UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS BIOLÓGICAS
Departamento de Bioquímica y Biología Molecular

TESIS DOCTORAL

Caracterización molecular de un regulador transcripcional del tipo Mga-AtxA en *Enterococcus faecalis*

Molecular characterization of an *Enterococcus faecalis* transcriptional regulator of the Mga-AtxA family

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Sofía Isabel Ruiz Cruz

Directora

Alicia Bravo García

Madrid, 2016

UNIVERSIDAD COMPLUTENSE MADRID

FACULTAD DE CIENCIAS BIOLÓGICAS

Departamento de Bioquímica y Biología Molecular I

Tesis Doctoral

CARACTERIZACIÓN MOLECULAR DE UN REGULADOR TRANSCRIPCIONAL DEL TIPO Mga/AtxA EN *Enterococcus faecalis*

MOLECULAR CHARACTERIZATION OF AN *Enterococcus faecalis* TRANSCRIPTIONAL REGULATOR OF THE Mga/AtxA FAMILY

Memoria presentada para optar al grado de

Doctora con Mención Europea por

Sofía Isabel Ruiz Cruz

Bajo la dirección de la Doctora

Alicia Bravo García

Centro de Investigaciones Biológicas

CSIC

Madrid, 2015

Este trabajo ha sido realizado por Sofía Isabel Ruiz Cruz bajo la dirección de la Dra. Alicia Bravo García, en el Departamento de Microbiología Molecular y Biología de las Infecciones del Centro de Investigaciones Biológicas (CIB) del Consejo Superior de Investigaciones Científicas (CSIC).
Financiación: Ministerio de Educación (Beca FPU AP2008-00105); Comunidad Autónoma de Madrid/CSIC (Proyecto CCG08-CSIC/SAL-3694); Ministerio de Ciencia e Innovación, Ministerio de Economía y Competitividad (Proyectos CSD2008-00013-INTERMODS, BFU2009-11868, BIO2013-49148-C2-2-R); CSIC (Proyecto PIE-201320E028).
La estancia de S. I. Ruiz Cruz en el Helmholtz Centre for Infection Research (Braunschweig, Alemania) se enmarcó dentro de una colaboración con el Dr. Oliver Goldmann y fue financiada por el CSIC (Proyecto PIE-201320E028).

Sofía Isabel Ruiz Cruz performed this work under the supervision of Dr. Alicia Bravo García at the Department of Molecular Microbiology and Infection Biology, Centro de Investigaciones Biológicas (CIB), Consejo Superior de Investigaciones Científicas (CSIC).
Funding: Ministry of Education (Fellowship FPU AP2008-00105); Community of Madrid/CSIC (Grant CCG08-CSIC/SAL-3694); Ministry of Science and Innovation, Ministry of Economy and Competitiveness (Grants CSD2008-00013-INTERMODS, BFU2009-11868, BIO2013-49148-C2-2-R); CSIC (Grant PIE-201320E028).
The stint of S. I. Ruiz Cruz at Helmholtz Centre for Infection Research (Braunschweig, Germany) was framed within cooperation with Dr. Oliver Goldmann and was funded by CSIC (Grant PIE-201320E028).

Opta al grado de Doctor **VºBº Directora de Tesis**

Sofía Isabel Ruiz Cruz **Alicia Bravo García**

TABLE OF CONTENTS

List of Tables

List of Figures

Abbreviations

aa	Amino acid (s)
APS	Ammonium persulfate
bp (s)	Base pair
BEA	Bile esculin azide
BSA	Bovine serum albumin
cDNA	Complementary DNA
Ci	Curies
cpm	Counts per minute
C_T	Threshold cycles
Da	Dalton
DAPI	4′,6′-diamidino-2-phenylindole
DNase I	Desoxirribonuclease I
dNTP	Deoxynucleotide triphosphate
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
EDTA	Ethylenediamine-tetra-acetic acid
EMSA	Electrophoretic mobility shift assay (s)
Em	Erythromycin
EtBr	Ethidium bromide
FC	Fold change
FPLC	Fast protein liquid chromatography
Fus	Fusidic acid
GAS	Group A *Streptococcus*
GFP	Green fluorescent protein
His	Histidine
HTH	Helix-turn-helix
IL	Interleukin
IPTG	Isopropyl-β-D-thiogalactosidase
IR	Inverted repeat
Kav	Partition coefficient
kb	Kilobase (s)
kDa	Kilo Dalton (s)
Km	Kanamycin
MCS	Multi-cloning site
mRNA	Messenger RNA

Mw,a	Average molecular mass
nt	Nucleotide (s)
OD	Optical density
PAA	Polyacrylamide
PAGE	Polyacrylamide gel electrophoresis
PEI	Polyethylenimine
PCR	Polymerase chain reaction
PRD	Phosphoenolpyruvate:carbohydrate phosphotransferase system regulation domain
PTS	Phosphoenolpyruvate:carbohydrate phosphotransferase system
Rif	Rifampicin
RT	Room temperature
RNAP	RNA polymerase
rpm	Revolutions per minute
RT-PCR	Reverse transcription-polymerase chain reaction
$S_{20,w}$	Standardized sedimentation coefficient
SD	Shine-Dalgarno
SDS	Sodium dodecyl sulfate
T4-PNK	Polynucleotide kinase of T4-bacteriophage
Tc	Tetracycline
TCSs	Two-component signal transduction systems
TEMED	N,N,N',N'-tetra-methylethylenediamine
Tm	Melting temperature
Tris	Tris-hydroxymethyl-aminomethane
UV	Ultraviolet
V	Volts
Van	Vancomycin
W	Watts
wHTH	Winged helix-turn-helix
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside

RESUMEN

INTRODUCCIÓN

Enterococcus faecalis es una bacteria Gram-positiva que se encuentra de forma comensal en el tracto gastrointestinal humano. Sin embargo, como patógeno oportunista puede causar graves infecciones como endocarditis y bacteriemia. En esta bacteria, la identificación de factores de virulencia aumentó considerablemente tras la publicación de la secuencia del genoma de la estirpe V583 (Paulsen *et al*., 2003). Sin embargo, el conocimiento de los mecanismos moleculares que controlan su virulencia sigue siendo escaso. En general, los reguladores globales que modulan la transcripción de múltiples genes en respuesta a señales ambientales específicas son esenciales en la adaptación de las bacterias patógenas a nichos particulares del hospedador. Mediante búsquedas en las bases de datos, encontramos que el gen EF3013 (llamado *maEfa* en este trabajo) de *E. faecalis* V583 codifica una proteína (MA*Efa*) que presenta similitud de secuencia con los reguladores globales de la familia Mga/AtxA. La estirpe OG1RF también codifica la proteína MA*Efa* (OG1RF_12293). La familia Mga/AtxA incluye a los reguladores de virulencia AtxA, Mga y Mga*Spn* de las bacterias Gram-positivas patógenas *Bacillus anthracis*, *Streptococcus pyogenes* y *Streptococcus pneumoniae*, respectivamente. Dichos reguladores controlan la expresión de numerosos genes y están implicados en la patogénesis de estas bacterias (Fouet, 2010; Hemsley *et al*., 2003; McIver, 2009; Solano-Collado *et al*., 2012). El trabajo de investigación presentado en esta Tesis ha estado centrado en la caracterización funcional de la proteína MA*Efa* de *E. faecalis*.

OBJETIVOS

Para determinar si MA*Efa* es un nuevo miembro de la familia Mga/AtxA de reguladores globales, hemos desarrollado los siguientes objetivos específicos:

1. Construcción de vectores plasmídicos para ensayar promotores y terminadores transcripcionales en *E. faecalis*

2. Construcción de vectores de expresión para *E. faecalis* basados en plásmidos

3. Análisis de la expresión del gen *maEfa* en células enterocócicas

4. Purificación de la proteína MA*Efa* y análisis de su estado de oligomerización

5. Estudio de la interacción de MA*Efa* con DNA lineal de cadena doble

6. Análisis del efecto de MA*Efa* en expresión génica global y virulencia

APORTACIONES FUNDAMENTALES Y CONCLUSIONES

1. Herramientas genéticas para *E. faecalis* basadas en plásmidos

A pesar del creciente interés clínico en *E. faecalis*, las herramientas disponibles para su manipulación genética son limitadas. Por ello, hemos construido un vector que permite ensayar promotores (plásmido pAST) y un vector que permite ensayar terminadores transcripcionales (plásmido pAS), ambos basados en el plásmido promiscuo pMV158 (del Solar *et al*., 1993). El gen reportero utilizado es una variante del gen *gfp* (proteína fluorescente verde). Para analizar la funcionalidad de estos vectores, insertamos diferentes señales promotoras y terminadoras en pAST y pAS, respectivamente. Mediante ensayos de fluorescencia demostramos que ambos vectores son de gran utilidad para evaluar la actividad de promotores y terminadores (ya sean homólogos o heterólogos) tanto en *E. faecalis* como en *S. pneumoniae*. Además, hemos desarrollado dos vectores de expresión (pDLF y pDLS) para *E. faecalis* basados en pDL287, un plásmido de amplio espectro de huésped (LeBlanc *et al*., 1993). El vector pDLF contiene un promotor de *E. faecalis* (*P2493*) y el vector pDLS contiene un promotor de *S. pneumoniae* (*PsulA*). La funcionalidad de estos vectores se confirmó mediante el clonaje del gen *maEfa.*

2. Transcripción *in vivo* del gen *maEfa*

Mediante experimentos de RT-PCR, hemos demostrado que hay transcripción del gen *maEfa* en diferentes estirpes de *E. faecalis* (V583, OG1RF y JH2-2) cuando las bacterias crecen en condiciones de laboratorio estándar. A partir de aproximaciones *in vivo* tales como el uso de fusiones transcripcionales (plásmido pAST-*Pma*) y experimentos de *primer extension,* hemos identificado: (i) el promotor del gen *maEfa* (*Pma*) y (ii) el inicio de transcripción de *maEfa*, que está localizado 15 nucleótidos *upstream* del codón de inicio predicho. Además, utilizando el vector pAS confirmamos que el promotor *Pma* está precedido por un terminador transcripcional intrínseco.

3. Características de la proteína MA*Efa*

Hemos establecido un protocolo para sobreproducir y purificar la proteína MA*Efa*. El protocolo de purificación consiste esencialmente en cuatro etapas: (i) precipitación de los ácidos nucleicos y MA*Efa* con polietilenimina (PEI) a baja fuerza iónica; (ii) elución de MA*Efa* del pellet de PEI empleando un tampón de mayor fuerza iónica; (iii) cromatografía de afinidad utilizando columnas de heparina y (iv)

cromatografía de filtración en gel. Mediante ensayos de filtración en gel y experimentos de ultracentrifugación analítica, demostramos que MA*Efa* forma dímeros en solución.

4. Interacción de la proteína MA*Efa* con DNA

Hemos analizado la interacción de MA*Efa* con DNA lineal de cadena doble mediante ensayos de retraso en gel (EMSA) y experimentos de protección frente a la digestión con DNasa I. Mediante EMSA mostramos que: (i) MA*Efa* genera complejos multiméricos sobre DNA lineal de cadena doble; (ii) MA*Efa* no parece reconocer una secuencia nucleotídica específica cuando interacciona con DNA lineal de cadena doble y (iii) MA*Efa*-His, una versión de MA*Efa* con cola de histidinas, también forma múltiples complejos proteína-DNA. En los ensayos de digestión con DNasa I, utilizamos fragmentos que contenían el promotor *Pma*. En dichos fragmentos, MA*Efa* reconoce preferentemente un sitio localizado *upstream* del promotor. Este sitio contiene una curvatura intrínseca potencial.

5. Efecto de MA*Efa* en expresión génica global y virulencia

Para investigar si MA*Efa* funciona como un regulador global, construimos la estirpe OG1RFΔ*maEfa*, un derivado de OG1RF que carece del gen *maEfa*. Mediante análisis de microarrays, RT-PCR cuantitativa y estudios de complementación, mostramos que MA*Efa* influye positivamente en la transcripción de numerosos genes. Muchos de ellos codifican componentes de transportadores PTS, componentes de transportadores ABC y proteínas implicadas en el metabolismo de fuentes de carbono. Asociado a este hecho, el crecimiento de la estirpe OG1RFΔ*maEfa* está afectado en medios de cultivo que contienen glicerol, maltosa o manitol. Además, estudiamos el papel de MA*Efa* en virulencia. Este estudio fue realizado en el Helmholtz Centre for Infection Research bajo la supervisión del Dr. Oliver Goldmann. Empleando un modelo de peritonitis murina descubrimos que la estirpe OG1RFΔ*maEfa* induce menor inflamación en la cavidad peritoneal comparada con la estirpe OG1RF. Nuestros resultados sugieren que MA*Efa* facilita la adaptación de *E. faecalis* controlando la transcripción de numerosos genes y, en consecuencia, contribuye a su posible virulencia.

SUMMARY

INTRODUCTION AND OBJECTIVES

The Gram-positive bacterium *Enterococcus faecalis* is a natural inhabitant of the human gastrointestinal tract. However, as an opportunistic pathogen, it is able to cause infections such as urinary tract infections, endocarditis, and bacteremia. Although the identification of virulence determinants increased considerably after the publication of the genome sequence of *E. faecalis* strain V583, a vancomycin-resistant clinical isolate (Paulsen et al., 2003), our understanding of the molecular mechanisms that control the virulence of this bacterium is still very limited. In general, global transcriptional regulators that activate and/or repress the transcription of multiple genes in response to specific environmental signals are essential in the adaptation of pathogenic bacteria to new host niches. Searching for homologies, we found that the EF3013 gene (named *maEfa* herein) of *E. faecalis* strain V583 encodes a protein (MA*Efa*; 482 residues) that has sequence similarity to global transcriptional regulators of the Mga/AtxA family. The enterococcal OG1RF genome also encodes the MA*Efa* protein (OG1RF_12293). The Mga/AtxA family includes the virulence regulators AtxA, Mga and Mga*Spn* from the Gram-positive pathogens *Bacillus anthracis, Streptococcus pyogenes* and *Streptococcus pneumoniae,* respectively. Such regulators control the expression of numerous genes and play an important role in the pathogenicity of these bacteria (Fouet, 2010; Hemsley *et al*., 2003; McIver, 2009; Solano-Collado *et al*., 2012). This Thesis has been focused on the functional characterization of MA*Efa.* To investigate whether MA*Efa* is a new member of the Mga/AtxA family of global regulators, we have worked on the following specific objectives:

1. Construction of promoter-probe and terminator-probe plasmid vectors in *E. faecalis*
2. Construction of expression plasmid vectors in *E. faecalis*
3. Analysis of *maEfa* gene expression in enterococcal cells
4. Purification of the MA*Efa* protein and analysis of its oligomerization state
5. Study of the interaction of MA*Efa* with linear double-stranded DNA
6. Analysis of the effect of MA*Efa* on global gene expression and virulence

RESULTS AND CONCLUSIONS

1. Plasmid-based genetic tools for *E. faecalis*

Despite the increasing clinical significance of *E. faecalis*, the tools available for its genetic manipulation are still very scarce. We have constructed a promoter-probe vector (pAST) and a terminator-probe vector (pAS) based on the broad-host range plasmid pMV158. As reporter gene, a *gfp* allele (green fluorescent protein) was used. To analyse the functionality of these vectors, we inserted promoter and terminator signals into pAST and pAS, respectively. We demonstrated by fluorescence assays that both vectors are suitable to assess the activity of promoters and terminators (both homologous and heterologous) not only in *E. faecalis* but also in *S. pneumoniae*. In addition we have developed two expression vectors based on the broad-host range plasmid pDL287: vector pDLF carries an enterococcal promoter (*P2493*) and vector pDLS carries a pneumococcal promoter (*PsulA*). The usefulness of these vectors was confirmed by cloning the *maEfa* gene.

2. *In vivo* transcription of the enterococcal *maEfa* gene

By RT-PCR experiments, we have demonstrated that the *maEfa* gene is transcribed in different *E. faecalis* strains (V583, OG1RF and JH2-2) under standard bacterial growth conditions Using *in vivo* approaches such as promoter-reporter fusions (plasmid pAST-*Pma*) and primer extension experiments we identified the promoter (*Pma*) of the *maEfa* gene. Its transcription initiation site is located 15 nucleotides upstream of the predicted translation codon. Moreover, using the terminator-probe vector pAS, we found that there is a functional transcriptional terminator upstream of the *Pma* promoter.

3. Features of the MA*Efa* protein

We have set up a protocol to overproduce and purify an untagged form of the MA*Efa* protein. It involved essentially four steps: (i) nucleic acids and MA*Efa* were precipitated with polyethylenimine (PEI) at a low ionic strength; (ii) MA*Efa* was eluted from the PEI pellet using a higher ionic strength buffer; (iii) the sample was subjected to heparin affinity chromatography, and (iv) the protein preparation was loaded onto a gel-filtration column. By gel filtration chromatography and analytical ultracentrifugation experiments we demonstrated that MA*Efa* form dimers in solution.

4. Interaction of the MA*Efa* protein with DNA

The interaction of MA*Efa* with linear double-stranded DNA was analysed by electrophoretic mobility shift assays (EMSA) and DNase I footprinting. By EMSA, we have shown that: (i) MA*Efa* is able to generate multimeric complexes on linear double-stranded DNA, (ii) MA*Efa* binds to linear double-stranded DNA with little or no sequence specificity, (iii) a His-tagged MA*Efa* protein (MA*Efa*-His) also forms multiple DNA-protein complexes, and (iv) the minimum DNA size required for MA*Efa* binding is between 26 bp and 32 bp. Moreover, by DNase I footprinting assays using DNA fragments that contain the *Pma* promoter we have shown that MA*Efa* recognizes preferentially a site located upstream of *Pma*. This site has a potential intrinsic curvature. The presence of such a binding site on a *Pma-gfp* transcriptional fusion does not affect the activity of *Pma* in cells overproducing MA*Efa* (cells carrying pAST-*Pma* and pDLF*maEfa*; fluorescence assays).

5. Effect of MA*Efa* on global gene expression and virulence

To investigate whether MA*Efa* functions as a global transcriptional regulator, we constructed an OG1RF derivative that is not able to synthesize MA*Efa* (OG1RFΔ*maEfa*). By genome-wide microarrays, quantitative RT-PCR assays, and complementation studies we have shown that MA*Efa* influences positively the transcription of numerous genes. Many of them encode components of PTS-type transporters, components of ABC-type transporters, and proteins involved in the metabolism of carbon sources. Associated to this fact, the growth of the *maEfa* deletion mutant strain is impaired in media containing glycerol, maltose or mannitol. Furthermore, the potential role of MA*Efa* in virulence was studied at the Helmholtz Centre for Infection Research under the supervision of Dr. Oliver Goldmann. Using a mouse peritonitis infection model, we found that the OG1RFΔ*maEfa* strain induces a significant lower degree of inflammation in the peritoneal cavity of mice compared to the wild-type strain. Thus, *E. faecalis* cells deficient in MA*Efa* are less virulent. Our results suggest that MA*Efa* facilitates the adaptation of *E. faecalis* to particular host niches through the transcriptional regulation of numerous genes and consequently, it contributes to its potential virulence.

INTRODUCTION

1.The opportunistic pathogen *Enterococcus faecalis*

The Gram-positive bacterium *Enterococcus faecalis* is a facultatively anaerobic organism, ovoid in shape, which appears in pairs or chains (Lebreton *et al*., 2014; Murray, 1990). It occupies a variety of habitats ranging from different environmental sources to the commensal presence in numerous animals and humans (Fisher and Phillips, 2009; Lebreton *et al*., 2014). However, as an opportunistic pathogen it is able to cause infections such as urinary tract infections, endocarditis and bacteremia, among other less frequent infections. Moreover, in the last decades, *E. faecalis* has emerged as a leading cause of nosocomial infections (Agudelo Higuita and Huycke, 2014; Fisher and Phillips, 2009). The ubiquitous nature of this bacterium is related with its metabolic versatility and capacity to withstand harsh conditions, along with its increasing antimicrobial resistance and its malleable genome (Arias and Murray, 2012; Ramsey *et al*., 2014).

E. faecalis is usually simply described as a lactic-acid-producing bacterium, but its metabolic potential is enormous. It can use numerous energy sources, including, among others, more than ten carbohydrates, glycerol, citrate and lactate (Ramsey *et al*., 2014). In addition, as other enterococci, it is able to grow in a wide range of temperature (from 10 to 45ºC) and pH (from 4.5 to 10). It tolerates high concentrations of NaCl (6.5%), bile salts (40%), metals and oxidants. Despite *E. faecalis* cannot form spores, it is highly resistant to desiccation (reviewed in (Lebreton *et al*., 2014; Ramsey *et al*., 2014, Fisher and Phillips, 2009). These characteristics as well as its increasing antibiotic resistance allow its survival in healthcare environments such as hospitals, thus contributing to the rise of enterococci as nosocomial pathogens (Arias and Murray, 2012). *E. faecalis* is inherently resistant to several antibiotics and can acquire new antibiotic resistances through sporadic mutations in intrinsic genes or by horizontal gene transfer. Additionally, they can disseminate the resistance genes to other pathogens (Arias and Murray, 2012; Hollenbeck and Rice, 2012). The ability of *E. faecalis* to exchange genetic material also results in the acquisition of new adaptive traits such as bacteriocins, metabolic genes and virulence determinants (Clewell *et al*., 2014).

E. faecalis is an organism which usually displays low levels of virulence (Arias and Murray, 2012). In fact, this bacterium is a natural inhabitant of the gastrointestinal tract in most humans and it has been used for decades as probiotic in humans as well as in food production (Agudelo Higuita and Huycke, 2014). However, it may possess

potential virulence determinants that contribute to cause diseases (reviewed in (Agudelo Higuita and Huycke, 2014; Arias and Murray, 2012; Garsin *et al*., 2014; Sava *et al*., 2010).

Several virulence determinants of *E. faecalis* are localized in the cell surface (Fig. 1). The enterococcal cell wall is a complex structure formed by peptidoglycan, proteins, polysaccharides, lipids, lipoproteins and several glycoconjugates (Sava *et al*., 2010). In Gram-positive bacteria, the polysaccharydes are usually involved in the evasion of the immune system. The enterococcal polysaccharide antigen (Epa) is widespread among *E. faecalis* strains. This carbohydrate is antigenic (Xu *et al*., 1997). The *epa* locus contains 16 genes involved in the biosynthesis of this polysaccharide (Xu *et al*., 2000). The disruption of this locus decreased biofilm formation and enterocyte translocation (Mohamed *et al*., 2005; Zeng *et al*., 2004); increased susceptibility to killing by polymorphonuclear neutrophils (Teng *et al*., 2002); and reduced virulence in mouse peritonitis and urinary tract infection models (Mohamed *et al*., 2005; Singh *et al*., 2009a ; Teng *et al*., 2002; Xu *et al*., 2000; Zeng *et al*., 2004). In contrast, the *cps* locus, which consists of 8 or 9 genes, has been found only in some *E. faecalis* strains (Hancock and Gilmore, 2002). This locus encodes a capsular polysaccharide and is present in serotypes C and D strains out of the four serogroups described by Hufnagel *et a*l. (2004a). The capsular polysaccharide of these serotypes masks lipoteichoic acid, and confers resistance to opsonophagocyitosis mediated by the complement (Hancock and Gilmore, 2002; Thurlow *et al*., 2009a). The lipoteichoic acid and the cell membrane glycolipids also seem to play a role in the pathogenesis of *E. faecalis* (Fabretti *et al*., 2006; Theilacker *et al*., 2009).

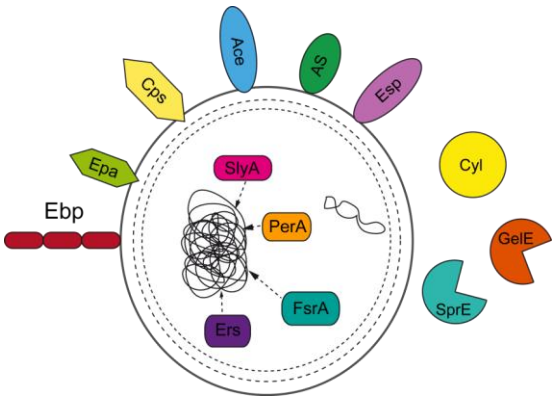


Figure 1. **Main virulence determinants of *E. faecalis.*** Important virulence determinants of *E. faecalis* include cell surface proteins (Ace, Esp and AS proteins); cell surface polysaccharides (Epa and Cps) and secreted factors (Cyl, GelE and SprE). The global regulators FsrA, PerA, Ers and SlyA are also involved in the virulence of *E. faecalis*.

In the early stages of *E. faecalis* infection, an important step is the attachment to proteins of the host extracellular matrix (Arias and Murray, 2012). The Ace protein, which belongs to the MSCRAMMs family (microbial surface components recognizing adhesive matrix molecules), binds to collagen and laminin (Nallapareddy *et al*., 2000a; Rich *et al*., 1999). This protein is produced by most of *E. faecalis* strains during human endocarditis and is antigenic (Nallapareddy *et al*., 2000b). In animal models, it has been shown that an *ace* deletion leads to attenuated endocarditis and urinary tract infections (Lebreton *et al*., 2009; Singh *et al*., 2009b). Another protein which has been involved in experimental urinary tract infections is the enterococcal surface protein Esp (Shankar *et al*., 2001), which is encoded on a pathogenicity island (Shankar *et al*., 2002). Ace and Esp contain LPXTG-like motifs, which are specific amino acid sequences necessary for the specific attachment of the protein to the cell wall peptidoglycan (Hendrickx *et al*., 2009). Such a motif is also present in the family of aggregation substance proteins (AS proteins). They are a group of surface-localized proteins encoded on pheromone-inducible conjugative plasmids that mediate binding of donor bacterial cells to recipient cells and promote conjugation (Sava *et al*., 2010, Arias and Murray, 2012, Olmsted, 1991). AS proteins are associated with an increase of *E. faecalis* binding to renal tubular cells (Kreft *et al*., 1992); internalization by intestinal cells (Sartingen *et al*., 2000); and survival within macrophages and neutrofils (Rakita *et al*., 1999; Süssmuth *et al*., 2000). In addition, these proteins also increased the virulence of *E. faecalis* in a rabbit model of infective endocarditis (Schlievert *et al*., 1998).

Pili of Gram-positive organisms have an important role in infection, participating in adhesion to human cells and biofilm formation. In *E. faecalis*, Ebp pili (endocarditis and biofilm-associated pili) and its associated sortase are involved in adhesion to components of the extracellular matrix and platelets (Nallapareddy *et al*., 2011). The Ebp subunits, which also contain LPXTG motifs, have been reported to be antigenic during human endocarditis (Sillanpää *et al*., 2004). In addition, its contribution to biofilm formation has been linked to the pathogenesis of experimental endocarditis and urinary tract infections (Kemp *et al*., 2007; Nallapareddy *et al*., 2006). Other genes related to biofilm formation have been shown to be important for virulence, such as the *bopABCD* operon (biofilm on plastic surfaces) or *bgsA* (biofilm-associated glycolipid synthesis A gene) (Hufnagel *et al*., 2004b; Theilacker *et al*., 2009).

E. faecalis is able to secrete proteins to the extracellular medium that contribute to the severity of the infection (Garsin *et al*., 2014). The cytolysin-haemolysin (Cyl) is a

toxin produced by some strains of *E. faecalis* (Arias and Murray, 2012). It is encoded on pheromone responsive plasmids or pathogenicity islands (Van Tyne *et al*., 2013). Cyl lyses a broad range of cells including Gram-positive bacteria and eukaryotic cells such as human erythrocytes (Van Tyne *et al*., 2013). Its contribution to the virulence *of E. faecalis* has been clearly demonstrated in all the animal models tested, ranging from invertebrates to mammals (Garsin *et al*., 2001; Ike *et al*., 1984; Singh *et al*., 1998). Epidemiological data has shown that Cyl is also associated with increased toxicity in human infections (Huycke and Gilmore, 1995; Huycke *et al*., 1991; Ike *et al*., 1987). Other secreted virulence factor is the metalloprotease gelatinase (GelE). It degrades host tissues and components of the host immune response (Park *et al*., 2008). It also activates an autolysin that leads to biofilm formation (Thomas *et al*., 2009). The *sprE* gene, which is located downstream of *gelE*, encodes another secreted protein, the serine proteinase (SprE). Both genes are cotranscribed (Qin *et al*., 2000) and have been shown to affect the pathogenesis of *E. faecalis* in numerous animal models and experimental infections (Engelbert *et al*., 2004; Singh *et al*., 2005; Singh *et al*., 1998 ; Thurlow *et al*., 2009c).

Even though a number of virulence determinants have been identified in *E. faecalis*, our understanding of the regulation mechanisms involved in its pathogenicity is still very limited. In general, the infection process can be conceived as an adaptation of the bacterium to particular host niches, and this adaptation requires a coordinated regulation in gene expression. In this context, global response regulators that activate and/or repress transcription of multiple genes in response to specific environmental signals are doubtless key elements. Among the global response regulators of *E. faecalis*, some of them form part of two-component signal transduction systems (TCSs). The TCSs consists of a membrane-associated histidine kinase receptor and a cytoplasmatic response regulator, which is usually a transcriptional regulator. In response to a specific signal, the histidine protein kinase is autophosphorylated. Then, it transfers the phosphoryl group to the cognate response regulator thus activating its function (Capra and Laub, 2012; Hancock and Perego, 2002). The well-studied *fsr* TCS has been implicated in the pathogenicity of *E. faecalis* (Mylonakis *et al*., 2002; Qin *et al*., 2000; Sifri *et al*., 2002). It consists of the *fsrA* and *fsrC* genes. Both genes, along with *fsrB* and *fsrD* constitute the Fsr quorum sensing system (Nakayama *et al*., 2006). The FsrC kinase responds to the signal peptide accumulation activating the global response regulator FsrA. Then, FsrA regulates the expression of numerous genes, including several virulence genes such as *gelE* and *sprE* (Bourgogne *et al*., 2006; Qin *et al*., 2000). In the last years, other global regulators not associated to sensor histidine kinases have been

reported as important for *E. faecalis* virulence. Among them, SlyA that belongs to the MarR/SlyA family (Michaux *et al*.), an AraC-type regulator designated PerA (Coburn *et al*., 2008; Maddox *et al*.), and Ers that is a member of the Crp/Fnr family (Riboulet-Bisson *et al*., 2008). Nevertheless, we are only beginning to unravel the regulatory networks of *E. faecalis*. To enhance our understanding of the gene regulation mechanisms that contribute to the pathogenesis of this bacterium, we have been working on the molecular characterization of MA*Efa* (this Thesis), a putative global regulator of the Mga/AtxA family. This family includes the virulence regulators AtxA, Mga and Mga*Spn* from the Gram-positive pathogens *Bacillus anthracis*, *Streptococcus pyogenes*, and *Streptococcus pneumoniae,* respectively.

2. The Mga/AtxA family of global response regulators

AtxA from *Bacillus anthracis*

B. anthracis, the etiological agent of anthrax, is a Gram-positive bacterium. Anthrax is mainly a disease of herbivores, but humans are also susceptible. Virulence of this bacterium is associated with the production of two major virulence factors, the tripartite toxin and the capsule. The toxin consists of three proteins, protective antigen, lethal factor and edema factor. The genes encoding these proteins (*pagA*, *lef* and *cya*, respectively) are located on the pXO1 virulence plasmid. The capsule biosynthesis operon, *capBCADE*, is located on pXO2, another virulence plasmid. AtxA (anthrax toxin activator) controls positively the expression of both virulence determinants. The *atxA* gene is located on plasmid pXO1, and its transcription level depends on temperature, redox potential, growth phase, and carbohydrate availability (reviewed in (Fouet, 2010)). Although some regulators have been implicated in the control of its transcription, only AbrB was shown to bind directly to the *atxA* promoter region (Strauch *et al*., 2005). At the post-transcriptional level, CO_2 has been shown to affect AtxA activity (Dai and Koehler, 1997; Hammerstrom *et al*., 2011). Moreover, it has been reported that CodY controls the intracellular accumulation of AtxA by an unknown mechanism (van Schaik *et al*., 2009). In addition to control the expression of key virulence genes, AtxA has been shown to function as a global regulator of gene expression. It regulates the expression of many genes located on the virulence plasmids and the chromosome. The regulated targets (positively or negatively) include genes predicted to encode secreted proteins and proteins with roles in transcriptional regulation and signaling (Bourgogne *et al*., 2003). Although the promoter regions of such genes do not seem to share sequence similarities, *in silico* and *in vitro* analyses

revealed that the promoter regions of several target genes are intrinsically curved (Hadjifrangiskou and Koehler, 2008). Currently, *in vitro* studies on the interaction of AtxA with DNA have not been reported and, therefore, its mechanism of action remains unknown.

AtxA is a 56 kDa protein (475 residues) whose three dimensional structure has been recently solved (Hammerstrom *et al*., 2015). AtxA crystallizes as a dimer and each monomer shows five distinct domains, which is in agreement with the previous predictions (Hammerstrom *et al*., 2011). The first two domains at the N-terminal region have structures indicative of DNA-binding. Residues 7 to 74 form a winged helix-turn-helix motif (HTH1 in Fig. 2) and residues 76 to 162 form a helix-turn-helix motif (HTH2 in Fig.2). The third and the fourth domains are phosphoenolpyruvate:carbohydrate phosphotranspherase system (PTS) regulation domains (PRDs), PRD1 (residues 170 to 262) and PRD2 (residues 276 to 391). The PRDs have been found in several transcriptional antiterminators and transcriptional activators that control the expression of genes involved in the uptake and metabolism of carbohydrates. The activity of such regulators is modulated by PTS-mediated phosphorylation. In response to the metabolic state of the cell and/or to the presence of a carbohydrate, the PTS components phosphorylate conserved histidine residues within the PRD domains (Deutscher *et al*.). In the case of AtxA, its activity is modulated by phosphorylation at H199 within PRD1 and H379 within PRD2 (Hammerstrom *et al*., 2015; Tsvetanova *et al*., 2007). At the C-terminal region of AtxA there is a domain (residues 391 to 475) that resembles an EIIB-like motif. Such a motif has been reported to be involved in AtxA multimerization (Hammerstrom *et al*., 2011)

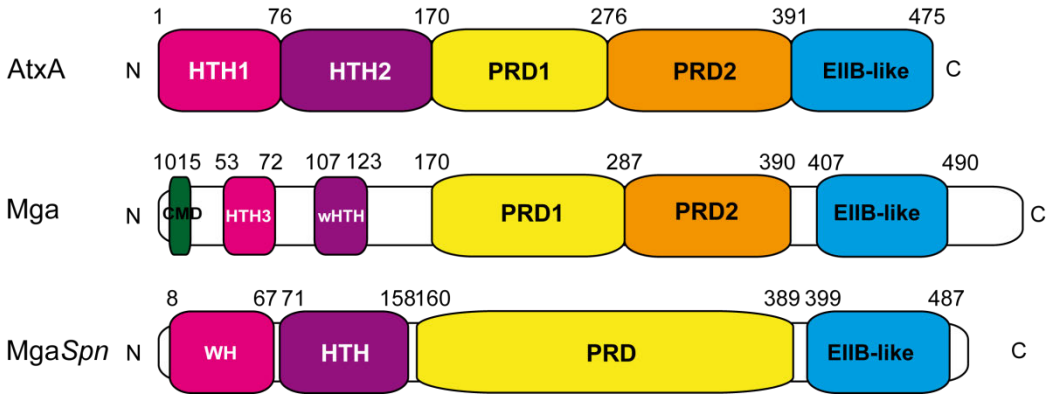


Figure 2. Organization of functional domains in AtxA, Mga and Mga*Spn*. AtxA has two N-terminal DNA-binding domains (HTH1 and HTH2). The third and fourth domains are PRDs. The C-terminal domain resembles an EIIB-like domain (Hammerstrom *et al*., 2015). Mga contains two helix-turn-helix motifs (HTH3 and HTH4) and a conserved region (CMD) at the N-terminus. The

central region has two putative PRDs. (PRD1 and PRD2). The C-terminal region contains a PTS EIIB-like domain (Hondorp *et al*., 2013). Mga*Spn* has two putative DNA-binding domains at the N-terminal region (wHTH and HTH motifs). The central region has structural homology to a PRD and the C-terminal region shows structural homology to an EIIB-like domain (Solano-Collado, 2014)

Mga from *Streptococcus pyogenes*

The Gram-positive bacterium *S. pyogenes* (group A Streptococcus (GAS)) is a strict human pathogen, causing a wide array of diseases ranging from trivial self-limiting infections to immune sequelae and life-threatening invasive disorders (Bisno *et al*., 2003; Cunningham, 2000). *S. pyogenes* is able to adapt and persist in various niches within the human body. In response to changing environmental conditions, a combination of TCSs and 'stand-alone' transcriptional regulators coordinate the control of gene expression (McIver, 2009). The term 'stand-alone' has been used to describe global response regulators that are not associated to a membrane-bound sensor histidine kinase. In many cases, their sensory components have yet to be fully defined. A key 'stand-alone' regulator of *S. pyogenes* is the global regulator Mga (multiple gene activator of GAS), which controls the expression of approximately 10% of the genome during the exponential phase of growth (Ribardo and McIver, 2006). The main Mga-activated regulon includes well established virulence genes, such as *emm* (M-protein), *scpA* (C5a peptidase), *sclA* (streptococcal collagen-like protein A), *fba* and *sof* (fibronecting binding proteins) and *sic* (streptococcal inhibitor of complement). These genes are important for adherence to host tissues, internalization into non-phagocytic cells and evasion of the host immune response (McIver, 2009; Ribardo and McIver, 2006). In addition, Mga activates or represses, possibly by an indirect mode, the expression of genes involved in the transport and utilization of carbohydrates as well as iron and amino acids (Ribardo and McIver, 2006).

The *mga* gene has been found in all strains examined and two divergent alleles have been identified (*mga-1* and *mga-2*). Expression of this regulatory gene has been linked to elevated CO_2, normal body temperature, increased iron levels and metabolizable sugars. Several regulators modulate the expression of *mga* (Hondorp and McIver, 2007; McIver, 2009). In addition, Mga activates directly the expression of its own gene (McIver *et al*., 1999).

Mga (62 kDa; 530 amino acids) binds to regions located upstream of its target promoters. The interaction of Mga with DNA was analyzed by EMSA and DNase I

protection experiments (McIver *et al*., 1995; McIver *et al*., 1999). In these experiments Mga fusion proteins were used, either Mga fused to a maltose binding protein (MBP) or a His-tagged Mga protein. Three categories of Mga-regulated promoters have been proposed based on the number of binding sites and their position relative to the start of transcription (Almengor and McIver, 2004). Sequence alignments of all established Mga-binding sites revealed that they exhibit very low sequence identity. Moreover, a mutational analysis in some target promoters indicated that Mga binds to DNA in a promoter-specific manner (Hause and McIver, 2012). The mechanism of transcriptional activation upon Mga binding to DNA remains unclear.

It has been predicted the organization of functional domains in Mga (Fig. 2) (Hondorp *et al*., 2013). At the N-terminal region, it has a conserved Mga domain (CMD, residues 10 to 15), a classical helix-turn-helix motif (HTH-3, residues 53 to 72) and a winged-helix-turn-helix motif (HTH-4, residues 107 to 126). Such motifs were shown to be involved in DNA binding and transcriptional activation. The HTH-4 motif is absolutely required for DNA binding to all Mga binding sites tested (Vahling and McIver, 2006). In the central region of Mga, two PRDs were identified: PRD1 (residues 170 to 287) and PRD2 (residues 288 to 390). Mga can be phosphorylated *in vitro* by components of the PTS system. This phosphorylation event was shown to inhibit the *in vitro* transcription of an Mga-regulated gene (Hondorp *et al*., 2013). Recently, it has been shown that Mga is phosphorylated *in vivo* at histidine residues (Sanson *et al*., 2015). Moreover, phosphorylation and non-phosphorylation mimicking substitutions at different histidines within the PRDs indicated that phosphorylation might modulate Mga activity (Hondorp *et al*., 2013; Sanson *et al*., 2015). The C-terminal region of Mga contains an EIIB-like domain. Elimination of this domain was reported to impair the ability of Mga to activate transcription *in vivo* and also its capacity to oligomerize *in vitro* (Hondorp *et al*., 2012).

Mga*Spn* from *Streptococcus pneumoniae*

S. pneumoniae usually resides in the nasopharynx of healthy humans. However, when the immune system weakens, it is also a major cause of life-threatening infections, such as pneumonia, meningitis and septicemia (Bogaert *et al*., 2004; Scott, 2007). In this Gram-positive bacterium, several genes important for its virulence have been identified by signature-tagged mutagenesis in the TIGR4 strain (Hava and Camilli, 2002). Among them, the *sp1800* gene, which encodes a protein that

was first named MgrA (Hemsley *et al*., 2003) and later Mga*Spn* (Solano-Collado *et al*., 2012). Mga*Spn* (493 residues) exhibits homology (42.6% similarity) to the Mga regulator of *S. pyogenes*. Mga*Spn* was shown to play a significant role in lung infection and nasopharyngeal carriage in a murine model. Additionally, Mga*Spn* was shown to represses the expression of several genes located within the *rlrA* pathogenicity islet (Hemsley *et al*., 2003). This islet has been found in a reduced number of strains (Paterson and Mitchell, 2006), whereas the *mgaSpn* gene is highly conserved among the pneumococcal genomes (Solano-Collado, 2014).

In Mga*Spn* (59 kDa), the predicted organization of functional domains is similar to that found in AtxA and Mga (Fig. 2). Mga*Spn* has two putative DNA-binding domains within the N-terminal region, the so-called HTH_Mga (residues 6 to 65) and Mga (residues 71 to 158) domains. The central region of Mga*Spn* has a putative PRD (residues 173 to 392) and the C-terminal region contains a putative EIIB-like domain (amino acids 399 to 487) (Solano-Collado, 2014). Mga*Spn* activates directly the transcription of a neighbouring four-gene operon of unknown function. This activation requires binding of Mga*Spn* to a site located upstream of the target promoter (Solano-Collado *et al*., 2012). Furthermore, the interaction of an untagged form of Mga*Spn* with linear double-stranded DNAs was analyzed by gel retardation, footprinting and electron microscopy experiments. These studies suggested that Mga*Spn* might recognize particular DNA conformations to achieve DNA-binding specificity. Upon binding to a primary site, Mga*Spn* is able to spread along the adjacent DNA regions generating multimeric protein-DNA-complexes (Solano-Collado *et al*., 2013).

3. The MA*Efa* protein of *Enterococcus faecalis*

Our understanding of the molecular mechanisms that control the virulence of *E. faecalis* is still very limited, even though the interest in the identification of virulence regulators increased considerably after the publication of the genome sequence of *E. faecalis* strain V583, the first vancomycin-resistant clinical isolate reported in the United States (Paulsen *et al*., 2003). Such a genome revealed that more than a quarter of the genome consists of mobile and/or exogenously acquired DNA, including probable integrated phage regions, insertion elements, conjugative transposons, a putative pathogenicity island, and integrated plasmid genes. Curiously, as pointed out by Paulsen *et al.* (2003), *E. faecalis* V583 possesses a modest collection of regulatory genes, in spite of the stress resistance capabilities of this bacterium. Several years

later, Bourgogne *et al*. (2008) published the genome sequence of *E. faecalis* strain OG1RF, a rifampicin and fusidic acid-resistant derivative of the OG1 human isolate. This strain has been largely used in the laboratories for genetic manipulation and virulence studies. Compared to V583, the OG1RF genome contains 227 unique open reading frames but has fewer mobile genetic elements. Furthermore, unlike strain V583, which harbours three plasmids, OG1RF is a plasmid-free strain. In 2008, searching for homologies, we found that the EF3013 gene (named *maEfa* herein) of the V583 strain encodes a protein (MA*Efa*) of 482 residues that has sequence similarity to regulators of the Mga/AtxA family. According to EMBOSS needle global sequence alignment (Rice *et al*., 2000), MA*Efa* has 40.7/22.1%, 31.3/19.6% and 38.8/23.7% of similarity/identity with AtxA (475 residues; strain Ames Ancestor), Mga (530 residues; strain MGAS10394) and Mga*Spn* (493 residues; strain R6), respectively. The enterococcal OG1RF genome also encodes the MA*Efa* protein (gene OG1RF_12293). Based on these observations, we proposed that MA*Efa* might function as a global transcriptional regulator involved in virulence, and since then we have been working on this hypothesis. First, we constructed several plasmid-based genetic tools for *E. faecalis* because, despite the increasing clinical significance of this bacterium, the tools available for its genetic manipulation are still very scarce. Then, we analyzed the expression of *maEfa* in enterococcal cells, studied the DNA-binding properties of MA*Efa*, and investigated its potential role in global regulation of gene expression. In 2011, when this Thesis was in progress, the three-dimensional structure of MA*Efa* was solved by X-ray crystallography and deposited in the Protein Data Bank (PDB 3SQN) (J. Osipiuk, R. Wu, S. Moy and A. Joachimiak, unpublished results). The functional characterization of MA*Efa* presented in this Thesis supports the notion that MA*Efa* is a new member of the Mga/AtxA family of global regulators. It might facilitate the adaptation of *E. faecalis* to particular host niches and, consequently, contribute to its potential virulence.

