Une affinité de MalR pour le maltose de l'ordre de mM ne semble par conséquent pas incompatible avec les conditions physiologiques potentiellement rencontrées lors du métabolisme du maltose et des maltodextrines chez *E. faecalis*.

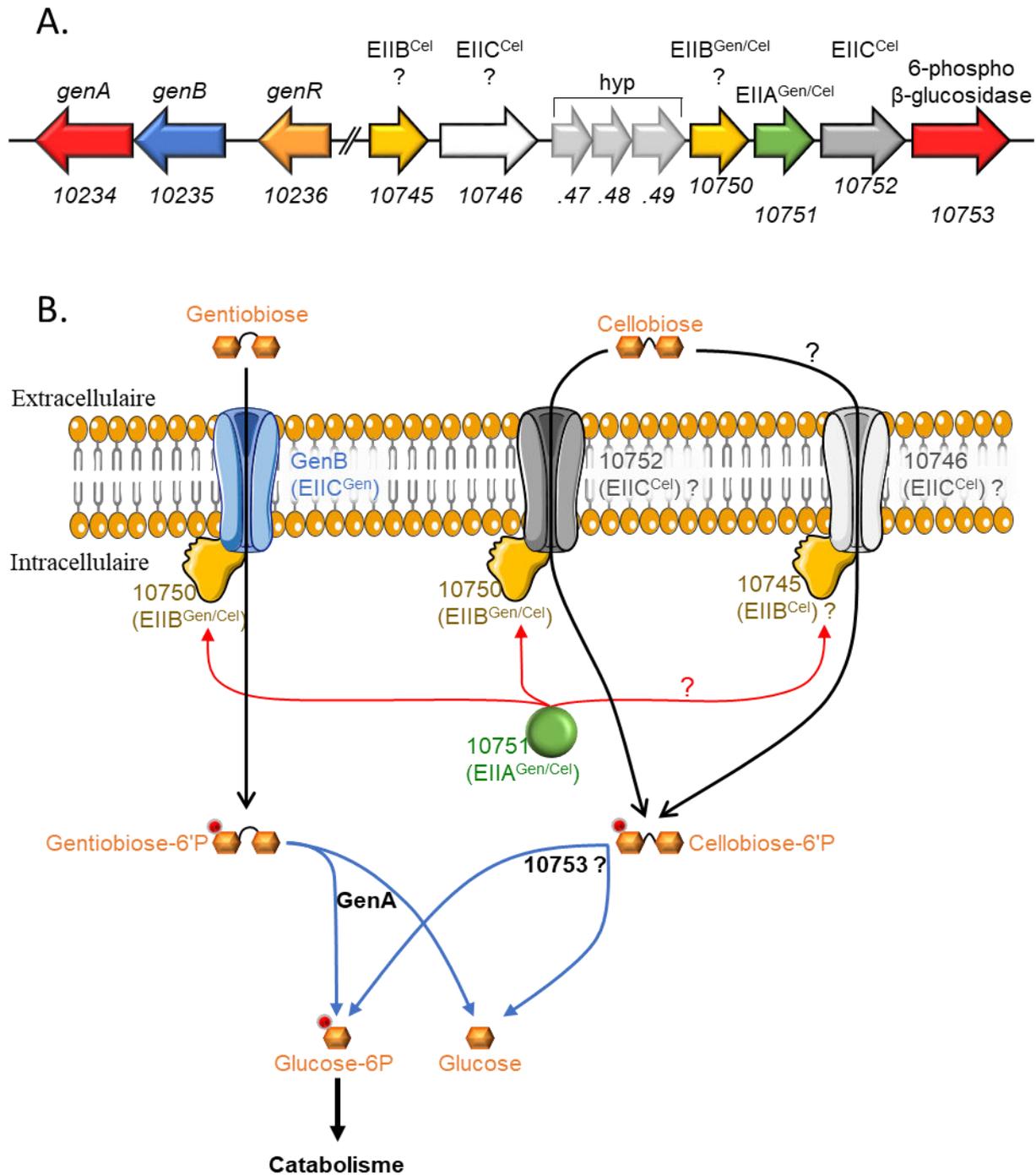
D'après toutes les données récoltées, la régulation du métabolisme du maltose et des maltodextrines chez *E. faecalis* semble reposer sur des équilibres dynamiques complexes médiés par des acteurs protéiques répondant à la concentration intracellulaire de certains sucres. Ce réseau de régulation permet ainsi à ces bactéries de rapidement s'adapter en fonction des nutriments présents dans leur environnement changeant, notamment par des phénomènes de répression catabolique. Une détermination de la structure tridimensionnelle du complexe MalR/P-Ser-HPr permettrait de caractériser l'interaction entre ces deux acteurs protéiques. L'identification des acides aminés mis en jeu ainsi que l'organisation locale de l'interface d'interaction pourrait permettre d'identifier de potentiels nouveaux partenaires d'interaction pour P-Ser-HPr. L'interaction de P-Ser-HPr avec des régulateurs transcriptionnels présentant un régulon réduit à quelques gènes spécifiques pourrait permettre d'expliquer une partie des régulations cataboliques indépendantes de CcpA encore incomprises. Cela renforcerait l'importance de P-Ser-HPr dans la régulation du métabolisme centrale des firmicutes lui allouant des effets de régulation globaux (via l'exclusion de l'inducteur ou comme co-facteur de CcpA) mais également des effets beaucoup plus ciblés (via l'interaction avec des régulateurs transcriptionnels très spécifiques comme MalR).

## **II. Le métabolisme du gentiobiose chez *E. faecalis***

Le second système sur lequel nous avons focalisé notre attention durant ce travail de thèse est représenté par les trois gènes *genA*, *genB* et *genR* chez *E. faecalis*. Comme décrit précédemment dans ce manuscrit, ces gènes sont indispensables pour permettre à cette bactérie d'utiliser le gentiobiose comme seule source de carbone. Le gentiobiose est un disaccharide assez peu répandu dans la nature mais on le retrouve cependant chez des végétaux dont certains fruits <sup>(288-291)</sup>. Ce sucre ne semble pas être utilisé chez ces espèces végétales comme une source de carbone mais

plutôt comme une molécule de signal. Le gentiobiose peut alors être présent dans l'alimentation humaine en fonction du régime alimentaire. L'utilisation de ce disaccharide par l'homme ou d'autres mammifères n'a été que très peu documentée mais on sait en revanche que les mammifères ne sont pas capables de dégrader et d'assimiler directement les fibres végétales. Ces dernières sont composées en grande partie de cellulose, un polymère de résidus glucosyl liés par des liaisons  $\beta$ -1,4, qui représente une partie importante de source de carbone contenu dans la matière végétale notamment au niveau de la paroi des cellules végétales. Aucune sous-unité de cellulose, y compris le disaccharide cellobiose représentant la plus courte possible, ne peut être assimilée par les cellules mammifères (<sup>274,292,293</sup>). Même les grands herbivores mammifères, dont le régime alimentaire se résume exclusivement à la consommation de matière végétale, ne peuvent pas dégrader ces sucres et nécessitent une flore microbienne capable de convertir pour eux ces sources de carbones en matière utilisable (<sup>292,293</sup>). On peut donc penser que le gentiobiose, assez proche du cellobiose ne peut pas non plus être assimilé par l'homme au niveau du petit intestin et pourrait donc atteindre le colon en cas de consommation (<sup>294</sup>). Ce disaccharide représenterait alors une potentielle source de carbone utilisable pour *E. faecalis*.

Les résultats présentés précédemment permettent d'établir un modèle présentant le métabolisme du gentiobiose chez *E. faecalis* (**Figure 50.B**). Ces résultats nous fournissent également des informations concernant le métabolisme du cellobiose chez cette bactérie (**Figure 50.A**). L'étude de la régulation de l'opéron *genBA* a abouti à un modèle de régulation assez simple présenté en **Figure 51**. Brièvement, l'expression de cette structure bicistronique est sous le contrôle direct de l'activateur transcriptionnel GenR qui nécessite la fixation d'un substrat phosphorylé pour être actif. L'expression de GenR semble, d'après les résultats du RNA-seq, faible et constitutive. En présence de glucose, cet opéron est réprimé de manière directe par le régulateur global CcpA. La régulation des gènes *OGIRF\_10745* à *10754* est encore inconnue mais semble indépendante du régulateur GenR.



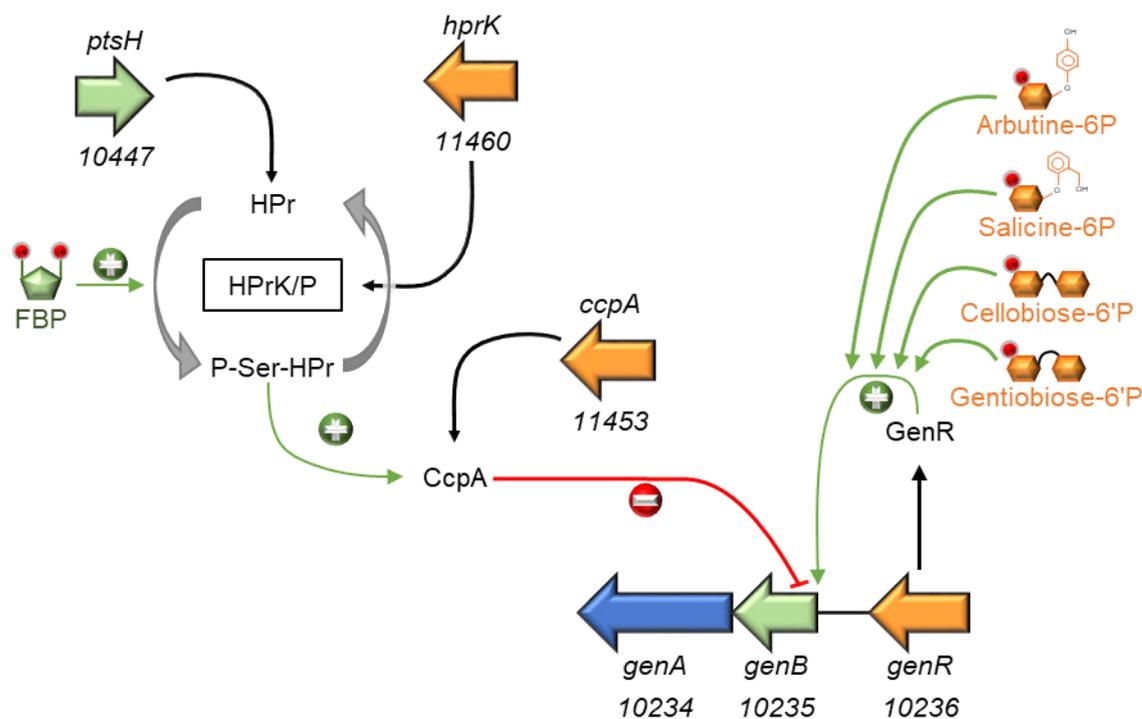
**Figure 50** : Schéma du métabolisme du gentiobiose et du cellobiose chez *E. faecalis*

**A.** Gènes impliqués dans le métabolisme du gentiobiose et suspectés être impliqués dans le métabolisme du cellobiose.

Lorsque les gènes n'ont officiellement pas de nom, leur fonction prédite est écrite au-dessus du gène.

**B.** Métabolisme du gentiobiose et hypothèses concernant le métabolisme du cellobiose.

Les flèches noires représentent des transports de molécules, les flèches bleues des modifications enzymatique et les flèches rouges les transferts de groupement phosphate entre les protéines EIIA et EIIB impliqués. La cascade complète de transferts de groupements phosphates des PTS n'est pas représentée dans un souci de clarté. Les activités prédites mais non vérifiées expérimentalement sont marquées d'un point d'interrogation (?)



**Figure 51** : Modèle de régulation de l'opéron *genBA* chez *E. faecalis* OG1RF

Les flèches vertes symbolisent des activations de transcription et/ou d'activité, les courbes rouges représentent des inhibitions.

La région formée par les gènes *OG1RF\_10745-54*, fonctionnellement liée aux gènes *gen*, comporte deux gènes codant des sous-unités EIIC de PTS. La détermination de la spécificité de substrat de ces deux EIIC permettrait vraisemblablement de mieux comprendre le rôle de cette région. Des résultats préliminaires semblent indiquer que les deux EIIC (*OG1RF\_10746* et *\_10752*) sont impliqués dans l'import de cello-oligosaccharides (glucose- $\beta$ -1,4-glucose). La différence entre ces deux protéines résiderait dans l'affinité pour les chaînes plus ou moins longues d'oligomère. Le domaine EIIC codé par le gène *OG1RF\_10752* semble impliqué, conformément à nos prédictions, dans l'import du cellobiose alors que la protéine EIIC, codée par le gène *OG1RF\_10476*, semble être essentielle pour le transport des oligomères plus longs comme le cellotétraose. Des expériences complémentaires restent cependant nécessaires pour confirmer ces premiers résultats.