OBJECTIVES

We are interested in the molecular mechanisms that control the virulence of the opportunistic pathogen *E. faecalis.* Searching for homologies, we found that the EF3013 gene (named *maEfa* herein) of *E. faecalis* strain V583 encodes a protein (MA*Efa*; 482 amino acids) that has sequence similarity to regulators of the Mga/AtxA family. The enterococcal OG1RF genome also encodes the MA*Efa* protein (gene OG1RF_12293). This Thesis has been focused on the functional characterization of MA*Efa*. Our main aim has been to investigate whether MA*Efa* is a new member of the Mga/AtxA family of global regulators. To this end, we have worked on the following specific objectives:

1. Construction of promoter-probe and terminator-probe plasmid vectors in *E. faecalis*
2. Construction of expression plasmid vectors in *E. faecalis*
3. Analysis of *maEfa* gene expression in enterococcal cells
4. Purification of the MA*Efa* protein and analysis of its oligomerization state
5. Study of the interaction of MA*Efa* with linear double-stranded DNA
6. Analysis of the effect of MA*Efa* on global gene expression and virulence

MATERIALS AND METHODS

MATERIALS

1. Bacterial strains

Table 1. Bacterial strains used in this work

Strains	Features	Source
E. faecalis OG1RF	*gelE+, sprE +, fsrABCD*+, GelE+, SprE+ Rif^R, Fus^R. Plasmid-free	(Dunny *et al.*, 1978)
E. faecalis JH2-2	*gel*E+, *spr*E+, *fsrABCD*-, GelE−, SprE- Rif^R; Fus^R. Plasmid-free	(Jacob and Hobbs, 1974)
E. faecalis V583	*gelE*+, *sprE*+, *fsrABCD*+, GelE+, SprE + Van^R	(Sahm *et al.*, 1989)
E. faecalis OG1RFΔ*maEfa*	Strain derived from OG1RF. It lacks the *maEfa* gene.	This work
E. faecalis JH2-2Δ*maEfa*	Strain derived from JH2-2. It lacks the *maEfa* gene.	This work
S. pneumoniae 708	*end*-1, *exo*-1, *trt*-1, *hex*-4, *malM*594. Plasmid-free	(Lacks and Greenberg, 1977)
E. coli BL21 (DE3)	λ DE3 (*lac*I *lac*UV5-T7 gene 1 *ind*1 *sam*7 *nin*5) F- *dcm omp*T *hsd*S(*r*B - *m*B+) *gal.* Plasmid-free.	(Studier and Moffatt, 1986)

2. Animals

The animals used in this work were female BALB/c mice (8–10 weeks old) purchased from Harlan Laboratories. Mice were housed in a pathogen-free animal facility at the Helmholtz Centre for Infection Research (Braunschweig, Germany) and maintained under standard conditions according to institutional guidelines.

3. Culture media

E. faecalis was grown in Brain Heart Infusion (BHI) medium unless otherwise noted. Enterococcal cells harbouring plasmids were grown in media supplemented with tetracycline (Tc; 4 µg/ml), kanamycin (Km; 250-500 µg/ml) or erythromycin (Em; 5 µg/ml). Competent cells were grown in SG-ESTY medium (ESTY medium supplemented with 0.5 M of sucrose and 7-8% of glycine depending on the strain). After transformation, cells were incubated in S-ESTY-MC medium (ESTY containing 0.5 M of sucrose, 10 mM of $MgCl_2$ and 10 mM of $CaCl_2$). Tryptone-yeast extract (TY) medium (1% tryptone, 0.5% yeast extract, 0.5% NaCl) (Maniatis *et al.*, 1982) was used during the construction of the deletion mutant strains. To evaluate the utilization of different carbon sources, bacteria were grown in M9 minimal medium (Maniatis *et al.*, 1982) supplemented with 0.5% yeast extract (M9YE medium) and the indicated carbon source (1% glycerol, 0.5% maltose or 0.5% mannitol). When glycerol was used as carbon source, the medium was also supplemented with 0.26% fumarate.

For bacterial growth on solid media different agar plates were used. TY plates (TY and 1.5% agar) were used for routine growth. The medium was supplemented with X-Gal (5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside) (100 µg/ml) and/or Em (5 µg/ml) when required to generate the chromosomal deletion. SR medium (Shepard and Gilmore, 1995) supplemented with the appropriate antibiotic was used to select transformant cells. Bile esculin azide (BEA) plates were used to determine the amount of viable bacteria recovered after the intraperitoneal mice infections.

S. pneumoniae cells were grown in AGCH medium (Lacks, 1966; Ruiz-Cruz *et al.*) supplemented with 0.2% yeast extract and 0.3% sucrose. The medium contained Tc (1 µg/ml) or Km (30-50 µg/ml) when the pneumococcal cells carried plasmids. Competent cells were grown in AGCH medium supplemented with 0.2% sucrose and 70 µM $CaCl_2$. Plates for bacterial growth in solid medium were freshly prepared as it is reported in Methods, Section 1.

E. coli cells were grown in tryptone-yeast extract (TY) medium. In the case of plasmid harbouring cells, the medium was supplemented with Km (30-50 µg/ml). Plates were prepared with TY and 1.5% agar. For competence of *E. coli*, SOB medium (Hanahan, 1983) was used. After transformation, cells were recovered in SOC medium (SOB medium containing 20 mM glucose).

4. Enzymes, chemical products and reactives

Restriction enzymes, T4 DNA ligase, T4 polynucleotide kinase and bovine serum albumin (BSA) were acquired from New England BioLabs. DNase I RNase-free, isopropyl-β-Dgalactopyranoside (IPTG), proteases inhibitor cocktail EDTA free complete, proteases inhibitor cocktail complete, High Pure Plasmid Isolation kit, dNTPs and Taq DNA polymerase were provided by Roche Applied Science. Phusion High-Fidelity DNA Polymerase (Finnzymes) was used. ThermoScript reverse transcriptase and protein pre-stained standard SeeBlue Plus were purchased to Invitrogen. NZYDNA Ladder III weight marker was from NZYTech. Proteinase K, RNase A, lysozyme, Tc, Km, Em, mutanolysin, fucose, sucrose, polyethylenimine (PEI), imidazole, His-select nickel affinity gel, acrylamide:bis-acrylamide 40% solution (29:1 ratio) and DMSO were from Sigma Aldrich. GelRed was adquired from Biotium. Ethanol absolute, hydrochloric acid, chloroform, isopropanol, methanol, aminoacids, vitamins, glucose and magnesium were purchased to Merck. The saturated phenol was from AppliChem. Culture media components were from Pronadisa, Merck, Sigma, BD and Difco. Low molecular weight protein marker and the HiLoad Superdex 200 column were from Amersham Biosciences. DNA sequencing kit (Sequenase Version 2.0) and X-Gal were from USB Corporation. Radioactive nucleotides were purchased to Perkin-Elmer or Hartmann. Agarose, acrylamide:bis-acrylamide 30% solution (37.5:1 ratio), ammonium persulfate (PSA), β-mercaptoethanol, TEMED, sodium dodecyl sulfate (SDS), Triton-X100, Bio-Safe Coomassie Stain, Aurum Total RNA Mini kit, iScript Select cDNA synthesis kit, iQ SYBR Green Supermix, Immun-blot PVDF membranes and Immun-StarTM HRP substrate kit were provided by Bio-Rad. Acrylamide:bis-acrylamide 40% solution (19:1 ratio) was from National Diagnostics. The HiTrap Heparin HP and HisTrap HP columns and Illustra MicroSpinTM G-25 columns were from GE Healthcare. QIAquick Gel Extraction kit and RNeasy Mini Kit were from QIAGEN. Bacterial genomic isolation kit was purchased to Norgen Biotek Corporation. Dialysis membranes were from Spectrum. Slide-A-Lyzer™ Dialysis Cassettes were adquired from Thermo Scientific. Nitrocellulose filters of pore size 0.22 and 0.45 µm were from Millipore. Autoradiography films were acquired from Kodak (X-Omat S). Cronex Lightning Plus amplifying X-ray screens were from Dupont. Imaging plates to visualize radioactive labelling using the Fujifilm Image Analyzer FLA-3000 were from Fuji. Centrifugal filters Micro-sep and Nano-sep for protein concentration were from Pall. 96-well plates for fluorescence measurements were from Millipore. Electroporation

cuvettes of 2 mm were from Cell Projects. RNAsin Plus RNAse Inhibitor and DTT were acquired from Promega.

5. Nucleic acids

5.1. Plasmids

Table 2. Plasmids used in this work.

Plasmid	Size (bp)	Description	Source
pLS1	4,408	Derivative of pMV158 that lacks the 1,132-bp *Eco*RI restriction fragment; non-mobilizable; Tc^R	(Stassi *et al.*, 1981)
pGreenTir	3,484	Derivative of pUC1813 that carries a *gfp* allele fused to an optimized translation initiation region; Amp^R	(Miller and Lindow, 1997)
pPR54	5,240	Expression vector for *Bacillus subtilis* based on plasmid pUB110. It contains the transcriptional termination sites of the *E. coli* ribosomal RNA operon (Brosius, 1981).	(Serrano-Heras *et al.*, 2005)
pAS	5,210	Terminator-probe vector. Derivative of pLS1 that carries the *gfp* reporter cassette of the pGreenTIR plasmid; Tc^R	This work
pAST	5,456	Promoter-probe vector. Derivative of pAS that carries the transcriptional termination sites T1T2 of the *E. coli rrnB* ribosomal RNA operon downstream of the *tetL* gene; Tc^R	This work
pAST*2*T1*rrnB*	5,456	Derivative of pAS that carries the transcriptional termination sites T1T2 of the *E. coli rrnB* ribosomal RNA operon in the opposite orientation (compared to the pAST vector); Tc^R	This work

Plasmid	Size (bp)	Description	Source
pAST*rsiV*	5,475	Derivative of pAS that carries the transcriptional terminator of the *E. faecalis rsiV* gene; Tc^R	This work
pAST*polA*	5,448	Derivative of pAS that carries the transcriptional terminator of the *S. pneumoniae polA* gene; Tc^R	This work
pAST-*PsulA*	5,622	Derivative of pAST that carries the promoter of the *S. pneumoniae sulA* gene; Tc^R	This work
pAST-*Pung*	5,615	Derivative of pAST that carries the promoter of the *S. pneumoniae ung* gene; Tc^R	This work
pAST-*PuppS*	5,620	Derivative of pAST that carries the promoter of the *E. faecalis uppS* gene; Tc^R	This work
pAST-*P2493*	5,616	Derivative of pAST that carries the promoter of the *E. faecalis* EF2493 gene or *cpsC* gene; Tc^R	This work
pAST-*P2962*	5,614	Derivative of pAST that carries the putative promoter of the *E. faecalis* EF2962 gene; Tc^R	This work
pAST-*PfcsK*	5,573	Derivative of pAST that carries the inducible promoter of the *S. pneumoniae fcsK* gene; Tc^R	This work
pAST-o*PfcsK*	5,573	Derivative of pAST that carries the inducible promoter of the *S. pneumoniae fcsK* gene in the opposite orientation (compared to pAST-*PfcsK*); Tc^R	This work

Plasmid	Size (bp)	Description	Source
pAST-*Pma*	5,670	Derivative of pAST that carries the promoter region of the *maEfa* gene; Tc^R	This work
pAST-*Pma*Δ19	5,653	Derivative of pAS that carries the *Pma* region, but the last 19 nucleotides have been removed, including the -10 box; Tc^R	This work
pAS-*Pma*Δ19	5,407	Derivative of pAST that carries the *Pma* region, but the last 19 nucleotides have been removed, including the - 10 box; Tc^R	This work
pET-24b	5,309	*E. coli* expression vector based on the Φ10 promoter of phage T7; Km^R	Novagen
pET-*maEfa*	6,703	Derivative of pET-24b that encodes an untagged MA*Efa* protein*;* Km^R	This work
pET-*maEfa*-His	6,790	Derivative of pET-24b that encodes a His-tagged MA*Efa* protein*;* Km^R	This work
pDL287	5,740	Derivative of pVA380-1; Km^R	(LeBlanc *et al.*, 1993)
pDLF	5,911	Expression vector. Derivative of pDL287 that carries the promoter of *E. faecalis* EF2493 or *cpsC* gene; Km^R	This work
pDLS	5,921	Expression vector. Derivative of pDL287 that carries the promoter of *S. pneumoniae sulA* gene; Km^R	This work
pDLF*maEfa*$_{V583}$	7,425	Derivative of pDLF that carries the *E. faecalis* V583 *maEfa* gene under the control of the promoter *P2493*; Km^R	This work

Plasmid	Size (bp)	Description	Source
pDLS*maEfa*$_{V583}$	~4,500	Derivative of pDLF that carries the *E .faecalis* V583 *maEfa* gene under the control of the promoter of *S. pneumoniae sulA* gene; Km^R	This work
pBVGh	5,980	Vector for chromosomal modification. Based on the pG+host replicon. It carries a constitutive β-galactosidase gene; Em^R	(Blancato and Magni, 2010)
pBVΔ*maEfa*	6,967	Derivative of pBVGh used to delete the *maEfa* gene. Em^R	This work
pDLF*maEfa*$_{OG1RF}$	7,425	Derivative of pDLF that carries the *E. faecalis* OG1RF *maEfa* gene under the control of the promoter *P2493*; Km^R	This work
pMV158GFP	6,900	Derivative of pMV158 that carries the *gfp* gene under the control of an inducible promoter *Pm* of *S. pneumoniae* ; Tc^R	(Nieto and Espinosa, 2003)

5.2 Oligonucleotides

The oligonucleotides used in this work were all synthesized at the CIB-Protein Chemistry Facility (Applied Biosystems 3400 synthesizer), purified by HPLC and resuspended in water to a final concentration of 100 μM.

Table 3. Oligonucleotides used in this study.

Name	Sequence (5'-3')	Applications
F-*gfp*	CCATGATTACGCC*AAGCTT*GG	Construction of pAS
R-*gfp*	CCCCCGGGTACC***A****AGCT****T***GAATTCC	
Up-*gfp*	CAAGAGGGCAATGGCTGATA	pAS sequencing
Dw-*gfp*	GCTCTTCTTGGTACTGTTTTC	
Int-*gfp*	CATCACCATCTAATTCAACAAG	Primer extension
F-*T1T2rrnB*	CGATGGTAGTGTGGG*GTC****GA****C*CCATGCGAGA	Construction of pAST
R-*T1T2rrnB*	TGACGACAGGAAGAGTTT*GT****C****GA****C***ACGCAA	
F-*TrsiV*	GGATGCCCCCAGTGATG***G****TC****G****AC*CCTTC	Construction of pAST*rsiV*
R-*TrsiV*	TGTTATGCTTAATTCTA*GT****CG****AC*GCTTCT	
F-*TpolA*	ACTAAGATGCTGTTACAA*GT****CG****AC*GATGAA	Construction of pAST*polA*
R-*TpolA*	CTTGGATTGATAAATT***G****T****C****GAC*TCATAG	
F-*PsulA*	ACATGATTGTTAATG*GGAT****CC***CTTTCTG	Construction of pAST-*PsulA*
R-*PsulA*	TCACTCCCTCAA*GGAT****CC***TCATCATAT	
F-*Pung*	CGAAAGAGGTAGTA***GG****AT****C****C*TTAATGAT	Construction of pAST-*Pung*
R-*Pung*	TGTTCCATAGCCGACT***GGA****TC****C***TTTTTACTGC CTC	
F-*PuppS*	AAAATTTTA*G****A****GCT****C***GGCAGATAC	Construction of pAST-*PuppS*
R-*PuppS*	CCCTCCATTCCAA***G****A****GC****T****C*TATCTTAATT	
F-*P2493*	AATTAAATAG*G****A****G****C****TC*GGATGTTAAATATC	Construction of pAST-*P2493*
R-*P2493*	CACGATTGAACAAG***G****A****GC****T****C***AAATACATTATT	
F-*P2962*	CAATTAAGGCCCT***GG****AT****C****C*AGCAAAAAGT	Construction of pAST-*P2962*
R-*P2962*	CGCGCCCATTCACC*GGA****T****C****C***CTTAATC	
F-*PfcsK*	TATTATAGCACAA***TC****T****A****G*AGGAATTTG	Construction of pAST-*PfcsK*
R-*PfcsK*	CCATTTTTCTTCTC*TCT****AGA***TCCTTGATTAAC	
Up*maEfa*	GCAAAAGGAGGTTTT***CAT****ATG*TACTCCATG	Construction of pET-*maEfa* and pET-*maEfa*-His
Dw*maEfa*	AGCCAAAAAAC*T****C****G****A****G*AATGTCCTCGCTAG	
Dw*maEfa*His	CCTCGCTAGTTC***TCGAG***AAAATAAGAATGA	

Name	Sequence (5'-3')	Applications
A	GATTAGCTGAATGGTCATCGTGG	RT-PCR
B	GTTAAAACGTGTGATAACGG	
C	GACCAATCCCCTTTTTATCCG	
D	AATGAAGAGTGAGCTCTGCTAG	
Up*Pma*	AGGAATCAGTG***GA****GCTC*CTGTCGGTAA	Construction of pAST-*Pma*, pAST-*PmaΔ19* and pAS-*PmaΔ19*
Dw*Pma*	GTACATGGCAA***GA******G****CTC*CTTTTGCTT	
Dw*PmaΔ19*	CTCCTTTTGCTTAAG***AGC******G****TC*GGATAAAAAG	
P2493Cla-F	AATAGGTG*ATCG****AT***TGTTAAATATCTG	Construction of pDLF
P2493Cla-R	AACTCTCAC***AT****CGAT*TGAACAA*GCATG****C***AAAA TAC	
PsulCla-F	TGTTAATGGGAT**C**G**A**TTTCTGTTTG	Construction of pDLS
PsulCla-R	GACAT*ATCG****A****T*CACTCCC***G****CA****T****G****C***ATTTTCAT C	
maSph-F	TTTTTATCCGTATTC**GC**A**T**GCAAAAGGAGG	Construction of pDLF*maEfa*, pDLF*maEfa*V and pDLS*maEfa*V
maSph-R	AACCAAACGAT**G**CATG**C**CGAAAGAAAGC	
3012EMb	AGGAATGGCTGTTGTAACCA	EMSA, footprintings
*ma*EMb	CAACTGTTCCAACAAACG	
*ma*c	AACAAACGAATTTGCCGAAGC	
0091G2	GGCTATTTTGATGCACATATCTG	EMSA
0092A2	CCCGCCTTCCTTCCCTTGCTC	
EM1	AGTTGAATGTTTAAAGAAATGATGG	
26A	TTCTTTGTGGTATAATTGCAAGAGGT	
Oligo 40	TATATCATGCTATACCTATTCTTTGTGGTATAA TTGCAAG	EMSA (small DNA fragments)
Oligo 40pC	CTTGCAATTATACCACAAAGAATAGGTATAGC ATGATATA	
26A	TTCTTTGTGGTATAATTGCAAGAGGT	
26B	ACCTCTTGCAATTATACCACAAAGAA	
Oligo 32A	TTCTTTGTGGTATAATTGCAAGAGGTTTAATC	
Oligo 32B	GATTAAACCTCTTGCAATTATACCACAAAGAA	
Oligo 2	TATATTGTCTCCGTAGTGTT	
Oligo 2C	AACACTACGGAGACAATATA	

Name	Sequence (5'-3')	Applications
Up12292	CTAATCCGCCAGCGAATGTCTTTAACGAGG	Chromosome DNA sequencing
Rv12294	GCAATGATTTTCGTCCATGTTGTGAGTCTG	
Int12294	ATACGATTGCTAAAGCCGTG	
12292*Nco*	CGCAAACCTTTC***C****ATGG*TCAATACGACGC	Construction of pBVΔ*maEfa*
*maCla*R	CCTCCTTTTGCTTAAGAAT*ATCGAT*AAAAAG	
*maCla*F	CAAGAATTAATG*A****T****CGA****T***CAAGCCGCAG	
12294*Nco*	ATAACCAAACGAT***C****CATG****G***CGAAAGAAAG	
bga-R	ACACTCCATCAGACGGTTCG	pBVΔ*maEfa* sequencing
recA-F	GCAACGAAATGGTGGAACAG	qRT-PCR OG1RF_12439
recA-R	AAGGCATCGGCAATCTCTAAG	
mtlD-F	GGCGTTCATTACGTTGCTGA	qRT-PCR OG1RF_10298
mtlD-R	TCGCTTCGTCAATGGTTTGAT	
lacR-F	AGACGCCAATTTGTTAGAACG	qRT-PCR OG1RF_10456
lacR-R	CACAACTAGCGGCTAAAGAAG	
map-F	CAGAAGATGGCTTACACATTACCT	qRT-PCR OG1RF_10683
map-R	GAGTAGGCTGTCCATGTCGCT	
exp5-F	TAATGGTCGTTGTGGCAGTA AAT	qRT-PCR OG1RF_10684
exp5-R	TGCCCAACCGATACTCT	
11135-F	CGTTCGTAGTTTTGCTGTCA	qRT-PCR OG1RF_11135
11135-R	GAAGGGACAAAGCCGATTTCT	
gldA-F	GAGGGTGGCTTTAGTGGAGA	qRT-PCR OG1RF_11146
gldA-R	TTCACCTTCTGCTACGACTT	
glpK-F	AACAAGCCGCCTTATTTGGT	qRT-PCR OG1RF_11592
glpK-R	GGTTCTTCGCCAGTGTTCAT	
treB-F	ATGCTTATTCAGTTGCGACGA	qRT-PCR OG1RF_11753
treB-R	CTAACATGGCTGGGATAACTTG	
11763-F	AACATCGGCGGTATCTTCAG	qRT-PCR OG1RF_11763
11763-R	ATGCCTACATCACCAGTAGC	
selD-F	CTTGTACTGATGTGACTGGGTT	qRT-PCR OG1RF_11948
selD-R	CCAACGCTCCTGTAATGGTT	
uxuA-F	CCCGACGAATGATTTGCCTA	qRT-PCR OG1RF_12398
uxuA-R	ATCTAACGAACCAGCGACACT	

Name	Sequence (5'-3')	Applications
gnd2-F	GAAGCATTGCGTTTGGAGAT	qRT-PCR OG1RF_12405
gnd2-R	AGAAGCGACCACTTTGTTTG	
citF-F	ACTTGTTCGTGGACGGATTC	qRT-PCR OG1RF_12564
citF-R	TGCAATGCCAACTTCTGTTA	

6. Buffers

All buffers and solutions used in this work are listed in Table 4.

Table 4. Buffers and solutions.

Buffer	Composition	Application
Anode buffer	200 mM Tris-HCl, pH 8.9	Protein electrophoresis: SDS-PAGE (Tris-Tricine)
AU	50 mM Tris-HCl, pH 7.6 1 mM EDTA 0.1 mM DTT 3% Glycerol 200 mM NaCl	Protein analysis by analytical ultracentrifugation
BXF	80% Deionised formamide 10 mM NaOH 0.1 % Bromophenol blue 0.1% Xylene cyanol 1 mM EDTA	Loading-dye for DNA electrophoresis in denaturing PAA gels
Buffer gel	3 M Tris-HCl, pH 8.45 0.3% SDS	Protein electrophoresis: SDS-PAGE (Tris-Tricine)

Buffer	Composition	Application
BXGE 10X	0.25% Bromophenol blue 0.25% Xylene cyanol 60% Glycerol 10 mM EDTA	Loading-dye for gel electrophoresis of DNA and DNA-protein complexes
Cathode buffer	100 mM Tris-HCl, pH 8.25 100 mM Tricine 0.1% SDS	Protein electrophoresis: SDS-PAGE (Tris-Tricine)
Electroporation buffer	500 mM Sucrose 10% Glicerol	Storage of electrocompetent cells of *E. faecalis*
EB	200 mM NaCl 20 mM Tris-HCl, pH 8 2 mM EDTA	Elution of DNA from native PAA (5%) gels
LB	10 mM Tris pH 8.0 1 mM EDTA pH 8.0 1 mg/ml Lysozyme	Lysis buffer for RNA extraction
LBD	50 mM Tris pH 8.0 100 mM NaCl 25% sucrose 28 µg/ml RNase A 10 mg/ml lysozyme	Lysis buffer for genomic DNA extraction
LBF	25 mM Tris pH 7.6 0.5 mM EDTA pH 8.0 0.2 mg/ml Lysozyme 260 U Mutanolysin	Lysis buffer for total extract preparations from *E.faecalis* for Western blot

Buffer	Composition	Application
LBP	50 mM Tris-HCl, pH 7.6 1 mM EDTA 50 mM NaCl 0.1% Deoxycholate	Lysis buffer for total extract preparations from *S. pneumoniae* for Western blot
OEB	10 mM Magnesium acetate 200 mM Sodium chloride 0.1% SDS	Elution of oligonucleotides from denaturing PAA gels.
PBS	10 mM Na_2HPO_4 2 mM KH_2PO_4 2.7 mM KCl 137 mM NaCl pH 7.4	Resuspension of cells for fluorescence measurements. Western blot: Washing buffer
SB	PBS 0.05% Tween20 0.2% Casein	Western blot: Saturation Buffer
SLB 5X	250 mM Tris-HCl, pH 7.2 10 % SDS 3.5 M β-mercaptoethanol 50% Glycerol 0.5% Bromophenol blue	Loading-dye for protein electrophoresis (SDS-PAGE)
STOP DNase I	2 M Ammonium acetate 0.8 mM Sodium acetate 0.15 M EDTA	Stop solution for DNase I digestion
TAE	40 mM Tris 20 mM Acetic acid 2 mM EDTA pH 8.1	DNA electrophoresis in agarose gels

Buffer	Composition	Application
TB	25 mM Tris	Western blot:Transfer Buffer
TBE	89 mM Tris 89 mM Boric acid 2.5 mM EDTA pH 8.3	DNA electrophoresis in polyacrylamide gel
TE	10 mM Tris pH 8.0 1 mM EDTA pH 8.0	Storage of DNA
TG	50 mM Tris-HCl, pH 8.3 300 mM Glycine 0.1 % SDS 2 mM EDTA	Protein electrophoresis: SDS-PAGE (Tris-Glycine)
S-His	10 mM Tris-HCl, pH 7.6 5% Glycerol 300 mM NaCl 1 mM DTT	Purification of MA*Efa*-His protein. Buffer S-His was supplemented with imidazole (from 10 mM to 250 mM)
VL	50 mM Tris-HCl, pH 7.6 5% Glycerol 1 mM DTT 1 mM EDTA	Purification and storage of MA*Efa* protein. Buffer VL was supplemented with different concentrations of NaCl
WB	PBS 0.05% Tween 20	Western blot, Washing Buffer

7. Acrylamide solutions

For protein electrophoresis, an acrylamide:bis-acrlylamide 40% solution (40:1) or 30% solution (37.5:1) was used. For nucleic acids, an acrylamide:bis-acrlylamide

30% solution (30:0.8) or 40% solution (29:1) for non-denaturing gels (native gels) and an acrylamide:bis-acrlylamide 38% solution (38:2) or 40% solution (19:1) for denaturing gels, were used.

8. Software

Table 5. Bioinformatic tools

Application	Program	Company/webpage
Analysis DNA sequences	Chromas	mb.mahidol.ac.th/pub/chromas
	ApE	biology.utah.edu/jorgensen/wayned/ape
Analysis oligonucleotides	OligoAnalyzer 3.1	eu.idtdna.com
	Tm calculator	finnzymes.fi/tm_determination.html
Analysis qRT-PCR	iQ™5 Optical System Software	Bio-Rad
Bibliography manager	EndNote X	Thomson Reuters
Data analysis, statistics and graphing	SigmaPlot	SigmaPlot
	GraphPad Prism 5	GraphPad Software
	Microsoft Excel 2007	Microsoft
Homologies finder	BLAST	blast.ncbi.nlm.nih.gov/Blast.cgi
Image and figures processing	Adobe Illustrator CS2 and CS3	Adobe Systems Inc.
	Adobe Photoshop CS2 and CS3	
Non-labelled DNA visualization	Quantity One	Bio-Rad

Application	Program	Company/webpage
Oligonucleotide design	Primer_3	frodo.wi.mit.edu
	Primer Blast	ncbi.nlm.nih.gov/tools/primer-blast
Operon prediction	Biocyc	www.biocyc.org
Prediction of intrinsic curvature	Bend.it	hydra.icgeb.trieste.it/dna/bend_it.html
Prediction of functional domains	CDD	ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml
	Pfam	pfam.sanger.ac.uk
	Phyre2	www.sbg.bio.ic.ac.uk/phyre2
Primary protein structure analysis	ProtParam	web.expasy.org/protparam
Promoter finder	BPROM	linux1.softberry.com/berry.phtml
Radiolabelled DNA visualization	Image-reader	Fuji
	MultiGauge	Fuji
	Quantity One	Bio-Rad
Restriction maps	ApE	biology.utah.edu/jorgensen/wayned/ape
Sequence alignments	ClustalW	ebi.ac.uk/Tools/clustalw2
	Clustal Omega	ebi.ac.uk/Tools/msa/clustalo
Sedimentation equilibrium	HeteroAnalysis	biotech.uconn.edu

Application	Program	Company/webpage
Sedimentation velocity	SEDFIT	analyticalultracentrifugation.com
	SEDNTERP	jphilo.mailway.com

9. Autoradiography and radioactive material

The radioactive DNA was visualized either by autoradiography using the X-OmatS films (Kodak) or using a Phosphor screen scanned with a Fujifilm Image Analyzer FLA-3000 (Fuji). When autoradiography was used, the radioactive signal was amplified using the intensifying screens Cronex Lightning Plus (Dupon).

METHODS

1. Bacterial growth conditions

E. faecalis and *S. pneumoniae* were grown in liquid medium under low aeration conditions at 37ºC, unless otherwise indicated. Bacteria were grown in flasks whose volume was twice the culture volume and incubated in a static bath. When enterococci were grown in M9YE medium, microtiter plates were used and 200 µl/well were applied. Microplates were incubated at 37ºC without shaking in a Thermo Scientific Varioskan Flash instrument. In solid medium, enterococcal cells were uniformly spread over TY agar, SR medium or BEA agar plates. In the case of pneumococci, cells and antibiotic (when required) were mixed with a basal layer (20 ml) of AGCH medium supplemented with sucrose (0.3%) and yeast extract (0.2%) plus 1% agar. Then, an over-layer (8 ml) of AGCH medium plus 0.75% agar was added covering the basal layer. Plates were incubated at 37ºC, except when specified.

E. coli was grown in TY medium with rotary shaking using Erlenmeyer flasks. The flask volume was 5-times larger than the culture volume in order to maintain a constant aeration. In solid medium, bacteria were grown at the surface of TY plates. Cells were incubated at 37ºC.

Bacterial growth in liquid medium was followed by turbidity at 650 nm for *E. faecalis* and *S. pneumoniae* or at 600 nm for *E. coli,* using a Bausch & Lomb

(Spectronic 20D+) spectrophotometer. When the Thermo Scientific Varioskan Flash instrument was used, growth of *E. faecalis* was measured at 600 nm. For preservation of all bacterial cultures, cells were grown to an optical density (OD) between 0.3 and 0.4, which correspond with the exponential phase of growth. Then, sterile glycerol was added to 1 ml of culture to a final concentration of 10%. The culture was kept at 37°C for 10 min and then on ice for 10 min. Finally, cultures were stored at -80°C.

2. Bacterial transformation

2.1. Preparation of competent cells

E. faecalis electrocompetent cells were prepared basically as reported by Shepard and Gilmore (1995). Cells were grown overnight in ESTY medium at 37ºC without aeration. Subsequently, the culture was 100-fold diluted into SG-ESTY medium(supplemented with antibiotic if required). The volume of the flask was 5-times larger than the culture volume, and the incubation continued as before. After 20-24 h, when the cultures reached an OD_{560} of 0.1-0.2, cells were collected by centrifugation at 6,000 rpm in an Eppendorf F-34-6-38 rotor for 10 minutes at 4ºC. Then, cells were washed twice with ice-cold electroporation buffer and resuspended in a volume of ice-cold electroporation buffer equal to 1/100 of the initial culture volume. Finally, aliquots of the cultures (40 µl) were stored at −80°C.

S. pneumoniae competent cells were prepared as reported previously (Lacks, 1966). Mainly, cells were grown in ACGH medium supplemented with sucrose (0.3%) and antibiotic (if required) at 37ºC to an OD_{650} of 0.3 (exponential phase). The culture was then diluted 1:40 with prewarmed medium and cells were grown under the same conditions to an OD_{650} of 0.3. This dilution step was repeated twice. Finally, glycerol was added to a final concentration of 10%, and aliquots of the culture (250 µl) were stored at -80°C.

In the case of *E. coli,* electrocompetent cells were prepared from cultures grown with rotary shaking in SOC medium to an OD_{600} of 0.5 (exponential phase). Then, the culture was cooled on ice and centrifuged at 5,000 rpm in an Eppendorf F-34-6-38 rotor for 15 min at 4°C. Cell pellet was washed several times with ice-cold sterile water. Finally, cells were resuspended in 10% glycerol, and aliquots (50 µl) were stored at -80°C.

2.2. Transformation

2.2.1. Electroporation

E. coli and *E. faecalis* electrocompetent cells were transformed as reported by (Dower *et al*., 1988; Shepard and Gilmore, 1995), respectively. DNA in water was mixed with electrocompetent cells in a 0.2 cm electroporation cuvette (CellProjects) pre-cooled on ice. The electric pulse was generated with a MicroPulser (Bio-Rad) (2.50 kV and 3-5 ms). In the case of *E. faecalis*, after the electric pulse, 1 ml of ice-cold SG-ESTY-MC medium was added to the cells. The cuvette was maintained at least 5 minutes on ice and then the cells were incubated at 37ºC for 2 h without aeration. Finally, the enterococcal cells were spread over SR plates supplemented with the appropriate antibiotic for selection of transformants. *E. coli* cells, after the pulse, were transferred to 0.8 ml SOB medium containing glucose (0.4%) and incubated at 37ºC with rotary shaking for 1 h. In order to select transformants, TY-agar plates containing the appropriate antibiotic were used.

2.2.2. Natural transformation

Competent cultures of *S. pneumoniae* were transformed following the method described by Lacks (1966). Briefly, cells were inoculated in AGCH medium supplemented with sucrose (0.2%) and $CaCl_2$ (70 μM) and incubated at 30°C for 20 min before addition of DNA. After addition of DNA, cultures were incubated at 30°C for 40 min. To enable the phenotypic expression (resistance to Tc), cultures were incubated at 37°C for 90 min. Transformants were selected in AGCH plates (described in Section 1) supplemented with Tc.

3. Construction of bacterial strains

3.1. Construction of strains OG1RFΔ*maEfa* and JH2-2Δ*maEfa*

Strain OG1RFΔ*maEfa* was constructed by deleting the chromosomal region that spans coordinates 2421575-2422640 (which corresponds to coordinates 2889057 and 2890122 in V583 genome, see Results Section 5.1. and Fig. 13) using the plasmid pBVGh (Blancato and Magni, 2010), kindly provided by Dr. C. Magni. The pBVGh plasmid carries a thermosensitive replication origin and a constitutively expressed β-galactosidase gene. This plasmid is suitable to delete specific chromosomal regions. It allows an easy modification of chromosomal genes and a rapid identification of

recombinant clones by colorimetric screenings. First, a 560-bp DNA region located upstream of the *maEfa* gene (locus_tag OG1RF_12293) was amplified by PCR using primers 12292*Nco* and *maCla*R, which introduced *Nco*I and *Cla*I restriction sites, respectively. Also, a 494-bp region that includes the 3´-end of *maEfa* was amplified using primers *maCla*F and 12294*Nco*, which introduced *Cla*I and *Nco*I sites, respectively. Both DNA fragments were digested with *Cla*I enzyme, mixed in equimolecular amounts and ligated with T4 DNA ligase. The ligation mixture was used as template for a PCR using primers 12292*Nco* and 12294*Nco.* The 1017-bp product was then digested with *Nco*I and ligated into the *Nco*I site of plasmid pBVGh. The ligation mixture was used to transform JH2-2 electrocompetent cells. After electroporation, cells were incubated for 150 min at 30ºC and plated on SR agar supplemented with X-Gal 100 µg/ml and Em 5 µg/ml. Plates were incubated at 30ºC for 36 h (permissive temperature for plasmid DNA replication) and several blue transformant colonies were obtained. Subsequently, plasmid DNA from blue colonies was isolated and analysed by restriction. The inserted region in the recombinant plasmid (pBVG*ΔmaEfa,* see Fig. 3) was verified by dye-terminator sequencing carried out at Secugen (Centro de Investigaciones Biológicas, Madrid). Then, plasmid pBVG*ΔmaEfa* was introduced into OG1RF electrocompetent cells.

To generate the *maEfa* deletion in *E. faecalis* (strains JH2-2 and OG1RF) a protocol based on the one described by Blancato and Magni (2010) was followed. First, cells carrying the recombinant plasmid (pBVG*ΔmaEfa*) were grown overnight in TY medium supplemented with Em (5 µg/ml) at 42ºC. Then, they were plated in TY agar containing Em (5 µg/m)l and X-gal (100 µg/ml). Plates were incubated at 42ºC overnight. In this step, cells were grown under selective pressure (Em) and at the non-permissive temperature for plasmid DNA replication. Thus, cells carrying the plasmid integrated into the chromosome were selected. The integration occurred by a single recombination event between homologous sequences present in the plasmid and the chromosome (Fig. 3). Integration was verified in several blue colonies by colony PCR (described in Section 4.3.2) using the oligonucleotides *bga*-R and Rv12294. One of the positive colonies was inoculated in TY medium supplemented with Em (5 µg/ml) and incubated overnight at 42ºC. Subsequently, the culture was 1500-fold diluted in TY medium without antibiotic. Cells were grown at 30ºC for 7 h to allow the excision of the integrated plasmid by a second recombination event (Fig. 3). Next, the culture was diluted 100-fold in fresh medium and incubated at 42ºC for 6 h, in order to lose the excised plasmid. After that, the culture was plated onto TY agar containing X-Gal (100 µg/ml) and incubated overnight at 42ºC. White colonies (no plasmid DNA) selected in

this step had either the *maEfa* deletion or the parental chromosome (see Fig. 3). To discriminate between both possibilities, colony PCR (described in Section 4.3.2) was performed using the oligonucleotides Up12292 and Rv12294. When cells carried the *maEfa* deletion, a 594-bp PCR product was detected. When cells had the wild-type chromosome, a 1654-bp PCR product was detected. Then, we confirmed the sensitivity of the deletion mutant cells to Em at both temperatures, 30ºC and 42ºC. Finally, the chromosomal deletion was confirmed by DNA sequencing. Specifically, chromosomal DNA was isolated from the candidate strains, and a 2,500-bp region that includes the deletion was amplified by PCR using oligonucleotides Up12292 and Int12294. The PCR product was then sequenced at Secugen (Centro de Investigaciones Biológicas, Madrid).

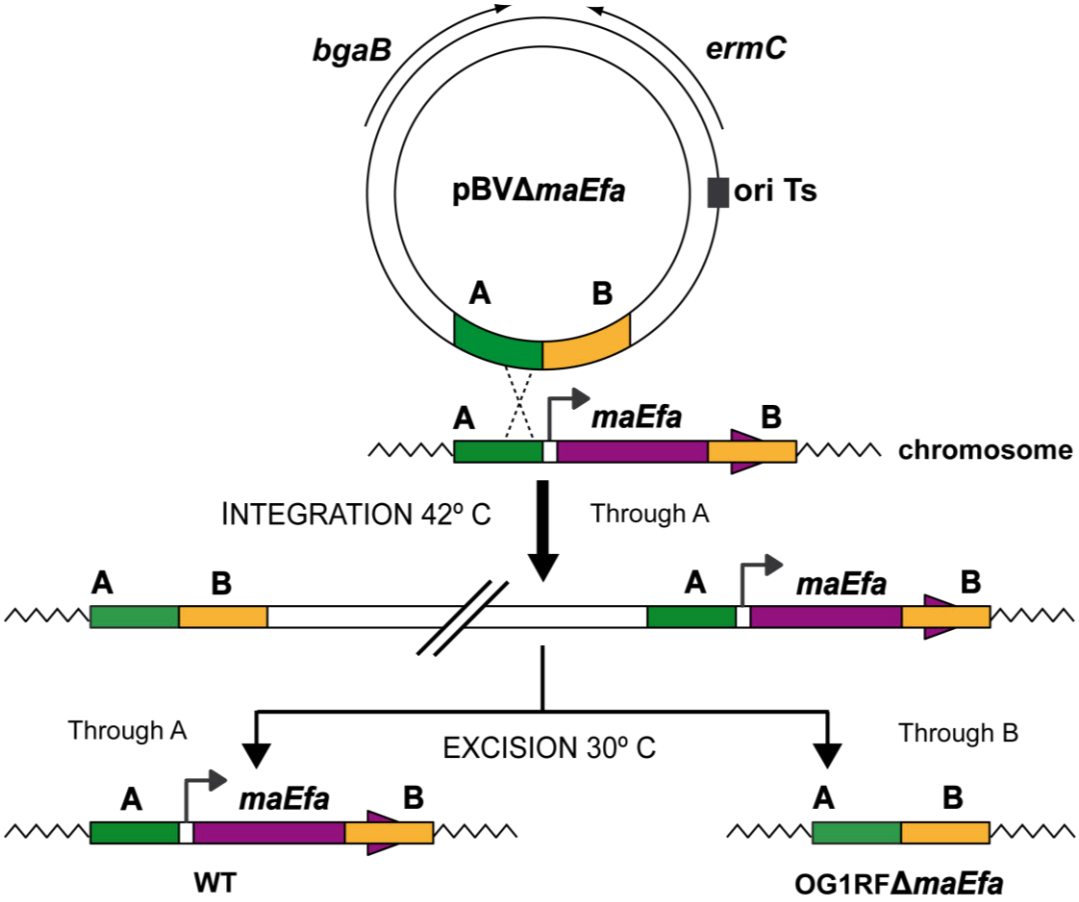


Figure 3. Deletion of the *maEfa* gene. First step: integration of pBVGΔ*maEfa* into the chromosome (selection at 42ºC with Em) ocurrs by a single recombination event through region A. Second step: the incubation at 30ºC stimulates a second recombination event between the duplicated homologous regions and leads to plasmid excision. Excision through region A restores the parental chromosome, whereas excision through region B gives rise to the deletion of *maEfa*. Subsequently, cells are incubated at 42ºC in order to lose the excised plasmid. Alternatively, an initial integration through B and excision through A also result in *maEfa* gene deletion. Modified from (Biswas *et al*., 1993)

4. DNA preparations

4.1. Plasmid DNA isolation

Plasmid DNA was isolated using the High Pure Plasmid Isolation Kit (Roche Applied Science) with some modifications in the composition of specif buffers when enterococcal and pneumococcal cells were used. The *Suspension Buffer* was supplemented with 50 mM glucose, 1.2 mg/ml lysozyme and 240 units/ml mutanolysin for enterococcus, or 50 mM glucose and 0.1% deoxycholate in the case of *S. pneumoniae*. For both bacteria, the *Lysis Buffer* was freshly prepared. It contained 0.17 M NaOH and 1% SDS.

4.2. Genomic DNA isolation

In order to isolate chromosomal DNA of *E. faecalis,* cells were grown in BHI medium containing 1.25% of glycine. Cells were incubated at 37ºC with rotary shaking in a flask whose volume was 5 times larger than the culture volume. When the cultures reached an OD_{650} of 1.2, cells were harvested and concentrated 10-fold in buffer LBD (Table 4) at 4ºC. Subsequently, 400 units of mutanolysin were added to 1 ml of the concentrated culture. After 20 min at 37ºC, SDS was added at a final concentration of 1%. The lysate was treated with proteinase K (240 µg/ml) for 15 min. DNA was further purified by extraction with phenol/chloroform, dialyzed against buffer TE and recovered by precipitation with ethanol. For small-scale preparations of chromosomal DNA, bacteria were grown under routine growth conditions and the Bacterial Genomic Isolation Kit (Norgen Biotek Corporation) was used.

4.3. Preparation of linear double-stranded DNA fragments

Linear double-stranded DNA (dsDNA) fragments used in this Thesis were obtained by digestion with restriction enzymes, by polymerase chain reaction (PCR) or by annealing of complementary oligonucleotides.

4.3.1. Digestion with restriction enzymes

DNA digestions using restriction enzymes were performed following the indications specified by the suppliers. When needed, 10 µg/ml of BSA was included in the digestion reaction. In general, enzymes were inactivated at 65ºC for 10 min.

4.3.2. Polymerase chain reaction

All PCR reactions were carried out in an iCycler Thermo Cycler (Bio-Rad). Analysis of transformants by PCR (colony PCR) was performed with Taq DNA polymerase (Roche). The transformant colonies were inoculated with a sterile toothpick in both growth medium (50 µl) and sterile water (50 µ)l. Then, 1 µl of the cells suspended in water was used as template for PCR. Reactions (25 µl) contained 10 mM Tris-HCl Ph 7.8, 50 mM KCl, 1.5 mM $MgCl_2$, 12.5 pmol of each primer, 200 µM of each dNTP and 0.75 unit of Taq DNA polymerase. PCR conditions were: initial denaturation step at 94ºC for 3 min, followed by 30 cycles including the next steps: denaturation at 94ºC for 45 s, annealing of the primers to the DNA template at around 55ºC (depending on the primer Tm) for 20-30 s followed by an extension at 72ºC for 20-40 s (depending on the amplicon length). A final extension step was performed at 72ºC for 10 min.

Phusion High-Fidelity DNA polymerase (Finnzymes) was used when high fidelity PCR products were required. Reaction mixtures (50 µl) contained 5-30 ng of template DNA, 50 pmol of each primer, 200 µM of each dNTP and 1 unit of DNA polymerase. PCR conditions were as described with Taq DNA polymerase, but the denaturation steps were performed at 98ºC.

4.3.3 Annealing of complementary oligonucleotides

Complementary oligonucleotides were annealed in buffer TE (2 mM Tris-HCl, pH 8, 0.2 mM EDTA) containing 50 mM NaCl. Equimolar amounts of the complementary oligonucleotides were used. Reaction mixtures (150 µl) were incubated at 95ºC for 10 min and then cooled down slowly to 37ºC. Then, they were kept at this temperature for 10 min and on ice for 10 min.

4.4. DNA purification

4.4.1 Linear double-stranded DNAs

When it was required, linear dsDNA fragments (obtained either by digestion with restriction enzymes or by PCR) were purified with the QIAquick PCR purification kit (QIAGEN). The linear dsDNA fragments radioactively labelled at 5'-ends (obtained by PCR) were purified from non-denaturing polyacrylamide (5%) gels. After electrophoresis, DNA was visualized by autoradiography. The desired band was excised with a clean scalpel and the gel slice was incubated overnight in elution buffer

(see Table 4) with continuous shaking (450 rpm) at 42ºC using a Thermo-shaker (TS-100, Biosan). The remains of the gel were removed using a Spin-X column (Costar). Finally, the eluted DNA was ethanol precipitated and dissolved in distilled water.

4.4.2. Oligonucleotides

Oligonucleotides used to obtain small dsDNA fragments (40-bp to 20-bp) were purified from denaturing (8 M urea) polyacrylamide (15-20%) gels (Maniatis *et al*., 1982). The oligonucleotide solution was mixed with deionised formamide (50% final concentration) and heated at 55ºC for 5 min. The sample was cooled down on ice before being loaded onto the sequencing gel. After electrophoresis, the gel was placed on a piece of saran wrap and then on a Fluor-coated TLC plate (Ambion) and shined with a short wavelength UV light. The DNA absorbs the UV light and casts a shadow against the fluorescent background allowing the DNA to be visualized. The area of the gel containing the full-length oligonucleotide was excised. To elute the oligonucleotide, the gel slice was incubated in buffer OEB (see Table 4) overnight at 42ºC with continuous shaking. The remains of gel in the eluted DNA solution were removed using a Spin-X column (Costar). Finally, the eluted oligonucleotide was ethanol precipitated, dissolved in 50 µl of distilled water and loaded onto a MicroSpinTM G-25 column (GE Healthcare) to eliminate the remains of salt.

4.5. Ligation

In general, for ligation of linear DNA fragments, a vector to insert molar ratio of 1:10 or 1:15. When the vectors were dephosphorylated, the vector to insert molar ratio was between 1:5 and 1:10. In all cases, 400 units of T4 DNA ligase (New England Biolabs) were added to a final reaction volume of 20-40 µl. Reaction mixtures were incubated overnight at room temperature.

4.6. Construction of recombinant plasmids

All the constructions described below were confirmed by dye-terminator sequencing at Secugen (Centro de Investigaciones Biológicas, Madrid).

4.6.1. Construction of plasmids pET24b-*maEfa* and pET24b-*maEfa*-His

To overproduce and purify the MA*Efa* and MA*Efa*-His proteins, plasmids pET24b-*maEfa* and pET24b-*MaEfa*-His were constructed (see Section 7). In both cases, the pET24b expression vector (Novagen) was used. To obtain pET24b-*maEfa*

plasmid, a 1,502-bp region of the V583 chromosme containing the *maEfa* gene was amplified by PCR using the Up*maEfa* and Dw*maEfa* primers. These primers contained a single restriction site for *Nde*I and *Xho*I, respectively. The amplified product was digested with both enzymes, and the 1,470-bp digestion product was inserted into the pET24b expression vector. To overproduce the MA*Efa*-His protein, the *maEfa* gene was engineered to encode a MA*Efa* protein fused to a C-terminal His_6-tag. Specifically, a 1,481-bp region of the V583 genome was amplified by PCR using the Up*maEfa* and Dw*maEfa*His primers. Dw*maEfa*His oligonucleotide was designed to change the stop codon of *maEfa* gene by a restriction site for *Xho*I. Subsequently, the PCR product was digested with *Nde*I and *Xho*I and the 1,448-bp digestion product was cloned into the pET24b expression vector.

4.6.2 Construction of the pAS terminator-probe vector and the pAST promoter-probe vector

To construct the terminator-probe vector, pAS, an 833-bp DNA region of the pGreenTIR plasmid (Miller and Lindow, 1997), which contains the *gfp* reporter cassette, was amplified by PCR using the oligonucleotides F-*gfp* and R-*gfp*. Both oligonucleotides include a *Hind*III restriction site. The PCR-synthesized DNA fragment was digested with *Hind*III and the 802-bp digestion product was inserted into *Hind*III-linearized plasmid pLS1 (Lacks *et al.*, 1986). The promoter-probe vector, pAST, was constructed introducing the transcriptional termination sites *T1T2* of the *E.coli rrnB* ribosomal RNA operon (Brosius, 1981) into the multicloning site of pAS plasmid. For PCR amplification of a DNA region containing such terminators we used the pPR54 plasmid as template (Serrano-Heras *et al.*, 2005) and the oligonucleotides F-*T1T2rrnB* and R-*T1T2rrnB* as primers. The 286-bp DNA fragment was further digested with *Sal*I, and the 246-bp digestion product was cloned into the *Sal*I site of plasmid pAS in both orientations, obtaining plasmid pAST (orientation *T1T2rrnB*; promoter-probe vector) and plasmid pAS-*T2T1rrnB* (opposite orientation).

4.6.3. Construction of pAS and pAST derivatives

To assess whether pAS functions as a terminator-probe vector, several derivatives were constructed, including pAST, whose construction was described in the previous section. To generate pAS-*TrsiV*, a 305-bp region of the enterococcal genomic DNA that contains the putative terminator of the *sigV–rsiV* operon (Benachour *et al.*, 2005) was amplified with the F-*TrsiV* and R-*TrsiV* primers. Then, the PCR-amplified DNA was digested with *Sal*I, and the 265-bp restricted fragment was inserted into the

*Sal*I site of the pAS vector. For the construction of plasmid pAS-*TpolA*, a 278-bp region of the pneumococcal genome containing the terminator of the *polA* gene (López *et al.*, 1989) was amplified with the F-*TpolA* and R-*TpolA* primers. After *Sal*I digestion, the resulting 238-bp fragment was cloned into the *Sal*I site of the pAS vector.

To analyse whether pAST is suitable as a promoter-probe vector, various DNA regions containing defined or predicted promoters were inserted into the multicloning site of the plasmid. From the enterococcal genome, two regions of 192-bp and 190-bp containing the promoter of the *uppS* and EF2493 genes (Hancock *et al.*, 2003), respectively, were amplified with the F-*PuppS* and R-*PuppS* primers or the F-*P2493* and R-*P2493* primers. After *Sac*I digestion, the 164-bp (*PuppS* promoter) and 160-bp (*P2493* promoter) restriction fragments were cloned into the *Sac*I site of the pAST vector, generating plasmids pAST-*PuppS* and pAST-*P2493*, respectively. Moreover, to construct plasmid pAST-*P2962*, a 191-bp region of the enterococcal genome that contains the putative promoter of the EF2962 gene was amplified with the F-*P2962* and R-*P2962* primers. After *Bam*HI digestion, the 158-bp restriction fragment (*P2962* promoter) was inserted into the *Bam*HI site of the pAST vector. Using pneumococcal genomic DNA as template, two regions of 199-bp and 195-bp containing the promoter of the *sulA* (Lacks *et al*., 1995; López *et al*., 1987) and *ung* (Méjean *et al.*, 1990) genes, respectively, were amplified with the F-*PsulA* and R-*PsulA* primers or the F-*Pung* and R-*Pung* primers. The PCR-synthesized DNAs were further digested with *Bam*HI. The 166-bp (*PsulA* promoter) and 159-bp (*Pung* promoter) digestion products were inserted into the *Bam*HI site of the pAST vector, generating plasmids pAST-*PsulA* and pAST-*Pung*, correspondingly. Regarding the inducible pneumococcal *PfcsK* promoter (Chan *et al.*, 2003), a 150-bp region was amplified using pneumococcal genomic DNA as template and the oligonucleotides F-*PfcsK* and R-*PfcsK* as primers. After *Xba*I digestion, the 117-bp restriction fragment (*PfcsK* promoter) was cloned into the *Xba*I site of the pAST vector in both orientations: plasmid pAST-*PfcsK* (gene *gfp* under the control of the *PfcsK* promoter) and plasmid pAST-o*PfcsK* (opposite orientation).

4.6.4. Construction of plasmids pAST-*Pma*, pAST-*PmaΔ19* and pAS-*PmaΔ19*

The pAST-*Pma* recombinant plasmid was used to identify the transcription start site of the *maEfa* gene. To construct pAST-*Pma*, a 242-bp region of the V583 enterococcal genome was amplified by PCR using the Up*Pma* and Dw*Pma* primers. Both oligonucleotides contained a restriction site for *SacI*. The amplified DNA was digested with *Sac*I, and the 214-bp digestion product (coordinates 288864-2889078)

was inserted into the *Sac*I site of pAST. In pAST-*Pma*, *gfp* expression is under the control of the *Pma* promoter. To construct pAST-*Pma*Δ19 and pAS-*Pma*Δ19, a 228-bp region of the V583 genome was amplified with the Up*Pma* and Dw*Pma2* primers. After *Sac*I digestion, the 197-bp restriction fragment (coordinates 2888864-2889059) was cloned into the *Sac*I site of both pAST (pAST-*PmaΔ19*) and pAS (pAS-*PmaΔ19*).

4.6.5. Construction of pDL287 derivatives plasmids

Several plasmids were constructed to provide the *maEfa* gene *in trans*. They are based on the pDL287 plasmid (LeBlanc *et al*., 1993)**,** a derivative of pVA380-1, that carries a Km resistance gene. First, two expression vectors for *E. faecalis* were developed*,* pDLF and pDLS. These plasmids contain an engineered unique restriction site for *Sph*I downstream of the *P2493* and *PsulA* promoters, respectively. For the construction of pDLF, a 194-bp region of the V583 enterococcal genome, which contains the *P2493* promoter, was amplified using the oligonucleotides *P2493Cla*-F and *P2493Cla*-R. Both oligonucleotides carry a restriction site for *Cla*I. The *P2493Cla*-R carries also a *Sph*I restriction site. The PCR product was digested with *Cla*I generating a 171-bp fragment which was inserted into *Cla*I-linearized pDL287. For the construction of pDLS, a 202-bp fragment that carries the *PsulA* promoter was amplified using pneumococcal chromosomic DNA and the primers *PsulCla*-F and *PsulCla*-R. Both primers contained a restriction site for *ClaI.* In addition, *PsulCla*-R includes a restriction site for *Sph*I. After *Cla*I digestion, the 181-bp DNA fragment was cloned into the *Cla*I site of pDL287. The next step was to introduce a promoterless *maEfa* gene from V583 strain (*maEfa*$_{V583}$) in the expression vectors. To this end, a 1,546-bp region of the V583 enterococcal genome was amplified using the oligonucleotides *maSph*-F and *maSph*-R. After *Sph*I digestion, the 1,514-bp restriction fragment was inserted into the *Sph*I site of pDLF and pDLS, obtaining pDLF*maEfa*$_{V583}$ and pDLS*maEfa*$_{V583}$, respectively. In addition, the *maEfa* gene of strain OG1RF (*maEfa*$_{OG1RF}$), which was amplified by PCR with the same primers, was introduced into pDLF, obtaining pDLF*maEfa*$_{OG1RF}$.

4.7. Radioactive labelling of DNA

Radiolabelled DNA was visualized either by autoradiography or using a Fujifilm Image Analyzer FLA-3000 (Phosphorimager). The intensity of the labelled DNA bands was quantified using the Quantity One software (Bio-Rad)

4.7.1. 5´-end labelling

Radioactive labelling of the oligonucleotides at their 5' ends was performed using [γ-^{32}P]-ATP and the T4 Polynucleotide Kinase (PNK) (New England Biolabs). Reactions mixtures (25 µl) contained 25 pmol of oligonucleotide, 2.5 µl of 10x kinase buffer (provided by the supplier), 41.5 pmol of [γ-^{32}P]-ATP (3,000 Ci/mmol; 10 µCi/µl) and 10 units of T4 PNK. After incubation at 37ºC for 30 min, additional T4 PNK (10 units) was added and the reaction mixtures were incubated at 37ºC for 30 min. Finally, to inactivate the enzyme, reaction mixtures were incubated at 65ºC for 20 min. Non incorporated nucleotide was removed using MicroSpinTM G-25 columns (GE Healthcare). The 5´-labelled oligonucleotides were used to obtain labelled-dsDNA by PCR amplification (labelling at the 5´ end of one DNA strand).

4.7.2. Internal labelling

The incorporation of a radiolabelled nucleotide ([α-^{32}P]-dATP, 3,000 Ci/mmol, 10 µCi/µL, Hartmann) was used for primer extension experiments and manual sequencing (using the Sequenase v2.0 kit, according to the indications of the supplier).

5. Analysis of DNA

5.1. DNA quantification

To quantify non-radiolabelled DNA, agarose gels were stained either with EtBr (1 µg/ml) or GelRed 1x (Biotium). DNA bands were visualized using a Gel Doc system (Bio-Rad) and quantified using a molecular weight marker (NZYDNA Ladder III) with the Quantity One program. In addition, the concentration of DNA samples was determined using a NanoDrop ND-2000 Spectrophotometer (Bio-Rad). For 5´-labelled DNA, the ionizing radiation was measured with a scintillation counter (Wallac1450 MicroBeta, TriLux). Knowing that 125 µCi of [γ-^{32}P]-ATP (41.5 pmol) are equivalent to 1.37x10^{8} cpm, we estimated the incorporation of [γ-^{32}P]-ATP in the labelling reaction and then the DNA concentration using the total cpm obtained.

5.2. DNA electrophoresis

5.2.1. Agarose gels

In general, DNA was analysed by horizontal agarose gel electrophoresis in TAE buffer. The agarose concentration used (generally 0.8-1.2 %) depended on the size of

the DNA analysed. DNA samples were mixed with BXGE buffer (see Table 4) and loaded onto the gel. After electrophoresis, gels were stained in a solution of EtBr (1 µg/ml in TAE buffer) or GelRed (1x in water supplemented with 0.1 M NaCl) for 15-30 min at room temperature. DNA was visualized using a short wavelength UV light (254 nm) with a Gel Doc XR system (Bio-Rad). The image obtained was captured with the Quantity One software and quantified when it was necessary.

5.2.2. Native polyacrylamide gels

Vertical polyacrylamide gels were run in TBE buffer (see Table 4) using a Mini Protean-III system (Bio-Rad). The concentration of acrylamide used (5-8%) depended on the size of the DNA analysed. DNA samples were mixed with BXGE buffer and loaded onto the gels. Gels were run at 100 V at room temperature or 4ºC. After electrophoresis, gels were soaked in GelRed 1x and visualized as described before. When radioactively labelled DNA was used, after electrophoresis, gels were fixed with acetic acid (10%), dried using a gel dryer (model 583, Bio-Rad) and the DNA was visualized by autoradiography or using a Fujifilm Image Analyzer FLA-3000.

5.2.3. Denaturing polyacrylamide gels

Oligonucleotides were purified (when needed) from denaturing polyacrylamide gels in TBE buffer and 8 M urea. The concentration of acrylamide used depended on the size of the oligonucleotide: 19% for 15-25 nt and 15% for 25-40 nt. After electrophoresis, oligonucleotides were purified as described before (see Section 4.4). Sequencing reactions, primer extension products, as well as footprinting reactions were run in vertical denaturing urea (8 M) polyacrylamide (6%) gels. A Sequi-Gen GT sequencing system and 21x50 cm gels (Bio-Rad) were used. Previous to load the DNA samples, gels were prerun to heat the gel up (50ºC). Then, gels were run at 50 W to maintain the gel temperature. After electrophoresis, gels were allowed to cool down. Siliconized glass plate was carefully removed from the gel assembly and, immediately, a large piece of Whatman™ 3MM paper filter was employed to cover the exposed gel. Gel was peeled off along the paper and covered with plastic SaranTM wrap, then vacuum-dried, and exposed to Phosphor screen. Bands were visualized using a Fujifilm Image Analyzer FLA-3000.

5.3. DNA sequencing

5.3.1 Manual DNA sequencing: Dideoxy chain-termination sequencing method (Sanger method)

Manual DNA sequencing was performed following the dideoxy chain-termination method (Sanger *et al.*, 1977). The Sequenase v2.0 DNA Sequencing kit (USB Corporation) was used following the supplier indications. When linear dsDNA was used some modifications were performed. The sequencing reactions which were run alongside the primer extension products were carried out using DNA from M13mp18 (Yanisch-Perron *et al.*, 1985) and the −40 M13 primer provided by the kit. In the annealing step, 1 µg of DNA was mixed with 1 pmol of the primer and the reaction mixture was incubated at 65ºC for 5 min and then cooled down slowly to 37ºC. Then, Sequenase DNA polymerase (3.25 units) and [α-^{32}P]-dATP (3,000 Ci/mmol; Hartmann) were added to the mixture. After 2-5 min at room temperature, the mixture was distributed into four tubes. Each tube contained one of the four ddNTPs (ddATP, ddCTP, ddGTP or ddTTP) pre-warmed at 37ºC. Mixtures were then incubated at 37ºC for 5 min and finally, the reactions were stopped by adding the Stop Solution of the kit (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol). For linear dsDNA sequencing, 2.5 pmol of a 5´ radiolabelled primer were mixed with 0.25 pmol of dsDNA (obtained by PCR). The mixture was heated at 95ºC for 3 min and placed on ice. Thereafter, the Sequenase was added and the reaction proceeded as described above.

5.3.2. Automated DNA sequencing

All the plasmid constructions as well as the chromosomal modifications were confirmed by dye-terminator sequencing at Secugen (Automated DNA Sequencing Service, CIB).

5.4. *In silico* prediction of intrinsic DNA curvature

The intrinsic curvature of the C and NC DNA fragments was predicted with the bend.it server (http://hydra.icgeb.trieste.it/dna/bend_it.html). It was calculated as degrees per helical turn (10.5º/helical turn = 1º/basepair). The curvature propensity plot was calculated using the consensus scale algorithm (DNase I + nucleosome positioning data) with a windows size of 20-bp.

6. RNA techniques

6.1. Total RNA isolation from *E. faecalis*

To extract total RNA from enterococcal cells either Aurum Total RNA mini Kit (Bio-Rad) or RNeasy mini Kit (QIAGEN) were used. For microarrays and qRT-PCR studies, cells were grown to an OD_{650} of 0.4 (exponential phase) and 2 ml of culture were centrifuged at 4°C. Then, cells were resuspended in 100 µl of LB buffer (Table 4) and 160 units of mutanolysin were added. The samples were incubated at 37°C for 10 min (cell lysis). After this step, the RNeasy mini Kit (QIAGEN) was used following the manufacturer's recommendations. To remove any residual DNA in the RNA preparations, the samples were treated with DNase I recombinant RNase free (Roche). After DNase I digestion, the RNA preparations were cleaned up using the kit columns. In the case of primer extension assays and RT-PCR experiments, cells were grown under the same conditions described for genomic DNA isolation (see Section 4.2) to an OD_{650} of 0.5. Then, 4 ml of culture were concentrated (40X) in LB buffer, followed by the addition of 160 units of mutanolysin. After incubation at 37ºC for 10 min, Aurum Total RNA mini Kit (Bio-Rad) was used. Finally, an additional DNase I digestion and the subsequent RNA purification (kit columns) were performed. In all the RNA preparations, the integrity and yield of rRNAs was checked by agarose gel electrophoresis and the RNA concentration was determined using a NanoDrop ND-1000 Spectrophotometer.

6.2. Primer extension

Primer extension analyses were carried out using total RNA isolated from the enterococcal strain JH2-2/pAST-*Pma.* The ThermoScript Reverse Transcriptase kit (Invitrogen) and [α-^{32}P]-dATP (3000 Ci/mmol, Perkin Elmer) were used. First, 4.5 µg of total RNA were mixed with 20 pmol of the corresponding primer and the dNTPs (100 µM of dCTP, dGTP and dTTP; 9.75 µM of dATP and 0.25 µM of [α-^{32}P]-dATP). The reaction mixture (12 µl) was incubated at 65º C for 5 min. Then, additional components of the kit and 15 units of ThermoScript Reverse Transcriptase were added. Extension reactions were carried out at 55ºC for 45 min. Reactions were stopped by heating at 85ºC for 5 min, and the non-incorporated nucleotides were removed using Illustra MicroSpin^{TM} G-25 columns (GE Healthcare). Finally, samples were ethanol precipitated and dissolved in BXF buffer (see Table 4). cDNA products were subjected to electrophoresis in a 8 M urea/ 6% polyacrylamide sequencing gel. To estimate the

length of the extension products, sequencing reactions obtained by the Sanger method were run in the same gel. Labelled products were visualized using a FUJIFILM Image Analyzer FLA-3000.

6.3. Reverse transcription polymerase chain reaction

For first-strand cDNA synthesis, the ThermoScript Reverse Transcriptase kit was used as recommended by the supplier. Specifically, 20 pmol of the specific primer were mixed with 175 ng of total RNA and 100 µM of each dNTP. The reaction mixture (12 µl) was incubated at 65ºC for 5 min. Then, 5 mM DTT, and cDNA Synthesis buffer (provided by the supplier) and the Thermoscript Reverse Transcriptase (15 units) were added. Reaction mixtures (20 µl) were incubated at 55ºC for 45 min. After the extension step, the enzyme was inactivated (incubation at 85ºC for 5 minutes). PCRs were then carried out as described in Section 4.3.2 using the cDNA as template (10% of the first-strand reaction) and Phusion High-Fidelity DNA Polymerase (Finnzymes). To rule out the presence of genomic DNA in the RNA preparations, the same reactions were performed in the absence of ThermoScript Reverse Transcriptase (negative control). As a positive control, PCRs were performed with genomic DNA. PCR products were analysed by agarose (0.8%) gel electrophoresis.

6.4. Quantitative reverse transcription polymerase chain reaction

Quantitative reverse polymerase chain reaction (qRT- PCR) was performed in two steps. First, we carried out the reverse transcription of RNA into cDNA. Then, the cDNA was used as template in a quantitative PCR based on SYBR® Green I detection. The synthesis of cDNA was performed either with ThermoScript Reverse Transcriptase kit or iScript Select cDNA Synthesis kit (Bio-Rad). In each case, total RNA from three biological replicates was used. ThermoScript Reverse Transcriptase was used for cDNA synthesis using specific primers (validation of microarray data) as described in Section 6.3 with some modifications. Specifically, 20 pmol of the specific primer and 250 ng of total RNA were used. In addition, the final reaction mixtures contained 40 units of RNasin® Plus RNase Inhibitor (Promega). Thereafter, samples were 2-fold diluted with sterile water and stored in aliquots at -80ºC. For cDNA synthesis with random primers (complementation assays), the iScript Select cDNA Synthesis kit was used according to manufacturer's instructions. The mixture reactions (20 µl) contained 1 µg of total RNA, 4 µl of iScript Select reaction mix, 2 µl of random primers and 1 µl of iScript Reverse Transcriptase. The reactions were incubated at 25 ºC for 5 min, then at

42 ºC for 30 min, and finally at 85ºC for 5 min. The cDNA samples were 3-fold diluted with sterile water and stored at -80ºC. Negative controls (the same reactions without adding Reverse Transcriptase) were performed with to ensure the absence of DNA

The primers for quantitative PCR (qPCR) were designed using the published genome sequence of the strain OG1RF (Bourgogne *et al*., 2008) and the Primer3 software (Rozen and Skaletsky, 2000) or Primer-BLAST (Ye *et al.*, 2012) to produce amplicons of similar lengths (100-150 bp). The analysis of each primer was performed with the OligoAnalizer 3.1 tool available at www.idtdna.com and its specificity was verified using the Basic Local Alignment Search Tool BLAST NCBI tool (Johnson *et al.*, 2008).

qPCRs were carried out using iQ SYBR Green Supermix (Bio-Rad), on a iCycler Thermal Cycler (Bio-Rad). PCR reactions were performed using as template 1 µl of cDNA obtained with specific primers or 3 µl of cDNA synthesized with random primers. In addition, PCR reactions were carried out using as template the negative controls of the reverse transcription reactions. The reaction mixtures (20 µl) also contained 10 µl of iQ SYBR Green Supermix 2x and 500 nM of each primer. The protocol consisted of an initial step of denaturation at 95ºC for 5 min followed by 40 cycles that included the next steps: (i) denaturation at 95ºC for 30 s; (ii) annealing at 55ºC or 60ºC for 30 s; and (iii) extension at 72ºC for 20 s. The final step was a melt curve analysis to confirm the specificity of the selected primers. Data was analysed with the software iQ™5 Optical System Software. Relative quantification of gene expression was performed using the comparative C_t method (Livak and Schmittgen, 2001; Schmittgen and Livak, 2008). The internal control gene employed was *recA*. The threshold cycles values (C_t) of the genes of interest and the control gene were used to calculate $2^{-\Delta C_t}$, where ΔC_t = (C_t gene of interest - C_t internal control). For a particular gene, the fold change in expression (FC) between two strains was obtained dividing the corresponding $2^{\Delta C_t}$ values.

6.5. Microarrays

For each strain under study, total RNA was isolated as described in Section 6.1 from four biological replicates. Microarrays experiments were performed at Bioarray S. L. Company (Alicante, Spain). The microarrays were specifically designed for *Enterococcus faecalis* OG1RF using eArray (Agilent) (ID 0 49739). The quality of the RNA was assessed using Tape Station and the kit R6K ScreenTape (Agilent). The

Two-Color Microarray-Based Prokaryote Analysis Fair Play III Labeling protocol (v. 1.3) (Agilent) was followed to label the samples. Bioinformatics analysis of microarray data were done using Bioconductor (www.bioconductor.org), with the following packages: Limma, Marray, affy, pcaMethods and EMA, in the R computing environment. Significant fold changes (FC) were considered when the $\log_2$FC of the wild-type strain *versus* the mutant strain was higher than 3.

7. Protein purification

The pET24b expression vector (Novagen) and the *E. coli* BL21 (DE3) strain were used for overproduction and purification of proteins. In pET24b, the gene of interest is expressed under control of the φ10 promoter of phage T7. *E. coli* BL21 (DE3) strain was used as host since it contains the T7 RNA polymerase (RNAP) encoding gene (T7 gene) under the control of the *lac*UV5 promoter and a chromosomal copy of the *lacI* gene. In the absence of IPTG, the LacI repressor binds to the operator region of the *lac*UV5 promoter and represses the transcription of the T7 gene. In the presence of IPTG, LacI is blocked and, therefore, there is synthesis of the T7 RNAP, which transcribes the gene of interest from the φ10 promoter.

7.1. Purification of MA*Efa*

E. coli BL21 (DE3) cells carrying the pET24b-*maEfa* plasmid were grown at 37ºC with rotary shaking in TY broth containing Km (30 μg/ml). When the culture reached an OD_{600} of 0.45, 1 mM IPTG was added to induce the expression of the *maEfa* gene. After 25 min under the same growth conditions, the culture was treated with rifampicin (200 μg/ml) for 60 min, which specifically inhibits bacterial RNAP. Cells were collected by centrifugation (8,000 rpm in an SLA-3000 rotor for 15 min at 4ºC), washed twice with buffer VL containing 200 mM NaCl and stored at -80ºC. The cell pellet was concentrated (40X) in buffer VL containing 200 mM NaCl and a protease inhibitor cocktail (Roche). Cells were lysed by two passages through a pre-chilled French pressure cell, and the whole-cell extract was centrifuged (10,000 rpm in an Eppendorf F-34-6-38 rotor for 40 min at 4ºC) to remove cell debris. The clarified extract was mixed with 0.2% polyethyleneimine (PEI), kept on ice for 30 min, and centrifuged at 9,000 rpm in an Eppendorf F-34-6-38 rotor for 20 min at 4ºC. Under these conditions, nucleic acids and some proteins, including MA*Efa*, precipitated. The pellet was washed twice with buffer VL containing 200 mM NaCl to remove any trapped protein. Then, the pellet was washed with buffer VL containing 400 mM NaCl to elute

the MA*Efa* protein. Subsequently, proteins present in the supernatant were precipitated with 70% (w/v) saturated ammonium sulphate, which was added slowly. The mixture was kept on ice with constant stirring for 60 min, followed by a centrifugation step (9,000 rpm in an Eppendorf F-34-6-38 rotor for 20 min at 4ºC). The precipitate was dissolved in buffer VL containing 200 mM NaCl and then dialyzed at 4ºC against the same buffer to remove the ammonium sulphate. The sample was loaded onto a heparin affinity column (HiPrep Heparin, GE Healthcare) equilibrated with buffer VL containing 200 mM NaCl. After washing the bound proteins with 5-column volumes of the same buffer, a linear gradient (200-600 mM NaCl) was applied to elute MA*Efa*. Protein fractions were analysed using Coomassie-stained SDS-polyacrylamide (12%) gels. Fractions containing MA*Efa* were pooled and dialyzed at 4ºC against VL buffer containing 200 mM NaCl. The protein preparation was concentrated by filtering through a 10-kDa-cutoff membrane (Macrosep; Pall). Then, the protein preparation was applied to a HiLoad Superdex 200 gel-filtration column (120 ml; 16 x 600 mm) (Amersham Biosciences) and subjected to fast-pressure liquid chromatography (FPLC) (Biologic DuoFlow; Bio-Rad) at 4ºC. The column was equilibrated with buffer VL containing 200 mM NaCl at a constant flow rate of 1 ml/min. Buffer VL containing 200 mM NaCl was used also as running buffer. Protein fractions were analysed as described above, and fractions containing pure MA*Efa* were pooled, concentrated and stored at -80ºC.

7.2. Purification of MA*Efa*-His

E. coli BL21 (DE3) cells harbouring the pET24b-*maEfa-His* plasmid were used to overproduce the MA*Efa-His* protein. Induction of the expression of *maEfa-His* was performed as described above (see Section 7.1). After the incubation with rifampicin, cells were sedimented by centrifugation, and washed twice with buffer S-His. The cell pellet was concentrated (30X) in buffer S-His supplemented with an EDTA-free protease inhibitor cocktail (Roche). Cells were disrupted by two passages through a prechilled cell-pressure French Press, and the cell lysate was clarified by centrifugation (8,000 rpm in an Eppendorf RF-34-6-38 rotor for 30 min at 4ºC). Then, imidazole was added to the clarified extract at a final concentration of 10 mM. This extract was loaded onto a nickel affinity column (HisTrap HP column, GE Healthcare) pre-equilibrated with buffer S-His containing 10 mM imidazole. After washing with the same buffer, MA*Efa*-His was eluted using a linear gradient of buffer S-His containing from 10mM to 250 mM imidazole. Fractions were analysed by electrophoresis on SDS polyacrylamide gels (10%) gels followed by staining with Coomassie Brilliant. Blue. Fractions containing MA*Efa*-His were pooled, dialyzed against buffer VL containing 200 mM NaCl and

concentrated by filtration through a 10-kDa-cutoff membrane (Macrosep; Pall). Then, the protein sample was loaded onto a Bio-Silect SEC 125-5 column (Bio-Rad) pre-equilibrated with buffer VL containing 200 mM NaCl, and subjected to fast-pressure liquid chromatography (FPLC; Biologic DuoFlow; Bio-Rad). Fractions were analysed as indicated above, and those containing MA*Efa*-His were pooled, concentrated, and stored at -80ºC.

8. Protein analysis

8.1. Determination of protein concentration

The theoretical molecular weight (Da) and the molar extinction coefficient (M^{-1} cm^{-1}) were calculated from the amino acid sequence of a protein. These values were used for measuring the concentration of a particular protein in solution using a Nanodrop (ND-1000) Spectrophotometer (Bio-Rad).

8.2. N-terminal sequencing

The purified MA*Efa* protein (in buffer VL containing 200 mM NaCl) was subjected to amino terminal sequencing by Edman degradation using a Procise 494 Sequencer (Perkin Elmer) (Protein Chemistry Facility; CIB).

8.3. Protein electrophoresis

8.3.1. Tris-Glycine SDS-PAGE

Protein samples were analysed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). The stacking gel contained 4% polyacrylamide in 0.125 M Tris-HCl, pH 6.8, and 0.1% SDS. The resolving gel contained 10-12% polyacrylamide in 0.374 M Tris-HCl, pH 8.8, and 0.1% SDS. TEMED and ammonium persulfate (APS) were used to catalyze polyacrylamide gel polymerization. Protein samples were mixed with SLB (see Table 4) and heated at 95ºC for 5 min prior to electrophoresis. Gels were run in a Mini-Protean III Electrophoresis System (Bio-Rad) using TG buffer (see Table 4). Electrophoresis was carried out at 80 V until the dye (bromophenol blue) migrated down to the bottom of the stacking gel. Then, the voltage was increased to 180 V. Gels were stained with Coomassie Brilliant Blue R-250 (Bio-Rad).

8.3.2. Tris-Tricine SDS-PAGE

Proteins smaller than 30 kDa were resolved by Tris-Tricine SDS-PAGE (Schägger and von Jagow, 1987). The stacking gel contained 4% polyacrylamide and 1 ml of Buffer gel (see Table 4). The resolving gel (5 ml) contained 16% polyacrylamide, 1.65 ml of Buffer gel and glycerol (13%). APS and TEMED were used to catalyze the polyacrylamide polymerization. Two different electrophoresis buffers were used, the cathode buffer and the anode buffer (see Table 4). Electrophoresis conditions were the same as described in the previous Section.

8.4. Gel filtration chromatography

Gel filtration chromatography was used to determine the size (Stokes radius) of the MA*Efa* protein. The protein preparation was subjected to gel filtration chromatography as described in Section 7.2, but the column was previously calibrated with a set of proteins of known Stokes radius. The standard proteins used were alcohol dehydrogenase (ADH; 45 Á), albumin (A; 35.5); ovalbumin (O; 30.5 Á) and carbonic anhydrase (CA; 20.1 Á). They were dissolved in buffer VL containing 200 mM NaCl at a final concentration of 4 mg/ml. The chromatographic runs were performed under the conditions described in Section 7.2. The elution volume of each protein was determined monitoring the absorbance at 280 nm. The K_{av}, an elution volume parameter, was calculated for each one as follows:

$$K_{av} = (V_e-V_0)/(V_t-V_0)$$

where V_e is the elution volume, V_0 is the void volume (equal to blue dextran elution volume) and V_t is the total volume of the packed bed.

8.5. Analytical ultracentrifugation

Analytical ultracentrifugation experiments were performed at the Analytical Ultracentrifugation and Macromolecular Interactions Facility of CIB. These experiments provided information about the molecular mass of MA*Efa* as well as the hydrodynamic behaviour of the protein. An Optima XL-I analytical ultracentrifuge (Beckman-Coulter) equipped with an UV-visible optical detection system was used to perform the experiments, using an An50Ti rotor and Epon-charcoal standard double sector centerpieces (12 mm optical path).

8.5.1. Sedimentation velocity

MA*Efa* was equilibrated in buffer AU (see Table 4) and two protein concentrations (5 and 10 µM; 350 µl) were analysed. Sedimentation velocity experiments were performed at 48,000 rpm and 12ºC. The sedimentation coefficient for MA*Efa* was estimated applying a direct linear least-squares boundary modelling of the sedimentation velocity data using the SEDFIT program (version 12.0) (Schuck and Rossmanith, 2000). The sedimentation coefficient was corrected to standard conditions to obtain the corresponding S_{20},w value using the SEDNTERP program (Laue *et al.*, 1992) The translational frictional coefficient (*f*) of MA*Efa* was determined from the molecular mass and sedimentation coefficient of the protein (van Holde, 1985), whereas the frictional coefficient of the equivalent hydrated sphere (f_0) was estimated using a hydration of 0.37 g H_2O/g protein (Pessen and Kumosinski, 1985). With these parameters the translational frictional ratio (f/f_0) was calculated, which allows an estimation of the hydrodynamic shape of MA*Efa.*

8.5.2. Sedimentation equilibrium

Sedimentation equilibrium experiments were performed at 12ºC. Two protein concentrations (5 and 10 µM) were analysed. Samples (90 µl) were centrifuged at two successive speeds (8,000 and 10,000 rpm) and absorbance readings were done after the sedimentation equilibrium was reached. The absorbance scans were taken at 280 and 291 nm, depending on MA*Efa* concentration. In all cases, the baseline signals were measured after high-speed centrifugation (40,000 rpm). Apparent average molecular masses of MA*Efa* protein were determined using the HETEROANALYSIS program (Cole and Michael, 2004). The partial specific volume of MA*Efa* was 0.742 ml/g, calculated from the amino acid composition with the SEDNTERP program (Laue *et al*., 1992).

8.6. Western blots

E. faecalis cells were grown to an OD_{650} of 0.3 (exponential phase) and 2 ml of culture were centrifuged to prepare whole-cell extracts. Bacteria were resuspended in 50 µl of LBW buffer and incubated at 37ºC for 10 min. Then, 13 µl of SLB buffer (5x) were added, and the samples were heated at 95ºC for 5 min. Subsequently, 6 µl of 10% SDS were added to the cell extracts and were heated as before. Equivalent amounts of the cell extracts (12 µl samples) were loaded onto a SDS-polyacrylamide (10%) gel to separate total proteins. Pre-stained proteins (SeeBlue Plus Invitrogen)

were run in the same gel as molecular weight markers. After electrophoresis, the gel was equilibrated with TB buffer (see Table 4). Immun-Blot PVDF membranes (Bio-Rad) were used for protein blotting. Prior to the transfer, the membrane was soaked in methanol followed by equilibration in TB buffer. Proteins were transferred electrophoretically to Immun-blot PVDF membranes in TB buffer using a Mini Trans Blot (Bio-Rad) at 100 mA and 4ºC for 120 min. Thereafter, the membrane was dried by immersion in methanol for 2 min. Membranes were probed with polyclonal antibodies against MA*Efa* raised in rabbits, which were diluted 1:1,000 in SB buffer (see Table 4). After 1 h of incubation, the membrane was rinsed five times with WB buffer. Subsequently, it was incubated with the secondary antibody for 1 h. The secondary antibody, anti rabbit IgG conjugated with horseradish peroxidase (HRP), was diluted 1:30,000 in SB buffer. The unbound secondary antibody was removed washing the membrane with PBS buffer (see Table 4) five times. Antigen-antibody complexes were detected using the Immun-StarTM HRP substrate kit (Bio-Rad). The peroxidase linked to the secondary antibody catalyzes the oxidation of the substrate (luminol) leading to light emission. The light signal was detected with a Luminescent Image Analyzer LAS-3000 (Fujifilm Life Science) or by autoradiography.

The pneumococcal whole-cell extracts were prepared from cultures grown to late logarithmic phase (OD_{650} of 0.6). Media containing 0.3% sucrose and different concentrations of fucose (0.1% to 1%) as carbon source were used. Bacteria were concentrated 40-fold in buffer LBP and incubated at 30ºC for 10 min. Then, samples (8 µl) were mixed with 2 µl of SLB buffer (5x) and heated at 95ºC for 5 min before being loaded onto SDS-polyacrylamide (14%) gels. Thus, equivalent amounts of the cell extracts (similar number of cells) were loaded onto the gel. The SeeBlue Plus 2 pre-stained protein standard (Invitrogen) was run in the same gel as a molecular weight marker. Protein transfer to the membrane was performed as described above (enterococcal whole cell-extracts). Then, primary and secondary antibody incubation were carried out as specified by the supplier. The primary antibody was Anti-GFP (Roche Applied Science), a mixture of two mouse monoclonal antibodies against the green fluorescent protein; and the secondary antibody was Peroxidase-conjugated AffiniPure Goat Anti-Mouse IgG (H+L) (Jackson ImmunoResearch).

9. DNA-protein interactions

9.1. Electrophoretic mobility shift assays

In general, standard binding reactions (10-20 µl) were performed mixing 10 nM of unlabelled DNA or 0.1-5 nM of 5´-labelled DNA with different amounts of purified MA*Efa* or MA*Efa*-His. The binding buffer contained 30 mM Tris-HCl, pH 7.6; 1 mM DTT; 0.2 mM EDTA; 1 % glicerol; 50 mM NaCl, and 0.5 mg/ml BSA. When indicated, non-labelled competitor calf thymus DNA was added simultaneously with the ^{32}P-labelled DNA to the binding reaction. The reaction mixture was incubated at room temperature for 20 min. Then, BXGE buffer was added and free and bound DNA forms were separated by electrophoresis on native polyacrylamide gels (5-8% PAA, depending on the DNA length) using TBE buffer. Gels and running buffer were maintained at 4ºC until the pre-electrophoresis was started. Then, gels were pre-run (20 min) and run at 100 V (room temperature or 4ºC). Labelled DNA was visualized by a Fujifilm Image Analyzer FLA-3000 and quantified with the Quantity One software (Bio-Rad).