L'opéron *genBA* est indéniablement essentiel pour permettre à la bactérie *E. faecalis* d'utiliser le gentiobiose comme source de carbone. Cependant ce sucre seul est vraisemblablement peu abondant dans la nature sous sa forme libre, utiliser ce disaccharide comme source de carbone ne devrait par conséquent fournir qu'un faible avantage sélectif à la bactérie. Le gentiobiose est en revanche plus régulièrement retrouvé au sein de l'amygdaline ou de la linustatine, deux hétérosides présents dans certains fruits ou graines. Le gentiobiose représente, dans ce cas, la partie glucidique des hétérosides (**Figure S.1**)<sup>(295)</sup>. L'opéron *genBA* est aussi induit en présence d'amygdaline ou de cellobiose. De plus, au vu de l'affinité de l'activateur transcriptionnel GenR, responsable de

l'induction de l'expression de l'opéron *genBA*, pour les hétérosides arbutine-6-P et salicine-6-P (**Figure S.1**), il est probable que l'arbutine et la salicine induisent également cet opéron. Ces données laisseraient supposer que l'opéron *genBA* pourrait posséder une spécificité de substrat plus large que le gentiobiose. Néanmoins, même si, comme le suggèrent les résultats de la publication n°3, la  $\beta$ -glucosidase GenA semble essentielle pour le métabolisme de l'amygdaline, aucun retard de croissance n'a été observé chez les mutants  $\Delta genA$  ou  $\Delta genB$  lors de croissance en présence de salicine ou d'esculine (données non montrées). De nombreuses  $\beta$ -glycosidases présentent des spécificités de substrats larges davantage dictées par le type de liaison que par la nature des monosaccharides (<sup>296,297</sup>). Il est donc envisageable que GenA puisse en réalité être impliquée dans le catabolisme de substrats encore non déterminés. La caractéristique retrouvée chez toutes les molécules connues pour être des substrats efficaces de GenA est la liaison  $\beta$ -1,6 présente entre les deux résidus glycosyl, ce type de liaison pourrait alors potentiellement représenter la spécificité de cette hydrolase.

Une telle liaison osidique  $\beta$ -1,6 nous offre au moins deux nouvelles pistes d'investigations. La première concerne la glycosylation des protéines de mammifères. Différents organismes peuvent en effet greffer des chaînes polysaccharidiques sur leurs protéines au niveau de certains acides aminés. Ces glycosylations peuvent assurer des rôles de régulation, de signalisation ou encore augmenter la résistance des protéines à la protéolyse (<sup>298</sup>). Diverse études ont d'ores et déjà montré qu'*E. faecalis* est capable de dégrader certaines glycosylations présentes sur des protéines de son hôte via des endoglycosidases sécrétées afin de libérer des oligosaccharides pouvant ensuite être utilisés comme source de carbone (<sup>299-302</sup>). Les liaisons osidiques  $\beta$ -1,6 ne sont pas majoritaires au niveau des glycosylations présentes sur les protéines de mammifères mais on peut toutefois en retrouver à la fois au niveau de O-glycosylations ou N-glycosylations. Les différents couples de monosaccharides impliqués dans ces liaisons sont répertoriés dans le **Tableau 10** (<sup>298,303</sup>). On peut alors émettre l'hypothèse d'une potentielle implication de l'opéron *genBA* dans le métabolisme d'oligosaccharides libérés suite à l'action des endoglycosidases. La confirmation expérimentale d'une telle activité de l'opéron *genBA* permettrait de confirmer son implication dans la virulence d'*E. faecalis* et expliquerait l'induction de son expression en contexte infectieux. Pour appuyer cette hypothèse, deux endoglycosidases EndoS et EndoF3 codées par les gènes *OGIRF\_10107* et *\_12167* semblent être plus exprimées lors d'une croissance sur gentiobiose (**Figure S2** de la publication n°3, pages 218 et 221 de ce manuscrit).

**Tableau 10 :** Couples de monosaccharides liés par des liaisons  $\beta$ -1,6 dans les glycosylations de protéines de mammifères

Disaccharides	Type de glycosylation
NAG <sup>a</sup> - $\beta$ -1,6-Mannose	N-glycosylations complexes
NAG- $\beta$ -1,6-GalNAc <sup>b</sup>	O-glycosylations
NAG- $\beta$ -1,6-GalNAc	O-glycosylations
GalNAc- $\beta$ -1,6-GalNAc	O-glycosylations

<sup>a</sup>NAG : *N*-acétylglucosamine

<sup>b</sup>GalNAc : *N*-acétylgalactosamine

Le métabolisme des  $\beta$ -glucosides a également été relié à la virulence de différentes espèces de Streptocoques (<sup>247,249,304,305</sup>). Chez ces dernières bactéries, le métabolisme des  $\beta$ -glucosides serait potentiellement impliqué dans la dégradation de glycoaminoglycanes (GAG), riches en liaisons  $\beta$ -1,4 et  $\beta$ -1,3, présents dans la matrice extracellulaire des mammifères. La bactérie *E. faecalis* étant incapable de se développer seule sur plusieurs GAG testés (données non montrées), il semble peu probable que l'opéron *genBA* soit impliqué dans le métabolisme des GAG. Il est cependant envisageable que les gènes *gen* aient un lien fonctionnel avec les endoglycosidases citées précédemment impliquées dans la dégradation des glycosylations présentes sur les protéines animales. Certains oligo-saccharides libérés suite à l'action de ces enzymes pourraient en effet être pris en charge par les gènes *gen* fournissant ainsi des sources de carbones utilisables à la bactérie lorsqu'elle se trouve au sein de son hôte. Mesurer l'impact des glycoprotéines sur l'expression de quelques gènes cibles dont les gènes *gen* permettrait de conforter l'implication de ces derniers dans ce processus avant d'envisager des études plus approfondies.

La seconde piste d'investigation pour comprendre le rôle de l'opéron *genBA* en condition d'infection concerne les oligosaccharides de poly- $\beta$ -1,6-N-acétylglucosamines (PNAG). Ces polymères, composés de monosaccharides de NAG liés par des liaisons  $\beta$ -1,6 représentent des constituants majeurs des exopolysaccharides (EPS) sécrétés au sein des biofilms de différentes bactéries pathogènes comme *Staphylococcus aureus*, *Staphylococcus carnosus*, *Staphylococcus epidermidis*, *Bordetella pertussis* ou encore *Escherichia coli* (<sup>306-308</sup>). Ces PNAG peuvent être partiellement déacétylés et ainsi être constitués d'un mélange de NAG et de glucosamine. Une étude récente a montré qu'*E. faecalis* était également capable de produire des EPS composés en grande partie de PNAG (<sup>309</sup>). Ces PNAG joueraient un rôle prépondérant dans la translocation d'*E. faecalis* à travers l'épithélium intestinal, représentant ainsi un facteur de virulence majeur. Si les oligomères de PNAG constituent effectivement un substrat efficace pour l'hydrolase

GenA, l'opéron *genBA* pourrait alors être impliqué dans le remodelage et/ou le recyclage de ces EPS. On peut également émettre l'hypothèse que cet opéron pourrait jouer un rôle dans l'utilisation de PNAG libérés à partir de biofilm d'autres bactéries au sein du tractus digestif. Ces différents rôles potentiels permettraient d'expliquer la forte induction observée pour cet opéron en condition d'infection dans un hôte murin. Chez la bactérie gram positive pathogène *Listeria monocytogenes*, une étude récente a positivement corrélié le métabolisme de  $\beta$ -glucosides dérivés de la chitine avec la production de biofilm (<sup>310</sup>). L'utilisation de ces sucres par cette bactérie constitue ainsi indirectement un facteur de virulence en favorisant la production de biofilm et augmentant ainsi la résistance de la cellule bactérienne aux molécules antimicrobiennes ainsi qu'aux défenses de son hôte. Chez *E. faecalis* V583, un opéron *ef\_0361-0362* (correspondant à l'opéron *OG1RF\_10250-10251*) prédit pour coder deux protéines avec des affinités pour la chitine dont une chitinase a également été rapporté comme induit lors de l'infection du péritoine chez la souris (<sup>20</sup>). Le métabolisme des  $\beta$ -glycosides apparaît ainsi positivement corrélié avec la colonisation d'un hôte mammifère chez différentes bactéries Gram-positives occupant des niches écologiques très différentes au sein de leur hôte comme *S. pneumoniae* (<sup>249,311</sup>), *L. monocytogenes* (<sup>310</sup>) ou encore *E. faecalis* (<sup>20</sup>). Mieux caractériser les métabolismes des  $\beta$ -glycosides chez les firmicutes s'avère ainsi d'un intérêt particulier pour mieux appréhender les relations étroites entre l'Homme et son microbiote. L'étude du biofilm représente ainsi une perspective d'étude intéressante. Cependant, la souche OG1RF d'*E. faecalis* est une souche produisant une quantité modérée de biofilm (<sup>312</sup>). L'utilisation de souches produisant plus de biofilm, en plus de la souche OG1RF, pourrait être indiquée pour étudier l'impact des différents gènes *gen* sur la production de biofilm d'*E. faecalis*.

Toutes ces pistes d'investigation représentent des perspectives intéressantes pour de futures études dans le but de mieux comprendre le caractère commensal mais également mieux appréhender le processus infectieux du pathogène opportuniste *E. faecalis*.

# **Références Bibliographiques**

1. Verneuil, N. *et al.* Implication of hypR in the virulence and oxidative stress response of *Enterococcus faecalis*. *FEMS Microbiol. Lett.* **252**, 137–141 (2005).
2. Verneuil, N. *et al.* Implication of (Mn)superoxide dismutase of *Enterococcus faecalis* in oxidative stress responses and survival inside macrophages. *Microbiol. Read. Engl.* **152**, 2579–2589 (2006).
3. La Carbona, S. *et al.* Comparative study of the physiological roles of three peroxidases (NADH peroxidase, Alkyl hydroperoxide reductase and Thiol peroxidase) in oxidative stress response, survival inside macrophages and virulence of *Enterococcus faecalis*. *Mol. Microbiol.* **66**, 1148–1163 (2007).
4. Bizzini, A., Zhao, C., Auffray, Y. & Hartke, A. The *Enterococcus faecalis* superoxide dismutase is essential for its tolerance to vancomycin and penicillin. *J. Antimicrob. Chemother.* **64**, 1196–1202 (2009).
5. Rana, N. F. *et al.* Redox balance via lactate dehydrogenase is important for multiple stress resistance and virulence in *Enterococcus faecalis*. *Infect. Immun.* **81**, 2662–2668 (2013).
6. Ladjouzi, R. *et al.* Loss of Antibiotic Tolerance in Sod-Deficient Mutants Is Dependent on the Energy Source and Arginine Catabolism in Enterococci. *J. Bacteriol.* **197**, 3283–3293 (2015).
7. Bizzini, A. *et al.* Glycerol is metabolized in a complex and strain-dependent manner in *Enterococcus faecalis*. *J. Bacteriol.* **192**, 779–785 (2010).
8. Sauvageot, N. *et al.* Aerobic glycerol dissimilation via the *Enterococcus faecalis* DhaK pathway depends on NADH oxidase and a phosphotransfer reaction from PEP to DhaK via EIIADha. *Microbiol. Read. Engl.* **158**, 2661–2666 (2012).
9. Le Jeune, A. *et al.* The extracytoplasmic function sigma factor SigV plays a key role in the original model of lysozyme resistance and virulence of *Enterococcus faecalis*. *PLoS One* **5**, e9658 (2010).
10. Benachour, A. *et al.* The lysozyme-induced peptidoglycan N-acetylglucosamine deacetylase PgdA (EF1843) is required for *Enterococcus faecalis* virulence. *J. Bacteriol.* **194**, 6066–6073 (2012).
11. Ladjouzi, R. *et al.* Analysis of the tolerance of pathogenic enterococci and *Staphylococcus aureus* to cell wall active antibiotics. *J. Antimicrob. Chemother.* **68**, 2083–2091 (2013).
12. Sinel, C. *et al.* Small RNAs in vancomycin-resistant *Enterococcus faecium* involved in daptomycin response and resistance. *Sci. Rep.* **7**, 11067 (2017).
13. Michaux, C. *et al.* Involvement of *Enterococcus faecalis* small RNAs in stress response and virulence. *Infect. Immun.* **82**, 3599–3611 (2014).
14. Michaux, C. *et al.* SlyA Is a Transcriptional Regulator Involved in the Virulence of *Enterococcus faecalis*. *Infect. Immun.* **79**, 2638–2645 (2011).

15. Michaux, C. *et al.* CspR, a cold shock RNA-binding protein involved in the long-term survival and the virulence of *Enterococcus faecalis*. *J. Bacteriol.* **194**, 6900–6908 (2012).
16. Martini, C. *et al.* The Polyamine N-Acetyltransferase-Like Enzyme PmvE Plays a Role in the Virulence of *Enterococcus faecalis*. *Infect. Immun.* **83**, 364–371 (2015).
17. Reffuveille, F., Leneveu, C., Chevalier, S., Auffray, Y. & Rincé, A. Lipoproteins of *Enterococcus faecalis*: bioinformatic identification, expression analysis and relation to virulence. *Microbiol. Read. Engl.* **157**, 3001–3013 (2011).
18. Reffuveille, F. *et al.* The prolipoprotein diacylglyceryl transferase (Lgt) of *Enterococcus faecalis* contributes to virulence. *Microbiol. Read. Engl.* **158**, 816–825 (2012).
19. Cacaci, M. *et al.* Expression profiling in a mammalian host reveals the strong induction of genes encoding LysM domain-containing proteins in *Enterococcus faecium*. *Sci. Rep.* **8**, 12412 (2018).
20. Muller, C. *et al.* The Intraperitoneal Transcriptome of the Opportunistic Pathogen *Enterococcus faecalis* in Mice. *PLoS One* **10**, e0126143 (2015).
21. Sauvageot, N. *et al.* *Enterococcus faecalis* uses a PTS permease and a host colonization-related ABC transporter for maltodextrin uptake. *J. Bacteriol.* (2017). doi:10.1128/JB.00878-16
22. MacCallum, W. G. & Hastings, T. W. A Case of Acute Endocarditis Caused by *Micrococcus Zymogenes* (nov. Spec.), with a Description of the Microorganism. *J. Exp. Med.* **4**, 521–534 (1899).
23. Thiercelin, E. Sur un Diplocoque saprophyte de l'intestin susceptible de devenir pathogène. in *Comptes rendus des séances de la société de biologie et de ses filiales* 269–271 (Masson et Cie, 1899).
24. Schleifer, K. H. & Kilpper-Bälz, R. Transfer of *Streptococcus faecalis* and *Streptococcus faecium* to the Genus *Enterococcus* nom. rev. as *Enterococcus faecalis* comb. nov. and *Enterococcus faecium* comb. nov. *Int. J. Syst. Evol. Microbiol.* **34**, 31–34 (1984).
25. Lebreton, F., Willems, R. J. L. & Gilmore, M. S. *Enterococcus* Diversity, Origins in Nature, and Gut Colonization. in *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection* (eds. Gilmore, M. S., Clewell, D. B., Ike, Y. & Shankar, N.) (Massachusetts Eye and Ear Infirmary, 2014).
26. Euzéby, J. P. List of Bacterial Names with Standing in Nomenclature: a folder available on the Internet. *Int. J. Syst. Bacteriol.* **47**, 590–592 (1997).
27. Parte, A. C. LPSN—list of prokaryotic names with standing in nomenclature. *Nucleic Acids Res.* **42**, D613–D616 (2014).
28. Parte, A. C. LPSN - List of Prokaryotic names with Standing in Nomenclature (bacterio.net), 20 years on. *Int. J. Syst. Evol. Microbiol.* **68**, 1825–1829 (2018).