9.2. DNase I footprinting assays

Binding reactions (50 µl) contained 2 nM of the 5'-labelled DNA fragment; 30 mM Tris-HCl, pH 7.6; 1 mM DTT; 0.2 mM EDTA; 1 % glicerol; 50 mM NaCl, 0.5 mg/ml BSA, 1 mM $CaCl_2$, 10 mM $MgCl_2$ and different concentrations of the MA*Efa* protein. The reactions mixtures were incubated at room temperature for 20 min. DNase I (stock solution) was diluted in 50% glycerol to 1:500. DNA was digested by adding 0.04 units of DNase I. After incubation for 5 min at room temperature, reactions were stopped by adding 25 µl of STOP DNase I buffer. DNA was precipitated with 187 µl of absolute ethanol, dried and dissolved in 5 µl of BXF buffer. Then samples were heated at 95ºC for 5 min and loaded onto sequencing gels (6% PAA, 8 M urea).

10. Fluorescence measurements

Enterococcal and pneumococcal cells harbouring plasmid (*gfp* reporter gene) were grown to an OD_{650} of 0.3, except pneumococcal cells carrying pAST-*PfcsK* or pAST-*oPfcsK*, which were grown to an OD_{650} of 0.6. Then, different volumes of the culture (from 25 µl to 1 ml) were centrifuged at 4ºC, and cells were resuspended in 200 µl of PBS buffer. Fluorescence was measured either on a LS-50B espectrophotometer

(Perkin-Elmer) or on a Thermo Scientific Varioskan Flash instrument (Perkin-Elmer) by excitation at 488 nm and detection of emission at 515 nm. The LS-50B espectrophotometer (Perkin-Elmer) was used for the experiments shown in Chapter 1. The Thermo Scientific Varioskan Flash instrument (Perkin-Elmer) was used for the experiments shown in Chapters 2 and 4. In all cases, three independent cultures were analysed. The fluorescence corresponding to 200 µl of PBS buffer without cells was also measured in each case.

11. Infection models

11.1. Bacterial adhesion to HEp-2 cells

The human epithelial cell line HEp-2 (ATCC CCL23) was cultured in DMEM medium (Gibco-ThermoFisher) supplemented with 2 mM glutamine, heat-inactivated 5% fetal bovine serum (Gibco-ThermoFisher), streptomycin 100 U/ml and penicillin 100 µg/ml. Cells were incubated at 37°C in an atmosphere containing 7% CO_2. Cells were passaged every 5 days. Prior to the infections, cells were seeded onto glass cover slips placed on the bottom of 4-well culture plates at a density of 2-5 x 10^5 cell/ well and cultivated for 24 h at 37ºC and 7% CO_2. To infect the cells, bacteria (strains OG1RF/pMV158GFP and OG1RFΔ*maEfa*/pMV158GFP) were grown to mid-exponential phase, centrifuged and resuspended into the cell medium to achieve a multiplicity of infection of 5:1. Then, bacteria suspensions were added to the epithelial cells and were incubated for 2 h at 37ºC and 7% CO_2. After that, the cells were washed twice with PBS to remove unbound bacteria and further incubated with fresh medium supplemented with gentamicin 100 µg/ml, streptomycin 100 U/ml and penicillin 100 µg/ml. After 2, 3 and 24 h of incubation, cells were fixed in PBS with 2.5% paraformaldehyde for 15 min and kept at 4ºC for 24 h. Then, cover slips were washed with PBS, mounted on glass slides with Moviol containing 4′,6′-diamidino-2-phenylindole (DAPI) and analysed by fluorescence microscopy. At each time point, five microscopic fields were randomly chosen and photographs were taken. The number of cells with bacteria attached was counted. Three independent experiments were performed.

11.2. Mice infection

The animal experiments were approved by the appropriate national ethics board (Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit,

Oldenburg, Germany). Mice were infected intraperitoneal with 5 x 10^7 CFU of live *E. faecalis* (strains OG1RF and OG1RFΔ*maEfa*) and euthanized by CO_2 inhalation at 24 h after bacterial inoculation. The infiltrating inflammatory cells were isolated from the side of infection by extensively rinsing with 2 ml of warm DMEM medium. The resulting cell suspension was counted using a Neubauer chamber and levels of inflammatory neutrophils were determined by flow cytometry analysis using anti-mouse Ly6G/C antibody (BD Pharmingen, San Diego, USA). The cell suspension was also used to determine the amount of viable bacteria by plating serial dilutions on bile esculin azide agar (Fluka) plates and the levels of inflammatory cytokines.

11.2.1 Cytokine determination

The determination of IL-6 levels in the peritoneal lavage was performed by specific enzyme-linked immunosorbent assay (ELISA), using matched antibody pairs and recombinant cytokines as standards. Briefly, 96-well microtiter plates were coated with the corresponding purified anti-murine capture monoclonal anti-IL-6 (Pharmingen, San Jose, CA) at a concentration of 2 µg/ml in sodium bicarbonate buffer overnight at 4°C. The wells were washed and then blocked with 1% bovine serum albumin-PBS before the serum samples and the appropriate standard were added to each well. Biotinylated rat monoclonal anti-IL-6 (BD Pharmingen) at 2 µg/ml was added as the second antibody. Detection was performed with streptavidin-peroxidase, and the plates were developed by use of TMB (3,3',5,5'-tetramethylbiphenyl-4,4'-diamine).

11.2.2. Statistical analysis

Data were analysed using Prism 5 (GraphPad). Statistical significance was determined by using the unpaired Student's t-test for the comparison of two groups. In all analyses, $P<0.05$ was considered statistically significant.

RESULTS

CHAPTER 1

Plasmid-based genetic tools for *Enterococcus faecalis*

Despite the increasing interest in a deeper understanding of the molecular biology of *E. faecalis*, the availability of plasmid-based genetic tools is still limited. We have developed several plasmid-based genetic tools which have been very useful to address essential questions in the study of the enterococcal *maEfa* gene. In this chapter, we describe the construction of these genetic tools and their validation in *E. faecalis.* Moreover, we show that they are also suitable for *S. pneumoniae*. The results concerning the utility of the pAS terminator-probe vector and the pAST promoter-probe vector in *E. faecalis* and *S. pneumoniae* were published in 2010 (Ruiz-Cruz *et al*., 2010).

1.1. The pAS terminator-probe vector

We have constructed a terminator-probe vector, pAS, based on the pMV158 replicon. The streptococcal plasmid pMV158 (5540 bp), which is the prototype of a family of rolling-circle replicating plasmids (del Solar *et al.*, 1998), is able to replicate and confer resistance to tetracycline (gene *tetL*) in a broad variety of bacterial hosts (del Solar et al., 1998; Lacks et al., 1986), including *E. faecalis* (Nieto and Espinosa, 2003). Sequence analysis of the region located just downstream of the *tetL* gene revealed the existence of a putative Rho-independent transcriptional terminator (Fig. 4). It consists of an inverted-repeat followed by a short stretch of thymine residues (Lacks *et al*., 1986). This sequence element is also present in plasmid pLS1, a pMV158-derivative that lacks the 1132-bp *Eco*RI restriction fragment (Lacks *et al*., 1986). The efficiency of the *tetL* inverted-repeat as transcriptional terminator was first investigated in our group by primer extension assays (Ruiz-Cruz *et al*., 2010). Such experiments were performed using total RNA isolated from pLS1-carrying *S. pneumoniae* cells. The results obtained suggested that the pneumococcal RNA polymerase (RNAP) does not recognize efficiently the *tetL* inverted-repeat as a stop signal (not shown).

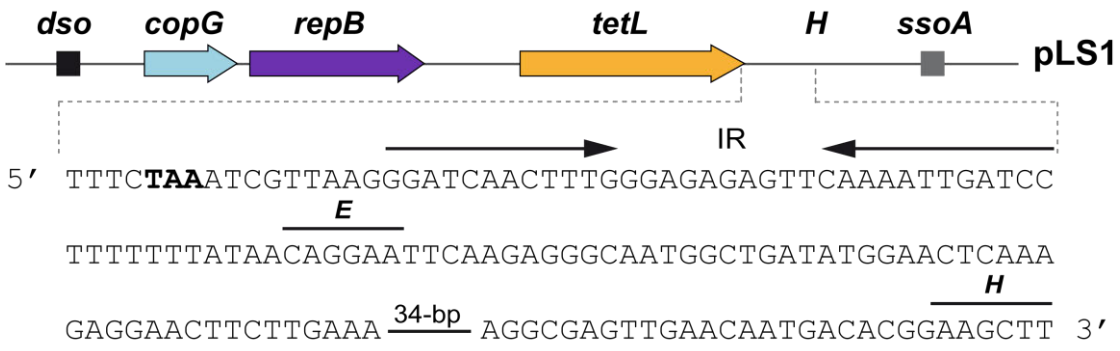


Figure 4. The transcriptional terminator of the *tetL* gene. Relevant features of pLS1 are indicated. *copG* and *repB* genes are involved in plasmid DNA replication. The location of the replication origins *dso* (double-strand origin) and *ssoA* (single-strand origin) is indicated. The *tetL* gene confers resistance to tetracycline. The nucleotide sequence of the region spanning the translation stop codon (TAA) of the *tetL* gene and the *Hind*III site (H) is shown. IR: inverted-repeat, *E*: *Eco*RI site.

To confirm this result, we cloned a *gfp* reporter cassette into the *Hind*III site of plasmid pLS1 (Fig. 5A). The *Hind*III site is located downstream of the *tetL* inverted-repeat. The *gfp* reporter cassette contains a multiple cloning site (MCS) upstream of a promoter-less *gfp* allele, which encodes a green fluorescent protein (GFP) that carries the F64L and S65T mutations (Cormack, 1996; Heim, 1995). The F64L mutation increases GFP solubility, while the S65T mutation increases GFP fluorescence and causes a red shift in the excitation spectrum. In addition, the *gfp* allele carries translation initiation signals (SD in Fig. 5) which are optimal for its expression in prokaryotes (Miller and Lindow, 1997). The *gfp* cassette was inserted into pLS1 in both orientations (plasmids pAS and pSA). In plasmid pAS, the *tetL* and *gfp* genes are located on the same DNA strand (Fig. 5A). Both plasmids were introduced into *S. pneumoniae* 708 and *E. faecalis* JH2-2. We analysed *gfp* expression in cells carrying pAS or pAS by measuring the intensity of fluorescence at 515 nm (excitation at 488 nm) (Fig. 5B). Expression of *gfp* was detected in pneumococcal and enterococcal cells carrying pAS. The fluorescence increased as a function of the culture volume in both cases. No *gfp* gene expression was observed in cells harbouring pSA, which confirmed the absence of promoter signals within the *gfp* reporter cassette. The fluorescence corresponding to 0.8 ml culture (OD_{650} = 0.3; pAS-containing cells) was 3-fold and 4.5-fold higher than the background level in *S. pneumoniae* and *E. faecalis*, respectively. Therefore, RNAP read through the *tetL* inverted repeat not only in *S. pneumoniae* but also in *E. faecalis*. As a consquence, there is transcription of the *gfp* reporter cassette. This fact and the presence of a MCS between the *Hind*III site and the promoter-less *gfp* gene suggested that plasmid pAS could be used as a terminator-

probe vector. This type of vectors makes possible to rapidly test whether a particular sequence functions as a transcriptional terminator.

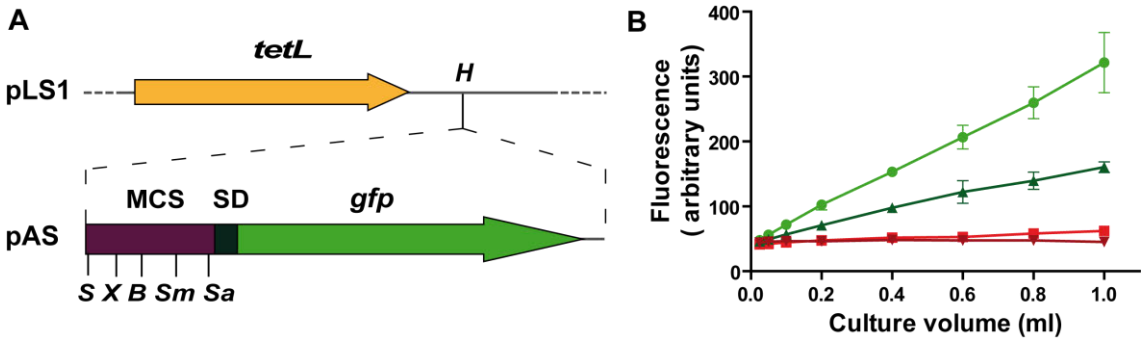


Figure 5. **The pAS terminator-probe vector**. **(A)** The *gfp* reporter cassette was inserted into the *Hind*III site (*H*) of the pLS1 plasmid. This cassette contains a multiple cloning site (MCS), translation initiation signals optimized for prokaryotes (SD) and a promoter-less *gfp* allele (Miller and Lindow, 1997). Plasmid pSA (control plasmid, not shown) carries the *gfp* reporter cassette inserted in the opposite orientation. *S: Sal*I, *H*: *Hind*III, *X*: *Xba*I, B: *Bam*HI, *Sm*: *Sma*I, *Sa*: *Sac*I. **(B)** *gfp* gene expression in plasmid-harbouring cells. *S. pneumoniae* carrying plasmid pAS (dark green) or pSA (dark red). *E. faecalis* carrying plasmid pAS (light green) or pSA (light red). The graph is the mean of three experiments.

To test this hypothesis we selected some predicted or experimentally determined Rho-independent terminators from different bacterial genomes. Specifically, we inserted independently the following DNA sequences (Fig. 6) into the *Sal*I site of the pAS plasmid (see Fig. 5A):

(i) A 246-bp region containing the tandem terminators *T1* and *T2* of the *E. coli rrnB* ribosomal RNA operon (Brosius, 1981). Such a fragment was inserted in both orientations (herein named *T1T2rrnB* and *T2T1rrnB* fragments, respectively). These terminators have been used frequently in the construction of plasmids (Brosius, 1984; Serrano-Heras *et al.*, 2005; Simons, 1987).

(ii) A 238-bp DNA fragment containing the transcriptional terminator of the *S. pneumoniae polA* gene (fragment *TpolA*). By mapping with S1 nuclease, it was shown that transcription of the *polA g*ene terminates at the palindrome shown in Fig. 6 (López *et al*., 1989).

(iii) A 265-bp *Sal*I region containing the putative Rho-independent terminator of the *E. faecalis sigV–rsiV* operon (fragment *TrsiV*). The *sigV* and *rsiV* genes encode members of the extracytoplasmic function subfamily of eubacterial RNAP sigma and anti-sigma factors, respectively (Benachour *et al*., 2005).

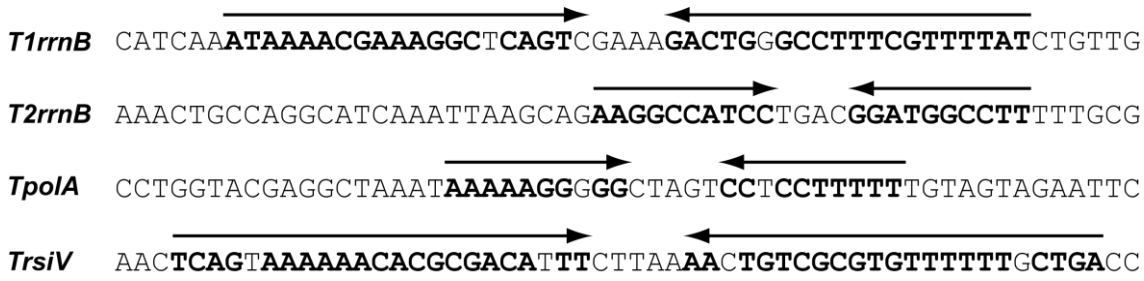


Figure 6. Palindromic sequences at the terminator regions analysed in this work. Arrows indicate nucleotide sequences corresponding to potential RNA hairpin structures. Complementary bases of the hairpin structures are shown in bold.

The recombinant plasmids pAST (fragment *T1T2rrnB*), pAS-*T2T1rrnB*, pAS-*TpolA* and pAS-*TrsiV* were introduced into *E. faecalis* JH2-2 and *S. pneumoniae* 708. The efficiency of the inserted fragments as transcriptional terminators was evaluated by monitoring *gfp* gene expression (Fig. 7). The fluorescence in enterococcal and pneumococcal cells carrying the pAS plasmid was 259.54 ± 24.49 and 139.54 ± 13.22, respectively. Compared to pAS-carrying cells, the *T1T2rrnB* and *TrsiV* fragments reduced the intensity of fluorescence to background values in both bacteria. In the case of the *T2T1rrnB* fragment, the fluorescence decreased 1.5-fold and 1.8-fold in enterococcus and pneumococcus, respectively. However, the *TpolA* fragment reduced the fluorescence in *E. faecalis* (3-fold) but not in *S. pneumoniae*. A further analysis of the *TpolA* fragment using the BPROM prediction program (Softberry, Inc.) revealed a near-consensus −10 hexamer (TAgAAT) located 5 nucleotides downstream of the *TpolA* palindrome, as well as a near-consensus extended −10 element (TGTa) (see Fig. 6). Thus, activity of this predicted promoter in *S. pneumoniae* but not in *E. faecalis* might explain why the terminator activity of the *TpolA* palindrome was only detected in *E. faecalis*. These results demonstrate that plasmid pAS can be used to assess whether particular sequences (homologous or heterologous) function as transcriptional terminators in *E. faecalis* and *S. pneumoniae*. Moreover, we have shown that the predicted *TrsiV* terminator of *E. faecalis* is active in both bacteria. In our system, it is as efficient as the tandem terminators *T1* and *T2* of *E*. *coli.*

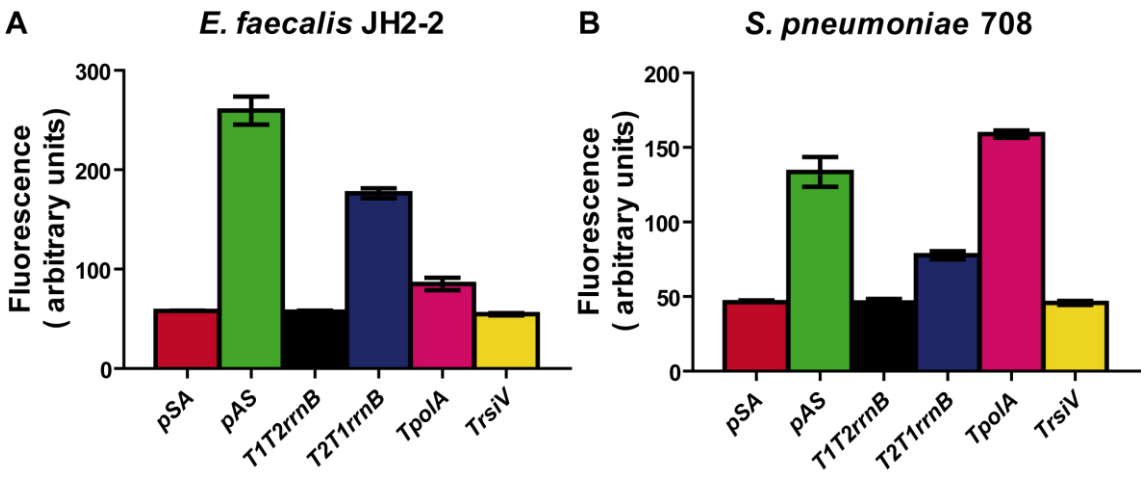


Figure 7. Use of plasmid pAS as a terminator-probe vector. The intensity of fluorescence in enterococcal **(A**) and pneumococcal cells **(B)** harbouring the indicated recombinant plasmids was measured. Cells were exponentially grown to an OD_{650}=0.3 and the fluorescence (arbitrary units) corresponds to 0.8 ml of culture. Each bar represents the mean ± SD of three independent experiments. The intensity of fluorescence in cells carrying the control plasmid pSA was 58.08 ± 0.64 and 46.32 ± 2.24, respectively.

1.2. The pAST promoter-probe vector

The bacterial RNAP holoenzyme is a complex of six subunits ($\alpha_2\beta\beta'\omega\sigma$). During initiation of transcription, most of the sequence-specific contacts of the RNAP with the promoter region are made by the σ subunit. In general, bacterial genomes encode diverse forms of the σ factor, which confers promoter specificity to the RNAP (Gruber and Gross, 2003; Wigneshweraraj *et al*., 2008). Most transcription in exponentially growing bacterial cells is initiated by RNAP carrying a housekeeping σ factor similar to *E. coli* σ^{70}. This holoenzyme recognizes promoters characterized by two main sequence elements, the –35 and –10 hexamers, whose consensus sequence is 5´-TTGACA-3´ and 5´-TATAAT-3´, respectively. The optimum spacer length between these elements is 17 nucleotides. Additionally, some of these promoters contain the extended –10 element, which is located one nucleotide upstream of the –10 hexamer. This element is more conserved in Gram-positive bacteria (5´-TRTG-3´ motif) than in *E. coli* (5´-TG-3´ motif) (Mitchell *et al*., 2003; Sabelnikov *et al*., 1995; Voskuil and Chambliss, 1998). Since the sequence elements at numerous promoters have evolved to diverge from the consensus, definitive identification of a promoter target for RNAP requires the use of diverse experimental strategies, such as the use of promoter-probe plasmid vectors (Minchin and Busby, 2009; Ross and Gourse, 2009).

Cloning of the *E. coli T1T2rrnB* terminator region into the *Sal*I site of the pAS terminator-probe vector generated plasmid pAST (5456 bp; see above). This derivative

conserves unique restriction sites for *Xba*I, *Bam*HI, *Sma*I and *Sac*I between the *T1T2rrnB* region (transcriptional terminator) and the promoter-less *gfp* gene (see Fig. 8). To examine whether plasmid pAST is suitable as a promoter-probe vector, we selected several DNA fragments containing a predicted or experimentally tested promoter from *E. faecalis* or *S. pneumoniae*. These regions were independently inserted into the *Bam*HI or *Sac*I site of pAST (Figs. 8 and 9). The recombinant plasmids were then introduced into *E. faecalis* JH2-2 and *S. pneumoniae* 708, and promoter activity was evaluated by monitoring *gfp* expression (Fig. 10).

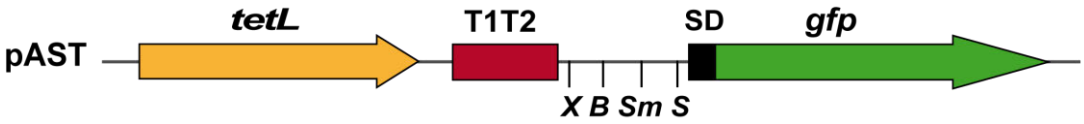


Figure 8. Relevant features of the pAST promoter-probe vector. The *tetL* gene confers resistance to tetracycline. T1T2: tandem terminators *T1* and *T2* of the *E. coli rrnB* rRNA operon. The translation initiation signals (SD) optimized for the expression of the *gfp* gene in prokaryotes are followed by a promoter-less *gfp* allele. *X*: *Xba*I, B: *Bam*HI, *Sm*: *Sma*I, *S*: *Sac*I.

Regarding the enterococcal promoters, we analysed the promoter region of three genes from *E. faecalis* V583: *uppS* (or *cpsA*; undecaprenyl diphosphate synthase) (Hancock and Gilmore, 2002; Thurlow *et al*., 2009b), *EF2493* (or *cpsC*; putative teichoic acid biosynthesis protein) (Hancock and Gilmore, 2002) and *EF2962* (putative LacI family transcriptional regulator) (Fig. 9). The *PuppS* and *P2493* promoters were identified by primer extension (Hancock *et al*., 2003). Promoter *PuppS* promoter has a consensus –10 hexamer and shows a 4/6 match at the –35 element, whereas *P2493* promoter has near-consensus –10 and –35 hexamers. In the case of the *EF2962* gene, the BPROM program (*Softberry, Inc.*) predicted a –10 hexamer (four consensus bases) located 56 nucleotides upstream of the initiation codon. This promoter has a near-consensus –10 extension and shows a 3/6 match at the –35 element. Compared to pAST-carrying cells, the *P2493* promoter was the strongest enterococcal promoter in both *E. faecalis* and *S. pneumoniae* (Fig. 10). The activity of such a promoter was higher than the *PuppS* and *P2962* promoters. It was 5.2 and 3.2-fold higher in enterococcus, and 1.6 and 2.3-fold higher in pneumococcus, respectively.

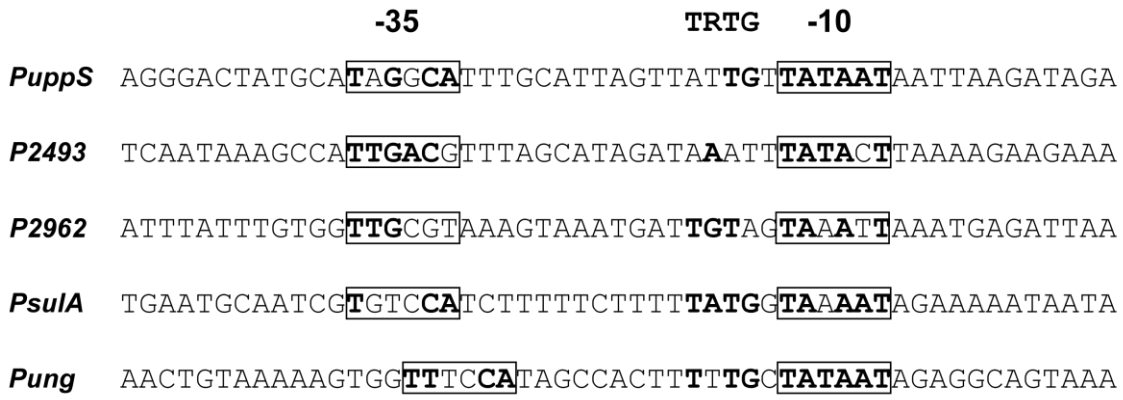


Figure 9. Main sequence elements at the promoter regions analysed in this work. The −35 and −10 hexamers are indicated. The position of the extended −10 element (5′-TRTG-3′ motif) is shown. Conserved nucleotides are indicated in bold.

We further analysed the promoter region of the pneumococcal *sulA* (dihydropteroate synthase) and *ung* (uracil-DNA glycosylase) genes. The *PsulA* promoter, which was identified by primer extension (Lacks *et al*., 1995; López *et al*., 1987), has a near-consensus –10 hexamer and a consensus –10 extension. In the case of the *ung* gene (Méjean *et al*., 1990), the BPROM prediction program (*Softberry, Inc.*) revealed a consensus –10 hexamer located 28 nucleotides upstream of the translation initiation codon and a near-consensus –10 extension (see Fig. 9). In enterococcus, the intensity of fluorescence increased only 2-fold when the *PsulA* promoter region was cloned into pAST (plasmid pAST-*PsulA*) (Fig 10). The activity of the *PsulA* promoter was slightly higher than that of the *Pung* promoter. Different results were obtained in pneumococcus. In this case, the fluorescence increased 10-fold in cells harbouring the pAST-*Pung* recombinant plasmid. The activity of the *Pung* promoter was 1.9-fold higher than that of the *PsulA* promoter. In conjunction, our results demonstrate that plasmid pAST is useful to assess whether particular sequences (both homologous and heterologous) function as promoter signals in *E. faecalis* and *S. pneumoniae*. Among the analysed promoters, we have shown that two predicted promoters, *Pung* and *P2962,* are active in both bacteria. Furthermore, we have demonstrated that, under our experimental conditions, the strongest promoters (10-fold increase in fluorescence) are the *P2493* promoter in enterococcus and the *Pung* promoter in pneumococcus.

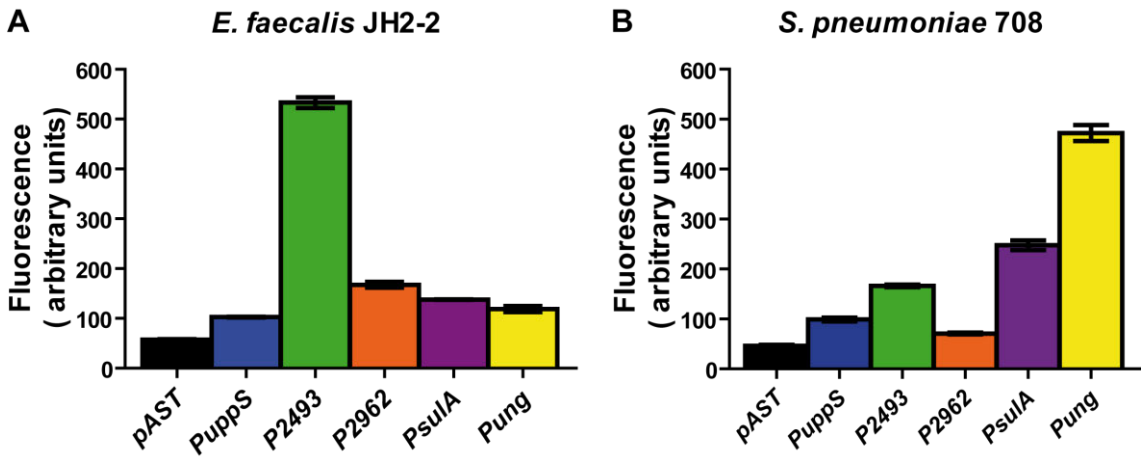
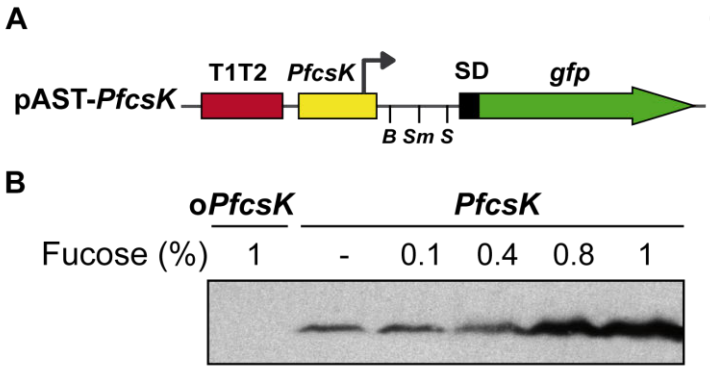


Figure 10. **Use of plasmid pAST as a promoter-probe vector.** The intensity of fluorescence in enterococcal **(A)** and pneumococcal cells **(B)** harbouring the indicated recombinant plasmids was measured. The fluorescence (arbitrary units) corresponds to 0.8 ml of culture. Each bar represents the mean ± SD of three independent experiments.

Next, we analysed whether plasmid pAST can be used for the study of regulated promoters. For this purpose, we selected the promoter of the pneumococcal fuculose kinase gene (*fcsK*), the first gene of the fucose operon, which is induced by fucose (Chan *et al*., 2003). This promoter (*PfcsK*) has a canonical −10 hexamer and a near-consensus −35 sequence (TTGAaA). Both sequence elements are separated by 17 nucleotides. According to primer extension experiments, transcription of the *fcsK* gene starts at an adenine residue located 24 nucleotides upstream of the initiation codon (Chan *et al*., 2003). A 117-bp DNA fragment containing the *PfcsK* promoter was inserted into the *Xba*I site of the pAST vector, generating plasmids pAST-*PfcsK* (gene *gfp* under the control of the *PfcsK* promoter) (Fig. 11A) and pAST-o*PfcsK* (fragment in the opposite orientation). Both recombinant plasmids were introduced into *S. pneumoniae* 708. Then, we examined whether fucose induces *gfp* expression in cells carrying the pAST-*PfcsK* plasmid. Cells harbouring pAST-o*PfcsK* were used as control. Since *S. pneumoniae* is unable to grow in media containing fucose as the sole carbon source (Chan *et al*., 2003) bacteria were grown in media containing 0.3% sucrose and different concentrations of fucose to late logarithmic phase (OD_{650}=0.6). The bacterial growth rate was similar under the various conditions assayed (not shown). In a first approach, *gfp* expression was analysed by Western blotting using monoclonal GFP antibodies (Fig. 11B). A protein band was detected in cells carrying pAST-*PfcsK* but not in control cells (plasmid pAST-o*PfcsK*). Since pre-stained proteins were run in the same gel, exposition of the blot to X-ray films allowed us to determine that such a band had the mobility expected for GFP (~28 kDa) (not shown). The Western blot analysis

revealed a basal level of *gfp* expression in cells grown without fucose. However, compared to cells grown without fucose, the intensity of the GFP band was 4.5-fold higher in cells grown with 1% fucose. Hence, the *PfcsK* promoter cloned into the pAST vector is activated by fucose.



C

Fucose (%)	Fluorescence (arbitrary units)
-	64.66 ± 5.95
0.1	91.96 ± 5.51
0.2	121.14 ± 10.17
0.4	205.45 ± 10.58
0.6	287.75 ± 50.97
1	339.41 ± 52.24

Figure 11. Fucose-induction of *gfp* expression in pneumococcal cells carrying plasmid pAST-*PfcsK*. Cells were grown in media containing 0.3% sucrose and the indicated amount of fucose to an OD_{650}=0.6. Cells harbouring plasmid pAST-o*PfcsK* were used as control. **(A)** Relevant features of the pAST-*PfcsK* plasmid. T1T2: tandem terminators *T1* and *T2* of the *E. coli rrnB* rRNA operon. The translation initiation signals (SD) are followed by a promoter-less *gfp* allele. B: *Bam*HI, *Sm*: *Sma*I, *S*: *Sac*I. **(B)** Western blot analysis using antibodies against GFP. Total proteins from cell extracts were separated by SDS-PAGE (14% polyacrylamide). Pre-stained proteins (Invitrogen) were run in the same gel as molecular weight markers (not shown). **(C)** Intensity of fluorescence in cultures (400 µl) of pneumococcal cells carrying the pAST-*PfcsK* plasmid. The intensity of fluorescence in cultures of cells carrying the pAST-o*PfcsK* plasmid was 43.22 ± 2.30 in the absence of fucose and 41.15 ± 1.81 in the presence of 1% fucose. In each case, three independent cultures were analysed.

These results were further confirmed by fluorescence assays (Fig. 11C). In the absence of fucose, the fluorescence in cells carrying pAST-*PfcsK* (64.66 ± 5.95 units) was slightly higher than in cells harbouring pAST-o*PfcsK* (43.22 ± 2.30; control cells). Thus, there is a low basal level of *gfp* expression. Moreover, the fluorescence in cells carrying pAST-*PfcsK* increased as a function of the fucose concentration (from 0.1% to 1%). Specifically, a 5-fold increase in fluorescence was observed when the medium was supplemented with 1% fucose (Fig. 11C). Therefore, we conclude that plasmid pAST is a useful vector for *in vivo* studies of regulated promoters.

1.3. The pDLF and pDLS expression vectors

To date, the availability of plasmid-based expression vectors to clone homologous and heterologous genes in *E. faecalis* is very scarce. Since such a tool was essential for *in vivo* studies of the *maEfa* gene, we constructed two expression vectors named pDLF and pDLS. They are based on the pDL287 plasmid, a pVA380-1

derivative which carries a kanamycin resistance gene and replicates via the rolling circle mode (LeBlanc *et al*., 1993). Vector pDLF (Fig. 12) was constructed by cloning a 194-bp that contains the enterococcal *P2493* promoter (Hancock *et al*., 2003) into the *Cla*I site of pDL287. To construct the pDLS vector (Fig. 12), a 202-bp region that contains the pneumococcal *PsulA* promoter (Lacks *et al*., 1995; López *et al*., 1987) was inserted into the *Cla*I site of pDL287. As shown in Figure 10, unlike in pneumococcal cells, the *P2493* promoter was stronger than the *PsulA* promoter in enterococcal cells. Both expression vectors contain an engineered restriction site for *Sph*I downstream of the inserted promoters (see Fig. 12). Promoter-less genes inserted into *Sph*I can be expressed not only in *E. faecalis* (see Results, Section 4.4) but also in *S. pneumoniae* (unpublished results).

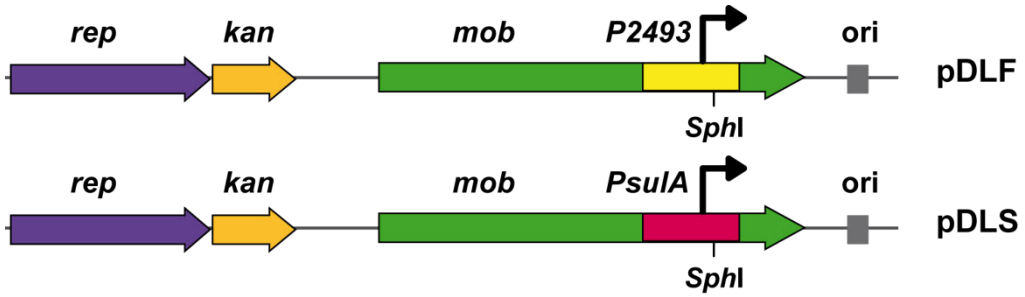


Figure 12. Expression vectors pDLF and pDLS. Promoters *P2493* (yellow) and *PsulA* (pink) were inserted into the *Cla*I site of the pDL287 plasmid (LeBlanc *et al*., 1993). Both plasmids carry an engineered restriction site for *Sph*I downstream of the promoter. Relevant features of pDL287 are indicated: *rep* gene is involved in plasmid DNA replication; *mob* gene encodes a mobilization protein; and *kan* gene confers kanamycin resistance. The location of the replication origin (ori) is indicated.

CHAPTER 2

In vivo transcription of the enterococcal *maEfa* gene

In this chapter, we show that the *maEfa* gene is transcribed in different *E. faecalis* strains under standard bacterial growth conditions. We have identified the promoter of the *maEfa* gene (promoter *Pma*) using several *in vivo* approaches: RT-PCR assays, promoter-reporter fusions (vector pAST) and primer extension experiments. Moreover, using the pAS terminator-probe vector, we have found that there is a functional transcriptional terminator upstream of the *Pma* promoter.

2.1. Genetic organization of the *maEfa* region

The complete genome sequence of the enterococcal V583 strain was published in 2003 (Paulsen *et al*., 2003) (GenBank AE016830.1). The EF3013 gene (here named *maEfa*) encodes a potential transcriptional regulator of the Mga/AtxA family. The ATG codon at coordinate 2889087 is likely the translation start site of the *maEfa* gene (locus_tag EF3013) (Fig.13). Translation from this ATG codon would generate a protein of 482 residues (MA*Efa*) as there is a stop codon (TAA) at coordinate 2890535. Downstream of the stop codon, there is a putative Rho-independent transcriptional terminator (coordinates 2890541- 2890576). The *maEfa* gene is flanked by two uncharacterized genes, the EF3012 (coordinates 2887814-2888395) and EF3014 genes (coordinates 2890541-2893093), which are predicted to encode a membrane protein and a cation transporter E1-E2 family ATPase, respectively. The *maEfa* and EF3014 genes are located on different strands. According to the BioCyc database (www.biocyc.org), the three genes would be transcribed as monocistronic mRNAs. In the enterococcal OG1RF strain, whose genome sequence was published in 2008 (Bourgogne *et al*., 2008) (GenBank CP002621.1), the *maEfa* gene (locus_tag OG1RF_12293; coordinates 2421605 to 2423053) is conserved. It encodes a protein that differs from the MA*Efa* protein of V583 strain in five amino acid residues (A37T, Q131L, M145T, S193N, I388S). Genes flanking the *maEfa* gene are also conserved in strain OG1RF. Regarding the enterococcal JH2-2 strain, we determined the nucleotide sequence of a chromosomal region (3383 nucleotides) that includes the corresponding EF3012 and *maEfa* genes. It also contains part of the corresponding EF3014 gene (the last 612 nucleotides). Compared to strain V583, MA*Efa* from strain JH2-2 has five amino acid changes (K133N, T139M, M145T, A152E, I388S).

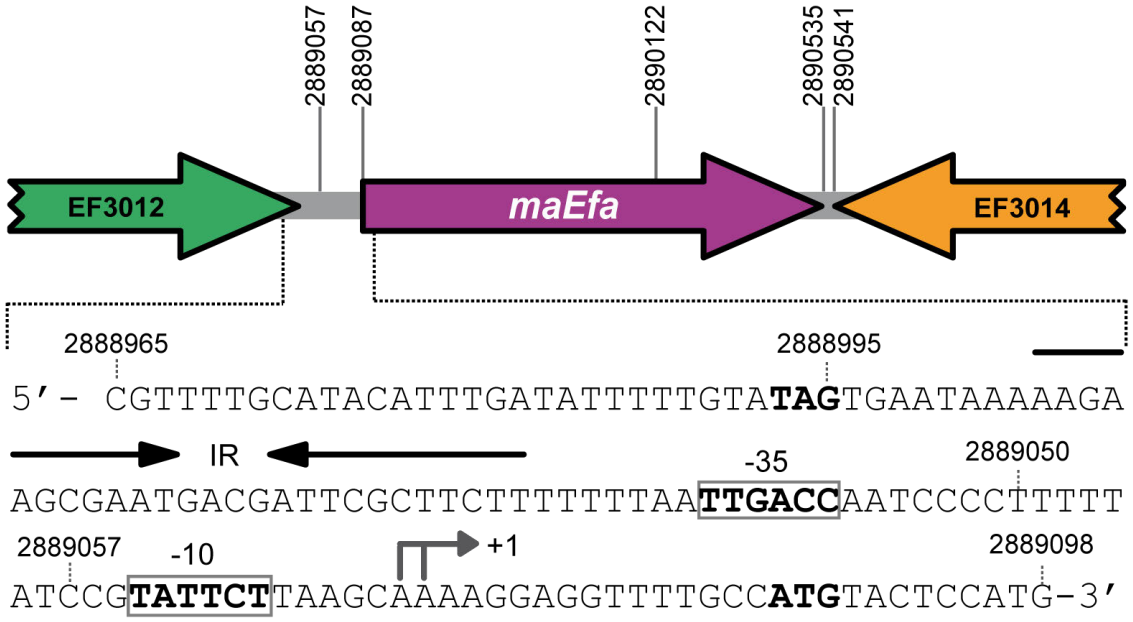


Figure 13. Genetic organization of the *maEfa* region (locus_tag EF3013) in the *E. faecalis* V583 genome. The coordinates of the predicted start and stop codons are shown. The nucleotide sequence of the region spanning coordinates 2889098 to 2888965 is shown. The stop codon (TAG) of EF3012 and the start codon (ATG) of *maEfa* are indicated in boldface letters. The transcription start site (+1 position) of the *maEfa* gene, and the main sequence elements (-35 box and -10 box) of the *Pma* promoter identified in this work are indicated. IR: inverted-repeat.

2.2. Identification of the *Pma* promoter

To examine whether the *maEfa* gene is transcribed in strains V583, OG1RF and JH2-2 under standard bacterial growth conditions, RT-PCR experiments were performed. Total RNA was isolated from the three strains. The oligonucleotide A (Fig. 14), which anneals to an internal region of the *maEfa* gene, was used for cDNA synthesis. The cDNA products were further amplified by PCR (Fig. 14). With oligonucleotides A and B, a PCR product that migrated at the position expected for a 439-bp DNA was synthesized. No PCR products were detected using oligonucleotides A and C, although they were able to amplify a fragment of 498-bp when chromosomal DNA was used as template (positive control). Results shown in Figure 14 correspond to strain V583. Similar results were obtained using total RNA from strains OG1RF and JH2-2. Moreover, we performed RT-PCR assays using oligonucleotide D for cDNA synthesis. Further amplification of the cDNA molecules with oligonucleotides D and B generated a product of 1427-bp (not shown). These results demonstrated that *maEfa* is transcribed in the three strains studied under standard laboratory conditions. In addition, we delimited that *maEfa* transcription starts between coordinates 2889039

and 2889098 of the V583 genome (Fig. 14). The nucleotide sequence of such a 60-bp region is identical in strains OG1RF and JH2-2.