29. Kristich, C. J., Rice, L. B. & Arias, C. A. Enterococcal Infection—Treatment and Antibiotic Resistance. in *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection* (eds. Gilmore, M. S., Clewell, D. B., Ike, Y. & Shankar, N.) (Massachusetts Eye and Ear Infirmary, 2014).
30. Ramsey, M., Hartke, A. & Huycke, M. The Physiology and Metabolism of Enterococci. in *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection* (eds. Gilmore, M. S., Clewell, D. B., Ike, Y. & Shankar, N.) (Massachusetts Eye and Ear Infirmary, 2014).
31. Mubarak, Z. & Soraya, C. The acid tolerance response and pH adaptation of *Enterococcus faecalis* in extract of lime *Citrus aurantiifolia* from Aceh Indonesia. *F1000Research* **7**, 287 (2018).
32. Bale, M. J., Bennett, P. M., Beringer, J. E. & Hinton, M. The survival of bacteria exposed to desiccation on surfaces associated with farm buildings. *J. Appl. Bacteriol.* **75**, 519–528 (1993).
33. Kramer, A., Schwebke, I. & Kampf, G. How long do nosocomial pathogens persist on inanimate surfaces? A systematic review. *BMC Infect. Dis.* **6**, 130 (2006).
34. Hollenbeck, B. L. & Rice, L. B. Intrinsic and acquired resistance mechanisms in enterococcus. *Virulence* **3**, 421–569 (2012).
35. Lebreton, F. *et al.* Tracing the Enterococci from Paleozoic Origins to the Hospital. *Cell* **169**, 849–861.e13 (2017).
36. Schloissnig, S. *et al.* Genomic variation landscape of the human gut microbiome. *Nature* **493**, 45–50 (2013).
37. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection.* (Massachusetts Eye and Ear Infirmary, 2014).
38. Cox, M. A. *et al.* Short-chain fatty acids act as antiinflammatory mediators by regulating prostaglandin E(2) and cytokines. *World J. Gastroenterol.* **15**, 5549–5557 (2009).
39. den Besten, G. *et al.* The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J. Lipid Res.* **54**, 2325–2340 (2013).
40. Mishra, A. K., Kumar, S. S. & Ghosh, A. R. Probiotic *Enterococcus faecalis* AG5 effectively assimilates cholesterol and produces fatty acids including propionate. *FEMS Microbiol. Lett.* **366**, (2019).
41. Cheng, D., Song, J., Xie, M. & Song, D. The bidirectional relationship between host physiology and microbiota and health benefits of probiotics: A review. *Trends Food Sci. Technol.* **91**, 426–435 (2019).
42. Guzman Prieto, A. M. *et al.* Global Emergence and Dissemination of Enterococci as Nosocomial Pathogens: Attack of the Clones? *Front. Microbiol.* **7**, (2016).
43. Agudelo Higueta, N. I. & Huycke, M. M. Enterococcal Disease, Epidemiology, and

Implications for Treatment. in *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection* (eds. Gilmore, M. S., Clewell, D. B., Ike, Y. & Shankar, N.) (Massachusetts Eye and Ear Infirmary, 2014).

44. Maugat, S. *et al.* Bilan 2001-2017 des signalements externes d'infections nosocomiales. Part des signalements impliquant une bactérie multirésistante, hautement résistante-émergente ou un *Clostridium difficile*. *Bull Epidemiol Hebd* (2018).
45. Tacconelli, E. *et al.* Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect. Dis.* **18**, 318–327 (2018).
46. Garsin, D. A. *et al.* Pathogenesis and Models of Enterococcal Infection. in *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection* (eds. Gilmore, M. S., Clewell, D. B., Ike, Y. & Shankar, N.) (Massachusetts Eye and Ear Infirmary, 2014).
47. Kao, P. H. N. & Kline, K. A. Jekyll and Mr. Hide: How *Enterococcus faecalis* Subverts the Host Immune Response to Cause Infection. *J. Mol. Biol.* (2019). doi:10.1016/j.jmb.2019.05.030
48. Hancock, L. E., Murray, B. E. & Sillanpää, J. Enterococcal Cell Wall Components and Structures. in *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection* (eds. Gilmore, M. S., Clewell, D. B., Ike, Y. & Shankar, N.) (Massachusetts Eye and Ear Infirmary, 2014).
49. Lin, E. C. Glycerol dissimilation and its regulation in bacteria. *Annu. Rev. Microbiol.* **30**, 535–578 (1976).
50. Sarantinopoulos, P., Kalantzopoulos, G. & Tsakalidou, E. Citrate metabolism by *Enterococcus faecalis* FAIR-E 229. *Appl. Environ. Microbiol.* **67**, 5482–5487 (2001).
51. Espariz, M. *et al.* Identification of malic and soluble oxaloacetate decarboxylase enzymes in *Enterococcus faecalis*. *FEBS J.* **278**, 2140–2151 (2011).
52. Deibel, R. H. Utilization of arginine as an energy source for the growth of *Streptococcus faecalis*. *J. Bacteriol.* **87**, 988–992 (1964).
53. Madigan, M. T., Martinko, J. M. & Brock, T. D. *Brock biology of microorganisms*. (Pearson, Prentice Hall, 2006).
54. Kim, B. H. & Gadd, G. M. *Bacterial Physiology and Metabolism*. (Cambridge University Press, 2008).
55. Kandler, O. Carbohydrate metabolism in lactic acid bacteria. *Antonie Van Leeuwenhoek* **49**, 209–224 (1983).
56. Buron-Moles, G., Chailyan, A., Dolejs, I., Forster, J. & Mikš, M. H. Uncovering carbohydrate metabolism through a genotype-phenotype association study of 56 lactic acid bacteria

- genomes. *Appl. Microbiol. Biotechnol.* **103**, 3135–3152 (2019).
57. Yun, J.-S., Wee, Y.-J. & Ryu, H.-W. Production of optically pure l(+)-lactic acid from various carbohydrates by batch fermentation of *Enterococcus faecalis* RKY1. *Enzyme Microb. Technol.* **33**, 416–423 (2003).
58. Subramanian, M. R., Talluri, S. & Christopher, L. P. Production of lactic acid using a new homofermentative *Enterococcus faecalis* isolate. *Microb. Biotechnol.* **8**, 221–229 (2015).
59. Giraffa, G. Enterococci from foods. *FEMS Microbiol. Rev.* **26**, 163–171 (2002).
60. M'hir, S., Minervini, F., Di Cagno, R., Chammem, N. & Hamdi, M. Technological, functional and safety aspects of enterococci in fermented vegetable products: a mini-review. *Ann. Microbiol.* **62**, 469–481 (2012).
61. Doi, Y. Glycerol metabolism and its regulation in lactic acid bacteria. *Appl. Microbiol. Biotechnol.* (2019). doi:10.1007/s00253-019-09830-y
62. Lolkema, J. S., Poolman, B. & Konings, W. N. Role of scalar protons in metabolic energy generation in lactic acid bacteria. *J. Bioenerg. Biomembr.* **27**, 467–473 (1995).
63. Ritchey, T. W. & Seeley, H. W. Cytochromes in *Streptococcus faecalis* var. *zymogenes* grown in a haematin-containing medium. *J. Gen. Microbiol.* **85**, 220–228 (1974).
64. Pritchard, G. G. & Wimpenny, J. W. Cytochrome formation, oxygen-induced proton extrusion and respiratory activity in *Streptococcus faecalis* var. *zymogenes* grown in the presence of haematin. *J. Gen. Microbiol.* **104**, 15–22 (1978).
65. Aue, B. J. & Deiel, R. H. Fumarate reductase activity of *Streptococcus faecalis*. *J. Bacteriol.* **93**, 1770–1776 (1967).
66. Huycke, M. M. *et al.* Extracellular superoxide production by *Enterococcus faecalis* requires demethylmenaquinone and is attenuated by functional terminal quinol oxidases. *Mol. Microbiol.* **42**, 729–740 (2001).
67. Schmidt, H. L., Stöcklein, W., Danzer, J., Kirch, P. & Limbach, B. Isolation and properties of an H<sub>2</sub>O-forming NADH oxidase from *Streptococcus faecalis*. *Eur. J. Biochem.* **156**, 149–155 (1986).
68. Riebel, B. R., Gibbs, P. R., Wellborn, W. B. & Bommarius, A. S. Cofactor Regeneration of NAD<sup>+</sup> from NADH: Novel Water-Forming NADH Oxidases. *Adv. Synth. Catal.* **344**, 1156–1168 (2002).
69. Palmer, K. L., van Schaik, W., Willems, R. J. L. & Gilmore, M. S. Enterococcal Genomics. in *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection* (eds. Gilmore, M. S., Clewell, D. B., Ike, Y. & Shankar, N.) (Massachusetts Eye and Ear Infirmary, 2014).
70. Bourgoigne, A. *et al.* Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF. *Genome Biol.* **9**, R110 (2008).

71. Oliver, D. R., Brown, B. L. & Clewell, D. B. Analysis of plasmid deoxyribonucleic acid in a cariogenic strain of *Streptococcus faecalis*: an approach to identifying genetic determinants on cryptic plasmids. *J. Bacteriol.* **130**, 759–765 (1977).
72. Gold, O. G., Jordan, H. V. & van Houte, J. The prevalence of enterococci in the human mouth and their pathogenicity in animal models. *Arch. Oral Biol.* **20**, 473-IN15 (1975).
73. Dunny, G. M., Brown, B. L. & Clewell, D. B. Induced cell aggregation and mating in *Streptococcus faecalis*: evidence for a bacterial sex pheromone. *Proc. Natl. Acad. Sci. U. S. A.* **75**, 3479–3483 (1978).
74. Busch, W. & Saier, M. H. The transporter classification (TC) system, 2002. *Crit. Rev. Biochem. Mol. Biol.* **37**, 287–337 (2002).
75. Elbourne, L. D. H. *et al.* Microbial Solute Transporters. in *Reference Module in Biomedical Sciences* (Elsevier, 2019). doi:10.1016/B978-0-12-801238-3.66200-1
76. Goswitz, V. C. & Brooker, R. J. Structural features of the uniporter/symporter/antiporter superfamily. *Protein Sci.* **4**, 534–537 (1995).
77. Rea, R., Angelis, M. G. D. & Baschetti, M. G. Models for Facilitated Transport Membranes: A Review. *Membranes* **9**, (2019).
78. Beis, K. Structural basis for the mechanism of ABC transporters. *Biochem. Soc. Trans.* **43**, 889–893 (2015).
79. Davidson, A. L., Dassa, E., Orelle, C. & Chen, J. Structure, Function, and Evolution of Bacterial ATP-Binding Cassette Systems. *Microbiol. Mol. Biol. Rev. MMBR* **72**, 317–364 (2008).
80. Dassa, E. & Bouige, P. The ABC of ABCS: a phylogenetic and functional classification of ABC systems in living organisms. *Res Microbiol* **152**, 211–229 (2001).
81. Zolnerciks, J. K., Andress, E. J., Nicolaou, M. & Linton, K. J. Structure of ABC transporters. *Essays Biochem.* **50**, 43–61 (2011).
82. Chen, J., Lu, G., Lin, J., Davidson, A. L. & Quioco, F. A. A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle. *Mol. Cell* **12**, 651–661 (2003).
83. Schneider, E. ABC transporters catalyzing carbohydrate uptake. *Res. Microbiol.* **152**, 303–310 (2001).
84. van der Heide, T. & Poolman, B. ABC transporters: one, two or four extracytoplasmic substrate-binding sites? *EMBO Rep.* **3**, 938–943 (2002).
85. Cui, J. & Davidson, A. L. ABC solute importers in bacteria. *Essays Biochem.* **50**, 85–99 (2011).
86. Petronilli, V. & Ames, G. F. Binding protein-independent histidine permease mutants. Uncoupling of ATP hydrolysis from transmembrane signaling. *J. Biol. Chem.* **266**, 16293–16296 (1991).