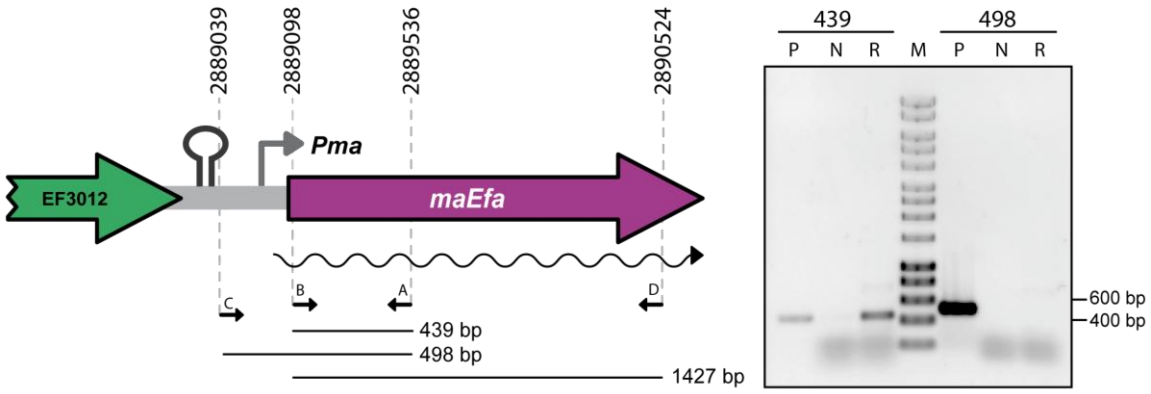


Figure 14. Transcription of *maEfa in vivo*. RT-PCR assays using total RNA from V583 cells. The position of the four primers (A, B, C and D) used in the experiment is shown. RT-PCRs samples (lanes R) were analysed by agarose (0.8%) gel electrophoresis. The sizes (in bp) of the DNA regions amplified by PCR using chromosomal DNA and the corresponding pair of primers (lanes P; positive control) are indicated. RT-PCRs without adding the reverse transcriptase were also performed (lanes N; negative control). The sizes (in bp) of DNA fragments (lane M) used as molecular weight markers (HyperLadder I, Bioline) are indicated on the right of the gel.

To determine the transcription initiation site of the *maEfa* gene, we performed primer extension assays using total RNA from V583 cells and different primers. Nevertheless, we failed to identify such a site suggesting that the amount of *maEfa* transcripts in the RNA preparations was small. To amplify the signal, a 214-bp region of the V583 genome (*Pma* region; coordinates 2888864 to 2889078) was inserted into the *Sac*I site of the pAST promoter-probe vector (Fig. 15). The pAST-*Pma* recombinant plasmid was then introduced into the enterococcal JH2-2 strain. The intensity of fluorescence in cells harbouring plasmid pAST-*Pma* was 2.2-fold higher than in pAST-containing cells (Fig. 15), indicating that the *Pma* region contained a promoter signal. Sequence analysis of the 214-bp *Pma* region revealed the existence of a putative promoter (*Pma*) (Fig. 13), which shows a 4/6 match at the -10 hexamer (5´-**TAT**TC**T**-3´) and a 5/6 match at the -35 hexamer (5´-**TTGAC**C-3´) (consensus residues are shown in bold). The -35 and -10 elements are separated by 17 nucleotides. Additional experiments confirmed that promoter *Pma* drives transcription of the *gfp* gene in plasmid pAST-*Pma*. Firstly, a deletion of 19 nucleotides in the *Pma* region (*Pma*Δ19 region, coordinates 2888864 to 2889059) reduced the intensity of fluorescence to background levels (cells harbouring plasmid pAST-*Pma*Δ19 *versus* cells harbouring pAST) (Figs. 13 and 15). Such a deletion removes the -10 element of the *Pma*

promoter (Fig. 15). Secondly, two cDNA extension products of 108 and 109 nucleotides were detected using total RNA from cells harbouring plasmid pAST-*Pma* and the INTgfp primer (it anneals to *gfp* transcripts) (Fig. 16). Thus, transcription of the *gfp* gene starts at the proper distance of 6-7 nucleotides from the -10 element of the *Pma* promoter. These results demonstrate that the *Pma* promoter is functional *in vivo* under our bacterial growth conditions.

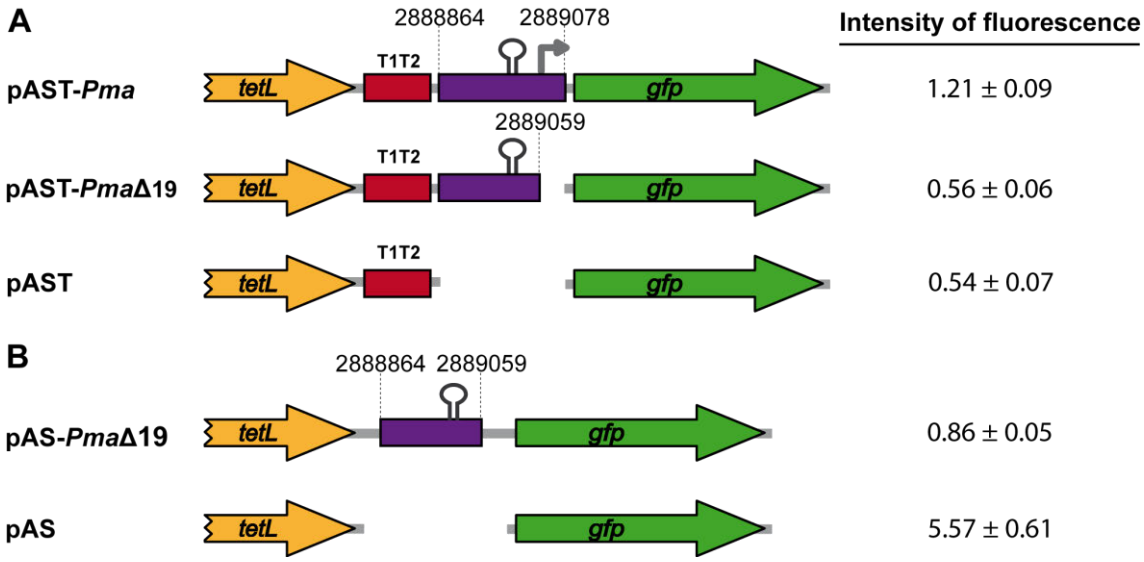


Figure 15. Fluorescence assays. **(A)** Activity of the *Pma* promoter. The promoter-probe vector pAST is described in this work (Fig. 8). Genes *tetL* (tetracycline resistance determinant) and *gfp* (green fluorescence protein) are indicated. The T1T2 box represents the tandem terminators T1 and T2 of the *E. coli rrnB* rRNA operon. The coordinates of the V583 DNA regions are indicated. The stem-loop structure represents the inverted-repeat (IR) identified upstream of the *Pma* promoter (this work). **(B)** Role of the IR element. The terminator-probe vector pAS is described in this work (Fig. 5). The intensity of fluorescence (arbitrary units) corresponds to 0.8 ml of culture (OD_{650}=0.3). In each case, three independent cultures were analysed.

Sequence analysis of the region located between the TAG stop codon of the locus_tag EF3012 (coordinate 2888995) and the *Pma* promoter revealed the existence of an inverted-repeat followed by a short stretch of thymine residues (IR element) (Fig. 13). This sequence element has the features of a Rho-independent transcriptional terminator. To examine the efficiency of this IR element as transcriptional terminator, we inserted the *Pma*Δ19 region (196-bp; coordinates 2888864 to 2889059) into the pAS terminator-probe vector. The recombinant plasmid pAS-*Pma*Δ19 (Fig. 15) was introduced into strain JH2-2. Compared to pAS-carrying cells, the intensity of fluorescence was 6.4-fold lower in cells harbouring plasmid pAS-*Pma*Δ19 (Fig. 15). Thus, there is a functional transcriptional terminator signal upstream of the *Pma* promoter.

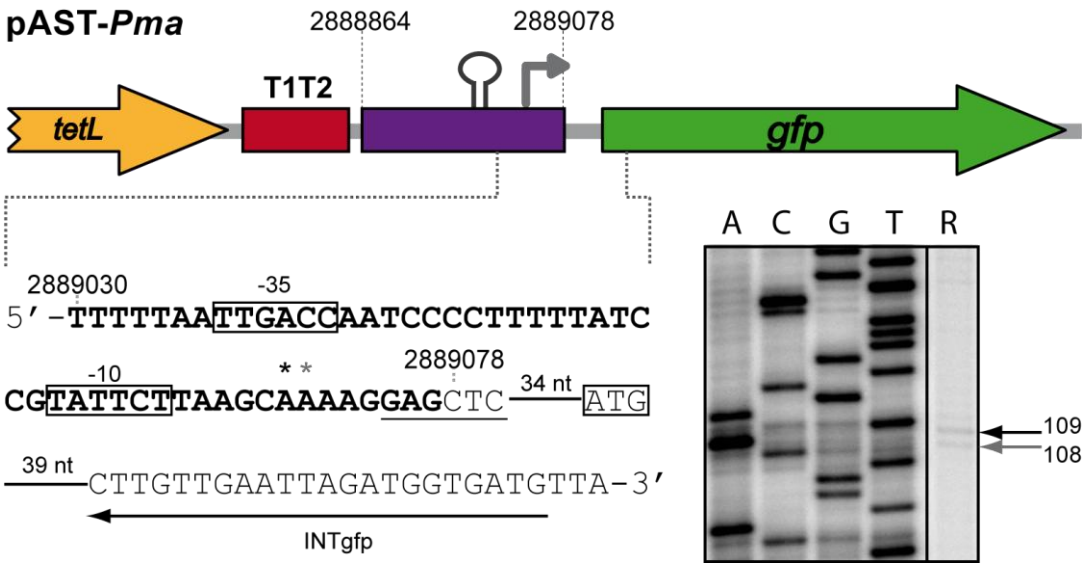


Figure 16. Initiation of transcription at the *Pma* promoter. Primer extension reactions were carried out using total RNA from JH2-2 cells harbouring plasmid pAST-*Pma*. The main sequence elements of the *Pma* promoter (-35 and -10 boxes), and the ATG initiation codon of the *gfp* gene are indicated. The *Sac*I site is underlined. The asterisks indicate the 3´-end of the cDNA products synthesized using the INTgfp primer. The sizes of the cDNA products (lane R) are indicated in nucleotides on the right of the gel. Dideoxy-mediated chain-termination sequencing reactions (M13mp18 DNA and primer -40 M13) were used as DNA size markers (lanes A, C, G, T).

CHAPTER 3

Features of the MA*Efa* protein

In this chapter, we show that the organization of predicted functional domains in MA*Efa* is similar to that found in the members of the Mga/AtxA family of global transcriptional regulators. We also describe a protocol to overproduce and purify an untagged form of the MA*Efa* protein. The purity grade of the protein preparations has allowed us to study the oligomerization state of MA*Efa* by gel filtration chromatography and analytical ultracentrifugation (sedimentation velocity and sedimentation equilibrium assays).

3.1. MA*Efa* is a potential member of the Mga/AtxA family of global transcriptional regulators

The organization of predicted functional domains in MA*Efa* resembles that found in the members of the Mga/AtxA family of transcriptional regulators (Mga, AtxA and Mga*Spn*). According to the conserved domain database of the National Center for Biotechnology Information (NCBI) (Marchler-Bauer *et al.*, 2015), MA*Efa* has two putative DNA-binding domains within the N-terminal region, the HTH_Mga (Family PF08280, residues 11 to 69) and Mga (Family PF05043, residues 76 to 164) domains, and an EIIB-like motif at the C-terminal region (residues 407 to 476) (Fig. 17). Two DNA-binding domains and an EIIB-like motif are also present in Mga, AtxA and Mga*Spn* (Hammerstrom *et al*., 2015; Hondorp *et al*., 2013; Solano-Collado, 2014; Solano-Collado *et al*., 2013). In Mga, the two DNA-binding domains were shown to be required for DNA binding and transcriptional activation (McIver and Myles, 2002; Vahling and McIver, 2006). Analysis of MA*Efa* with the protein structure prediction server Phyre2 (Kelley and Sternberg, 2009) revealed that the central region of MA*Efa* has structural homology to PTS regulatory domains (PRDs) (residues 171 to 389) (Fig. 17). The Mga, AtxA and Mga*Spn* proteins are also PRD-containing global regulators (Hammerstrom *et al*., 2015; Hondorp *et al*., 2013; McIver and Myles, 2002; Solano-Collado, 2014; Vahling and McIver, 2006). It has been shown that the activity of AtxA is modulated by phosphorylation of histidine residues within the PRDs (Hammerstrom *et al*., 2015; Tsvetanova *et al*., 2007), and this is likely a general feature of the Mga/AtxA family of regulators. In fact, Mga can be phosphorylated *in vitro* by components of the PTS (Hondorp *et al*., 2013).

Figure 17. Predicted functional domains in MA*Efa*. Predicted domains according to *in silico* analysis using Conserved Domain Database (CDD) and Phyre2 programs (Kelley and Sternberg, 2009; Marchler-Bauer *et al*., 2015). MA*Efa* contains at the N-terminal region two putative DNA binding domains: the HTH-Mga and Mga motifs. The central region is predicted to have structural homology to PTS regulatory domains (PRD). The C-terminal region has a putative EIIB-like motif.

The three-dimensional structures of MA*Efa* (EF3013; PDB 3SQN) ((Osipiuk, 2011), unpublished results) and AtxA (PDB 4R6I) (Hammerstrom *et al*., 2015) have been solved. Both proteins crystallize as dimers. The closest structural homolog for the AtxA C-terminal region (EIIB-like motif) is the C-terminal region of MA*Efa* (Hammerstrom *et al*., 2015). Moreover, it has been reported that the closest overall protein structure identified with similarity to Mga was the crystal structure of MA*Efa* (Hondorp *et al*., 2013). In conjunction, these observations suggest that MA*Efa* is a potential member of the Mga/AtxA family of transcriptional regulators.

3.2. Purification of an untagged form of MA*Efa*

The protocol used to overproduce and purify an untagged form of MA*Efa* is essentially based on the one described previously for Mga*Spn* (Solano-Collado *et al*., 2013). This protocol is explained in detail in Methods (Section 7.1). The *maEfa* gene from the enterococcal strain V583 was cloned into the inducible expression vector pET24b, and the recombinant plasmid (pET-*maEfa)* was introduced into the *E. coli* BL21 (DE3) strain. The *maEfa* gene is under the control of a promoter that is recognized by the T7 RNAP. The BL21 (DE3) strain carries the T7 RNAP-encoding gene fused to the *lac*UV5 promoter, which is an IPTG-inducible promoter. Thus, expression of the *maEfa* gene is induced when IPTG is added to the bacterial cultures (Fig. 18, lanes 1 and 2). Various bacterial growth conditions and IPTG concentrations were assayed to define the optimal conditions (TY medium, 37ºC and 1 mM IPTG) for *maEfa* gene expression. The large-scale purification of the MA*Efa* protein (Fig. 18) involved essentially four steps: (*i*) nucleic acids and some DNA-binding proteins, including MA*Efa*, were precipitated with PEI at a low ionic strength; (*ii*) the MA*Efa* protein was eluted from the PEI pellet using a higher ionic strength buffer; (*iii*) the sample was subjected to heparin affinity chromatography and (*iv*) the protein preparation was loaded onto a gel-filtration column. Protein fractions were analysed in each step of the purification process using Coomassie-stained SDS-Tris-Glycine

polyacrylamide (12%) gels (Fig. 18). Under these conditions, MA*Efa* migrated between the 45 and 66 kDa bands of the molecular weight marker (Fig. 18), which agrees with the theoretical molecular weight of the MA*Efa* monomer (56,247 Da; 482 residues). The final protein preparation was more than 95% pure as it was shown using Coomassie-stained overloaded gels. Additionally, SDS-Tricine-PAGE (16%) was performed to rule out the presence of proteins smaller than 30 kDa. Determination of the N-terminal amino acid sequence of MA*Efa* by Edman degradation showed that the first Met residue was not removed in the purified protein.

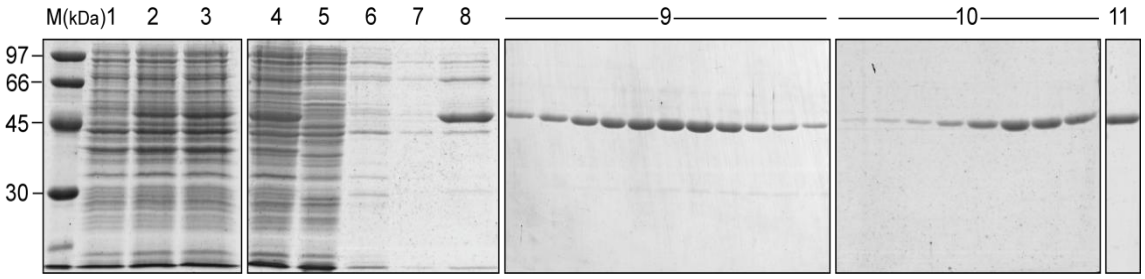


Figure 18. **Purification of MA*Efa*.** Fractions of the different purification steps were analysed by electrophoresis on SDS-Tris-Glycine polyacrilamide (12%) gels. From lane 1 to 3: induction of *maEfa* gene expression; (1) uninduced cultures; (2) *maEfa* expresion induced with IPTG for 30 min; (3) after treatment with rifampicine for 1 h. From lanes 4 to 11: purification steps: (4) supernatant of a clear lysate (5) supernatant after PEI precipitation at low ionic strength; (6-7) proteins eluted from the PEI pellet using the same low ionic strength buffer; (8) proteins eluted from the PEI pellet using a higher ionic strength buffer; (lanes group 9) heparin affinity chromatography: proteins retained in the column were eluted with a salt gradient (0.2-0.6 M NaCl); (lanes group 10) gel filtration chromatography, (11) final protein preparation after concentration. M indicates the molecular weight standards (in kDa).

3.3. MA*Efa* forms dimers in solution

To study the oligomerization state of MA*Efa* in solution and its hydrodynamic properties, we performed gel filtration chromatography and analytical ultracentrifugation experiments. First, to determine the molecular size (Stokes radius) of MA*Efa*, we carried out gel filtration chromatography using a running buffer that contained 200 mM of NaCl. The elution profile is shown in Figure 19. MA*Efa* eluted from the column as a single peak, and the elution volume was used to calculate the K_{av} value, (see Methods Section 8.4). To calibrate the column, several standard proteins of known Stokes radius were loaded and their corresponding values of K_{av} was calculated. The Stokes radius of MA*Efa* determined from the calibration curve, $((-\log K_{av})^{1/2}$ *versus* the Stokes radius) (Fig. 19B) was 43 Å, which is slightly lower than the Stokes radius of the alcohol dehydrogenase protein (45 Å), with a molecular weight of 150 kDa. Since the

molecular weight of the MA*Efa* monomer calculated from its amino acid sequence is 56,247 Da, the protein appears to be a dimer under our experimental conditions

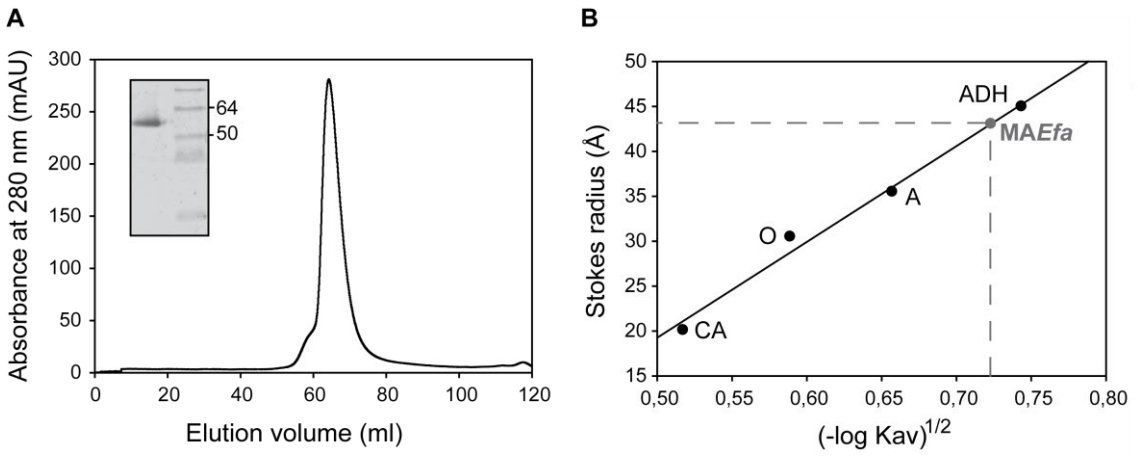


Figure 19. Determination of the Stokes radius of MA*Efa*. **(A)** Elution profile of MA*Efa* on the gel filtration column HiLoad Superdex 200 (Amersham) at a flow rate of 1 ml/min. The inset shows an analysis of the eluted protein by SDS-polyacrilamide (12%) gel electrophoresis. The molecular weight (in kDa) of proteins used as markers. **(B)** Calibration plot using proteins of known Stokes radii: alcohol dehydrogenase (ADH; 45 Å), albumin (A; 35 Å), ovalbumin (O; 30.5 Å) and carbonic anhydrase (CA; 20.1 Å). Stokes radius of MA*Efa* was determined by interpolation.

To corroborate the previous result, analytical ultracentrifugation experiments (sedimentation velocity and sedimentation equilibrium) were carried out at 5 and 10 µM (Fig. 20). In both cases, sedimentation velocity profiles showed a major peak (92-98%, respectively), with an $S_{20,w}$ value of 6 S. Samples were also analysed by sedimentation equilibrium. At 5 µM, the experimental data are best fit to an average molecular mass ($M_{w,a}$) of 118,000 ± 1,000 Da, a value that corresponds with the theoretical mass of a MA*Efa* dimer (112,495 Da). A similar average molecular mass was determined at 10 µM (121,000 ± 1,000 Da). The frictional ratio (f/f_0) calculated was 1.34, indicating that the dimer of MA*Efa* deviates from the behaviour expected for a rigid spherical particle (f/f_0=1). From these results, we conclude that, under the conditions tested, MA*Efa* forms dimers in solution.

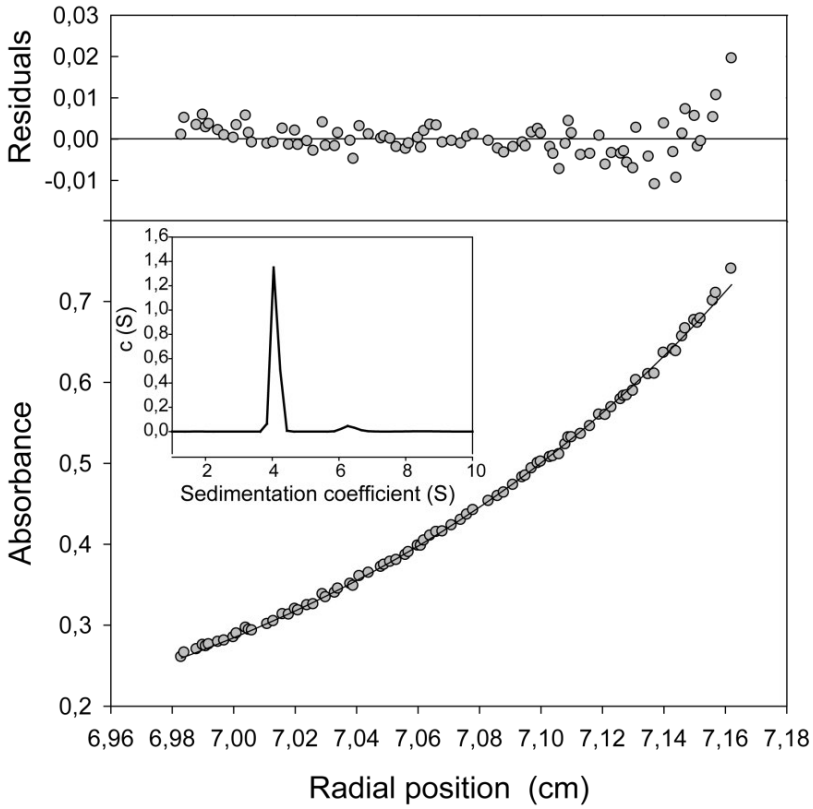


Figure 20. Analytical ultracentrifugation analysis of MA*Efa*. Sedimentation equilibrium profile of MA*Efa* (5 µM) at 10,000 rpm and 12ºC. The lower part shows the experimental data (circles) and the best fit (solid line) to a single species with a $M_{w,a}$=118,000. The upper part shows the residuals to the fit. The inset shows the sedimentation coefficient distribution determined for the same MA*Efa* protein sample at 48,000 rpm and 12ºC.

CHAPTER 4

Interaction of the MA*Efa* protein with DNA

In this chapter, we analyse the interaction of MA*Efa* with DNA by gel retardation and footprinting experiments. We demonstrate that MA*Efa* binds to linear double-stranded DNA generating multimeric complexes. MA*Efa* does not seem to recognize a specific nucleotide sequence. Moreover, we show that MA*Efa* binds preferentially to a site located upstream of the promoter of its own gene (promoter *Pma*). This site contains a potential intrinsic curvature.

4.1. MA*Efa* generates multimeric complexes on linear double-stranded DNA

The Mga global transcriptional regulator from *S. pyogenes* was shown to activate the transcription of its own gene by binding to a region located upstream of the major promoter (McIver *et al*., 1999). In addition, the Mga*Spn* protein of *S. pneumoniae* is able to bind to the promoter of its own gene as primary binding site (Solano-Collado *et al*., 2013), although it has not yet been determined whether binding of Mga*Spn* to such a site influences the expression of its own gene *in vivo*. For these reasons to investigate the DNA binding properties of the untagged MA*Efa* protein, we performed electrophoretic mobility shift assays (EMSA) using, firstly, a linear double-stranded DNA (dsDNA) fragment that contains the promoter of the *maEfa* gene (*Pma*) (see Fig.13 in Chapter 2). This DNA fragment (217-bp; coordinates 2888932-2889148) was radioactively labelled at the 5'-end of the non-coding strand. The labelled DNA (2 nM) was incubated with increasing concentrations of MA*Efa* (20 to 360 nM) in the presence of non-labelled competitor calf thymus DNA (5 µg/ml) (Fig. 21A). The binding conditions were those described in Methods, Section 9.1. Free and bound DNAs were separated by electrophoresis on a native polyacrylamide (6%) gel. Labelled DNA bands were visualized using a Fujifilm Image Analyzer (FLA-3000). At the lower concentration of MA*Efa*, a DNA-protein complex was observed. When the protein concentration was raised, higher-order complexes appeared gradually, while complexes moving more rapidly disappeared. This result suggests that multimeric DNA-protein complexes are generated by sequential binding of several MA*Efa* units to the DNA molecule.

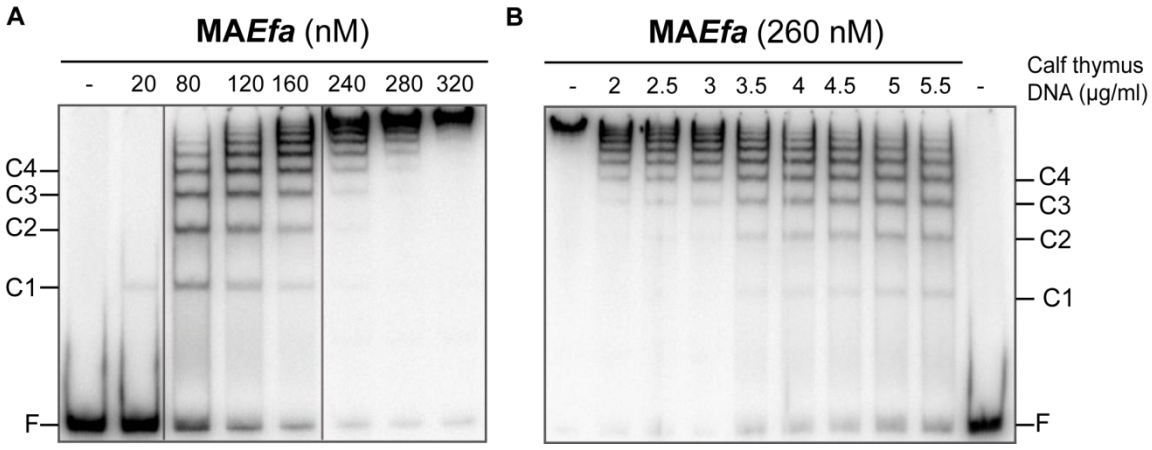


Figure 21. Detection of MA*Efa*-DNA complexes by EMSA. The 217-bp DNA fragment (coordinates 2888932- 2889148) was radiolabelled at the 5'-end of the non-coding strand. Free and bound DNAs were resolved on a native polyacrilamide gel (6%). Labelled bands were visualized with a Fujifilm Image Analyzer (FLA- 3000). **(A)** Formation of multimeric MA*Efa*-DNA complexes: the labelled DNA fragment (2 nM) was incubated with increasing amounts of MA*Efa* in the presence of non-labelled calf thymus DNA (5 µg/ml); **(B)** Dissociation experiments: the labelled DNA (2 nM) was mixed with 260 nM of MA*Efa* (formation of higher-order complexes), and then competed with the indicated amount of non-labelled calf thymus DNA.

Furthermore, we performed dissociation experiments using the ^{32}P-labelled 217-bp DNA fragment and non-labelled competitor calf thymus DNA (Fig. 21B). In this assay, the 217-bp DNA (2 nM) was mixed with MA*Efa* (260 nM) and incubated for 20 min at room temperature (formation of higher-order complexes). Then, increasing amounts of the competitor DNA were added to the reaction mixtures and they were incubated for 5 minutes at the same temperature. As the concentration of the competitor was increased, higher-order complexes disappeared, whereas complexes moving faster and free DNA appeared progressively. Thus, the dissociation of the protein units from the higher-order complexes was sequential. The appearance of free 217-bp DNA molecules indicated that the competitor DNA was able to displace all the protein units bound to the DNA molecule. We have also performed dissociation experiments using supercoiled plasmid DNA (pMV158) as competitor DNA (not shown). Our results show that MA*Efa* binds not only to linear dsDNA but also to supercoiled DNA.

To further analyse the DNA-binding properties of MA*Efa*, we determined the time required for the reaction mixture to reach the equilibrium (Fig. 22A). In this experiment, MA*Efa* (100 nM) was mixed with the non-labelled 217-bp DNA fragment (10 nM) under standard conditions (50 mM NaCl). After different incubation times (2.5–8.5 min), the reactions mixtures were loaded onto running gels (200 V). When the last

sample entered the gel, the voltage was reduced to 100 V. Under these conditions, the loading process itself acts to quench the association reaction, since the free DNA begins to be separated from complexes as soon as the sample is loaded onto the running gel (Carey, 1991). A faint band of free DNA and six protein-DNA complexes were detected at 2.5 min. The amount of the different complexes did not change at longer incubation times, indicating that the reactions had already reached the equilibrium at 2.5 min. A similar experiment was performed using a higher concentration of NaCl (200 mM). In this assay, the incubation time varied from 2 min to 10 min (not shown). Free DNA and six complexes were detected at 2 min. Their amount was similar at 10 min, confirming that the protein-DNA binding reaction reaches the equilibrium very fast.

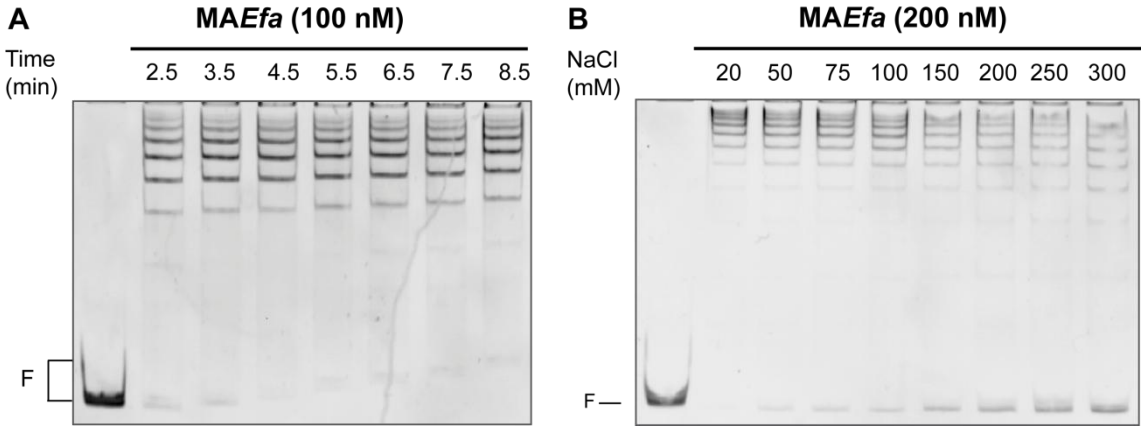


Figure 22. Binding of MA*Efa* to DNA under different conditions. The non-labelled 217-bp DNA fragment was used in both experiments. Free and bound DNAs were separated by native polyacrylamide (6%) gel electrophoresis. Bands were stained with GelRed (Biotium) and visualized using a Gel Doc system (Bio-Rad). **(A)** Time course formation of the DNA-protein complexes: the DNA fragment (10 nM) was mixed with 100 nM of MA*Efa*. Then, samples were incubated for the indicated time (from 2.5 to 8.5 min) and loaded onto the running gel (200 V). **(B)** Effect of NaCl concentration on the formation of MA*Efa*-DNA complexes: MA*Efa* (200 nM) was incubated with DNA (10 nM) in the presence of the indicated concentration of NaCl.

The DNA-binding ability of some proteins is highly dependent on the salt concentration of the reaction mixture (Laniel *et al.*, 2001). We have analysed the effect of NaCl concentration on the MA*Efa*-DNA binding reaction. The non-labelled 217-bp DNA fragment (10 nM) was mixed with MA*Efa* (200 nM). The reaction mixtures contained different concentrations of NaCl (20-300 mM). As shown in Figure 22B, at high salt concentrations (150-300 mM), a slight increase in the amount of free DNA and a small decrease in the amount of higher order-complexes was observed. Thus, high concentrations of NaCl affect slightly the binding of MA*Efa* to DNA. The results were

similar when the experiment was performed at a lower concentration of MA*Efa* (100 nM) (not shown).

4.2. MA*Efa* does not appear to recognize a specific nucleotide sequence

MA*Efa* is thought to be a member of the Mga/AtxA family of global response regulators. A general feature of these regulators seems to be their ability to bind to DNA with little or no sequence specificity (Hause and McIver, 2012; Solano-Collado *et al*., 2013). Therefore, we analysed the binding of MA*Efa* to additional dsDNA fragments by EMSA (Fig. 23). Specifically, we used a 321-bp DNA from the enterococcal V583 chromosome (coordinates 94488–94808) and a 322-bp DNA from the pneumococcal R6 chromosome (coordinates 1598010–1598331). Although these fragments have a similar A+T content (72.3% and 71.1%, respectively), sequence similarities are not apparent. Both DNA fragments were radioactively labelled. The labelled DNAs (1 nM) were incubated with increasing concentrations of MA*Efa* in the presence of non-labelled competitor calf thymus DNA (1 µg/ml). In both cases, MA*Efa* generated a pattern of DNA-protein complexes similar to that generated on the 217-bp DNA (Fig. 21A). Thus, MA*Efa* does not appear to recognize a specific nucleotide sequence. Moreover, these results support that MA*Efa* is able to form multimeric complexes on linear dsDNA.

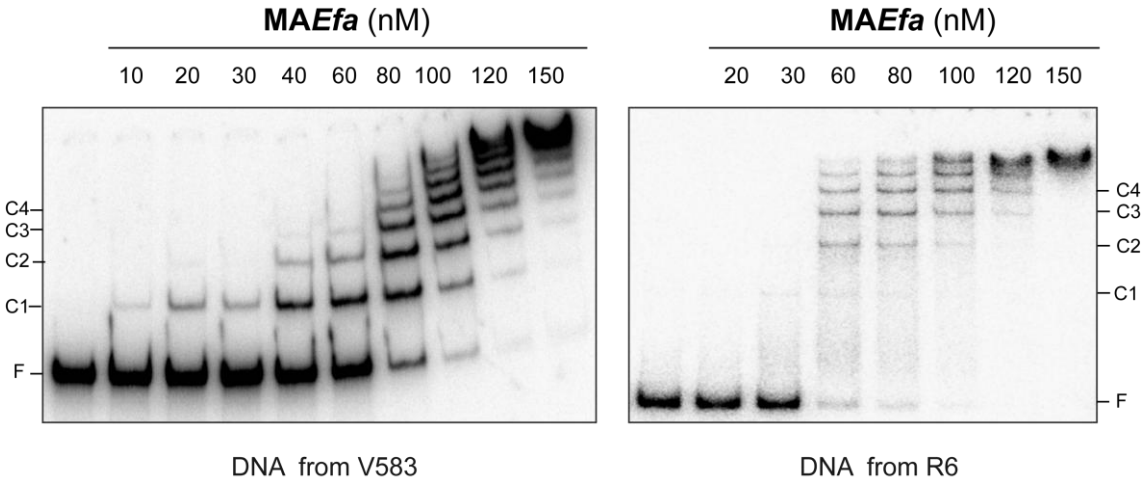


Figure 23. MA*Efa* binds to linear dsDNAs with low sequence specificity. EMSAs were carried out using a 321-bp DNA fragment from the enterococcal V583 chromosome and a 322-bp DNA fragment from the pneumococcal R6 chromosome. Both DNA fragments (1 nM) were radiolabelled and incubated with increasing amounts of MA*Efa* in the presence of non-labelled calf thymus DNA (1 µg/ml). Bands corresponding to free DNA (F) and MA*Efa*-DNA complexes (C1 to C4) are indicated.

4.3. MA*Efa*-His also generates multimeric complexes

We have also overproduced and purified a His-tagged MA*Efa* protein (MA*Efa*-His). It carries the Leu-Glu-6xHis tag fused to its C-terminus. To investigate whether this variant of the MA*Efa* protein binds to linear dsDNA, we performed EMSA experiments using two different DNA fragments from the enterococcal V583 chromosome: a 260-bp DNA fragment (coordinates 2888858-2889117), which includes the *Pma* promoter, and a 321-bp DNA fragment (coordinates 94488–94808) from a region unrelated to the *maEfa* gene. Apparently, these fragments do not have sequence identity. As shown in Figure 24, MA*Efa*-His binds to both DNA fragments forming a pattern of complexes which is similar to that formed by MA*Efa* protein on such DNAs (not shown). This result indicates that the eight amino acid peptide fused to the C-terminal end of MA*Efa* does not affect its ability to generate multiple DNA-protein complexes.

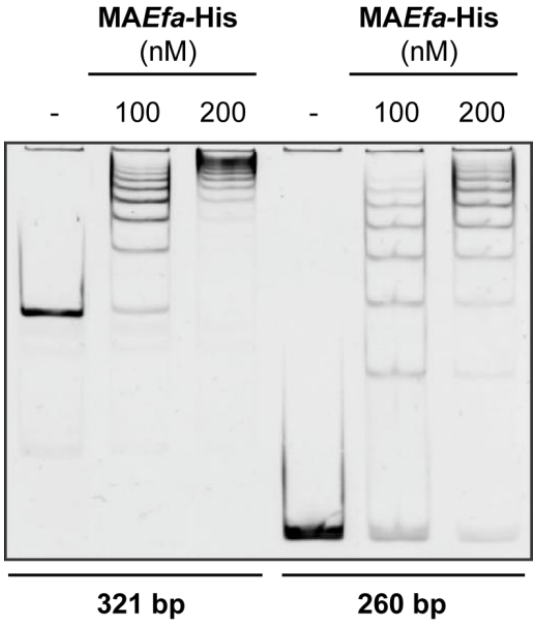


Figure 24. **MA*Efa*-His binds to linear dsDNA forming multiple complexes.** EMSA analysis of the interaction of MA*Efa*-His with two DNA fragments of the V583 chromosome. The non-labelled 321-bp and 260-bp DNAs were mixed with the indicated amounts of MA*Efa*-His. DNA was stained with GelRed (Biotium) and visualized using a Gel Doc system (Bio-Rad).

4.4. Binding of MA*Efa* to small DNA fragments

Next, we analysed the minimum DNA size required for MA*Efa* binding by EMSA experiments. We used DNA fragments of 20, 26, 32 and 40 bp. These fragments, which were obtained by annealing of complementary oligonucleotides, had been previously used in our laboratory to determine the minimum DNA size required for Mga*Spn* binding (Solano-Collado, 2014; Solano-Collado *et al*., 2013). The oligonucleotides carry sequences from the pneumococcal R6 chromosome. The non-labelled DNA fragments were incubated with increasing concentrations of MA*Efa*. In

the case of the 20-bp DNA fragment, a faint band was observed at 0.3 µM of MA*Efa* (Fig 25A). However, its intensity did not change significantly as the protein concentration was increased. Most of the DNA molecules moved as free DNA even at high concentrations (4.8-9 µM). A similar result was obtained with the 26-bp DNA fragment. As shown in Figure 25B, most of the 26-bp DNA moved as free DNA at 4 µM of MA*Efa*. Different results were obtained with the 32-bp and 40-bp DNA fragments (Fig. 25B and C). In both cases, as the protein concentration was increased, the amount of free DNA decreased. Furthermore, a MA*Efa*-DNA complex was detected at 0.2 µM of MA*Efa*, and the amount of such a complex increased at high protein concentrations (4-5 µM). From these results we conclude that the minimum DNA size required for MA*Efa* binding is between 26 and 32 bp.

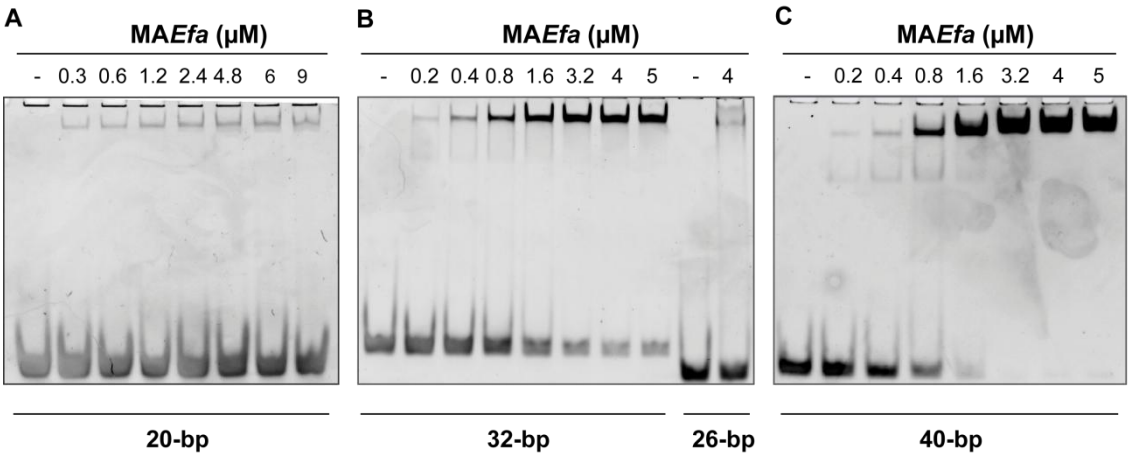


Figure 25. Binding of MA*Efa* to small DNA fragments. EMSA experiments were performed with DNA fragments of 20-bp **(A)**, 32-bp and 26-bp **(B)** and 40-bp **(C)**. The indicated concentration of MA*Efa* was mixed with 300 nM of DNA (20-bp and 26-bp DNA fragments) or with 200 nM of DNA (32 bp and 40 bp DNA fragments). Binding reactions were analysed by native polyacrilamide (8%)gels electrophoresis DNA was stained with GelRed (Biotium) and visualized using a Gel Doc system (Bio-Rad).

4.5. MA*Efa* recognizes a site located upstream of the *Pma* promoter

To further investigate the interaction of MA*Efa* with the promoter region of its own gene, we carried out DNase I footprinting experiments. We used a 227-bp DNA fragment (coordinates 2888932-2889158) and a 217-bp DNA fragment coordinates (2888932-2889148). Both fragments contain the promoter (*Pma*) promoter and the transcription initiation site of *maEfa* (Fig. 26). The 227-bp DNA fragment was radioactively labelled at the 5´-end of the coding strand, while the 217-bp DNA fragment was radioactively labelled at the 5´-end of the non-coding strand. We used the 217-bp DNA because we had technical problems to label the non-coding strand of

the 227-bp DNA. Binding reactions contained 2 nM of the labelled DNA fragment and different concentrations of MA*Efa.* On the coding strand and at 80 nM of MA*Efa*, two regions were protected against DNase I digestion (Fig. 27). They were located between the positions -74 to -80 and -88 to -110 with respect to the transcription initiation site of the *maEfa* gene. At higher protein concentrations, changes in the DNase I sensitivity (diminished cleavages) were observed along the DNA fragment. On the non-coding strand and at 120 nM, diminished cleavages were observed from -73 to -80 and from -88 to -101. When the protein concentration was increased, a slight protection against DNase I cleavage was observed in the regions flanking such positions. These results indicate that MA*Efa* recognizes preferentially a site on the linear DNA fragments that contain the *Pma* promoter. Such a primary binding site is located upstream of the *Pma* promoter, between the positions -73 and -110 relative to the transcription start site. Moreover, the protections observed at higher concentrations of MA*Efa*, together with the pattern of DNA-protein complexes observed by EMSA (Fig. 21), suggest that MA*Efa* might be able to spread along the DNA fragment subsequent to its initial binding.

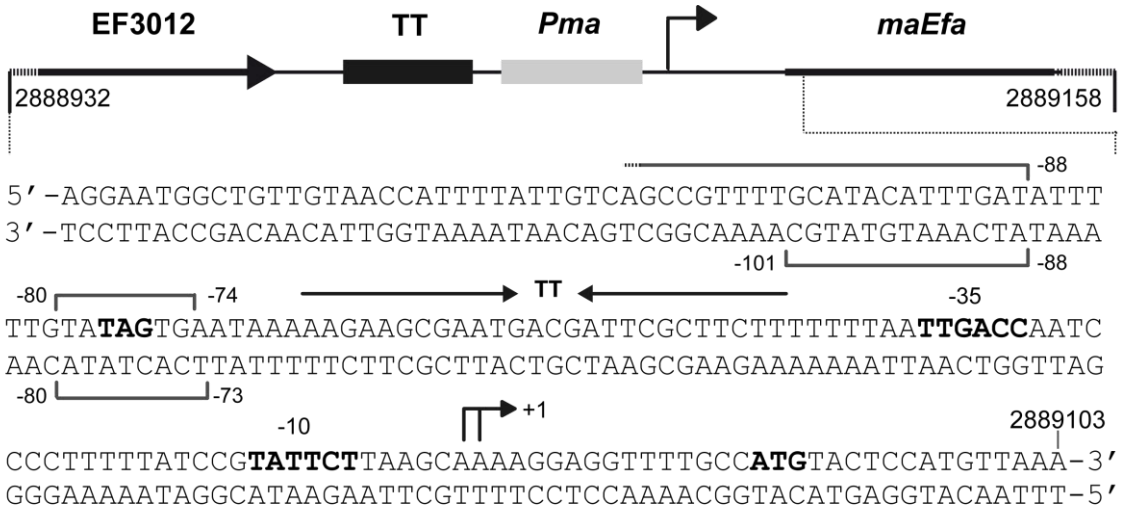


Figure 26. Relevant features of the two DNA fragments used for DNase I footprinting assays. The DNAs fragments were radiolabelled at the 5´-end of either the coding (227-bp DNA fragment, coordinates 2888932-2889158) or the non-coding strand (217-bp DNA fragment, coordinates 2888932-2889148) relative to the *maEfa* gene. The nucleotide sequence of the region spanning coordinates 2888932-2889103 of the V583 genome is shown. It includes the predicted stop codon (TAG) of EF3012, the start codon (ATG) of *maEfa,* the transcriptional terminator (TT), the *Pma* promoter (-35 and -10 elements) and the transcription initiation site of *maEfa*. The site recognized by MA*Efa* is indicated with brackets.

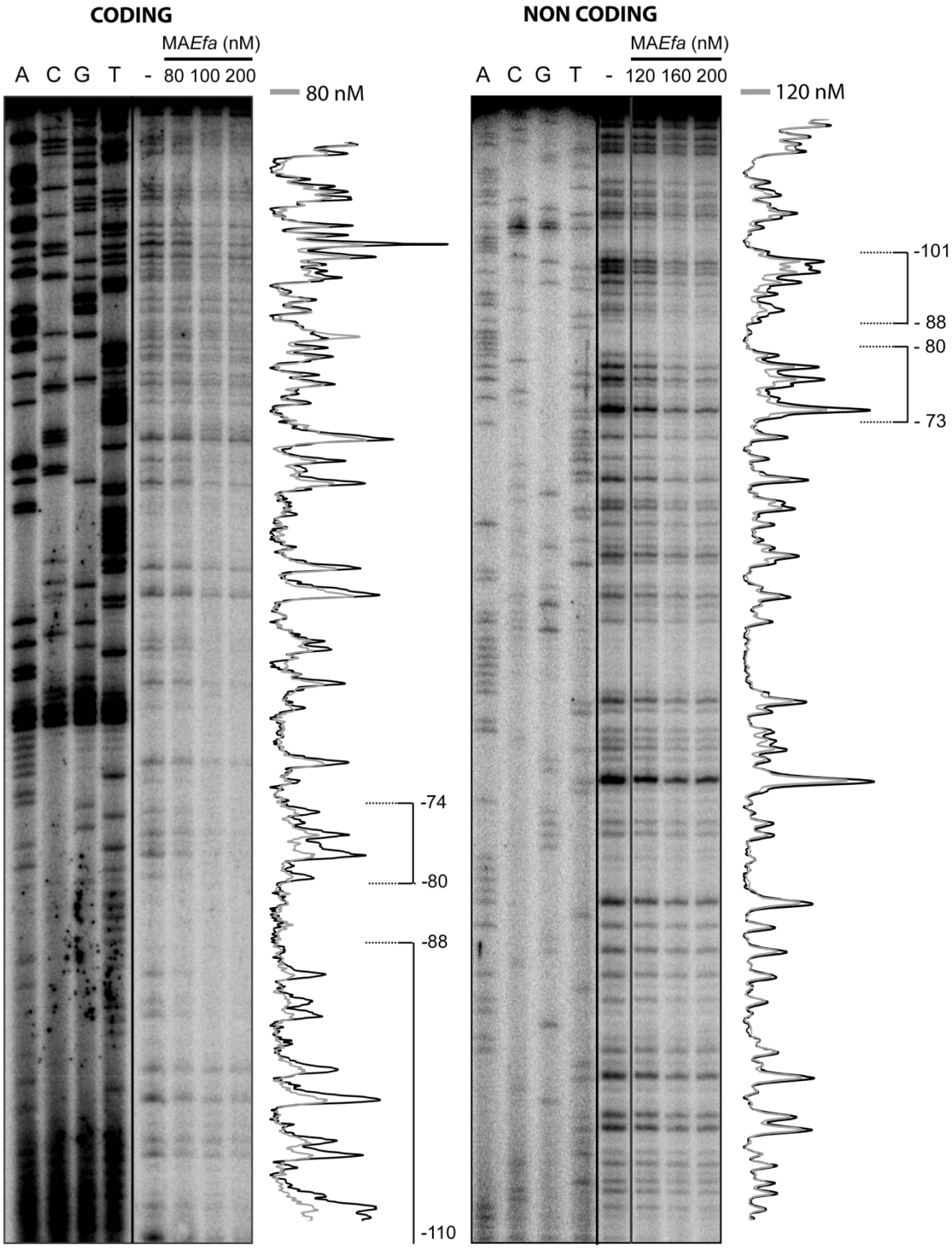


Figure 27. **DNase I footprints of complexes formed by MA*Efa* on linear DNA**. Coding and non-coding strands relative to the *Pma* promoter were ^{32}P-labelled at the 5'-end. Dideoxy-mediated chain termination sequencing reactions were run in the same gel (lanes A, C, G, T). All the lanes displayed came from the same gel. Densitometer scans corresponding to DNA without protein (black line) and DNA with the indicated concentration of protein (grey line) is shown. In this assay, the concentration of DNA was 2 nM. Brackets show the regions protected at the indicated concentration. The indicated positions are relative to the transcription start site of *maEfa*.

In vitro studies showed that the pneumococcal Mga*Spn* regulator interacts with DNA sites that contain a potential intrinsic curvature (Solano-Collado *et al.*, 2013). Moreover, the promoter regions of several target genes of AtxA are intrinsically curved (Hadjifrangiskou and Koehler, 2008). Therefore, we calculated the curvature propensity plot of the 227-bp DNA fragment using the bend.it program (Vlahovicek *et al.*, 2003). The profile of the DNA fragment contains one peak of potential sequence-dependent curvature at position -104 (Fig. 28), which is within the primary binding site of MA*Efa* (positions -110 to -73). The magnitude of the curvature propensity (14.13) is within the range calculated for experimentally tested curved motifs (Gabrielian *et al.*, 1997). Thus, the site recognized by MA*Efa* contains a potential intrinsic curvature.

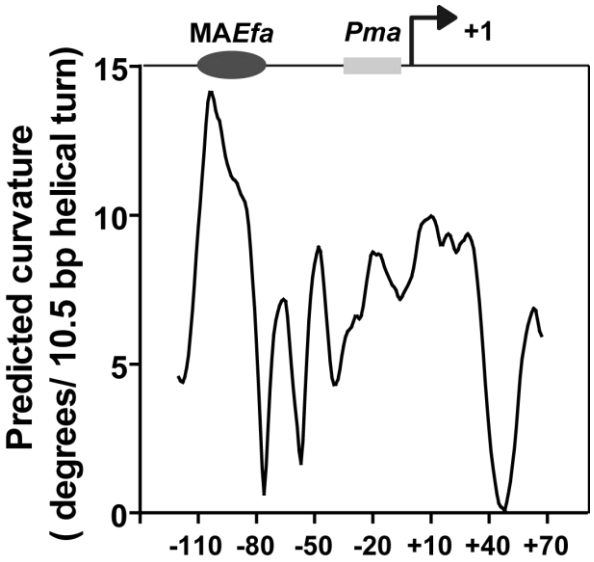


Figure 28. Curvature propensity plot of the 227-bp DNA fragment according to the bend.it program (Vlahovicek *et al.*, 2003). The primary binding site of MA*Efa* is indicated (positions -110 to -73). The *Pma* promoter and the transcription initiation site are also indicated.

Since MA*Efa* is able to recognize a site located upstream of the *Pma* promoter, we decided to analyse whether MA*Efa* influences the expression of its own gene *in vivo*. To this end, we constructed two plasmids to overproduce MA*Efa.* We inserted the *maEfa* gene from the V583 strain (*maEfa*$_{V583}$) (coordinates 2889087-2890535) into the expression vectors pDLF (Fig. 29A) and pDLS (this work, see Results 1.3.). In the pDLF*maEfa*$_{V583}$ and pDLS*maEfa*$_{V583}$ recombinant plasmids, the *maEfa* gene is under the control of the *P2493* and *PsulA* promoters, respectively. In *E. faecalis*, the *P2493* promoter is stronger than the *PsulA* promoter (see Chapter 1, Section 1.2). Both plasmids were introduced into JH2-2 cells. To confirm that MA*Efa* is synthesized in cells carrying the recombinant plasmids, we carried out Western blot assays using polyclonal antibodies against MA*Efa* (Fig. 29B). As expected, the amount of MA*Efa* in cells harbouring pDLF*maEfa*$_{V583}$ was higher than in those cells carrying pDLS*maEfa*$_{V583}$. In JH2-2 cells harbouring the pDLF vector, MA*Efa* was not detected. This was expected as we were not able to detect MA*Efa* in JH2-2 cells without plasmid

(not shown). Then, to analyse the effect of the MA*Efa* overproduction on the activity of the *Pma* promoter, we tried to introduce pDLF, pDLS*maEfa*$_{v583}$ and pDLFm*aEfav*$_{583}$ into JH2-2 cells harbouring plasmid pAST-*Pma*. However, we failed to obtain strain JH2-2/pAST-*Pma*/pDLS*maEfa* transformants. Thus, we continued our study with strains JH2-2/pAST-*Pma*/pDLF (negative control) and JH2-2/pAST-*Pma*/pDLF*maEfa*. Plasmid pAST-*Pma* carries the *Pma*::*gfp* fusion gene (see Results, Section 2.2). Fluorescence assays were carried out to evaluate the activity of the *Pma* promoter (*gfp* expression). As shown in Fig. 29C, the intensity of fluorescence was similar in both strains. Moreover, it was similar to the fluorescence of JH2-2 cells harbouring only plasmid pAST-*Pma*. Hence, the activity of the *Pma* promoter was not affected by the overproduction of MA*Efa*.

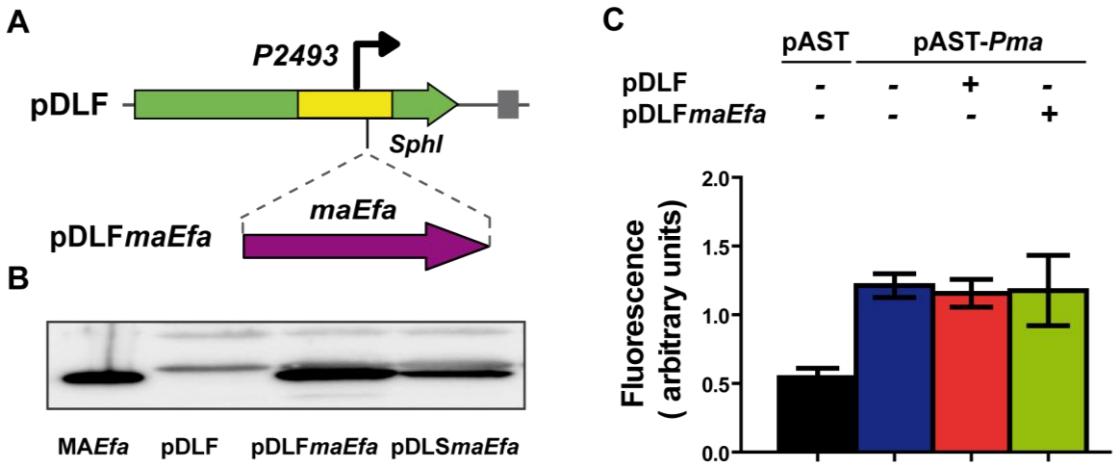


Figure 29. Overproduction of MA*Efa* in *E. faecalis* JH2-2 cells and its effect on the activity of the *Pma* promoter. (A) Construction of the pDLF*maEfa*$_{v583}$ plasmid. The promoter-less *maEfa* gene was inserted into the *Sph*I site of the expression vector pDLF (see Results, Section 1.3), generating pDLF*maEfa*. **(B)** Detection of MA*Efa* by Western blotting using polyclonal antibodies against MA*Efa*. Equivalent amounts of total proteins from JH2-2 cells carrying pDLF, pDLFm*aEfa* and pDLS*maEfa* were separated by SDS-PAGE (12%). MA*Efa* and pre-stained proteins (Invitrogen) were run in the same gel **(C)** The intensity of fluorescence in JH2-2 cells harbouring pAST (black); pAST-*Pma* (blue); pAST-*Pma* and pDLF (pink); or pAST-*Pma* and pDLF*maEfa* (green). Cells were exponentially grown to an OD_{650}=0.3 and the fluorescence (arbitrary units) corresponds to 0.8 ml of culture. Each bar represents the mean ± SD of three independent experiments.

CHAPTER 5

Effect of MA*Efa* on global gene expression and virulence

In this chapter, we show that MA*Efa* influences positively the transcription of many genes by genome-wide microarrays, quantitative RT-PCR assays, and complementation studies. The most significant target genes encode components of PTS-type membrane transporters, components of ABC-type membrane transporters, and proteins involved in the metabolism of carbon sources. Furthermore, we show that a MA*Efa*-lacking strain induces a significant lower degree of inflammation in the peritoneal cavity of mice, indicating that enterococcal cells deficient in MA*Efa* are less virulent.

5.1. Deletion of the *maEfa* gene in strain OG1RF

The Mga protein of *S. pyogenes* regulates, directly or indirectly, the expression of about 10% of the genome during exponential growth (Ribardo and McIver, 2006). The Mga-regulated genes are mainly involved in virulence, carbohydrate metabolism and other metabolic processes (Ribardo and McIver, 2006). The *B. anthracis* AtxA protein controls the expression of more than a hundred genes, including the anthrax toxin genes, the capsule synthesis genes and genes implicated in transcriptional regulation and signaling (Fouet, 2010). To investigate whether MA*Efa* functions as a global transcriptional regulator, we constructed an OG1RF derivative, named OG1RFΔ*maEfa,* which is not able to synthesize MA*Efa*. We selected the OG1RF strain for this study because (i) its genome has been totally sequenced (Bourgogne *et al*., 2008); (ii) it does not carry plasmids, (iii) it is readily transformable by electroporation, and (iv) it is sensitive to common antibiotics. These features facilitate its genetic manipulation. For the construction of the OG1RFΔ*maEfa* strain (see details in Methods, Section 3.1), the chromosomal region between coordinates 2421575 and 2422640 was deleted (coordinates 2889057 and 2890122 in the V583 genome) (see Fig. 30). This deletion removes the -10 element of the *Pma* promoter and the first 344 codons of the *maEfa* gene but not the IR element located upstream of *Pma* (transcriptional terminator signal), as it was confirmed by DNA sequencing. Moreover, by RT-PCR assays, we confirmed the presence of *maEfa* transcripts in strain OG1RF but not in OG1RFΔ*maEfa* under standard bacterial growth conditions.

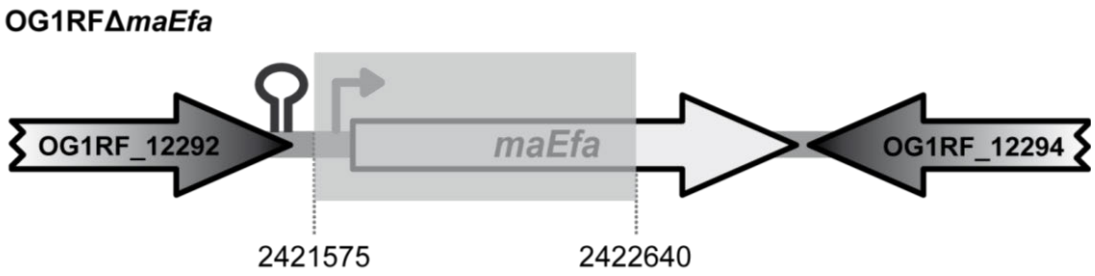


Figure 30. Strain OG1RFΔ*maEfa*. It lacks the chromosomal between coordinates 2421575 an 2422640. The -10 hexamer of the *Pma* promoter and the first 344 codons of the *maEfa* gene have been removed, but not the transcriptional terminator of OG1RF_12292. OG1RFΔ*maEfa* is not able to synthesize MA*Efa*.

5.2. MA*Efa* influences transcription of numerous enterococcal genes

By genome-wide microarrays, we obtained the transcriptional profiles of strains OG1RF and OG1RFΔ*maEfa.* The OG1RF genome has 2,658 predicted genes (Bourgogne *et al*., 2008) and the total number of genes represented in the array was 2,629. Bacteria were grown to mid-log phase under standard laboratory conditions. In MA*Efa*-lacking cells, 90 genes were significantly differentially expressed (P value < 0.05): 87 genes were down-regulated (log_2FC of mutant strain *versus* wild-type strain was lower than -3) (Annexes, Table 1) and three genes were up-regulated (log_2FC of mutant strain *versus* wild-type strain was about 4). The latter genes were OG1RF_10454 (*fruA*, PTS family fructose porter, IIABC component), OG1RF_10455 (*fruk2*, 1-phosphofructokinase) and OG1RF_10456 (*lacR*, lactose PTS family porter repressor), which constitute an operon (*lacR-fruk2-fruA*). Thus, MA*Efa* influences negatively the transcription of such an operon. Among the 87 down-regulated genes, 15 genes encode components of PTS-type membrane transporters and nine genes encode components of ABC-type membrane transporters (Annexes, Table 1). Moreover, and according to KEGG (Kyoto Encyclopedia of Genes and Genomes) annotations, 18 genes encode enzymes involved in carbon source metabolism.

Our microarray analysis showed that MA*Efa* influences positively the expression of numerous PTS genes. In the absence of MA*Efa*, the highest reduction in gene expression corresponded to *mtlA2*, *mltF2* and *mtlD* (OG1RF_10296-98) (log_2FC about -6). These three genes constitute an operon involved in mannitol utilization (Table 6). Bacteria usually take up mannitol via the mannitol PTS transporter (*mtlA2* and *mltF2*), which phosphorylates mannitol to mannitol-1-phosphate during the transport process. Mannitol-1-phosphate dehydrogenase (*mtlD*) converts intracellular mannitol-1-phosphate into fructose-6-phosphate, which is then metabolized via

glycolysis. A high reduction in gene expression (log_2FC about -5) was also observed in a gene that encodes the EIIBC component of a trehalose PTS transporter (OG1RF_11753; *treB*). Among the mannose family of PTS transporters, *E. faecalis* encodes a gluconate specific EII system, which consists of four components (Zúñiga *et al.*, 2005). It is part of a predicted metabolic pathway that consists of two operons, (OG1RF_12399-97 and OG1RF_12405-00). Such a pathway facilitates gluconate uptake and catabolism via the mannonate route. Both operons (nine genes in total) were down-regulated in MA*Efa*-lacking cells (log_2FC values between -4.8 and -3.7) (Table 6). The putative operon OG1RF_11616-11, which encodes four components of another mannose-class PTS transporter, was also down-regulated in the absence of MA*Efa*. (log_2FC values between -4.8 and -3.3). Thus, MA*Efa* has a positive effect on the expression of several PTS transporters.

TABLE 6 Operons down-regulated in MA*Efa*-lacking cells

Operon	Description	Gene validated by qRT-PCR
OG1RF_10296-98	Mannitol utilization (PTS transporter)	OG1RF_10298 (*mtlD*), mannitol-1-phosphate 5-dehydrogenase
OG1RF_10683-80	Maltose metabolism	OG1RF_10683 (*malP*), maltose phosphorylase
OG1RF_11003-05	ABC transporter (unknown substrate)	ND
OG1RF_11135-33	Sugar ABC transporter	OG1RF_11135, sugar-binding protein
OG1RF_11146-49	Glycerol metabolism	OG1RF_11146 (*gldA*), NAD^+-dependent glycerol dehydrogenase
OG1RF_11181-88	Molybdenum cofactor biosynthesis (ABC transporter)	ND
OG1RF_11592-90	Glycerol metabolism	OG1RF_11592 (*glpK*), glycerol kinase
OG1RF_11616-11	Mannose-class PTS transporter	ND
OG1RF_11763-61	Carbohydrate ABC transporter	OG1RF_11763, membrane protein
OG1RF_11951-44	Selenocompound metabolism	OG1RF_11948 (*selD*), selenide water dikinase
OG1RF_12399-97	Gluconate utilization (PTS transporter)	OG1RF_12398 (*uxuA*), mannonate dehydratase

Operon	Description	Gene validated by qRT-PCR
OG1RF_12405-00	Gluconate utilization	OG1RF_12405 (*gnd2*), 6-phosphogluconate dehydrogenase
OG1RF_12571-60	Citrate utilization	OG1RF_12564 (*citF*), citrate lyase
OG1RF_12572-73	Citrate utilization	ND

ND: no determined

MA*Efa* influences positively the expression of some ABC transporters (Table 6). Transcription of genes OG1RF_11763-61, which encode components of a carbohydrate ABC transporter, and transcription of the putative operon OG1RF_11135-33 (sugar ABC transporter) were highly reduced in MA*Efa*-lacking cells (log_2FC about -5 and -4, respectively). Furthermore, the OG1RF_11003-05 operon (ABC transporter of unknown substrate; log_2FC about -3) and genes OG1RF_11186-88 (molybdenum substrate) were also down-regulated. The latter genes are part of an eight-gene operon (OG1RF_11181-88) that includes genes involved in molybdenum cofactor biosynthesis. Seven genes of such an operon were differentially expressed in the absence of MA*Efa* (log_2FC between -2.5 and -3.7).

Numerous genes involved in carbon source utilization were found to be down-regulated in MA*Efa*-lacking cells. Among them, genes implicated in the utilization of maltose, citrate and glycerol were found. The *malPBMR* operon (OG1RF_10683-80) and its neighbouring OG1RF_10684 gene (*malT*, PTS transporter) are divergently transcribed and were shown to be involved in maltose utilization (Le Breton *et al.*, 2005). In the absence of the MA*Efa* regulator, *malP* (OG1RF_10683), *malB* (OG1RF_10682) and *malT* (OG1RF_10684) were down-regulated (log_2FC about -4). The *malPBMR* operon was originally designated *bopABCD* (biofilm on plastic surfaces) because it was found to influence biofilm formation (Hufnagel *et al*., 2004b). Genes OG1RF_11951 (selenium-dependent molybdenum hydroxylase) and OG1RF_11948 (*selD*; selenide, water dikinase) were also down-regulated in the absence of MA*Efa* (log_2FC about -3). The latter genes form part of a putative eight-gene operon (OG1RF_11951-44 operon) also reported to contribute positively to biofilm formation (Ballering *et al*., 2009; Srivastava *et al*., 2011).

E. faecalis is able to use citrate as the sole source of carbon and energy. Citrate metabolism has been extensively investigated in this bacterium (Blancato *et al*., 2008; Repizo *et al*., 2013; Suárez *et al*., 2011). The *cit* locus is constituted by two divergent

operons: *citHO* and *oadHDB-citCDEFX-oadA-citMG*. Our microarray analysis revealed that the *cit* locus (OG1RF_12572-73 and OG1RF_12571-60) was down-regulated in the absence of the MA*Efa* regulator. With the exception of *citO* (OG1RF_12573; log_2FC -1.66) and *citG* (OG1RF_12560; log_2FC -2.46), the log_2FC value of such genes was between -5.38 (*oadH*, OG1RF_12571) and -3.07 (*citM*, OG1RF_12561) (Annexes, Table 1). Therefore, MA*Efa* has a positive effect on the transcription of the two operons involved in citrate utilization.

In *E. faecalis*, as in many Gram-positive bacteria, glycerol is an essential precursor for the synthesis of lipids and (lipo)teichoic acids. Glycerol can also be a carbon/energy source for several pathogenic bacteria (Ramsey *et al.*, 2014). Two pathways for glycerol catabolism are present in *E. faecalis*. In the GlpO/GlpK pathway, glycerol is first phosphorylated by glycerol kinase (*glpK*) and then oxidized by glycerol-3-phosphate oxidase (*glpO*). In the GldA/DhaK pathway, glycerol is first oxidized by a NAD^+-dependent glycerol dehydrogenase (*gldA1*) and then phosphorylated by dihydroxyacetone kinase (*dhaK*). The final product in both pathways is dihydroxyacetone phosphate. The *glpK, glpO* and *glpF* genes constitute an operon (OG1RF_11592-90). Gene *glpF* encodes a membrane protein involved in glycerol transport. Our microarray analysis showed that expression of the *glpKOF* operon was down-regulated in MA*Efa*-lacking cells (log_2FC about -3). A higher reduction in gene expression was also observed in the OG1RF_11146-49 operon, which includes genes *gldA* and *dhaK* (log_2FC about -5).

E. faecalis encodes several putative glycosidases but little is known about their biological function. The endo-β-*N*-acetylglucosaminidase EndoE (OG1RF_10107) was shown to cleave the N-linked glycans of both the human immunoglobulin G (Giard *et al.*, 2006) and the human glycoprotein lactoferrin (Riboulet-Bisson *et al.*, 2009), which is considered to be a component of innate immunity. Gene OG1RF_12167 also encodes an endo-β-*N*-acetylglucosaminidase activity (Ballering *et al.*, 2009). Our microarray analysis revealed that transcription of both genes (OG1RF_10107 and OG1RF_12167) was highly reduced in the absence of MA*Efa* (log_2FC about -4)

5.3. Validation of the microarray results by qRT-PCR.

To validate the microarray results, we performed qRT-PCR assays using the comparative C_T method (Schmittgen and Livak, 2008). Specifically, we selected 12 potential MA*Efa* target genes involved in metabolism (*mtlD*, *malP*, *gldA*, *glpK*, *selD*, *uxuA*, *gnd2* and *citF*), regulation (*lacR*) or transport (*malT*, OG1RF_11135,

OG1RF_11763) (Table 6). These genes are transcribed from different promoters. We determined their relative expression in OG1RF and OG1RFΔ*maEfa* using *recA* as internal control gene (Fig. 31). Except for *lacR*, gene expression was reduced in strain OG1RFΔ*maEfa* compared to OG1RF. In all the cases, the fold change in gene expression (log_2FC) due to the lack of MA*Efa* (OG1RFΔ*maEfa* versus OG1RF) was comparable to the value obtained in the microarray experiments (Fig. 31). Additionally, we performed qRT-PCR assays using total RNA isolated from strains JH2-2 and JH2-2Δ*maEfa*. In this case, we determined the relative expression of genes *lacR*, *mtlD*, *gldA* and *citF*. The results obtained confirmed that MA*Efa* has a negative effect on the transcription of *lacR* but a positive effect on the transcription of *mtlD*, *gldA* and *citF* (not shown).

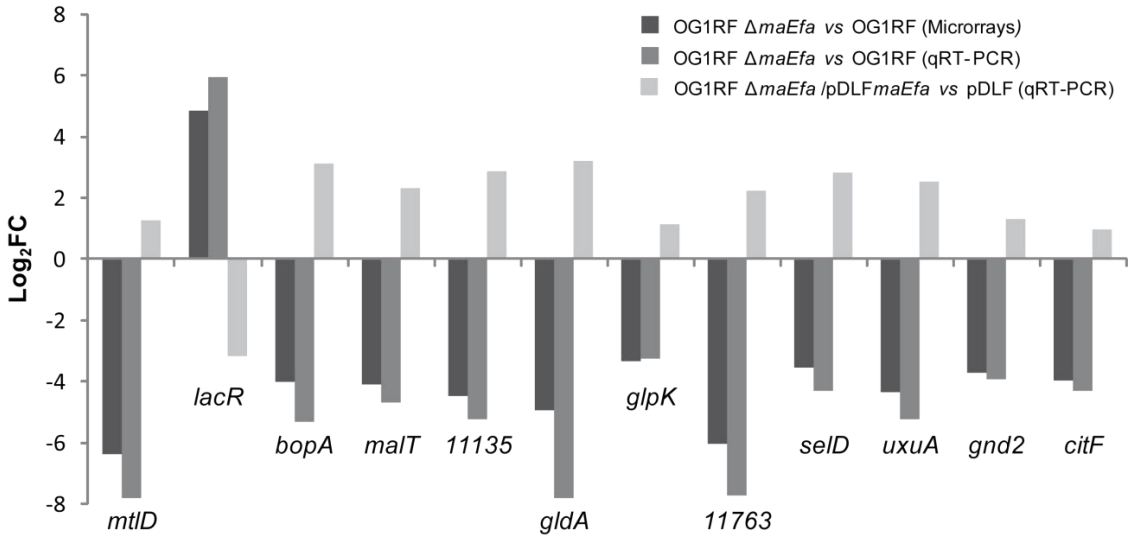


Figure 31. Effect of MA*Efa* on gene expression. Log_2 fold change in the expression of the indicated genes due to the lack of MA*Efa* (OG1RFΔ*maEfa* *vs* OG1RF) determined by microarray analysis and by qRT-PCR. Log_2 fold change in gene expression due to the presence of the MA*Efa*-encoding plasmid (OG1RFΔ*maEfa*/pDLF$maEfa_{OG1RF}$ *vs* OG1RFΔ*maEfa*/pDLF).

5.4. Genetic complementation studies in strain OG1RFΔ*maEfa*

For genetic complementation of the *maEfa* deletion, we inserted the *maEfa* gene of OG1RF ($maEfa_{OG1RF}$) into the pDLF expression vector (see Results Section 1.3), and introduced the pDLF$maEfa_{OG1RF}$ recombinant plasmid into strain OG1RFΔ*maEfa*. By qRT-PCR, we determined the relative expression of the 12 genes mentioned above in strain OG1RFΔ*maEfa* harbouring either pDLF$maEfa_{OG1RF}$ (plasmid-encoded MA*Efa*) or pDLF (absence of MA*Efa*). The fold change in gene expression (log_2FC) due to the plasmid-encoded MA*Efa* is shown in Figure 31.

Compared to cells harbouring pDLF, the relative expression of *lacR* was reduced in cells harbouring pDLF*maEfa*$_{OG1RF}$. On the contrary, the relative expression of the other 11 genes was increased to a different extent. Hence, plasmid-encoded MA*Efa* influences positively the transcription of the 11 genes in the *maEfa* deletion mutant strain.

5.5. Potential curvatures in promoter regions of MA*Efa*-regulated operons

MA*Efa* is thought to be a member of the Mga/AtxA family of global response regulators. These regulators appear to bind DNA with low sequence specificity. *In vitro* studies showed that Mga binds to regions located upstream of the target promoters. Sequence alignments of all established Mga-binding regions revealed that they exhibit only 13.4% identity (Hause and McIver, 2012). In the case of AtxA, *in silico* and *in vitro* analyses revealed that the promoter regions of several target genes are intrinsically curved (Hadjifrangiskou and Koehler, 2008). Furthermore, *in vitro* studies showed that Mga*Spn* interacts with DNA sites that contain a potential intrinsic curvature. (Solano-Collado *et al*., 2012). Recently, Innocenti *et al.* (2015) mapped numerous transcription start sites on the V583 genome using a modified RNA-seq approach. Their study allowed us to locate the transcription star sites of six MA*Efa*-regulated operons on the OG1RF genome. The first gene of each operon and its transcription start site are listed in Table 7. We calculated the curvature propensity plots of the corresponding promoter regions (positions -1 to -200) using the bend.it program (Vlahovicek *et al.*, 2003). Compared with the global A+T content (62.2%) of the OG1RF genome, five out of the six promoter regions display a high A+T content (74-77%). Moreover, they contain one peak of potential sequence-dependent curvature with a magnitude higher than 14º/helical turn. The peak of such a curvature is located at position -171 in gene *glpK*, and between positions -63 and -86 in genes *gldA*, OG1RF_11763, *gnd2* and OG1RF_12571 (Table 7). Further work is required to elucidate whether MA*Efa* is able to bind to DNA containing such potential intrinsic curvatures.

TABLE 7. Potential intrinsic curvatures within promoter regions

Locus tag (gene)	**TSS position** [a]	**A+T** [b] (%)	**Curvature** [c] (º/helical turn)	**Location** [d]
OG1RF_10296 (*mtlA2*)	310152	64.5	12.02	-40
OG1RF_11146 (*gldA*)	1197479	75.5	13.16 16.16	-179 -63
OG1RF_11592 (*glpK*)	1660384	77	17.35	-171
OG1RF_11763	1846229	74.5	14.64 15.01	-86 -78
OG1RF_12405 (*gnd2*)	2542744	74	12.92 14.27	-136 -76
OG1RF_12571	2733071	76.5	14.83 13.71	-85 -24

[a] Position of the transcription start site on the OG1RF genome
[b] A+T content of the promoter region (from position -1 to position -200)
[c] Predicted curvatures with a magnitude >12 degrees/10.5 bp helical turn
[d] Location of peaks of potential curvatures within the promoter regions (relative to the TSS)

5.6. MA*Efa* plays a positive role in the utilization of different carbon sources

Two operons for glycerol metabolism are present in *E. faecalis*. Both of them were down-regulated in MA*Efa*-lacking cells (Table 6, Fig. 31). To evaluate the effect of MA*Efa* on the utilization of glycerol, we analysed the growth of strains OG1RF and OG1RFΔ*maEfa* in M9YE medium supplemented or not with glycerol (Fig. 32A). In the absence of glycerol, both cultures reached a similar OD_{600} (about 0.2). However, in glycerol-supplemented medium, the final OD_{600} in cultures of the mutant strain was lower than in cultures of the wild-type strain. Furthermore, compared to the wild-type strain, the MA*Efa*-lacking strain showed a diminished growth in M9YE medium supplemented with maltose (Fig. 32B), which correlates with the lower expression found in genes *malP*, *malB* and *malT* (maltose utilization) (Table 6, Fig. 31). The growth of the mutant strain was also reduced in M9YE medium supplemented with mannitol (Fig. 32C), which is linked to the reduction in expression of the OG1RF_10296-98 operon (mannitol utilization) (Table 6, Fig. 31). Therefore, MA*Efa* has a positive effect on the utilization of glycerol, maltose and mannitol.

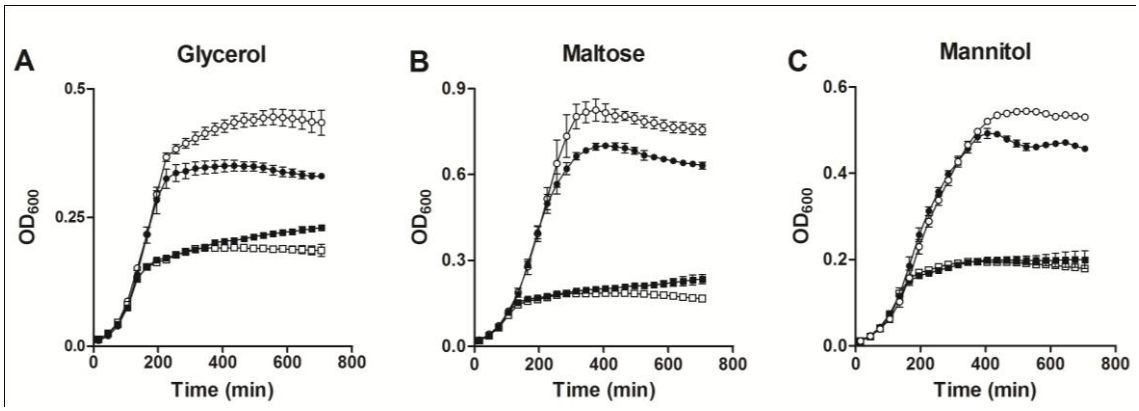


Figure 32. Effect of MA*Efa* on the utilization of different carbon sources. Bacteria were grown in M9YE medium supplemented (circles) or not (squares) with glycerol **(A**), maltose **(B)** or mannitol **(C)**. White symbols correspond to strain OG1RF. Black symbols correspond to strain OG1RFΔ*maEfa*. Data correspond to the mean values of three independent experiments. Error bars indicate standard error.

5.7. Contribution of MA*Efa* to the virulence of *E. faecalis*

In order to study whether the MA*Efa* protein contributes to the virulence of *E. faecalis,* we used first an *in vitro* infection model. We compared the adhesion capacity of strains OG1RF/pMV158GFP (wild-type) and OG1RFΔ*maEfa*/pMV158GFP (mutant) to the human epithelial cell line HEp-2. Both strains carry the *gfp* gene inserted into plasmid pMV158GFP and, therefore, they synthesize constitutively the green fluorescent protein (Nieto and Espinosa, 2003). Epithelial cells were infected with bacteria and, after different incubation times, the number of cells with bacteria attached was determined (fluorescence microscopy). No significant differences were observed between both bacterial strains. Representative images taken after 2 and 24 h of infection are shown in Figure 33.

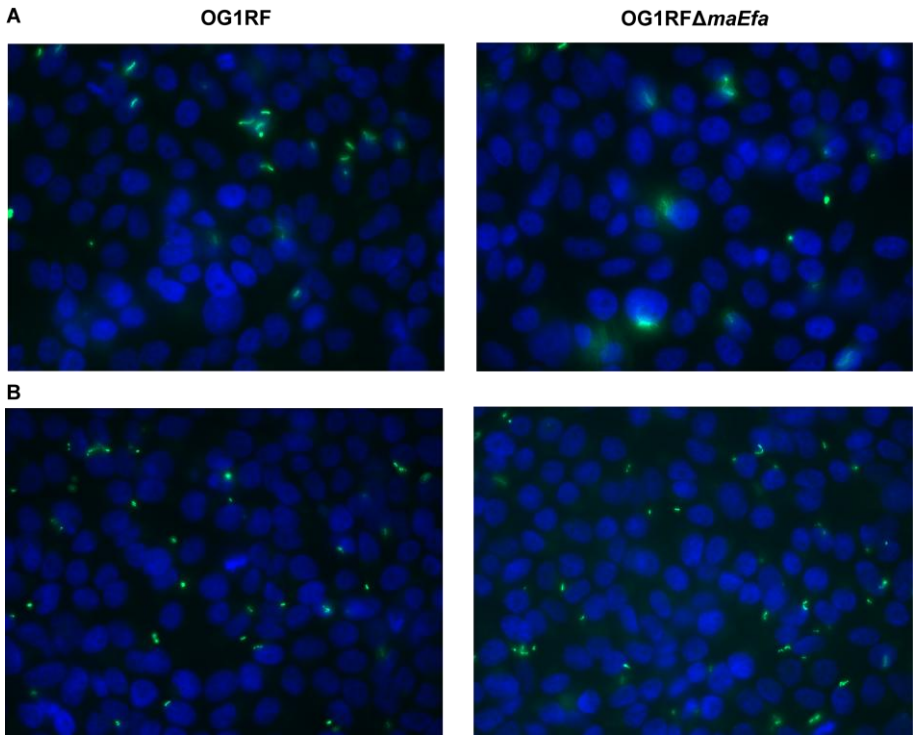


Figure 33. Attachment of enterococcal bacteria to human epithelial cells *in vitro* (fluorescence microscopy). The epithelial cells were infected with OG1RF/pMV158GFP or OG1RFΔ*maEfa*/pMV158GFP for 2 h **(A)** and 24 h **(B).** Bacteria synthesize constitutively the green fluorescent protein. The epithelial cells nuclei are stained by DAPI (blue). The number of cells with bacteria attached was determined in each case.

To investigate the influence of the MA*Efa* regulator *in vivo*, a peritonitis infection model was used comparing the *maEfa* deletion mutant strain (OG1RFΔ*maEfa*) and the corresponding isogenic wild-type strain (OG1RF). Our results indicate a role of the MA*Efa* regulator in bacterial virulence. A significant higher degree of inflammation could be obtained in the peritoneal cavity of mice infected with the wild-type bacteria compared to the *maEfa* deletion mutant. This finding was indicated on the cytokine level by a higher release of IL-6 (Fig. 34A) an important systemic inflammatory cytokine, and on the cellular basis by higher numbers of infiltrating neutrophils (Fig. 34B) to the side of infection. It is important to state that even though the number of infiltrating neutrophils was 2.5 times higher when mice were infected with the wild-type bacteria compared to the mutant strain, no significant reduction in the bacterial loads within the peritoneal cavity could be observed (data not shown); this also highlights the role of the MA*Efa* regulator in bacterial virulence. Taken together, our observations clearly demonstrate an *in vivo* role of the MA*Efa* regulator in the course of infections caused by *E. faecalis*.