87. Davidson, A. L., Shuman, H. A. & Nikaido, H. Mechanism of maltose transport in *Escherichia coli*: transmembrane signaling by periplasmic binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 2360–2364 (1992).
88. Karlskås, I. L., Saleihan, Z., Holo, H., Mathiesen, G. & Eijsink, V. G. H. EF0176 and EF0177 from *Enterococcus faecalis* V583 are substrate-binding lipoproteins involved in ABC transporter mediated ribonucleoside uptake. *Microbiol. Read. Engl.* **161**, 754–764 (2015).
89. Kilstrup, M., Hammer, K., Ruhdal Jensen, P. & Martinussen, J. Nucleotide metabolism and its control in lactic acid bacteria. *FEMS Microbiol. Rev.* **29**, 555–590 (2005).
90. Henderson, R. K., Fendler, K. & Poolman, B. Coupling efficiency of secondary active transporters. *Curr. Opin. Biotechnol.* **58**, 62–71 (2018).
91. Forrest, L. R., Krämer, R. & Ziegler, C. The structural basis of secondary active transport mechanisms. *Biochim. Biophys. Acta BBA - Bioenerg.* **1807**, 167–188 (2011).
92. Boudker, O. & Verdon, G. Structural perspectives on secondary active transporters. *Trends Pharmacol. Sci.* **31**, 418–426 (2010).
93. Blancato, V. S., Magni, C. & Lolkema, J. S. Functional characterization and Me ion specificity of a Ca-citrate transporter from *Enterococcus faecalis*. *FEBS J.* **273**, 5121–5130 (2006).
94. Drew, D. & Boudker, O. Shared Molecular Mechanisms of Membrane Transporters. *Annu. Rev. Biochem.* **85**, 543–572 (2016).
95. Kim, J. W. *et al.* Structural insights into the elevator-like mechanism of the sodium/citrate symporter CitS. *Sci. Rep.* **7**, 2548 (2017).
96. Pickl, A., Johnsen, U. & Schönheit, P. Fructose degradation in the haloarchaeon *Haloferax volcanii* involves a bacterial type phosphoenolpyruvate-dependent phosphotransferase system, fructose-1-phosphate kinase, and class II fructose-1,6-bisphosphate aldolase. *J. Bacteriol.* **194**, 3088–3097 (2012).
97. Cai, L. *et al.* Analysis of the transcriptional regulator GlpR, promoter elements, and posttranscriptional processing involved in fructose-induced activation of the phosphoenolpyruvate-dependent sugar phosphotransferase system in *Haloferax mediterranei*. *Appl. Environ. Microbiol.* **80**, 1430–1440 (2014).
98. Deutscher, J. *et al.* The bacterial phosphoenolpyruvate:carbohydrate phosphotransferase system: regulation by protein phosphorylation and phosphorylation-dependent protein-protein interactions. *Microbiol. Mol. Biol. Rev. MMBR* **78**, 231–256 (2014).
99. Deutscher, J., Francke, C. & Postma, P. W. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol. Mol. Biol. Rev. MMBR* **70**, 939–1031 (2006).
100. Alpert, C. A., Frank, R., Stüber, K., Deutscher, J. & Hengstenberg, W.

- Phosphoenolpyruvate-dependent protein kinase enzyme I of *Streptococcus faecalis*: purification and properties of the enzyme and characterization of its active center. *Biochemistry (Mosc.)* **24**, 959–964 (1985).
101. Postma, P. W., Lengeler, J. W. & Jacobson, G. R. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol. Rev.* **57**, 543–594 (1993).
  102. Hu, K.-Y. & Saier, M. H. Phylogeny of phosphoryl transfer proteins of the phosphoenolpyruvate-dependent sugar-transporting phosphotransferase system. *Res. Microbiol.* **153**, 405–415 (2002).
  103. Lee, B. R., Lecchi, P., Pannell, L., Jaffe, H. & Peterkofsky, A. Identification of the N-terminal domain of enzyme I of the *Escherichia coli* phosphoenolpyruvate:sugar phosphotransferase system produced by proteolytic digestion. *Arch. Biochem. Biophys.* **312**, 121–124 (1994).
  104. van Nuland, N. A. *et al.* The high-resolution structure of the histidine-containing phosphocarrier protein HPr from *Escherichia coli* determined by restrained molecular dynamics from nuclear magnetic resonance nuclear Overhauser effect data. *J. Mol. Biol.* **237**, 544–559 (1994).
  105. Madej, T. *et al.* MMDB and VAST+: tracking structural similarities between macromolecular complexes. *Nucleic Acids Res.* **42**, D297–303 (2014).
  106. Barabote, R. D. & Saier, M. H. Comparative genomic analyses of the bacterial phosphotransferase system. *Microbiol. Mol. Biol. Rev. MMBR* **69**, 608–634 (2005).
  107. Morabbi Heravi, K. & Altenbuchner, J. Cross-talk among transporters of the phosphoenolpyruvate-dependent phosphotransferase system in *Bacillus subtilis*. *J. Bacteriol.* (2018). doi:10.1128/JB.00213-18
  108. Le Breton, Y., Pichereau, V., Sauvageot, N., Auffray, Y. & Rincé, A. Maltose utilization in *Enterococcus faecalis*. *J. Appl. Microbiol.* **98**, 806–813 (2005).
  109. Cases, I., Velázquez, F. & de Lorenzo, V. The ancestral role of the phosphoenolpyruvate–carbohydrate phosphotransferase system (PTS) as exposed by comparative genomics. *Res. Microbiol.* **158**, 666–670 (2007).
  110. Galinier, A. [Carbon catabolite repression or how bacteria choose their favorite sugars]. *Med. Sci. MS* **34**, 531–539 (2018).
  111. Deutscher, J. & Saier, M. H. ATP-dependent protein kinase-catalyzed phosphorylation of a seryl residue in HPr, a phosphate carrier protein of the phosphotransferase system in *Streptococcus pyogenes*. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 6790–6794 (1983).
  112. Kravanja, M. *et al.* The *hprK* gene of *Enterococcus faecalis* encodes a novel bifunctional enzyme: the HPr kinase/phosphatase. *Mol. Microbiol.* **31**, 59–66 (1999).

113. Nessler, S. The bacterial HPr kinase/phosphorylase: a new type of Ser/Thr kinase as antimicrobial target. *Biochim. Biophys. Acta* **1754**, 126–131 (2005).
114. Fieulaine, S. *et al.* X-ray structure of HPr kinase: a bacterial protein kinase with a P-loop nucleotide-binding domain. *EMBO J.* **20**, 3917–3927 (2001).
115. Monedero, V. *et al.* Mutations lowering the phosphatase activity of HPr kinase/phosphatase switch off carbon metabolism. *EMBO J.* **20**, 3928–3937 (2001).
116. Deutscher, J. & Engelmann, R. Purification and characterization of an ATP-dependent protein kinase from *Streptococcus faecalis*. *FEMS Microbiol. Lett.* **23**, 157–162 (1984).
117. Mijakovic, I. *et al.* Pyrophosphate-producing protein dephosphorylation by HPr kinase/phosphorylase: a relic of early life? *Proc. Natl. Acad. Sci. U. S. A.* **99**, 13442–13447 (2002).
118. Deutscher, J., Kessler, U., Alpert, C. A. & Hengstenberg, W. Bacterial phosphoenolpyruvate-dependent phosphotransferase system: P-Ser-HPr and its possible regulatory function? *Biochemistry (Mosc.)* **23**, 4455–4460 (1984).
119. Jones, B. E. *et al.* Binding of the catabolite repressor protein CcpA to its DNA target is regulated by phosphorylation of its corepressor HPr. *J. Biol. Chem.* **272**, 26530–26535 (1997).
120. Moreno, M. S., Schneider, B. L., Maile, R. R., Weyler, W. & Saier, M. H. Catabolite repression mediated by the CcpA protein in *Bacillus subtilis*: novel modes of regulation revealed by whole-genome analyses. *Mol. Microbiol.* **39**, 1366–1381 (2001).
121. Opsata, M., Nes, I. F. & Holo, H. Class IIa bacteriocin resistance in *Enterococcus faecalis* V583: The mannose PTS operon mediates global transcriptional responses. *BMC Microbiol.* **10**, 224 (2010).
122. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
123. Homburg, C. *et al.* Inducer exclusion in Firmicutes: Insights into the regulation of a carbohydrate ATP binding cassette transporter from *Lactobacillus casei* BL23 by the signal transducing protein P-Ser46-HPr. *Mol. Microbiol.* (2017). doi:10.1111/mmi.13680
124. Wuttge, S., Licht, A., Timachi, M. H., Bordignon, E. & Schneider, E. Mode of Interaction of the Signal-Transducing Protein EIIA(Glc) with the Maltose ABC Transporter in the Process of Inducer Exclusion. *Biochemistry (Mosc.)* **55**, 5442–5452 (2016).
125. Galinier, A. & Deutscher, J. Sophisticated regulation of transcriptional factors by the bacterial phosphoenolpyruvate: Sugar phosphotransferase system. *J. Mol. Biol.* (2017). doi:10.1016/j.jmb.2017.02.006
126. Ye, J. J., Reizer, J., Cui, X. & Saier, M. H. Inhibition of the phosphoenolpyruvate:lactose phosphotransferase system and activation of a cytoplasmic sugar-phosphate phosphatase in

- Lactococcus lactis by ATP-dependent metabolite-activated phosphorylation of serine 46 in the phosphocarrier protein HPr. *J. Biol. Chem.* **269**, 11837–11844 (1994).
127. Ye, J.-J., Minarcik, J. & Saier, M. H. Inducer Expulsion and the Occurrence of an HPr(Ser-P)-Activated Sugar-Phosphate Phosphatase in Enterococcus Faecalis and Streptococcus Pyogenes. *Microbiology* **142**, 585–592 (1996).
  128. Reizer, J., Novotny, M. J., Panos, C. & Saier, M. H. Mechanism of inducer expulsion in Streptococcus pyogenes: a two-step process activated by ATP. *J. Bacteriol.* **156**, 354–361 (1983).
  129. Kearns, D. B., Cook, G. M. & Russell, J. B. Inducer expulsion is not a determinant of diauxic growth in Streptococcus bovis. *Curr. Microbiol.* **32**, 221–224 (1996).
  130. Cook, G. M., Ye, J. J., Russell, J. B. & Saier, M. H. Properties of two sugar phosphate phosphatases from Streptococcus bovis and their potential involvement in inducer expulsion. *J. Bacteriol.* **177**, 7007–7009 (1995).
  131. Dossonnet, V. *et al.* Phosphorylation of HPr by the bifunctional HPr Kinase/P-ser-HPr phosphatase from Lactobacillus casei controls catabolite repression and inducer exclusion but not inducer expulsion. *J. Bacteriol.* **182**, 2582–2590 (2000).
  132. Ye, J. J. & Saier, M. H. Purification and characterization of a small membrane-associated sugar phosphate phosphatase that is allosterically activated by HPr(Ser(P)) of the phosphotransferase system in Lactococcus lactis. *J. Biol. Chem.* **270**, 16740–16744 (1995).
  133. Stülke, J., Arnaud, M., Rapoport, G. & Martin-Verstraete, I. PRD — a protein domain involved in PTS-dependent induction and carbon catabolite repression of catabolic operons in bacteria. *Mol. Microbiol.* **28**, 865–874 (1998).
  134. Deutscher, J. The mechanisms of carbon catabolite repression in bacteria. *Curr. Opin. Microbiol.* **11**, 87–93 (2008).
  135. Graille, M. *et al.* Activation of the LicT transcriptional antiterminator involves a domain swing/lock mechanism provoking massive structural changes. *J. Biol. Chem.* **280**, 14780–14789 (2005).
  136. Joyet, P. *et al.* Transcription regulators controlled by interaction with enzyme IIB components of the phosphoenolpyruvate:sugar phosphotransferase system. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1834**, 1415–1424 (2013).
  137. Leboeuf, C., Leblanc, L., Auffray, Y. & Hartke, A. Characterization of the ccpA Gene of Enterococcus faecalis: Identification of Starvation-Inducible Proteins Regulated by CcpA. *J. Bacteriol.* **182**, 5799–5806 (2000).
  138. Roy, D. J., Casabon, I., Vaillancourt, K., Huot, J. L. & Vadeboncoeur, C. Streptococci and lactococci synthesize large amounts of HPr(Ser-P)(His~P). *Can. J. Microbiol.* **54**, 941–949

- (2008).
139. Weickert, M. J. & Adhya, S. A family of bacterial regulators homologous to Gal and Lac repressors. *J. Biol. Chem.* **267**, 15869–15874 (1992).
  140. Francke, C., Kerkhoven, R., Wels, M. & Siezen, R. J. A generic approach to identify Transcription Factor-specific operator motifs; Inferences for LacI-family mediated regulation in *Lactobacillus plantarum* WCFS1. *BMC Genomics* **9**, 145 (2008).
  141. Swint-Kruse, L. & Matthews, K. S. Allosterity in the LacI/GalR family: variations on a theme. *Curr. Opin. Microbiol.* **12**, 129–137 (2009).
  142. Lewis, M. *et al.* Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* **271**, 1247–1254 (1996).
  143. Huffman, J. L. & Brennan, R. G. Prokaryotic transcription regulators: more than just the helix-turn-helix motif. *Curr. Opin. Struct. Biol.* **12**, 98–106 (2002).
  144. Brennan, R. G. & Matthews, B. W. The helix-turn-helix DNA binding motif. *J. Biol. Chem.* **264**, 1903–1906 (1989).
  145. Felder, C. B., Graul, R. C., Lee, A. Y., Merkle, H. P. & Sadee, W. The Venus flytrap of periplasmic binding proteins: an ancient protein module present in multiple drug receptors. *AAPS PharmSci* **1**, E2 (1999).
  146. Wilson, C. J., Zhan, H., Swint-Kruse, L. & Matthews, K. S. The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. *Cell. Mol. Life Sci. CMLS* **64**, 3–16 (2007).
  147. Bell, C. E. & Lewis, M. The Lac repressor: a second generation of structural and functional studies. *Curr. Opin. Struct. Biol.* **11**, 19–25 (2001).
  148. Tungtur, S., Parente, D. J. & Swint-Kruse, L. Functionally Important Positions Can Comprise the Majority of a Protein's Architecture. *Proteins* **79**, 1589–1608 (2011).
  149. Sørensen, K. I. & Hove-Jensen, B. Ribose catabolism of *Escherichia coli*: characterization of the rpiB gene encoding ribose phosphate isomerase B and of the rpiR gene, which is involved in regulation of rpiB expression. *J. Bacteriol.* **178**, 1003–1011 (1996).
  150. Poulsen, T. S., Chang, Y. Y. & Hove-Jensen, B. D-Allose catabolism of *Escherichia coli*: involvement of alsI and regulation of als regulon expression by allose and ribose. *J. Bacteriol.* **181**, 7126–7130 (1999).
  151. Jaeger, T. & Mayer, C. The transcriptional factors MurR and catabolite activator protein regulate N-acetylmuramic acid catabolism in *Escherichia coli*. *J. Bacteriol.* **190**, 6598–6608 (2008).
  152. Kohler, P. R. A., Choong, E.-L. & Rossbach, S. The RpiR-like repressor IolR regulates inositol catabolism in *Sinorhizobium meliloti*. *J. Bacteriol.* **193**, 5155–5163 (2011).