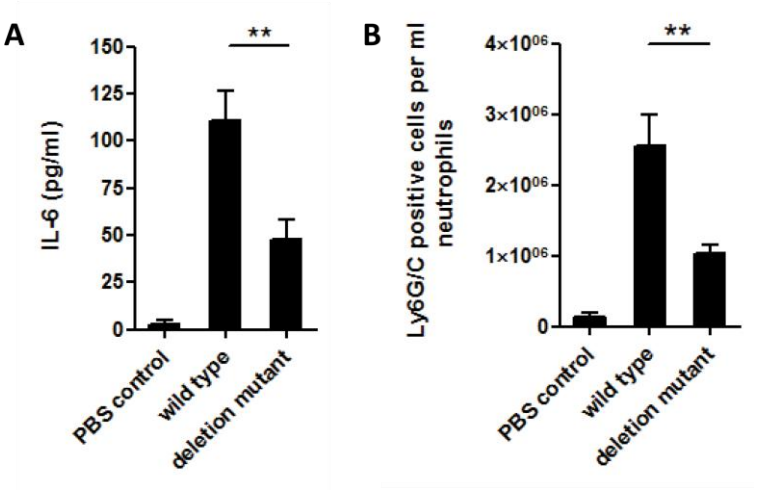


Figure 34. Determination of IL-6 levels and number of infiltrating neutrophils. Levels of IL-6 **(A)** and number of infiltrating LyG6C positive cells **(B)** in the peritoneal cavity of BALB/c mice after intraperitoneal infection with *E. faecalis* OG1RF wild type and OG1RF∆*maEfa* mutant strain. PBS treated animals served as negative controls. Each bar represents the mean ± SD of three independent experiments. **,$p < 0.05$.

DISCUSSION

New plasmid-based genetic tools in *E. faecalis*

Identification of promoters on the bacterial genomes is essential to understand the regulation of gene expression. In bacteria, many genes are organized in operons and, therefore, they are transcribed from the same promoter into a single polycistronic mRNA molecule. Promoters can be located at intergenic regions or within the coding region of adjacent genes. Although many algorithms have been developed for the prediction of promoter sequences in genomic DNAs (de Jong *et al*., 2012; Lin *et al*., 2014; Song, 2012), definitive identification of promoters requires the use of several experimental approaches, both *in vivo* and *in vitro* (Ross and Gourse, 2009). Among the *in vivo* approaches, the use of promoter-probe plasmid vectors facilitates the identification of sequence elements that are essential for promoter activity. Moreover, they make possible to rapidly examine whether a particular promoter is active in different genetic backgrounds and under diverse environmental stimuli. Promoter-probe plasmid vectors are particularly necessary when dealing with bacterial genomes that have a high A+T content, as it is the case of *E. faecalis* and *S. pneumoniae,* whose genomes have about 60% of A+T content. Sequences resembling the consensus sequence of the promoter element -10 (5´-TATAAT-3´) are frequent in both genomes. Such an element is present in promoters recognized by RNAPs that carry a housekeeping σ factor similar to the *E. coli* σ^{70} factor. Despite this fact, plasmid-based genetic tools are still very scarce in *E. faecalis* and *S. pneumoniae.* In this Thesis, we have constructed the pAST promoter-probe vector, and we have demonstrated that it can be used to assess the activity of specific promoter sequences (homologous and heterologous) not only in *E. faecalis* but also in *S. pneumoniae.* Among the promoters tested, we have shown that two predicted promoters, the pneumococcal *Pung* promoter and the enterococcal *P2962* promoter, are functional in both bacteria. In addition, the pAST promoter-probe vector has allowed us to identify the transcription start site of the *maEfa* regulatory gene. This site was not identified in the study performed by Innocenti *et al.* (2015) using a modified RNA-seq approach (tagRNA-seq), which highlights the usefulness of pAST for the identification of promoters. Up to now, plasmid pAST has been successfully used in the characterization of promoters activated by the pneumococcal Mga*Spn* regulator (Solano-Collado *et al*., 2012), and in studies of the pneumococcal *yefM-yoeB* toxin-antitoxin system (Chan *et al*., 2014). The pAST promoter-probe vector is based on the *S. agalactiae* pMV158 plasmid, which is one of the most promiscuous replicons reported so far. It has been established in nearly 30 different bacterial species (M. Espinosa, unpublished observations). Hence, it is very likely that pAST can be used in a number of different Gram-positive bacteria.

Plasmid pAST is also a valuable tool for the study of regulated promoters, and for the identification of bacterial growth conditions that influence the activity of a particular promoter. This conclusion is supported by our results concerning the study of the pneumococcal *PfcsK* promoter, which is induced by fucose (Chan *et al*., 2003). Specifically, we constructed plasmid pAST-*PfcsK* and demonstrated that the activity of the *PfcsK* promoter depends on the concentration of fucose added to the bacterial growth medium. These experiments were performed with the *S. pneumoniae* strain 708, and we found that the *PfcsK* promoter was not totally repressed in the absence of fucose. This finding indicates that a single chromosomal copy of the putative fucose regulator gene (*fcsR*) is not sufficient for total repression of the *PfcsK* promoter placed on a pLS1 derivative (pLS1 has ~22 copies per genome equivalent) (del Solar *et al.*, 1993). Nevertheless, since gene *fcsR* is widely conserved in *S. pneumoniae* (Chan *et al.*, 2003), we believe that plasmid pAST-*PfcsK* could be used as an inducible expression vector in pneumococcus. The promoter-less gene of interest would be inserted into the multiple cloning site located between the *PfcsK* promoter and the *gfp* reporter gene. Compared to bacteria grown without fucose, we found that the activity of the *PfcsK* promoter increases about 5-fold (fluorescence assays) when the bacterial medium contains 1% fucose. Therefore, this expression system would be useful when cells expressing different levels of a particular gene are needed.

The studies performed with the pAST promoter-probe vector allowed us to design new plasmid-based expression vectors suitable for *E. faecalis*. Expression vectors are useful to carry out genetic complementation studies or to assess the effect of transcriptional regulators on specific promoters. In order to develop new plasmid-based expression vectors, we selected the pneumococcal *PsulA* promoter and the enterococcal *P2493* promoter, which are functional not only in *E. faecalis* but also in *S. pneumoniae* (this work). In *E. faecalis*, promoter *P2493* is stronger than promoter *PsulA*. Both promoters were inserted into plasmid pDL287 (replicon pVA380-1), which is compatible with the pAST promoter probe vector (replicon pMV158). Thus, vector pDLF (based on promoter *P2493*) and vector pDLS (based on promoter *PsulA*) make possible to analyse the effect of particular regulatory genes on promoters cloned into pAST. In our case, the *maEfa* gene was cloned into both expression vectors, and enterococcal cells producing different levels of the MA*Efa* protein were obtained. It is likely that both pDLF and pDLS could be used in other Gram-positive bacteria, since they are based on plasmid pDL287, which is a broad host-range plasmid (LeBlanc *et al.*, 1993). In fact, the *mgaSpn* regulatory gene was fused to the *PsulA* promoter, and the fusion gene was inserted into pDL287. Pneumococcal cells carrying the

recombinant plasmid were shown to overproduce Mga*Spn* (Solano-Collado *et al.*, 2012). In *S. pneumoniae*, promoter *PsulA* is stronger than promoter *P2493* (this work). The pDLF and pDLS expression vectors have only a single restriction site (*Sph*I) downstream of *P2493* and *PsulA*, respectively. Insertion of a multiple cloning site into the *Sph*I would give more versatility to both expression vectors.

Bacterial gene expression can be controlled at the transcription termination process. The RNAP can terminate transcription efficiently at Rho-independent signals, which are active in the nascent transcript. The Rho-independent terminators (also known as intrinsic terminators) typically consist of a G:C rich stem–loop structure, followed by a short stretch of U residues. The stem–loop structure halts the RNAP and leads to its release. Thus, transcription termination occurs near the end of the poly(U) region. Transcription attenuation is a highly conserved regulatory mechanism used by bacteria. Attenuators are usually located at the 5′ untranslated regions of genes or operons and combine an intrinsic terminator with an RNA element that senses specific environmental stimuli (Merino and Yanofsky, 2005; Naville and Gautheret, 2009). Several algorithms are able to detect intrinsic terminators in genomic DNAs (Mitra et al., 2010; Naville *et al*., 2011). Nevertheless, some intrinsic terminating sequences deviate from the common motif and, consequently, the availability of terminator-probe plasmid vectors makes possible to rapidly test whether a particular sequence functions as a terminator. Such genetic tools are still very scarce in *E. faecalis*. In this work, we have constructed the pAS terminator-probe vector, and demonstrated that it can be used to examine whether particular sequences (homologous or heterologous) function as transcriptional terminators in *E. faecalis* and *S. pneumoniae*. Moreover, we have shown that the predicted *TrsiV* terminator of *E. faecalis* is active in both bacteria. In our system, such terminator is as efficient as the tandem terminators *T1* and *T2* of *E. coli*, which have been used frequently in the construction of plasmid vectors (Brosius, 1984; Serrano-Heras *et al*., 2005; Simons, 1987). The *TrsiV* transcriptional terminator has been used recently in our laboratory to improve the features of the pAST promoter-probe vector. Specifically, the *TrsiV* terminator has been inserted downstream of the *gfp* reporter gene, and the new pAST derivative has been shown to be useful to study strong promoters (unpublished results).

Expression of the *maEfa* gene in *E. faecalis*

By RT-PCR, we have demonstrated that *maEfa* is transcribed in three different strains of *E. faecalis* (V583, OG1RF and JH2-2) under standard laboratory conditions (BHI medium, 37ºC, without aeration). Moreover, we have shown that genes EF3012 and *maEfa* (EF3013) are not cotranscribed. The *maEfa* gene is also transcribed in *E. faecalis* OG1RF cells grown in the laboratory under aeration conditions, and during *in vitro* macrophage infection (not shown). Under standard conditions, we found that the amount of *maEfa* transcripts in V583 cells is low, as we were unable to detect a primer extension cDNA product using total RNA preparations. In agreement with this, we could not detect MA*Efa* in whole-cell extracts by Western blotting using polyclonal antibodies. These observations suggest that *maEfa* expression might be activated under specific external signals.

In order to identify the transcription start site of the *maEfa* gene, we cloned a region that contained a predicted promoter into the pAST promoter-probe vector. Using total RNA from strain JH2-2 harbouring the recombinant plasmid (pAST-*Pma*), we demonstrated that the predicted promoter (promoter *Pma*) is functional under standard laboratory conditions (primer extension assays). This conclusion was supported by fluorescence assays, as plasmid pAST-*Pma* carries the *gfp* gene under control of the *Pma* promoter. We believe that plasmid pAST-*Pma* could be useful to identify environmental signals that influence the activity of the *Pma* promoter and, therefore, the expression of the *maEfa* gene. Sequence analysis of the intergenic region (91 nucleotides) between EF3012 and *maEfa* revealed the existence of a putative Rho-independent transcriptional terminator upstream of the *Pma* promoter. Using the pAS terminator-probe vector, we have demonstrated that such a terminator sequence is functional, although it is not as efficient as the tandem terminators *T1* and *T2* of *E. coli*. In addition, Merino and Yanofski (2005) predicted that an alternative RNA secondary structure might function as an anti-terminator signal of *maEfa*. If this were the case, transcription initiated upstream of the anti-terminator signal might continue through the transcriptional terminator sequence. This finding suggests that transcriptional attenuation might control the expression of *maEfa* from a promoter located upstream of the *Pma* promoter. According to the BPROM prediction program (*Softberry, Inc.*), there is a putative promoter sequence upstream of the EF3012 gene. Therefore, under certain conditions, genes EF3012 and *maEfa* might be cotranscribed. In the case of the regulatory genes *mga* (*S. pyogenes*) and *atxA* (*B. anthracis*), two promoters have been identified. The major promoter (*P2*) of *mga* is located close to the coding sequence,

whereas the secondary promoter (*P1*) is located about 300-bp upstream of the *P2* promoter (Okada *et al.*, 1993). Regarding *atxA*, two transcription start sites have been identified. They are separated by 650-bp. Moreover, the distal promoter is located within the coding region of a divergently transcribed gene (Bongiorni *et al.*, 2008).

The Mga regulator activates the transcription of its own gene during exponential growth (McIver *et al.*, 1999). Mga fused to the maltose binding protein was shown to bind to two sites located upstream of the *P2* promoter. Such sites are centred around the -104 and -185 positions, respectively (McIver *et al.*, 1999). By DNase I footprinting analysis, we have shown that an untagged form of MA*Efa* recognizes a site located upstream of the *Pma* promoter. This site is centred at the position -90. The DNA fragment used included from the position -140 to the position +87. This finding suggested that binding of MA*Efa* to such a site might control the expression of its own gene. To test this hypothesis, we used plasmid pAST-*Pma*, in which the *gfp* gene is under control of the *Pma* promoter. In this transcriptional fusion, the *Pma* promoter region includes from the position –207 to the position +7. We determined the activity of the *Pma* promoter in JH2-2 cells overproducing MA*Efa*, and found that its activity was independent on the intracellular amount of MA*Efa*. Thus, either MA*Efa* does not influence the activity of the *Pma* promoter *in vivo* or sequence elements that are not present in the *Pma::gfp* transcriptional fusion are required for a MA*Efa*-mediated regulation. Moreover, MA*Efa* could require a factor that is not at the amount needed when MA*Efa* is overproduced. It will be also interesting to investigate the putative role of MA*Efa* in self-regulation using different *E. faecalis* strains. For instance, Mga does not activate the transcription of its own gene in the serotype M4 of GAS (Hondorp *et al*., 2012; Vahling and McIver, 2006). In addition, the global regulator Ers of *E. faecalis* was shown to repress the expression of one of its target genes in strain JH2-2 but not in strain OG1RF (Cohen *et al.*, 2013). Regarding other regulators of the Mga/AtxA family, AtxA does not seem to control the expression of its own gene (Bongiorni *et al.*, 2008), and Mga*Spn* was shown to bind to the promoter of its own gene (positions -23 to +21) (Solano-Collado *et al*., 2013), but it has not yet been determined whether this binding influences the expression of *mgaSpn in vivo*.

MA*Efa* is a new member of the Mga/AtxA family of global regulators

AtxA controls the expression of numerous genes on the *B. anthracis* plasmids and chromosome (Bourgogne *et al.*, 2008). Sequence similarities in the promoter regions of AtxA-regulated genes are not apparent, but *in silico* modeling and *in vitro* experiments revealed that the promoter regions of several target genes are intrinsically curved. Thus, intrinsic curvature might play a role in recognition of the promoters by AtxA (Hadjifrangiskou and Koehler, 2008). The crystal structure of AtxA has revealed two structures indicatives of DNA-binding: a winged helix-turn-helix motif and an adjacent helix-turn-helix motif. Despite the fact that possible models of AtxA-DNA complexes have been suggested (Hammerstrom *et al*., 2015), the molecular basis for AtxA function as a transcriptional regulator has not been established, since there are no reports on DNA binding activity. In the case of the Mga global regulator of *S. pyogenes*, several DNA-binding sites have been identified using fusion proteins (Hondorp and McIver, 2007). Three categories of Mga-regulated promoters were proposed based on the number of binding sites and their position relative to the initiation transcription site (Almengor and McIver, 2004). Nevertheless, comparing the nucleotide sequence of the different Mga binding sites, it was found that they exhibit only 13.4% identity with no discernible symmetry. A further mutagenesis analysis established that Mga binds to DNA in a promoter-specific manner (Hause and McIver, 2012). Regarding the pneumococcal Mga*Spn* regulator, its interaction with linear dsDNA was studied, and two Mga*Spn* binding sites have been identified (sites I and II) (Solano-Collado *et al*., 2013). Site I is located upstream of a target promoter (positions -60 to -90 of the *P1623B* promoter), whereas site II overlaps the promoter of the *mgaSpn* gene (positions -23 to +21). These sites were shown to have a low sequence identity and a potential intrinsic curvature. Moreover, Mga*Spn* showed a high affinity for a naturally occurring curved DNA. It has been suggested that Mga*Spn* might recognize particular DNA conformations to achieve DNA-binding specificity. On binding to the primary site (site I or II), multiple Mga*Spn* units bind orderly along the DNA molecule generating multimeric complexes (Solano-Collado *et al*., 2013).

In this Thesis, we have purified an untagged form of the MA*Efa* protein and shown that it interacts with linear dsDNAs. To our knowledge, neither Mga nor AtxA have been purified without tags. By DNase I footprinting assays and using DNA fragments that contain the *Pma* promoter (promoter of the *maEfa* gene), we have shown that MA*Efa* recognizes preferentially a site that is located upstream of the *Pma* promoter, between the positions -73 and -110. According to the bend.it program

(Vlahovicek *et al*., 2003), this site contains a potential intrinsic curvature, suggesting that MA*Efa*, like other regulators of the Mga/Atxa family, might recognize structural features of the DNA rather than a specific nucleotide sequence. In fact, by EMSA, we found that MA*Efa* binds to various DNA fragments that do not have sequence identity. Furthermore, using the bend.it program, we analysed the promoter region of six operons whose transcription increases in the presence of MA*Efa* (microarray results). Five out of the six promoter regions contain a potential intrinsic curvature located upstream of the -35 element (between positions -63 and -171). The magnitude of such curvatures (>14 degrees/helical turn) is within the range calculated for experimentally tested curved motifs (Gabrielian *et al*., 1997). It will be interesting to analyse whether MA*Efa* recognizes such curvatures to influence the expression of the target genes. For some proteins, recognition of DNA shape is the mechanism used to achieve DNA-binding specificity. This mechanism called *shape readout* is less known than the *base readout* mechanism (when the protein recognizes the unique chemical signatures of the DNA bases), but it has been described for numerous transcription factors (Rohs *et al*., 2010). For instance, the Fis nucleoid-associated protein, which functions as a global regulator and is conserved in most Gram-negative bacteria, was shown to recognize the shape of the minor groove in some DNA targets (Stella *et al*., 2009). With a preference for intrinsically curved AT-rich regions, H-NS is known to function as a global modulator of gene expression (Gruber, 2014). In the Gram-positive bacterium *Staphylococcus aureus*, the global regulator SarA was shown to have a preference for AT-rich binding sites (Sterba *et al*., 2003) and SarA was proposed to recognize topological features of the DNA (Fujimoto *et al*., 2000). Nevertheless, Rohs *et al*. (2010) pointed out that any one DNA-binding protein is likely to use a combination of readout mechanisms to achieve DNA-binding specificity.

By EMSA and using increasing amounts of protein, we found that MA*Efa* is able to generate multimeric complexes on linear double-stranded DNAs. Higher-order MA*Efa*-DNA complexes appeared gradually before the total disappearance of unbound DNA. The pattern of complexes detected in those experiments is compatible with a sequential binding of several MA*Efa* units on the same DNA molecule. This interpretation has been supported by DNase I footprinting assays. Specifically, using DNA fragments that contain the *Pma* promoter, we observed binding of MA*Efa* to a specific site and then, when protein concentration was increased, binding of MA*Efa* to the regions flanking such a site. These results indicate that the DNA-binding properties of MA*Efa* are similar to those described for the pneumococcal Mga*Spn* regulator

(Solano-Collado *et al.*, 2012). Mga*Spn* recognizes a specific site on particular linear DNAs and, on binding to such a site, it is able to spread (presumably through oligomerization) along the adjacent DNA regions(Solano-Collado *et al*., 2012). In solution (without DNA), proteins Mga*Spn* (Solano-Collado *et al*., 2012) and MA*Efa* (this work) form dimers. Formation of multimeric protein-DNA complexes has not been shown for other regulators of the Mga/AtxA family. Nevertheless, AtxA has been reported to form homo-multimers by co-affinity purification, non-denaturing PAGE and cross-linking experiments. In these experiments, several strains producing native and epitope-tagged AtxA were used. In addition, AtxA dimerization was shown to enhance under conditions that promote expression of AtxA-regulated genes (cells grown with elevated CO_2/bicarbonate) (Hammerstrom *et al*., 2011). Later, it was shown that AtxA mutants affected in dimerization are also impaired in transcriptional activity (Hammerstrom *et al*., 2015). Similar to MA*Efa* (PDB 3SQN), AtxA crystallizes as a dimer in the absence of DNA (Hammerstrom *et al.*, 2015). Regarding the Mga regulator, it was shown to form oligomers in solution (Hondorp *et al*., 2012). Its ability to oligomerize was found to correlate with transcriptional activation, thus DNA binding is necessary but not sufficient for full Mga activity (Hondorp *et al*., 2013; Hondorp *et al*., 2012). Hause *et al.* (2012) suggested that Mga interacts with DNA as a dimer. They supported their hypothesis, in addition to their results, with the crystal structure of MA*Efa* (PDB 3SQN). These authors compared the main DNA binding domain of Mga (wHTH-4) to the amino-terminal wHTH domain present in each monomer of the MA*Efa* dimer. They estimated that the distance between the predicted wHTH recognition helices was 95 Å to 100 Å, corresponding to about 30 nucleotides. In this work, we have experimentally determined that the minimum DNA size required for MA*Efa* binding is between 26 and 32 bp. Thus, based on these observations, we propose that MA*Efa* binds to DNA as a dimer.

MA*Efa* might facilitate the adaptation of *E. faecalis* to particular host niches

The adaptation of a pathogenic bacterium to new host niches requires a coordinated regulation in the expression of numerous genes. Global transcriptional regulators that respond to specific signals are essential in the adaptation process. In Gram-positive pathogenic bacteria, various global response regulators have been implicated in virulence. Among them, the regulators of the Mga/AtxA family are found. For instance, the *B. anthracis* regulator AtxA was reported to have both positive and negative effects on gene expression (Bourgogne *et al*., 2003). It controls the

expression of chromosomal genes, in addition to genes located on the virulence plasmids pXO1 (anthrax toxin genes) and pXO2 (capsule synthesis genes). Other AtxA-regulated genes encode secreted proteins and proteins implicated in transcriptional regulation and signaling (Bourgogne *et al.*, 2003). Regarding the pneumococcal Mga*Spn* regulator, microarray experiments showed that it influences negatively the expression of several genes located within the *rlrA* pathogenicity islet (Hemsley *et al*., 2003). Moreover, Mga*Spn* activates directly the expression of a four-gene operon of unknown function (Solano-Collado *et al*., 2012). In *S. pyogenes*, the global regulator Mga is known to regulate, directly or indirectly, the expression of approximately 10% of its genome. It activates the transcription of numerous genes that enable the bacterium to colonize specific host tissues and evade the host immune response (Ribardo and McIver, 2006). Additionally, Mga has a positive or negative effect on the expression of various genes involved in the transport and utilization of sugars and other metabolites (Hondorp and McIver, 2007). In this Thesis, we have demonstrated that the MA*Efa* protein of *E. faecalis,* like the regulators of the Mga/AtxA family, influences the transcription of numerous genes. By genome-wide microarrays, quantitative RT-PCR assays, and complementation studies, we have shown that MA*Efa* increases the expression of many genes that encode components of PTS transporters, components of ABC transporters, and proteins involved in the metabolism of carbon sources.

In bacteria, the PTS catalyzes the transport and phosphorylation of a variety of sugars and sugars derivatives (Deutscher *et al*., 2014). The PTS is usually composed of two general components, enzyme I (EI) and HPr, and the sugar-specific-components, enzyme IIA (EIIA), enzyme IIB (EIIB), enzyme IIC (EIIC). In mannose-type PTS there is an additional component, enzyme IID (EIID). Genes encoding the sugar-specific components are normally organized in an operon, which often contains the genes required for the catabolism of the transported sugar (Deutscher *et al*., 2014). The sugars most commonly metabolized by enterococci are PTS substrates (Ramsey *et al*., 2014). In addition, the PTS carries out regulatory functions (Deutscher *et al*., 2014). In some pathogenic streptococci and *Salmonella*, mutagenesis screenings identified some PTS genes as virulence factors (Hava and Camilli, 2002; Jones *et al*., 2000; Turner *et al*., 1998). Our microarray analysis showed that MA*Efa* influences positively the expression of numerous PTS genes, as *mtlA2* and *mltF2* (mannitol PTS transporter). Both genes and *mtlD* constitute the OG1RF_10296-98 operon. Associated to this result, the growth of the mutant strain (OG1RFΔ*maEfa*) was slightly reduced in M9YE medium supplemented with mannitol (compared to the OG1RF wild-type strain).

Gene *mtlD* was reported to be significantly up-regulated during growth of *E. faecalis* in the intestinal tract of mice (Lindenstraüss *et al.*, 2014). In this study, genes encoding a mannose-class PTS transporter (the gluconate PTS transporter) were also up-regulated (Lindenstraüss *et al.*, 2014). Such genes are located within two operons (nine genes) involved in gluconate utilization (OG1RF_12399-97 and OG1RF_12405-00). Our results showed that MA*Efa* has a positive effect on the expression of both operons. The expression of such operons was also enhanced during the growth of *E. faecalis* in blood (Vebø *et al*., 2009). Furthermore, MA*Efa* has a positive effect on the expression of the putative operon OG1RF_11616-11, which encodes another mannose-class PTS transporter**.** Zúñiga *et al*. (2005) suggested that mannose-class PTS transporters might play a significant role in the adaptation of bacteria to epithelial surfaces. In addition to PTS transporters, MA*Efa* influences positively the expression of several operons encoding ABC transporters. Among them, the OG1RF_11135-33 operon (sugar substrate) was found to be up-regulated during growth of *E. faecalis* in blood (Vebø *et al*., 2009). Therefore, MA*Efa* increases the expression of PTS and ABC transporters that are apparently involved in the metabolic adaptation of *E. faecalis* to various environments.

MA*Efa* has also a positive effect on the transcription of genes involved in the utilization of citrate (OG1RF_12572-73 and OG1RF_12571-60) and maltose (*malP, malB* and *malT*). Such genes were found to be up-regulated during growth of *E. faecalis* in blood (Vebø *et al*., 2009), as well as several genes of both pathways were up-regulated during its growth in human urine (Vebø *et al.*, 2010). In our study, additionally, we observed that the OG1RFΔ*maEfa* strain showed a diminished growth in M9YE medium supplemented with maltose compared to the OG1RF strain. Regarding other carbon sources, MA*Efa* influences positively the transcription of the two operons involved in glycerol metabolism (the GlpO/GlpK and GldA/DhaK pathways), although the activator effect seems to be greater on the genes of the GldA/DhaK pathway. Moreover, we found that the growth of the OG1RFΔma*Efa* strain is impaired in M9YE medium supplemented with glycerol. Vebo *et al.* (2009) found that the *glpKOF* operon (GlpO/GlpK pathway) was highly up-regulated during growth of *E. faecalis* in blood, and they suggested that glycerol and other C3-glycerides from blood serve as a source of energy. Glycerol seems to be also an important substrate for this bacterium in the intestinal tract of mice, as it was proposed by Lindenstraüss *et al*. (2014). They observed that both glycerol catabolic operons were up-regulated during the adaptation of *E. faecalis* to the intestinal tract of mice. Thus, we hypothesize that

MA*Efa* may facilitate the adaptation of *E. faecalis* to particular host niches activating the transcription of genes involved in the utilization of specific carbon sources.

The potential role of MA*Efa* in virulence was studied at the Helmholtz Centre for Infection Research, under the supervision of Dr. Oliver Goldmann (Infection and Immunology Group). In a first approach, we used different *in vitro* infection models, as well as the wild-type strain (OG1RF) and the *maEfa* deletion mutant strain (OG1RFΔ*maEfa*). First, we compared several parameters of bacterial killing by both murine immortalized macrophages and bone-marrow derived macrophages (not shown). Then, by fluorescence microscopic approaches, we compared the adherence of both strains to human epithelial cells. In all cases, no significant differences were found between wild-type and mutant strains. Therefore, we tested a mouse peritonitis model. Using this model, some global regulators were shown to contribute to the pathogenesis of *E. faecalis*, for instance, FsrA. (Qin *et al*., 2000), PerA (Coburn *et al*., 2008) and the PrfA-like regulator Ers (Giard *et al*., 2006). In this Thesis, we have shown that *E. faecalis* cells lacking MA*Efa* are less virulent in a mouse peritonitis model. Compared to the wild-type strain, the *maEfa* deletion mutant strain induced a more moderate inflammation in terms of infiltration of neutrophils and cytokine response (IL-6), but no significant reduction in the bacterial loads within the peritoneal cavity was observed. However, the precise underlying molecular mechanisms of how this regulatory axis is contributing to bacterial virulence need to be elucidated in further studies. Recently, Muller *et al.* (2015) analysed the transcriptome of *E. faecalis* isolated from an infection site (mouse peritonitis). Among the genes differentially expressed, they found that the OG1RF_10454-56 operon (*lacR-fruk2-fruA*) was down-regulated, whereas the OG1RF_11135-33 operon (ABC transporter) and the two operons involved in glycerol catabolism were highly induced. Moreover, Muller *et al*. (2015) found that mutants unable to metabolize glycerol are affected in organ colonization in a systemic murine infection model. Interestingly, MA*Efa* has a negative effect on the expression of the *lacR-fruk2-fruA* operon, but a positive effect on the expression of the other mentioned operons. Thus, some MA*Efa*-regulated genes are differentially expressed during the intraperitoneal infection process. Taken together, our results suggest that MA*Efa* facilitates the adaptation of *E. faecalis* to particular host niches through the transcriptional regulation of numerous genes and, consequently, it contributes to its potential virulence.

CONCLUSIONS

The functional characterization of the *E. faecalis* MA*Efa* protein presented in this Thesis supports that MA*Efa* is a new member of the Mga/AtxA family of global transcriptional regulators. The main conclusions of this work are:

1. We have developed new plasmid-based genetic tools suitable for *E. faecalis* and *S. pneumoniae*: a promoter-probe vector (plasmid pAST), a terminator-probe vector (pAS) and two expression vectors (pDLF and pDLS). All of them are based on broad host-range plasmids.

2. The *maEfa* gene is transcribed *in vivo* from the *Pma* promoter. The transcription start site of this gene is located 15 nucleotides upstream of the predicted translation start codon.

3. Upstream of the *Pma* promoter there is a functional intrinsic transcriptional terminator.

4. We have set up a procedure to purify an untagged form of the MA*Efa* protein. This protein forms dimers in solution.

5. MA*Efa* binds to linear double-stranded DNAs with little or no sequence specificity, generating multimeric DNA-protein complexes.

6. The presence of a His-tag at the C-terminus of MA*Efa* does not affect the formation of multimeric DNA-protein complexes.

7. MA*Efa* recognizes a site located upstream of the *Pma* promoter (positions -73 to -110). This site has a potential intrinsic curvature.

8. MA*Efa* influences positively the expression of numerous genes involved in the utilization of carbon sources, including genes that encode components of PTS and ABC membrane transporters. Associated to this fact, the growth of a *maEfa* deletion mutant strain is impaired in media containing glycerol, maltose or mannitol.

9. *E. faecalis* cells deficient in MA*Efa* are less virulent in a mouse peritonitis model.

REFERENCES

Agudelo Higuita, N., and Huycke, M. (2014). Enterococcal disease, epidemiology, and implications for treatment. In Gilmore MS, Clewell DB, Ike Y, *et al*, editors Enterococci: From Commensals to Leading Causes of Drug Resistant Infection [Internet] (Boston: Massachusetts Eye and Ear Infirmary), pp. 1-27.

Almengor, A.C., and McIver, K.S. (2004). Transcriptional activation of *sclA* by Mga requires a distal binding site in *Streptococcus pyogenes*. *J Bacteriol* **186**, 7847-7857.

Almengor, A.C., Walters, M.S., and McIver, K.S. (2006). Mga is sufficient to activate transcription in vitro of *sof-sfbX* and other Mga-regulated virulence genes in the group A *Streptococcus*. *J Bacteriol* **188**, 2038-2047.

Arias, C.A., and Murray, B.E. (2012). The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Micro* **10**, 266-278.

Ballering, K.S., Kristich, C.J., Grindle, S.M., Oromendia, A., Beattie, D.T., and Dunny, G.M. (2009). Functional genomics of *Enterococcus faecalis*: multiple novel genetic determinants for biofilm formation in the core genome. *J Bacteriol* **191**, 2806-2814.

Benachour, A., Muller, C., Dabrowski-Coton, M., Le Breton, Y., Giard, J., Rincé, A., Auffray, Y., and Hartke, A. (2005). The *Enterococcus faecalis* SigV protein Is an extracytoplasmic function sigma factor contributing to survival following heat, acid, and ethanol treatments. *J Bacteriol* **187**, 1022-1035.

Bisno, A.L., Brito, M.O., and Collins, C.M. (2003). Molecular basis of group A streptococcal virulence. *Lancet Infect Dis* **3**, 191-200.

Biswas, I., Gruss, A., Ehrlich, S.D., and Maguin, E. (1993). High-efficiency gene inactivation and replacement system for gram-positive bacteria. *J Bacteriol* **175**, 3628-3635.

Blancato, V.S., and Magni, C. (2010). A chimeric vector for efficient chromosomal modification in *Enterococcus faecalis* and other lactic acid bacteria. *Lett Applied Microbiol* **50**, 542-546.

Blancato, V.S., Repizo, G.D., Suárez, C.A., and Magni, C. (2008). Transcriptional regulation of the citrate gene cluster of *Enterococcus faecalis* involves the GntR family transcriptional activator CitO. *J Bacteriol* **190**, 7419-7430.

Bogaert, D., de Groot, R., and Hermans, P.W. (2004). *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. *Lancet Infect Dis* **4**, 144-154.

Bongiorni, C., Fukushima, T., Wilson, A.C., Chiang, C., Mansilla, M.C., Hoch, J.A., and Perego, M. (2008). Dual promoters control expression of the *Bacillus anthracis* virulence factor AtxA. *J Bacteriol* **190**, 6483-6492.

Bourgogne, A., Drysdale, M., Hilsenbeck, S.G., Peterson, S.N., and Koehler, T.M. (2003). Global effects of virulence gene regulators in a *Bacillus anthracis* strain with both virulence plasmids. *Infect Immun* **71**, 2736-2743.

Bourgogne, A., Garsin, D.A., Qin, X., Singh, K.V., Sillanpaa, J., Yerrapragada, S., Ding, Y., Dugan-Rocha, S., Buhay, C., Shen, H., Chen G., Williams G., Muzny D., Maadani A., Fox K.A., Gioia J., Chen L., Shang Y., Arias C.A., Nallapareddy S.R., Zhao M., Prakash V.P., Chowdhury S., Jiang H., Gibbs R.A., Murray B.E., Highlander S.K., and Weinstock G.M. (2008). Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF. *Genome Biol* **9**, R110-R110.

Bourgogne, A., Hilsenbeck, S.G., Dunny, G.M., and Murray, B.E. (2006). Comparison of OG1RF and an isogenic *fsrB* deletion mutant by transcriptional analysis: the Fsr system of *Enterococcus faecalis* is more than the activator of gelatinase and serine protease. *J Bacteriol* **188**, 2875-2884.

Brosius, J. (1984). Plasmid vectors for the selection of promoters. *Gene* **27**, 151-160.

Brosius, J., Dull, T.J., Sleeter, D.D., and Noller, H.F. (1981). Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. *J Mol Biol* **148**, 107-127.

Capra, E., J., and Laub, M., T. (2012). Evolution of Two-Component Signal Transduction Systems. *Annu Rev Microbiol* **66**, 325-347.

Carey, J. (1991). Gel retardation. *Methods Enzymol* **208**, 103-117.

Clewell, D., Weaver, K., Dunny, G., Coque, T.M., Francia, M.V., and Hayes, F. (2014). Extrachromosomal and mobile elements in enterococci: transmission, maintenance, and epidemiology. In Gilmore M.S., Clewell D.B., Ike Y., *et al*, editors Enterococci: From Commensals to Leading Causes of Drug Resistant Infection [Internet] (Boston: Massachusetts Eye and Ear Infirmary).

Coburn, P.S., Baghdayan, A.S., Dolan, G.T., and Shankar, N. (2008). An AraC-type transcriptional regulator encoded on the *Enterococcus faecalis* pathogenicity island contributes to pathogenesis and intracellular macrophage survival. *Infect Immun* **76**, 5668-5676.

Cohen, A.L.V., Roh, J.H., Nallapareddy, S.R., Hook, M., and Murray, B.E. (2013). Expression of the collagen adhesin *ace* by *Enterococcus faecalis* strain OG1RF is not repressed by Ers but requires the Ers box. *FEMS Microbiol Lett* **344**, 18-24.

Cole, J.L. (2004). Analysis of heterogeneous interactions. *Methods Enzymol* **384,** 212-232.

Cormack, B.P., Valdivia, R.H., Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33-38.

Cunningham, M.W. (2000). Pathogenesis of group A streptococcal infections. *Clin Microbiol Rev* **13**, 470-511.

Chan, P.F., O'Dwyer, K.M., Palmer, L.M., Ambrad, J.D., Ingraham, K.A., So, C., Lonetto, M.A., Biswas, S., Rosenberg, M., Holmes, D.J., Zalacain, M. (2003). Characterization of a novel fucose-regulated promoter (*PfcsK*) suitable for gene essentiality and antibacterial mode-of-action studies in *Streptococcus pneumoniae*. *J Bacteriol* **185**, 2051-2058.

Chan, W.T., Yeo, C.C., Sadowy, E., and Espinosa, M. (2014). Functional validation of putative toxin-antitoxin genes from the Gram-positive pathogen *Streptococcus pneumoniae*: *phd-doc* is the fourth *bona-fide* operon. *Front Microbiol* **5***:* 677*.*

Dai, Z., and Koehler, T.M. (1997). Regulation of anthrax toxin activator gene (*atxA*) expression in *Bacillus anthracis*: temperature, not CO_2/bicarbonate, affects AtxA synthesis. *Infect Immun* **65**, 2576-2582.

de Jong, A., Pietersma, H., Cordes, M., Kuipers, O., and Kok, J. (2012). PePPER: a webserver for prediction of prokaryote promoter elements and regulons. *BMC Genomics* **13**, 299.

del Solar, G., Giraldo, R., Ruiz-Echevarria, M.J., Espinosa, M., and Diaz-Orejas, R. (1998). Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* **62**, 434-464.

del Solar, G., Kramer, G., Ballester, S., and Espinosa, M. (1993). Replication of the promiscuous plasmid pLS1: a region encompassing the minus origin of replication is associated with stable plasmid inheritance. *Mol Gen Genet* **241**, 97-105.

Deutscher, J., Aké, F.M.D., Derkaoui, M., Zébré, A.C., Cao, T.N., Bouraoui, H., Kentache, T., Mokhtari, A., Milohanic, E., and Joyet, P. (2014). The bacterial phosphoenolpyruvate:carbohydrate phosphotransferase system: regulation by protein phosphorylation and phosphorylation-dependent protein-protein interactions. *Microbiol Mol Biol Rev* **78**, 231-256.

Dower, W.J., Miller, J.F., and Ragsdale, C.W. (1988). High efficiency transformation of *E.coli* by high voltage electroporation. *Nucleic Acids Res* **16**, 6127-6145.

Dunny, G.M., Brown, B.L., and Clewell, D.B. (1978). Induced cell aggregation and mating in *Streptococcus faecalis*: evidence for a bacterial sex pheromone. *Proc Natl Acad Sci USA* **75**, 3479-3483.

Engelbert, M., Mylonakis, E., Ausubel, F.M., Calderwood, S.B., and Gilmore, M.S. (2004). Contribution of gelatinase, serine protease, and *fsr* to the pathogenesis of *Enterococcus faecalis* endophthalmitis. *Infect Immun* **72**, 3628-3633.

Fabretti, F., Theilacker, C., Baldassarri, L., Kaczynski, Z., Kropec, A., Holst, O., and Huebner, J. (2006). Alanine esters of enterococcal lipoteichoic acid play a role in biofilm formation and resistance to antimicrobial peptides. *Infect Immun* **74**, 4164-4171.

Fisher, K., and Phillips, C. (2009). The ecology, epidemiology and virulence of *Enterococcus*. *Microbiology* **155** 1749-1757.

Fouet, A. (2010). AtxA, a *Bacillus anthracis* global virulence regulator. *Res Microbio*, **161**, 735-742.

Fujimoto, D.F., Brunskill, E.W., and Bayles, K.W. (2000). Analysis of genetic elements controlling *Staphylococcus aureus lrgAB* expression: potential role of DNA topology in SarA regulation. *J Bacteriol* **182**, 4822-4828.

Gabrielian, A., Vlahovicek, K., and Pongor, S. (1997). Distribution of sequence-dependent curvature in genomic DNA sequences. *FEBS Lett* **406**, 69-74.

Garsin, D., Frank, K., and Silanpää J.(2014). Pathogenesis and models of enterococcal infection. In Gilmore MS, Clewell DB, Ike Y, *et al*, editors Enterococci: From Commensals to Leading Causes of Drug Resistant Infection [Internet] (Boston: Massachusetts Eye and Ear Infirmary).

Garsin, D.A., Sifri, C.D., Mylonakis, E., Qin, X., Singh, K.V., Murray, B.E., Calderwood, S.B., and Ausubel, F.M. (2001). A simple model host for identifying Gram-positive virulence factors. *Proc Natl Acad Sci USA* **98**, 10892-10897.

Giard, J.C., Riboulet, E., Verneuil, N., Sanguinetti, M., Auffray, Y., and Hartke, A. (2006). Characterization of Ers, a PrfA-like regulator of *Enterococcus faecalis*. *FEMS Immunol Med Microbiol* **46**, 410-418.

Gruber, S. (2014). Multilayer chromosome organization through DNA bending, bridging and extrusion. *Curr Opin Microbiol* **22**, 102-110.

Gruber, T.M., and Gross, C.A. (2003). Multiple sigma subunits and the partition of the transcriptional space. *Annu Rev Microbiol* **57**, 441-466.

Hadjifrangiskou, M., and Koehler, T.M. (2008). Intrinsic curvature associated with the coordinately regulated anthrax toxin gene promoters. *Microbiology* **154**, 2501-2512.

Hammerstrom, T.G., Horton, L.B., Swick, M.C., Joachimiak, A., Osipiuk, J., and Koehler, T.M. (2015). Crystal structure of *Bacillus anthracis* virulence regulator AtxA and effects of phosphorylated histidines on multimerization and activity. *Mol Microbiol* **95**, 426-441.

Hammerstrom, T.G., Roh, J.H., Nikonowicz, E.P., and Koehler, T.M. (2011). *Bacillus anthracis* virulence regulator AtxA: oligomeric state, function and CO_2 -signalling. *Mol Microbiol* **82**, 634-647.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557-580.

Hancock, L., and Perego, M. (2002). Two-Component signal transduction in *Enterococcus faecalis*. *J Bacteriol* **184**, 5819-5825.

Hancock, L.E., and Gilmore, M.S. (2002). The capsular polysaccharide of *Enterococcus faecalis* and its relationship to other polysaccharides in the cell wall. *Proc Natl Acad Sci USA* **99**, 1574-1579.

Hancock, L.E., Shepard, B.D., and Gilmore, M.S. (2003). Molecular analysis of the *Enterococcus faecalis* serotype 2 polysaccharide determinant. *J Bacteriol* **185**, 4393-4401.

Hause, L.L., and McIver, K.S. (2012). Nucleotides critical for the interaction of the *Streptococcus pyogenes* Mga virulence regulator with Mga-regulated promoter sequences. *J Bacteriol* **194**, 4904-4919.

Hava, D.L., and Camilli, A. (2002). Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol Microbiol* **45**, 1389-1406.

Heim, R., Cubitt, A.B., Tsien, R.Y. (1995). Improved green fluorescence. *Nature* **373**, 663-664.

Hemsley, C., Joyce, E., Hava, D.L., Kawale, A., and Camilli, A. (2003). MgrA, an orthologue of Mga, acts as a transcriptional repressor of the genes within the *rlrA* pathogenicity islet in *Streptococcus pneumoniae*. *J Bacteriol* **185**, 6640-6647.

Hendrickx, A.P, Willems, R.J., Bonten, M.J., and van Schaik, W. (2009). LPxTG surface proteins of enterococci. *Trends Microbiol* **17**, 423-430.

Hollenbeck, B.L., and Rice, L.B. (2012). Intrinsic and acquired resistance mechanisms in *Enterococcus*. *Virulence* **3**, 421-569.