153. Guzmán, K. *et al.* Characterization of the gene cluster involved in allantoate catabolism and its transcriptional regulation by the RpiR-type repressor HpxU in *Klebsiella pneumoniae*. *Int. Microbiol. Off. J. Span. Soc. Microbiol.* **16**, 165–176 (2013).
154. Bateman, A. The SIS domain: a phosphosugar-binding domain. *Trends Biochem. Sci.* **24**, 94–95 (1999).
155. Teplyakov, A., Obmolova, G., Badet-Denisot, M. A., Badet, B. & Polikarpov, I. Involvement of the C terminus in intramolecular nitrogen channeling in glucosamine 6-phosphate synthase: evidence from a 1.6 Å crystal structure of the isomerase domain. *Struct. Lond. Engl.* **1993** **6**, 1047–1055 (1998).
156. King, M. W. *Integrative Medical Biochemistry examination and board review.* (2014).
157. Boos, W. & Shuman, H. Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism, and regulation. *Microbiol. Mol. Biol. Rev. MMBR* **62**, 204–229 (1998).
158. Randall-Hazelbauer, L. & Schwartz, M. Isolation of the Bacteriophage Lambda Receptor from *Escherichia coli*. *J. Bacteriol.* **116**, 1436–1446 (1973).
159. Freundlieb, S., Ehmman, U. & Boos, W. Facilitated diffusion of p-nitrophenyl- $\alpha$ -D-maltohexaoside through the outer membrane of *Escherichia coli*. Characterization of LamB as a specific and saturable channel for maltooligosaccharides. *J. Biol. Chem.* **263**, 314–320 (1988).
160. Nikaido, H. & Vaara, M. Molecular basis of bacterial outer membrane permeability. *Microbiol. Rev.* **49**, 1–32 (1985).
161. Schneider, E., Freundlieb, S., Tapio, S. & Boos, W. Molecular characterization of the MalT-dependent periplasmic  $\alpha$ -amylase of *Escherichia coli* encoded by malS. *J. Biol. Chem.* **267**, 5148–5154 (1992).
162. Hengge, R. & Boos, W. Maltose and lactose transport in *Escherichia coli*. Examples of two different types of concentrative transport systems. *Biochim. Biophys. Acta* **737**, 443–478 (1983).
163. Chen, J. Molecular mechanism of the *Escherichia coli* maltose transporter. *Curr. Opin. Struct. Biol.* **23**, 492–498 (2013).
164. Shuman, H. A. & Silhavy, T. J. Identification of the malK gene product. A peripheral membrane component of the *Escherichia coli* maltose transport system. *J. Biol. Chem.* **256**, 560–562 (1981).
165. Morbach, S., Tebbe, S. & Schneider, E. The ATP-binding cassette (ABC) transporter for maltose/maltodextrins of *Salmonella typhimurium*. Characterization of the ATPase activity associated with the purified MalK subunit. *J. Biol. Chem.* **268**, 18617–18621 (1993).
166. Kellermann, O. & Szmelcman, S. Active transport of maltose in *Escherichia coli* K12.

- Involvement of a 'periplasmic' maltose binding protein. *Eur. J. Biochem.* **47**, 139–149 (1974).
167. Ferenci, T., Brass, J. & Boos, W. The role of the periplasmic maltose-binding protein and the outer-membrane phage lambda receptor in maltodextrin transport of *Escherichia coli*. *Biochem. Soc. Trans.* **8**, 680–681 (1980).
168. Schwartz, M. & Hofnung, M. [Maltodextrin phosphorylase of *Escherichia coli*]. *Eur. J. Biochem.* **2**, 132–145 (1967).
169. Becker, S., Palm, D. & Schinzel, R. Dissecting differential binding in the forward and reverse reaction of *Escherichia coli* maltodextrin phosphorylase using 2-deoxyglucosyl substrates. *J. Biol. Chem.* **269**, 2485–2490 (1994).
170. O'Reilly, M., Watson, K. A., Schinzel, R., Palm, D. & Johnson, L. N. Oligosaccharide substrate binding in *Escherichia coli* maltodextrin phosphorylase. *Nat. Struct. Biol.* **4**, 405–412 (1997).
171. Dippel, R. & Boos, W. The maltodextrin system of *Escherichia coli*: metabolism and transport. *J. Bacteriol.* **187**, 8322–8331 (2005).
172. Decker, K. *et al.* Maltose and maltotriose can be formed endogenously in *Escherichia coli* from glucose and glucose-1-phosphate independently of enzymes of the maltose system. *J. Bacteriol.* **175**, 5655–5665 (1993).
173. Lu, M. & Kleckner, N. Molecular cloning and characterization of the *pgm* gene encoding phosphoglucomutase of *Escherichia coli*. *J. Bacteriol.* **176**, 5847–5851 (1994).
174. Reyes, M., Treptow, N. A. & Shuman, H. A. Transport of p-nitrophenyl-alpha-maltoside by the maltose transport system of *Escherichia coli* and its subsequent hydrolysis by a cytoplasmic alpha-maltosidase. *J. Bacteriol.* **165**, 918–922 (1986).
175. Tapio, S., Yeh, F., Shuman, H. A. & Boos, W. The *malZ* gene of *Escherichia coli*, a member of the maltose regulon, encodes a maltodextrin glucosidase. *J. Biol. Chem.* **266**, 19450–19458 (1991).
176. Peist, R., Schneider-Fresenius, C. & Boos, W. The MalT-dependent and *malZ*-encoded maltodextrin glucosidase of *Escherichia coli* can be converted into a dextrinyltransferase by a single mutation. *J. Biol. Chem.* **271**, 10681–10689 (1996).
177. Palmer, T. N., Ryman, B. E. & Whelan, W. J. The action pattern of amyloamylase from *Escherichia coli*. *Eur. J. Biochem.* **69**, 105–115 (1976).
178. Pugsley, A. P. & Dubreuil, C. Molecular characterization of *malQ*, the structural gene for the *Escherichia coli* enzyme amyloamylase. *Mol. Microbiol.* **2**, 473–479 (1988).
179. Park, J.-T. *et al.* Role of maltose enzymes in glycogen synthesis by *Escherichia coli*. *J. Bacteriol.* **193**, 2517–2526 (2011).
180. Nguyen, D. H. D. *et al.* Characterization of the Transglycosylation Reaction of 4- $\alpha$ -

- Glucanotransferase (MalQ) and Its Role in Glycogen Breakdown in Escherichia coli. *J. Microbiol. Biotechnol.* (2019). doi:10.4014/jmb.1811.11051
181. Rousset, J. P., Gilson, E. & Hofnung, M. malM, a new gene of the maltose regulon in Escherichia coli K12. II. Mutations affecting the signal peptide of the MalM protein. *J. Mol. Biol.* **191**, 313–320 (1986).
  182. Gilson, E., Rousset, J.-P., Charbit, A., Perrin, D. & Hofnung, M. malM, a new gene of the maltose regulon in Escherichia coli K12: I. malM is the last gene of the malK-lamB operon and encodes a periplasmic protein. *J. Mol. Biol.* **191**, 303–311 (1986).
  183. Schneider, E., Francoz, E. & Dassa, E. Completion of the nucleotide sequence of the ‘maltose B’ region in Salmonella typhimurium: the high conservation of the malM gene suggests a selected physiological role for its product. *Biochim. Biophys. Acta* **1129**, 223–227 (1992).
  184. Hatfield, D., Hofnung, M. & Schwartz, M. Genetic analysis of the maltose A region in Escherichia coli. *J. Bacteriol.* **98**, 559–567 (1969).
  185. Débarbouillé, M., Shuman, H. A., Silhavy, T. J. & Schwartz, M. Dominant constitutive mutations in malT, the positive regulator gene of the maltose regulon in Escherichia coli. *J. Mol. Biol.* **124**, 359–371 (1978).
  186. Raibaud, O., Vidal-Ingigliardi, D. & Richet, E. A complex nucleoprotein structure involved in activation of transcription of two divergent Escherichia coli promoters. *J. Mol. Biol.* **205**, 471–485 (1989).
  187. Raibaud, O. & Richet, E. Maltotriose is the inducer of the maltose regulon of Escherichia coli. *J. Bacteriol.* **169**, 3059–3061 (1987).
  188. Richet, E. & Raibaud, O. MalT, the regulatory protein of the Escherichia coli maltose system, is an ATP-dependent transcriptional activator. *EMBO J.* **8**, 981–987 (1989).
  189. Boos, W. & Böhm, A. Learning new tricks from an old dog: MalT of the Escherichia coli maltose system is part of a complex regulatory network. *Trends Genet. TIG* **16**, 404–409 (2000).
  190. Joly, N., Böhm, A., Boos, W. & Richet, E. MalK, the ATP-binding cassette component of the Escherichia coli maltodextrin transporter, inhibits the transcriptional activator malt by antagonizing inducer binding. *J. Biol. Chem.* **279**, 33123–33130 (2004).
  191. Panagiotidis, C. H., Boos, W. & Shuman, H. A. The ATP-binding cassette subunit of the maltose transporter MalK antagonizes MalT, the activator of the Escherichia coli mal regulon. *Mol. Microbiol.* **30**, 535–546 (1998).
  192. Zdych, E., Peist, R., Reidl, J. & Boos, W. MalY of Escherichia coli is an enzyme with the activity of a beta C-S lyase (cystathionase). *J. Bacteriol.* **177**, 5035–5039 (1995).

193. Schlegel, A., Danot, O., Richet, E., Ferenci, T. & Boos, W. The N terminus of the Escherichia coli transcription activator MalT is the domain of interaction with MalY. *J. Bacteriol.* **184**, 3069–3077 (2002).
194. Joly, N., Danot, O., Schlegel, A., Boos, W. & Richet, E. The Aes protein directly controls the activity of MalT, the central transcriptional activator of the Escherichia coli maltose regulon. *J. Biol. Chem.* **277**, 16606–16613 (2002).
195. Decker, K., Plumbridge, J. & Boos, W. Negative transcriptional regulation of a positive regulator: the expression of malT, encoding the transcriptional activator of the maltose regulon of Escherichia coli, is negatively controlled by Mlc. *Mol. Microbiol.* **27**, 381–390 (1998).
196. Schiefner, A. *et al.* The crystal structure of Mlc, a global regulator of sugar metabolism in Escherichia coli. *J. Biol. Chem.* **280**, 29073–29079 (2005).
197. Lee, S. J., Boos, W., Bouché, J. P. & Plumbridge, J. Signal transduction between a membrane-bound transporter, PtsG, and a soluble transcription factor, Mlc, of Escherichia coli. *EMBO J.* **19**, 5353–5361 (2000).
198. Nam, T. W. *et al.* The Escherichia coli glucose transporter enzyme IICB(Glc) recruits the global repressor Mlc. *EMBO J.* **20**, 491–498 (2001).
199. Nam, T.-W. *et al.* Analyses of Mlc-IIBGlc interaction and a plausible molecular mechanism of Mlc inactivation by membrane sequestration. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 3751–3756 (2008).
200. Chapon, C. Role of the catabolite activator protein in the maltose regulon of Escherichia coli. *J. Bacteriol.* **150**, 722–729 (1982).
201. Chapon, C. & Kolb, A. Action of CAP on the malT promoter in vitro. *J. Bacteriol.* **156**, 1135–1143 (1983).
202. Zubay, G., Schwartz, D. & Beckwith, J. Mechanism of activation of catabolite-sensitive genes: a positive control system. *Proc. Natl. Acad. Sci. U. S. A.* **66**, 104–110 (1970).
203. De Crombrughe, E. *et al.* Role of cyclic adenosine 3',5'-monophosphate and the cyclic adenosine 3',5'-monophosphate receptor protein in the initiation of lac transcription. *J. Biol. Chem.* **246**, 7343–7348 (1971).
204. Kühnau, S., Reyes, M., Sievertsen, A., Shuman, H. A. & Boos, W. The activities of the Escherichia coli MalK protein in maltose transport, regulation, and inducer exclusion can be separated by mutations. *J. Bacteriol.* **173**, 2180–2186 (1991).
205. Webb, A. J., Homer, K. A. & Hosie, A. H. F. A phosphoenolpyruvate-dependent phosphotransferase system is the principal maltose transporter in Streptococcus mutans. *J. Bacteriol.* **189**, 3322–3327 (2007).
206. Shelburne, S. A. *et al.* Molecular characterization of group A Streptococcus maltodextrin

- catabolism and its role in pharyngitis. *Mol. Microbiol.* **69**, 436–452 (2008).
207. Bidossi, A. *et al.* A functional genomics approach to establish the complement of carbohydrate transporters in *Streptococcus pneumoniae*. *PLoS One* **7**, e33320 (2012).
208. Mokhtari, A. *et al.* *Enterococcus faecalis* utilizes maltose by connecting two incompatible metabolic routes via a novel maltose 6'-phosphate phosphatase (MapP). *Mol. Microbiol.* **88**, 234–253 (2013).
209. Puyet, A. & Espinosa, M. Structure of the maltodextrin-uptake locus of *Streptococcus pneumoniae*. Correlation to the *Escherichia coli* maltose regulon. *J. Mol. Biol.* **230**, 800–811 (1993).
210. Abbott, D. W. *et al.* The molecular basis of glycogen breakdown and transport in *Streptococcus pneumoniae*. *Mol. Microbiol.* **77**, 183–199 (2010).
211. Webb, A. J., Homer, K. A. & Hosie, A. H. F. Two closely related ABC transporters in *Streptococcus mutans* are involved in disaccharide and/or oligosaccharide uptake. *J. Bacteriol.* **190**, 168–178 (2008).
212. Shelburne, S. A. *et al.* MalE of Group A *Streptococcus* Participates in the Rapid Transport of Maltotriose and Longer Maltodextrins. *J. Bacteriol.* **189**, 2610–2617 (2007).
213. Martin, S. A. & Russell, J. B. Transport and phosphorylation of disaccharides by the ruminal bacterium *Streptococcus bovis*. *Appl. Environ. Microbiol.* **53**, 2388–2393 (1987).
214. Sato, Y., Okamoto-Shibayama, K. & Azuma, T. The malQ gene is essential for starch metabolism in *Streptococcus mutans*. *J. Oral Microbiol.* **5**, (2013).
215. Medda, S. & Smith, E. E. A radioisotope method for assays of amyloamylase and D-enzyme. *Anal. Biochem.* **138**, 354–359 (1984).
216. Hardy, G. G., Caimano, M. J. & Yother, J. Capsule biosynthesis and basic metabolism in *Streptococcus pneumoniae* are linked through the cellular phosphoglucomutase. *J. Bacteriol.* **182**, 1854–1863 (2000).
217. Hardy, G. G., Magee, A. D., Ventura, C. L., Caimano, M. J. & Yother, J. Essential role for cellular phosphoglucomutase in virulence of type 3 *Streptococcus pneumoniae*. *Infect. Immun.* **69**, 2309–2317 (2001).
218. Hytönen, J., Haataja, S. & Finne, J. *Streptococcus pyogenes* glycoprotein-binding streptadhesin activity is mediated by a surface-associated carbohydrate-degrading enzyme, pullulanase. *Infect. Immun.* **71**, 784–793 (2003).
219. Saburi, W., Mori, H., Saito, S., Okuyama, M. & Kimura, A. Structural elements in dextran glucosidase responsible for high specificity to long chain substrate. *Biochim. Biophys. Acta* **1764**, 688–698 (2006).
220. Saburi, W. *et al.* Crystallization and preliminary X-ray analysis of *Streptococcus mutans*

- dextran glucosidase. *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun.* **63**, 774–776 (2007).
221. Satoh, E., Uchimura, T., Kudo, T. & Komagata, K. Purification, characterization, and nucleotide sequence of an intracellular maltotriose-producing alpha-amylase from *Streptococcus bovis* 148. *Appl. Environ. Microbiol.* **63**, 4941–4944 (1997).
222. Simpson, C. L. & Russell, R. R. Intracellular alpha-amylase of *Streptococcus mutans*. *J. Bacteriol.* **180**, 4711–4717 (1998).
223. Shelburne, S. A. *et al.* Contribution of AmyA, an extracellular alpha-glucan degrading enzyme, to group A streptococcal host-pathogen interaction. *Mol. Microbiol.* **74**, 159–174 (2009).
224. Sumbly, P., Whitney, A. R., Graviss, E. A., DeLeo, F. R. & Musser, J. M. Genome-wide analysis of group A streptococci reveals a mutation that modulates global phenotype and disease specificity. *PLoS Pathog.* **2**, e5 (2006).
225. Miyoshi-Akiyama, T. *et al.* Use of DNA arrays to identify a mutation in the negative regulator, *csrR*, responsible for the high virulence of a naturally occurring type M3 group A streptococcus clinical isolate. *J. Infect. Dis.* **193**, 1677–1684 (2006).
226. Nieto, C., Espinosa, M. & Puyet, A. The maltose/maltodextrin regulon of *Streptococcus pneumoniae* Differential promoter regulation by the transcriptional repressor MalR. *J. Biol. Chem.* **272**, 30860–30865 (1997).
227. Nieto, C., Puyet, A. & Espinosa, M. MalR-mediated Regulation of the *Streptococcus pneumoniae* malMP Operon at PromoterPM: influence of a proximal divergent promoter region and competition between MalR and RNA polymerase proteins. *J. Biol. Chem.* **276**, 14946–14954 (2001).
228. Jobe, A. & Bourgeois, S. lac repressor-operator interaction: VI. The natural inducer of the lac operon. *J. Mol. Biol.* **69**, 397–408 (1972).
229. Shelburne, S. A. *et al.* Regulation of polysaccharide utilization contributes to the persistence of group A streptococcus in the oropharynx. *Infect. Immun.* **75**, 2981–2990 (2007).
230. Shelburne, S. A. *et al.* Niche-specific contribution to streptococcal virulence of a MalR-regulated carbohydrate binding protein. *Mol. Microbiol.* **81**, 500–514 (2011).
231. Afzal, M., Shafeeq, S., Manzoor, I. & Kuipers, O. P. Maltose-Dependent Transcriptional Regulation of the mal Regulon by MalR in *Streptococcus pneumoniae*. *PloS One* **10**, e0127579 (2015).
232. Ben-Zvi, R. & Schramm, M. A Phosphoglucomutase Specific for  $\beta$ -Glucose 1-Phosphate. *J. Biol. Chem.* **236**, 2186–2189 (1961).
233. Qian, N., Stanley, G. A., Hahn-Hägerdal, B. & Rådström, P. Purification and

- characterization of two phosphoglucomutases from *Lactococcus lactis* subsp. *lactis* and their regulation in maltose- and glucose-utilizing cells. *J. Bacteriol.* **176**, 5304–5311 (1994).
234. Qian, N., Stanley, G. A., Bunte, A. & Rådström, P. Product formation and phosphoglucomutase activities in *Lactococcus lactis*: cloning and characterization of a novel phosphoglucomutase gene. *Microbiol. Read. Engl.* **143** ( Pt 3), 855–865 (1997).
235. Nakai, H. *et al.* The maltodextrin transport system and metabolism in *Lactobacillus acidophilus* NCFM and production of novel alpha-glucosides through reverse phosphorolysis by maltose phosphorylase. *FEBS J.* **276**, 7353–7365 (2009).
236. Joyet, P. *et al.* The enzymes required for maltodextrin catabolism in *Enterococcus faecalis* exhibit several novel activities. *Appl. Environ. Microbiol.* (2017). doi:10.1128/AEM.00038-17
237. Zhang, X. *et al.* A LacI-Family Regulator Activates Maltodextrin Metabolism of *Enterococcus faecium*. *PLoS ONE* **8**, e72285 (2013).
238. Palmer, R. E. & Anderson, R. L. Metabolism of gentiobiose in *Aerobacter aerogenes*. *J. Bacteriol.* **112**, 1316–1320 (1972).
239. Kricker, M. & Hall, B. G. Directed evolution of cellobiose utilization in *Escherichia coli* K12. *Mol. Biol. Evol.* **1**, 171–182 (1984).
240. Kachroo, A. H., Kancherla, A. K., Singh, N. S., Varshney, U. & Mahadevan, S. Mutations that alter the regulation of the *chb* operon of *Escherichia coli* allow utilization of cellobiose. *Mol. Microbiol.* **66**, 1382–1395 (2007).
241. Hall, B. G., Yokoyama, S. & Calhoun, D. H. Role of cryptic genes in microbial evolution. *Mol. Biol. Evol.* **1**, 109–124 (1983).
242. Tamburini, E. & Mastromei, G. Do bacterial cryptic genes really exist? *Res. Microbiol.* **151**, 179–182 (2000).
243. Joseph, A. M., Sonowal, R. & Mahadevan, S. A comparative study of the evolution of cellobiose utilization in *Escherichia coli* and *Shigella sonnei*. *Arch. Microbiol.* **199**, 247–257 (2017).
244. Thompson, J., Ruvinov, S. B., Freedberg, D. I. & Hall, B. G. Cellobiose-6-phosphate hydrolase (CelF) of *Escherichia coli*: characterization and assignment to the unusual family 4 of glycosylhydrolases. *J. Bacteriol.* **181**, 7339–7345 (1999).
245. Verma, S. C. & Mahadevan, S. The *chbG* gene of the chitobiose (*chb*) operon of *Escherichia coli* encodes a chitooligosaccharide deacetylase. *J. Bacteriol.* **194**, 4959–4971 (2012).
246. Plumbridge, J. & Pellegrini, O. Expression of the chitobiose operon of *Escherichia coli* is regulated by three transcription factors: NagC, ChbR and CAP. *Mol. Microbiol.* **52**, 437–449 (2004).
247. Shafeeq, S., Kuipers, O. P. & Kloosterman, T. G. Cellobiose-mediated gene expression in

- Streptococcus pneumoniae*: a repressor function of the novel GntR-type regulator BguR. *PLoS One* **8**, e57586 (2013).
248. McKessar, S. J. & Hakenbeck, R. The two-component regulatory system TCS08 is involved in cellobiose metabolism of *Streptococcus pneumoniae* R6. *J. Bacteriol.* **189**, 1342–1350 (2007).
249. Shafeeq, S., Kloosterman, T. G. & Kuipers, O. P. CelR-mediated activation of the cellobiose-utilization gene cluster in *Streptococcus pneumoniae*. *Microbiol. Read. Engl.* **157**, 2854–2861 (2011).
250. Zeng, L. & Burne, R. A. Transcriptional regulation of the cellobiose operon of *Streptococcus mutans*. *J. Bacteriol.* **191**, 2153–2162 (2009).
251. Bertani, G. Studies on lysogenesis : the mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* **62**, 293–300 (1951).
252. Leenhouts, K. *et al.* A general system for generating unlabelled gene replacements in bacterial chromosomes. *Mol. Gen. Genet. MGG* **253**, 217–224 (1996).
253. Studier, F. W. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**, 113–130 (1986).
254. Terzaghi, B. E. & Sandine, W. E. Improved medium for lactic streptococci and their bacteriophages. *Appl. Microbiol.* **29**, 807–813 (1975).
255. Arnaud, M., Chastanet, A. & Debarbouille, M. New Vector for Efficient Allelic Replacement in Naturally Nontransformable, Low-GC-Content, Gram-Positive Bacteria. *Appl. Environ. Microbiol.* **70**, 6887–6891 (2004).
256. Frère, J., Benachour, A., Novel, M. & Novel, G. Identification of the theta-type minimal replicon of the *Lactococcus lactis* subsp. *lactis* CNRZ270 lactose protease plasmid pUCL22. *Curr. Microbiol.* **27**, 97–102 (1993).
257. Suárez, C. A., Blancato, V. S., Poncet, S., Deutscher, J. & Magni, C. CcpA represses the expression of the divergent *cit* operons of *Enterococcus faecalis* through multiple *cre* sites. *BMC Microbiol.* **11**, 1 (2011).
258. Platteeuw, C., Simons, G. & de Vos, W. M. Use of the *Escherichia coli* beta-glucuronidase (*gusA*) gene as a reporter gene for analyzing promoters in lactic acid bacteria. *Appl. Environ. Microbiol.* **60**, 587–593 (1994).
259. Brückner, R., Wagner, E. & Götz, F. Characterization of a sucrase gene from *Staphylococcus xylosus*. *J. Bacteriol.* **175**, 851–857 (1993).
260. Dower, W. J., Miller, J. F. & Ragsdale, C. W. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.* **16**, 6127–6145 (1988).
261. Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**,

557–580 (1983).

262. Aranda, P. S., LaJoie, D. M. & Jorcyk, C. L. Bleach gel: A simple agarose gel for analyzing RNA quality. *ELECTROPHORESIS* **33**, 366–369 (2012).
263. Wang, L., Wang, S. & Li, W. RSeQC: quality control of RNA-seq experiments. *Bioinforma. Oxf. Engl.* **28**, 2184–2185 (2012).
264. Tartof, K. D. & Hobbs, C. A. Improved media for growing plasmid and cosmid clones. *Bethesda Res Lab Focus* **9**, 12 (1987).
265. Vadeboncoeur, C., Brochu, D. & Reizer, J. Quantitative determination of the intracellular concentration of the various forms of HPr, a phosphocarrier protein of the phosphoenolpyruvate: Sugar phosphotransferase system in growing cells of oral streptococci. *Anal. Biochem.* **196**, 24–30 (1991).
266. Ferreira, M. J. & Sá-Nogueira, I. de. A multitask ATPase serving different ABC-type sugar importers in *Bacillus subtilis*. *J. Bacteriol.* **192**, 5312–5318 (2010).
267. Marciniak, B. C. *et al.* High- and low-affinity cre boxes for CcpA binding in *Bacillus subtilis* revealed by genome-wide analysis. *BMC Genomics* **13**, 401 (2012).
268. Passalacqua, K. D., Charbonneau, M.-E. & O’Riordan, M. X. D. Bacterial Metabolism Shapes the Host-Pathogen Interface. *Microbiol. Spectr.* **4**, (2016).
269. Richardson, A. R., Somerville, G. A. & Sonenshein, A. L. Regulating the Intersection of Metabolism and Pathogenesis in Gram-positive Bacteria. *Microbiol. Spectr.* **3**, (2015).
270. Oren, A. & Rupnik, M. *Clostridium difficile* and *Clostridioides difficile*: Two validly published and correct names. *Anaerobe* **52**, 125–126 (2018).
271. Collins, J. *et al.* Dietary trehalose enhances virulence of epidemic *Clostridium difficile*. *Nature* **553**, 291–294 (2018).
272. Perry, G. H. *et al.* Diet and the evolution of human amylase gene copy number variation. *Nat. Genet.* **39**, 1256–1260 (2007).
273. Kaoutari, A. E., Armougom, F., Gordon, J. I., Raoult, D. & Henrissat, B. The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat. Rev. Microbiol.* **11**, 497–504 (2013).
274. Koropatkin, N. M., Cameron, E. A. & Martens, E. C. How glycan metabolism shapes the human gut microbiota. *Nat. Rev. Microbiol.* **10**, 323–335 (2012).
275. Deibel, R. H., Lake, D. E. & Niven, C. F. PHYSIOLOGY OF THE ENTEROCOCCI AS RELATED TO THEIR TAXONOMY. *J. Bacteriol.* **86**, 1275–1282 (1963).
276. Joyet, P. *et al.* Enzymes Required for Maltodextrin Catabolism in *Enterococcus faecalis* Exhibit Novel Activities. *Appl. Environ. Microbiol.* **83**, (2017).
277. Sauvageot, N. *et al.* *Enterococcus faecalis* Uses a Phosphotransferase System Permease and