Hondorp, E.R., Hou, S.C., Hause, L.L., Gera, K., Lee, C.-E., and McIver, K.S. (2013). PTS phosphorylation of Mga modulates regulon expression and virulence in the group A streptococcus. *Mol Microbiol* **88**, 1176-1193.

Hondorp, E.R., Hou, S.C., Hempstead, A.D., Hause, L.L., Beckett, D.M., and McIver, K.S. (2012). Characterization of the Group A Streptococcus Mga virulence regulator reveals a role for the C-terminal region in oligomerization and transcriptional activation. *Mol Microbiol* **83**, 953-967.

Hondorp, E.R., and McIver, K.S. (2007). The Mga virulence regulon: infection where the grass is greener. *Mol Microbiol* **66**, 1056-1065.

Hufnagel, M., Hancock, L.E., Koch, S., Theilacker, C., Gilmore, M.S., and Huebner, J. (2004a). Serological and genetic diversity of capsular polysaccharides in *Enterococcus faecalis*. *J Clin Microbiol* **42**, 2548-2557.

Hufnagel, M., Koch, S., Creti, R., Baldassarri, L., and Huebner, J. (2004b). A putative sugar-binding transcriptional regulator in a novel gene locus in *Enterococcus faecalis* contributes to production of biofilm and prolonged bacteremia in mice. *J Infect Dis* **189**, 420-430.

Huycke, M.M., and Gilmore, M.S. (1995). Frequency of aggregation substance and cytolysin genes among enterococcal endocarditis isolates. *Plasmid* **34**, 152-156.

Huycke, M.M., Spiegel, C.A., and Gilmore, M.S. (1991). Bacteremia caused by hemolytic, high-level gentamicin-resistant *Enterococcus faecalis*. *Antimicrob Agents Chemother* **35**, 1626-1634.

Ike, Y., Hashimoto, H., and Clewell, D.B. (1984). Hemolysin of *Streptococcus faecalis* subspecies *zymogenes* contributes to virulence in mice. *Infect Immun* **45**, 528-530.

Ike, Y., Hashimoto, H., and Clewell, D.B. (1987). High incidence of hemolysin production by *Enterococcus* (*Streptococcus*) *faecalis* strains associated with human parenteral infections. *J Clin Microbiol* **25**, 1524-1528.

Innocenti, N., Golumbeanu, M., Fouquier d'Hérouël, A., Lacoux, C., Bonnin, R.A., Kennedy, S.P., Wessner, F., Serror, P., Bouloc, P., Repoila, F., Aurell E**.** (2015). Whole-genome mapping of 5' RNA ends in bacteria by tagged sequencing: a comprehensive view in *Enterococcus faecalis*. *RNA* **21**, 1018-1030.

Jacob, A.E., and Hobbs, S.J. (1974). Conjugal transfer of plasmid-borne multiple antibiotic resistance in *Streptococcus faecalis* var. *zymogenes*. *J Bacteriol* **117**, 360-372

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., and Madden, T.L. (2008). NCBI BLAST: a better web interface. *Nucleic Acids Res* **36**, W5-9.

Jones, A.L., Knoll, K.M., and Rubens, C.E. (2000). Identification of *Streptococcus agalactiae* virulence genes in the neonatal rat sepsis model using signature-tagged mutagenesis. *Mol Microbiol* **37**, 1444-1455.

Kelley, L.A., and Sternberg, M.J.E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4**, 363-371.

Kemp, K.D., Singh, K.V., Nallapareddy, S.R., and Murray, B.E. (2007). Relative contributions of *Enterococcus faecalis* OG1RF sortase-encoding genes, *srtA* and *bps* (*srtC*), to biofilm formation and a murine model of urinary tract infection. *Infect Immun* **75**, 5399-5404.

Kreft, B., Marre, R., Schramm, U., and Wirth, R. (1992). Aggregation substance of *Enterococcus faecalis* mediates adhesion to cultured renal tubular cells. *Infect Immun* **60**, 25-30.

Lacks, S., and Greenberg, B. (1977). Complementary specificity of restriction endonucleases of *Diplococcus pneumoniae* with respect to DNA methylation. *J Mol Biol* **114**, 153-168.

Lacks, S.A. (1966). Integration efficiency and genetic recombination in pneumococcal transformation. *Genetics* **53**, 207-235.

Lacks, S.A., Greenberg, B., and Lopez, P. (1995). A cluster of four genes encoding enzymes for five steps in the folate biosynthetic pathway of *Streptococcus pneumoniae*. *J Bacteriol* **177**, 66-74.

Lacks, S.A., Lopez, P., Greenberg, B., and Espinosa, M. (1986). Identification and analysis of genes for tetracycline resistance and replication functions in the broad-host-range plasmid pLS1. *J Mol Biol* **192**, 753-765.

Laniel, M.-A., Béliveau, A., and Guérin, S.L. (2001). Electrophoretic mobility shift assays for the analysis of DNA-Protein interactions. In DNA-Protein Interactions (Humana Press), pp. 13-30.

Laue, T.M., Shah, B.D., Ridgeway, T.M., and Pelletier, S.L. (1992). Analytical ultracentrifugation in biochemistry and polymer sciences. Harding S.E., Rowe A., Horton J.C editors Cambridge: Royal Society of Chemistry, pp. 90-125

LeBlanc, D.J., Chen, Y.Y., and Lee, L.N. (1993). Identification and characterization of a mobilization gene in the streptococcal plasmid, pVA380-1. *Plasmid* **30**, 296-302.

Lebreton, F., Riboulet-Bisson, E., Serror, P., Sanguinetti, M., Posteraro, B., Torelli, R., Hartke, A., Auffray, Y., and Giard, J.C. (2009). *ace*, which encodes an adhesin in

Enterococcus faecalis, is regulated by Ers and is involved in virulence. *Infect Immun* **77**, 2832-2839.

Lebreton, F., Willems, R., and Gilmore, M. (2014). *Enterococcus* diversity, origins in nature, and gut colonization. In Gilmore MS, Clewell DB, Ike Y, *et al*, editors Enterococci: From Commensals to Leading Causes of Drug Resistant Infection [Internet] (Boston: Massachusetts Eye and Ear Infirmary).

Le Breton, Y., Pichereau, V., Sauvageot, N., Auffray, Y., and Rincé, A. (2005). Maltose utilization in *Enterococcus faecalis*. *J Appl Microbiol* **98**, 806-813.

Lin, H., Deng, E.-Z., Ding, H., Chen, W., and Chou, K.-C. (2014). iPro54-PseKNC: a sequence-based predictor for identifying sigma-54 promoters in prokaryote with pseudo k-tuple nucleotide composition. *Nucleic Acids Res* **42**, 12961-12972.

Lindenstraüss, A.G., Ehrmann, M.A., Behr, J., Landstorfer, R., Haller, D., Sartor, R.B., and Vogel, R.F. (2014). Transcriptome analysis of *Enterococcus faecalis* toward its adaption to surviving in the mouse intestinal tract. *Arch Microbiol* **196**, 423-433.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the C_T Method. *Methods* **25**, 402-408.

López, P., Espinosa, M., Greenberg, B., and Lacks, S.A. (1987). Sulfonamide resistance in *Streptococcus pneumoniae*: DNA sequence of the gene encoding dihydropteroate synthase and characterization of the enzyme. *J Bacteriol* **169**, 4320-4326.

López, P., Martínez, S., Díaz, A., Espinosa, M., and Lacks, S.A. (1989). Characterization of the *polA* gene of *Streptococcus pneumoniae* and comparison of the DNA polymerase I it encodes to homologous enzymes from *Escherichia coli* and phage T7. *J Biol Chem* **264**, 4255-4263.

Maddox, S.M., Coburn, P.S., Shankar, N., and Conway, T. (2012). Transcriptional regulator PerA influences biofilm-associated, platelet binding, and metabolic gene expression in *Enterococcus faecalis*. *PLoS ONE* **7**, e34398.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). Molecular cloning: a laboratory manual. *Cold Spring Harbor Laboratory Press New York*.

Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I., Lanczycki, C.J.,Lu, F.,Marchler, G.H., Song, J.S.,Thanki, N.,Wang, Z.,Yamashita, R. A., Zhang, D.,Zheng, C., and Bryant, S. H.(2015). CDD: NCBI's conserved domain database. *Nucleic Acids Res* **43**, D222-D226.

McIver, K.S. (2009). *Stand-alone* response regulators controlling global virulence networks in *Streptococcus pyogenes*. *Contrib Microbiol* **16**, 103-119.

McIver, K.S., Heath, A.S., Green, B.D., and Scott, J.R. (1995). Specific binding of the activator Mga to promoter sequences of the *emm* and *scpA* genes in the group A *Streptococcus*. *J Bacteriol* **177**, 6619-6624.

McIver, K.S., and Myles, R.L. (2002). Two DNA-binding domains of Mga are required for virulence gene activation in the group A streptococcus. *Mol Microbiol* **43**, 1591-1601.

McIver, K.S., Thurman, A.S., and Scott, J.R. (1999). Regulation of *mga* transcription in the Group A *Streptococcus*: specific binding of Mga within its own promoter and evidence for a negative regulator. *J Bacteriol* **181**, 5373-5383.

Méjean, V., Rives, I., and Claverys, J.P. (1990). Nucleotide sequence of the *Streptococcus pneumoniae ung* gene encoding uracil-DNA glycosylase. *Nucleic Acids Res* **18**, 6693.

Merino, E., and Yanofsky, C. (2005). Transcription attenuation: a highly conserved regulatory strategy used by bacteria. *Trends Genet* **21**, 260-264.

Michaux, C., Sanguinetti, M., Reffuveille, F., Auffray, Y., Posteraro, B., Gilmore, M.S., Hartke, A., and Giard, J.-C. (2011). SlyA Is a transcriptional regulator involved in the virulence of *Enterococcus faecalis*. *Infect Immun* **79**, 2638-2645.

Miller, W.G., and Lindow, S.E. (1997). An improved GFP cloning cassette designed for prokaryotic transcriptional fusions. *Gene* **191**, 149-153.

Minchin, S.D., and Busby, S.J.W. (2009). Analysis of mechanisms of activation and repression at bacterial promoters. *Methods* **47**, 6-12.

Mitchell, J.E., Zheng, D., Busby, S.J.W., and Minchin, S.D. (2003). Identification and analysis of 'extended –10' promoters in *Escherichia coli*. *Nucleic Acids Res* **31**, 4689-4695.

Mitra, A., Kesarwani, A.K., Pal, D., and Nagaraja, V. (2010). WebGeSTer DB-a transcription terminator database. *Nucleic Acids Res* **39,** D129-D135**.**

Mohamed, J.A., Huang, W., Nallapareddy, S.R., Teng, F., and Murray, B.E. (2005). Influence of origin of isolates, especially endocarditis isolates, and various genes on biofilm formation by *Enterococcus faecalis*. *Infect Immun* **72**, 3658-3663.

Muller, C., Cacaci, M., Sauvageot, N., Sanguinetti, M., Rattei, T., Eder, T., Giard, J.-C., Kalinowski, J., Hain, T., and Hartke, A. (2015). The intraperitoneal transcriptome of the opportunistic pathogen *Enterococcus faecalis* in mice. *PLoS ONE* **10**, e0126143.

Murray, B.E. (1990). The life and times of the *Enterococcus*. *Clin Microbiol Rev* **3**, 46-65.

Mylonakis, E., Engelbert, M., Qin, X., Sifri, C.D., Murray, B.E., Ausubel, F.M., Gilmore, M.S., and Calderwood, S.B. (2002). The *Enterococcus faecalis fsrB* gene, a key component of the *fsr* quorum-sensing system, is associated with virulence in the rabbit endophthalmitis model. *Infect Immun* **70**, 4678-4681.

Nakayama, J., Chen, S., Oyama, N., Nishiguchi, K., Azab, E.A., Tanaka, E., Kariyama, R., and Sonomoto, K. (2006). Revised model for *Enterococcus faecalis fsr* quorum-sensing system: the small open reading frame *fsrD* encodes the gelatinase biosynthesis-activating pheromone propeptide corresponding to staphylococcal AgrD. *J Bacteriol* **188**, 8321-8326.

Nallapareddy, S.R., Qin, X., Weinstock, G.M., Höok, M., and Murray, B.E. (2000a). *Enterococcus faecalis* adhesin, Ace, mediates attachment to extracellular matrix proteins collagen type IV and laminin as well as collagen type I. *Infect Immun* **68**, 5218-5224.

Nallapareddy, S.R., Sillanpää, J., Mitchell, J., Singh, K.V., Chowdhury, S.A., Weinstock, G.M., Sullam, P.M., and Murray, B.E. (2011). Conservation of Ebp-type pilus genes among enterococci and demonstration of their role in adherence of *Enterococcus faecalis* to human platelets. *Infect Immun* **79**, 2911-2920.

Nallapareddy, S.R., Singh, K.V., Duh, R.-W., Weinstock, G.M., and Murray, B.E. (2000b). Diversity of *ace*, a gene encoding a microbial surface component recognizing adhesive matrix molecules, from different strains of *Enterococcus faecalis* and evidence for production of Ace during human infections. *Infect Immun* **68**, 5210-5217.

Nallapareddy, S.R., Singh, K.V., Sillanpää, J., Garsin, D.A., Höök, M., Erlandsen, S.L., and Murray, B.E. (2006). Endocarditis and biofilm-associated pili of *Enterococcus faecalis*. *J Clini Invest* **116**, 2799-2807.

Naville, M., and Gautheret, D. (2009). Transcription attenuation in bacteria: theme and variations. *Brief Funct Genomic Proteomic* **8**, 482-492.

Naville, M., Ghuillot-Gaudeffroy, A., Marchais, A., and Gautheret, D. (2011). ARNold: a web tool for the prediction of Rho-independent transcription terminators. *RNA Biol* **8**, 11-13.

Nieto, C., and Espinosa, M. (2003). Construction of the mobilizable plasmid pMV158GFP, a derivative of pMV158 that carries the gene encoding the green fluorescent protein. *Plasmid* **49**, 281-285.

Okada, N., Geist, R.T., and Caparon, M.G. (1993). Positive transcriptional control of *mry* regulates virulence in the group A streptococcus. *Mol Microbiol* **7**, 893-903.

Olmsted, S.B., Kao, S.M., van Putte, L.J., Gallo, J.C., and Dunny, G.M. (1991). Role of the pheromone-inducible surface protein Asc10 in mating aggregate formation and conjugal transfer of the *Enterococcus faecalis* plasmid pCF10. *J Bacteriol* **173**, 7665-7672

Osipiuk, J., Wu, R., Jedrzejczak, R., Moy, S., and Joachimiak, A. (2011). Putative Mga family transcriptional regulator from *Enterococcus faecalis*.

Park, S.Y., Shin, Y.P., Kim, C.H., Park, H.J., Seong, Y.S., Kim, B.S., Seo, S.J., and Lee, I.H. (2008). Immune evasion of *Enterococcus faecalis* by an extracellular gelatinase that cleaves C3 and iC3b. *J Immunol* **181**, 6328-6336.

Paterson, G.K., and Mitchell, T.J. (2006). The role of *Streptococcus pneumoniae* sortase A in colonisation and pathogenesis. *Microbes Infec* **8**, 145-153.

Paulsen, I.T., Banerjei, L., Myers, G.S.A., Nelson, K.E., Seshadri, R., Read, T.D., Fouts, D.E., Eisen, J.A., Gill, S.R., Heidelberg, J.F., Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M., Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W., Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H., Utterback T., Radune D., Ketchum K.A., Dougherty B.A., and Fraser C.M. (2003). Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* **299**, 2071-2074.

Pessen, H., and Kumosinski, T.F. (1985). Measurements of protein hydration by various techniques. *Methods Enzymol* **117**, 219-255.

Qin, X., Singh, K.V., Weinstock, G.M., and Murray, B.E. (2000). Effects of *Enterococcus faecalis fsr* genes on production of gelatinase and a serine protease and virulence. *Infect Immun* **68**, 2579-2586.

Rakita, R.M., Vanek, N.N., Jacques-Palaz, K., Mee, M., Mariscalco, M.M., Dunny, G.M., Snuggs, M., Van Winkle, W.B., and Simon, S.I. (1999). *Enterococcus faecalis* bearing aggregation substance is resistant to killing by human neutrophils despite phagocytosis and neutrophil activation. *Infect Immun* **67**, 6067-6075.

Ramsey, M., Hartke, A., and Huycke, M. (2014). The physiology and metabolism of Enterococci. In Gilmore MS, Clewell DB, Ike Y, *et al* Enterococci: From Commensals to Leading Causes of Drug Resistant Infection [Internet] (Boston: Massachusetts Eye and Ear Infirmary).

Repizo, G.D., Blancato, V.S., Mortera, P., Lolkema, J.S., and Magni, C. (2013). Biochemical and genetic characterization of the *Enterococcus faecalis* oxaloacetate decarboxylase complex. *Appl Environ Microbiol* **79**, 2882-2890.

Ribardo, D.A., and McIver, K.S. (2006). Defining the Mga regulon: comparative transcriptome analysis reveals both direct and indirect regulation by Mga in the group A *Streptococcus*. *Mol Microbiol* **62**, 491-508.

Riboulet-Bisson, E., Hartke, A., Auffray, Y., and Giard, J.-C. (2009). Ers controls glycerol metabolism in *Enterococcus faecalis*. *Curr Microbiol* **58**, 201-204.

Riboulet-Bisson, E., Sanguinetti, M., Budin-Verneuil, A.l., Auffray, Y., Hartke, A., and Giard, J.-C. (2008). Characterization of the Ers regulon of *Enterococcus faecalis*. *Infect Immun* **76**, 3064-3074.

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet* **16**, 276-277.

Rich, R.L., Kreikemeyer, B., Owens, R.T., LaBrenz, S., Narayana, S.V.L., Weinstock, G.M., Murray, B.E., and Höok, M. (1999). Ace is a collagen-binding MSCRAMM from *Enterococcus faecalis*. *J Biol Chem* **274**, 26939-26945.

Rohs, R., Jin, X., West, S.M., Joshi, R., Honig, B., and Mann, R.S. (2010). Origins of Specificity in Protein-DNA Recognition. *Annu Rev Biochem* **79**, 233-269.

Ross, W., and Gourse, R.L. (2009). Analysis of RNA polymerase-promoter complex formation. *Methods* **47**, 13-24.

Rozen, S., and Skaletsky, H. J. (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386.

Ruiz-Cruz, S., Solano-Collado, V., Espinosa, M., and Bravo, A. (2010). Novel plasmid-based genetic tools for the study of promoters and terminators in *Streptococcus pneumoniae* and *Enterococcus faecalis*. *J Microbiol Methods* **83**, 156-163.

Sabelnikov, A.G., Greenberg, B., and Lacks, S.A. (1995). An extended -10 promoter alone directs transcription of the DpnII operon of *Streptococcus pneumoniae*. *J Mol Biol* **250**, 144-155.

Sahm, D., Kissinger, J., Gilmore, M.S., Murray, P.R., Mulder, R., Solliday, J, and Clarke B. (1989). In vitro susceptibility studies of vancomycin-resistant *Enterococcus faecalis*. *Antimicrob Agents Chemother* **33**, 1588-1591.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463-5467.

Sanson, M., Makthal, N., Gavagan, M., Cantu, C., Olsen, R.J., Musser, J.M., and Kumaraswami, M. (2015). Phosphorylation events in the multiple gene regulator

of Group A *Streptococcus* significantly influence global gene expression and virulence. *Infect Immun* **83**, 2382-2395.

Sartingen, S., Rozdzinski, E., Muscholl-Silberhorn, A., and Marre, R. (2000). Aggregation substance increases adherence and internalization, but not translocation, of *Enterococcus faecalis* through different intestinal epithelial cells *in vitro*. *Infect Immun* **68**, 6044-6047.

Sava, I.G., Heikens, E., and Huebner, J. (2010). Pathogenesis and immunity in enterococcal infections. *Clin Microbiol Infec* **16**, 533-540.

Scott, J.A.G. (2007). The preventable burden of pneumococcal disease in the developing world. *Vaccine* **25**, 2398-2405.

Schägger, H., and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166**, 368-379.

Schlievert, P.M., Gahr, P.J., Assimacopoulos, A.P., Dinges, M.M., Stoehr, J.A., Harmala, J.W., Hirt, H., and Dunny, G.M. (1998). Aggregation and binding substances enhance pathogenicity in rabbit models of *Enterococcus faecalis* endocarditis. *Infect Immun* **66**, 218-223.

Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C_T method. *Nat Protoc* **3**, 1101-1108.

Schuck, P., and Rossmanith, P. (2000). Determination of the sedimentation coefficient distribution by least-squares boundary modeling. *Biopolymers* **54**, 328-341.

Serrano-Heras, G., Salas, M., and Bravo, A. (2005). A new plasmid vector for regulated gene expression in *Bacillus subtilis*. *Plasmid* **54**, 278-282.

Shankar, N., Baghdayan, A.S., and Gilmore, M.S. (2002). Modulation of virulence within a pathogenicity island in vancomycin-resistant *Enterococcus faecalis*. *Nature* **417**, 746-750.

Shankar, N., Lockatell, C.V., Baghdayan, A.S., Drachenberg, C., Gilmore, M.S., and Johnson, D.E. (2001). Role of *Enterococcus faecalis* surface protein Esp in the pathogenesis of ascending urinary tract infection. *Infect Immun* **69**, 4366-4372.

Shepard, B.D., and Gilmore, M.S. (1995). Electroporation and efficient transformation of *Enterococcus faecalis* grown in high concentrations of glycine. *Methods Mol Biol* **47**, 217-226.

Sifri, C.D., Mylonakis, E., Singh, K.V., Qin, X., Garsin, D.A., Murray, B.E., Ausubel, F.M., and Calderwood, S.B. (2002). Virulence effect of *Enterococcus faecalis* protease genes and the quorum-sensing locus *fsr* in *Caenorhabditis elegans* and mice. *Infect Immun* **70**, 5647-5650.

Sillanpää, J., Xu, Y., Nallapareddy, S.R., Murray, B.E., and Höök, M. (2004). A family of putative MSCRAMMs from *Enterococcus faecalis*. *Microbiology* **150**, 2069-2078.

Simons, R.W.H., F. and Kleckner, N. (1987). Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene* **53**, 85-96.

Singh, K.V., Lewis, R.J., and Murray, B.E. (2009a). Importance of the *epa* locus of *Enterococcus faecalis* OG1RF in a mouse model of ascending urinary tract infection. *J Infect Dis* **200**, 417-420.

Singh, K.V., Nallapareddy, S.R., Nannini, E.C., and Murray, B.E. (2005). Fsr-independent production of protease(s) may explain the lack of attenuation of an *Enterococcus faecalis fsr* mutant *versus* a *gelE-sprE* mutant in induction of endocarditis. *Infect Immun* **73**, 4888-4894.

Singh, K.V., Nallapareddy, S.R., Sillanpää, J., and Murray, B.E. (2009b). Importance of the collagen adhesin Ace in pathogenesis and protection against *Enterococcus faecalis* experimental endocarditis. *PLoS Pathog* **6**, e1000716.

Singh, K.V., Qin, X., Weinstock, G.M., and Murray, B.E. (1998). Generation and testing of mutants of *Enterococcus faecalis* in a mouse peritonitis model. *J Infect Dis* **178**, 1416-1420.

Solano-Collado, M.V. (2014). Caracterización molecular del regulador transcripcional Mga*Spn* de *Streptococcus pneumoniae.* Doctoral Thesis , Universidad Complutense de Madrid

Solano-Collado, V., Espinosa, M., and Bravo, A. (2012). Activator role of the pneumococcal Mga-like virulence transcriptional regulator. *J Bacteriol* **194**, 4197-4207.

Solano-Collado, V., Lurz, R., Espinosa, M., and Bravo, A. (2013). The pneumococcal Mga*Spn* virulence transcriptional regulator generates multimeric complexes on linear double-stranded DNA. *Nucleic Acids Res* **41**, 6975-6991.

Song, K. (2012). Recognition of prokaryotic promoters based on a novel variable-window Z-curve method. *Nucleic Acids Res* **40**, 963-971.

Srivastava, M., Mallard, C., Barke, T., Hancock, L.E., and Self, W.T. (2011). A selenium-dependent xanthine dehydrogenase triggers biofilm proliferation in *Enterococcus faecalis* through oxidant production. *J Bacteriol* **193**, 1643-1652.

Stassi, D.L., Lopez, P., Espinosa, M., and Lacks, S.A. (1981). Cloning of chromosomal genes in *Streptococcus pneumoniae*. *Proc Natl Acad Sci USA* **78**, 7028-7032.

Stella, S., Cascio, D., and Johnson, R.C. (2009). The shape of the DNA minor groove directs binding by the DNA-bending protein Fis. *Genes Dev* **24**, 814-826.

Sterba, K.M., Mackintosh, S.G., Blevins, J.S., Hurlburt, B.K., and Smeltzer, M.S. (2003). Characterization of *Staphylococcus aureus* SarA binding sites. *J Bacteriol* **185**, 4410-4417.

Strauch, M.A., Ballar, P., Rowshan, A.J., and Zoller, K.L. (2005). The DNA-binding specificity of the *Bacillus anthracis* AbrB protein. *Microbiology* **151**, 1751-1759.

Studier, F.W., and Moffatt, B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**, 113-130.

Suárez, C., Blancato, V., Poncet, S., Deutscher, J., and Magni, C. (2011). CcpA represses the expression of the divergent *cit* operons of *Enterococcus faecalis* through multiple *cre* sites. *BMC Microbiol* **11**, 227-240.

Süssmuth, S.D., Muscholl-Silberhorn, A., Wirth, R., Susa, M., Marre, R., and Rozdzinski, E. (2000). Aggregation substance promotes adherence, phagocytosis, and intracellular survival of *Enterococcus faecalis* within human macrophages and suppresses respiratory burst. *Infect Immun* **68**, 4900-4906.

Teng, F., Jacques-Palaz, K.D., Weinstock, G.M., and Murray, B.E. (2002). Evidence that the enterococcal polysaccharide antigen gene (*epa*) cluster is widespread in *Enterococcus faecalis* and influences resistance to phagocytic killing of *E. faecalis*. *Infect Immun* **70**, 2010-2015.

Theilacker, C., Sánchez-Carballo, P., Toma, I., Fabretti, F., Sava, I., Kropec, A., Holst, O., and Huebner, J. (2009). Glycolipids are involved in biofilm accumulation and prolonged bacteraemia in *Enterococcus faecalis*. *Mol Microbiol* **71**, 1055-1069.

Thomas, V.C., Hiromasa, Y., Harms, N., Thurlow, L., Tomich, J., and Hancock, L.E. (2009). A fratricidal mechanism is responsible for eDNA release and contributes to biofilm development of *Enterococcus faecalis*. *Mol Microbiol* **72**, 1022-1036.

Thurlow, L.R., Thomas, V.C., Fleming, S.D., and Hancock, L.E. (2009a). Enterococcus faecalis capsular polysaccharide serotypes C and D and their contributions to host innate immune evasion. *Infect Immun* **77**, 5551-5557.

Thurlow, L.R., Thomas, V.C., and Hancock, L.E. (2009b). Capsular polysaccharide production in *Enterococcus faecalis* and contribution of CpsF to capsule serospecificity. *J Bacteriol* **191**, 6203-6210.

Thurlow, L.R., Thomas, V.C., Narayanan, S., Olson, S., Fleming, S.D., and Hancock, L.E. (2009c). Gelatinase contributes to the pathogenesis of endocarditis caused by *Enterococcus faecalis*. *Infect Immun* **78**, 4936-4943.

Tsvetanova, B., Wilson, A.C., Bongiorni, C., Chiang, C., Hoch, J.A., and Perego, M. (2007). Opposing effects of histidine phosphorylation regulate the AtxA virulence transcription factor in *Bacillus anthracis*. *Mol Microbiol* **63**, 644-655.

Turner, A.K., Lovell, M.A., Hulme, S.D., Zhang-Barber, L., and Barrow, P.A. (1998). Identification of *Salmonella typhimurium* genes required for colonization of the chicken alimentary tract and for virulence in newly hatched chicks. *Infect Immun* **66**, 2099-2106.

Vahling, C.M., and McIver, K.S. (2006). Domains required for transcriptional activation show conservation in the Mga family of virulence gene regulators. *J Bacteriol* **188**, 863-873.

van Holde, K.E. (1985). *Physical Biochemistry*. 2nd edn. Prentice Hall: Englewoods Cliffs

van Schaik, W., Château, A., Dillies, M.-A., Coppée, J.-Y., Sonenshein, A.L., and Fouet, A. (2009). The global regulator CodY regulates toxin gene expression in *Bacillus anthracis* and is required for full virulence. *Infect Immun* **77**, 4437-4445.

Van Tyne, D., Martin, M.J., and Gilmore, M.S. (2013). Structure, function, and biology of the *Enterococcus faecalis* cytolysin. *Toxins* **5**, 895-911.

Vebø, H.C., Snipen, L., Nes, I.F., and Brede, D.A. (2009). The transcriptome of the nosocomial pathogen *Enterococcus faecalis* V583 reveals adaptive responses to growth in blood. *PLoS ONE* **4**, e7660.

Vebø, H.C., Solheim, M., Snipen, L., Nes, I.F., and Brede, D.A. (2010). Comparative genomic analysis of pathogenic and probiotic *Enterococcus faecalis* isolates, and their transcriptional responses to growth in human urine. *PLoS ONE* **5**, e12489.

Vlahovicek, K., Kaján, L., and Pongor, S. (2003). DNA analysis servers: plot.it, bend.it, model.it and IS. *Nucleic Acids Res* **31**, 3686-3687.

Voskuil, M.I., and Chambliss, G.H. (1998). The -16 region of *Bacillus subtilis* and other gram-positive bacterial promoters. *Nucleic Acids Res* **26**, 3584-3590.

Wigneshweraraj, S., Bose, D., Burrows, P.C., Joly, N., Schumacher, J.r., Rappas, M., Pape, T., Zhang, X., Stockley, P., Severinov, K., Buck, M. (2008). *Modus operandi* of the bacterial RNA polymerase containing the σ^{54} promoter-specificity factor. *Mol Microbiol* **68**, 538-546.

Xu, Y., Jiang, L., Murray, B.E., and Weinstock, G.M. (1997). *Enterococcus faecalis* antigens in human infections. *Infect Immun* **65**, 4207-4215.

Xu, Y., Singh, K.V., Qin, X., Murray, B.E., and Weinstock, G.M. (2000). Analysis of a gene cluster of *Enterococcus faecalis* involved in polysaccharide biosynthesis. *Infect Immun* **68**, 815-823.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103-119.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., and Madden, T.L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinform* **13**, 134-134.

Zeng, J., Teng, F., Weinstock, G.M., and Murray, B.E. (2004). Translocation of *Enterococcus faecalis* strains across a monolayer of polarized human enterocyte-like T84 cells. *J Clin Microbiol* **42**, 1149-1154.

Zúñiga, M., Comas, I., Linaje, R., Monedero, V., Yebra, M.J., Esteban, C.D., Deutscher, J., Pérez-Martínez, G., and González-Candelas, F. (2005). Horizontal gene transfer in the molecular evolution of mannose PTS transporters. *Mol Biol Evol* **22**, 1673-1685.

ANNEXES

Table 1: Genes differentially expressed in MA*Efa*-lacking cells

Locus tag	V583	Gene	Log_2FC	Description	KEGG pathways
OG1RF_10064	EF0071		-4.68	glycoside hydrolase	ND
OG1RF_10107	EF0114		-4.52	family 20 glycosyl hydrolase	Other glycan degradation Amino sugar and nucleotide sugar metabolism Metabolic pathways Biosynthesis of secondary metabolites
OG1RF_10108	EF0115		-3.89	endoribonuclease L-PSP	ND
OG1RF_10198	EF0253	*aldA*	-3.52	aldehyde dehydrogenase	ND
OG1RF_10296	EF0411	*mtlA2*	-6.40	PTS family mannitol porter, EIICB component	Fructose and mannose metabolism Phosphotransferase system (PTS)
OG1RF_10297	EF0412	*mltF2*	-6.40	PTS family fructose/mannitol porter component IIA	Fructose and mannose metabolism Phosphotransferase system (PTS)
OG1RF_10298	EF0413	*mtlD*	-6.37	mannitol-1-phosphate 5-dehydrogenase	Fructose and mannose metabolism
OG1RF_10433	EF0695		-3.67	PTS family fructose/mannitol (fru) porter component IIA	Fructose and mannose metabolism Metabolic pathways Microbial metabolism in diverse environments Phosphotransferase system (PTS)
OG1RF_10434	EF0696	*lacD1*	-3.44	tagatose-bisphosphate aldolase	Galactose metabolism Metabolic pathways
OG1RF_10682	EF0956	*pgmB*	-3.98	beta-phosphoglucomutase	Starch and sucrose metabolism
OG1RF_10683	EF0957	*map*	-4.00	family 65 glycosyl hydrolase	Starch and sucrose metabolism Metabolic pathways
OG1RF_10684	EF0958	*exp5*	-4.11	PTS family porter component IIABC	Glycolysis/Gluconeogenesis Starch and sucrose metabolism Amino sugar and nucleotide sugar metabolism Phosphotransferase system (PTS)
OG1RF_10746	EF1013		-4.50	PTS family lactose/cellobiose porter component IIC	Phosphotransferase system (PTS)
OG1RF_10747	EF1014		-3.70	hypothetical protein	ND
OG1RF_10748	EF1015		-4.09	hypothetical protein	ND
OG1RF_10749	EF1016		-4.06	hypothetical protein	ND
OG1RF_10979	EF1207		-3.37	CCS family citrate carrier protein	ND
OG1RF_11004	EF1233		-3.61	ABC superfamily ATP binding cassette transporter, membrane protein	ABC transporters

OG1RF_11005	EF1234		-3.72	ABC superfamily ATP binding cassette transporter, substrate-binding protein	ABC transporters
OG1RF_11008	EF1237		-3.17	endonuclease/exonuclease/phosphatase	ND
OG1RF_11009	EF1238	*bglX*	-3.46	putative beta-glucosidase	Cyanoamino acid metabolism Starch and sucrose metabolism Metabolic pathways Biosynthesis of secondary metabolites
OG1RF_11133	EF1343		-4.20	sugar ABC transporter ATP-binding protein	ABC transporters
OG1RF_11134	EF1344		-4.20	sugar ABC transporter ATP-binding protein	ABC transporters
OG1RF_11135	EF1345		-4.48	sugar ABC superfamily ATP binding cassette transporter, sugar-binding protein	ABC transporters
OG1RF_11136	EF1347	*nplT*	-3.73	neopullulanase	ND
OG1RF_11137	EF1348	*dexB*	-4.11	glucan 1,6-alpha-glucosidase	ND
OG1RF_11138	EF1349	*malL*	-4.25	oligo-1,6-glucosidase	Galactose metabolism Starch and sucrose metabolism Metabolic pathways
OG1RF_11146	EF1358	*gldA*	-4.92	glycerol dehydrogenase	Glycerolipid metabolism Metabolic pathways
OG1RF_11147	EF1359		-5.04	glycerone kinase PTS family porter component IIA	ND
OG1RF_11148	EF1360	*dhaK*	-5.01	dihydroxyacetone kinase	Glycerolipid metabolism Metabolic pathways
OG1RF_11149	EF1361		-4.95	dihydroxyacetone kinase	Glycerolipid metabolism Metabolic pathways
OG1RF_11150	EF1362		-3.20	hypothetical protein	ND
OG1RF_11184	EF1395	*moaB*	-3.30	molybdenum cofactor biosynthesis protein B	ND
OG1RF_11185	EF1396		-3.10	molybdopterin-binding domain protein	ND
OG1RF_11186	EF1397	*modA*	-3.70	molybdenum ABC superfamily ATP binding cassette transporter, binding protein	ABC transporters
OG1RF_11584	EF1920		-3.43	C4-dicarboxylate anaerobic carrier	ND
OG1RF_11585	EF1921		-3.47	Putative ribosylpyrimidine nucleosidase	ND
OG1RF_11590	EF1927	*glpF2*	-3.26	MIP family glycerol uptake facilitator protein GlpF	ND
OG1RF_11591	EF1928	*glpO*	-3.31	glycerol-3-phosphate oxidase	ND

OG1RF_11592	EF1929	*glpK*	-3.32	glycerol kinase	Glycerolipid metabolism Metabolic pathways
OG1RF_11611	EF1950		-3.84	phosphosugar isomerase	ND
OG1RF_11612	EF1951		-4.03	phosphosugar-binding protein	ND
OG1RF_11613	EF1952		-3.34	PTS system mannose/fructose/sorbose transporter subunit IID	Fructose and mannose metabolism Amino sugar and nucleotide sugar metabolism Metabolic pathways Phosphotransferase system (PTS)
OG1RF_11614	EF1953		-4.81	PTS family mannose/fructose/sorbose porter component IIC	Fructose and mannose metabolism Amino sugar and nucleotide sugar metabolism Metabolic pathways Phosphotransferase system (PTS)
OG1RF_11615		*ulaB*	-3.66	PTS family ascorbate porter, IIB component	Fructose and mannose metabolism Amino sugar and nucleotide sugar metabolism Metabolic pathways Phosphotransferase system (PTS)
OG1RF_11616			-4.14	PTS system mannose/fructose/sorbose transporter subunit IIA	Fructose and mannose metabolism Amino sugar and nucleotide sugar metabolism Metabolic pathways Phosphotransferase system (PTS)
OG1RF_11753	EF2213	*treB*	-5.09	PTS family trehalose porter, IIBC component	Starch and sucrose metabolism Phosphotransferase system (PTS)
OG1RF_11761	EF2221		-4.98	ABC superfamily ATP binding cassette transporter, binding protein	ABC transporters
OG1RF_11762	EF2222		-5.39	carbohydrate ABC superfamily ATP binding cassette transporter, membrane protein	ABC transporters
OG1RF_11763	EF2223		-6.05	ABC superfamily ATP binding cassette transporter, membrane protein	ABC transporters
OG1RF_11775	EF2235	*nifJ2*	-3.28	glucuronyl hydrolase	ND
OG1RF_11940	EF2559		-3.12	pyruvate:ferredoxin oxidoreductase	Metabolic pathways Microbial metabolism in diverse environments Carbon metabolism
OG1RF_11944	EF2563		-3.03	YqeB family selenium-dependent molybdenum hydroxylase system protein	ND
OG1RF_11948	EF2567	*selD*	-3.53	selenide, water dikinase	Selenocompound metabolism Metabolic pathways
OG1RF_11949	EF2568		-3.15	putative cysteine desulfurase	ND
OG1RF_11950	EF2569	*ygfJ*	-3.12	YgfJ family molybdenum hydroxylase accessory protein	ND

OG1RF_11951	EF2570		-3.13	selenium-dependent molybdenum hydroxylase 1	ND
OG1RF_11956	EF2575	*arcC3*	-3.21	carbamate kinase	Purine metabolism Arginine and proline metabolism Nitrogen metabolism Microbial metabolism in diverse environments Carbon metabolism
OG1RF_11957	EF2577	*ygeW*	-3.39	carbamoyltransferase YgeW	ND
OG1RF_11958	EF2578	*ygeY*	-3.88	M20/DapE family protein YgeY	ND
OG1RF_11959	EF2579	*dpaL*	-3.21	diaminopropionate ammonia-lyase	ND
OG1RF_12167	EF2863	*endOF3*	-4.43	mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase	ND
OG1RF_12249	EF2966		-3.17	BglG family transcriptional antiterminator	ND
OG1RF_12303	EF3023		-4.49	family 8 polysaccharide lyase	ND
OG1RF_12397	EF3134	*eda*	-4.17	2-dehydro-3-deoxyphosphogluconate aldolase	Pentose phosphate pathway Glyoxylate and dicarboxylate metabolism Metabolic pathways Microbial metabolism in diverse environments Carbon metabolism
OG1RF_12398	EF3135	*uxuA*	-4.34	mannonate dehydratase	Pentose and glucuronate interconversions Metabolic pathways
OG1RF_12399	EF3136	*sorA*	-4.60	PTS system mannose/fructose/sorbose transporter subunit IIA	Fructose and mannose metabolism Amino sugar and nucleotide sugar metabolism Metabolic pathways Phosphotransferase system (PTS)
OG1RF_12400	EF3137	*sorB2*	4.64	PTS family mannose/fructose/sorbose porter, IIB component	Fructose and mannose metabolism Amino sugar and nucleotide sugar metabolism Metabolic pathways Phosphotransferase system (PTS)
OG1RF_12401	EF3138	*sorD*	-4.87	PTS system mannose/fructose/sorbose transporter subunit IID	Fructose and mannose metabolism Amino sugar and nucleotide sugar metabolism Metabolic pathways Phosphotransferase system (PTS)
OG1RF_12402	EF3139	*sorC*	-4.35	PTS family mannose/fructose/sorbose porter component IIC	Fructose and mannose metabolism Amino sugar and nucleotide sugar metabolism Metabolic pathways Phosphotransferase system (PTS)
OG1RF_12403	EF3140	*gldA2*	-3.94	putative glycerol dehydrogenase	ND
OG1RF_12404	EF3141		-3.97	D-isomer specific 2-hydroxyacid dehydrogenase	ND

OG1RF_12405	EF3142	*gnd2*	-3.73	6-phosphogluconate dehydrogenase	Pentose phosphate pathway Glutathione metabolism Metabolic pathways Biosynthesis of secondary metabolites Microbial metabolism in diverse environments Carbon metabolism
OG1RF_12425	EF3157	*trePP*	-4.07	glycosyl hydrolase	ND
OG1RF_12426	EF3158	*yvdM*	-3.97	beta-phosphoglucomutase	Starch and sucrose metabolism
OG1RF_12561	EF3316	*mdh2*	-3.07	malate dehydrogenase (oxaloacetate-decarboxylating) (NADP(+))	Pyruvate metabolism Carbon metabolism Two-component system
OG1RF_12562	EF3317	*oadA*	-3.6	oxaloacetate decarboxylase	Pyruvate metabolism Metabolic pathways Two-component system
OG1RF_12563	EF3318	*citX*	-3.86	2-(5''-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase	Two-component system Two-component system
OG1RF_12564	EF3319	*citF*	-3.98	citrate (pro-3S)-lyase	Two-component system
OG1RF_12565	EF3320	*citE*	-4.19	citrate (Pro-3S)-lyase	Two-component system
OG1RF_12566	EF3321	*citD*	-4.30	citrate lyase acyl carrier protein	ND
OG1RF_12567	EF3322	*citC*	-4.39	[citrate [pro-3S]-lyase] ligase	Pyruvate metabolism Metabolic pathways
OG1RF_12568	EF3323		-3.06	hypothetical protein	ND
OG1RF_12569	EF3324	*gcdB*	-4.68	glutaconyl-CoA decarboxylase	ND
OG1RF_12570	EF3325		-4.91	sodium ion-translocating decarboxylase	ND
OG1RF_12571	EF3326		-5.38	hypothetical protein	
OG1RF_12572	EF3327		-4.66	citrate transporter	

PUBLICATIONS ASSOCIATED TO THIS WORK :

Ruiz-Cruz, S., Solano-Collado, V., Espinosa, M., and Bravo, A. (2010). Novel plasmid-based genetic tools for the study of promoters and terminators in *Streptococcus pneumoniae* and *Enterococcus faecalis*. *J Microbiol Methods* **83**, 156-163.

Fernández-López C, Bravo A, **Ruiz-Cruz S**, Solano-Collado V, Garsin DA, Lorenzo-Díaz F, Espinosa M (2014) Mobilizable Rolling-Circle Replicating Plasmids from Gram-Positive Bacteria: A Low-Cost Conjugative Transfer. *Microbiol Spectr* **2**(5):PLAS-0008-2013.

Ruiz-Cruz, S., Espinosa, M., Goldmann, O., and Bravo, A. (2015). Contribution of the MA*Efa* protein to global gene expression and virulence in *Enterococcus faecalis. Submitted*