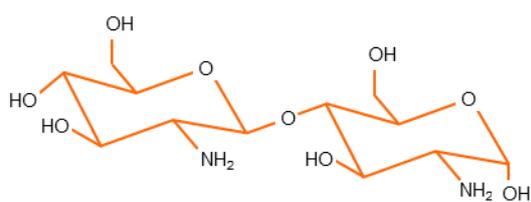
- a Host Colonization-Related ABC Transporter for Maltodextrin Uptake. *J. Bacteriol.* **199**, (2017).
278. Turrone, F. *et al.* Glycan Utilization and Cross-Feeding Activities by Bifidobacteria. *Trends Microbiol.* **0**, (2017).
279. Barbier, C. S., Short, S. A. & Senear, D. F. Allosteric Mechanism of Induction of CytR-regulated Gene Expression CytR REPRESSOR-CYTIDINE INTERACTION. *J. Biol. Chem.* **272**, 16962–16971 (1997).
280. Schumacher, M. A., Glasfeld, A., Zalkin, H. & Brennan, R. G. The X-ray Structure of the PurR-Guanine-purF Operator Complex Reveals the Contributions of Complementary Electrostatic Surfaces and a Water-mediated Hydrogen Bond to Corepressor Specificity and Binding Affinity. *J. Biol. Chem.* **272**, 22648–22653 (1997).
281. Geanacopoulos, M. & Adhya, S. Functional characterization of roles of GalR and GalS as regulators of the gal regulon. *J. Bacteriol.* **179**, 228–234 (1997).
282. Hars, U., Horlacher, R., Boos, W., Welte, W. & Diederichs, K. Crystal structure of the effector-binding domain of the trehalose-repressor of escherichia coli, a member of the LacI family, in its complexes with inducer trehalose-6-phosphate and noninducer trehalose. *Protein Sci.* **7**, 2511–2521 (1998).
283. Choi, P. J., Cai, L., Frieda, K. & Xie, X. S. A Stochastic Single-Molecule Event Triggers Phenotype Switching of a Bacterial Cell. *Science* **322**, 442–446 (2008).
284. Procházková, K. *et al.* Structure of the effector-binding domain of the arabinose repressor AraR from *Bacillus subtilis*. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 176–185 (2012).
285. Waugh, D. S. Making the most of affinity tags. *Trends Biotechnol.* **23**, 316–320 (2005).
286. Thielges, M. C., Chung, J. K., Axup, J. Y. & Fayer, M. D. Influence of Histidine Tag Attachment on Picosecond Protein Dynamics. *Biochemistry (Mosc.)* **50**, 5799–5805 (2011).
287. Buchholz, A., Hurlbauss, J., Wandrey, C. & Takors, R. Metabolomics: quantification of intracellular metabolite dynamics. *Biomol. Eng.* **19**, 5–15 (2002).
288. Badenhuis, N., Bose, R. J., Kirkwood, S., Lewis, B. A. & Smith, F. Isolation of Gentiobiose from Gentian Root. *J. Org. Chem.* **29**, 2079–2080 (1964).
289. Dumville, J. C. & Fry, S. C. Gentiobiose: a novel oligosaccharin in ripening tomato fruit. *Planta* **216**, 484–495 (2003).
290. Takahashi, H. *et al.* The Gentio-Oligosaccharide Gentiobiose Functions in the Modulation of Bud Dormancy in the Herbaceous Perennial *Gentiana*. *Plant Cell* **26**, 3949–3963 (2014).
291. Akšić, M. F. *et al.* Comparison of Sugar Profile between Leaves and Fruits of Blueberry and Strawberry Cultivars Grown in Organic and Integrated Production System. *Plants* **8**, 205 (2019).

292. Salyers, A. A. & Leedle, J. A. Z. CHAPTER 6 - Carbohydrate Metabolism in the Human Colon. in *Human Intestinal Microflora in Health and Disease* (ed. Hentges, D. J.) 129–146 (Academic Press, 1983). doi:10.1016/B978-0-12-341280-5.50012-4
293. Wolin, M. J. & Miller, T. L. CHAPTER 7 - Carbohydrate Fermentation. in *Human Intestinal Microflora in Health and Disease* (ed. Hentges, D. J.) 147–165 (Academic Press, 1983). doi:10.1016/B978-0-12-341280-5.50013-6
294. Crittenden, R. G. & Playne, M. J. Production, properties and applications of food-grade oligosaccharides. *Trends Food Sci. Technol.* **7**, 353–361 (1996).
295. Cressey, P. & Reeve, J. Metabolism of cyanogenic glycosides: A review. *Food Chem. Toxicol. Int. J. Publ. Br. Ind. Biol. Res. Assoc.* **125**, 225–232 (2019).
296. Bhatia, Y., Mishra, S. & Bisaria, V. S. Microbial  $\beta$ -Glucosidases: Cloning, Properties, and Applications. *Crit. Rev. Biotechnol.* **22**, 375–407 (2002).
297. Srivastava, N. *et al.* Microbial Beta Glucosidase Enzymes: Recent Advances in Biomass Conversation for Biofuels Application. *Biomolecules* **9**, 220 (2019).
298. Warren, C. E. Glycosylation. *Curr. Opin. Biotechnol.* **4**, 596–602 (1993).
299. Collin, M. & Fischetti, V. A. A novel secreted endoglycosidase from *Enterococcus faecalis* with activity on human immunoglobulin G and ribonuclease B. *J. Biol. Chem.* **279**, 22558–22570 (2004).
300. Koutsoulis, D., Landry, D. & Guthrie, E. P. Novel endo- $\alpha$ -N-acetylgalactosaminidases with broader substrate specificity. *Glycobiology* **18**, 799–805 (2008).
301. Bøhle, L. A., Mathiesen, G., Vaaje-Kolstad, G. & Eijsink, V. G. H. An endo- $\beta$ -N-acetylglucosaminidase from *Enterococcus faecalis* V583 responsible for the hydrolysis of high-mannose and hybrid-type N-linked glycans. *FEMS Microbiol. Lett.* **325**, 123–129 (2011).
302. Garbe, J. *et al.* EndoE from *Enterococcus faecalis* hydrolyzes the glycans of the biofilm inhibiting protein lactoferrin and mediates growth. *PLoS One* **9**, e91035 (2014).
303. Hossler, P., Khattak, S. F. & Li, Z. J. Optimal and consistent protein glycosylation in mammalian cell culture. *Glycobiology* **19**, 936–949 (2009).
304. Kiliç, A. O. *et al.* Involvement of *Streptococcus gordonii* Beta-Glucoside Metabolism Systems in Adhesion, Biofilm Formation, and In Vivo Gene Expression. *J. Bacteriol.* **186**, 4246–4253 (2004).
305. Terra, V. S., Zhi, X., Kahya, H. F., Andrew, P. W. & Yesilkaya, H. Pneumococcal 6-Phospho- $\beta$ -Glucosidase (BglA3) Is Involved in Virulence and Nutrient Metabolism. *Infect. Immun.* **84**, 286–292 (2015).
306. Jabbouri, S. & Sadovskaya, I. Characteristics of the biofilm matrix and its role as a possible target for the detection and eradication of *Staphylococcus epidermidis* associated with medical

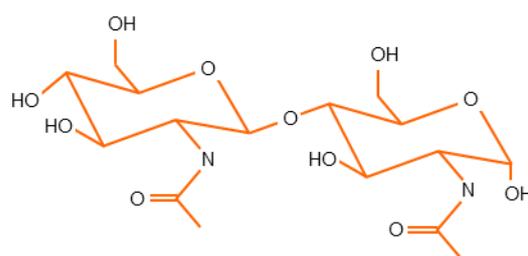
- implant infections. *FEMS Immunol. Med. Microbiol.* **59**, 280–291 (2010).
307. Arciola, C. R., Campoccia, D., Ravaoli, S. & Montanaro, L. Polysaccharide intercellular adhesin in biofilm: structural and regulatory aspects. *Front. Cell. Infect. Microbiol.* **5**, (2015).
308. Low, K. E. & Howell, P. L. Gram-negative synthase-dependent exopolysaccharide biosynthetic machines. *Curr. Opin. Struct. Biol.* **53**, 32–44 (2018).
309. Ramos, Y. *et al.* PolyGlcNAc-containing exopolymers enable surface penetration by non-motile *Enterococcus faecalis*. *PLOS Pathog.* **15**, e1007571 (2019).
310. Khan, F., Jeong, M.-C., Park, S., Kim, S.-K. & Kim, Y.-M. Contribution of chitooligosaccharides to biofilm formation, antibiotics resistance and disinfectants tolerance of *Listeria monocytogenes*. *Microb. Pathog.* 103673 (2019). doi:10.1016/j.micpath.2019.103673
311. Hava, D. L. & Camilli, A. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol. Microbiol.* **45**, 1389–1406 (2002).
312. Mohamed, J. A., Huang, W., Nallapareddy, S. R., Teng, F. & Murray, B. E. Influence of origin of isolates, especially endocarditis isolates, and various genes on biofilm formation by *Enterococcus faecalis*. *Infect. Immun.* **72**, 3658–3663 (2004).

# Annexes

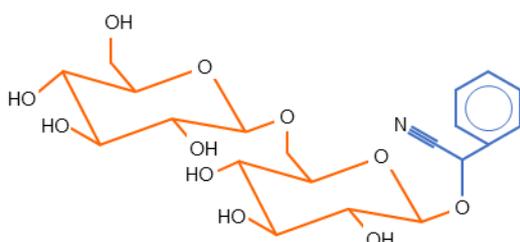
## I. Figures supplémentaires



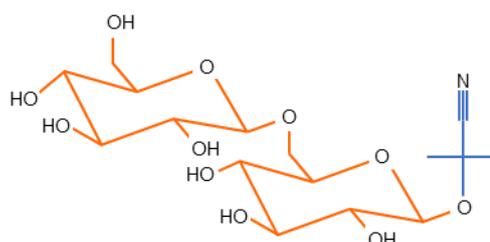
Chitosane (dimère)  
(Glucosamine- $\beta$ -1,4-Glucosamine)



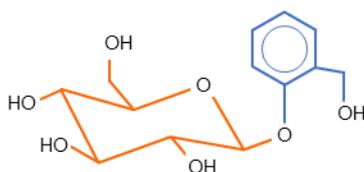
Chitobiose  
(N-acetylglucosamine- $\beta$ -1,4-N-acetylglucosamine)



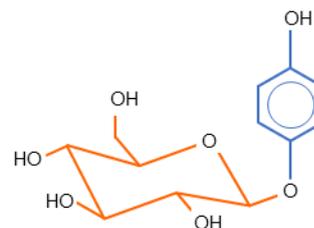
Amygdaline  
(Glucose- $\beta$ -1,6-Glucose- $\beta$ -Mandelonitrile)  
Gentiobiose



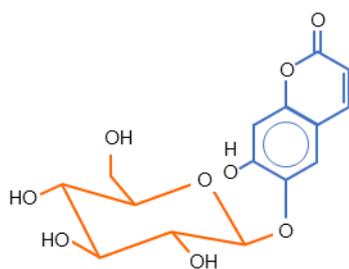
Linustatine  
(Glucose- $\beta$ -1,6-Glucose- $\beta$ -Methylpropanenitrile)  
Gentiobiose



Salicine  
(Glucose- $\beta$ -saligénine)



Arbutine  
(Glucose- $\beta$ -hydroquinone)



Esculine  
(Glucose- $\beta$ -hydroxycoumarine)

**Figure S.1** : Représentation de  $\beta$ -glycosides utilisés dans cette étude.

Les diholosides chitosane (dimère) et chitobiose ainsi que des hétérosides amygdaline, linustatine, salicine, arbutine et esculine. Les hexoses sont représentés en conformation chaise.

La partie glycosidique des hétérosides est représentée en orange et leur partie aglycone en bleu

## II. Compositions des milieux de cultures utilisés dans cette étude

**Tableau S.1** : Composition du milieu de culture LB

Constituant	Concentration
Tryptone USP	10 g/L
Extrait de levure	5 g/L
Chlorure de sodium	5 g/L
pH 7,1	

**Tableau S.2** : Composition du milieu de culture M17

Constituant	Concentration
Tryptone USP	5 g/L
Peptone de Soja	5 g/L
Extrait de viande	5 g/L
Extrait de levure	2,5 g/L
$\beta$ -glycérophosphate de sodium	19 g/L
Acide L-ascorbique	0,5 g/L
Sulfate de magnésium (heptahydrate)	0,25 g/L

**Tableau S.3** : Composition du milieu de culture ccM17

Constituant	Concentration
Tryptone USP	5 g/L
Peptone de Soja	5 g/L
Extrait de viande	5 g/L
Extrait de levure	2,5 g/L
MOPS	42 g/L
Acide L-ascorbique	0,5 g/L
Sulfate de magnésium (heptahydrate)	0,25 g/L

**Tableau S.4** : Composition du milieu de culture SOB

<b>Constituant</b>	<b>Concentration</b>
Tryptone USP	20 g/L
Extrait de levure	5 g/L
Chlorure de sodium	0,5 g/L
Chlorure de potassium	0,2 g/L
Chlorure de magnésium (anhydre)	1 g/L
Sulfate de magnésium (heptahydrate)	2,5 g/L

**Tableau S.5** : Composition du milieu de culture SOC

<b>Constituant</b>	<b>Concentration</b>
Tryptone USP	20 g/L
Extrait de levure	5 g/L
Chlorure de sodium	0,5 g/L
Chlorure de potassium	0,2 g/L
Chlorure de magnésium (anhydre)	1 g/L
Sulfate de magnésium (heptahydrate)	2,5 g/L
Glucose	3,6 g/L

**Tableau S.6** : Composition du milieu de culture Terrific Broth

<b>Constituant</b>	<b>Concentration</b>
Extrait de levure	24 g/L
Tryptone USP	12 g/L
Phosphate de potassium dibasique ( $K_2HPO_4$ )	9,4 g/L
Phosphate de potassium monobasique ( $KH_2PO_4$ )	2,2 g/L
pH 7,2	

### III. Amorces utilisées

**Tableau S.7 :** Amorces utilisées pour les purifications de protéines recombinantes

Nom de l'amorce	Séquence (5'→3')	Description
OG-MalR-NdeI OG-MalR-XhoI	GCAC <u>CATATG</u> GCAATTACAGTAAAAGATGTCGC CGCTCTCGAGTTAATGAATAGGAATAACTGTTTCACGC	Clonage du gène <i>malR</i> dans le vecteur d'expression pET28b
GenA_pQE30_for GenA_pQE30_rev	CTGAGGATCCATGTCAATTTTGAAAAATGATTTTC GGCCAAGCTTTTATAATTCTTCCCCATTCGTAGC	Clonage du gène <i>genA</i> dans le vecteur d'expression pQE30
GenR_pQE30_for GenR_pQE30_rev	GAGGGGATCCTTGTGTTGATTTAGAGAAGGTTTCAG TGAAAAGCTTCTAGAAAGGTGTTTCATCTAGTTC	Clonage du gène <i>genR</i> dans le vecteur d'expression pQE30

Les sites de restrictions introduits grâce à ces amorces sont soulignés

**Tableau S.8 :** Amorces utilisées pour les expériences de RT-qPCR

Nom de l'amorce	Séquence (5'→3')	Nom du gène étudié
ef_0006L ef_0006R	GATGGGGAAATCAGGGATTC TCTTTTCCATTTCGGCATTTC	<i>gyrA</i>
EF0291L EF0291R	AGAAGACCGCGAAGCAATTA TGGAAGTCGGGATTGATAGC	<i>genA</i>
EF0292L EF0292R	GCCATGGTGGTTTTTCATTTT GCCAGCTGTCCAATCTTTGT	<i>genB</i>
EF0293L EF0293R	TGAACCGTGGCATCATAAAA CAAAATGGGCTATGCTGGTT	<i>genR</i>
ef_0957L ef_0957R	CAGTAAAGCCTTTTGGTCCG AGATGATGAAGCCCAACAAG	<i>malP</i>
ef_0958L ef_0958R	TGATTGAACAAGCTGGCAAAA CGCAGTGGTGGTTCTTAGTT	<i>malT</i>
ef_1345L ef_1345R	ATGCCTTCCTTCTGGCCCT CCATTGCCAAAATCAATGGT	<i>mdxE</i>
ef_1347L ef_1347R	CAACCAGGTGTCCCTTGTCT TGCTGATTTTCCTCCCAAAC	<i>mmdH</i>

**Tableau S.9** : Amorces utilisées pour la création et la complémentation des mutants

Nom de l'amorce	Séquence (5'→3')	Description
mut_malR_for mut_malR_rev pMAD_malR_for pMAD_malR_rev verif_malR	GAATTTGGGTTAACATATGTGAAGTGCCATTG CAATGGCACTTCACATATGTAAACCCAAATTC ATCATATCCATGGTGGCTCAAAGG AAGAGCGTCGACATAGAAAGAAGCGTAAGC GGAACCAGCTCAATTTGATGC	Création du mutant <i>malR</i>
ccpArevSal ccpA_Sac_for ccpA_Sac_rev ccpA_JH_for_Nco	AAGGGGTCGACACGAGCTAAAGAAAGAGCTTG GCCATGAGCTCATTTCCTCTCTAGCAACATC ATATCGAGCTCGTGTCAATGCGCTTATTAAC GTAGGATCCATGGCATTAAAAGGTTTCGTC	Création du mutant $\Delta$ <i>ccpA</i>
pMAD_genA_for pMAD_genA_rev Mut_genA_for Mut_genA_rev verif_genA	TTTTGGATCCATTTTAAACGGTCAC GGCCGTCGACTTTAGACGAAGGCC TTGCTAGATCTTGATCACTCGTTACCAG ATCAAGATCTAGCAAATTAGACAAAAAATC TTGTCTAATTTGCTAGATCTTGA	Création du mutant $\Delta$ <i>genA</i>
pMAD_genB_for pMAD_genB_rev Mut_genB_for Mut_genB_rev verif_genB	CCAGGGATCCAAATGATTGTTCAAGC ATCGGTCGACCGGTTTCTTCTGCGTAG GATCTAGATCTTAAGATGCGGTTTCTGGC CTTAAGATCTAGATCACATAAACAAAATG TTTTGTTTATGTGATCTAGATCTT	Création du mutant $\Delta$ <i>genB</i>
pMAD_genR_for pMAD_genR_rev Mut_genR_for Mut_genR_rev verif_genR	GCTTGGATCCTCTCGCCCTATGAG TTGCGTCGACCCATTACAGAAATAGC TTGAAGATCTTAATATGCAGTGAAAGAAG ATTAAGATCTTCAAAAACCAGCATAGCC TATGCTGGTTTTGAAGATCTT	Création du mutant $\Delta$ <i>genR</i>
pMAD_1017for ef1019_rev mut_1018for mut_1018rev verif_1017 verif_1019rev	TTAATGAATTCTGTTATTTTTGATATC AAATGGATCCAAGAAAGGTTTGTTTAC GTATTGCTAGCGAGTAATCTCGTTGTC CCCATGCTAGCTTCTAAGTTCTGTTGTC AGCAGATGAATTTTACACGC TTGCCCTGTTCCACCTAAG	Création du mutant pour le gène <i>OGIRF_10751</i>
1F 2R	TCTAGCTAATGTTACGTTACAC TCATAATGGGAAGGCCATC	Vérification des clonages dans le plasmide pMAD
pRB473_for(bis) pRB473_rev(bis)	TTAAGGGCACCAATAACTGCC TGTTCAGAACGCTCGGTTG	Vérification des clonages dans le plasmide pRB473
pRB473_for pRB473_rev	GTCGACCTGCAGGCATGCAAGCTTGG GATGAGTGGCAGGGCGGGGCG	Amplification du plasmide pRB473
malR_(comp_pRB473)_for malR_(comp_pRB473)_rev	AAAAAATTACGCCCCGCCCTGCCACTCATCGAA CCAGCTCAATTTGATGC AGAATCCAAGCTTGCATGCCTGCAGGTCGACTC CAACGCTATCCAACCGG	Complémentation du mutant <i>malR</i>
ccpA_(comp_pRB473)_for ccpA_(comp_pRB473)_rev	AAAAAATTACGCCCCGCCCTGCCACTCATCTTCT TTCTTTTATTCTGAAG GAATCCAAGCTTGCATGCCTGCAGGTCGACATTT TAATTTTCAATCTTAC	Complémentation du mutant $\Delta$ <i>ccpA</i>

Les sites de restrictions introduits grâce à ces amorces sont soulignés

**Tableau S.10** : Amorces utilisées pour les expériences de 5'RACE-PCR

Nom de l'amorce	Séquence (5'→3')	Nom du gène étudié
Race_0954-1	TCCTTACTATCTGAATACAC	<i>malR</i>
Race_0954-2	TCAAAACTTTTGGCAGTTGC	
Race_0954-3	CGTTCTTTTCGAATCTAAAGG	
genB_sp1	GCCAGCTGTCCAATCTTTGT	<i>genB</i>
genB_sp2	TAGGCAATACCGATGACCGC	
genB_sp3	AAACACGGAACCAACAATGG	
genR_sp1	TGAACCGTGGCATCATAAAA	<i>genR</i>
genR_sp2	TCGGAAAACCAGCATAGCC	
genR_sp3	GGTGGAGACATGACAGGAAGTAG	

**Tableau S.11** : Autres amorces utilisées dans cette étude

Nom de l'amorce	Séquence (5'→3')	Description
PmalR_BamHI_for	AAAGGATCCTTTAATCCTACGAACCACGC	Clonage du promoteur <i>PmalR</i> dans le plasmide pNZ273
PmalR_PstI_rev	AAACTGCAGAAGGTTAAATTGATAAACCG	
PmalR_BamHI_rev (ctrl-)	AAAGGATCCAAGGTTAAATTGATAAACCG	Clonage de la séquence inversé du promoteur <i>PmalR</i> dans le plasmide pNZ273
PmalR_PstI_for (ctrl-)	AAACTGCAGTTTAATCCTACGAACCACGC	
pNZ273_MCS_for	ATGAGATAATGCCGACTGTAC	Vérification des clonages dans le plasmide pNZ273
pNZ273_MCS_rev	CGATCCAGACTGAATGCC	
PmalP-T_for	ACGTTTGATTTGTTTCATGAGAATC	Amplification de la région intergénique entre les gènes <i>malP</i> et <i>malt</i>
PmalP-T_rev	CTCAAACTAAACATTTTTTTCATTC	
PmalT-I_for (bis)	AGGATTAAGTAGTAGGGAGTTG	
PmalP-I_rev	TCAAATAAAGCTCCTTCTAACCTAC	
GusA-D	GAAATCATGGAAGTAAGACTGC	Amplification d'un fragment interne du gène <i>gusA</i>
GusA-G	GTGTGGGTCAATAATCAGG	
PmdxE-mH_for	AATCTTTTTCCAATTCGCTTTC	Amplification de la région intergénique entre les gènes <i>mdxE</i> et <i>mmdH</i>
PmdxE-mH_rev	GCAGTGTTCAATTAATTCTCTCC	
PmdxE-I_rev	ACTTTAAACAAACAGTATGTTATTTATAAC	
PmdxE-I_for	AAGTGGTGTATACCACTAATTAAGC	
pU	TGTAAAACGACGGCCAGT	Vérification des clonages dans les plasmides pUCB300 et pGEM-T
pR (M13-rev)	CAGGAAACAGCTATGACC	
Promo_genB_for	AGGAACTAGATGAAACACCTTCC	Amplification du promoteur du gène <i>genB</i>
Promo_genB_rev	GGAGGTAATAATTTCTCGTTAATCC	

Les sites de restrictions introduits grâce à ces amorces sont soulignés

## Résumé

### Régulation des opérons maltose/maltodextrines et gentiobiose induits en contexte d'infection chez *Enterococcus faecalis*

Les entérocoques sont des bactéries commensales de l'Homme majoritairement rencontrées dans le tractus digestif. En dépit du caractère bénéfique pour leur hôte, ces microorganismes sont retrouvés au deuxième rang des bactéries responsables d'infection nosocomiales en France ces dernières décennies. Diverses études tendent à montrer que le métabolisme énergétique constitue un facteur crucial pour le processus infectieux des microorganismes. Lors de ce travail, nous nous sommes intéressés à l'étude des métabolismes de différents polymères de glucoses chez *Enterococcus faecalis* : les maltodextrines et le gentiobiose. L'utilisation du maltose et des maltodextrines est, chez cette bactérie, directement coordonnée au niveau transcriptionnel par le répresseur MalR. L'activité de ce régulateur est rapidement modulée par le maltose qui représente l'inducteur du système et par un corépresseur protéique : la protéine P-Ser-HPr qui, à l'inverse, favorise la répression exercée par MalR. Le métabolisme des maltodextrines complexes, mais pas le métabolisme du maltose, est également réprimé par le régulateur pléiotrope CcpA en coordination avec son cofacteur P-Ser-HPr en présence de glucose. La répression catabolique de l'opéron *genBA*, impliqué dans le métabolisme du  $\beta$ -glycoside gentiobiose, est aussi assurée par ce régulateur CcpA en présence de glucose. Cet opéron *genBA* est responsable de l'import du gentiobiose par un PTS ainsi que de son catabolisme grâce à une hydrolase. L'expression de cette structure opéronique nécessite la présence de l'activateur transcriptionnel GenR actif en présence de l'inducteur gentiobiose-6'-P.

*Mots clés : Enterococcus, métabolisme, régulation, maltose, maltodextrine,  $\beta$ -glycoside, polysaccharide*

---

## Abstract

### Regulation of maltose/maltodextrin and gentiobiose operons induced during infection in *Enterococcus faecalis*

Enterococci are commensal bacteria of Humans predominantly encountered in the digestive tract. Despite their beneficial activity for their host, these microorganisms became the second leading bacterial cause of hospital acquired infections in France for last decades. Some studies showed that the central metabolism is a critical factor for microorganisms infection process. In this study, we worked on the characterisation of metabolisms of the different glucose polymers maltodextrins and gentiobiose in *Enterococcus faecalis*. The maltose and maltodextrins utilization is coordinated in this bacterium transcriptionally by the MalR repressor. The MalR activity is rapidly modulated by the inducer maltose and by the co-repressor P-Ser-HPr which strengthens the MalR DNA binding. The metabolism of long maltodextrins is also repressed by the pleiotropic regulator CcpA in complex with its essential cofactor P-Ser-HPr in presence of glucose. The Catabolite repression of the operon *genBA*, involved in metabolism of the  $\beta$ -glycoside gentiobiose, is assumed by CcpA in presence of glucose. This operon *genBA* allows the gentiobiose uptake with a PTS and its catabolism by a hydrolase. The expression of this latter operon requires both the GenR transcriptional activator and the inducer gentiobiose-6'-P.

*Key words: Enterococcus, metabolism, regulation, maltose, maltodextrin,  $\beta$ -glycoside, polysaccharide*